



Aston University

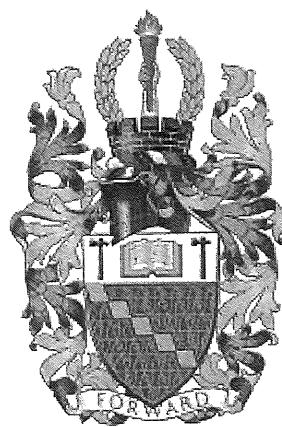
**Some pages of this thesis may have been removed for copyright restrictions.**

If you have discovered material in AURA which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown Policy](#) and [contact the service immediately](#)

# **MODELLING DATA AND VOICE TRAFFIC OVER IP NETWORKS USING CONTINUOUS-TIME MARKOV MODELS**

**AHMED DASTAGIR SHAIKH**

**Doctor of Philosophy**



**ASTON UNIVERSITY**

**October 2009**

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

# MODELLING DATA AND VOICE TRAFFIC OVER IP NETWORKS USING CONTINUOUS-TIME MARKOV TRAFFIC MODELS

AHMED DASTAGIR SHAIKH

Doctor of Philosophy, 2009

## Thesis Summary

As major network operators migrate from the traditional circuit-switched networks to the converged packet-switched networks, global IP traffic is set to grow at a very rapid pace. This calls for efficient methods of analysis, monitoring and modelling of IP traffic, which is crucial for the design of network applications and enhancing the capacity of networks. This analysis of IP traffic is performed with the aid of mathematical traffic models which can be set up for *Monte-Carlo* simulations in order to simulate the network traffic sources and statistics. In order to provide reliable, valid and realistically approximate modelling data for analysis, the network traffic models used, should capture as many features of the network as possible through measured results. Common approaches to traffic modelling have featured the use of stochastic models, based on the *Markov* property, which can be classified into *black box* and *white box* models based on the approach used for modelling traffic. Recently, *black box* models, such as *Markov Modulated Poisson Processes* (MMPP) and *Hidden Markov Models* (HMM), have gained popularity over the traditional Markov and Poisson models, owing to the *bursty* nature of IP traffic. These models are often aimed at reproducing the measured first-order statistics and while they benefit from the statistical advantage that they can be trained to replicate the traffic statistics, they do have a key disadvantage, as the physical reasons behind these statistics, cannot be examined due to the incredibly complex and opaque structure of the models. On the other hand, *white box* models, are simple to understand, transparent and have a physical meaning attributed to each of the associated parameters. To exploit this key advantage, this thesis explores the use of simple classic continuous-time Markov models based on a *white box* approach, to model, not only the network traffic statistics but also the source behaviour with respect to the network and application. The thesis is divided into two parts:

The first part focuses on the use of simple Markov and Semi-Markov traffic models, starting from the simplest two-state model moving upwards to n-state models with Poisson and non-Poisson statistics. The thesis then introduces the convenient to use, mathematically derived, Gaussian Markov models which are used to model the measured network IP traffic statistics. As one of the most significant contributions, the thesis establishes the significance of the second-order density statistics as it reveals that, in contrast to first-order density, they carry much more unique information on traffic sources and behaviour, such as the sequences and periodic structure of packets. The thesis then exploits the use of Gaussian Markov models to model these unique features and finally shows how the use of simple classic Markov models coupled with use of second-order density statistics provides an excellent tool for capturing maximum traffic detail, which in itself is the essence of good traffic modelling.

The second part of the thesis, studies the ON-OFF characteristics of VoIP traffic with reference to accurate measurements of the ON and OFF periods, made from a large multi-lingual database of over 100 hours worth of VoIP call recordings. The impact of the language, prosodic structure and speech rate of the speaker on the statistics of the ON-OFF periods is analysed and relevant conclusions are presented. Finally, an ON-OFF VoIP source model with log-normal transitions is contributed as an ideal candidate to model VoIP traffic and the results of this model are compared with those of previously published work.

**Keywords:** *Markov models, Semi-Markov models, Internet Traffic Modelling, Gaussian distribution, Joint Density, ON-OFF model, Voice over IP*

*To my dearest wife, son and parents*

# Acknowledgements

Pursuing a Ph.D. has been my dream for a long time, and while a lot of hard work and sacrifices have been put into fulfilling this dream, the experience has been very rewarding. Great effort has gone into presenting quality research work in this thesis. However, the thesis would be incomplete without thanking those who made it possible in the first place.

First of all, I would like to thank Dr. Marc A. Eberhard for his excellent supervision, which offered me endless guidance, patience, positivity, encouragement and tips to be an excellent researcher. I also owe my utmost gratitude to Prof. Keith J. Blow for the invaluable advice, mentoring, generosity and the constant support which I received from him at every stage of my time here at Aston. I am also very grateful to him for showing a lot of patience in reading and commenting on the drafts of this thesis. I am very thankful to the UK DTI (Department of Trade and Industry) along with my supervisors, for granting me a research studentship, as part of the HIPNet project. I also wish to gratefully acknowledge the HEFCE (Higher Education Funding Council of England) for granting me the competitive ORSAS (Overseas Research Student Academic Scholarship) award for three years. With these two grants, I was able to pursue my Ph.D. without any financial constraints. Further thanks also go to Prof. Akihiro Maruta, Dr. Yasufumi Morioka, and the GCOE team of Osaka University for providing me a fantastic research internship opportunity in Japan. I will always remember Ms. Shihoku Kurosaki for being an excellent host during my stay in Japan.

This Ph.D. would not have been possible without the love, unconditional support, patience and sacrifices of my dearest wife, Kauser. She bore the brunt of all the difficult and frustrating stages of my time here and has always been on my side. I wish to thank her for all her belief in me. My son Abdullah was the most beautiful thing to have happened to me during my Ph.D. His cheer and exuberance have always inspired me and kept me relaxed even during stressful times. I am greatly indebted to my beloved parents for all their sacrifices, encouragement and love they have offered me throughout my life. They stood by me for every little thing. Without their support, I would not be where I am today. I would also like to thank my in-laws for their endless love and support.

I wish to thank my colleagues and friends who have never hesitated in lending a helping hand, in particular: Dr. Scott Fowler, for all his contributions as a research partner, and Qasim Iqbal, Siham Muhyaldin and Raja Pappu, for their moral support and help at both professional and personal levels. I also deeply appreciate the support I received for my work from members of the ACNRG, Ericsson UK (Coventry) and other teams of the HIPNet consortium. Finally, I would like to offer special thanks to countless other friends and family members, who have loved and supported me throughout my life.

# The HIPNet Project

This Ph.D. has been a part of the HIPNet (Heterogeneous IP Networks) Project, a £ 20 million initiative, by the UK DTI (Department of Trade and Industry) and the EPSRC (Engineering & Physical Sciences Research Council), involving a collaboration between Ericsson, Freescale Semiconductors, Emerson, University of Wales, University of Cambridge, University of Essex, University of Leeds and Aston University. The primary goal of this project focuses on the design and modelling of Next Generation IP networks.

The tremendous growth of the Internet and the rapid advances in IP technology has resulted in the push for an IP core network. Network operators are now gearing up to introduce All-IP networks, which consist of heterogeneous services operating over a converged IP medium. However as these networks grow exponentially, it is important to be able to predict the behaviour of complex networks and develop rules which ensure that networks are built in such a way that they can be scaled to meet the needs of new and demanding services and applications running on the IP network. This needs to be addressed by accurately measuring, analysing and modelling the characteristics of IP traffic for validating networks and use the data to generate traffic in network test beds.

Voice services have long been dominant over the cellular technology market, but with the advent of the IMS and the growth of 3G networks, there will be strong necessity for voice services to be provided on the same technology as that of the multimedia services with better quality. There are several advantages of Voice over IP services over the traditional PSTN or 2G cellular systems. First and foremost, they are a lot cheaper. Secondly they are portable. Besides a VoIP service can provide almost all features which a client can desire from a current generation PSTN or 2G cellular system. However there are drawbacks as well. Because of the nature of the IP network, the Voice Packets can have variable transit time through the network. The delays experienced are larger than that of the Circuit Switched network and if the networks are congested it would mean that there are lost packets. Providing VoIP services over today's networks as well as future networks has been a challenge because of the many issues which need to be tackled such as efficient bandwidth management.

Aston's role in the HIPNet project has been to produce an accurate model for Voice over IP traffic over wireless and wired IP networks, and also characterise IP Multimedia traffic streams using stochastic processes and higher order statistics for the purpose of capacity and design enhancement of Next Generation IP Networks (NGN).

# Table of Contents

Acknowledgements .....	4
The HIPNet Project .....	5
List of Tables.....	12
List of Figures .....	13
Chapter One .....	18
Introduction .....	18
1.1 Background .....	18
1.1.1 The Internet .....	18
1.1.2 Packet Switched Networks .....	18
1.1.3 The Internet Protocol (IP) .....	19
1.1.4 Traffic Modelling .....	20
1.2 Common Approaches to Traffic Modelling .....	22
1.2.1 History.....	22
1.2.2 Black Box approaches .....	24
1.2.3 White Box approaches .....	26
1.3 Motivations and Objectives .....	28
1.3.1 Understanding and Modelling of First and Second order statistics of measured IP traffic flows using simple Markov Models .....	28
1.3.2 Modelling Voice traffic based on measurements from a multi-lingual database of VoIP calls .....	30
1.4 Contributions of the Thesis .....	31
1.5 Outline of the Thesis .....	33
Chapter Two .....	36
Understanding Markov and Semi-Markov Packet Traffic Models .....	36
2.1 Mathematical Background on Markov Models .....	36
2.1.1 What are Markov Models?.....	36
2.1.2 Elements of a Markov Model .....	37

2.1.3	Using Markov Models to model Packet traffic.....	41
2.1.4	Setting up Markov Monte Carlo Simulation to generate IPT statistics from the Markov Models.....	42
2.2	Markov Traffic Models with Poisson Statistics .....	45
2.2.1	A two-state Poisson Markov Traffic Model .....	45
2.2.2	A three-state Poisson Markov Traffic Model .....	55
2.2.3	n-state Poisson Markov Traffic Models .....	61
2.3	Markov Traffic Models with non-Poisson Statistics .....	67
2.3.1	Two state Markov Traffic Model with one state emitting packets .....	67
2.3.2	n-state Markov Traffic Models with only one state emitting packets .....	73
2.4	Semi-Markov Models: A Gamma Markov Model .....	78
2.4.1	30-state Model with non- Poisson statistics .....	78
2.4.2	Simplifying the 30 state model with a 3 state Gamma Markov Model.....	79
2.4.3	Using the 3-state Gamma Markov Model in a Monte Carlo Simulation .....	81
2.5	Semi-Markov Models: A Gaussian Markov Model .....	82
2.5.1	Disadvantages of using a Gamma Markov traffic model .....	82
2.5.2	Using a Gaussian Markov traffic model.....	84
2.5.3	Analysing the Gaussian Monte Carlo simulation results .....	86
2.6	Summary .....	87
<b>Chapter Three</b>	.....	88
Modelling Measured Statistical Results with Gaussian Markov Models	.....	88
3.1	Measuring IP traffic – Methodology .....	88
3.1.1	Approach.....	88
3.1.2	Traffic Measurement Tools Used .....	91
3.2	Measured Results for IPT of IP Traffic.....	92
3.2.1	Example 1: Measurement at Port 80 (TCP) - IP Radio Broadcast .....	92
3.2.2	Example 2: Measurement at Port 80 (UDP) - IP TV Broadcast .....	94
3.2.3	Example 3: Measurement at Port 15010 (UDP) –Telephony Traffic .....	95
3.3	Estimating Gaussian Markov Model Parameters.....	97
3.3.1	Understanding the Gaussian distribution in terms of modelling the peaks – Estimating the mean parameters.....	97

3.3.2	Estimating the state visiting probabilities of the Gaussian Markov Models .....	99
3.3.3	Estimating the standard deviation using the EM algorithm .....	99
3.3.4	Using the amplitudes of the Gaussian components of the Gaussian Markov Models to improve estimation .....	100
3.4	Producing Gaussian Markov Traffic Models for Measured Results .....	100
3.4.1	Modelling Measured Results – Port 80 (TCP) IP Radio .....	100
3.4.2	Modelling Measured Results – Port 80 (UDP) IP TV/Video .....	107
3.4.3	Modelling Measured Results – Port 15010 (UDP) IP Voice/Video Telephony .....	109
3.5	Summary .....	111
<b>Chapter Four</b>	.....	<b>112</b>
	Understanding and modelling Joint Density Results – Packet Sequences .....	112
4.1	The Joint Density results and Packet Sequences .....	112
4.1.1	Joint Density Plots of Modelled Results .....	112
4.1.2	Packet Sequence Information .....	115
4.2	Packet Sequences and Markov Models .....	116
4.2.1	State holding time distribution parameters and model setup .....	116
4.2.2	Case 1: Markov Model ‘A’ with Packets in Random Order .....	117
4.2.3	Case 2: Markov Model ‘B’ with Packets in Sequence (Loop) .....	122
4.3	The Probability Density Function: A projection of the Joint Density Function .....	125
4.4	The Symmetrical and Asymmetrical nature of Joint Density results .....	127
4.4.1	Detecting and filtering out Symmetric and Asymmetric structure from the Joint Density Results .....	127
4.4.2	Symmetric and Asymmetric nature of the modelled results .....	128
4.4.3	Symmetrical nature of measured results .....	130
4.5	Impact of introducing Packet Loss in Markov Traffic Models .....	131
4.5.1	Creating Random packet loss in the Markov Monte Carlo Simulation .....	131
4.5.2	Impact of packet loss on the Joint Density of IPT .....	133
4.6	Summary .....	134

<b>Chapter Five .....</b>	<b>136</b>
Understanding and Modelling Joint Density Results – Periodic Processes.....	136
5.1 Elements of Periodicity in the Joint Density Results.....	137
5.1.1 Presence of periodic processes in results associated with real time streaming protocols .....	138
5.1.2 Principle behind the presence of the Curve of Periodicity .....	139
5.2 Using Markov Models to model the periodic elements .....	140
5.2.1 Presence of periodic processes in results associated with real time streaming protocols .....	140
5.2.2 Setting up the Markov Monte Carlo Simulation Process for modelling Periodic events .....	141
5.3 Analysis of results of Monte Carlo Simulation of Markov models representing periodic processes.....	142
5.3.1 Case 1: Small values of $\mu$ , with varying $\Delta$ and fixed $\sigma$ .....	143
5.3.2 Case 2: Impact of Varying $\sigma$ of Models ‘C’ and ‘D’ .....	147
5.3.3 Case 3: Choosing $\mu_1$ and $\mu_2$ far away from the origin .....	150
5.3.4 Summary of Analysis of Results of Monte Carlo simulations .....	153
5.4 Confirming the presence of only two periodic processes associated with a curve of periodicity .....	153
5.4.1 Test Case 1: Introducing one additional periodic process to the Monte Carlo Simulation. ....	154
5.4.2 Test Case 2: Including two additional periodic processes to the Monte Carlo Simulation .....	155
5.4.3 Conclusion on Test results.....	157
5.5 Relating the modelled Joint Density results to the measured Joint Density results .....	158
5.5.1 Example 1: UDP Streaming Traffic Results .....	159
5.5.2 Example 2: Voice/Video streaming over IM clients .....	160
5.6 Phased nature of IP Traffic Transmission .....	161
5.7 Summary .....	162

<b>Chapter Six .....</b>	164
Analysis of Voice over IP traffic from a multi-lingual VoIP call database .....	164
6.1 Background on Voice Traffic Modelling.....	164
6.2 Measuring the $T_{ON}$ and $T_{OFF}$ periods for VoIP traffic .....	167
6.2.1 The VoIP Call database .....	167
6.2.2 The Silence Detection Process .....	167
6.3 Analysis of $T_{ON}$ and $T_{OFF}$ periods and the impact of language, speech rate and caller behaviour .....	173
6.4 Summary.....	178
 <b>Chapter Seven .....</b>	179
An ON-OFF VoIP Source Model with Lognormal Transitions .....	179
7.1 Analysis of Composite Statistical Samples of the $T_{ON}$ and $T_{OFF}$ periods.....	179
7.1.1 Logarithmic Analysis and Modelling of $T_{OFF}$ periods.....	180
7.1.2 Logarithmic Analysis and Modelling of $T_{ON}$ periods .....	183
7.2 The ON-OFF Source Model .....	185
7.3 Simulated Results and Analysis .....	186
7.3.1 Analysis of simulated results of $T_{OFF}$ periods.....	187
7.3.2 Analysis of simulated results of $T_{ON}$ periods .....	187
7.3.3 Conclusion on the measured and simulated results of the $T_{ON}$ and $T_{OFF}$ periods .....	188
7.4 Summary.....	189
 <b>Chapter Eight.....</b>	191
Conclusions and Future Work .....	191
8.1 Conclusions .....	191
8.1.1 Analysis of IPT statistics based on simple Markov traffic models with Poisson and non-Poisson statistics .....	191
8.1.2 Use of Gaussian Markov Models for modelling traffic.....	192
8.1.3 Significance of the Joint Density results for understanding and modelling network IPT statistics .....	192
8.1.4 Markov modelling with Joint Density results: A powerful tool for analysis and an alternative to Deep Packet Inspection (DPI).....	194

8.1.5	Impact of languages, speech rate and prosodic structure on the ON-OFF characteristics of Voice traffic over IP networks .....	195
8.1.6	An ON-OFF model with Log-normal transitions and its significance .....	195
8.2	Future Work.....	196
8.2.1	The fourth dimension: Packet Sizes (PS) .....	196
8.2.2	Applying the use of Joint Density results and Markov models to model and understand specific networking protocols .....	197
<b>List of Publications Related to the Thesis .....</b>		199
<b>List of References.....</b>		200

# List of Tables

Table 3.1: Measurement Scenario 1 – TCP Port 80 – IP Radio Broadcast.....	93
Table 3.2: Measurement Scenario 2 – UDP Port 80 – IP Tv Broadcast .....	95
Table 3.3: Measurement Scenario 3 – UDP Port 15010 – IP Telephony Traffic .....	97
Table 7.1: Estimated values of parameters for Lognormal densities for $T_{OFF}$ periods.....	183
Table 7.2: Estimated values of parameters for Lognormal densities for $T_{ON}$ periods.....	186

# List of Figures

Figure 1.1: The IP Protocol with reference to the OSI and TCP/IP models .....	19
Figure 1.2: Encapsulation of an IP Packet .....	20
Figure 1.3: Example of a Simple Two State Markov Traffic Model.....	27
Figure 1.4: Example of ON-OFF model for Voice Packet traffic.....	28
Figure 1.5: Structure of Thesis with chapter inter-dependencies .....	34
Figure 2.1: State Transition Diagram of a Three State Markov Model .....	37
Figure 2.2: A simple two state Packet Traffic Model .....	42
Figure 2.3: State Transition Diagram of a Symmetric Two-State Markov Traffic Model .....	46
Figure 2.4: PDF of Network IPT times based on the symmetric case of 2-state Markov Model.....	48
Figure 2.5: Joint Density Plot for 2- state Markov Model (Symmetric Case) – Analytical result.....	50
Figure 2.6: Joint Density Plot for a 2- state Markov Model (Symmetric Case) – Monte Carlo result...	51
Figure 2.7: State Transition Diagram of a Two State Markov Model with Asymmetric Rates .....	51
Figure 2.8: PDF for Network IPT times for a Two - State Markov Model (Asymmetric Case).....	53
Figure 2.9: Joint Density Plot for 2- state Markov Model (Asymmetric Case) – Analytical result.....	54
Figure 2.10: Joint Density Plot for 2- state Markov Model (Asymmetric Case) – Monte Carlo results	54
Figure 2.11: State Transition Diagram for a Three - State Markov Model - Symmetric Case .....	55
Figure 2.12: Calculating State visiting probabilities – Three State Markov Model.....	57
Figure 2.13: PDF for Network IPT for a 3-state Markov Model (Symmetric Case) .....	58
Figure 2.14: State Transition Diagram for 3-state Markov Model with Asymmetric Rates .....	59
Figure 2.15: Histogram of Counts of visits to states for 3-state Markov model simulation .....	60
Figure 2.16: PDF for IPT for a 3-state Markov Model (Asymmetric Case) .....	61
Figure 2.17: State Transition diagram for (a) Three-state Markov Model with all states emitting packets (b)Four-state Loop model (c) Four-state Loop model with diagonal transitions (d) Five state Markov Model and (e) Seven state Markov Model .....	62
Figure 2.18: State Transition Diagram of Four-State Markov Model with diagonal transitions .....	63
Figure 2.19: Calculating State Transition Probabilities for a Four-state Model with diagonal transitions.....	63
Figure 2.20: Histogram of counts of visits to states 1,2,3 and 4 by 4-state Monte Carlo Simulation ..	65
Figure 2.21: PDF Results for (a) 3-state (All branch symmetric) and (b) 5-state (Asymmetric Model) ..	66
Figure 2.22: 2- state Markov Model with only one state emitting packets (Symmetric) .....	68
Figure 2.23: PDF for IPT for 2-state Model (Symmetric Case) with only state 1 emitting packets.....	69
Figure 2.24: JDF for 2-state Model (one state emitting packet) (Symmetric Case) - Analytical result.	70
Figure 2.25: JDF for 2-state Model (one state emitting packet) (Symmetric Case) – Monte Carlo result .....	70
Figure 2.26: 2- state Markov Model with only one state emitting packets (Asymmetric case).....	71
Figure 2.27: PDF for IPT for 2-state model (Asymmetric Case) - with only state 1 emitting packets...	71
Figure 2.28: JDF for 2-state Model (one state emitting packet) (Asymmetric Case) – Analytical result.....	72
Figure 2.29: JDF for IPT 2-state Model (one state emitting packet) (Asymmetric Case) – Monte Carlo .....	73
Figure 2.30: (a) Three-state and (b) Four-state loop Markov model with only state 1 emitting packets.....	74

Figure 2.31: PDF (Linear scale) for IPT for ( $n < 11$ ) states loop Markov Models - With only state 1 emitting packets.....	77
Figure 2.32: 30-state Markov Model with non-Poisson statistics .....	79
Figure 2.33: Simplified 3-state Gamma Markov Model.....	80
Figure 2.34: Comparison of Monte Carlo and analytical results for 30-state and 3-state Gamma Model .....	81
Figure 2.35: State Transition Diagram of a 4-state Gamma Markov Model .....	82
Figure 2.36: PDF for IPT of 4-state Gamma Markov Model .....	83
Figure 2.37: Comparison of a Gamma distribution with ' $n = 500$ ' with a Gaussian distribution. ....	85
Figure 2.38: PDF for IPT results for the 4-state Gaussian Markov model.....	87
Figure 3.1: Traditional approach to Internet traffic measurement .....	89
Figure 3.2: Revised Approach to IP traffic measurement .....	90
Figure 3.3: Probability Density Function of IPT for IP Packets on TCP Port 80 – IP Radio Broadcast ...	93
Figure 3.4: Joint Density Function of IPT for IP packets on TCP Port 80 – IP Radio Broadcast .....	93
Figure 3.5: Probability Density Function of IPT for IP Packet Traffic on UDP Port 80 – IP Video Broadcast.....	94
Figure 3.6: Joint Density Function for IPT for IP packets on UDP Port 80 – IP TV Broadcast.....	95
Figure 3.7: Probability Density Function of IPT for IP Packets on UDP Port 15010 – IP Telephony ....	96
Figure 3.8: Joint Density Function of IPT for IP Packets on UDP Port 15010 – IP Telephony.....	96
Figure 3.9: a) A Gaussian peak buried within another Gaussian peak b) Two fitted Gaussian components .....	98
Figure 3.10: PDF for IPT for Measured Results for Port 80 TCP for IP Radio Traffic. ....	101
Figure 3.11: PDF and CDF for IPT for Measured results on TCP Port 80- IP Radio .....	102
Figure 3.12: A 5 + 1 State Gaussian Markov Model to model measured results for IPT of IP Radio packets.....	105
Figure 3.13: PDF Plots of modelled, analytical and measured results of IPT for IP Radio traffic .....	106
Figure 3.14: PDF and CDF for IPT for Measured results on UDP Port 80- IP Tv Packet Traffic.....	107
Figure 3.15: PDF Plots of modelled, analytical and measured results of IPT for IP TV traffic .....	108
Figure 3.16: PDF and CDF for IPT for Measured results on UDP Port 80- IP Voice/Video Telephony	109
Figure 3.17: PDF Plots of modelled, analytical and measured results of IPT for IP Telephony Traffic	110
Figure 3.18: Comparison of CDF plots for measured and modelled results for IPT of IP Telephony Traffic. ....	111
Figure 4.1: PDF Plots of modelled, analytical and measured results of IPT for IP Telephony Traffic ..	113
Figure 4.2: Joint Density Plot of IPT of Gaussian Markov model results of IPT for IP Telephony Traffic .....	113
Figure 4.3: Joint Density Function for IPT for IP Packets on UDP Port 15010 – IP Telephony .....	114
Figure 4.4: Sketch showing association of IPT Gaussian peaks in Joint density (sub-set) with packet sequences. ....	115
Figure 4.5: Markov Model ‘A’ with packets A,B and C in random order.....	117
Figure 4.6: PDF plot of IPT for Markov Model ‘A’ with packets in Random Order.....	118
Figure 4.7: JDF plot of IPT for Markov Model ‘A’ with packets in Random Order.....	119
Figure 4.8: Markov Model ‘B’ with packets A,B and C in sequence.....	122
Figure 4.9: PDF plot of IPT for Markov Model ‘B’ with packets in sequence .....	123
Figure 4.10: Comparison of PDFs of IPT for Markov Models ‘A’ and ‘B’ .....	123
Figure 4.11: JDF plot of IPT for Markov Model ‘B’ with packets in sequence .....	124

Figure 4.12: The PDF is a Projection of the Joint Densities of the two Markov Models 'A' and 'B' ....	126
Figure 4.13: Asymmetric structure of Joint density Result for IPT of Markov Model 'A'.....	128
Figure 4.14: Symmetric Structure of Joint density result for IPT of Markov Model 'A' .....	129
Figure 4.15: Asymmetric structure of Joint Density Result for IPT of Markov Model 'B' .....	129
Figure 4.16: Symmetric structure of JDF of IPT of Markov Model 'B'.....	130
Figure 4.17: Symmetric Structure of Joint Density Results of IPT at Port 15010 .....	130
Figure 4.18: Asymmetric Structure of the Joint Density result for IPT at Port 15010.....	131
Figure 4.19: Illustration of calculating IPTs in the Markov Monte Carlo Simulation with and without Packet Loss.....	132
Figure 4.20: Joint Density Results for Markov Model 'B' with 10% packet loss.....	133
Figure 4.21: Joint Density Results (Contour version) for Markov Model 'B' with 90% packet loss.....	134
Figure 5.1: Joint Density Result of IPT for IP telephony traffic with a curve of periodicity highlighted.....	137
Figure 5.2: Joint Density Result of IPT for a UDP traffic stream with curves of periodicity highlighted.....	137
Figure 5.3: Joint Density Results of IPT for traffic of (a) Windows Live Messenger client measured at Port 20012 and (b) Yahoo Messenger client measured at Port 44579, with curve of periodicity indicated by black curves.....	138
Figure 5.4: Explanation of the Curve of Periodicity in an example of a Joint Density result .....	139
Figure 5.5: Markov Models 'C' and 'D' for modeling Periodic Processes.....	141
Figure 5.6: Markov Monte Carlo Setup for Modelling Periodic Events.....	142
Figure 5.7: PDF for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.645$ and $\Delta = 0.2$ .....	143
Figure 5.8: JDF (Linear version) for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.645$ and $\Delta = 0.2$ .....	144
Figure 5.9: JDF (Logarithmic version) for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.645$ and $\Delta = 0.2$ .....	144
Figure 5.10: PDF for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 3.7$ , $\sigma = 0.645$ and $\Delta = 1.2$ .....	145
Figure 5.11: JDF for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 3.7$ , $\sigma = 0.645$ and $\Delta = 1.2$ .....	145
Figure 5.12: PDF for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 9.7$ , $\sigma = 0.645$ and $\Delta = 7.2$ .....	146
Figure 5.13: JDF (Three dimensional Linear) for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 9.7$ , $\sigma = 0.645$ and $\Delta = 7.2$ .....	146
Figure 5.14: JDF (Logarithmic) IPT, of $\mu_1 = 2.5$ , $\mu_2 = 9.7$ , $\sigma = 0.645$ and $\Delta = 7.2$ .....	147
Figure 5.15: PDF for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.1$ and $\Delta = 0.2$ .....	148
Figure 5.16- JDF (Linear X and Y axes; three dimensional) for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.1$ and $\Delta = 0.2$ .....	148
Figure 5.17: JDF (Logarithmic X and Y axes, two dimensional) for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.1$ , $\Delta = 0.2$ .....	148
Figure 5.18: JDF (Logarithmic X and Y axes, two dimensional) for IPT, of $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.01$ , $\Delta = 0.2$ .....	149
Figure 5.19: Joint Density Function (Three dimensional) – Poisson like statistics for large $\sigma$ .....	150
Figure 5.20: JDF for IPT, of $\mu_1 = 9.5$ , $\mu_2 = 9.7$ , $\sigma = 0.645$ and $\Delta = 0.2$ .....	150
Figure 5.21 - JDF for IPT, of $\mu_1 = 9.5$ , $\mu_2 = 19.7$ , $\sigma = 0.645$ and $\Delta = 10.2$ .....	151
Figure 5.22: JDF for IPT, of $\mu_1 = 20.5$ , $\mu_2 = 100.7$ , $\sigma = 0.645$ and $\Delta = 80.2$ .....	152
Figure 5.23: PDF for IPT, of $\mu_1 = 20.5$ , $\mu_2 = 100.7$ , $\sigma = 0.645$ and $\Delta = 80.2$ .....	152
Figure 5.24: PDF for 3 periodic Processes with $\mu_1 = 2.5$ , $\mu_2 = 3.0$ , $\mu_3 = 3.5$ , $\sigma = 0.645$ and $\Delta = 0.5$ .....	154
Figure 5.25: JDF (3d Linear) for 3 periodic Processes, $\mu_1 = 2.5$ , $\mu_2 = 3.0$ and $\mu_3 = 3.5$ , $\Delta = 0.5$ .....	154
Figure 5.26: JDF (2d Logarithmic) for 3 periodic Processes, $\mu_1 = 2.5$ , $\mu_2 = 3.0$ and $\mu_3 = 3.5$ , $\Delta = 0.5$ .....	155
Figure 5.27: PDF for 4 periodic Processes with $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5$ , $\sigma = 0.645$ and $\Delta = 0.5$ .....	156

Figure 5.28: JDF (Linear scale) for 4 periodic Processes with $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5, \sigma = 0.645$ and $\Delta = 0.5$	156
Figure 5.29: JDF (2d Logarithmic) for 4 periodic Processes, $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5$ .....	157
Figure 5.30: Poisson statistics (Joint density) for IPT of Two-state Symmetric Markov Model from Chapter 2 (Fig.2.6).....	158
Figure 5.31: (a) The $S\Delta$ Model and (b) The $L\Delta$ Model .....	159
Figure 5.32: Example Result showing how two statistical features of the Joint Density of IPT have been captured by the Markov Monte Carlo simulation of the $L\Delta$ and the $S\Delta$ models ...	159
Figure 5.33: Example Results of Joint Density statistics of traffic on Windows Live Messenger and Yahoo Messenger IM clients showing how their statistical features have been captured by the Markov Model sets ' $L\Delta$ ' and ' $S\Delta$ ' representing periodic processes with large and small $\Delta$ respectively.....	160
Figure 5.34: Joint Density Plot of UDP Traffic for only the first 30% of the captured packets.....	161
Figure 6.1: The 8-state ON-OFF model proposed by Stern et al.....	166
Figure 6.2: Example of waveform of extract of a speech signal $S(t)$ with the $T_{ON}$ and $T_{OFF}$ periods shown .....	168
Figure 6.3: Zoomed section of a waveform of extract of a speech signal $S(t)$ showing a small $T_{OFF}$ period .....	169
Figure 6.4: An illustration of the Overlap Add method used for Low pass filtering the speech signal	170
Figure 6.5: The Acoustic Envelope $m(t)$ of a speech signal $S(t)$ .....	171
Figure 6.6: Impact of change in value of $f_c$ on the $T_{OFF}$ periods .....	172
Figure 6.7: Impact of change in value of TCV on the $T_{OFF}$ periods .....	173
Figure 6.8: Comparison of monolog silence lengths ( $T_{OFF}$ periods) for Chinese, French and English Languages .....	174
Figure 6.9: Cumulative distribution of silence lengths ( $T_{OFF}$ periods) for various language samples of two-way VoIP calls in contrast to measurement by [81] .....	175
Figure 6.10: Cumulative distribution of silence lengths ( $T_{OFF}$ periods) for eight different calls for the Spanish language.....	176
Figure 6.11: Cumulative distribution of silence lengths ( $T_{ON}$ periods) for call samples for different languages.....	177
Figure 7.1: Tri-modal Gaussian distribution fit to Probability Density of Composite Sample of $T_{OFF}$ periods.....	181
Figure 7.2: Confirming the heavy-tailed nature of the $T_{OFF}$ periods .....	183
Figure 7.3: Bi-modal Gaussian distribution fit to Probability Density of Composite Sample of $T_{ON}$ periods.....	184
Figure 7.4: The ON-OFF VoIP source Markov model with Log-normal transitions.....	186
Figure 7.5 : Cumulative distributions for $T_{OFF}$ periods - Measured values (Composite) vs. Our Log-normal model and results from [81] and [79] .....	187
Figure 7.6: Cumulative distributions for $T_{ON}$ periods - Measured values (Composite) vs. Our Log-normal model and results from [81] and [79] .....	188
Figure 7.7: Cumulative distributions for $T_{ON}$ periods – with a composite sample excluding silences less than 10 ms.....	189

## **PART – I**

# **Modelling IP Traffic using simple Markov Traffic Models**

# Chapter One

## Introduction

### 1.1 Background

#### 1.1.1 The Internet

The *Internet* is a global system of inter-networked computer systems based on a unique addressing system monitored by the Internet Protocol (IP) and associated extensions. From its origins in the 1960s, as a distributed computer network built for military research projects of the United States, the internet has come a long way and is now a revolutionary medium that carries massive amount of information and provides essential services such as E-mail, IP Telephony, Audio and Video Streaming, File transfers and Data Services. The first explosive growth of the internet was witnessed in the 1990s with the inception of the World Wide Web (WWW). Since then, the number of users over the internet has been growing exponentially fast. It was reported that by mid - 2009, about 1.6 billion people have been using the internet [1] . This accounts for almost a quarter of the world's population. The continuous increase in the number of Internet users also means that there is a continuous increase in the Internet traffic generated by the users. In order to accommodate this growing traffic, it becomes a necessity to measure, analyse and model this traffic in order to build provisions for enhancing the performance and capacity of the infrastructure associated with the internet. This thesis focuses on the various aspects of statistical analysis and modelling of packet traffic associated with data and voice services over the internet.

#### 1.1.2 Packet Switched Networks

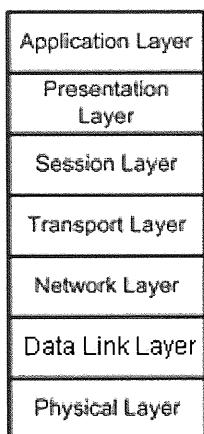
The concept of Packet Switching in networks originated in the early 1960s and it served as a base for building the internet [2]. The fundamental idea was that each data message would be divided into a group of packets which are sent through a shared network of communication nodes to a destination determined by a unique destination address. The Transmission Control Protocol / Internet Protocol (TCP/IP) Suite otherwise known as the *Internet Protocol Suite* is the best known example based on a

Packet Switched Network. The Internet Protocol (IP) looks after the addressing and control information of the data packets, while the Transmission Control Protocol (TCP) ensures reliable delivery of these packets to their destinations. The packet switched network contrasts with the *circuit switched networks* where a dedicated circuit (often an electric circuit) is established in advance, between the two communicating nodes for the duration of the transmission of the message, after which it is released. In a packet switched network, the transfer of packets is through a connectionless medium where bandwidth is acquired and released as and when necessary for the transfer. The packets may be routed across any route in the network to get to their destination. Packet Switched networks are superior to circuit switched networks because they offer several advantages, such as robustness of communications, high utilization level, optimized use of bandwidth, and efficient transmission of burst mode data.

### 1.1.3 The Internet Protocol (IP)

The Internet Protocol (IP) is one of the most primary and vital elements of the Internet. It was formally developed and defined in 1981 for use in Packet Switched Networks [3]. This protocol belongs to the Network Layer and the Internet Layer of the OSI Reference Model (Open System Interconnection Reference Model) [4] and the TCP/IP Suite [5] respectively as shown in Fig. 1.1. The primary responsibility of the Internet Protocol is to route the packets with the use of a logical addressing system which assigns the source and destination addresses. In other words, it actually tells the data packets where to go in the network.

OSI Model



TCP/IP Model

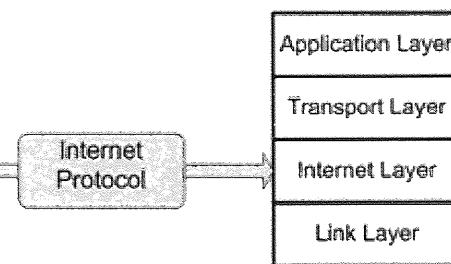


Figure 1.1: The IP Protocol with reference to the OSI and TCP/IP models

The Internet Protocol encapsulates the data from the upper layers of the models into an IP packet as shown in Fig. 1.2. An IP packet has two principal blocks: the block of the data for transmission received from the upper layers, and the IP blocks. The typical size of an IP packet is 1500 bytes while the header is of 20 bytes. The IP header contains controlling instructions related to the data block being transmitted. This typically includes the source and destination IP addresses, the number of packets the message is divided into, identification of the packets, information on the protocol associated with the packet data (i.e. TCP, UDP, and ICMP), packet lengths, checksums and other control information associated with the data.

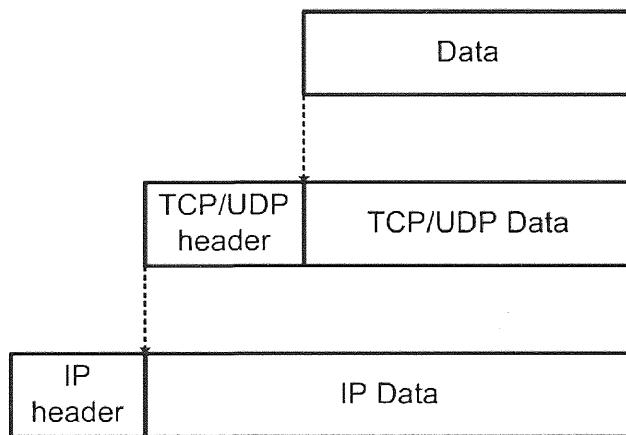


Figure 1.2: Encapsulation of an IP Packet

Because of the abstract nature of the encapsulation process of the Internet Protocol, it can be used over a *heterogeneous network*, which consists of a group of networks with common network TCP/IP connections and different Physical and Medium Access protocols. When taken a step further, this concept will form the basis of a Next Generation Network or an ‘All IP’ network, which is essentially one common network that will provide multiple data and media services, the raw data of all of which is encapsulated into IP packets [6]. Thus, the IP packets are the most important and fundamental units in the design, operation and performance analysis of the internet. Throughout this thesis we will focus on IP Packet level network traffic modelling with the use of advanced and appropriate statistical tools and probabilistic models.

### 1.1.4 Traffic Modelling

As networks evolve to support various types of services related to data, voice and multimedia, the traffic generated associated with these applications carries different statistical properties unique to

each type of application or service. In addition to this, the increasing geographical, behaviour dependent and time varying complexity of the network traffic makes it an absolute necessity for the network service providers and the research community to carefully study these traffic characteristics either collectively or service dependently. This helps in solving technical challenges related to performance and capacity enhancement of the networks, and also enables to emulate or simulate the traffic on the network for further analytical purposes.

Traffic Modelling and analysis is often performed with the aid of mathematical models which can be setup for simulations. In order to achieve reliable, realistic and valid modelling data, these mathematical models used must capture as many features of the network from the measured results for efficient modelling. The traffic modelling process usually consists of three main steps which are discussed below:

- a) **Measured Data:** In order to propose a unique mathematical model associated with a particular type of traffic stream or network packet events, we often need unique measured statistical results which suggest a degree of anomaly with previously recorded results or understanding of network traffic of that particular kind. The two most common parameters used for modelling traffic are the Inter-Packet times (IPT) and the Packet Sizes (PS). In our work, all our statistics are limited to the Inter-Packet times of the traffic streams measured. The Inter-Packet time is defined as the difference between the recorded absolute time stamps of two successive packets received by the machine where the measurement takes place. The measurement procedure used for obtaining statistical results of the IPT related to the various traffic streams and types considered is covered in detail in Chapter 3.
- b) **Proposing the Model:** The measured statistical results are then used to obtain information on the timing structure of the packet events and this gives several clues and hints about the behaviour of that type of traffic on the internet. Using this information, a model can be fitted to the measured data, with each component of the model relating to a particular element observed in the measured statistical results. After the underlying structure of the model is formed, the parameters of each significant component of the model are estimated with the aid of measured results and sophisticated numerical-estimation algorithms.
- c) **Simulating the Model:** Once the model is mathematically ready, it can be immediately implemented for simulation. The simulation is iterated for several times until the desired amount of statistics is achieved. Such a process is called a *Monte-Carlo* Simulation. Results

from the simulation are obtained in the same manner as that with the measured results.

Finally, the simulated results are then compared with the measured results to validate the proposed model. The deviation between the measured results and the simulated results helps to analyse how well the model fits the data. Based on this, conclusions can be drawn on the statistical properties and behaviour of the traffic measured.

Most of the mathematical models that are used in the traffic modelling of packet switched networks are based on the network layer or the internet layer and do not necessarily rely on any other information from any of the upper layer protocols. In the next section, we will briefly look at the most commonly used approaches for network traffic modelling.

## 1.2 Common Approaches to Traffic Modelling

### 1.2.1 History

The most commonly used mathematical methods for Traffic Modelling are based on the use of stochastic models where the components of these models are based on some unique probability distributions and involve the use of random number generation processes.

#### 1.2.1.1 Poisson Model

The first attempt on telecommunication traffic modelling was done by a Danish mathematician by the name of A.K. Erlang in 1909 [7] and his pioneering work had modelled the traffic related to telephone conversations by a Poisson model. The Poisson model follows the Poisson process which defines that the inter-arrival times of events are exponentially distributed. Since the very beginning, Poisson models have been dominantly used for analysing traffic on telecommunication networks [8] [9] [10]. The model was based on the assumptions that the number of traffic sources were infinite and that the traffic followed a random pattern. Poisson models have also been used in the modelling of network queuing systems and were viewed as a role model in bringing the fields of mathematical modelling and traffic management together.

### 1.2.1.2 Pareto Model

However, during the late 80s and the 90s, it was established that packet arrivals in a network were not necessarily Poisson [9] [11] [12]. Poisson Models were also deemed inefficient when it came to modelling burst mode traffic [13] [14]. Subsequent research work established the *Self-Similar* and *Heavy tailed* nature of the Packet traffic [15] [16]. This paved way for the use of Pareto Models to model traffic that has self-similar arrivals [17]. The Pareto distribution is a two-parameter distribution that obeys the *power law* and produces independent and identically distributed Inter-Arrival times [15]. It has now become a standard model to be used in modelling traffic which shows non-exponential, heavy tailed and self-similar characteristics.

### 1.2.1.3 Weibull Model

The Weibull Model is another type of traffic model based on the Weibull Process which has been recently used to model long tail characteristics especially for traffic related to multimedia content [15] [18]. The Weibull traffic model offers the advantage of modelling heavy-tailed and self-similar traffic by multiplexing the sources of the active and inactive periods for the network traffic [19].

So far we have looked at models which are based on a particular probability distribution for the packet traffic arrival and inter-packet times (IPT). The other most commonly used type of traffic models are based on Markov Processes. These are stochastic processes based on the Markov Property which states that the future event depends only on the current event and is independent of the past events. In other words these are memory less systems. The models based on the Markov Processes are used to model the activities of network traffic sources with a finitely limited number of states. In the following sections we will deal with the broad classification of the two types of categories of Markov Traffic Models with a brief summary on how they work. A detailed mathematical background on the use of classic packet source Markov traffic models for modelling packet traffic is covered in Chapter 2. Markov Traffic Models can be generally classified into *Black Box* and *White Box* Models. Some of the popular types of Markov based models belonging to these two types of categories are covered in the following two sections.

## 1.2.2 Long range dependence of traffic

One key area of research, particularly in the late 90's was on the long range dependence of traffic statistics. Long-range dependence is associated with the phenomenon of observing slower decay of correlations of measured data at different time scales, in particular longer time scales, in the time series analysis of traffic. This phenomenon suggests that the rate of statistical dependence is slower than the traditional exponential decays observed in Poisson processes [20]. The seminal study of Leland et al [21], proved the failure of Poisson models in modelling traffic, introducing the presence of long range dependence and self similarity in internet traffic. Since then several contributions have reported long range dependence in traffic [22] [23] [24] [14] [25] [26] [27]. While it has been established that the Internet traffic statistics does exhibit long-range dependence, some recent work such as [28] established the complex behaviour of network traffic making it tedious for the correct analysis of dependency and its associated estimation processes. Long-range dependence is often an integral part of averaged statistics with regards to Internet traffic modelling. However, in this thesis we focus only on short term statistics as we intend to model and identify protocol behaviour and its associated traffic patterns based on individual connection flows, as will be discussed in detail in Chapter 3.

## 1.2.3 Black Box approaches

A *black box* model is a type of mathematical model whose internal structure is often unknown or is *opaque* in terms of the statistical activity being processed within the model. Also, often for these kinds of models there is very little amount of *a priori* information available as input. Sections 1.2.2.1 and 1.2.2.2 cover two widely used examples of such types of models for modelling packet traffic.

### 1.2.3.1 Markov Modulated Poisson Processes (MMPP)

The *Markov Modulated Poisson Process* is a special type *Markovian Arrival Process* and is a doubly stochastic process. It consists of 'N' number of finite super positioned states acting as packet sources as part of a Markov Process which spends an exponential amount of time in each state. In addition it

also consists of a Poisson Process that generates arrivals with a rate that is dependent on the current state of the Markov Process. In other words, the Poisson Process is modulated by the Markov Process, hence the name. The Markov Process serves as an auxiliary process to the Poisson Process. These Markov Modulated Poisson Processes (MMPP) are black box models as their internal structure is not explicitly modelled. They have been widely used in modelling traffic, especially of the *bursty* nature [15] [29] [30] [31] [32]. A variety of other models based on the Markov Modulated Processes such as *Batch Markovian Arrival Process* (BMAP), *Markov Modulated Rate Process* (MMRP), *Markov Modulated Bernoulli Process* (MMBP) and *Markov Modulated Fluid Models*(MMFM) have also been used for traffic modelling. For a detailed explanation on these types of models, one can refer to [33].

### 1.2.3.2 Hidden Markov Models (HMM)

Another type of Markov Process based traffic models that have a black box approach are *Hidden Markov Models* (HMM). These are Markov processes where the states or at least some of the states or events are not observable and are *hidden*. However the output associated with these states is observable. A MMPP can also be viewed as a special case of HMM, where the associated arrival processes  $A(t)$  are observable and the underlying Continuous-time Markov Chains are hidden [34]. There are three main sets of parameters of the HMM that are related to the observed element sequence  $M$  and the number of hidden states  $N$ , and they are: i) an initial probability distribution of events, ii) the state transition matrix and iii) a final output distribution. Each of the observable values is associated with a probability distribution of a particular state of the model. The sequence of observed events along with the model parameters can be used to estimate the sequence of state transitions and output probabilities that have most likely created the sequence of observed events. This is often done by using a sophisticated algorithm called the *Baum-Welch Algorithm*, a special and efficient type of an Expectation-Maximization algorithm [35]. The estimation process does require however, a good amount of prior knowledge on the structure of the proposed model based on ideas of the application being modelled.

From a network point of view, a HMM can be used to predict or trace protocol behaviour from the observed statistics of network *events*, leading to a possibility of simulating a particular type of traffic based on measurements from traffic streams. The *events* could be related to sequences of packet sizes or the inter-packet times. Hidden Markov Models have also been recently used for modelling

packet level traffic [36] [37] [38]. One advantage of using the HMMs is that these models can be trained iteratively to obtain model parameters that are closest to the observed elements and this training process can be altered to better suit the requirements of modelling. However it does have disadvantages as well. For instance, it is possible that the limited amount of prior knowledge we have based on our expectations about the underlying protocol may not be adequate. Also, the model training may not converge with the data very well. Even when a proposed model fits data, it is difficult to judge source behaviour of a state with respect to other states, because of the ‘hidden’ factor. Besides they are mathematically and computationally more complicated than conventional Markov and most other modelling methods.

### 1.2.4 White Box approaches

Models that adopt the *Black Box* approach are often aimed at merely reproducing the statistics of traffic with the use of a large number of parameters, most of which have no physical meaning in terms of the traffic process being modelled. The actual behaviour of the source contributing to the statistics is concealed by the ‘black box’ factor [39]. *White box models* on the other hand are models whose internal structure is transparent. The modelling parameters associated with the model also have a physical meaning to them and they help in understanding and relating the structure of the model to the network events and protocols that generate particular type of traffic statistics. It also brings confidence in predicting future events and evaluating similar statistical patterns based on the transparent knowledge of the structure of the model. Given below are two examples, which are more or less related to each other in terms of packet traffic generation, and which follow the white box model approach.

#### 1.2.4.1 Markov and Semi-Markov Monte Carlo Models

These models are based on the conventional Markov Processes and they have a finite and small number of states. Each state can be modelled after a packet source that emits packets or a packet train with a particular rate. The models follow the *Markov Property* and are based on a continuous time system. The inter-packet times for the Markov source models are exponentially distributed, which in other words means that the model spends an exponential amount of time in each state before it emits a packet and moves on to the next state. These models are simple to understand and

are very useful as analytical tools for statistical analysis. The properties of each state can be closely related to the properties of an actual traffic source in the network. The *Monte Carlo* process involves repeated sampling from various distributions associated with each state until the statistical results converge to the expected results or until the number of packets required to be simulated is reached. The disadvantage however is the model becomes increasingly complex with increase in number of states. A Semi-Markov Monte Carlo model is similar to the formerly discussed model but with the difference that it can have any distribution apart from the exponential distribution for the inter-packet times. The simplest Markov Monte Carlo Packet source model is a two-state Markov Model that is based on the inter-packet times that can be associated with a client-server model where a REQ packet is acknowledged with an ACK packet as shown in Fig. 1.3. These models are used throughout the thesis as important tools for analytical understanding, and modelling of traffic streams. Examples of Markov packet source models and their relation with simplified Semi-Markov Non-Poisson models and associated statistical analysis is covered in the chapters 2 and 3.

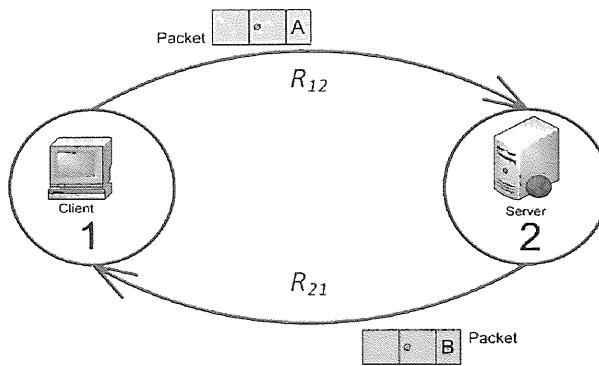


Figure 1.3: Example of a Simple Two State Markov Traffic Model

#### 1.2.4.2 ON-OFF Packet Source Model

The ON-OFF models were first introduced to capture the scaling behaviour of traffic and analysis of the structure of IP traffic. These models have either two or two sets of states that represent 'On' and 'Off' periods of traffic respectively. The model is often designed to alternate between the ON and OFF states. If the transitions are exponentially distributed, the model follows a Markov Process. For non-exponential transitions it follows a Semi-Markov Process. ON-OFF models have been popularly used with modelling Voice Traffic [40], where the periods of speech or talk spurts are modelled by the ON state and the periods of silences are modelled by the OFF state. Unlike the Markov Models, the On states and the Off states can both have different distributions. Fig. 1.4 shows an example of an ON-OFF model for Voice traffic. In our work, we propose an ON-OFF model for voice traffic in Chapter 7.

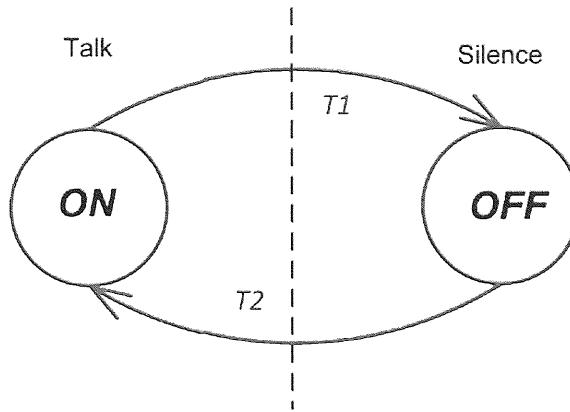


Figure 1.4: Example of ON-OFF model for Voice Packet traffic

## 1.3 Motivations and Objectives

The growth of internet traffic has been tremendous, and with it comes the need of efficient network management and maintaining the quality of service levels which the end user expects. As the Internet was designed on the basis of a connection less medium, it has evolved into a very massive and complex structure with ever increasing bounds that makes it difficult to control. To enable a systematic control mechanism, one must understand the behaviour of the network being monitored. This is where traffic analysis takes a centre stage and acts as a vital component in helping understand the features and capacity of the network. In the process of traffic analysis, one studies the pattern of communications and traffic flows and then with the aid of statistical analysis and protocol knowledge, proposes mathematical models that can be simulated, not just for testing and design purposes, but also for understanding the source behaviour. A good traffic model, as mentioned earlier, must be a one which captures the entire complexity of the measured traffic.

The motivations and associated objectives of this thesis, though revolving around a common theme on traffic modelling, have been based on two fronts: (1) Understanding and modelling of first order and second order statistics of IP Traffic flows, using Simple Markov and Semi-Markov models and (2) Analysis and modelling of Voice traffic over the internet based on measurements of a large multilingual database of VoIP calls.

### 1.3.1 Understanding and Modelling of First and Second order statistics of measured IP traffic flows using simple Markov Models

Traditionally, traffic on the telephone network has been modelled using simple Markov and Poisson traffic models based on elements of queuing theory. But with the advent of packet switched [28]

---

networks with ‘bursty’ data, these simple Markov and Poisson models, based on experimental results, were deemed unsuitable for modelling traffic characteristics of packet switched networks [14], and hence the use of simple Markov models fell out of favour. This eventually paved way for black box approaches for modelling traffic described in section 1.2.2. These models have the disadvantage, that while they can be trained into replicating captured traffic statistics, the physical reasons behind the statistics cannot be examined, because of the opaqueness of the models. Models based on white box approaches, as previously pointed out, are analytically and mathematically tractable and allow us to interpret the physical meaning of the traffic statistics with respect to the traffic source. Classic Markov models, as mentioned earlier, are a very good example of white box approach and apart from being analytically tractable, offer the advantage of providing the ability to derive mathematical expressions from it, that can be associated to the complete or part of the statistics observed that is being modelled. Even under the circumstances of not being able to achieve the mathematical and analytical tractability of these models, the models can be setup for Monte Carlo simulation and the simulated results can then be used for comparison with observed data and then judgements can be made on how the models can be modified further.

Another disadvantage of the currently popular black box modelling procedures is that they are based on first order statistics only, namely the probability density function of Inter-Packet times and Packet sizes. The Probability Density Functions for the Inter-packet statistics only tells one about the timing structure of the traffic, but does not give any additional information on the source behaviour. Consequently, we believe that by the use of additional statistical tools namely, the second order density plots, we can achieve much more information on the traffic, such as packet sequences, presence of periodicity and the packet source behaviour and topology based on measured Inter-packet times. This when coupled with the use of classic Markov models for advantages mentioned above, provides an excellent combination of tools required for building simple traffic models that capture the maximum amount of details.

Our main objectives on this front consist of first studying and analysing the first and second order density results of the measured IPT of traffic. We will then consider simple Markov Models with Poisson and non-Poisson statistics, perform a mathematical analysis on them and verify our understanding of these models with the numerical results from the Markov Monte Carlo Simulations. We will then look at fitting these simple models to first order density results followed by establishing

---

the meaning of the unique features of the IPT statistics observed in the Joint Density results and then explore the possibility of modelling each of these unique features with these Markov traffic models.

### **1.3.2 Modelling Voice traffic based on measurements from a multilingual database of VoIP calls**

The increasing popularity and commercial success of Voice over IP (VoIP) solutions has led to a steady rise in VoIP traffic over the past few years. As major network operators migrate from the traditional circuit switched networks to the converged IP network, VoIP traffic is poised to grow even further, constituting a large portion of global IP traffic [41]. This calls for accurate modelling of statistical properties of Voice traffic over IP networks for efficient bandwidth management. Often VoIP traffic has been modelled using ON-OFF source models, as described in section 1.2.3.2, where periods of speech activity, talk spurts are modelled as ON periods separated by silence lengths modelled as OFF periods. The voice packet transmission occurs only for the duration of the ON periods and the traffic sources are turned off during the OFF periods of a VoIP call, thus making the OFF periods in a VoIP traffic stream crucial for efficient traffic generation.

While several ON-OFF models have been proposed in the past, majority of these used an exponential distribution for the ON and OFF periods associated with the talk spurts and silence lengths. However recently, there have been hints that this may not be the case, and that the distributions are heavy tailed. However these results did not take into consider the statistically significant short silences (<200 ms). The inclusion of short silences in OFF period statistics is vital for efficient bandwidth management of Voice traffic. Apart from this, the voice database, commonly referred to for the measured ON-OFF statistics, is quite old and consist of analogue voice recordings of limited time durations based on the English language and did not consider other languages.

To this end, as part of our objectives for the work on Voice Traffic Modelling, we consider a large database of VoIP call recordings from multilingual speakers and using sophisticated FFT based filtering, make accurate measurements of the ON and OFF statistics, including gathering the most shortest silences possible, for different speakers of different languages. We then do an in-depth analysis on the differences in the ON and OFF period statistics for each language, and finally aim to

propose our own ON-OFF source model for Voice traffic. The traffic analysis done in this aspect is slightly different from the one done with the simple Markov models, as the traffic considered here is modelled from an Application layer point of view rather than the IP layer.

## 1.4 Contributions of the Thesis

This thesis has made the following contributions:

### **Modelling First and Second order density results of Inter-Packet times using Simple Markov models**

- We consider simple Markov models with Poisson statistics and derive the mathematical parameters for these Markov models and confirm the analytical results with the numerical results from the Markov Monte Carlo simulations. We then modify these results and present simple Markov models with non-Poisson statistics and again verify our analytical results with the simulated results. We then derive the Gamma Markov model from the n-state Markov model with non-Poisson statistics as a simplified model serving as a replacement to complex n-state models with non-Poisson statistics. We take the model derivation work further, by deriving convenient to use Gaussian Markov models from the Gamma Markov models and prove its legitimate use for packet traffic modelling with non-Poisson statistics.
- The use of Gaussian Markov models is then exploited to model the first order density results and demonstrate how these simple to estimate Gaussian Markov Models can be neatly fit to the measured first order density results.
- We establish the significance of the use of Joint Density results in measuring and modelling the IPT results. We compare the Joint density results of the model fitted to a first order density plot with that of the measured ones, and highlight that the modelled first order density results do not capture all the information that the Joint density results provide, thus proving the significance of the Joint Density results. We look at some of the unique features observed in the Joint Density results such as information on packet sequences and symmetrical nature of IPT data, and we model those Joint Density results using simple

Gaussian Markov Models. In another significantly important contribution we prove how two different Markov models with different topologies can have the same probability density function but yet completely different Joint Density results. We also demonstrate that the Probability density results are just a projection of the Joint Density result viewed from the X or Y axis of the Joint Density.

- We then also study in detail another unique feature of the Joint Density results of the measured IPT associated with the traffic. This feature is the ‘curve of periodicity’. Upon indicating the presence of the curve of periodicity in the measured Joint Density result, we establish the reason why this curve occurs, and again we use two simple Markov Models in a Monte Carlo simulation to model two periodic process, and we then completely model the joint density results of IP traffic using a combination of two sets of simple two-state Markov models, thus proving further that these simple Markov models can go a long way in modelling from some of the most simplest to some of the most complex and primitive features observed in the first and second order densities of the IPT of IP traffic.

### **Modelling Voice Traffic with reference to measurements from a large multilingual database of recorded VoIP calls**

- We consider a large multilingual database of over 100 hours of recorded VoIP calls with native speakers speaking in their languages, and we use a sophisticated FFT Low pass filtering based measurement technique to measure and isolate the silences and talk spurts in the recorded voice call files such that even the smallest silence and talk spurt is statistically accounted. We also consider a multilingual monolog database to study the short silence lengths for monolog bursts in VoIP calls. Using this database, we reveal that the incidence of short silence (<200 ms) is unique to each language and that this observation is independent of the samples chosen for the languages. More specifically, using the monolog database, we prove that the language dependent silence lengths are statistically significant and that voice models based on monolog speech, should therefore take the language into consideration as the silence length structure of the voice traffic greatly depends on the prosodic structure of the language.

- However, in our analysis of the two-way VoIP calls, we observe that the statistical differences in languages are overshadowed by the random dynamics of the interaction of speakers. In our measurements, we record an average of 30% more short silences than any of the previous studies that included short silences. We observe from the results that although there are differences in distributions of silence lengths in the voice call samples for each language, the frequent presence of non-grammatical silence lengths such as those related to listening, thinking and hesitation in the voice call samples, cause complex behaviour inconsistent with modelling of voice call silences on a linguistic or speech rate basis alone.
- We then form the composite statistics of the ON and OFF periods, and analyse them and since the human perception of speech time is logarithmic in nature, we do a logarithmic statistical fit to the silence and talk periods. This reveals a tri-modal and bi-modal Gaussian profile for the Silence and talk spurt periods respectively, thus signifying that the ON and OFF periods of the Voice over IP traffic are best modelled by a tri-modal and bi-modal Log-normal distribution. Using our knowledge gained from the understanding of Semi-Markov models and the Gaussian Markov model fitting, we propose a five-state Semi-Markov ON-OFF model with log-normal transitions based on the distributions associated with the composite ON and OFF periods.

## 1.5 Outline of the Thesis

This thesis starts with Chapter 1 offering a literature review and an introduction to the work done. In this chapter, we have seen a background on Internet traffic modelling along with the motivations and objectives of this thesis. The rest of thesis is then divided into two major blocks: The first block, consisting of chapters 2, 3, 4 and 5, deals with the modelling of IP traffic using simple Markov source models with the use of first order and second order joint density plots. The second block consists of chapters 6 and 7 which explain the analysis and modelling of VoIP traffic from a large multi-lingual database of VoIP call recordings. Finally the thesis concludes with Chapter 8. Fig. 1.5 schematically explains the structure of thesis with the sequence of the chapters and their interdependencies.

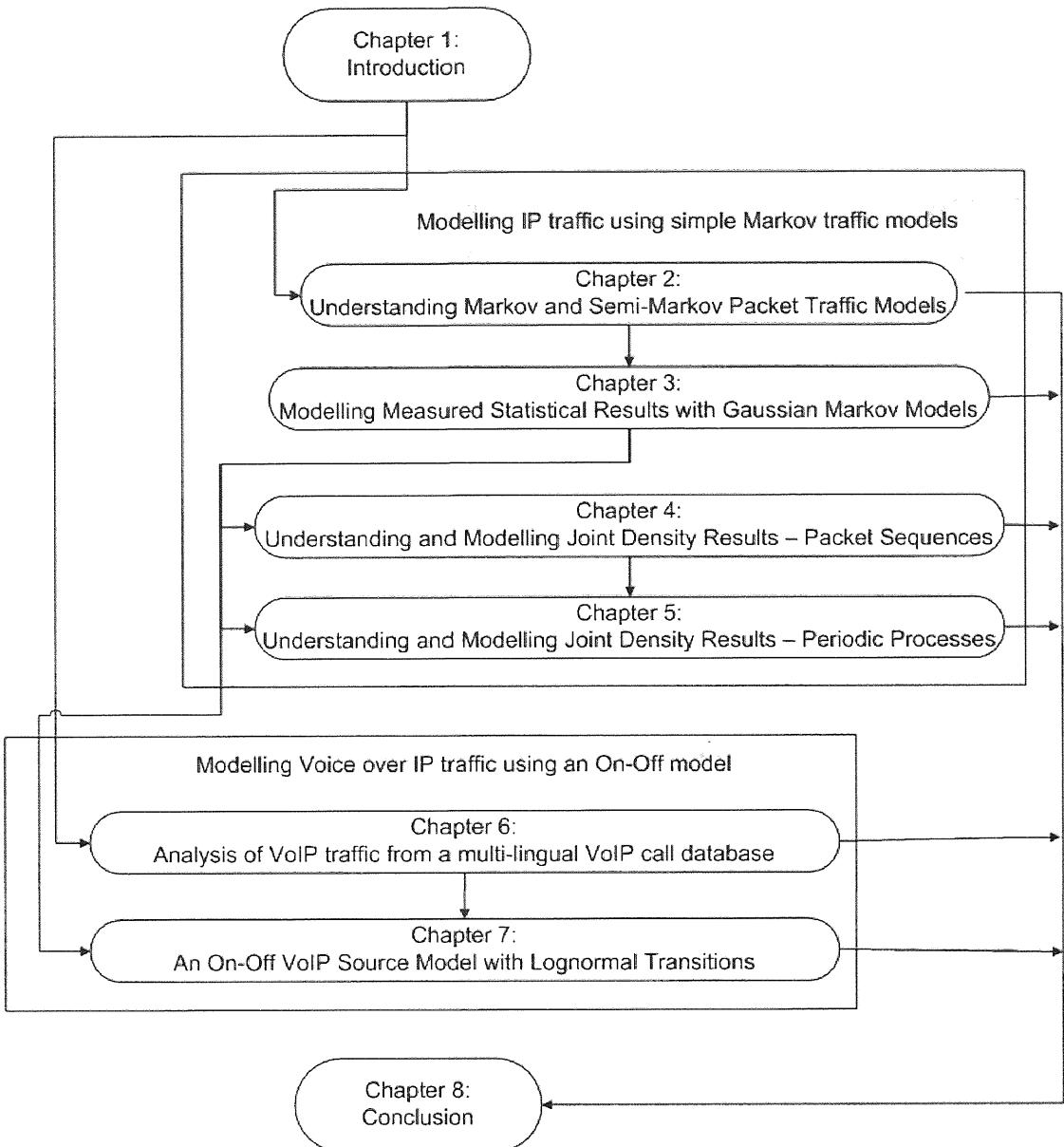


Figure 1.5: Structure of Thesis with chapter inter-dependencies

A brief introduction to the contents of each of the remaining chapters is as follows:

Chapter 2 provides a detailed mathematical background of Markov models, their elements, and the procedure of setting them up for Markov Monte Carlo simulations. We look at simple Markov models with Poisson and non-Poisson statistics and verify our analytical understanding of the models with their numerical counterparts obtained from the Monte Carlo simulations. We then derive the Gamma and Gaussian Markov models and highlight their advantages.

Chapter 3 looks at the use of the Gaussian Markov models to model the first order density results. We look at the detailed process of fitting the Gaussian models along with the estimation techniques of the parameters.

In Chapter 4, we look at the significance of using the Joint Density plots for the modelled and the measured results. We look at the process of gathering information on packet sequences from the Joint Density results and then use Markov model to model the packet sequence events. We also look at the symmetric and asymmetric nature of the Joint density results.

Chapter 5 also continues to look at the Joint Density results, this time with reference to periodic events. The features representing the presence of periodic events in the Joint Density plots are discussed in great detail along with how they can be modelled using Markov Models

Chapter 6 looks at the ON and OFF statistics of Voice traffic with respect to a large multi-lingual database of VoIP call recordings. The procedure of measurement of silences and talk spurts from this database is discussed and the impact of the prosodic structure of the various languages on the statistics of ON and OFF periods is discussed in elaborate details.

Chapter 7 looks at the distributions of the composite statistics of ON and OFF periods followed by the proposal of an ON-OFF source model that includes short silence lengths and has heavy tailed log-normal transitions.

Finally in Chapter 8, we draw conclusions on the various contributions of the thesis, also covering points not mentioned in the preceding chapters and then make recommendations for future work.

# Chapter Two

## Understanding Markov and Semi-Markov Packet Traffic Models

In the previous chapter, we have seen briefly, the various common mathematical approaches to network traffic modelling. In this chapter, we will cover in great detail, the mathematical background on the use of Markov and Semi-Markov Monte Carlo Packet Traffic models, as analytical tools for understanding the behaviour of network packet sources, in various simple topologies.

### 2.1 Mathematical Background on Markov Models

#### 2.1.1 What are Markov Models?

A *Markov Model*, also known as a *Markov Chain* or a *Markov Process*, simply put, is a stochastic model that consists of a finite number of states representing the behaviour and random evolution of a memory less system. Each state represents an event or a set of events and the system can be in any of this finite number of states at a given time. It progresses from one state to another with a fixed probability called the *state transition* probability. The model and its underlying stochastic process is based on the *Markov property* which was proposed by a Russian Mathematician by the name of Andrey Markov [42]. This property defined that the future behaviour of the system (i.e. the progression of the system from one state to another state) is completely independent of the past behaviour and is dependent solely on the present state of the system. Michael Baron [43] explained this property in the form of a simple relation as follows:

$$P \{ \text{future} | \text{past, present} \} = P \{ \text{future} | \text{present} \}$$

From a general mathematical modelling perspective, the Markov models can be broadly classified into two categories: *Discrete time Markov Models* and *Continuous time Markov Models*. In a Discrete time Markov model, the state transitions occur at integral values of time and the associated random variable  $X$ , which represents the time spent in the state, follows a geometric distribution. If there are no transitions for the discrete time model then  $X = 1$ . On the other hand, in a Continuous time Markov Model, the transition behaviour is different. Each state has a number of events which cause the transition to occur and this happens after the system spends an exponential amount of time in

that state. In other words the associated random variable ‘X’ follows an exponential distribution. As observed often with most types of traffic cases [15], we know that the network events happen at random times following a stochastic process and that the inter-arrival times of events have a unique distribution often exponential, and so we use Continuous time Markov Models in our work. Continuous time Markov Models are a special case of the more general class of Semi-Markov models, where the amount of time spent in each state is exponentially distributed. The general class of Semi-Markov models are those which can have any standard distribution for the state holding times apart from the exponential distributed. From now on we shall refer to the exponential based Continuous time Markov models as ‘Markov Models’ and use them for analysis as shown in sections 2.3 and 2.4.

## 2.1.2 Elements of a Markov Model

The key mathematical elements associated with a Markov Model are described as follows:

### State Transition Diagram

A Markov Model is often illustrated by using a *state transition diagram* which shows the state, transitions among the states, and the associated transition rates  $R_{ij}$ . A good example of a state transition diagram is shown in Fig. 2.1 which illustrates a three state model. The states labelled from 1 to 3 are shown in circles, the transitions are marked by arrows and the associated transition rates are indicated as  $R_{ij}$  close to the corresponding transition arrows. Any missing arrows indicate that the probability of that transition happening is zero. Another important point to note, with these continuous time Markov Models is that there is no such thing as a state transiting to itself, unlike some other versions of Markov Models. Also from an application perspective of the network traffic being modelled, there is a very high probability of a packet being followed by another packet as observed with most protocols, hence the traffic model does not have self transitions.

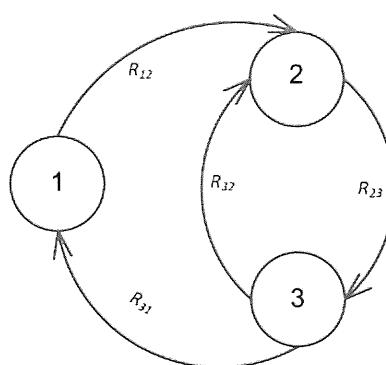


Figure 2.1: State Transition Diagram of a Three State Markov Model

### State Transition Probability Matrix, $P_{ij}$

The Markov model transits from one state (state  $i$ ) to another state (state  $j$ ) with a fixed probability  $P_{ij}$  called a transition probability. The matrix  $P$  whose  $ij^{\text{th}}$  entry is  $p_{ij}$ , called the *State Transition Probability Matrix* is as shown:

$$P_{ij} = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1j} \\ p_{21} & p_{22} & \dots & p_{2j} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ p_{i1} & p_{i2} & \dots & p_{ij} \end{pmatrix}, \text{ where } \sum_j P_{ij} = 1 \text{ and } p_{ii} = 0.$$

The State Transition Probability Matrix, also known as a *Jump Matrix* is a right stochastic matrix where the sum of the elements of its row must equal unity. If the Markov Model is in state  $i$ , the transition to the next state  $j$  takes place after an exponential amount of time with the *Rate parameter*  $R_{ij}$ .

### State Transition Rate Matrix

Another crucial element of a Continuous Time Markov Model is the *State Transition Rate Matrix* also known as the *Generator Matrix* or simply as the *Rate Matrix*. The matrix  $Q_{ij}$ , has the  $ij^{\text{th}}$  entry  $R_{ij}$  which is defined as the rate of the transition of the model from state  $i$  to state  $j$ . This is the reciprocal of the mean amount of time taken by the model to transit from state  $i$  to state  $j$ ,  $T_{ij}$ .

$$Q_{ij} = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1j} \\ R_{21} & R_{22} & \dots & R_{2j} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ R_{i1} & R_{i2} & \dots & R_{ij} \end{pmatrix}, \text{ where } \sum_j R_{ij} = 0 \text{ and } R_{ii} = 1/T_{ii}.$$

The diagonal elements or the elements  $R_{ii}$  of the matrix  $Q_{ij}$  are defined as:

$$R_{ii} := - \sum_{j \neq i} R_{ij}. \quad (2.1)$$

This implies that the sum of the elements of each row of the Rate matrix must be zero. The relationship between the elements of the state transition probability matrix  $P_{ij}$  and the elements of the rate matrix  $R_{ij}$  is given by:

$$p_{ij} := \begin{cases} R_{ij} / \sum_{k \neq i} R_{ik}, & \text{if } i \neq j \\ 0, & \text{otherwise} \end{cases}. \quad (2.2)$$

Thus it is possible to derive the State Transition Probability Matrix  $P_{ij}$  from the Rate Matrix  $Q_{ij}$  using this equation (2.2).

### State holding time Probability density

The parameter  $P_n(t)$ , where  $n$  is the number of the current state and  $t$  is the time (continuous time variable), is defined as the probability density function (PDF) associated with the state holding time, which in other words, is the time spent by the model in the state. This is the probability distribution with a mean parameter  $R_n$ , where  $R_n$  is the sum of all outward transition rate parameters from the  $n^{\text{th}}$  state. So for example, if the model has exponential inter arrivals, the holding time distribution for State  $n$  would be as follows:

$$P_n(t) = R_n e^{-R_n t} \quad , \text{where } R_n = \sum_{j=1}^M R_{ij} ; i = n \text{ and } M = \text{no. of states.} \quad (2.3)$$

### State Visiting Probability

The State Visiting Probability, defined by the parameter  $V_n$ , where ' $n$ ' is the number of current state, is the probability value of the model visiting that state in a Monte Carlo Simulation. It is the ratio of the number of visits to each state, to the total number of simulated steps. The sum of all state visiting probabilities must equal to unity. The state visiting probability is given by the limiting distribution of the Markov Model. A detailed section on the mathematical derivation and numerical computation of the state visiting probabilities is covered in section 2.2.3 with examples.

### 2.1.3 State Visiting Probabilities and ergodic Markov models

In this thesis, the Markov traffic models used in modelling the traffic IPT statistics are ergodic Markov models. An ergodic Markov model is defined as a Markov model which has a recurring finite number of states and is irreducible. These ergodic Markov models have states such that it is possible to transit from one state to any state (in any number of transitions). Ergodic Models have a very unique and useful property where regardless of the initial state the model starts in, it reaches a unique stationary probability distribution after  $N$  steps. This unique distribution is also called the Equilibrium probability distribution or the limiting distribution of the model. This equilibrium probability distribution is extremely useful as it defines the total number of visits to each state thus providing

the exact values of the state visiting probabilities. In Section 2.2, we refer to this property again and demonstrate again with examples how the state visiting probabilities can be calculated using the limiting probability distribution.

From a traffic modelling point of view, with the use of Markov models on modelling IPT statistics, only ergodic Markov models can be used to model IPT statistics. It is completely unusual to consider a non-ergodic Markov model to model the IPT traffic since each traffic source sends packets to different sources and hence must be visited recurrently, as otherwise it would mean that the traffic source did not contribute to the IPT statistics at all. Hence only ergodic Markov models have been considered in this thesis.

#### **Relationship between the state visiting probabilities and the equilibrium probability distribution:**

Let us consider  $P$  – the state transition matrix described earlier. Let  $N$  denote the number of state transition steps for the Markov model. Let  $V$  be defined as the limiting probability distribution of the ergodic Markov model and  $W$  be the limiting matrix (after  $N$  transitions) of the Markov model. Then as defined in standard texts of Queuing theory and Markov models [44] [45], as  $N$  approaches infinity, the powers  $P^N$  culminate into the limiting matrix  $W$  where each row is given by the row vector  $V$  which is nothing but the limiting distribution of the Markov model.

$$W = \lim_{n \rightarrow \infty} P^n \quad (2.4)$$

and,

$$V \cdot P = V \quad (2.5)$$

where,

$V$  is a common row vector or is a constant multiple of the limiting matrix  $W$ .

This vector  $V$  consists of the elements  $V_1, V_2, V_3, V_4, \dots, V_n$  such that  $V_1 + V_2 + V_3 + V_4 + \dots + V_n = 1$ .

These elements are nothing but the state visiting probabilities of the Markov model.

$$(V_1 \ V_2 \ \dots \ V_n) \begin{pmatrix} p_{11} & p_{12} & \cdot & \cdot & p_{1j} \\ p_{21} & p_{22} & \cdot & \cdot & p_{2j} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{i1} & p_{i2} & \cdot & \cdot & p_{ij} \end{pmatrix} = (V_1 \ V_2 \ \dots \ V_n). \quad (2.6)$$

The solutions for the equation (2.6) can be computed using the ‘balance equations’ approach as suggested in [10] [44]. This procedure is demonstrated further in section 2.2.3.3.

In section 2.2, we look at a few examples of Markov traffic models where we graphically estimate the state visiting probabilities using graph theory, and compare the results with those obtained from the Markov Monte Carlo simulations.

#### 2.1.4 Using Markov Models to model Packet traffic

As mentioned earlier in Chapter 1, Markov Models can be used to model the traffic sources of a network, by using a finite number of states. Each state once visited, executes an event. From a network perspective, this event can be emitting or receiving packets or packet trains depending on the network protocol being modelled. Once the state is visited, the model waits for a certain amount of time before the event is executed and it then moves to the next state with a fixed probability based on the transition rate. Fig. 2.2 shows a simple example of a two state Markov Packet Traffic Model based on a Client – Server model following a REQ-ACK session. State 1 can be modelled as a client sending a packet A to the Server to which the State 2, modelled as a server, replies back with a packet B.

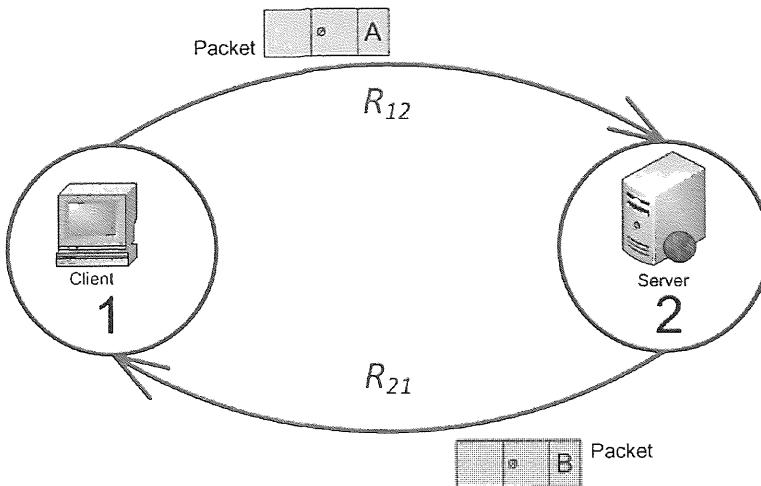


Figure 2.2: A simple two state Packet Traffic Model

We had briefly mentioned earlier that Packet-level traffic models usually express their flows in terms of the Inter-Packet times (IPT) and Packet Size (PS). These are simpler to study and a variety of other factors like Loss, Delay and Jitter can be observed using these two parameters. They also offer the advantage of avoiding any assumptions regarding the application layer protocol characteristics and the possibility to study a mixture of a variety of traffic sources [37]. The scope of this thesis is limited to detailed statistical and analytical work on the modelling of the Inter-Packet times (IPT) related to network packet traffic. Throughout we will be focussing on the *probability density* and *joint density* of the IPT derived analytically as well as those results from the Monte Carlo simulations of the Markov models associated with various simple network topologies.

### 2.1.5 Setting up Markov Monte Carlo Simulation to generate IPT statistics from the Markov Models

#### Markov Monte Carlo Simulation

A Monte Carlo simulation of a Markov Model refers to a simulation which involves repeated sampling of a variable from the holding time distributions  $P_N(t)$  associated with the states in the model. Each state emits a packet and the samples generated out of the Monte Carlo simulation correspond to the IPT (Inter-Packet times) being generated out of the model. The main parameters required for the Monte Carlo simulation of the Markov model, are the transition probability matrix  $P_{ij}$  derived from the rate matrix  $Q_{ij}$ , the initial state distribution, which indicates which state the Markov Model should begin the sampling process with, and finally ' $N$ ' the total number of transitions the

Monte Carlo Simulation needs to run for. This  $N$  also corresponds to the total number of IPT samples generated from the model.  $N$  should be typically a high value to achieve good statistics.

## Statistical Plots

One of the main purposes of having the Markov Monte Carlo Simulations is to generate statistical results from the Markov Models and compare it to the measured and analytically derived results to confirm our understanding on the protocol of the traffic type being modelled. In this thesis, we will use two main types of the statistical plots: the Probability Density Plot and the Joint Density Plot.

### Probability Density Function (PDF) – Statistical Plot

The primary use of the Probability Density Plot is to quantize the IPT samples in units of time. This characterises a continuous probability distribution. The X-axis of the plot represents the IPT samples in units of time, and the Y-axis of the plot represents the density of the probability for the range of values of the IPT times. The curve associated with this plot is the probability density function. The total area under the probability density curve should be equal to unity. The probability that the continuous time variable ‘ $t$ ’ associated with the IPT takes some value between two time limits is given by the area of the probability density function between those limits. The probability density function can be represented as a function  $f(t)$ , which is defined as:

$$P[a < T \leq b] = \int_a^b f(t)dt \quad (2.7)$$

### Cumulative Density Function (CDF) – Statistical Plot

The *Cumulative Density Function*, also occasionally used in our work, is the integral of the probability density function. Consequently for this plot, the values on the Y – Axis read from 0 to 1. The Cumulative Density Function (CDF) plot determines probability of observing a value less than the given value. The CDF can be represented as  $F(t)$  as shown in equation (2.8):

$$F(t) = \int_{-\infty}^t f(x)dx \quad \text{and} \quad F(t) = P(x < t) \quad (2.8)$$

From a Markov Monte Carlo perspective, the CDFs are useful in visually estimating the values of the State Visiting Probabilities, as we shall see in the next chapter.

### Joint Density Function (JDF) – Statistical Plot

A Joint Density Function (JDF) is a two variable density function that is three dimensional in nature. It is also known as a *Bi-variate Distribution Function*. The plot shows the density for a pair-wise sequence of two random variables representing times  $t_1$  and  $t_2$  plotted on the X and Y Axes respectively and the probability density being plotted on the Z axis. A two dimensional surface represents the Joint Density and again as with the probability density, the total volume under the surface is equal to unity. The Joint Density function can be defined as:

$$f_{T1,T2}(t_1, t_2) = f_{T1|T2}(t_2|t_1)f_{T1}(t_1) = f_{T2|T1}(t_1|t_2)f_{T2}(t_2), \quad (2.9)$$

Such that:

$$\int_{t_1} \int_{t_2} f_{T1,T2}(t_1, t_2) dt_1 dt_2 = 1. \quad (2.10)$$

In terms of the numerical computation, the Joint Density Function, is computed in the same way as the Probability Density Plot, except for the fact that a Bi-dimensional histogram is taken into account and the counts in each of the bins is normalised in such a way that the three dimensional volume covered by the plot is equal to unity. In our case of the Markov Monte Carlo simulations, the X axis is the IPT which we denote by the continuous time variable as  $t_n$  or simply as  $N$ . The second time variable is not a separately generated random variable, but is rather a right cyclic shift of values in variable  $t_n$  or  $N$  and this can be denoted as  $t_{n+1}$  or  $N+1$ . In other words the two variables used for representing the X and Y axes respectively are structured in such a way that the first element of  $t_{n+1}$  or  $N+1$  vector of IPT samples is the second element of the  $t_n$  or  $N$  vector of IPT samples respectively. The Joint Density Plot is very useful in traffic modelling as, it shows unique information about the traffic source behaviour which the probability density function does not show. The significance of the use of these Joint density plots shall be revealed and discussed in detail in chapters 4 and 5.

In the next few sections, we will consider some very simple Markov packet traffic models and use the Monte Carlo simulation to compare our analytically calculated results, with the results obtained from the Markov Monte Carlo Simulation which use different sample sizes  $N$  (Number of samples): 10000 (ten thousand), 100000 (hundred thousand), 1000000 (1 million), and we will see the impact of changes in the value of  $N$  on the statistical results of the Markov model. In a general observation made in the process of using the Markov Monte Carlo simulation, the error rate deviation in terms of

the comparison of the analytical and numerical results is < 2% for sample sizes larger than one hundred thousand, hence the choice of  $N$  as mentioned.

Once the Markov Monte Carlo simulations are executed, the system Probability Density Function (PDF) and the Joint Density Function (JDF) are plotted and compared to their analytically derived equivalents to confirm our analytical understanding of the Model.

## 2.2 Markov Traffic Models with Poisson Statistics

In this section, we introduce the exposition of background of modelling traffic with Poisson statistics using simple Markov traffic models. As highlighted earlier, one key difference with our approach when compared with the conventional approach of Markov modelling is that we use the time spent in each state to generate the statistics and the actual packet events of the network are marked by the transitions, whereas in the conventional approach the events are those that are executed within the state and the statistics are based on these events. While standard texts on Markov models such as [45] [46] [47] [48] cover the basics of Markov Models, in this chapter we will be considering simple Markov traffic models and analysing them in a Network and IPT context and these results will form a strong foundation for further understanding of protocol features and traffic patterns observed from the measured IPT traffic.

In our process of analysis and understanding of how simple Poisson Markov Packet Traffic Models work, we will first consider a simple two state Poisson Markov Model, the same example which was referred to in Section 2.1.3. By a Poisson Markov Model we mean a model, which has exponential distributions for the state holding times, that is, the IPT carry an exponential distribution. We will cover two variants of this 2-state Markov Packet models: Symmetric rates and Asymmetric rates.

### 2.2.1 A two-state Poisson Markov Traffic Model

#### 2.2.1.1 Symmetric Case

Let us consider the two – state Markov Model as shown in Fig 2.3. The Model is labelled symmetric because both the state transition rates ‘R’, of the two states transitioning from one to another are the same as seen in Fig.2.3, State 1 emits a Packet A and State 2 emits packet B, so the packet sequence will be ‘ABABABABABAB....’and so on. The packets are indicated with three sectioned rectangles which are often used to denote data packets.

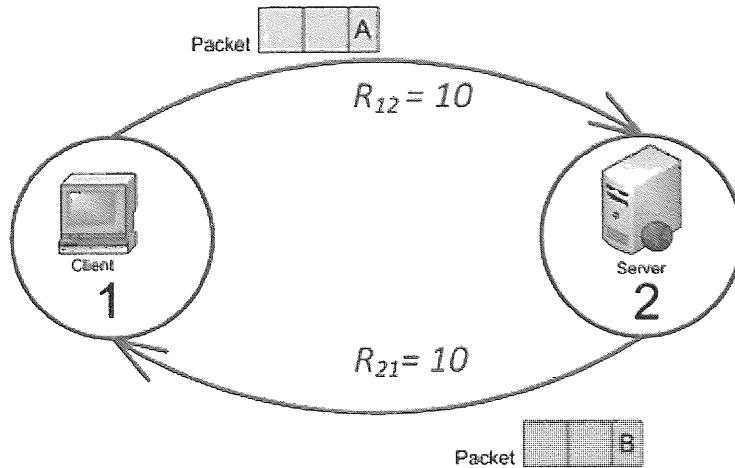


Figure 2.3: State Transition Diagram of a Symmetric Two-State Markov Traffic Model

Before we proceed with the numerical simulation, we will look at a mathematical analysis and derive the result to verify against the simulation and confirm our analytical understanding of the model.

#### 2.2.1.1.1 Probability Density Function – Mathematical Analysis

We begin by identifying a packet event with each state transition. Let  $p_{12}$  and  $p_{21}$  be the probability of transiting to state 1 and state 2 respectively.  $R_{12}$  is the transition rate from state 1 to state 2 and  $R_{21}$  is the transition rate from state 2 to state 1 respectively. The important assumption here is that  $R$  does not depend on the history of states occupied by the system. This is based on the Markov property discussed earlier. Now let us consider the behaviour of the system and calculate the probabilities under the assumption that the system is in a steady state.

Using equation (2.2), the probabilities  $p_{12}$  and  $p_{21}$  can be calculated as:

$$p_{12} = \frac{R_{12}}{R_{21} + R_{12}}, p_{21} = \frac{R_{21}}{R_{22} + R_{21}}.$$

#### State Holding time distribution, $P(t)$ :

Next, we will look at the probability distribution of the state holding time,  $P_n(t)$ , described in section 2.1.2. We can define  $P_1(t)$  as the probability that the system is in state 1 for time  $t$ . This Probability Density representing the holding time for State 1 for the two-state Poisson Markov Model is given by an exponential random variable with the rate  $R_{12}$  as shown in equation (2.11):

$$P_1(t) = R_{12} e^{-R_{12} t}, \quad (2.11)$$

And similarly we also have for State 2:

$$P_2(t) = R_{21} e^{-R_{21}t} \quad . \quad (2.12)$$

### **Rate Matrix $Q_{ij}$ and Transition Probability Matrix $P_{ij}$ :**

For the two-state Poisson Symmetric model, we have the matrix  $R_{ij}$  as follows:

$$R_{ij} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} = \begin{pmatrix} 0 & 10 \\ 10 & 0 \end{pmatrix}.$$

Using equation (2.1), we then have the Rate Matrix  $Q_{ij}$  given as:

$$Q_{ij} = \begin{pmatrix} -10 & 10 \\ 10 & -10 \end{pmatrix}$$

Then using equation (2.2), we can obtain the Transition Probability Matrix  $P_{ij}$ :

$$P_{ij} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

This transition matrix  $P_{ij}$ , is fed into the Markov Monte Carlo Simulation for creating the transition chain, while the  $i^{\text{th}}$  diagonal elements of the CT Rate Matrix  $Q_{ij}$  are used for generating the exponential random variables associated with the exponential amount of time that the model waits between before transiting to the next possible state.

### **State Visiting Probabilities, $V_n$ :**

Calculating the state visiting probability for the two state Markov model is straight forward. As stated earlier, Packet A is always followed by packet B and vice versa. When in state 1, the next state the model has to go to is state 2 and then return to state 1 again. There are no other possibilities of state transitions. So it is clear from this, that after a simulation of  $N$  samples, the model spends half the time in visiting state 1 and the other half of the time in visiting state 2.

So it is clear that the visiting probabilities of the two-state Markov model will be:

$$V_1 = 0.5, \quad \text{and } V_2 = 0.5.$$

The sum of all visiting probabilities should of course be equal to 1. So,

$$V_1 + V_2 = 1.$$

### Deriving the PDF Equation:

The Probability Density Function for the entire system is the sum of the probability densities of the time being spent in the individual states weighted by their state visiting probabilities.

$$PDF_{(2-state)} = V_1 P_1(t) + V_2 P_2(t) \quad (2.13)$$

Where  $P_1(t)$  and  $P_2(t)$  are given by (2.11) and (2.12) respectively.

Substituting the values of rates of the two state symmetric Markov model, we have:

$$PDF_{(2-state,sym)} = 0.5 * 10 e^{-10t} + 0.5 * 10 e^{-10t}. \quad (2.14)$$

Equation (2.14) is the derived analytical equation for the PDF of the mean IPT times generated by the two-state symmetric case Markov model, which are of course based on the mean holding times of any of the two states. Now we shall compare this to the numerical result, obtained for the same model by executing the Monte Carlo Simulation discussed earlier.

#### 2.2.1.1.2 Probability Density Function - Markov Monte Carlo Simulated Results

In Fig. 2.4, we show the numerical result for the Markov Monte Carlo simulation of the two-state symmetric Markov model for three different values of  $N$ . The black line represents the analytical solution in equation (2.10).

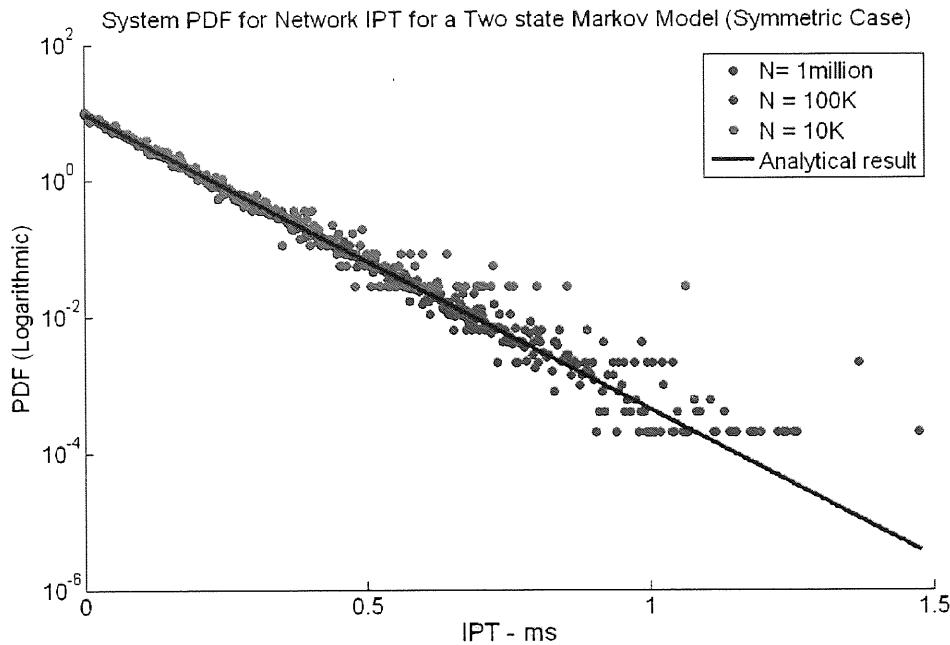


Figure 2.4: PDF of Network IPT times based on the symmetric case of 2-state Markov Model

The PDF is plotted on a logarithmic scale to clearly observe the differences between the results of the three different values of  $N$ . What we see here is simple Poisson statistics. The coloured dots represent the numerical results from the Markov Monte Carlo simulation. We clearly see that the numerical results follow our analytically derived results well and this confirms our analytical understanding of the probability density of the model. The red dots represent the results from the Markov Monte Carlo Simulation for  $N = 10,000$  samples. Likewise the blue dots represent results for  $N = 100,000$  samples and the green dots represent results for  $N = 1$  million samples. We observe that at lower values of the density, we begin to see Poisson noise, also known as counting noise. We see that as we increase the number of samples, the noise occurs at lower values of density, which means we can always improve the statistics by increasing the sample size  $N$ . Having made this conclusion, we will from here onwards only use  $N = 1$  million samples for further scenarios of Markov Models. The results are exponentially linear because of the uniform rate of transition.

#### 2.2.1.1.3 Joint Density Function – Mathematical Analysis

We previously derived the general PDF equation for the IPT for the two-state model given by (2.14). Now for the Joint Density Function for the IPT based on the symmetric case of the two-state Markov Model, we know that State 1 always follows State 2 and State 2 always follows State 1 and both happen equally often.

The joint density would be symmetric under the interchange of  $t_1$  and  $t_2$ . If we consider any three packets in sequence, then  $t_1$  is the time between the first two packets and  $t_2$  is the time between the last two packets. The Joint Density therefore represents the probability of observing an IPT of  $t_2$  when the previous IPT was  $t_1$ . Joint Density PDF function is then given by the equation (2.15):

$$\text{PDF}_{(2\text{-state})}(t_1, t_2) = V_1 P_1(t_1)P_2(t_2) + V_2 P_2(t_1) P_1(t_2). \quad (2.15)$$

Where  $V_1$  and  $V_2$  are the state visiting probabilities,  $P_1(t_1)$  and  $P_1(t_2)$  are the state holding time distributions for State 1 for times  $t_1$  and  $t_2$  respectively, and  $P_2(t_1)$  and  $P_2(t_2)$  are the state holding time distributions for State 2 for times  $t_1$  and  $t_2$  respectively, the values of which are given by:

$$V_1 = V_2 = 1/2$$

$$P_1(t_1) = R_{12} e^{-R_{12}t_1},$$

$$P_1(t_2) = R_{12} e^{-R_{12}t_2},$$

$$P_2(t_1) = R_{21} e^{-R_{21}t_1},$$

$$P_2(t_2) = R_{21} e^{-R_{21}t_2}.$$

Substituting all of those values in Equation (2.15), we get:

$$PDF_{(2-state)}(t_1, t_2) = \frac{R_{12}R_{21}}{2} [e^{-(R_{12}t_1+R_{21}t_2)} + e^{-(R_{12}t_2+R_{21}t_1)}]. \quad (2.16)$$

Using the rates of the symmetric example in consideration,  $R_{12} = R_{21} = 10$ , and plotting equation (2.16), we obtain the three dimensional plot as seen in Fig. 2.5.

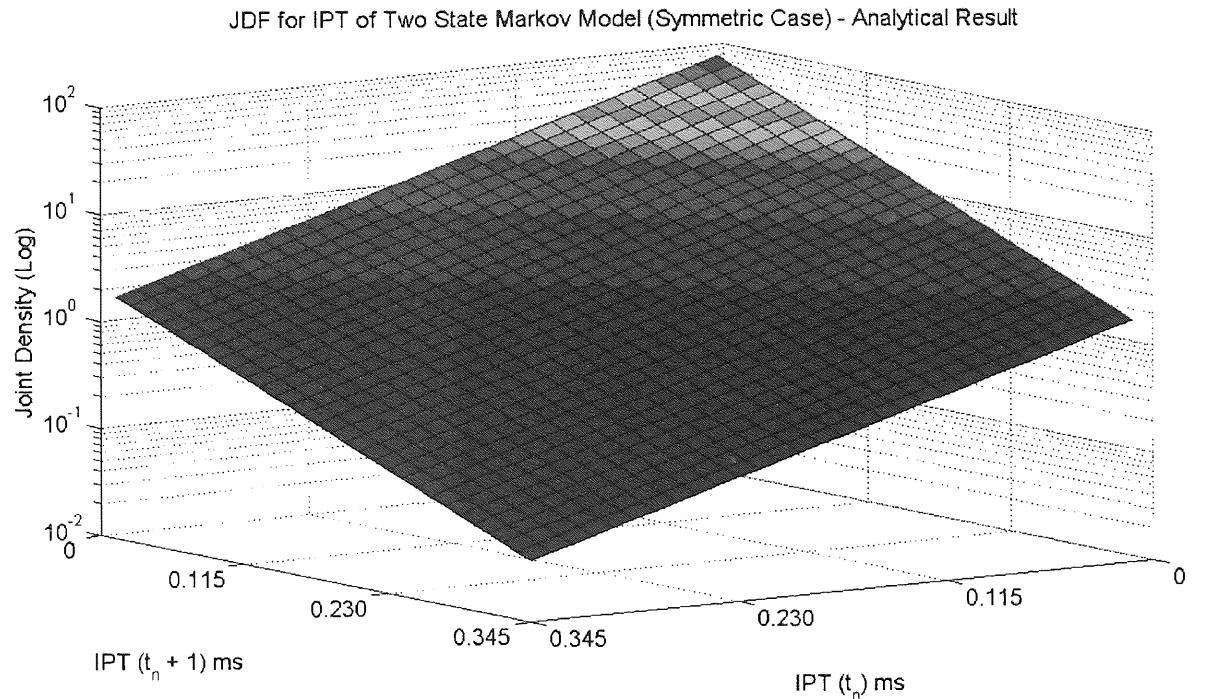


Figure 2.5: Joint Density Plot for 2- state Markov Model (Symmetric Case) – Analytical result

The Z-axis represents the density on a logarithmic scale whereas the X and Y axes, represent the values of  $t_n$  ( $t_1$ ) and  $t_{n+1}$  ( $t_2$ ) in time (ms), where the width of each bin used for plotting the Joint Density is a constant fraction of a second. The plot in Fig. 2.5 shows only the first 30 bins (limited to 0.345 ms) that is only a part of the actual 100 bin plot. This is to bring clarity into the interesting area of the plot. As can be seen from the Fig. 2.5, the graph shows a linear relationship between the values on the two axes because of the symmetric rates. Simply put, the result reflects three dimensional linear Poisson statistics similar in nature to that observed with the PDF earlier.

#### 2.2.1.1.4 Joint Density Function - Markov Monte Carlo Simulated Results

Now we compare the analytical result shown in Fig. 2.5 with the numerical result (Fig. 2.6) obtained from the Markov Monte Carlo Simulation for  $N = 1$  million samples. This is again shows only the first 30-bins of what is actually a 100 bin sized joint density plot.

As can be observed in Fig. 2.6, the result is in very close agreement to the analytical result with the exception of a noisy surface, which proves that our understanding of the interaction of the states with symmetric rates is right. The noisy surface arises due to sampling error, the effect of which is similar to that observed in the case of the PDF in Fig. 2.4.

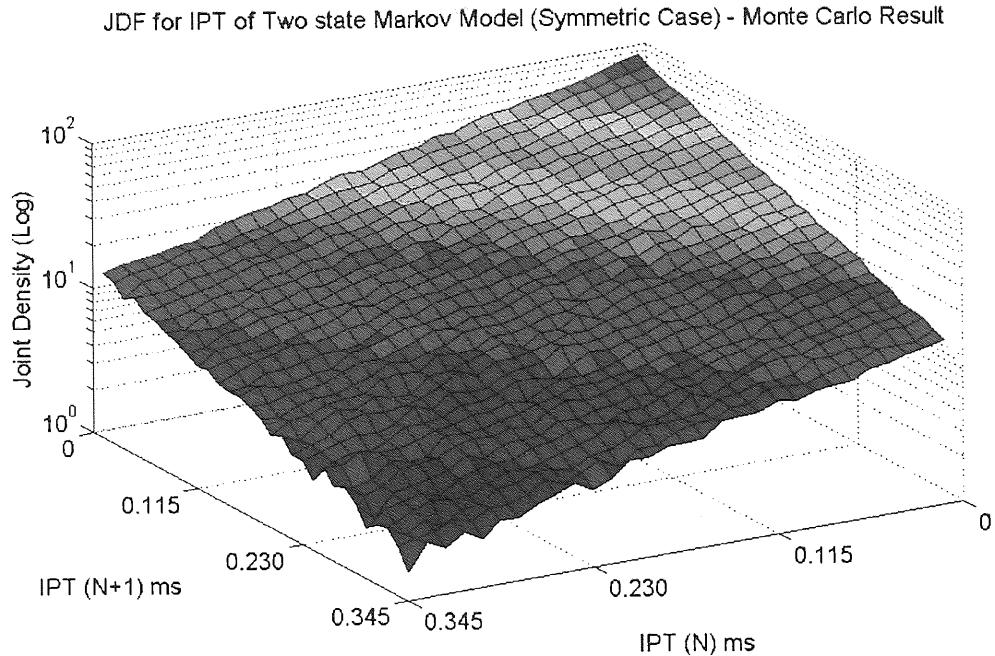


Figure 2.6: Joint Density Plot for a 2- state Markov Model (Symmetric Case) – Monte Carlo result

### 2.2.1.2 Asymmetric Case

Now let us consider the Asymmetric case of the 2 – state Markov Model. By asymmetric rates, we mean, that one of the transition rates is faster than the other one as shown in the state transition diagram in Fig.2.7. Here the rates are  $R_{12} = 10$  and  $R_{21} = 100$  respectively.

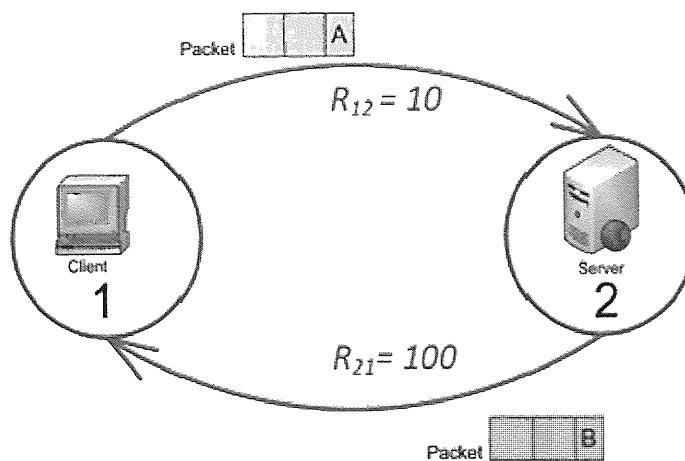


Figure 2.7: State Transition Diagram of a Two State Markov Model with Asymmetric Rates

### 2.2.1.2.1 Probability Density Function – Mathematical Analysis

The analytical equation for the Probability Density Function can be derived in the same manner as for the previous case of the model.

#### Deriving the Analytical Equation:

Being a two-state Markov Model again, the equations for the holding time distributions for both the states remain the same as given by equations (2.11) and (2.12).

The Rate matrix  $Q_{ij}$ , with the rates substituted is given by:

$$Q_{ij} = \begin{pmatrix} -10 & 10 \\ 100 & -100 \end{pmatrix}$$

And the Transition Probability Matrix is the same as before,

$$P_{ij} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Let us again identify a packet emission event with each state transition of the model. The visiting probabilities  $V_1$  and  $V_2$  also remain the same i.e. 0.5 each. Substituting the new asymmetric rates in the holding time distributions of  $P_1(t)$  and  $P_2(t)$  in the equation (2.16), we obtain:

$$\boxed{PDF_{(2-state,asym)} = 0.5 * 10 e^{-10t} + 0.5 * 100 e^{-100t}} \quad (2.17)$$

### 2.2.1.2.2 Probability Density Function – Markov Monte Carlo Simulated Results

Taking into account the new asymmetric rates, we run the simulation again for the two states Markov Model. This time for reasons stated earlier in section 2.2.1.1.2, we will be only considering the numerical results for a sample size of  $N = 1$  million. Fig. 2.6 shows the numerical results generated from the Markov Monte Carlo Simulation of the asymmetric two-state Markov Model, along with the curve representing the analytical solution in equation (2.17).

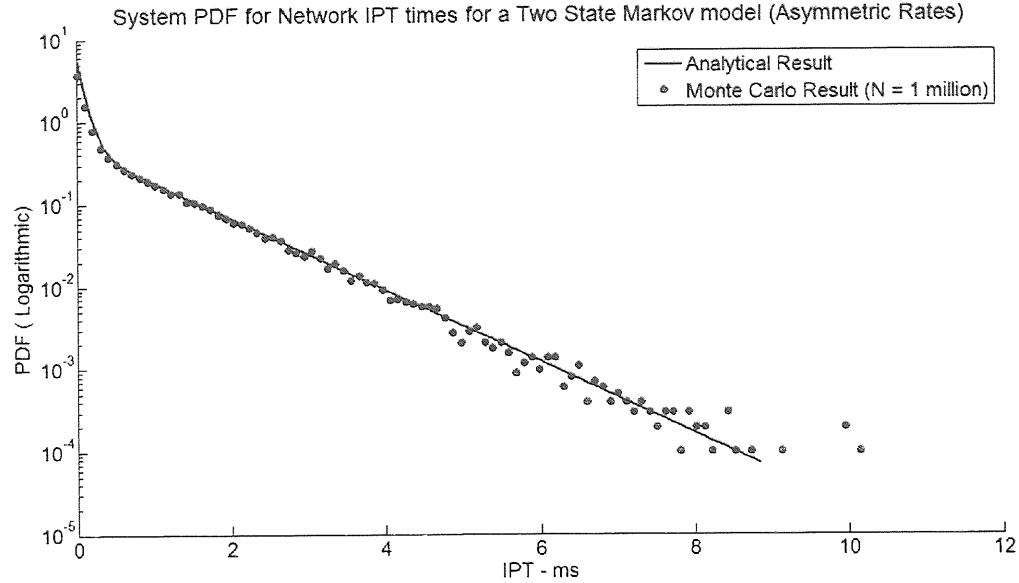


Figure 2.8: PDF for Network IPT times for a Two - State Markov Model (Asymmetric Case)

Again we observe that the numerical result is in agreement with the analytical result and is Poisson in nature. While we had observed an exponentially linear nature of the result with the symmetric case of the model, here we see a bend in the curve of the PDF, which indicates the transition of the system from the faster rate to the slower rate. This graph shows that around 50% of the packets with IPT ranging from 0 – 0.08 ms have been emitted with the faster rate, while the remaining 50% of the packets come from the state with the slower outward transition rate.

#### 2.2.1.2.3 Mathematical Analysis – Joint Density Function

For the Joint Density Function, again the same equation (2.16) holds good for the asymmetric case, where the rates  $R_{12}$  and  $R_{21}$  are 10 and 100 respectively.

Substituting those asymmetric rates in Equation (2.16) and plotting the Joint Density, we obtain the following analytical result as shown in Fig. 2.9:

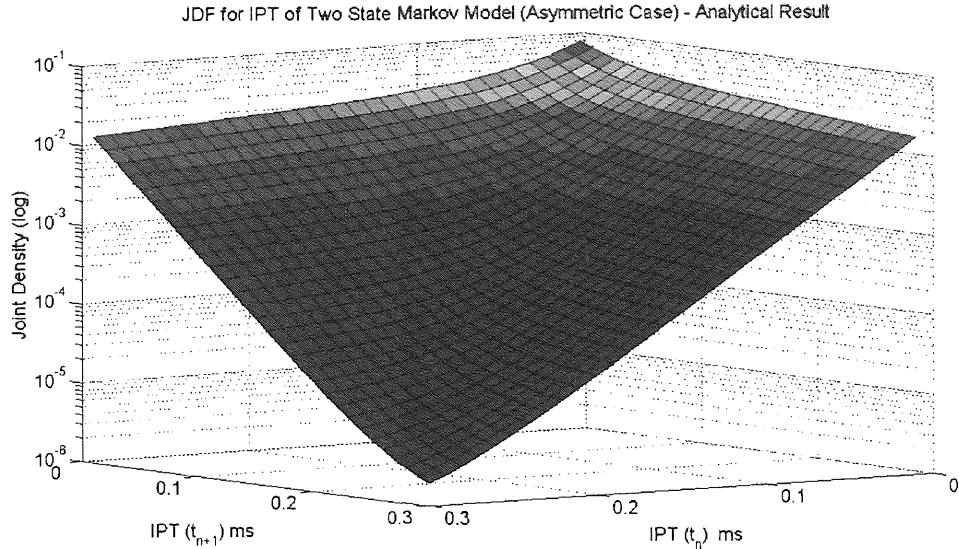


Figure 2.9: Joint Density Plot for 2- state Markov Model (Asymmetric Case) – Analytical result

The edges of the plot towards X and Y axes, are similar to that observed with the PDF curve for the Asymmetric rates. The transition from the faster rates to the slower rates is observed here as well, with the plot having a wing like structure by the walls indicating the slower transitions at the ends. The structure of the Poisson surface is like that of a plateau where slower transitions between the states are indicated by the deep pit and faster transitions are indicated by flattening end of the plot.

#### 2.2.1.2.4 Joint Density Function – Markov Monte Carlo Simulated Results

Comparing the analytical result in Fig. 2.9 with the simulated result in Fig.2.10, we observe that the

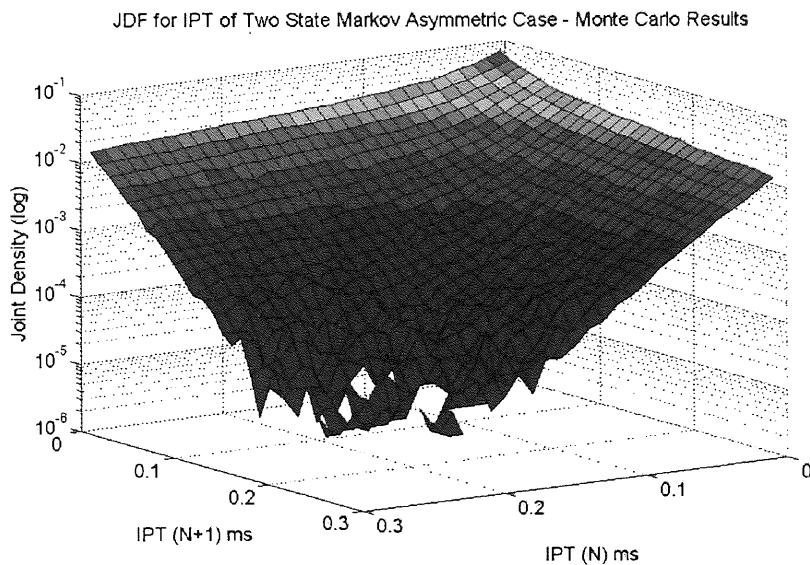


Figure 2.10: Joint Density Plot for 2- state Markov Model (Asymmetric Case) – Monte Carlo results

analytical result is in very close agreement again with the numerical result. As before, we can also observe the sampling noise at the lower values of Joint Density. An experimental check reveals that the RMSE (Root Mean Square Error) between the analytical and numerical results is  $R = 0.001246$  and the maximum of the error is  $\max = 0.004528$ , which confirm the agreement of both the results. This agreement of the analytical and simulated results now establishes confidence that we have understood the way the two state Markov model works in terms of generating the IPT samples and we now also know what differences to expect based on the choice of transition rates we choose for the Markov Model to simulate results based on measured IPT data from the traffic of a data network.

## 2.2.2 A three-state Poisson Markov Traffic Model

We will now move on to an example of a three-state Markov model whose state transition diagram is shown in Fig. 2.11. This model can be considered as an extension of the two state models where state 2 sends a packet and receives a response from state 3, but for 80% of the time, the packet has to be re-routed back through state 1 instead of going directly from state 3 to state 2. From a network perspective, this could be an extension to the two-state Markov model, where the server (state 3) occasionally re-routes the packet for the client (state 2) through another node (state 3).

### 2.2.2.1 Mathematical Notes

Let us first consider the symmetric case of the Three State Markov Model. For n-state Markov models where n is greater than 2, the “Symmetric” here is defined as being the property of the model such that the diagonal elements of the transition rate matrix Q of the model are identical. This means the sum of all outgoing stat transition parameters of each state is the same as can also be observed from the example shown in Fig. 2.11:

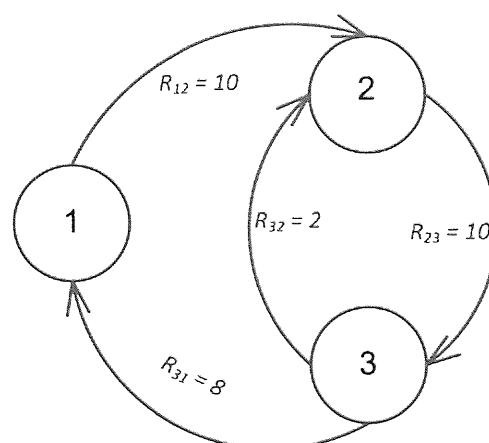


Figure 2.11: State Transition Diagram for a Three - State Markov Model - Symmetric Case

Now, with reference to equation (2.3), the State holding time probability distributions for a generic 3-state Markov Model is given by:

$$P_1(t) = (R_{12} + R_{13})e^{-(R_{12}+R_{13})t},$$

$$P_2(t) = (R_{21} + R_{23})e^{-(R_{21}+R_{23})t},$$

$$P_3(t) = (R_{31} + R_{32})e^{-(R_{31}+R_{32})t}.$$

The Rate Matrix  $Q_{ij}$  is:

$$Q_{ij} = \begin{pmatrix} -(R_{12} + R_{13}) & R_{12} & R_{13} \\ R_{21} & -(R_{21} + R_{23}) & R_{23} \\ R_{31} & R_{32} & -(R_{31} + R_{32}) \end{pmatrix} = \begin{pmatrix} -10 & 10 & 0 \\ 0 & -10 & 10 \\ 8 & 2 & -10 \end{pmatrix}.$$

It would be worthwhile to note that for higher order Markov model, an easier definition of symmetry of states would imply that the  $i^{\text{th}}$  diagonal elements of the CT Rate Matrix  $Q_{ij}$  would be identical.

And using Equation (2.2) on  $Q_{ij}$  again we have, the transition Probability Matrix,  $P_{ij}$ :

$$P_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.8 & 0.2 & 0 \end{pmatrix}.$$

### 2.2.2.2 State Visiting Probability

The calculation of the state Visiting Probabilities,  $V_1$ ,  $V_2$ , and  $V_3$  for the three-state Markov Model is slightly more complicated than that of the two-state model, as far as the graphical approach is concerned. We will consider the three-state model considered in the example, and calculate the state visiting probabilities using graph theory.

To begin with, let us consider the *branching* probabilities of State 3,  $p$  and  $q$  with reference to the model shown in Fig. 2.11:

$$p = \frac{R_{31}}{R_{31} + R_{32}} = \frac{8}{8+2} = 0.8,$$

$$q = 1 - p = 0.2.$$

Using graph theory, we can produce the graphical representation of the sequence of states for the three-state model as shown in Fig. 2.12. Suppose we start at state 2 (it does not matter which state we start from), then we have (Fig. 2.12):

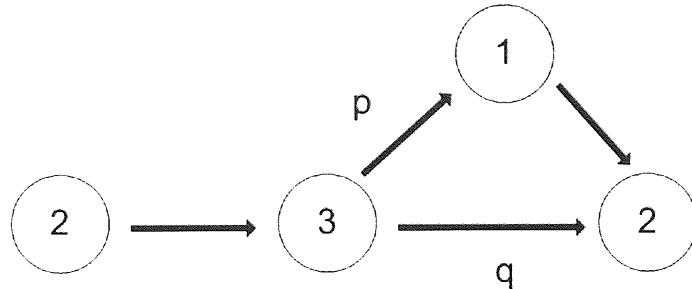


Figure 2.12: Calculating State visiting probabilities – Three State Markov Model

From the figure, we know that in the sequence of states, state 3 always follows state 2. So we have,

$$V_2 = V_3 , \quad (2.18)$$

$$V_1 = p \times V_3 = p \times V_2. \quad (2.19)$$

And since sum of probabilities should be unity we have,

$$V_1 + V_2 + V_3 = 1 . \quad (2.20)$$

Substituting (2.18) and (2.19) in (2.20), we have:

$$p \times V_3 + V_3 + V_3 = 1 . \quad (2.21)$$

Solving this, yields:

$$V_2 = V_3 = \frac{1}{2+p} , \quad (2.22)$$

$$V_1 = \frac{p}{2+p} . \quad (2.23)$$

Substituting the value of  $p = 0.8$ , in (2.19) and (2.20) we have the state visiting probabilities for the three-state Markov model:

$$V_2 = V_3 = 1 / 2.8 = 0.3571 ,$$

And,

$$V_1 = 0.8 / 2.8 = 0.2857.$$

### 2.2.2.3 Probability Density Function – Symmetric Case

The general PDF equation for the network IPT times based on the 3-state Markov Model is given by:

$$PDF_{(3\text{-state})} = V_1 P_1(t) + V_2 P_2(t) + V_3 P_3(t). \quad (2.24)$$

Substituting the values of the probability densities, rates and visiting probabilities in 2.21, the PDF equation for the symmetric case of the 3-state Markov Model will be:

$$PDF_{(3\text{-state},\text{sym})} = 0.2857 * 10 e^{-10t} + 0.3571 * 10 e^{-10t} + 0.3571 * 10 e^{-10t}. \quad (2.25)$$

Plotting the analytical solution (2.25) against the numerical results from the Markov Monte Carlo simulation as seen in Fig. 2.13, we observe that the results look similar to the Poisson natured results

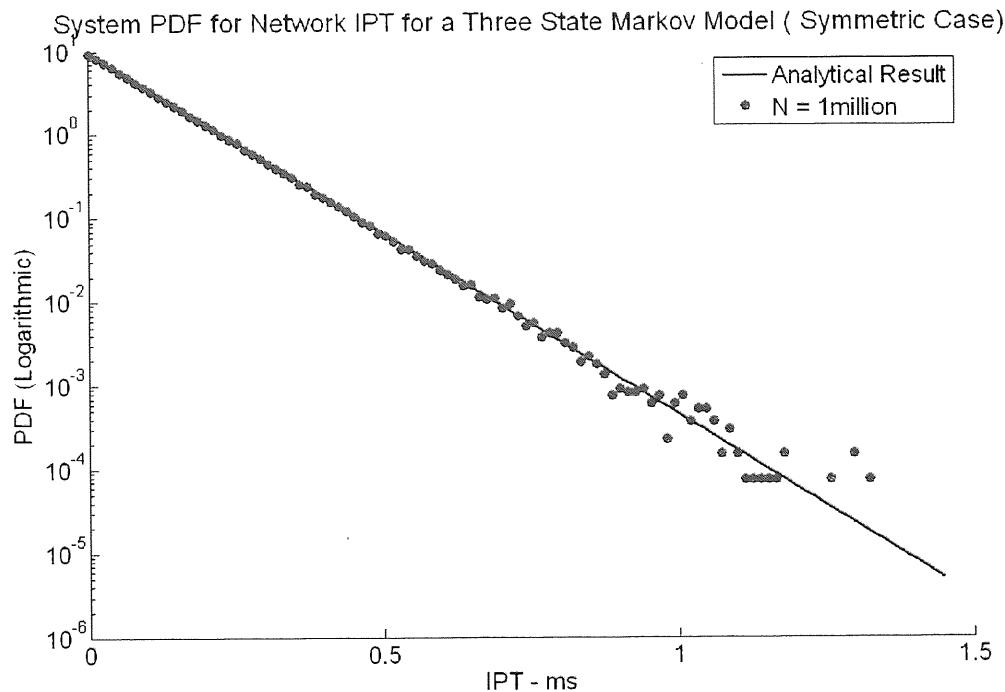


Figure 2.13: PDF for Network IPT for a 3-state Markov Model (Symmetric Case)

observed from the two-state symmetric model, owing to the symmetric nature of the rates at which the states emit the packets. We see that the analytical solution represented by the equation (2.25) is in total agreement with the numerical results achieved by the Markov Model simulation. Next we move on to the Asymmetric Case of the three- state Markov Model.

#### 2.2.2.4 Probability Density Function - Asymmetric Case

Let us consider the same three-state model again, but with asymmetric rates as shown in Fig. 2.14.

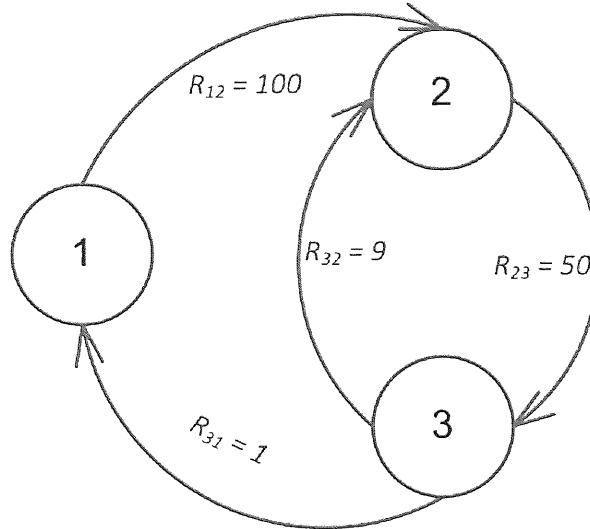


Figure 2.14: State Transition Diagram for 3-state Markov Model with Asymmetric Rates

With these changes in rates, the model acts like a two state model (States 2 and 3) 90% of the time. The packets from state 3 are re-routed occasionally via state 1, the remaining 10% of the time.

The Rate Matrix  $Q_{ij}$  for this model is:

$$Q_{ij} = \begin{pmatrix} -(R_{12} + R_{13}) & R_{12} & R_{13} \\ R_{21} & -(R_{21} + R_{23}) & R_{23} \\ R_{31} & R_{32} & -(R_{31} + R_{32}) \end{pmatrix} = \begin{pmatrix} -100 & 100 & 0 \\ 0 & -50 & 50 \\ 1 & 9 & -10 \end{pmatrix}.$$

And using Equation (2.2) on  $Q_{ij}$  again we have, the transition Probability Matrix,  $P_{ij}$ :

$$P_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.1 & 0.9 & 0 \end{pmatrix}.$$

As explained with the symmetric case, we can calculate the state visiting probabilities using the same approach explained in section 2.2.2.2.

The branching probabilities of State 3,  $p$  and  $q$  are given as follows:

$$p = \frac{R_{31}}{R_{31} + R_{32}} = \frac{1}{1+9} = 0.1,$$

$$q = 1 - p = 0.9.$$

Again, as before, we have:

$$V_2 = V_3 = \frac{1}{2+p},$$

$$V_1 = \frac{p}{2+p}.$$

Substituting  $p = 0.1$ , we get:

$$V_2 = V_3 = 0.476 \text{ and } V_1 = 0.0476.$$

We can verify these values with, the use of counts of state visits from the Monte Carlo Simulation of the asymmetric case of the three-state Markov Model. Fig. 2.15 shows a histogram of the counts of visits to each of the three states for the asymmetric 3-state Markov model. As mentioned before, the total number of iterations for the Monte Carlo simulations is a million.

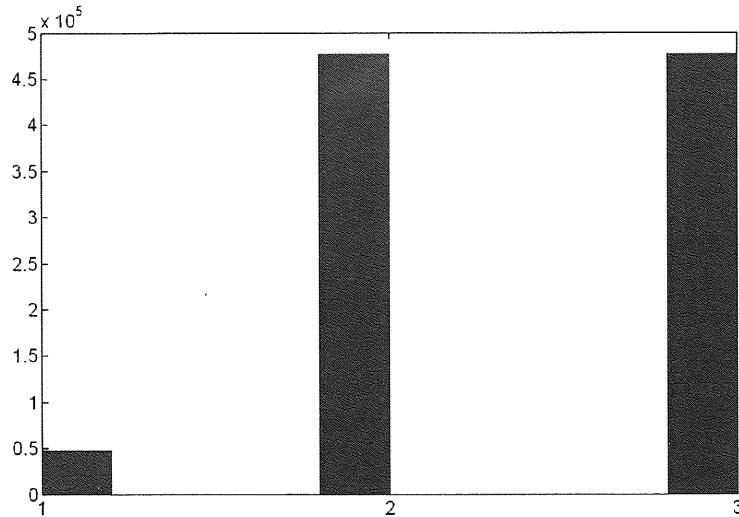


Figure 2.15: Histogram of Counts of visits to states for 3-state Markov model simulation

The X-axis represents the states and the Y-axis represents the count of visits to each of these states. Upon calculating the percentage of visits to each state, we find that the visits to state 1 represent 4.76 %, and states 2 and 3 have 47.6 % of visits each. Upon normalising these values to 1, we see that

these numbers tally with the previously calculated state visiting probabilities, thus verifying our approach.

Substituting the values of the state visiting probabilities, the rates and the holding time probability distributions in equation (2.24), we get:

$$\boxed{PDF_{(3-state, \text{asym})} = 0.0476 * 100 e^{-100t} + 0.476 * 50 e^{-50t} + 0.476 * 10 e^{-10t}} \quad (2.26)$$

Plotting this analytical solution with the numerical result as shown in Fig. 2.16, we see that the results are in agreement with each other and we see that this is again similar to the Poisson natured results of the asymmetric case of the two-state Markov model. The bend in the curve demonstrates transition of model from the fastest rate (100) to a slower rate (50) and then to the slowest rate (10).

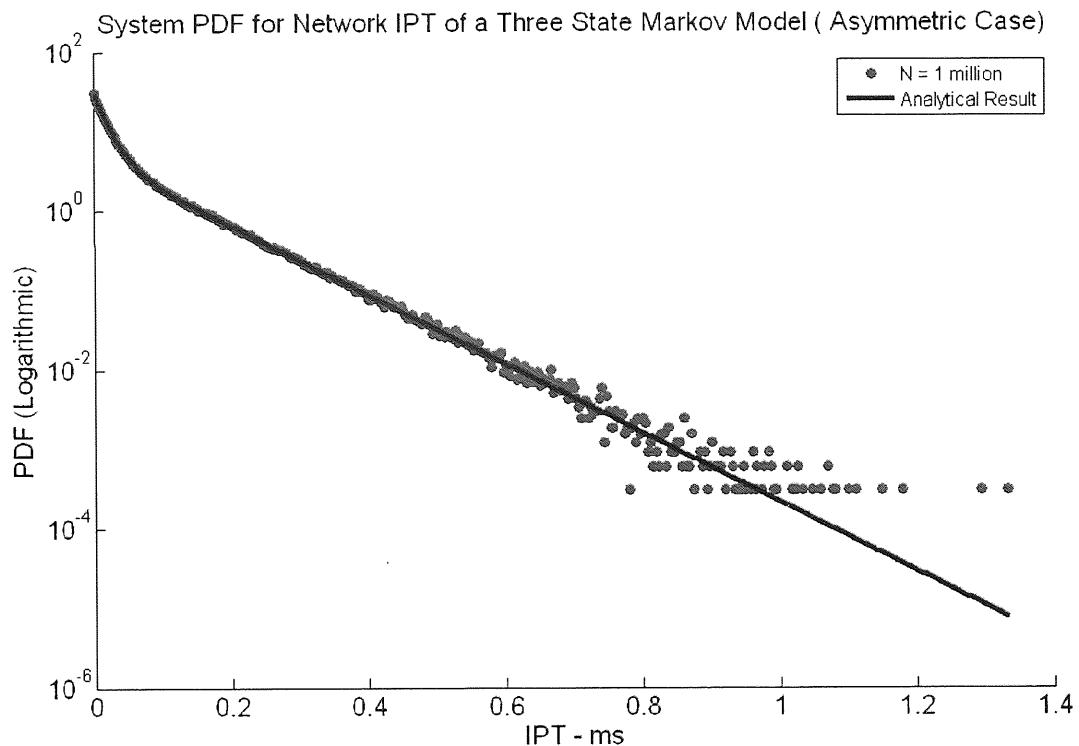


Figure 2.16: PDF for IPT for a 3-state Markov Model (Asymmetric Case)

### 2.2.3 n-state Poisson Markov Traffic Models

So far we have considered, studied and understood the results of 2-state and 3-state Poisson Markov Traffic models. To see if the mathematical procedure used to derive the general PDF equation of the

two state and three state Markov model holds good for all cases of the threes state model, we considered a variety of other examples of models where every state emits a packet.

### 2.2.3.1 Some Examples

The state transition diagrams for some of the higher-order Poisson Markov Traffic models considered are shown in Fig. 2.17 (a-e):

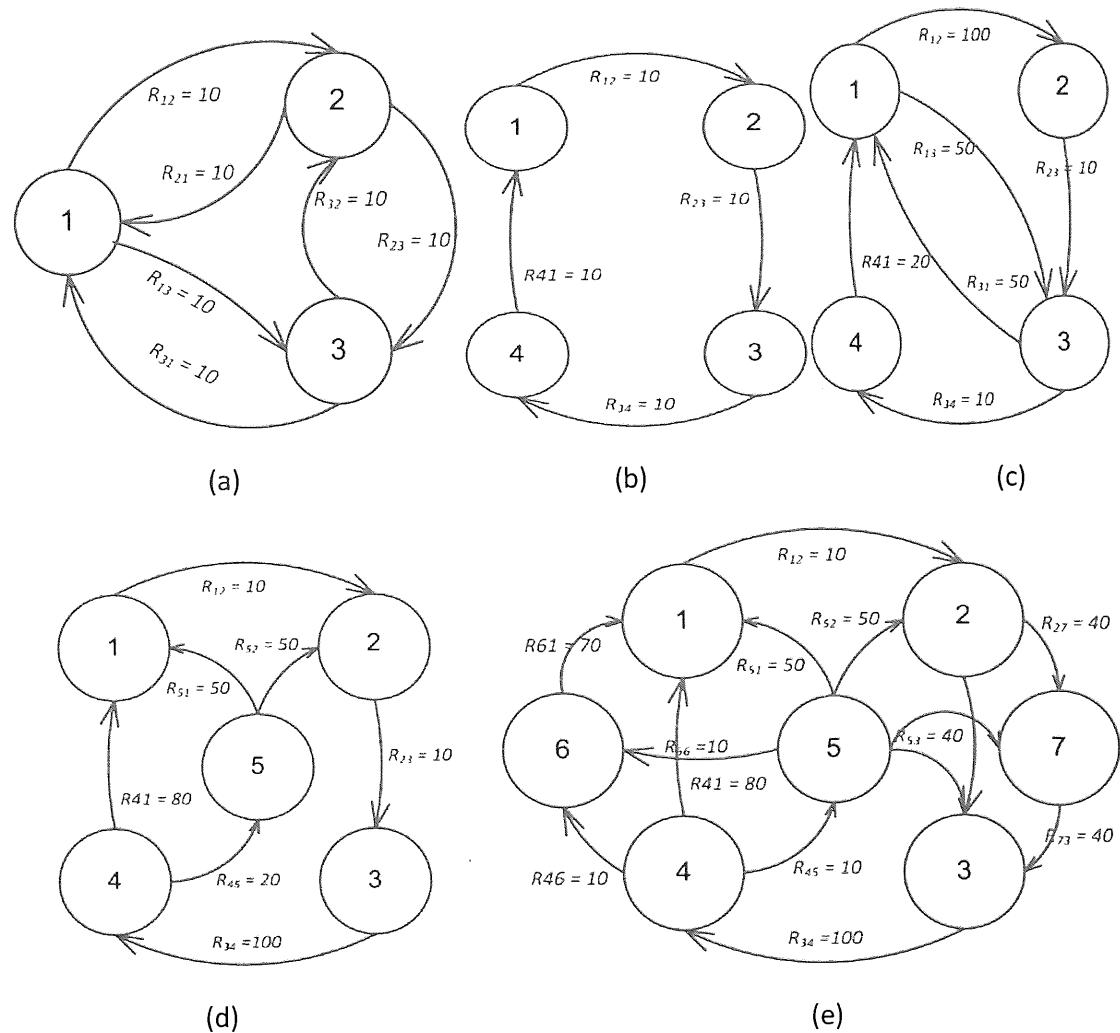


Figure 2.17: State Transition diagram for (a) Three-state Markov Model with all states emitting packets (b)Four-state Loop model (c) Four-state Loop model with diagonal transitions (d) Five state Markov Model and (e) Seven state Markov Model

Figure 2.17 (a) shows a 3-state Markov Model which can represent a network topology of 3 nodes, all sending packets to each other with equal probability. Fig. 2.17 (b) represents a Four state Markov Model which can be modelled as a token ring structure with four packet emitting nodes and Fig. 2.17 (c) is another representation of the four-node token ring structure but with diagonal transitions. Fig. 2.17 (d) represents a Five state Markov Model that can be modelled on a media server

communicating with nodes rest of the network and lastly, Fig. 2.17 (e) is loosely based on a SIP (Session Initiation Protocol) session, where the SIP server (State 5) is trying to establish a SIP based VoIP session between two end-user clients (States 6, 7) via their Proxy (States 1 and 3) and Re-direct (States 2, 4) server respectively. More details on SIP can be found in references such as [49].

### 2.2.3.2 Calculating State Visiting Probabilities for higher order Markov Models

#### Graphical Approach (the drawback)

In this sub-section, we will look at deriving the state visiting probabilities by considering a complex example of a Markov traffic model. We will use the four-state Markov traffic model show in Fig. 2.18:

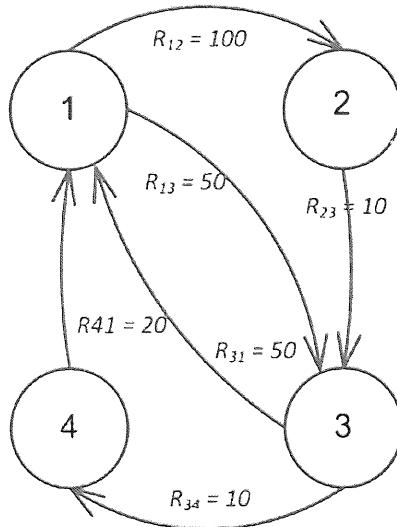


Figure 2.18: State Transition Diagram of Four-State Markov Model with diagonal transitions

Let the branching probabilities of State 1 be  $x$  and  $y$  and State 3 be  $p$  and  $q$  respectively as shown in Fig. 2.19. We will now derive the visiting probabilities  $V_1, V_2, V_3$  and  $V_4$  respectively using graph theory.

If we start in state 1, the state sequence can be illustrated as shown in Fig. 2.19.

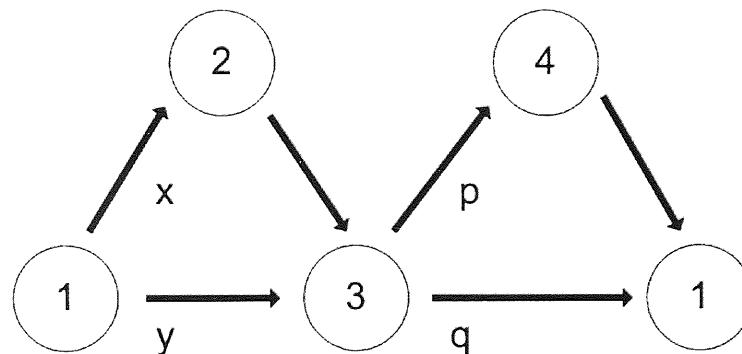


Figure 2.19: Calculating State Transition Probabilities for a Four-state Model with diagonal transitions  
[63]

From Fig. 2.19, we have:

State 3 and states 1 have equal weights as they follow each other. So,

$$V_1 = V_3,$$

From the state transitions we have,

$$V_2 = y \times V_1.$$

$$\text{where } y = 1 - x = 1 - \frac{R_{13}}{R_{13} + R_{12}} = 100/150,$$

and

$$V_4 = p \times V_3 = p \times V_1.$$

$$\text{where } p = \frac{R_{34}}{R_{34} + R_{31}} = 10/60.$$

Therefore, we obtain:

$$V_1 + y \times V_1 + V_1 + p \times V_1 = 1.$$

This gives:

$$V_1 = 1 / (1 + \frac{100}{150} + 1 + \frac{10}{60}) = 0.353.$$

Substituting  $V_1$ , we get  $V_2$ ,  $V_3$  and  $V_4$  as follows:

$$V_2 = y \times 0.353 = 0.235,$$

$$V_3 = 0.353,$$

$$V_4 = p \times 0.353 = 0.05899.$$

Next, we compare these state visiting probability values to a plot of the histogram of the counts of visits to the states 1, 2, 3 and 4 in the Monte Carlo simulation of the Markov Model as shown in Fig. 2.20. Again, the number of iterations of the Monte Carlo simulation are  $N = 1$  million.

If we normalize the counts of visits to 1, we obtain the values  $V_1 = V_3 = 0.353$ ,  $V_2 = 0.235$  and  $V_4 = 0.059$ , as can be visually clarified from the histogram and this proves that the calculation of the visiting probabilities is right.

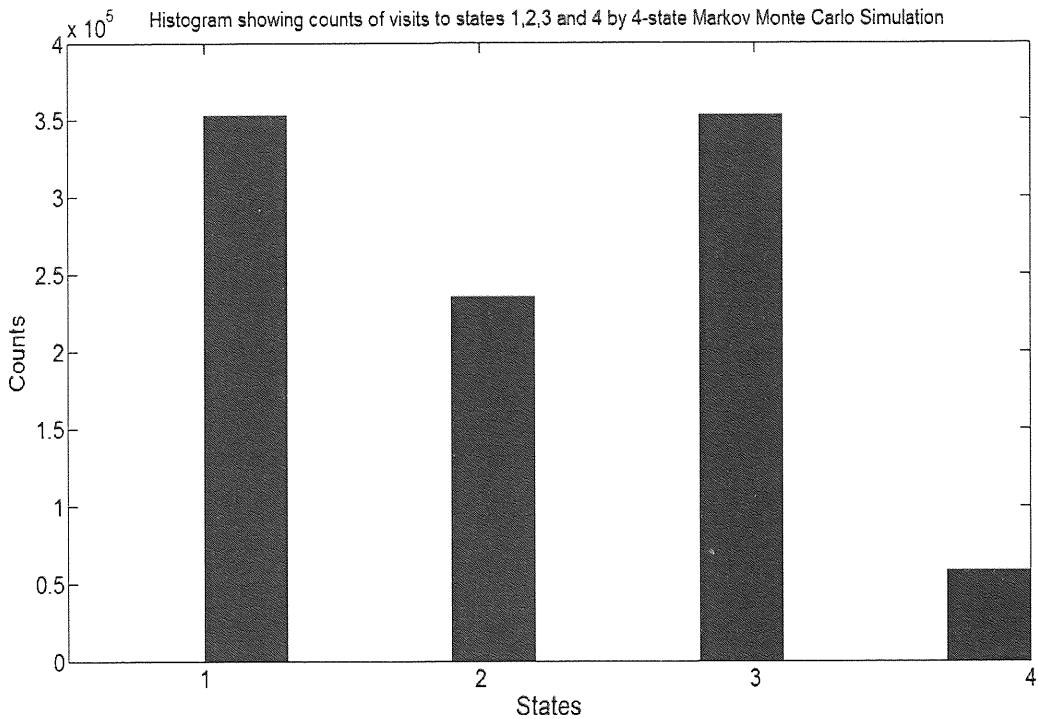


Figure 2.20: Histogram of counts of visits to states 1,2,3 and 4 by 4-state Monte Carlo Simulation

### 2.2.3.3 Using Balance equations to calculate the state visiting probabilities

It is clear from the considered four-state Markov traffic model that as the number of states and transitions increase, the calculation of the state visiting probabilities becomes very complicated with the graphical approach. So it is important that in order to achieve modelling efficiency, the state visiting probabilities must be computed efficiently using an efficient method that can be programmed. The calculation of the state visiting probabilities as the number of transitions approach infinity can be done by using the standard *balance equations* described in standard texts [45] and also introduced briefly in section 2.1.3. As mentioned before in section 2.1.3, the state visiting probabilities can be calculated from the limiting distribution of the Markov model. The limiting distribution and its probability values can be computed using the balance equations, which balance the probabilities of leaving and exiting from a state when the model is in a steady state. These balance equations are given as follows:

$$V_i \sum_{j \neq i} P_{ij} = \sum_{j \neq i} P_{ij} V_j \text{ where, } i \in S.$$

In the vector notation, this can be written as:

$$V \cdot P = V.$$

Where the ‘.’ represents the multiplication of a vector by a matrix and  $V$  is the limiting vector containing the row elements of  $V_i$ . This again can be rewritten as:

$$V \cdot (P - I) = 0.$$

Where ‘I’ is the identity matrix and ‘0’ is a zero matrix. Now, we need to solve for ‘V’ such that the sum of its elements is 1, in other words we need to find the normalised left Eigen vector of the transition matrix  $P$ , with Eigen value 1. Setting up a program for this, and computing the vector values for the symmetric case of the four state Markov Model, we find  $V = [0.3535, 0.2351, 0.3535, 0.05790]$ , which are the visiting probabilities  $V_1, V_2, V_3$  and  $V_4$  which are the same as those observed by the graph theory method as well as consistent with observations from Fig. 2.20. This procedure can be applied to all  $n$ -state ergodic Markov Poisson models, where  $n$  is a finite number.

#### 2.2.3.4 Probability Density Function for n-state Markov Models

Fig. 2.21 shows some examples of the nature of the Probability Density results for IPT obtained for the 3-state and the 5-state models shown in Fig. 2.17 (a) and Fig. 2.17(d) respectively.

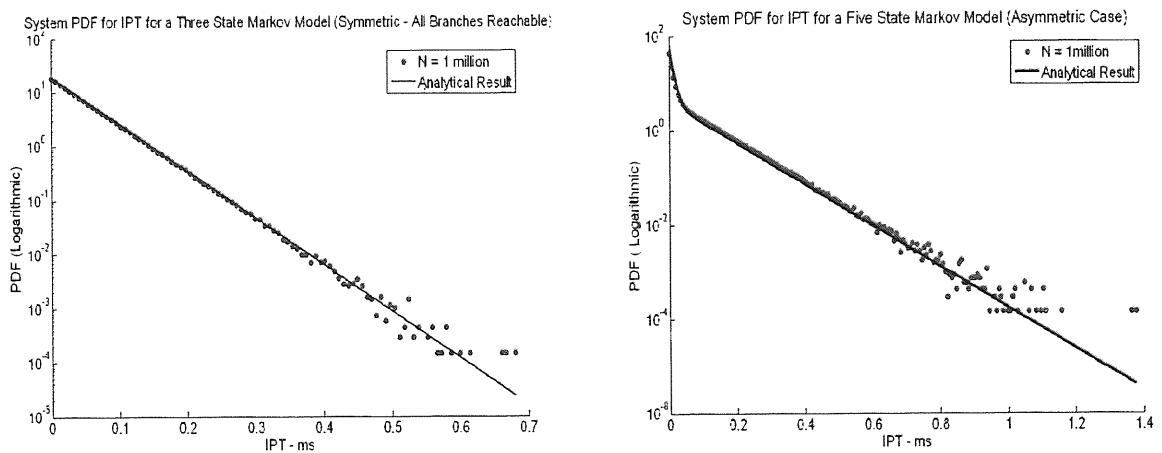


Figure 2.21: PDF Results for (a) 3-state (All branch symmetric) and (b) 5-state (Asymmetric Model)

We have seen with all the variety of cases tried so far that the density results are Poisson in nature. Also, the analytical PDF equations for network IPT for all Markov Models have been in perfect

agreement with all the numerical results confirming our analytical understanding of the Poisson Markov Traffic Models. So this can be confirmed to be rightly applicable for all cases of n-state Markov Models where  $n$  is a finite number.

So, then, the general equation for the n-state Model can be written as:

$$\text{PDF}_{(n\text{-state})} = V_1 P_1(t) + V_2 P_2(t) + V_3 P_3(t) \dots \dots \dots + V_n P_n(t). \quad (2.27)$$

And if the equation represents a model where the model behaves in a loop chain or has all of its states accessible by every of the other states with equal probability then (2.27) can be simplified and written as:

$$\text{PDF}_{(n\text{-state})} = [P_1(t) + P_2(t) + P_3(t) \dots \dots \dots + V_n P_n(t)]/n \quad (2.28)$$

where,  $n$  is the number of states. The equations (2.24) and (2.26) hold good for models where every state is emitting a packet and is *eventful* in some manner. Of course this is again all true for Ergodic models only. We do not consider the Joint density results for the higher n-state models as they have similar properties to those of the two-state model.

## 2.3 Markov Traffic Models with non-Poisson Statistics

So far, we have looked at simple Markov Models which have simple and straight forward Poisson statistics. However, in our measurements, which will be covered in detail in the next chapter, we have observed other non-Poisson elements such as a ‘peak’. In this section we will be looking at Markov models with one or more states that are *uneventful* i.e. they do not emit packets when they are visited but yet they do spend an exponential amount of time in that state before moving on to the next state. We will start with a simple two state model.

### 2.3.1 Two state Markov Traffic Model with one state emitting packets

#### 2.3.1.1 Symmetric Case

For ease of understanding, we will consider the same two-state Markov model considered in the previous section but this time with a difference. One of the states does not emit a packet. The state with the outward transition represented by a dotted arrow line instead of a solid one represents the state with no packet being emitted on that path. From a network perspective, for this model, the client sends a packet to the server, but the server does not respond back.

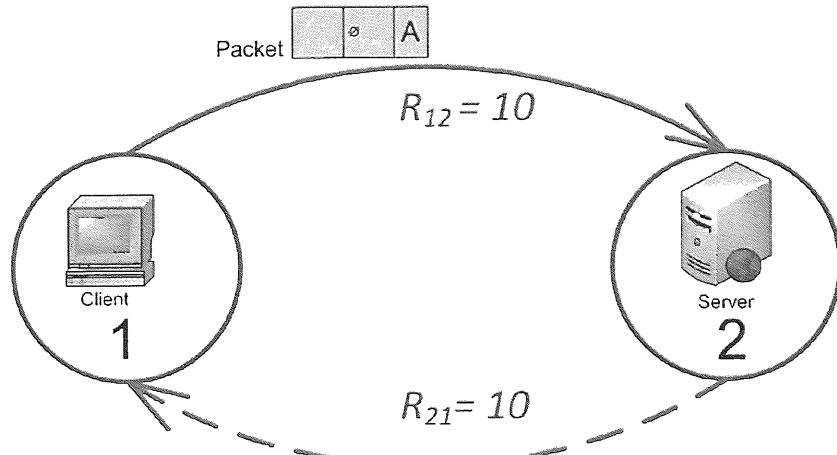


Figure 2.22: 2- state Markov Model with only one state emitting packets (Symmetric)

### 2.3.1.1.1 Probability Density Function – Mathematical Analysis

As can be seen from Fig. 2.22, state 2 is not emitting a packet. As before, the Probability distribution of the occupation times of the states are represented by Equations (2.11) and (2.12) respectively.

We have the Rate Matrix,  $Q_{ij}$  again given as:

$$Q_{ij} = \begin{pmatrix} -10 & 10 \\ 10 & -10 \end{pmatrix}$$

Now using equation (2.2), we have the Transition Probability Matrix  $P_{ij}$ :

$$P_{ij} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The analytical equation now differs from what we have seen in the previous cases. Because we identify a packet transmission with only a single change of state i.e. from state 1 to state 2, then the PDFs are modified as the time in state 2 now has to be included. The sequence of states '1212' corresponds to two packets sent and happens when the model spends an unknown time  $t'$  in state 2 and a known time  $t-t'$  in state 1 and so the PDF for the IPT would be a convolution of the probability densities of times spent  $t'$  in state 2 and times  $t-t'$  spent in state 1. The equation (2.29) represents the analytical solution for the PDF of IPT for this model.

$$PDF = \int_0^t P_2(t') P_1(t-t') dt'$$

(2.29)

Where,

$$P_1(t) = R_{12} e^{-R_{12}(t-t')}$$

$$P_2(t) = R_{21} e^{-R_{21} t'}.$$

### 2.3.1.1.2 Probability Density Function – Markov Monte Carlo Results

Substituting these probability densities and evaluating the integral and then plotting the analytical solution with the numerical solution obtained from the Markov Model Simulation we see the graph as shown in Fig. 2.23:

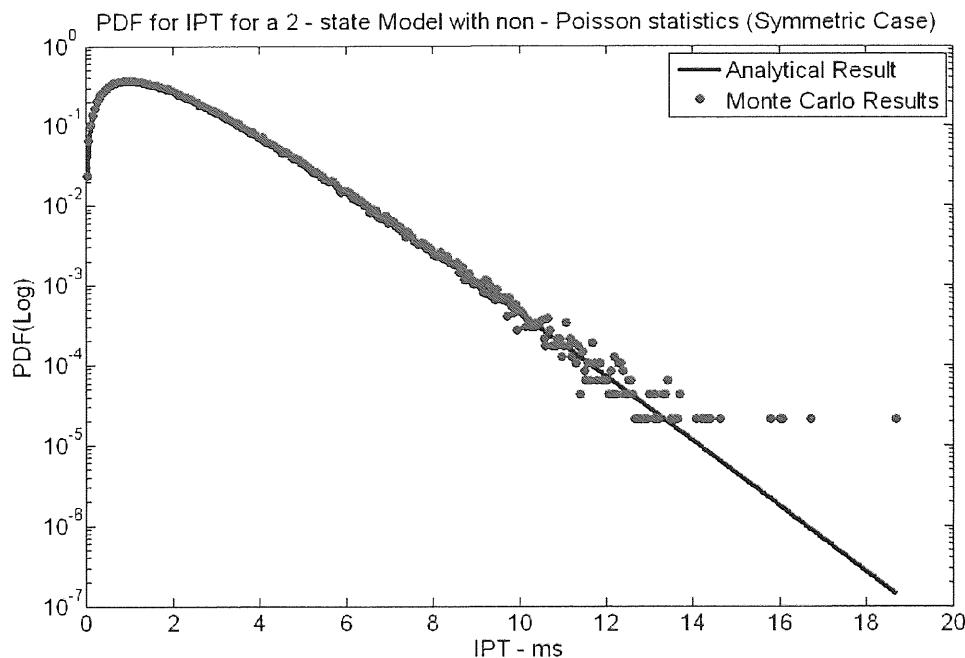


Figure 2.23: PDF for IPT for 2-state Model (Symmetric Case) with only state 1 emitting packets

As can be seen from the graph, we now have a peak. The maximum is no longer at the origin. The initial part of the PDF curve ranging from 0 to 2 seconds represents those IPT times which have come from state 2, where the state is actually not emitting a packet. In this case of the model, the exponential amount of time spent in state 2 gets added to state 1, and so the curve takes a downward trend because of the absorption of the inter-arrival times based on the rate  $R_{21}$  into state 1. Let us now look at the Joint Density Function.

### 2.3.1.1.3 Joint Density Function – Mathematical Analysis

In order to determine the analytical equation, we know that there is only one packet generation process associated with this two state Markov Model. Then, the JDF will be given by the equation (2.30):

$$P(t_1, t_2) = P(t_1)P(t_2) \quad (2.30)$$

Equation (2.27) is simply the product of the two state holding time distributions for times  $t_1$  and  $t_2$ .

A plot of the analytical solution represented by equation (2.30) is shown in Fig. 2.24.

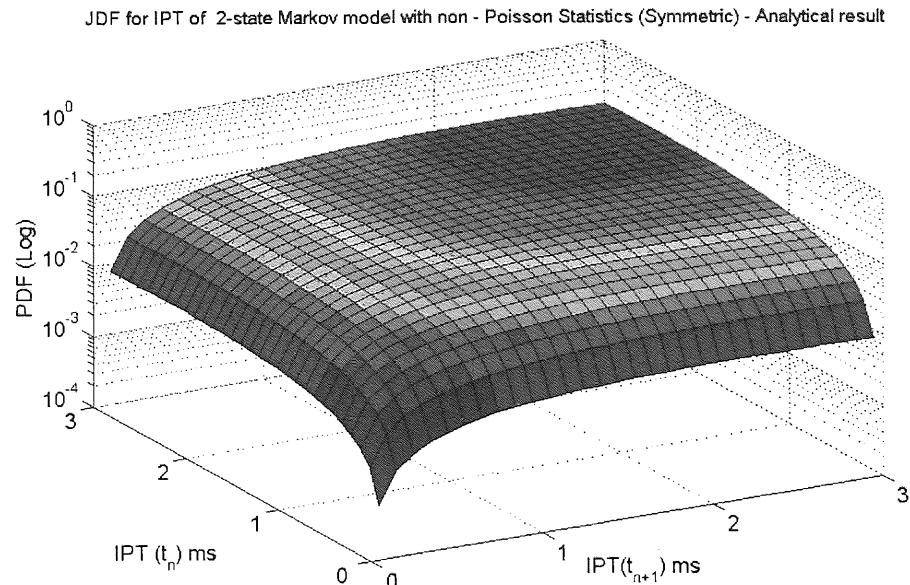


Figure 2.24: JDF for 2-state Model (one state emitting packet) (Symmetric Case) - Analytical result

#### 2.3.1.1.4 Joint Density Function – Monte Carlo Results

Using the same procedure as with the Poisson 2-state Markov Model, we obtain the numerical results for the Joint Density function as shown in Fig. 2.25.

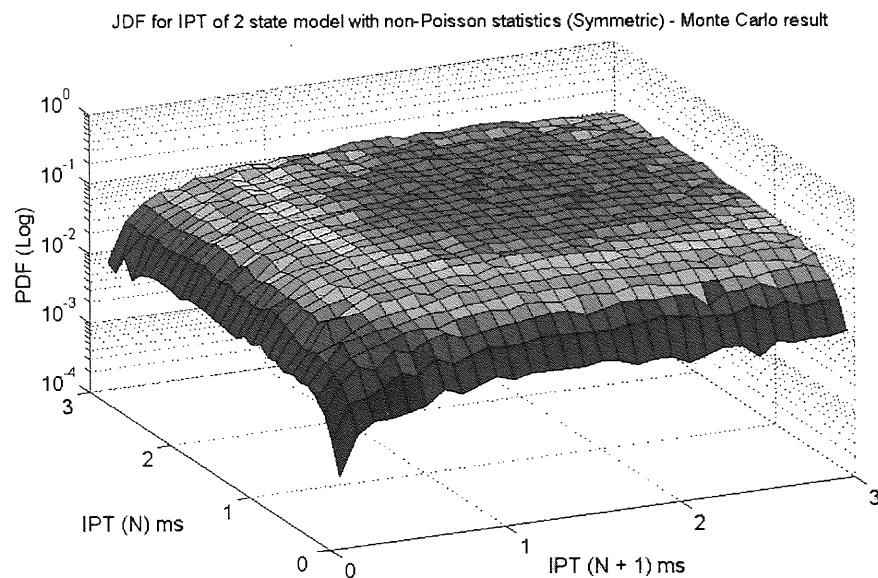


Figure 2.25: JDF for 2-state Model (one state emitting packet) (Symmetric Case) – Monte Carlo result

Comparing Figs. 2.24 and 2.25 we find that they are in close agreement with each other. The top section of the three dimensional peak can be closely observed. If set on a linear density scale, the shape of the surface would look like that of a quadrant of a hemisphere. Next we will look at the results of the asymmetric case as shown in Fig. 2.25.

### 2.3.1.2 Asymmetric Case

Now let us consider the asymmetric case for the same two-state Markov Model as shown in Fig. 2.26.

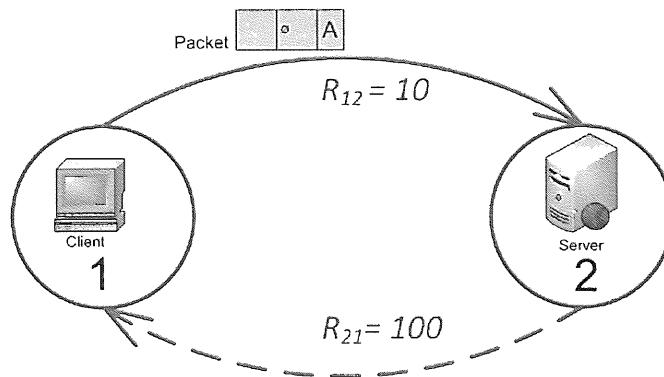


Figure 2.26: 2- state Markov Model with only one state emitting packets (Asymmetric case)

#### 2.3.1.2.1 Probability Density Function - Analytical and Monte Carlo Results

Once again, the analytical solution for the PDF would be the same as that shown in equation (2.29).

Substituting the new rates into the equation, evaluating the integral and plotting the equation

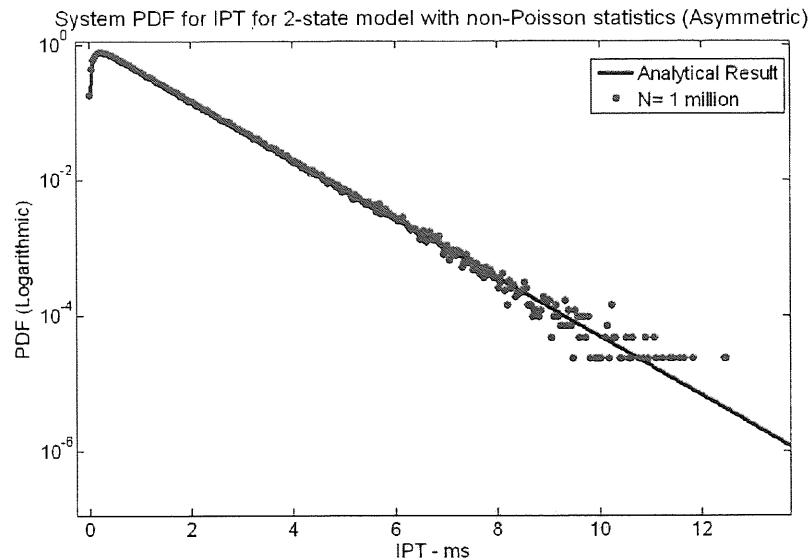


Figure 2.27: PDF for IPT for 2-state model (Asymmetric Case) - with only state 1 emitting packets

against the numerical results as seen in Fig. 2.27, we see that the numerical results and analytical results agree with each other proving that the general equation given by (2.29) holds good for any two-state Markov model with one of the states not emitting packets. The difference between the case of the asymmetric case and the symmetric case is also noticeable where for the asymmetric case the bending of the curve is sharper as it represents the sum of a higher transition rate and a slower transition rate, where as on the other hand, it is smoother for the symmetric case as it is a sum of two equal transition rates.

### 2.3.1.2.2 Joint Density Function – Analytical and Monte Carlo Results

The Equation (2.30) is used again with the new rates, and the analytical and simulated results are shown in Figs.2.28 and 2.29 respectively.

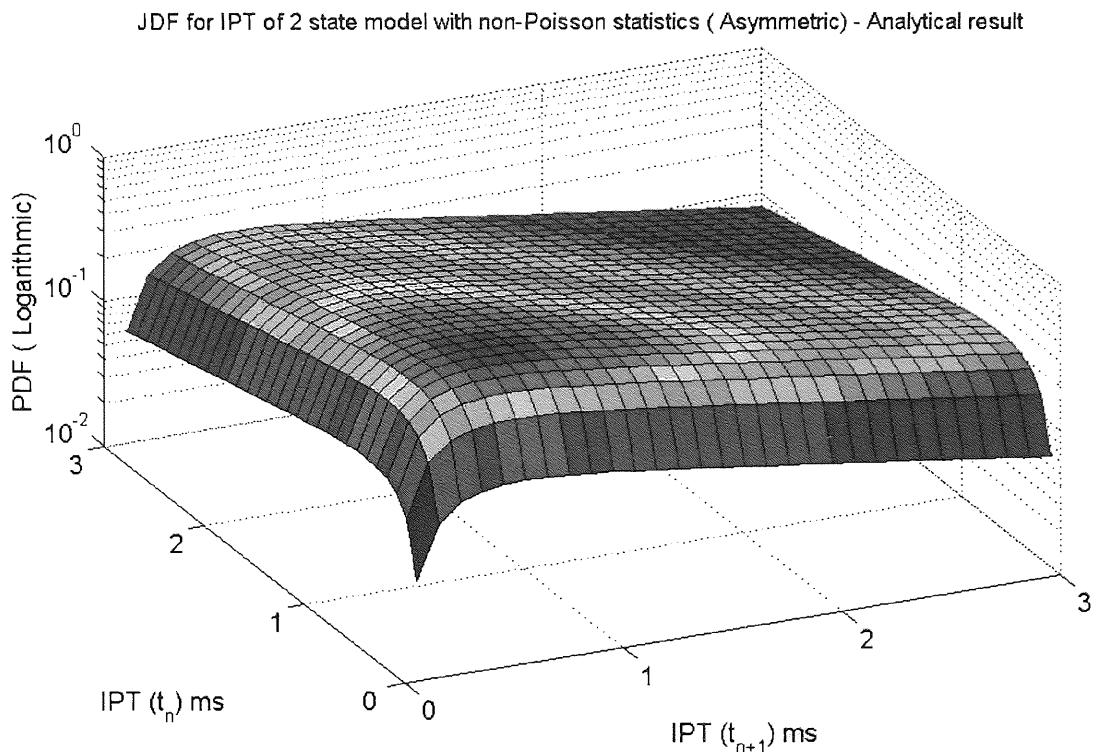


Figure 2.28: JDF for 2-state Model (one state emitting packet) (Asymmetric Case) – Analytical result

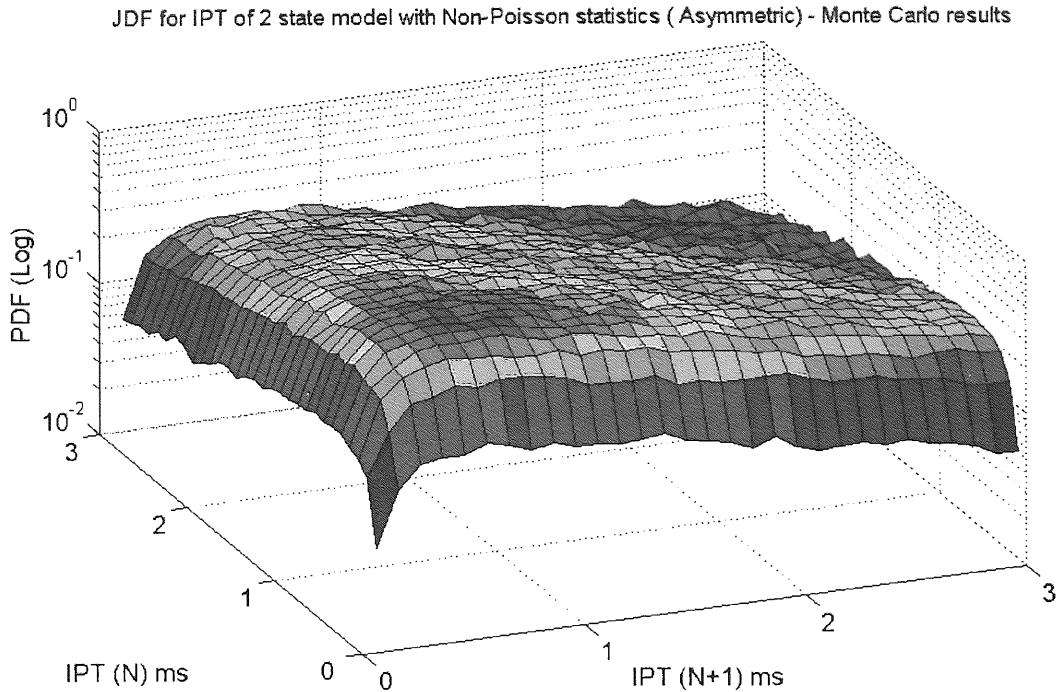


Figure 2.29: JDF for IPT 2-state Model (one state emitting packet) (Asymmetric Case) – Monte Carlo

Again we notice that the numerical and the analytical results of the Joint Densities of the asymmetric versions are in agreement with each other. The surface has a steeper end towards the front end side, because of the sum of the exponential times based on the asymmetric transition rates.

### 2.3.2 n-state Markov Traffic Models with only one state emitting packets

#### 2.3.2.1 Some Examples

Now that we have understood the non-Poisson statistics of the two-state Markov Traffic Model, we can move further, by considering  $n$ -state non-Poisson Markov Traffic Models by adding  $n-1$  uneventful states in a loop fashion. As examples, shown in Fig.2.30(a) and (b), are 3-state and 4-state loop non-Poisson traffic Models, of which only the state 1 emits a packet with a common transition rate.

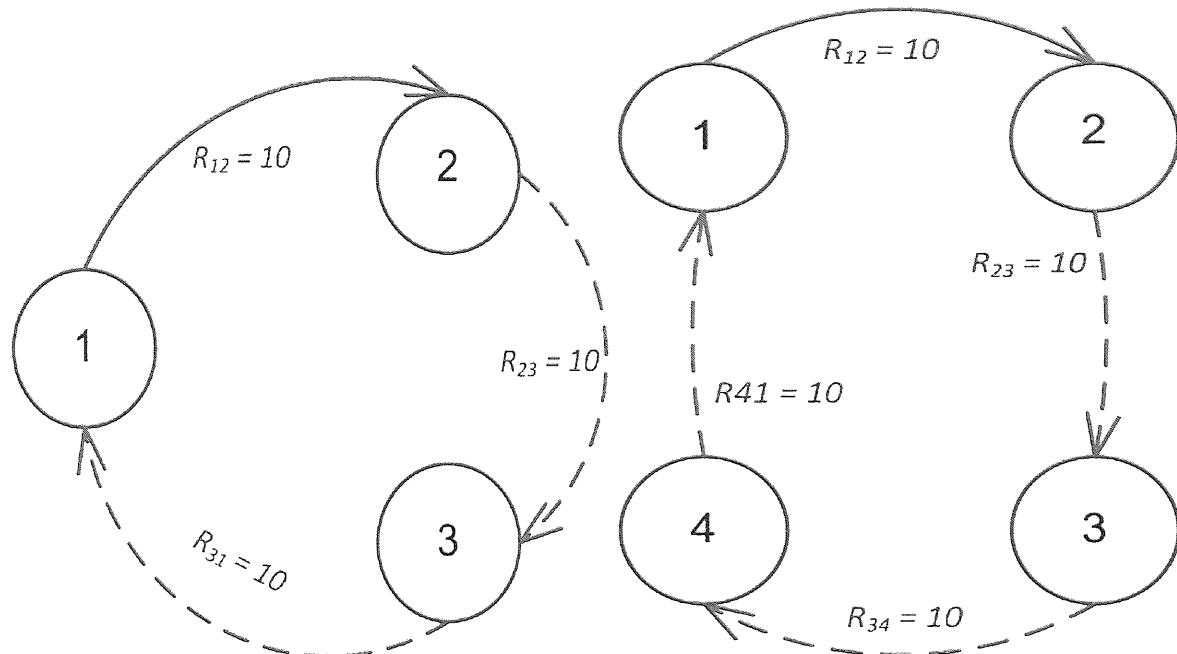


Figure 2.30: (a) Three-state and (b) Four-state loop Markov model with only state 1 emitting packets

### 2.3.2.2 Probability Density Function for n-state non-Poisson loop Markov Model

#### 2.3.2.2.1 Analytical and Monte Carlo simulated results

Let us first consider the three-state loop Markov model shown in Fig 2.30(a).

The Rates involved are  $R_{12} = R_{23} = R_{31} = 10$ .

Now for time  $t$ ,

$$P_1(t) = R_{12} e^{-R_{12}t}$$

$$P_2(t) = R_{23} e^{-R_{23}t}$$

$$P_3(t) = R_{31} e^{-R_{31}t}$$

The analytical solution of the PDF of IPT for the three state Markov model can be derived with the aid of the analytical solution of the two-state loop model shown in equation (2.29). Let us assume that the model spends time  $t''$ ,  $t-t'$ , and  $t-t'-t''$  in three states respectively. This establishes that for the three- state model where only state one is emitting a packet, the equation can be obtained as:

$$PDF = P_{123}(t) = \int_0^t P_{12}(t') P_3(t - t') dt' \quad (2.31)$$

The Equation (2.28) represents the PDF of state one emitting a packet and is a convolution of the terms  $P_{12}(t')$  and  $P_3(t-t')$ , where  $P_{12}(t')$  is given by:

$$P_{12}(t') = \int_0^{t'} P_1(t'') P_2(t' - t'') dt'' \quad (2.32)$$

Now substituting the equation 2.32 in equation 2.31, we get for the three-state loop model, the following PDF equation:

$$\begin{aligned} PDF_{(3\text{state-loop})} &= P_{123}(t) \\ &= \int_0^t \int_0^{t'} R_{12} e^{-R_{12}t''} R_{23} e^{-R_{23}(t'-t'')} dt'' R_{31} e^{-R_{31}(t-t')} dt' \end{aligned} \quad (2.33)$$

Finally, substituting the values of the rates, i.e  $R_{12} = R_{23} = R_{31} = 10$  in equation (2.33) and solving it, gives us the expression:

$$PDF_{(3\text{-state loop})} = 500 t^2 e^{-10t} \quad (2.34)$$

Next we will add another non-packet emitting state to the three-state loop model, thus creating a four-state loop model example as shown in Fig. 2.30 (b). Following the same process as with the case of the two-state and three-state loop examples, the PDF for the four-state loop model is given by the three step convolution of the four state holding time distributions as shown in equation (2.35).

$$PDF = P_{1234}(t) = \int_0^t P_{123}(t') P_4(t - t') dt' \quad (2.35)$$

where,

$$P_{123}(t') = \int_0^{t'} P_{12}(t'') P_3(t' - t'') dt''$$

and

$$P_{12}(t'') = \int_0^{t''} P_1(t''') P_2(t'' - t''') dt'''$$

Upon substituting the rates  $R_{12} = R_{23} = R_{34} = R_{41} = 10$ , and the probability densities in equation (2.35), we have for the four-state Markov model, the following numerical equation:

$$PDF_{(4\text{-state loop})} = \frac{5000}{3} t^3 e^{-10t} \quad (2.36)$$

Here we see that the difference between the equation of the four-state loop Model (Equation (2.36)) and the three-state loop model (Equation(2.34)) is that of the power of 't' and the amplitude.

Similarly if we continue with the addition of another non-packet emitting state to the model to make it a 5-state Markov Model, we then have the general equation for the PDF as:

$$PDF = P_{12345}(t) = \int_0^t P_{1234}(t') P_5(t - t') dt' \quad (2.37)$$

Where  $P_{1234}(t')$  is given by equation (2.35).

Substituting the rates and densities in equation (2.37) gives:

$$PDF_{(5\text{-state})} = \frac{12500}{3} t^4 e^{-10t} \quad (2.38)$$

This process can be further extended to any n-state loop models, where  $n$  is a finite number.

So for a ten-state loop model, the general PDF equation will be:

$$PDF = P_{123456789 \rightarrow 10}(t) = \int_0^t P_{1234569}(t') P_{10}(t-t') dt' \quad (2.39)$$

Evaluating the equation (2.39) gives:

$$PDF_{(10-state)} = \frac{15625000}{567} t^9 e^{-10t} \quad (2.40)$$

Next, we plot all together, the analytical solutions and Monte Carlo simulation results ( $N = 1$  million) of the PDF for IPT based on the three, four, five, six, seven, eight and ten-state loop models with only State 1 emitting packets, as shown in Fig. 2.31. The density is on a linear scale.

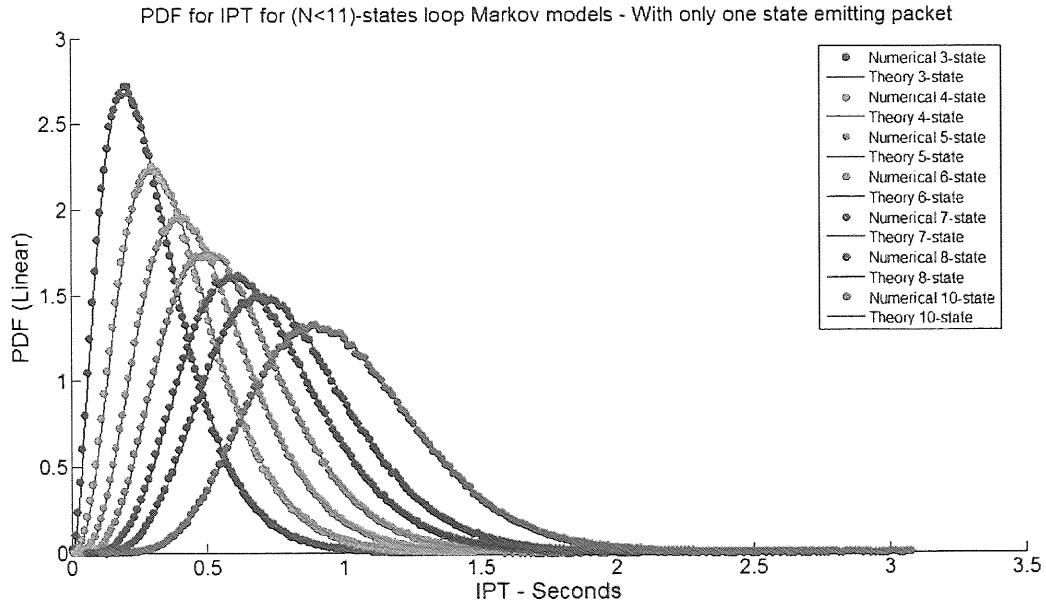


Figure 2.31: PDF (Linear scale) for IPT for ( $n < 11$ ) states loop Markov Models - With only state 1 emitting packets

These results show that as the number of states increase in the loop chain, where only State 1 is emitting packets, we notice that the peak gets pushed further away from the origin. In addition, the amplitude of the peak decreases and the width of the peak increases.

### 2.3.2.2.2 General PDF Equation for the n-state loop non-Poisson model – The Gamma distribution

Following the results obtained from the equations (2.29), (2.34), (2.36), (2.38) and (2.40) for the PDF of the IPT based on the two, three, four, five and ten-state loop non-Poisson models, we conclude that the general n-state IPT PDF equation could be generalised to:

$$PDF_{(Loop, N-state)} = \frac{R^n}{(n-1)!} \cdot t^{n-1} e^{-Rt} \quad (2.41)$$

where,  $n$  is the number of states,  $R$  is the transition rate parameter, and  $t$  is the time variable.

With reference to statistical theory, equation (2.41) actually represents a Gamma distribution.

The Gamma distribution is defined as a distribution representing the sum of  $n$  number of independent exponentially distributed random variables with a common rate parameter  $R$ . From the packet model point of view, the packet once emitted from the originating state of the n-state loop, and then spends an exponential amount of time in each of the rest of the states before the next packet is emitted. Thus the Inter-packet time for this case is the sum of exponential times associated with all the  $n$  states in the n-state loop model which have a common transition rate parameter  $R$ . Thus using this fundamental theory it is possible to reduce an n-state loop model to a single state which has a Gamma holding time distribution. This is explored in section 2.4.

## 2.4 Semi-Markov Models: A Gamma Markov Model

### 2.4.1 30-state Model with non- Poisson statistics

We will start by looking at a 30 state Model, which consists of 3 sets of 10-state loop models with only one state emitting packets. The model is configured in such a way such that it has a common transition rate parameter unique to each set of 10-states loop model and the 10<sup>th</sup> state of each set of loop model is linked to the first state of the next 10-state loop model. Only the first of each ten-state sets of the 30-state model emit packets. The state transition diagram of the 30-state model is shown in Fig. 2.32. For the sake of understanding, each 10 state block has been represented as one cloud of states. The rates for first, second and third set of ten states are 10, 20 and 30 respectively.

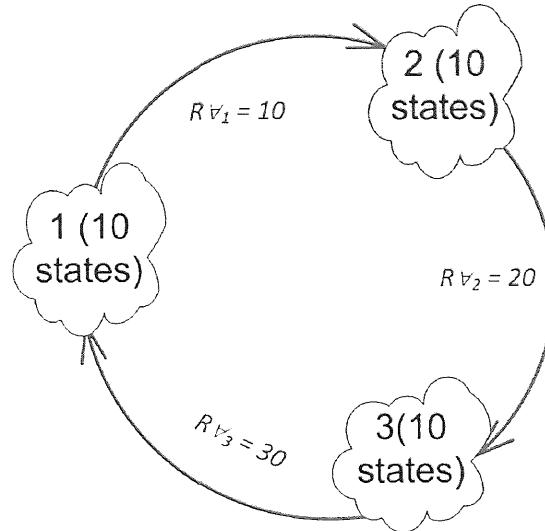


Figure 2.32-state Markov Model with non-Poisson statistics

In the first cloud of ten states, State 1 emits a packet and the model transits through the remaining 9 states with rates  $R\forall 1 = 10$  till it reaches state 11 which emits a packet and then the model transits with rates  $R\forall 2 = 20$  through the next nine states again till it reaches state 21 where this state emits a packet and again the model transits through the last remaining nine states till it reaches state 1 back again, with a rate  $R\forall 3 = 30$ . So, there are three packets per cycle.

#### 2.4.2 Simplifying the 30 state model with a 3 state Gamma Markov Model

We know that for the first cloud of ten states shown in Fig. 2.32, the distribution or PDF can be represented by a Gamma distribution which takes the form:

$$PDF_{(Loop, N-state)} = \frac{R^n}{(n-1)!} \cdot t^{n-1} e^{-Rt} \quad (2.41)$$

Upon substituting the transition rates from Fig.2.32, the PDF for the first 10 states is given by:

$$PDF_{(10-states)} = 27557 \cdot t^9 e^{-10t} \quad (2.42)$$

For the next ten states for  $R\forall 2 = 20$ , we have:

$$PDF_{(2nd 10-states)} = 28218694 \cdot t^9 e^{-20t} \quad (2.43)$$

And finally for the last ten states for  $R\forall 3 = 30$ , we have:

$$PDF_{(3rd 10-states)} = 1.627232142 \cdot t^9 e^{-30t} \quad (2.44)$$

Based on the three equations (2.42), (2.43) and (2.44), we can model these as a three-state Gamma Loop Model replacing each ten states shown in Fig. 2.32 with a single state which carries a Gamma holding time distribution, resulting in the model as shown in Fig. 2.33.

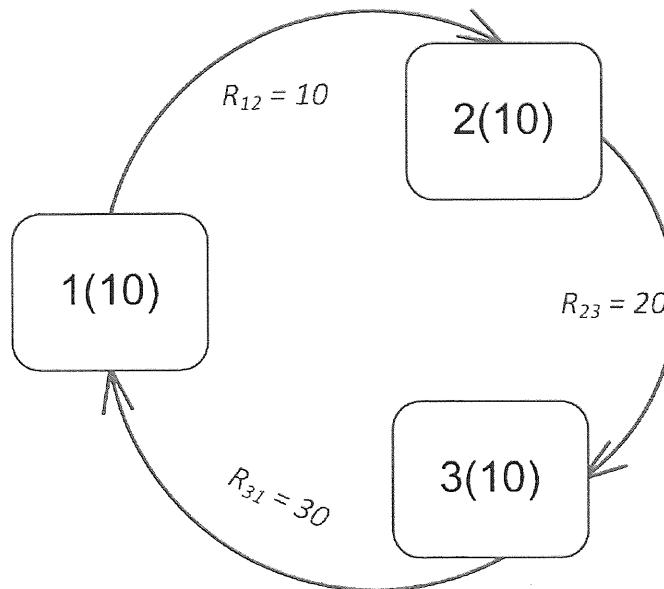


Figure 2.33: Simplified 3-state Gamma Markov Model

The three states in the Gamma Markov Model have now been represented by curved cornered rectangles instead of the circles to distinguish the difference between the earlier exponentially distributed Markov models with these Gamma distributed ones. The number of states, that is 'n', to be associated with the Gamma distribution of each state is given in brackets by the side of the state identifier number.

Using equation (2.41), the Gamma holding time distribution for each of the three states can be simply represented in the form  $\Gamma(n, R)$  as follows:

$$P_1(t) = \Gamma(n, R) = \Gamma(10, 10) \quad (2.45)$$

Similarly for State2 and from Equation (2.43):

$$P_2(t) = \Gamma(10, 20) \quad (2.46)$$

And finally for State 3 and from Equation (2.44):

$$P_3(t) = \Gamma(10, 30) \quad (2.47)$$

Thus, we have presented that the Probability distribution of occupation times of the three states can be represented by their relevant Gamma distributions respectively.

## Probability Density Function for the 3-state Gamma Markov Model

Since the model is now reduced to a 3-state model where every state is emitting a packet, we can use equation (2.27) for the PDF. Because it is a loop model, the three states are visited equally often, and hence  $V_1 = V_2 = V_3 = 1/3$ . Substituting the holding time distributions given in equations (2.45 – 2.47) and the state visiting probabilities in equation (2.27) gives us the analytical solution for the 3-state Gamma Markov model shown in equation (2.48).

$$PDF_{(3\text{-state Gamma})} = \frac{1}{3}\Gamma(10,10) + \frac{1}{3}\Gamma(10,20) + \frac{1}{3}\Gamma(10,30) \quad (2.48)$$

### 2.4.3 Using the 3-state Gamma Markov Model in a Monte Carlo Simulation

The Monte Carlo simulation process used for simulating the 3-state Gamma Markov Model is similar to the earlier approaches, with the only exception that a Gamma distribution is used instead of an exponential for generating the IPT samples. Fig. 2.34 shows the Monte Carlo Simulated results for the PDF of IPT observed from both the 30-state model and the 3-state Gamma Markov Model shown in Figs. 2.33 and 2.34 respectively, along with the analytical solution represented by equation (2.48).

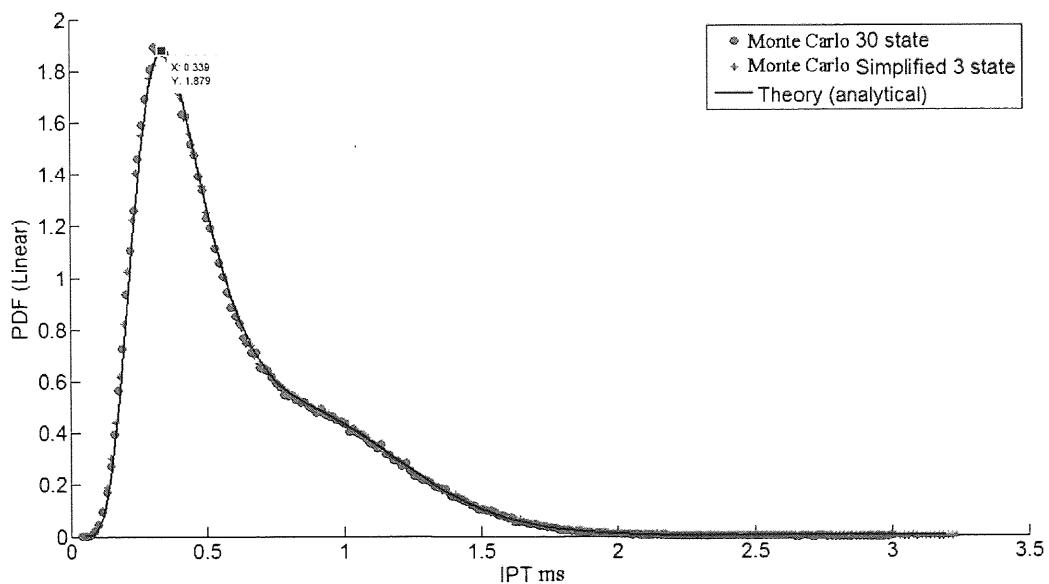


Figure 2.34: Comparison of Monte Carlo and analytical results for 30-state and 3-state Gamma Model

We see that the results of the 3-state simplified Gamma Markov Model overlap with the results of the 30-state loop model, and also agree with the analytical solution, thus justifying the use of these Gamma Markov Models for generating non-Poisson statistics.

## 2.5 Semi-Markov Models: A Gaussian Markov Model

### 2.5.1 Disadvantages of using a Gamma Markov traffic model

While there is nothing wrong with using the Gamma Markov Models, they do have a slight disadvantage when it comes to modelling the various peaks that are associated with the various states. To study this disadvantage, we will consider a four-state Gamma Markov Model as shown in Fig. 2.35. This hypothetical traffic model example is loosely based on some random measured IPT PDF result.

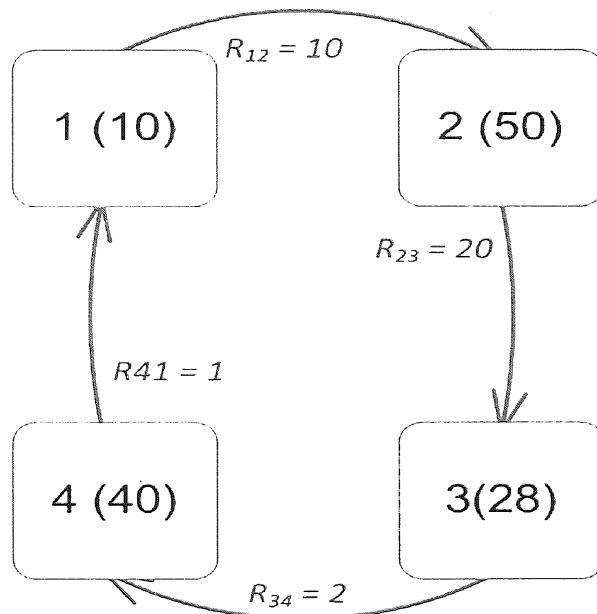


Figure 2.35: State Transition Diagram of a 4-state Gamma Markov Model

The equation for the PDF of IPT for the four-state Gamma Markov mode is given as follows:

$$PDF_{(4\text{-state Gamma})} = \frac{1}{4}\Gamma(10,10) + \frac{1}{4}\Gamma(50,20) + \frac{1}{4}\Gamma(28,2) + \frac{1}{4}\Gamma(40,1) \quad (2.49)$$

Plotting the analytical result solution given by (2.49) and the Monte Carlo simulation Results for the four-state Gamma Markov model together, we get the graph shown in Fig. 2.36:

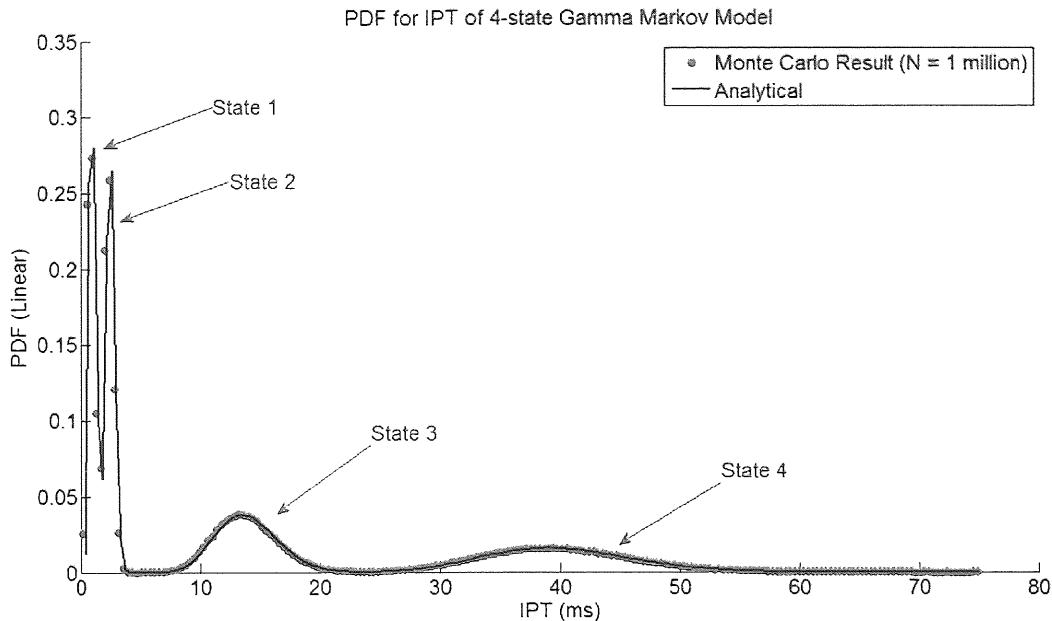


Figure 2.36: PDF for IPT of 4-state Gamma Markov Model

The peaks associated with each Gamma state are marked clearly in the figure. The process of modelling traffic requires that a model is produced based on the PDF where a single Gamma distribution is fitted to each peak. In order to do this, estimating the two parameters 'n' and 'R' of the Gamma distribution is required such that the estimated values determine the location of the 'maximum' of each peak. The location of the maximum of each location is given by  $n/R$ .

For State 1:  $n_1 = 10, R_1 = 10$

The location of the maximum of the peak =  $n/R = 1$  ms

For State 2:  $n_2 = 50, R_2 = 20$

The location of the maximum of the peak =  $n/R = 2.5$  ms

For State 3:  $n_3 = 28, R_3 = 2$

The location of the maximum of the peak =  $n/R = 14$  ms

For State 4:  $n_4 = 40, R_4 = 1$

The location of the maximum of the peak =  $n/R = 40$  ms

The width and the amplitude of the peak also varies with the values chosen for 'n' and 'R'.

Thus, as seen, in order to fit a Gamma component, the values ‘n’ and ‘R’ should be adjusted such that the peak, location, width and amplitude match well to our requirements. This is a complicated procedure for modelling, as opposed to a case of using the Gaussian distribution, where estimation of a peak component is much simpler, as the mean of the Gaussian distribution ‘ $\mu$ ’ determines the location of the peak maximum and the standard deviation ‘ $\sigma$ ’ determines the width of the peak. In the next sub-section, we will take our work further and see how we can use Gaussian Markov Models instead of Gamma Markov Models.

### 2.5.2 Using a Gaussian Markov traffic model

The equation for the Gaussian holding time distribution for each state can be derived from the Gamma holding time distribution of a state. It is known that the limiting case of a Gamma distribution is a Gaussian distribution [50]. This means that in equation (2.41), as ‘n’ approaches infinity; the Gamma distribution takes the form of a Gaussian distribution as shown in equation (2.42).

Let us consider Equation (2.38) again:

$$P(t)_{(Gamma)} = \frac{R^n}{(n-1)!} \cdot t^{n-1} e^{-Rt} \quad (2.41)$$

Now applying a limit of ‘n’ towards infinity:

$$\lim_{n \rightarrow \infty} \frac{R^n}{(n-1)!} \cdot t^{n-1} e^{-Rt}$$

gives us a Gaussian distribution:

$$P(t)_{(Gaussian)} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(t-\mu)^2}{2\sigma^2}} \quad (2.42)$$

where:

Mean,  $\mu = n/R$  and Variance,  $\sigma^2 = n/R^2$

The standard deviation is the square root of the variance and is thus given by:  $\sigma = \sqrt{n}/R$

To illustrate, this limiting principle, we will take a small example. We generate a million samples from a Gamma distribution given by equation (2.41), where we set 'n' to a high value that is 500, and fix 'R' to a constant rate. And we then compare this to a similar Gaussian distribution where the mean  $\mu$  is given by  $n/R$  and the variance  $\sigma^2$  is given by  $n/R^2$  as shown in Fig. 2.37

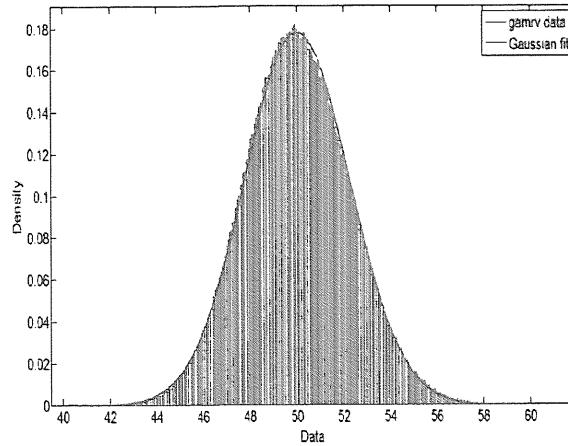


Figure 2.37: Comparison of a Gamma distribution with 'n = 500' with a Gaussian distribution.

As can be seen, as 'n' of the Gamma distribution takes a large enough value, the distribution tends to look more like a Gaussian bell curve. For small or moderate values of 'n' the Gamma distribution shows a right skew to the Gaussian distribution, which reduces as the value of 'n' keeps increasing.

### Probability Density Function for an n-state Gaussian Markov Model

Following suit, as with the Gamma Markov model, and with reference to equation (2.27), the PDF equation for an n-state Gaussian Markov Model can be written as follows:

$PDF_{(n-state, Gaussian)}$

$$= V_1 Gaussian(\mu_1, \sigma_1) + V_2 Gaussian(\mu_2, \sigma_2) + \dots + V_n Gaussian(\mu_n, \sigma_n) \quad (2.50)$$

Where the Gaussian parameters for each of the states are given as:

Gaussian parameters for State 1:

$$\mu_1 = n_1/R_1 = 1, \sigma_1 = \sqrt{n_1/R_1} = 0.316$$

Gaussian parameters for State 2:

$$\mu_2 = n_2/R_2 = 2.5, \sigma_2 = \sqrt{n_2/R_2} = 0.353$$

Gaussian parameters for State 3:

$$\mu_3 = n_3/R_3 = 14, \sigma_3 = \sqrt{n_3/R_3} = 2.645$$

Gaussian parameters for State 4:

$$\mu_4 = n_4/R_4 = 40, \sigma_4 = \sqrt{n_4/R_4} = 6.324$$

The visiting probabilities, of course, remain the same:  $V_1 = V_2 = V_3 = V_4 = 1/4$

Substituting all of these in the equation (2.50), we get the PDF equation for the IPT for the Four-state Gaussian Markov Model as follows:

$$\begin{aligned} PDF_{(4-state, Gaussian)} &= \frac{1}{4} Gaussian(1, 0.316) + \frac{1}{4} Gaussian(2.5, 0.353) + \frac{1}{4} Gaussian(14, 2.645) \\ &\quad + \frac{1}{4} Gaussian(40, 6.324) \end{aligned} \tag{2.51}$$

### 2.5.3 Analysing the Gaussian Monte Carlo simulation results

Since we have derived the Gaussian parameters, we can now set up the Gaussian Markov model for a Monte Carlo Simulation such that a Gaussian distribution is now used for the state holding time distributions.

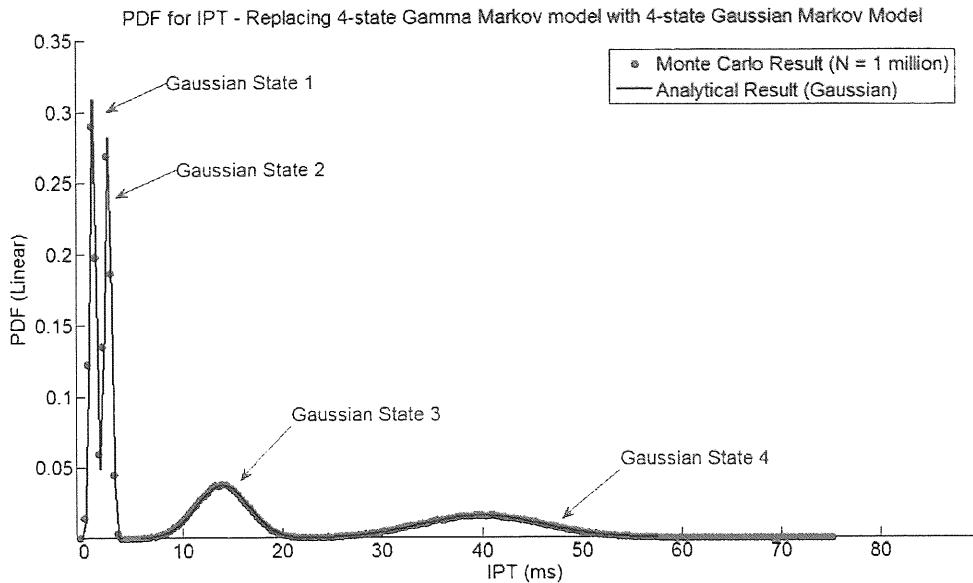


Figure 2.38: PDF for IPT results for the 4-state Gaussian Markov model

Comparing the result of the Four-state Gaussian Markov Model shown in Fig. 2.38 to the result of the Four-state Gamma Markov Model shown in Fig. 2.36, we find that they are very similar. The comparison also establishes the ease and advantage of choosing Gaussian modelling parameters especially with peaks closer to the origin (i.e. for smaller 'n' values of a Gamma distribution). This result paves way for the use of Gaussian Markov Models as a viable tool for modelling measured results. A detailed explanation on the use of Gaussian Markov Models for modelling measured results is discussed in the next chapter.

## 2.6 Summary

In this chapter, we began with explaining the mathematical basics of Markov Models and how they can be used in IP traffic modelling. In section 2.2 we looked at the principle of operation of simple finite-state Markov Models with Poisson statistics, starting with a simple-two state Markov model, moving upwards to higher n-state Poisson models and deriving the general equation for the PDF of the IPT of the models where every state is emitting a packet. We then moved onto consider loop Markov models with non-Poisson statistics in section 2.3, with only one state emitting a packet and concluded that an n-state loop model can be modelled as a single state with a Gamma distribution. Using this important conclusion, we introduced the use of Gamma Markov models in section 2.4 and highlighted their disadvantages over the more convenient Gaussian Markov Models in section 2.5.

# Chapter Three

## Modelling Measured Statistical Results with Gaussian Markov Models

As discussed in Chapter 1, the first step in modelling IP traffic is to obtain measured statistical results for network IP traffic. Dr. Scott Fowler, a Post-doctoral co-researcher, has contributed to the measured traffic statistical results which have been used for traffic modelling and analysis in chapters 3-5 of this thesis. In this chapter, an introduction to the methodology used for measurement along with a few key measured results have been highlighted in sections 3.1 and 3.2 respectively, followed by the modelling procedure using Gaussian Markov Models discussed in section 3.3. Finally, examples of the modelled results are shown in section 3.4 followed by an overall summary in 3.5.

### 3.1 Measuring IP traffic – Methodology

#### 3.1.1 Approach

The approach to measuring IP traffic comprised of having a host PC connected to the University's network, communicating with other server nodes on the internet using different applications that generate IP traffic. Tools and scripts discussed in section 1.1.2, were used to capture packets, via the *Network Interface Card* of the PC, filtered to a particular protocol and specific network *ports*. These *Ports* are defined as application specific end points for network communications that serve as an interface between the host node and the rest of the network. The measurements for the encapsulated IP traffic were made on the Transport layer, filtered using ports associated with the Transport layer protocols namely the *Transmission Control Protocol* (TCP) and the *User Datagram Protocol* (UDP). The measurements were based on the absolute *timestamps* associated with every packet received. The statistical results were based on the Inter-packet times of the packets, generated by calculating the successive differences in the timestamps of every packet captured in sequence of time. Again, as mentioned in Chapter 1, this thesis intends to look at short-term statistics rather than study the averaged long range dependence of traffic.

In an initial experimental attempt to measure IP traffic, a conventional approach was considered, where the packets based on a particular application and *port number* were captured collectively and

the Inter-packet times were calculated as the difference between the time stamps of all packets regardless of the connection as shown in Fig.3.1. The figure shows an example of two connections  $A_i$  and  $C_i$ . Connection  $A_i$  represents packets going from Source port A to Destination port B and vice versa. Connection  $C_i$  represents packets from Source port C to Destination Port D and vice versa. In the conventional approach, the Inter-packet times  $IPT(A_i, C_i)$  were calculated collectively for every packet received as indicated by the single IPT stream at the bottom of the figure. It was observed in the results generated by this approach, the statistics were such that the original data appeared to be Poisson like and the higher order statistical plots revealed no structure other than a single peak at zero time.

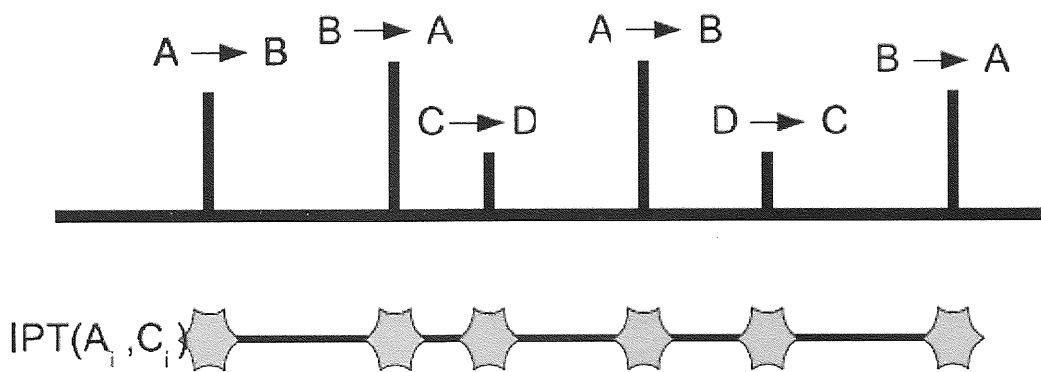
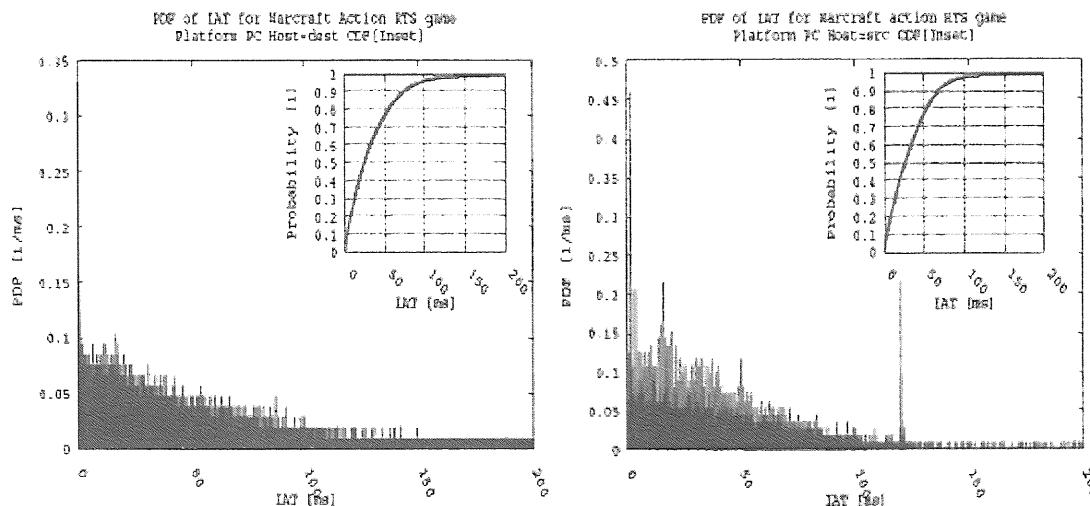


Figure 3.1: Traditional approach to Internet traffic measurement

Examples of plots (PDF) of IPT statistics for this approach are shown below in Fig. 3.2 and 3.3 respectively. This were measured for a largely Poisson in nature, for a wide majority of the packet traffic measurements of different traffic content , as was also observed in studies such as [51].



3.1 a - Two Sample PDFs (Online Gaming) showing Poisson like statistics for measurements using approach in Fig. 3.1

Fig. 3.1a shows the Poisson like nature of the PDF IPT statistics which was measured based on the approach suggested in Fig. 3.1 for traffic based on online gaming applications.. Fig. 3.1b, shows similar Poisson like nature of IPT statistics for Online video streaming traffic (Joost), again based on the approach suggested in Fig. 3.1. As clearly seen from both the sets of measured plots, they all have a peak close to zero, and have identical nature of statistics.

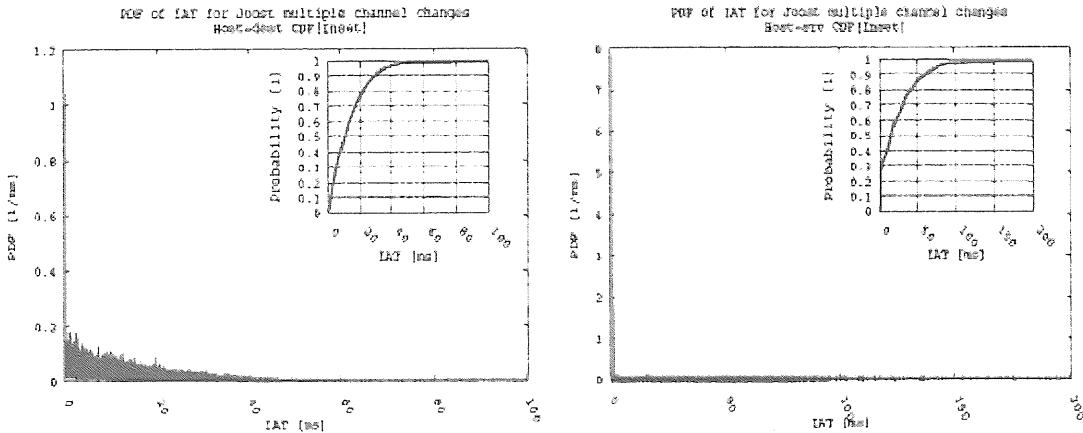


Figure 3.1b - Two Sample PDFs (Joost video) showing Poisson like statistics for measurements using approach in Fig. 3.1

Consequently, a revised approach was considered, as shown in Fig. 3.2, where the Inter-packet times were measured based on individual connection flows. The Inter-packet times, for this approach, are unique for connections  $A \leftrightarrow B$  and  $C \leftrightarrow D$ , as marked by  $\text{IPT}(A_i)$  and  $\text{IPT}(C_i)$  respectively in Fig. 3.2. This approach resulted in unique statistics for traffic based on individual connections for different types of applications, providing a good opportunity to study the packet traffic source behaviour that is unique to the type of application and protocol being analysed. A few example sets of results generated using this revised approach are covered and analysed in details in the following sections.

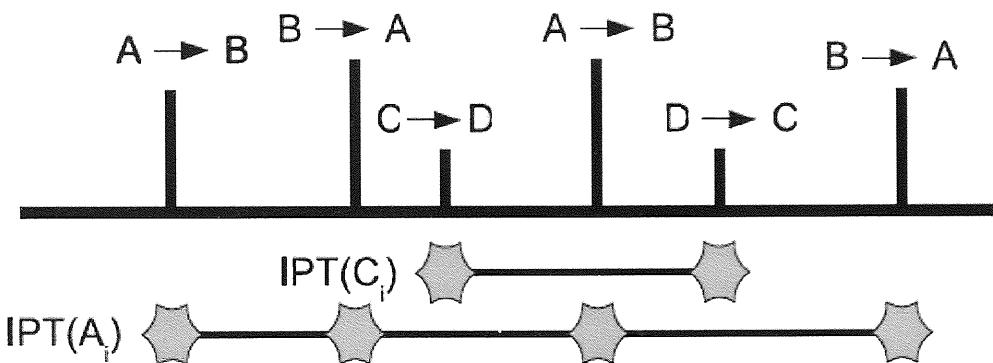


Figure 3.2: Revised Approach to IP traffic measurement

### 3.1.2 Traffic Measurement Tools Used

#### 3.1.2.1 Tcpdump

Packet Analyzers are software-based measurement tools that typically modify the kernel of a host computer's network interface to give it packet-capture capability. It places the network interface device in a *promiscuous mode* such that it captures all passing packet traffic, in addition to packets that were meant only for that computer. One such widely used tool is *tcpdump* [52], which uses the Berkeley Packet Filter architecture to capture TCP/IP packets [53]. Tcpdump allows one to capture a network's IP packets and filter the captured traffic streams based on specific IP addresses, port numbers and protocol types [54]. It has been widely used to study Internet applications and growth trends in Internet traffic over time, as well as in traffic measuring [55] [56] [57].

#### 3.1.2.2 Wireshark

Wireshark formerly known as Ethereal is another variant of a Packet Analyzer tool, which is very similar to Tcpdump, but has a graphical user-interface with advanced functionality [58]. Wireshark has also been recently widely used for monitoring and measuring [59] [60] IP traffic.

#### 3.1.2.3 Libpcap

Libpcap refers to a Unix environment packet capture library, that was developed by the people who developed tcpdump, as an application programming interface (API), independent of the system, that can capture passing packets, and provide the capability of collecting information, filtered to specific detail, and written to a file of a specific format that can be then taken as input for statistical analysis of the network [61]. The Libpcap library was used extensively in our work as part of scripts and tailor made tools to perform packet traffic measurements suited to our requirements.

#### 3.1.2.4 Scripts for Statistical Analysis and Computation

Apart from the aforementioned tools, a number of programming scripts written in C and C++, were used for filtering numerical data according to port numbers and IP addresses. This data was then arranged in a convenient manner to calculate the Inter-packet times (IPT). The IPT data was then arranged in to two columns 'N' and 'N+1', as was done with the simulated data of a Markov Monte Carlo Simulation explained in Chapter 2. This IPT data would be then used for producing the

statistical plots for analysis and modelling, namely the Probability Density Function (PDF) and the Joint Density Function (JDF).

## 3.2 Measured Results for IPT of IP Traffic

In this section we will be looking at a few selected examples of measured results, which were measured based on the alternative approach mentioned in section 3.1.1.

### 3.2.1 Example 1: Measurement at Port 80 (TCP) - IP Radio Broadcast

The first example for IP traffic measurement is of IPT statistics for IP traffic associated with packet traffic captured on port 80 having connections based on the TCP protocol. Port 80 which supports both TCP and UDP transport layer protocols, is one of the most popularly used internet communication ports. Often protocols with a pre-defined way of communicating use Port 80. The most common communication protocol known to use port 80 is the Hypertext Transfer Protocol (HTTP) which is used for web communication. Hence it dominantly handles traffic associated with the World Wide Web (WWW). On a web server, port 80 listens for incoming connection requests from web clients. Apart from web traffic, a lot of IP Radio broadcasting services have also been known to use the Port 80 [62]. One of the most prominent broadcasting services which uses Port 80 is the BBC Iplayer's radio streaming service [63]. The measurements for the IPT statistics in this example are based on a 30 minute capture of IP traffic stream associated with an IP Radio program, broadcast on the BBC IPlayer. Table 3.1 shows details of the packet capture.

Port Number	Transport Protocol	Application	Number of IP packets
80	TCP	BBC IPlayer - Radio	32137

Table 3.1: Measurement Scenario 1 – TCP Port 80 – IP Radio Broadcast

## Probability Density Function

Fig. 3.3 shows the Probability Density Function (PDF) for IPT for IP packets captured on port 80/TCP. From the graph, we notice two significant peaks and two small peaks. The significant peaks show that most of the packet traffic sent was sent at every 0.1 ms and 0.3 ms with other peaks indicating that occasionally packets were sent at every 9 ms and 100 ms. We also observe that the results show non-Poisson statistics with several peaks.

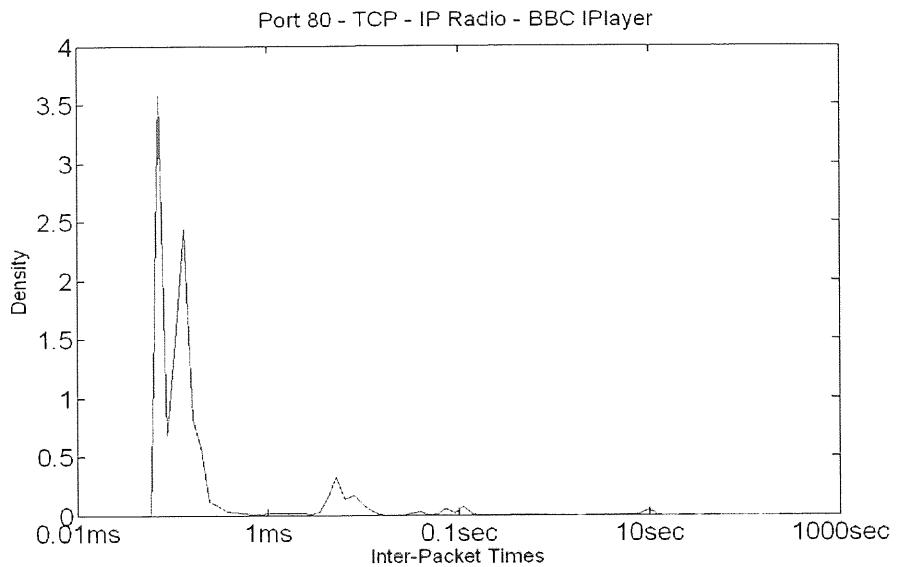


Figure 3.3: Probability Density Function of IPT for IP Packets on TCP Port 80 – IP Radio Broadcast

## Joint Density Function

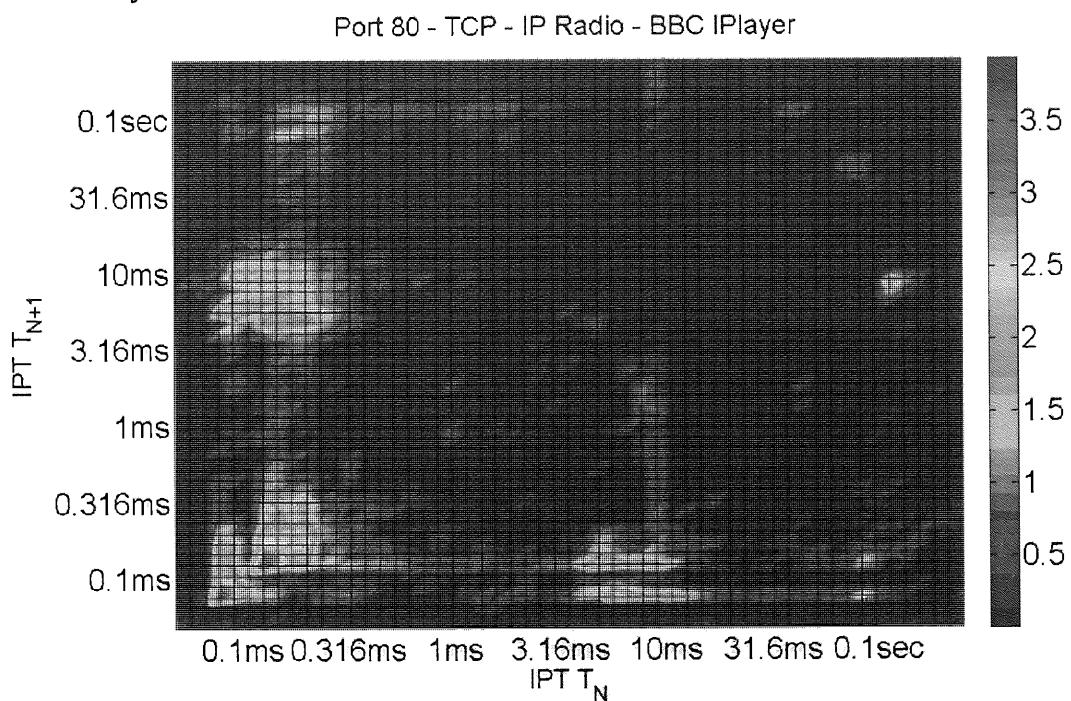


Figure 3.4: Joint Density Function of IPT for IP packets on TCP Port 80 – IP Radio Broadcast

The Joint Density results shown in Fig. 3.4 show that for the vast majority of times a packet arriving at 0.1 ms is followed by another packet after 0.1 ms and similarly a packet of 0.3 ms is followed by another packet after 0.3 ms. The cloud at 10ms on the Y-axis represents a group of packets, with inter-packet times distributed around a mean of 10 ms and were followed by packets with a period of 0.1 – 0.3 ms. Thus, the Joint Density results reveal a lot more interesting features on the timing structure of packet traffic than observed with the PDF. The significance of the Joint Density Plots is discussed in elaborate detail in the next chapter.

### 3.2.2 Example 2: Measurement at Port 80 (UDP) - IP TV Broadcast

In the second measurement example, we will again look at measured results from the Port 80, but this time for a UDP connection carrying IP TV video stream traffic of the BBC Iplayer. The video packet traffic streaming of the BBC IPlayer uses the UDP transport protocol for the transport of video data packets. The IPT statistics recorded are for video IP traffic for thirty minutes of run time. The capture details are shown in Table 3.2.

Port	Transport Protocol	Application	Number of IP packets
80	UDP	BBC IPlayer - IP TV	81897

Table 3.2: Measurement Scenario 2 – UDP Port 80 – IP TV Video Broadcast

#### Probability Density Function

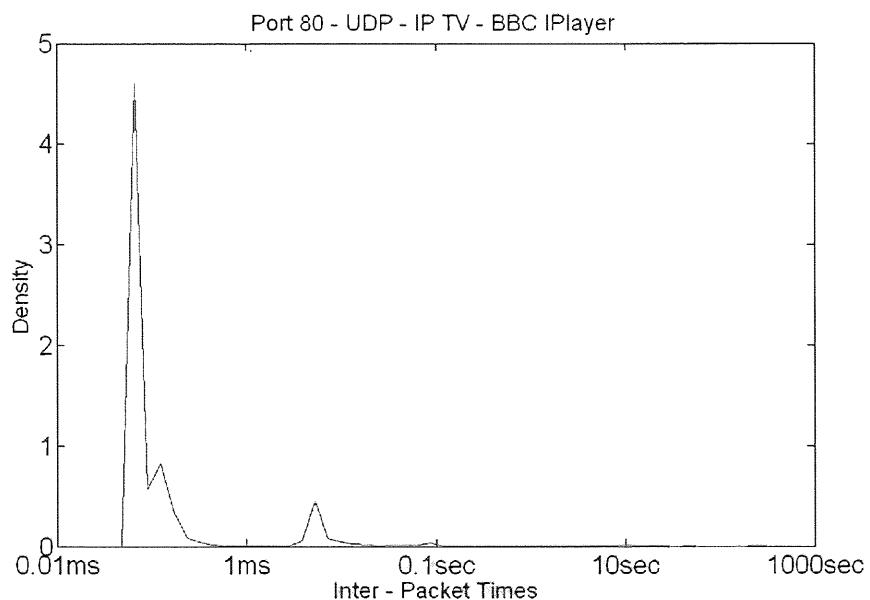


Figure 3.5: Probability Density Function of IPT for IP Packet Traffic on UDP Port 80 – IP Video Broadcast

The probability density function show in Fig. 3.5 again shows that packets most packet traffic associated with the IP TV stream from the BBC Iplayer, occurred at intervals of 0.1 ms with other peaks at 0.3 ms and 10 ms.

### Joint Density Function

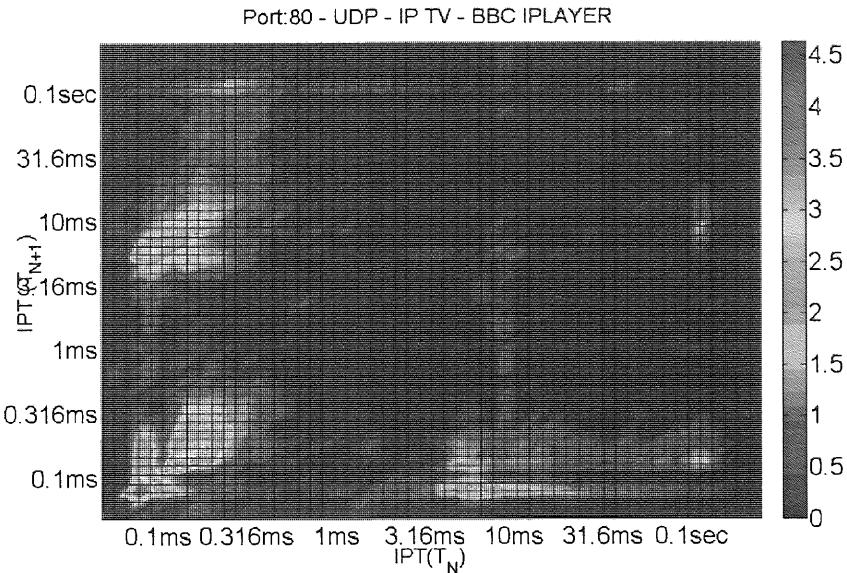


Figure 3.6: Joint Density Function for IPT for IP packets on UDP Port 80 – IP TV Broadcast

Fig. 3.6 shows the Joint Density Results for IPT of Video data IP traffic associated with the BBC IPlayer. When compared to the Joint Density Plot of the IPT for IP Radio of the BBC Iplayer (Fig. 3.4), we see that they both are very similar, thus signifying that they both have similar packet source behaviour, where again for the majority of the time, a packet arriving at 0.1 ms is followed by another packet arriving at 0.1 ms.

### 3.2.3 Example 3: Measurement at Port 15010 (UDP) -Telephony Traffic

The third example of measured result highlights the IPT statistics for traffic associated with a VoIP client. Measurements were recorded at Port 15010 and were based on the UDP transport protocol. No specific data is available for this port, and it is a randomly picked port for communication used for establishing a two-way real time communication session based on using an appropriate signalling protocol such as the Session Initiation Protocol (SIP) which streams the IP packets associated with the voice and video chat using a streaming protocol such as the *RTP Control Protocol* (RTCP) and Real Time Streaming Protocol (RTSP). Details of the packet capture at Port 15010 are shown in Table 3.3.

Port	Transport Protocol	Application	Number of IP packets
15010	UDP	IP Telephony	169317

Table 3.3: Measurement Scenario 3 – UDP Port 15010 – IP Telephony traffic

### Probability Density Function

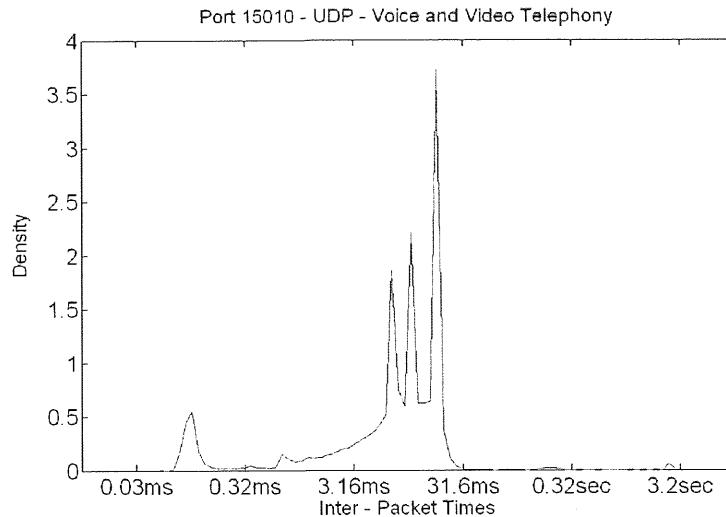


Figure 3.7: Probability Density Function of IPT for IP Packets on UDP Port 15010 – IP Telephony

The PDF shown in Fig. 3.4 shows that the traffic related to IP telephony have different IPT statistics with three significant peaks indicating that most of the packet traffic is focussed at intervals of 30ms, 17 ms and 10 ms, with a minor amount of packet traffic occurring at intervals of around 0.17 ms. The three significant peaks show that there could be three different streams of traffic associated with this IP telephony sample.

### Joint Density Function

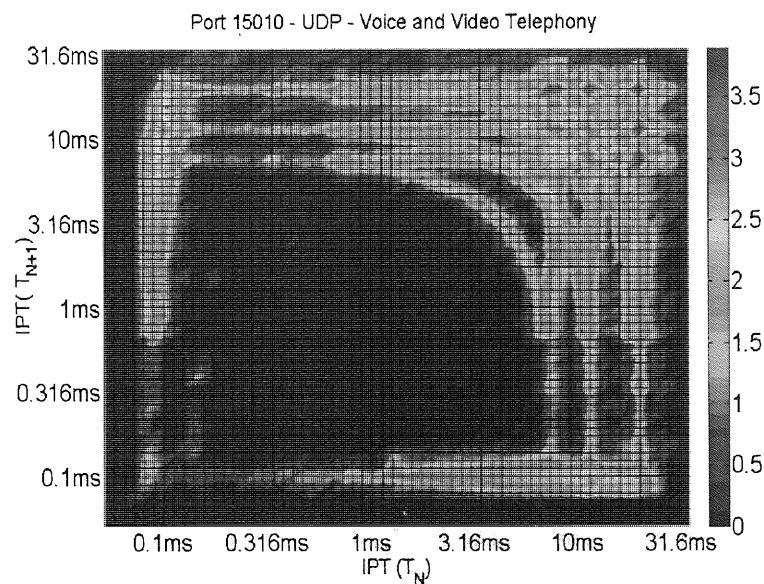


Figure 3.8: Joint Density Function of IPT for IP Packets on UDP Port 15010 – IP Telephony

The Joint Density result, shown in Fig. 3.8, reveals some unique features, which we have particularly seen with telephony applications. One of these unique features which is elaborately covered in Chapter 5, is the ‘curve of periodicity’, which gives hints on the presence of elements of periodicity in the packet traffic. These are characteristics of packet streams that are relayed by streaming protocols such as RTSP or RTCP. It is clear again from the Joint Density Plot, that there are three different packet streams as observed in the top right corner, and the three curves of periodicity associated with them.

Work on modelling packet source behaviour observed through the Joint Density plots is covered in detail in Chapters 5 and 6.

### 3.3 Estimating Gaussian Markov Model Parameters

In the previous chapter, we had introduced the Gaussian Markov Models. In this chapter we will look at how we can use the Gaussian Markov Models to model the first order statistics of the measured results.

#### 3.3.1 Understanding the Gaussian distribution in terms of modelling the peaks – Estimating the mean parameters

In Chapter 2, we had seen that the general PDF equation for the IPT for an n-state Gaussian Markov Model is given by the following equation (3.1):

$$PDF_{(n-state, \text{ Gaussian})} = V_1 G_1(\mu_1, \sigma_1, t) + V_2 G_2(\mu_2, \sigma_2, t) + \dots V_n G_n(\mu_n, \sigma_n, t) \quad (3.1)$$

Where each Gaussian component is given by the equation (3.2):

$$G_i(\mu_i, \sigma_i, t) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{\frac{-(t-\mu_i)^2}{2\sigma_i^2}}, i = 1 \text{ to } n \quad (3.2)$$

The first step towards modelling a measured result would be to judge the number of significant peaks that the model can capture, and associate a Gaussian component with each of these peaks. From the analysis in the previous chapter we know that in case of the Gaussian Markov Model, the location of the Gaussian peak maximum is given by the mean  $\mu$  and the width of the Gaussian distribution is determined by the standard deviation parameter,  $\sigma$ . So in order to propose a Gaussian Markov model, we need to estimate these parameters for each Gaussian distribution and fit a model

to the measured PDF results, the analytical solution of which is given by a convex combination of these Gaussian distributions, where each Gaussian distribution represents a significant peak, such that the model when simulated, exhibits similar PDF results to the measured result.

Often, when we are estimating the number of Gaussian components of the model based on the peaks that we can visualise and differentiate, we might come across peaks that have a part of it buried within other peaks. One strong clue to recognise such a structure is that the taller peak in that structure does not follow a perfect Gaussian bell curve and has a shoulder extending out of the curve, indicating the presence of another peak within. An example of such a scenario is given in Fig. 3.9 (a). Through this structure, it is easy to visually estimate the mean  $\mu_1$  of the taller peak, which is around 9.5. The location for the maximum of the second peak requires careful judgement based on the amount of the peak protruding out of the other peak. In this example, the mean of the second peak appears to be at around 4.8. Once the means are estimated, the amplitudes and state visiting probabilities can be estimated through mathematical evaluation and an estimation algorithm such as the Expectation–Maximization algorithm respectively. This is covered in detail in the sections 3.3.2 and 3.3.3. Finally, the standard deviations are estimated and that completes the estimation of all parameters of the two Gaussian distributions, and thus a mix of two Gaussian distributions can be fitted to the structure in Fig. 3.9 (a) as shown in Fig. 3.9 (b).

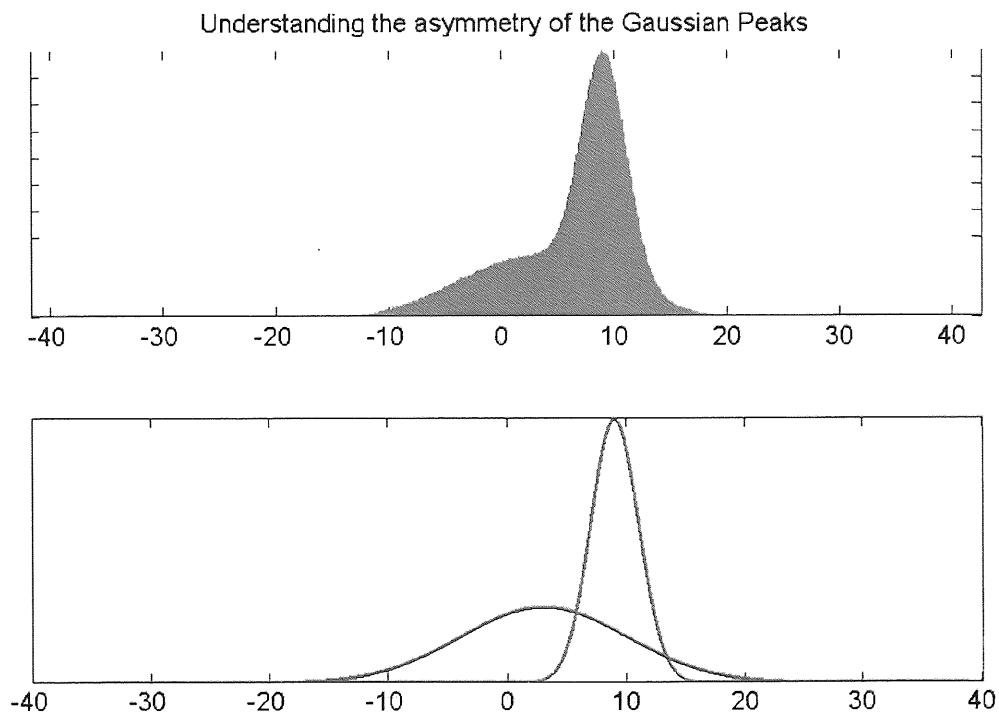


Figure 3.9: a) A Gaussian peak buried within another Gaussian peak b) Two fitted Gaussian components

### 3.3.2 Estimating the state visiting probabilities of the Gaussian Markov Models

The CDF (Cumulative Distribution Function) of the IPT is very instrumental in estimating the state visiting probabilities. This CDF when plotted together with the PDF as shown in Fig. 3.10, for one of the measured results, is useful to provide rough estimates of the Gaussian parameters associated with each Gaussian peak. The CDF lets one know the probability ratio of mixture of each Gaussian component. In other words, it provides an estimate of what percentage of each peak contributes to the entire PDF. The points of change in slope provide the probability value associated with each peak defining the contribution of each Gaussian component, ultimately providing with the probability value for visiting the states in the Monte Carlo simulation of the Markov model. Once this is estimated, it paves way for the estimation of the standard deviations and amplitude of the Gaussian component to be fitted.

### 3.3.3 Estimating the standard deviation using the EM algorithm

Once the means and state visiting probabilities for each of the Gaussian component are fitted, the next parameter to be calculated in the process, is the standard deviation of each of the Gaussian distributions. Considering the general PDF as shown in equation (3.1), after the estimation of the number of Gaussian components with their estimated means and state visiting probabilities as input, the standard deviations in the PDF equation can be estimated using the EM algorithm.

The EM (Expectation Maximization) algorithm was developed in 1977 [64] as a procedure to estimate mathematical model parameters given that there is limited observed data or incomplete data to complete the model. The Expectation Maximization algorithm uses a Maximum-likelihood estimation process where the missing parameters are estimated such that the model with those estimated parameters produces the PDF result for IPT of traffic with the highest likelihood. The algorithm works in two main steps: The Expectation step and the Maximization step. In the Expectation step, the missing parameter is estimated based on the observed results and the initial estimates of other parameters (the PDF equation with the rates and state visiting probabilities). In the Maximization step, the algorithm optimizes the bound values for the estimated parameter thus providing improved estimation in the following E-step. Eventually after several iterations of the algorithm, the results of the model with estimated parameters are sure to converge with the observed results [65]. For more mathematical details on the EM algorithm, please refer to [66].

### 3.3.4 Using the amplitudes of the Gaussian components of the Gaussian Markov Models to improve estimation

We know that the Gaussian distribution is given by equation (3.2). Once the standard deviation of a peak is estimated using the EM algorithm, the amplitude of the peak can be used to match the peak closely with the peak in the measured result.

The amplitude of the peak for a Gaussian distribution is given by

$$A = V \times 1/\sigma\sqrt{2\pi} \quad (3.3)$$

Since, the amplitude of the Gaussian peak, is known from the measured result, the amplitudes of the peak are another set of observable inputs which can be used to improve the estimation of  $V_i$ , the visiting probability for the states associated with that Gaussian distribution. Using the estimated values of the means ( $\mu_i$ ) as reference for locations, the amplitude values  $A$  along with sigma  $\sigma$ , are paired for each peak that can be visually segregated from the entire PDF (based on the values of  $\mu$ ) to compute the new estimated values of  $V_i$  for each corresponding peak using the equation 3.3 to further improve estimation of PDF in the modelling process.

In the next section, we will put the estimation procedure discussed in this section to practical use in producing Gamma Markov Models to model the three examples of the measured results discussed in section 3.2.

## 3.4 Producing Gaussian Markov Traffic Models for Measured Results

### 3.4.1 Modelling Measured Results – Port 80 (TCP) IP Radio

We will start with the PDF results for IPT obtained from the TCP Port 80- IP Radio measurements. Applying the procedure for modelling, which we discussed in the previous section, we will first start with analysing the PDF result for estimating the number of Gaussian components for the Gaussian Markov Traffic Modelling.

#### 3.4.1.1 Analysing the Probability Density Function for Gaussian Components

We will first consider the probability density plot for the results, shown in Fig. 3.10.

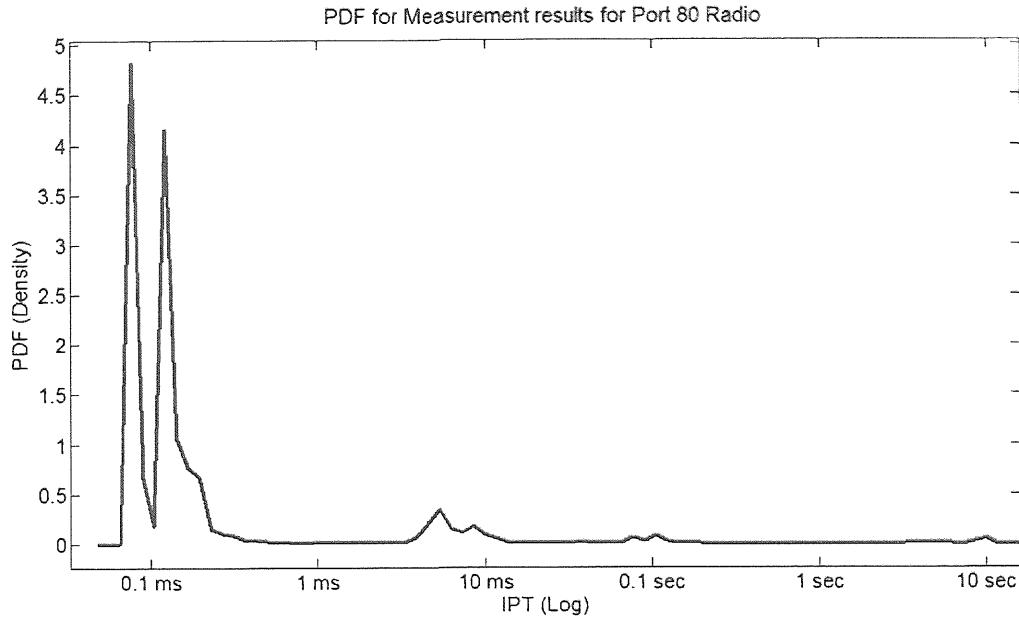


Figure 3.10: PDF for IPT for Measured Results for Port 80 TCP for IP Radio Traffic.

To start with the modelling process, it is useful to note that we intend to model only the peaks that contribute the most to the total IPT data. In other words we want to model only the peaks that are significant. In Fig. 3.10, we see five significant peaks (from left to right), the location of means for which are given as follows:

Peak 1,  $\mu_1 = 0.08$  ms

Peak 2,  $\mu_2 = 0.126$  ms

Peak3,  $\mu_3 = 0.178$  ms (Peak 3 is buried along with Peak2)

Peak 4,  $\mu_4 = 6.19$  ms

Peak 5,  $\mu_5 = 8.57$  ms

Thus accordingly a 5-state Gaussian Markov Model can be proposed.

### 3.4.1.2 Estimating the state visiting probabilities for the Gaussian Markov Model:

In order to make the initial assumptions on the values of the state visiting probabilities associated with the states of the model, we plot the CDF along with the PDF of IPT of the measured results as shown in Fig.3.11.

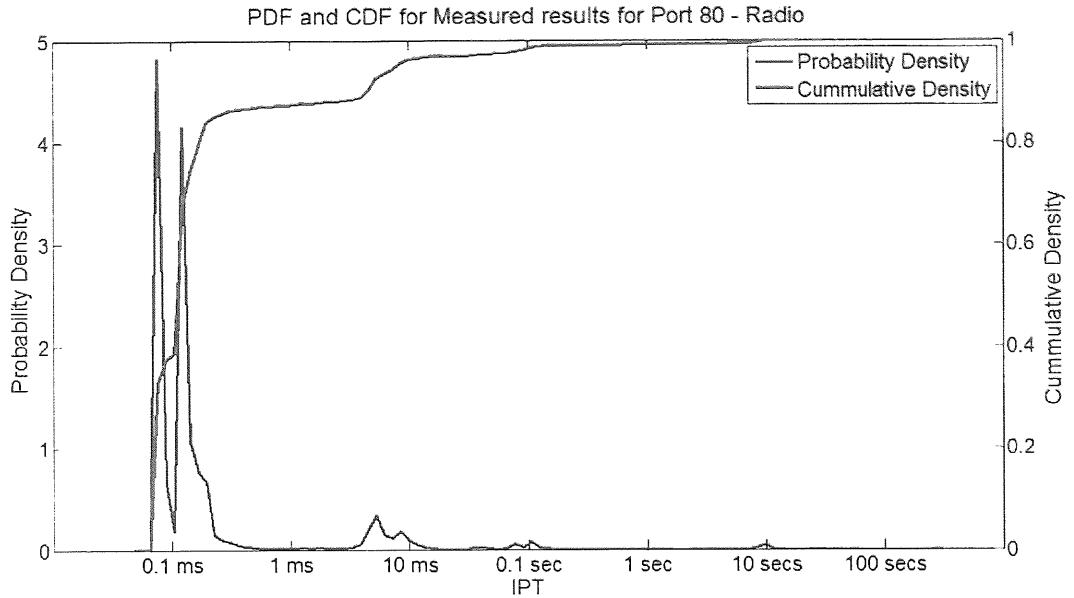


Figure 3.11: PDF and CDF for IPT for Measured results on TCP Port 80- IP Radio

The changes in slope of the CDF curve associated with each peak of the PDF shown in Fig. 3.11, establishes the initial assumptions for values of the state visiting probabilities.

Initial assumptions for state visiting probabilities using CDF for IPT:

$V_1 = 0.39$ , because the first change in slope on first peak is at 0.39 of the cumulative probability axis.

The cumulative probability corresponding to the second peak is approximately equal to the difference in slope up till 0.7 from 0.39, and thus the contribution of IPT events for that peak is  $0.7 - 0.39 = 0.31$ . Therefore, we have  $V_2 = 0.31$ .

The contribution of IPT events for the third peak is given by  $0.8 - 0.7 = 0.1$ . Therefore  $V_3 = 0.1$

Similarly we have,

$$V_4 = 0.04$$

$$V_5 = 0.03$$

### 3.4.1.3 Estimating the standard deviations for the Gaussian Markov Model using the EM algorithm

Using the equation (3.1) as input where  $n = 5$ , and substituting the mean parameters and the state visiting probabilities, the EM algorithm is executed for a number of iterations to estimate the values of the standard deviations for all the Gaussian components which are obtained as the following:

$$\sigma_1 = 0.0232, \sigma_2 = 0.0311, \sigma_3 = 0.0617, \sigma_4 = 0.0571 \text{ and } \sigma_5 = 0.07701$$

Please note that these estimated values of standard deviation parameters are in the logarithmic domain (for ease of computation) and do not scale with the time axis shown in Fig. 3.10 and 3.11.

### **3.4.1.4 Using the peak amplitudes for improved estimates of the state visiting probabilities and standard deviation:**

Since we now have rough estimates of the standard deviation and the state visiting probabilities, we can now use the amplitudes of the peaks from the measured result, as previously described in section 3.3.4, to improve the estimation process and achieve a Gaussian Markov model that fits very closely as follows.

The amplitudes of the five significant peaks obtained from the measured results are:

$$A_1 = 4.8, A_2 = 4.38, A_3 = 0.9, A_4 = 0.34 \text{ and } A_5 = 0.17$$

Now based on equation (3.3), we have

For State 1,

$$A_1 = 4.8 = V_1 \times 1 / 0.0232 \times 2.506$$

Therefore,

$$V_1 = 0.385$$

Similarly for State 2,

$$A_2 = 4.38 = V_2 \times 1 / 0.0311 \times 2.506$$

Therefore,

$$V_2 = 0.3424$$

Similarly, we have for States 3, 4 and 5 of the model:

$$V_3 = 0.1298$$

$$V_4 = 0.0489$$

$$V_5 = 0.0325$$

The Equation (3.3) can also be used for re-estimating standard deviation values when required.

### 3.4.1.5 Constructing the Gaussian Markov Model – Analytical Equation

We know, that for the Gaussian Markov Model,

$V_1 + V_2 + V_3 + V_4 + V_5$  must be equal to 1. However most of the time, the estimation process results in  $V_1 + V_2 + V_3 \dots + V_n = 0.9 \sim 0.99$ , as we cannot possibly produce a fit which is 100% accurate, as we model only the significant peaks. Therefore the sum of the estimated state visiting probabilities need not necessarily be 1.

Therefore in order to appropriately define the transition Probability Matrix  $P_{ij}$ , we need to normalise these values, so that they add up to unity. So, we calculate a value  $N_v$  as shown (for this example):

$$N_v = 1 - (V_1 + V_2 + V_3 + V_4 + V_5) = 0.0627$$

We then use the value  $N_v$  to normalise the other values of  $V_1$  to  $V_5$ .

So the new and appropriate values for the state visiting probabilities are:

$$V_1 = 0.4091, V_2 = 0.3698, V_3 = 0.1385, V_4 = 0.0522, V_5 = 0.0034$$

And this now gives,

$$V_1 + V_2 + V_3 + V_4 + V_5 = 1$$

At this stage, we have all the parameter values needed to determine the analytical solution for the PDF of the IPT for the 5-state Gaussian Markov Model. Substituting the estimated means, standard deviations and the state visiting probabilities of the model in equation (3.1) gives:

$$\begin{aligned}
 PDF_{(5\text{-state}, \text{ Gaussian (Port 80-IP Radio)}} \\
 &= 0.4091 \times G_1(0.08, 0.02) + 0.3698 \times G_2(0.126, 0.03) + 0.1385 \\
 &\quad \times G_3(0.178, 0.0617) + 0.0522 \times G_4(6.19, 0.05) + 0.0034 \\
 &\quad \times G_5(8.57, 0.077)
 \end{aligned} \tag{3.4}$$

Equation (3.4) is the analytical solution for the PDF of IPT based on the proposed 5-state Gaussian Markov Traffic model for modelling observed from the measured PDF result of the IP packets associated with and IP Radio broadcast packet traffic on TCP Port 80.

### 3.4.1.6 Constructing the Gaussian Markov Model – Monte Carlo Setup

In the previous chapter, we had considered loop model examples for the Gaussian Markov model, where the state visiting probabilities were all equal since the states are visited equally often. But in the current scenario, we have to model states, which have non-linear state visiting probabilities. In other words,  $V_1 \neq V_2 \neq V_3 \neq \dots \neq V_n$ . In order to produce the state visits in the Monte Carlo Simulation in such a way, such that the right state visiting probabilities are generated, the state transitions used for the model must be chosen such that every state must transit to every other state with a specific unique probability. But clearly, this is a mathematically complicated procedure for a large number of states. Apart from this, from a mathematical perspective, as explained in Chapter 2, the states of a continuous time Markov model cannot have a self-transition. This is also true in the network perspective as well, as there can be no such thing as a source sending a packet to itself as per the traffic model. Hence an alternative approach is required.

An easier alternative approach is to create transitions that match the visiting probabilities by using an additional ‘feeder’ state. This *feeder* state creates transition such that each state is visited (from the feeder state) with their respective visiting probabilities and once the model spends the Gaussian distributed amount of time in the state it is visiting, the state returns back to the feeder state with a transition probability 1. Fig.3.12 shows a state transition diagram for the proposed 5-state Gaussian model with an additional sixth ‘feeder’ state.

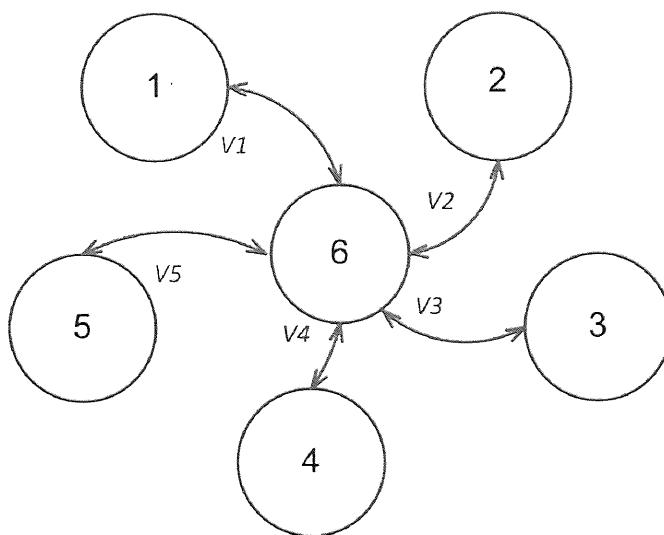


Figure 3.12: A 5 + 1 State Gaussian Markov Model to model measured results for IPT of IP Radio packets

The sixth feeder state generates transitions to the other states with transition probabilities same as the state visiting probabilities. The bi-directional transition arrows indicate that the model visits states 1 to 5 with their state visiting probabilities and returns to state 6 with a transition probability 1. It is important to note that the statistics of the IPT is not affected by the presence of the feeder state, as the time the model spends in the 6<sup>th</sup> state is not included in the statistics of the IPT, but rather its presence is only for generating the state visits according to their visiting probabilities, thus producing statistics of a 5-state Gaussian Markov Model. Mathematically, one can imagine that the time spent in the sixth feeder state is so small that it is negligible to the statistics.

### 3.4.1.7 Results of the Monte Carlo Simulation of the Proposed 5-state Gaussian Markov Model:

Fig. 3.13 shows the results of PDF for IPT from the Monte Carlo simulation of the proposed 5-state (5+1) Gaussian plotted alongside with the analytical solution represented by the equation (3.4) and the original measured result for the PDF of the IPT of packet traffic associated with an IP radio broadcast.

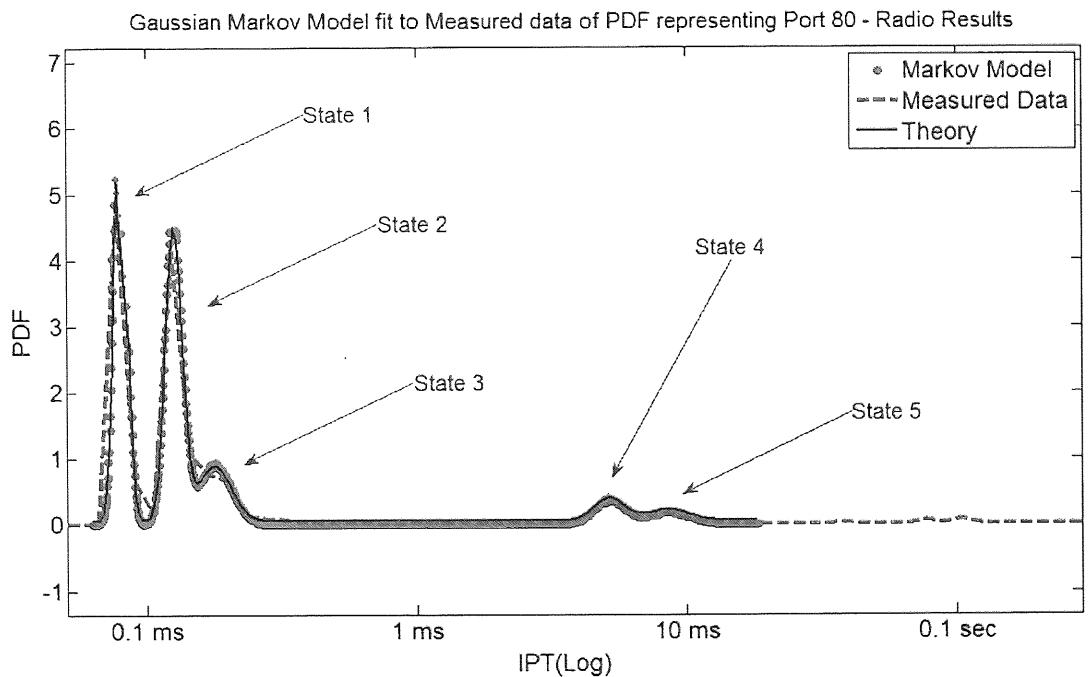


Figure 3.13: PDF Plots of modelled, analytical and measured results of IPT for IP Radio traffic

As shown in Fig. 3.13, the modelled results are in good agreement with the measured results as well as the analytical result for the model. The five states associated with the five Gaussian peaks have been clearly highlighted in the figure.

### 3.4.2 Modelling Measured Results – Port 80 (UDP) IP TV/Video

Next, we move on to the second set of measured results, for IP TV/Video. Since, we have already discussed the mathematical detail of the steps involved in creating a proposed n-state Gaussian Markov model for measured IPT statistics, with the previous set of measured results, we will skip the mathematical detail of obtaining the model as it is exactly the same procedure as with the earlier modelled result. Instead we will focus on the proposed model for this example and its results.

#### 3.4.2.1 Modelled Parameters:

Fig. 3.13 shows the probability density function (PDF) and the cumulative density function (CDF) of the IPT plotted alongside each other for estimating the mean, standard deviation and the state visiting probability parameters for the proposed Gaussian Markov Traffic model.

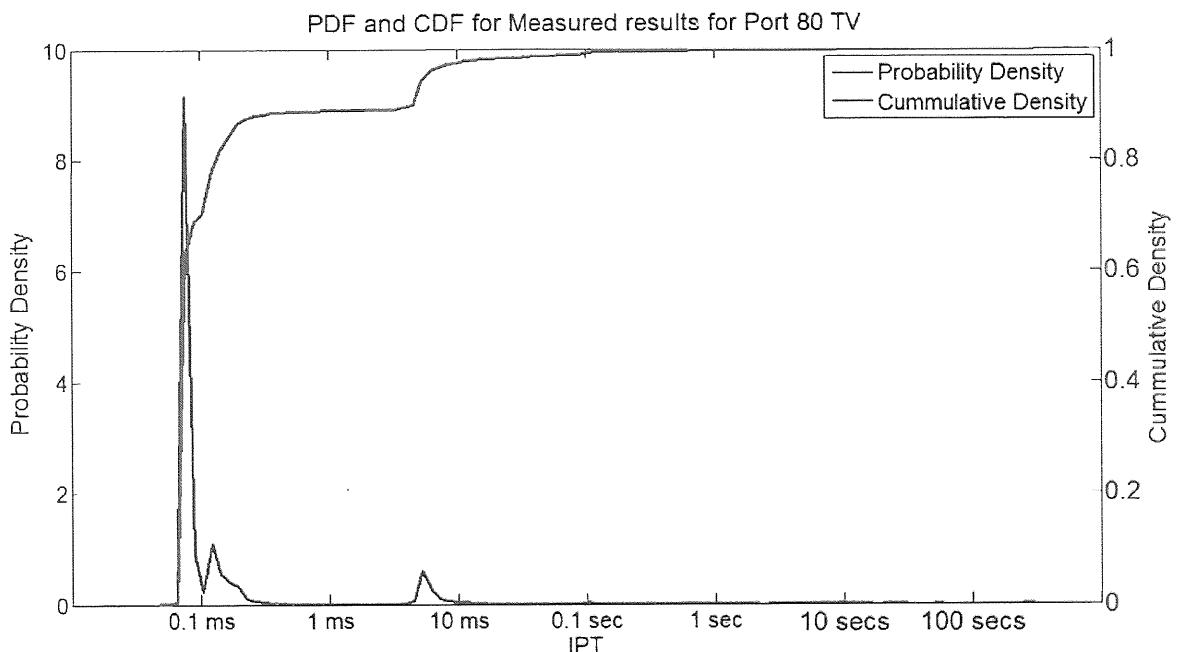


Figure 3.14: PDF and CDF for IPT for Measured results on UDP Port 80- IP Tv Packet Traffic

As can be seen from Fig. 3.14 and 3.15, there are five significant peaks and thus a 5-state Gaussian Markov model very similar to the one introduced previously is proposed again for the IP TV statistics. Upon using the same mathematical procedure as for the previous case for proposing and constructing the model, a 5-state Gaussian Markov Model, for which the analytical solution for the PDF for IPT is defined in equation (3.5), was fitted to the measured results.

$$\begin{aligned}
 PDF_{(5\text{-state}, \text{ Gaussian (Port 80-IP Tv)}} &= 0.4832 \times G_1(0.079, 0.02) + 0.105 \times G_2(0.123, 0.036) + 0.214 \\
 &\quad \times G_3(0.1603, 0.11) + 0.0345 \times G_4(5.5, 0.033) + 0.0048 \\
 &\quad \times G_5(7.063, 0.039)
 \end{aligned} \tag{3.5}$$

As before, for reasons stated earlier, the means in the equation have been represented in ms whereas the standard deviations are not in unit of times.

### 3.4.2.2 Monte Carlo Simulation Results:

Following the same Monte Carlo Simulation procedure as with the previously proposed model, we consider a 5 state Markov Model with an additional sixth ‘feeder’ state that does not contribute to IPT statistics as shown in Fig. 3.12. The Monte-Carlo simulated results of the proposed 5-state Gaussian Markov Model are shown in Fig. 3.15 along with the analytical result represented by the equation (3.5) and the original measured result for the PDF of the IPT of packet traffic associated with an IP TV broadcast based on the BBC Iplayer application. The states representing each peak modelled have also been clearly marked for easier understanding.

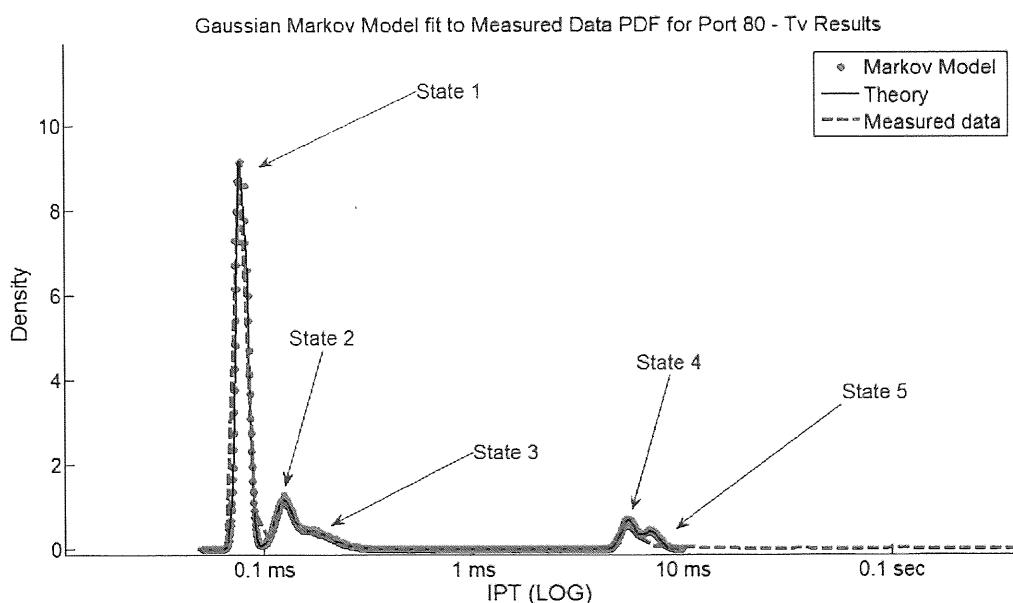


Figure 3.15: PDF Plots of modelled, analytical and measured results of IPT for IP TV traffic  
We notice that the results of the proposed 5-state Gaussian Markov model closely agree with the analytical and measured PDF results for the IPT of IP TV packet traffic measured at UDP Port 80.

### 3.4.3 Modelling Measured Results – Port 15010 (UDP) IP Voice/Video Telephony

Finally we move on to the last set of measured results, where we model the IPT results based on traffic captured at the UDP Port 15010 for an IP telephony application running both voice and video services.

#### 3.4.3.1 Modelled Parameters:

Fig. 3.16 shows the PDF and CDF results for the measurement data, based on which the mean, standard deviation and the state visiting probability parameters for the proposed Gaussian Markov Traffic model were estimated.

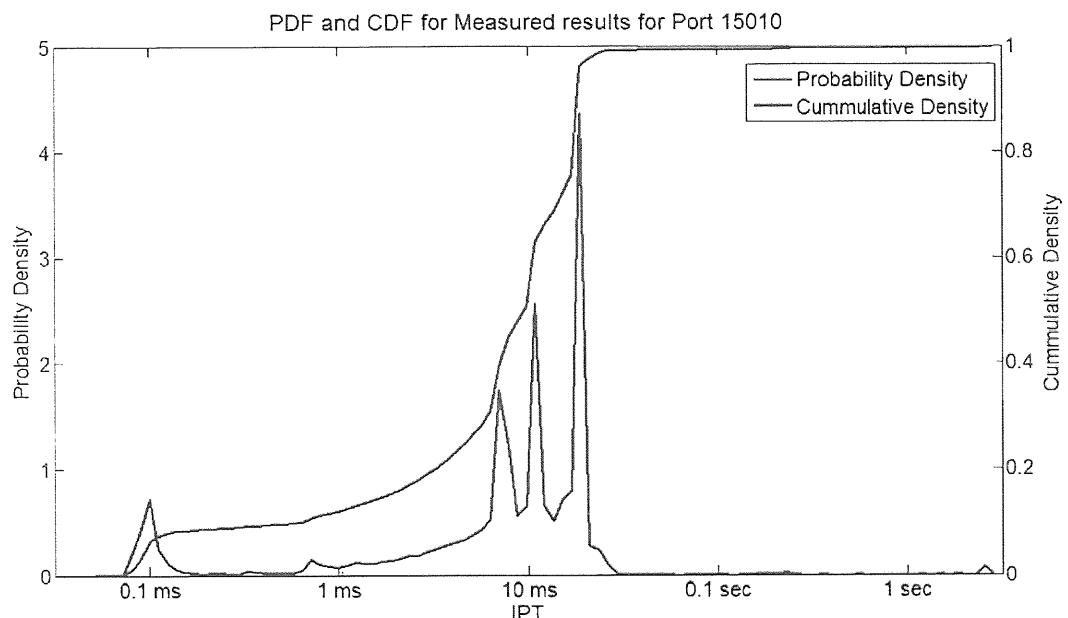


Figure 3.16: PDF and CDF for IPT for Measured results on UDP Port 80- IP Voice/Video Telephony  
This time we have six significant peaks that can be modelled, as shown in Fig. 3.16 and 3.17 and this leads to a proposed 6-state (6+1) Gaussian Markov Model to model the results, the PDF equation of which is given below:

$$\begin{aligned}
 & PDF_{(5\text{-state}, \text{ Gaussian (Port 15010 -IP Telephony)}} \\
 & = 0.08 \times G_1(0.97, 0.04) + 0.39 \times G_2(6.5, 0.386) + 0.1203 \times G_3(7.2, 0.03) \\
 & + 0.1408 \times G_4(10, 0.02) + 0.2301 \times G_5(18.5, 0.0199) + 0.0349 \\
 & \times G_6(14.5, 0.3867)
 \end{aligned} \tag{3.6}$$

### 3.4.3.2 Monte Carlo Simulation Results

Again a similar setup is adopted as before, for the Monte Carlo simulation, except that a 6 + 1 state model is now used in the simulation, where the 7<sup>th</sup> state connects to all other states and does not account for the IPT statistics. The Monte-Carlo simulated results of the proposed 6-state Gaussian Markov Model are plotted in Fig. 3.17 alongside with the analytical solution represented by the equation (3.6) and the original measured result for the PDF of the IPT of packet traffic associated with an IP Telephony application. The states representing each peak modelled have also been indicated.

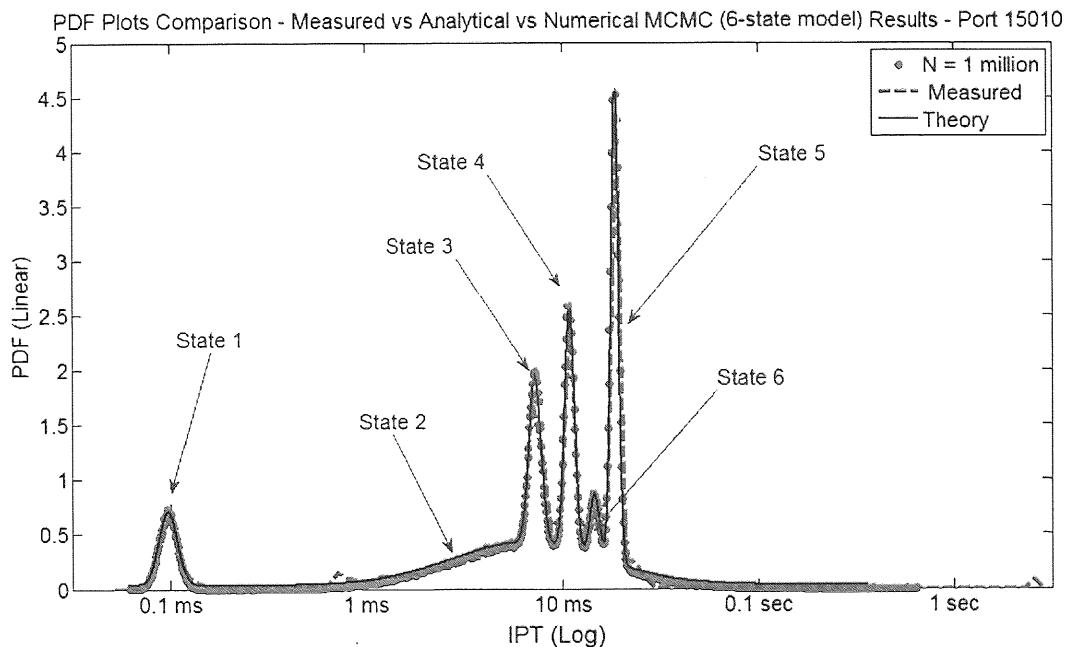


Figure 3.17: PDF Plots of modelled, analytical and measured results of IPT for IP Telephony Traffic

As shown in Fig. 3.17, the 6-state Gaussian Markov Model nearly fits perfectly to the PDF results achieved from the measurement data of IPT based on traffic from an IP Telephony application. The goodness of fit can also be visualised by looking at the CDF plots for both the measured result and the modelled result plotted together as shown in Fig. 3.18.

We observe that the modelled data closely follows the measured IPT result, and we can conclude that these Gaussian Markov Models, which are simple to estimate, implement and have a white box approach, are very useful for modelling traffic IPT PDF results.

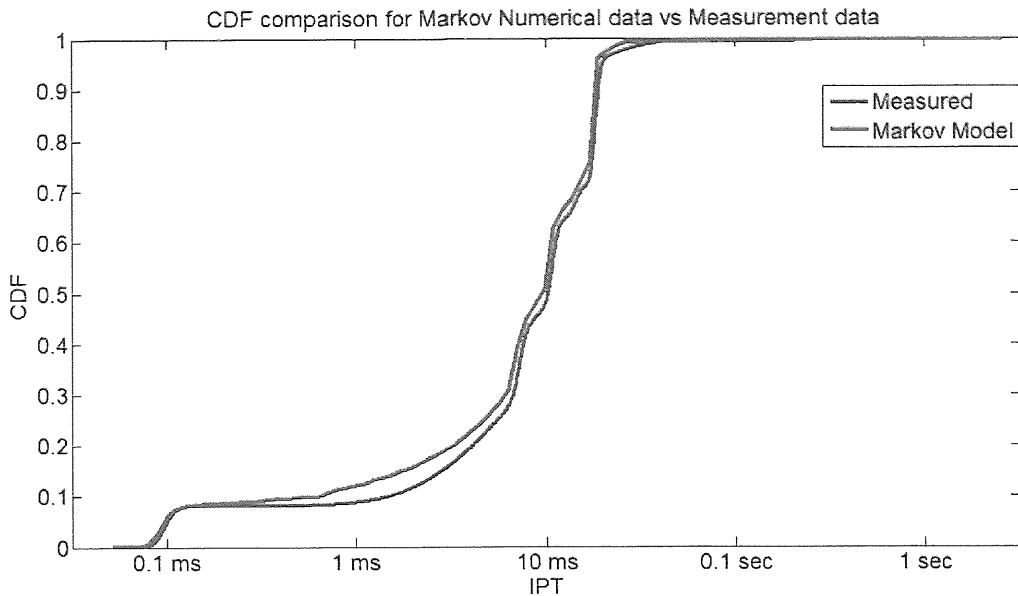


Figure 3.18: Comparison of CDF plots for measured and modelled results for IPT of IP Telephony Traffic.

## 3.5 Summary

In this chapter, we have looked at how we can model the first order statistics of IPT by using Gaussian Markov Models. We started by looking at the various tools and steps that were required in measuring IP traffic statistics and then summarised three examples of measured results, before looking at a detailed procedure for fitting Gaussian Markov Models to the measured statistics. We have modelled the three examples of measured results associated with IP traffic from IP radio broadcasting, IP TV broadcasting and IP telephony services respectively. Finally, upon comparing the modelled result with the measured results, we conclude that the Gaussian Markov Models are good candidates for modelling network IPT statistics. While we have looked at the measured Joint Density results, we have only considered modelling the Probability Density (first order statistics) using the Gaussian Markov Models. However, the Joint Density Plots are extremely useful and significant, as they provide a lot more information than the Probability Density plots. In the next two chapters we will look at how we can use these Gaussian Markov Models to model interesting elements observed in the measured Joint Density results, and also establish the significance of using Joint Density plots in traffic modelling.

# Chapter Four

## Understanding and modelling Joint Density Results – Packet Sequences

In the previous chapter, we had seen how we could model the first order statistics, that is, the Probability density function (PDF) of the measured IPT for IP packet traffic, using simple Gaussian Markov Models. It was observed that the PDFs of the modelled results fit the PDFs of the measured results very well. Often in modelling traffic, the PDF has been very widely used, to produce mathematical models of measured results [67] [68] [69] [70] [71]. So far, we have considered the Probability density function for modelling the measured result and we will now extend the work further by looking at the effect of using the Joint Density Function (JDF), in section 4.1, and see whether the same Gaussian Markov models used to model the Probability Density Function, can also be used to model the Joint Density Function of the measured results. In section 4.2, we prove the significance and the uniqueness of the characteristics of packet traffic statistics as observed from the Joint Density plot, by considering two packet traffic models of two completely different topologies which have two different Joint Density Functions (JDF), but yet the same Probability Density Function (PDF). Section 4.3 looks at the symmetrical and asymmetrical nature of the modelled and measured IPT Joint density results and finally section 4.4 looks at the impact of introducing random packet loss in the modelled results.

### 4.1 The Joint Density results and Packet Sequences

#### 4.1.1 Joint Density Plots of Modelled Results

Let us start by considering the 6-state (6+1 state) Gaussian Markov model that was proposed in section 3.4.3 of the previous chapter to model the statistics of IPT associated with traffic generated from an IP telephony application, measured at the UDP Port 15010.

The PDF of the measured and modelled result which was shown in Fig 3.16, has been produced here as Fig.4.1 for convenience.

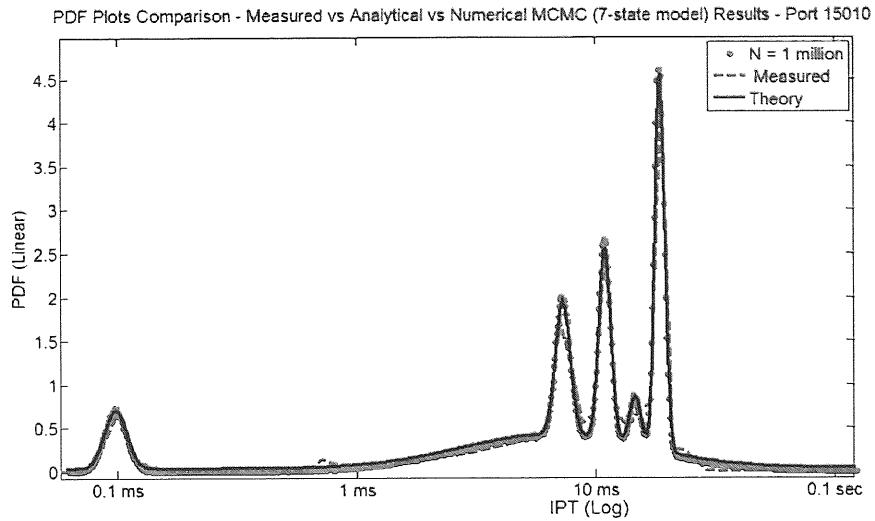


Figure 4.1: PDF Plots of modelled, analytical and measured results of IPT for IP Telephony Traffic  
The PDF plot conveys that the model has captured the timing structure of the packets very well. However the next question is whether the proposed model, models the Joint Density (JDF) result as well. Fig. 4.2 shows the Joint Density Function (JDF) of the simulated result from the same 6-state Gaussian Markov Model which was proposed in section 3.4.3 to model the IPT PDF statistics of IP Telephony traffic measured at Port 15010.

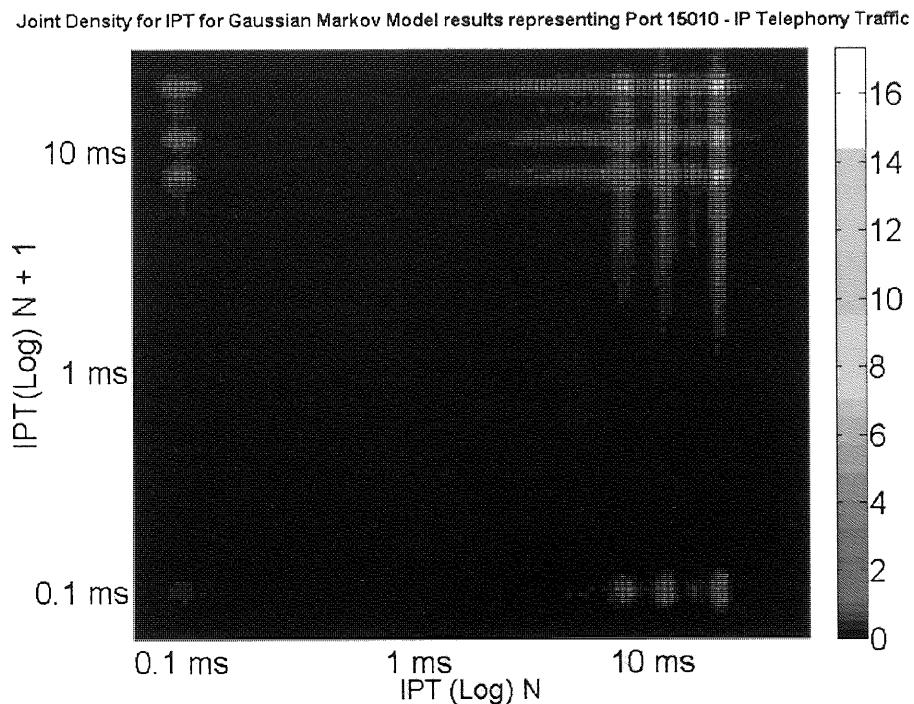


Figure 4.2: Joint Density Plot of IPT of Gaussian Markov model results of IPT for IP Telephony Traffic

Comparing the modelled JDF result shown in Fig. 4.2 with the measured JDF result shown in Fig. 4.3, we observe that the modelled result is inconsistent with the measured result, in the sense that the Gaussian Markov Model which was modelled on the PDF of the measured results, has failed to capture all the features observed in the measured Joint Density result.

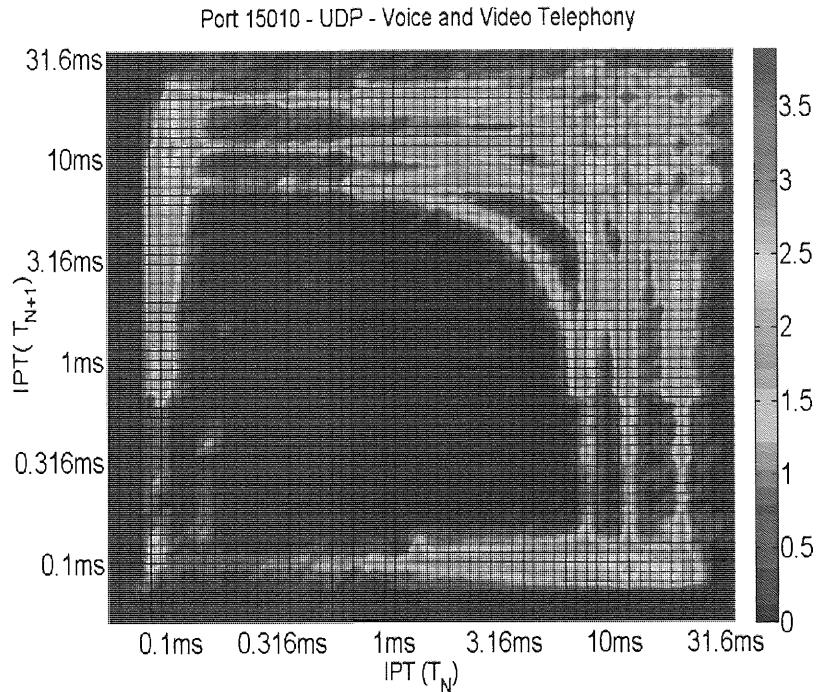


Figure 4.3: Joint Density Function for IPT for IP Packets on UDP Port 15010 – IP Telephony

This shows that a model based on the PDF, merely reproduces the statistics shown by the PDF, but does not account for the details of the timing structure and sequence of times related to the packet flows, which the Joint density explains. As observed by comparing Fig. 4.2 and Fig. 4.3, the Gaussian Markov Model has only captured two sets of symmetric structures: the structure close to the axes lines, and the symmetric structure in the top right corner. The symmetric structure in the top right corner indicates that one of the set of six time frames of IPT is always followed by any of the set of six time frames of IPT. This shall become clearer to understand, as we consider some simple hypothetical examples of Markov traffic models and study the packet sequences and the symmetric structure of the Joint Density plots which shall be covered in sections 4.2 and 4.3 respectively.

### 4.1.2 Packet Sequence Information

One of the major reasons for considering these Joint Density plots, as described in Chapter 2, is to understand in great detail, the features associated with packet source behaviour. One such feature which the Probability Density results do not show us, is the *packet sequence*. The idea of having the IPT samples vector 'N', cyclic shifted '(N+1)' and having it plotted against the original IPT samples vector 'N', gives us the possibility of knowing what inter-packet time (IPT) is going to follow after the previous inter-packet time (IPT) for all packets captured, for which the statistics were plotted. This way, looking at the Joint Density results, we can classify the IPTs into certain groups or sets of time and assume a Gaussian distribution for the IPTs for each significant peak that appears in the Joint Density results. Having done that, we get a clearer picture on the number of possible packet sources that have generated the traffic, based on the IPT statistics that have been measured or simulated. With the aid of Fig. 4.4, we can understand the association of Gaussian peaks, observed in the Joint Density results, with the packet sequences.

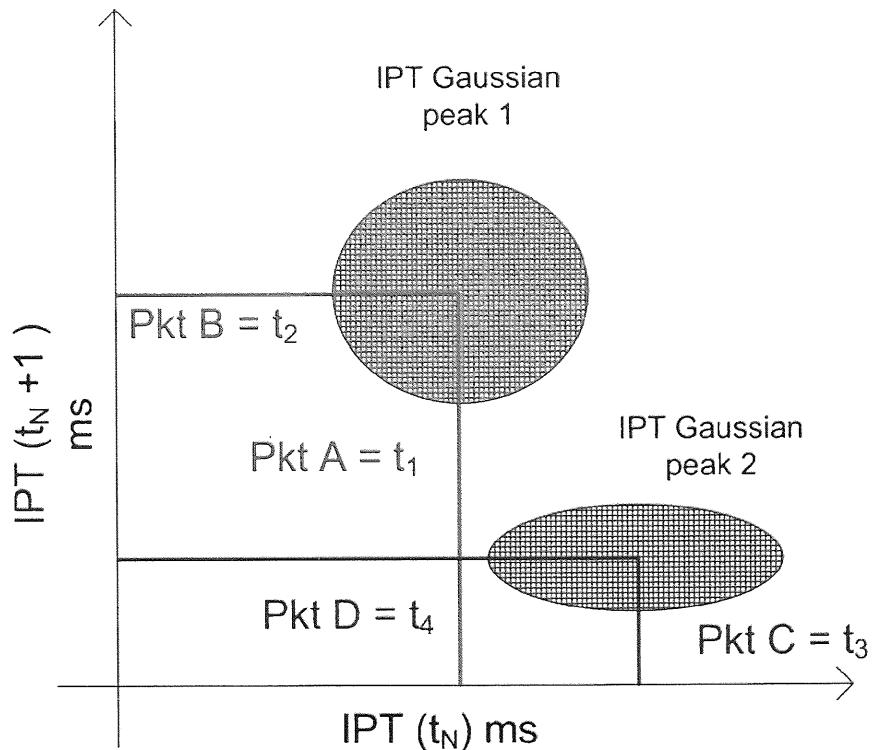


Figure 4.4: Sketch showing association of IPT Gaussian peaks in Joint density (sub-set) with packet sequences.

Fig 4.4 shows a sketch of a sub-set of a Joint Density result as an example, where the X-axis represents the current sample of IPT, the Y-axis represents the next sample in time, and the Z-axis (given by intensity) represents the density. Let us consider the two Gaussian peaks shown in Fig. 4.4.

These Gaussian peaks show that the IPTs are concentrated in two groups of times. One group of times, consists of IPT samples distributed around a mean value  $(t_1, t_2)$  as shown by the Gaussian peak 1, and another group of times consists of IPT samples that are distributed around a mean value of  $(t_3, t_4)$  as shown by Gaussian peak 2. The two Gaussian peaks are not linked to each other in terms of the time axis (for this subset), so there are four different time sets, hence the possibility of four different packets arriving at four different time intervals. The Gaussian peak 1 shows that a packet 'A' arriving at a time  $t_1$  after the previous packet gets followed by a packet 'B' after a time  $t_2$ . Similarly, the Gaussian peak 2 represents a packet 'C' arriving after a time  $t_3$  of the previous packet, gets followed by a packet 'D' after a time  $t_4$ . In this particular example, the times  $(t_1, t_2)$  and  $(t_3, t_4)$  are assumed to be independent of each other, although this is not always the case with the results as we shall see in the next section.

## 4.2 Packet Sequences and Markov Models

In this section, we will look at some very interesting results with which we will establish the significance of the Joint Density plots. We will consider two hypothetical examples of network traffic Markov Models of different topologies but with the same number of states having the same means, standard deviations and state visiting probabilities. We use these two models in Monte Carlo simulations and then plot both the Probability Density and Joint Density results for statistical analysis. First, we will define the state holding time distribution parameters and describe the other features of the models.

### 4.2.1 State holding time distribution parameters and model setup

As mentioned earlier, we will consider two versions of the hypothetical example of a three - state Gaussian Markov Traffic model that contains the following common state distribution parameters for each of the three states, defined by equations (4.1), (4.2) and (4.3) respectively:

State 1 emits Packet A, with the following state holding time distribution:

$$P_1(t)_{(Gaussian)} = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{\frac{-(t-\mu_1)^2}{2\sigma_1^2}} \quad (4.1)$$

where  $\mu_1 = 2.5$  ms and  $\sigma_1 = 0.75$  ms.

State 2 emits Packet B, with the following state holding time distribution:

$$P_2(t)_{(Gaussian)} = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{\frac{-(t-\mu_2)^2}{2\sigma_2^2}} \quad (4.2)$$

where  $\mu_2 = 14$  ms and  $\sigma_2 = 2.645$  ms.

State 3 emits Packet C, with the following state holding time distribution:

$$P_3(t)_{(Gaussian)} = \frac{1}{\sigma_3 \sqrt{2\pi}} e^{\frac{-(t-\mu_3)^2}{2\sigma_3^2}} \quad (4.3)$$

where  $\mu_3 = 40$  ms and  $\sigma_3 = 6.32$  ms.

State visiting Probabilities:

To keep both the models very simple, we shall assume equal state visiting probabilities:

$$V_1 = V_2 = V_3 = 1/3 \quad (4.4)$$

#### 4.2.2 Case 1: Markov Model ‘A’ with Packets in Random Order

The first version of the Markov Model, which we shall call Markov Model ‘A’, is described by the state transition diagram shown in Fig. 4.5.

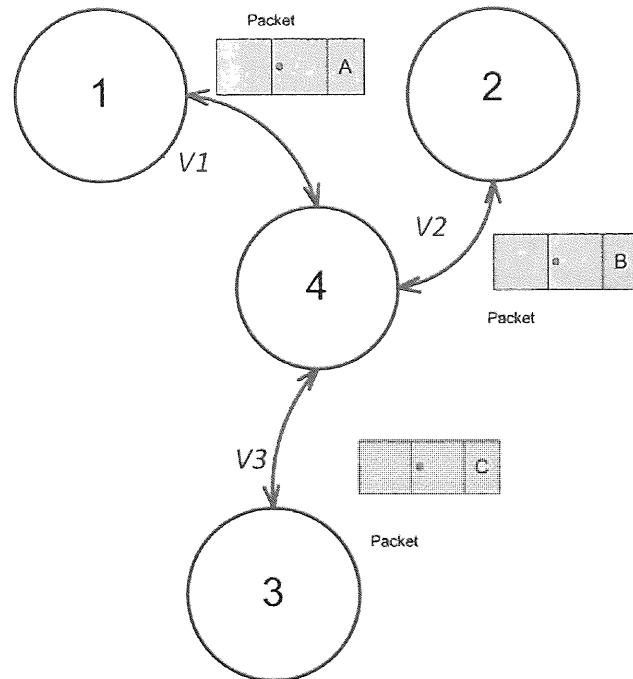


Figure 4.5: Markov Model ‘A’ with packets A,B and C in random order

As before, for reasons stated earlier in Chapter 3, we have a fourth *feeder* state for generating the state visiting probabilities. The time spent in this *feeder* state is not included in the IPT statistics as it is present only for generating transitions to states in a random manner based on the equal branching probabilities to each state. Visiting any of the three states is equally likely, because of the equal state visiting probabilities. As shown in Fig. 4.5, the model can transit from state 4 to any of the other three states with a probability of  $1/3$  and return back to state 4 with a probability 1, after which, again it can transit to any of the three states. Thus the packets A, B and C are in complete random order, as shown in the example extract of a Monte Carlo simulation packet sequence:

Packet Sequence in Monte Carlo Simulation for Markov Model 'A' = ACCBCAAACBCBBACBBCCAB..

#### 4.2.2.1 Probability Density Plot for IPT of Markov Model 'A'

For statistical analysis, we will first look at the Probability Density plot for IPT of the Markov Model 'A' as shown in Fig. 4.6.

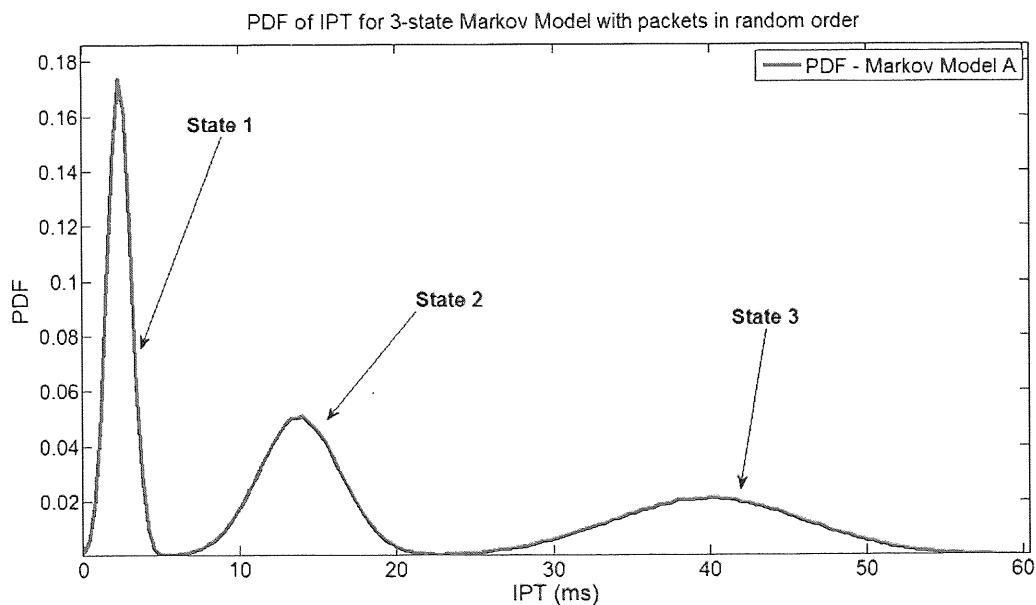


Figure 4.6: PDF plot of IPT for Markov Model 'A' with packets in Random Order

The PDF result is very straightforward and we clearly observe three peaks associated with the three Gaussian distributions of the three states located at the mean values of 2.5 ms, 14 ms and 40 ms respectively. Next, it will be more interesting to look at the Joint Density result.

#### 4.2.2.2 Joint Density Plot for IPT of Markov Model ‘A’

Fig. 4.7 shows the Joint Density result of IPT for Markov Model ‘A’. We see that it has a symmetric surface with nine peaks. The feeder state, i.e. state 4, has equally probable transitions to States 1, 2 and 3 which emit packets A, B and C respectively. Any of the three packets can be followed by any of the three packets, and hence we have,  $3 \times 3 = 9$ , possibilities of having peaks associated with all nine combinations of packet sequence pairs for the three packets. A more detailed analysis on the peaks and knowledge of the packet sequences from the Joint Density plot for the Markov Model ‘A’ follows in section 4.2.2.3.

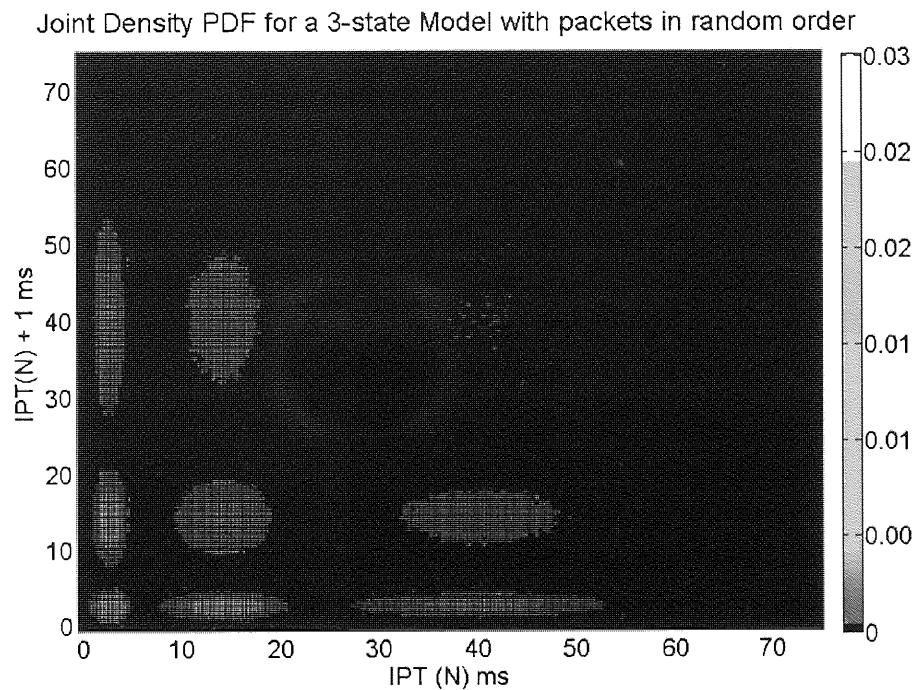


Figure 4.7: JDF plot of IPT for Markov Model ‘A’ with packets in Random Order

#### 4.2.2.3 Analysis of Packet Sequences from Joint Density for Markov Model ‘A’

Let us number the peaks in Fig. 4.7, from 1 to 9 in order, starting from the top left to right and then bottom in the same manner. Beginning with peak 1 in the top left corner, and moving down in that order, the packet sequence information we get from the peaks is as follows:

**Peak 1: Packet A followed by C**

Looking at the position of the peak maximum on both the axes, it corresponds to a mean of  $\mu_1 = 2.5$  on the X-axis (N) and  $\mu_3 = 40$  on the Y-Axis (N+1). This means that an IPT sample from the distribution associated with State 1 was followed by an IPT sample from the distribution associated with State 3. Since we know, that State 1 emits packet A and State 3 emits packet C, we can conclude that the peak represents a packet A being followed by a packet C.

State 1 → State 3 : Packet A → Packet C

Similarly from Fig. 4.7, the following can also be revealed from the other peaks.

**Peak 2: Packet B followed by C**

The peak is represented by  $\mu_2 = 14$  on the X-axis (N) and  $\mu_3 = 40$  on the Y-Axis (N+1).

This means, an IPT from State 2 was followed by an IPT from State 3.

Therefore this peak shows that, Packet B was followed by C.

State 2 → State 3 : Packet B → Packet C

**Peak 3: Packet C followed by C**

The peak is represented by  $\mu_3 = 40$  on the X-axis (N) and  $\mu_3 = 40$  on the Y-Axis (N+1).

This means, an IPT from State 3 was followed by an IPT from State 3.

Therefore this peak shows that, Packet C was followed by C.

State 3 → State 3 : Packet C → Packet C

**Peak 4: Packet A followed by B**

The Peak is represented by  $\mu_1 = 2.5$  on the X-axis (N) and  $\mu_2 = 14$  on the Y-Axis (N+1).

This means, an IPT from State 1 was followed by an IPT from State 2.

Therefore this peak shows that, Packet A was followed by B.

State 1 → State 2 : Packet A → Packet B

**Peak 5: Packet B followed by B**

The peak is represented by  $\mu_2 = 14$  on the X-axis (N) and  $\mu_2 = 14$  on the Y-Axis (N+1).

This means, an IPT from State 2 was followed by an IPT from State2.

Therefore this peak shows that, Packet B was followed by B.

State 2 → State 2 : Packet B → Packet B

**Peak 6: Packet C followed by B**

The peak is represented by  $\mu_3 = 40$  on the X-axis (N) and  $\mu_2 = 14$  on the Y-Axis (N+1).

This means, an IPT from State 3 was followed by an IPT from State 2.

Therefore this peak shows that, Packet C was followed by B.

State 3 → State 2 : Packet C → Packet B

**Peak 7: Packet A followed by A**

The peak is represented by  $\mu_1 = 2.5$  on the X-axis (N) and  $\mu_1 = 2.5$  on the Y-Axis (N+1).

This means, an IPT from State 1 was followed by an IPT from State 1.

Therefore this peak shows that, Packet A was followed by A.

State 1 → State 1 : Packet A → Packet A

**Peak 8: Packet B followed by A**

The peak is represented by  $\mu_2 = 14$  on the X-axis (N) and  $\mu_1 = 2.5$  on the Y-Axis (N+1).

This means, an IPT from State 2 was followed by an IPT from State1.

Therefore this peak shows that, Packet B was followed by A.

State 2 → State 1 : Packet B → Packet A

### Peak 9: Packet C followed by A

And finally, Peak 9 is represented by  $\mu_3 = 40$  on the X-axis (N) and  $\mu_1 = 2.5$  on the Y-Axis (N+1).

This means, an IPT from State 3 was followed by an IPT from State 1.

Therefore this peak shows that, Packet C was followed by A.

State 3 → State 1 : Packet C → Packet A

Thus the nine possibilities of packet sequence pairs corresponding to the nine peaks that have been observed in the Joint Density plot of IPT for the Markov Model 'A' have been clearly explained.

### 4.2.3 Case 2: Markov Model 'B' with Packets in Sequence (Loop)

Next, we move on to the other version of the Gaussian Markov model, called Markov Model 'B', where the states, as before, emit packets A, B and C, but this time in a strict sequence. The model has a three-state loop structure. Once the model enters State 1, the next state it must transit to is State 2 and then from State 2 it must transit to State 3 and so on. Therefore the packet sequence would be A followed by B and B followed by C and so on.

Packet Sequence in Monte Carlo Simulation for Markov Model 'B' = ABCABCABCABCABCABCABC..

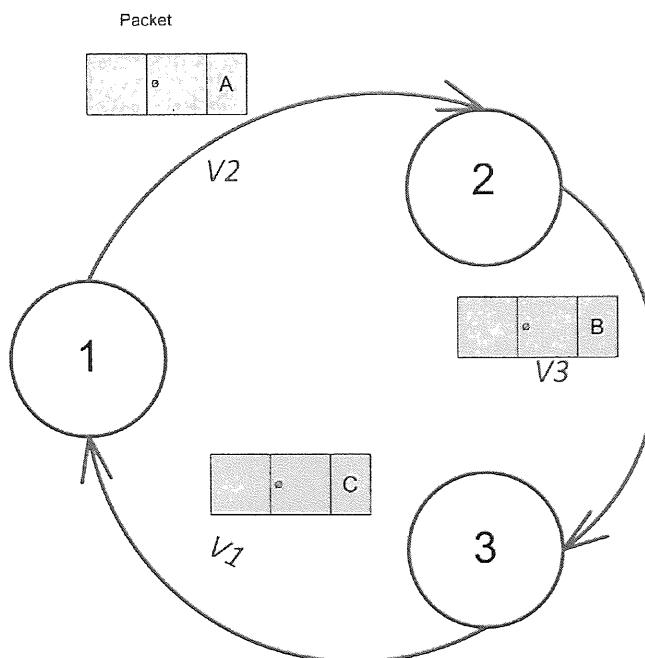


Figure 4.8: Markov Model 'B' with packets A,B and C in sequence

#### 4.2.3.1 Probability Density Plot for IPT of Markov Model ‘B’

Looking at the PDF of the IPT for Markov Model ‘B’ as shown in Fig 4.9, we observe that it looks the same as that of the Markov Model ‘A’.

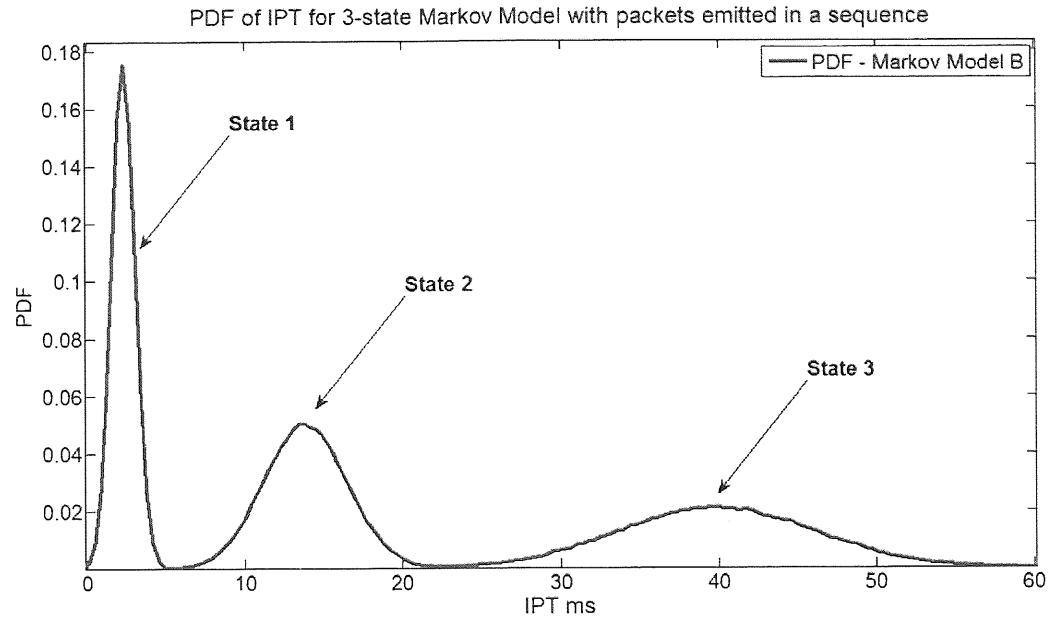


Figure 4.9: PDF plot of IPT for Markov Model ‘B’ with packets in sequence

The PDFs of both the models are identical by construction and to show this the two are plotted together in Fig. 4.10.

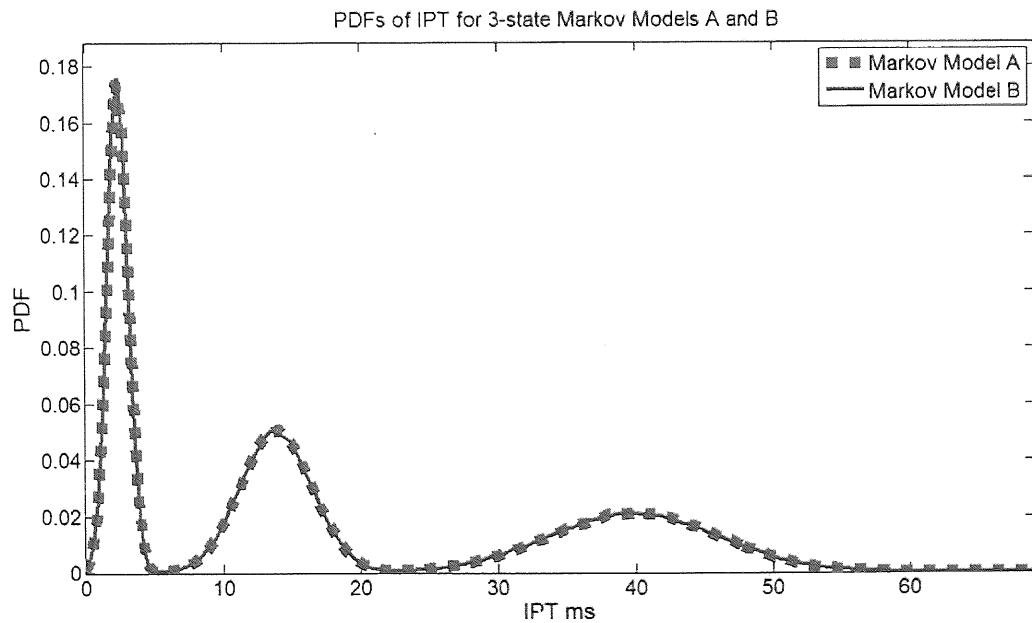


Figure 4.10: Comparison of PDFs of IPT for Markov Models ‘A’ and ‘B’

#### 4.2.3.2 Joint Density Plot for IPT of Markov Model ‘B’

Next, we move on to the Joint Density Function for the 3-state Gaussian Markov Model ‘B’.

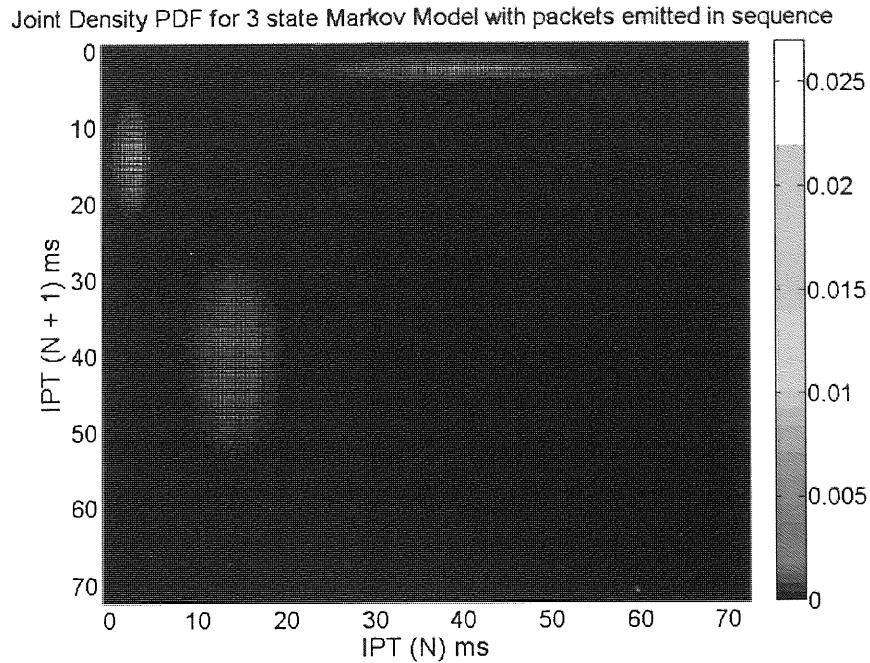


Figure 4.11: JDF plot of IPT for Markov Model ‘B’ with packets in sequence

Here, we notice a significant difference when compared to the Joint Density result of the previous model ‘A’. We see only three peaks as opposed to the nine peaks observed in the previous case. This is clearly because the states are visited in a loop and the order of the states ‘ $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$ ’ is always maintained. This is explained in further detail as follows:

#### 4.2.3.3 Analysis of Packet Sequences from Joint Density for Markov Model ‘B’

Let us number the peaks 1, 2 and 3 (from top left) approaching in an anti-clockwise direction:

##### Peak 1: Packet A followed by B

The peak is represented by  $\mu_1 = 2.5$  on the X-axis ( $N$ ) and  $\mu_2 = 14$  on the Y-Axis ( $N+1$ ).

This means, an IPT from State 1 was followed by an IPT from State 2.

Therefore this peak shows that, Packet A was followed by B.

State 1 $\rightarrow$ State 2	:	Packet A $\rightarrow$ Packet B
-------------------------------	---	---------------------------------

### Peak 2: Packet B followed by C

The peak is represented by  $\mu_2 = 14$  on the X-axis (N) and  $\mu_3 = 40$  on the Y-Axis (N+1).

This means, an IPT from State 2 was followed by an IPT from State 3.

Therefore this peak shows that, Packet B was followed by C.

State 2 → State 3 : Packet B → Packet C

### Peak 3: Packet C followed by A

The peak is represented by  $\mu_3 = 40$  on the X-axis (N) and  $\mu_1 = 2.5$  on the Y-Axis (N+1).

This means, an IPT from State 3 was followed by an IPT from State 1.

Therefore this peak shows that, Packet C was followed by A.

State 3 → State 1 : Packet C → Packet A

Thus the Joint Density Function (JDF) establishes the cyclic sequence of the packets from A to B and B to C and so on. The conclusion from the example of the two Markov models is very significant. We have seen how two Markov Traffic Models having different topologies have exactly the same PDF but different JDFs. This again highlights the significance of the Joint Density results, as they give much more information than the PDF on its own.

## 4.3 The Probability Density Function: A projection of the Joint Density Function

Another important observation made from the work with the Joint density results is that the Probability Density Function is a projection of the Joint Density results. In a simple analogy, the PDF can be thought of as a shadow on the wall with the three-dimensional surface of the Joint Density Function (JDF) placed in front of a lamp. In other words the PDF is a two dimensional projection of a three dimensional Joint Density plot. In mathematical terms, the PDF is the integral of Joint Density along one axis. Fig. 4.12 proves this phenomenon with reference to the joint density results of IPT of Markov Models ‘A’ and ‘B’ and their common PDF.

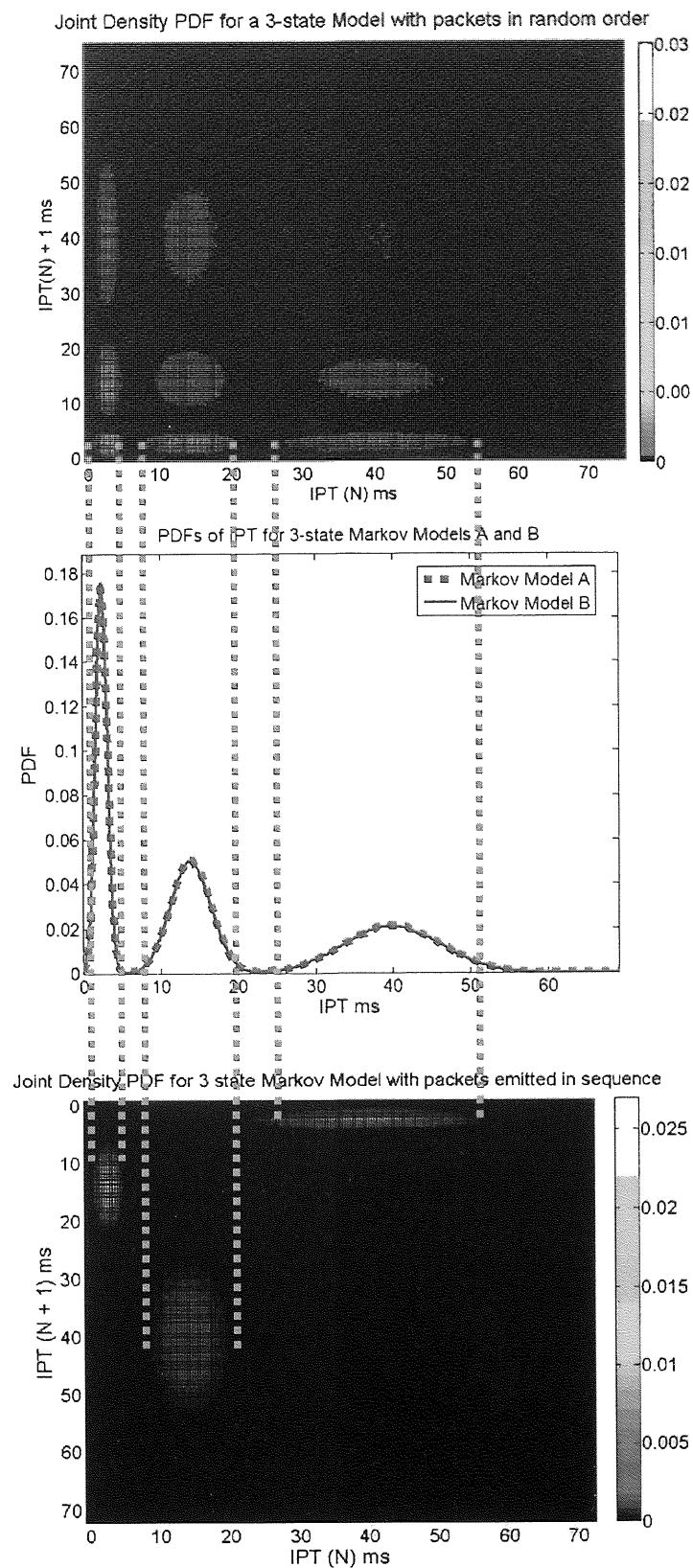


Figure 4.12: The PDF is a Projection of the Joint Densities of the two Markov Models 'A' and 'B'

## 4.4 The Symmetrical and Asymmetrical nature of Joint Density results

Another observation made from the study of the Joint Density results of both the measured and modelled results, was on the symmetrical and asymmetrical nature of the Joint Density results. Referring to Fig. 4.8, it is shown, that the Joint Density results of the Markov Model 'A' are symmetric in nature, in that the results are identical when reflected through the line  $t_n=t_{n+1}$ . On the other hand, from Fig. 4.11, we observe that the Joint Density results of the Markov Model 'B' are asymmetric in nature. From the use of Markov Model 'A' and Markov Model 'B', we have seen that the symmetry arises with the condition that any of a set of finite number of packets is followed by any of the set of finite number of packets, whereas the asymmetric nature of the results is due to a strict sequence of packet flow. It would be useful to detect whether the measured IPT statistics are of symmetric or of asymmetric nature, as this also gives us further information on the behaviour of the network packet sources while the measurements were being made. In order to detect and filter out the symmetric and asymmetric part of a measured or modelled result, a simple algorithm described in section 4.4.1 was used.

### 4.4.1 Detecting and filtering out Symmetric and Asymmetric structure from the Joint Density Results

The algorithm used for detecting and filtering out the asymmetric and symmetric structure of the Joint Density result was based on a simple arithmetical approach. The first step in the process is to use the Joint Density result and create a mirror image of the result, by producing another Joint Density plot with the axes swapped. The asymmetric structure of the original density plot is then the positive difference (absolute value) of the original density result and the density result with the axes swapped. Once the asymmetric structure is obtained, the symmetric structure can also be achieved by subtracting the asymmetric structure from the original Joint Density result. Equations (4.4) and (4.5), explain this procedure mathematically.

$$Asym(N, N + 1) = |(JDF(N, N + 1) - JDF(N + 1, N))| \quad (4.4)$$

And:

$$Sym(N, N + 1) = [JDF(N, N + 1) - Asym(N + 1, N)] \quad (4.5)$$

Where:

$\text{Asym}(N, N+1)$  = Asymmetric Part of the Joint Density Function

$\text{JDF}(N, N+1)$  = The original JDF result

$\text{JDF}(N+1, N)$  = The JDF result with the axes swapped

$\text{Sym}(N, N+1)$  = Symmetric Part of the Joint Density Function

#### 4.4.2 Symmetric and Asymmetric nature of the modelled results

To confirm the symmetric and asymmetric nature of the modelled results, let us again consider the Joint Density results of Markov Model ‘A’ and Markov Model ‘B’. We shall put the symmetric and asymmetric filter algorithm to test to verify the asymmetry and symmetry of the results for these models. This time, we will produce the Joint Density plots with a three dimensional view of the peaks rather than the two-dimensional view that we had been considering so far.

##### 4.4.2.1 Markov Model ‘A’

We will first consider the JDF for Markov model ‘A’, and upon executing the first step of our filtering process given by equation (4.4), we obtain Fig. 4.13. As expected, the Joint Density results of Markov Model ‘B’ being completely symmetric in nature do not have any asymmetric components. The z-axis has been scaled to the original maximum density level to reflect the zero plane of the asymmetric result. The only minute amount of elements present in the figure is the asymmetric content of the noise associated with random number generator that results as part of the Monte Carlo simulation.

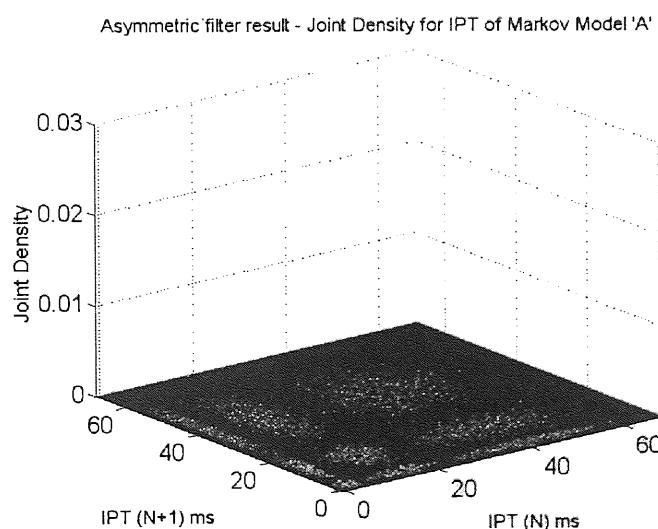


Figure 4.13: Asymmetric structure of Joint density Result for IPT of Markov Model ‘A’

Executing the next step, guided by equation (4.5), gives the symmetric structure of the Joint Density Result for Markov Model ‘A’, and as seen in Fig. 4.14, the entire original structure of the Joint Density result is retained as the original result is purely symmetric.

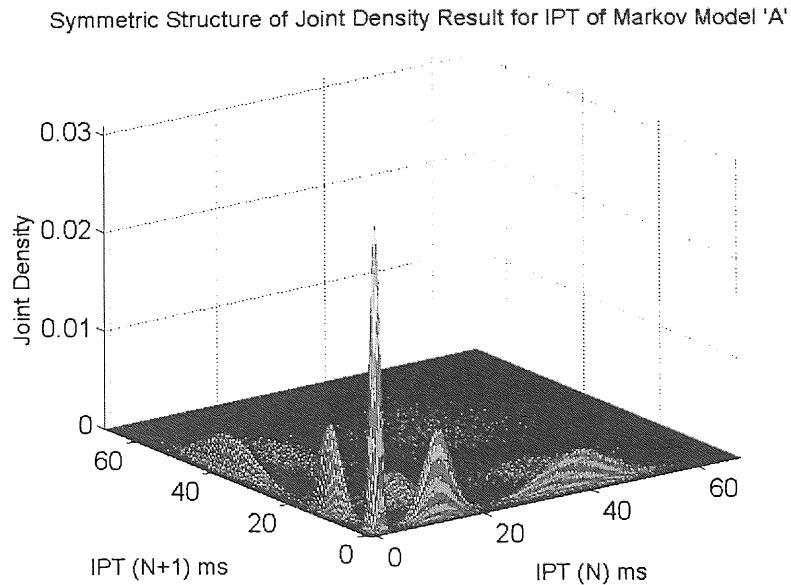


Figure 4.14: Symmetric Structure of Joint density result for IPT of Markov Model ‘A’

#### 4.4.2.2 Markov Model ‘B’

Repeating the same procedure with the simulated Joint Density result for Markov Model ‘B’, we first extract the asymmetric structure of the result given in Fig. 4.15. This time, as expected, the original result being asymmetric in nature has the original structure entirely retained after asymmetric filter.

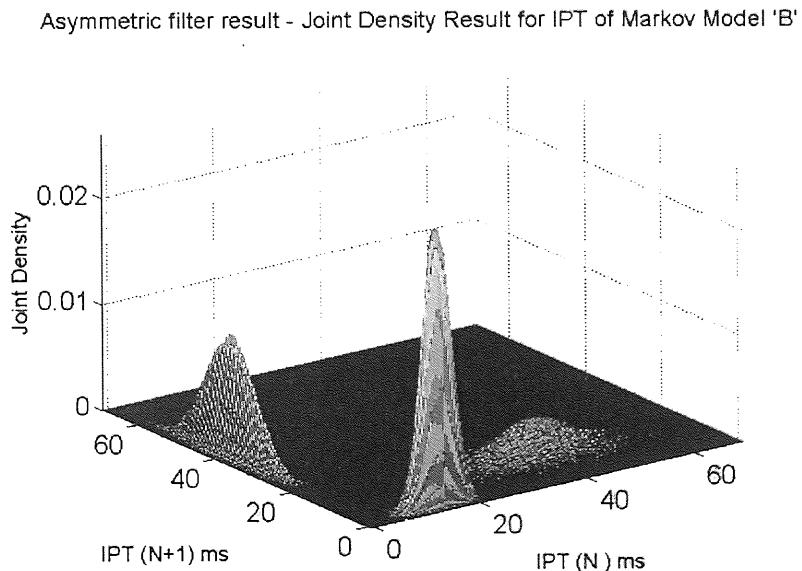


Figure 4.15: Asymmetric structure of Joint Density Result for IPT of Markov Model ‘B’

And of course for the symmetric test result, we observe only the zero plane except for what is only a minute amount of symmetric noise associated with the random number generator.

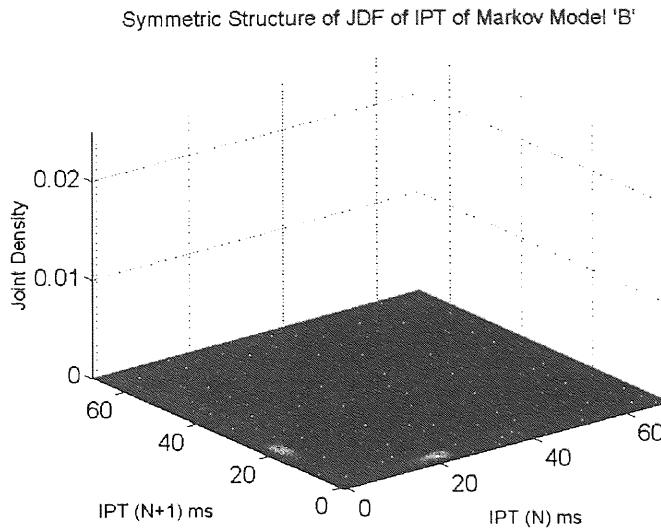


Figure 4.16: Symmetric structure of JDF of IPT of Markov Model 'B'

#### 4.4.3 Symmetrical nature of measured results

Having tested the algorithm for detecting and filtering symmetric and asymmetric structures and also having understood the cause of symmetric and asymmetric conditions, we used the test on all of the measured Joint density results, and upon analysing it was observed that almost all the Joint Density results for the measured IP traffic were symmetrical in nature. Fig. 4.17 shows the residual symmetric structure, obtained after the test, for the Joint Density results of IPT for traffic of IP telephony measured at UDP Port 15010.

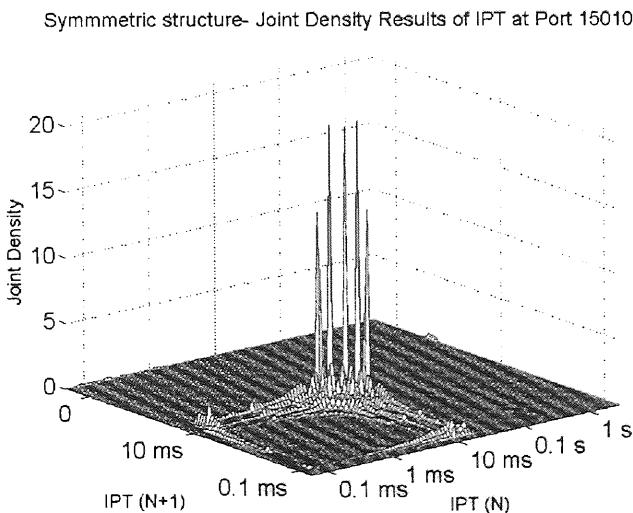


Figure 4.17: Symmetric Structure of Joint Density Results of IPT at Port 15010

While the symmetric structure contains the entire Joint Density surface, the asymmetric part displays an almost zero plane, thus confirming that the measured result is purely symmetric in nature.

Asymmetric Filter result - Joint Density Result for IPT at Port 15010

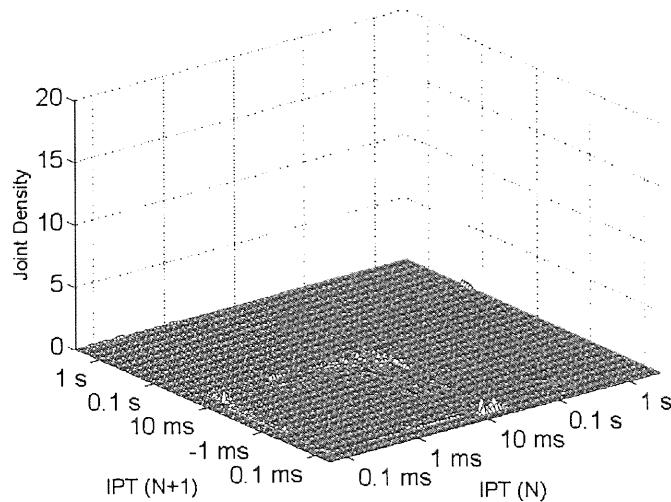


Figure 4.18: Asymmetric Structure of the Joint Density result for IPT at Port 15010

The fact that the majority of the measured results have been found to be of symmetrical nature, reveals that the packet traffic sources generate a chain of packets where the packets are in random order and not in a specific sequence. Again, we have attained this knowledge by studying the Joint Density plots.

## 4.5 Impact of introducing Packet Loss in Markov Traffic Models

Having studied the packet sequences and the symmetrical and asymmetrical nature of the Joint Density Results so far, we will now look at the impact of introducing random packet loss in the Markov Monte Carlo simulation. We will investigate the effect of introducing minor and major packet losses on the Joint Density results.

### 4.5.1 Creating Random packet loss in the Markov Monte Carlo Simulation

The random packet loss is created in the Markov Monte Carlo simulation by randomly dropping some packets irrespective of which source or state it originates from. A random number generator is used to decide on dropping the packets based on an assigned probability percentage level of packet

loss. A 10% packet loss reflects that one in ten packets are dropped randomly, whereas a 50% packet loss means five in ten packets are dropped randomly. For balancing the statistics, the Markov Monte Carlo simulation program was modified to balance the loss by generating more samples to adjust  $N = 1$  million. From a network perspective, the extra packets generated can be considered as retransmissions made by the sources, as in a real network. As mentioned earlier, the IPT is calculated between the last packet received and the current one. So if we lose a packet, the IPT is the sum of the IPT for the lost packet and the one for the previous packet. If we lose two packets in a row, the IPT is the sum of three IPT values all together. This revised calculation technique based on the packet loss, which was implemented in the Markov Monte Carlo simulation, along with the original calculation method is illustrated in Fig. 4.19.

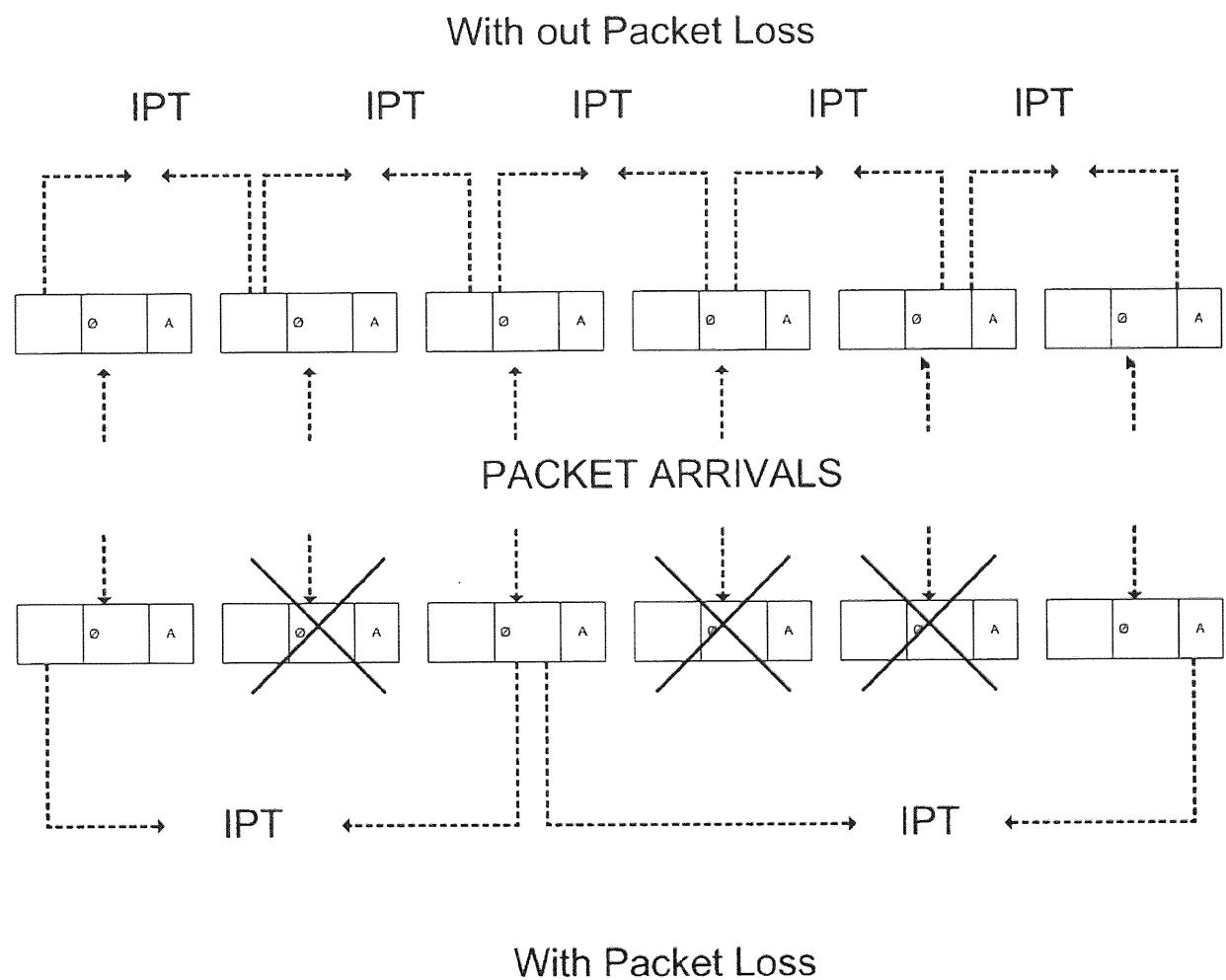


Figure 4.19: Illustration of calculating IPTs in the Markov Monte Carlo Simulation with and without Packet Loss.

### 4.5.2 Impact of packet loss on the Joint Density of IPT

To keep things simple and easy for understanding, we will again consider the IPT Joint Density of the simulated result obtained from the Markov Monte Carlo Simulation of Markov Model ‘B’. The original Joint Density in the three dimensional form looks like as shown in Fig. 4.15 (section 4.4.2.2). Keeping that in mind, we will now look at introducing 10% random packet loss in the Markov Monte Carlo Simulation and plot the Joint Density results for the new set of IPTs as shown in Fig.4.20.

Impact of 10% Random Packet loss - Joint Density Results for IPT for Markov Model

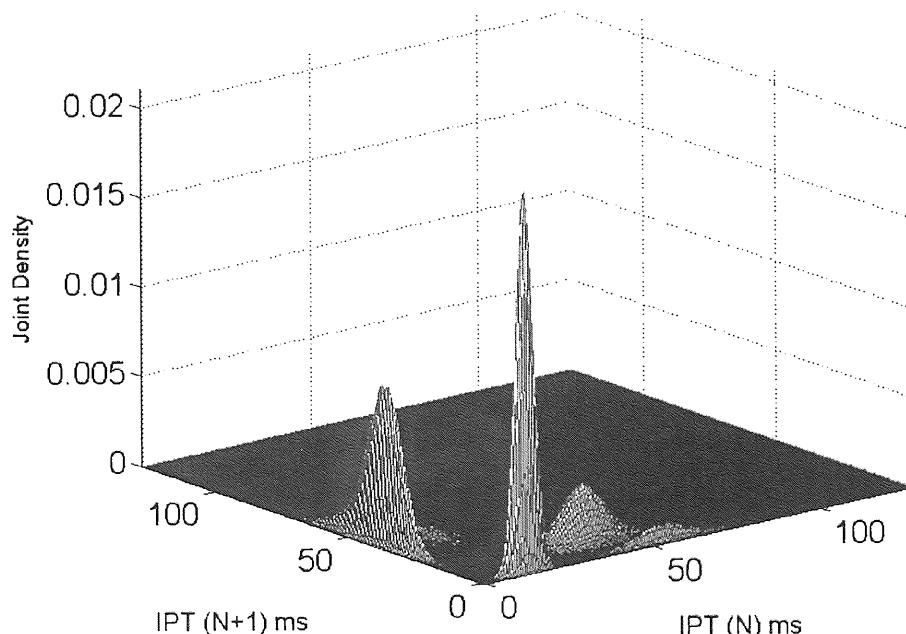


Figure 4.20: Joint Density Results for Markov Model ‘B’ with 10% packet loss

Referring to Fig. 4.20 and looking at the impact of 10% packet loss, we can immediately notice that the original asymmetric structure has changed, with new peaks being introduced at higher IPT time scales. This is because the sum of IPTs results in additional peaks at larger values. The change in the original structure, because of the packet loss can be associated with the percentage of packet loss introduced into the Monte Carlo simulation. A 50% packet loss results in more peaks at larger times and around 50% deflection to the original asymmetric structure causing it to dissolve into the new peaks originating at larger times. Let us look at one more result of the other extreme which uses 90% packet loss. Fig. 4.21 shows a contour version (for sake of clarity) of the Joint Density plot exhibiting the impact of 90% packet loss to the IPT Joint Density Statistics.

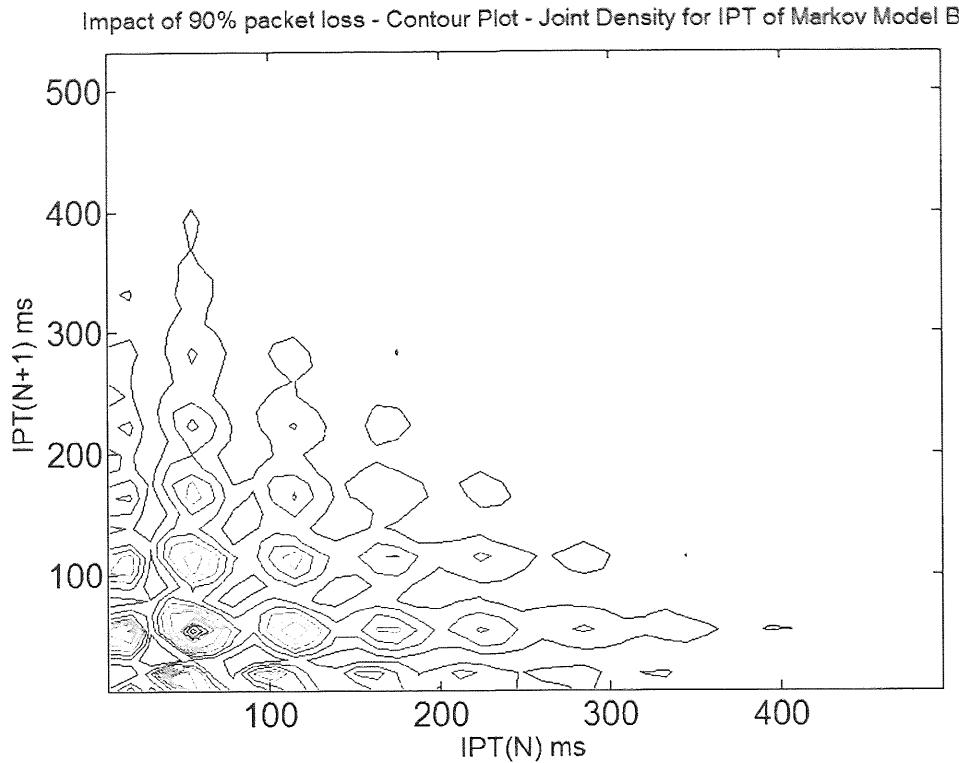


Figure 4.21: Joint Density Results (Contour version) for Markov Model ‘B’ with 90% packet loss

We observe that IPTs are as long as 400 ms. Another important observation made from this result is that the original asymmetric structure has now turned into a widely spread symmetric structure. This observation has been consistent with other results with high packet loss. Hence, we conclude that packet loss introduces elements of symmetry into the IPT statistics.

## 4.6 Summary

Throughout this Chapter, we have highlighted in detail, the significance and the advantages of using Joint Density Plots for the IPT statistics. We had considered the 6-state Gaussian Markov model proposed for modelling the IP telephony traffic statistics based on the PDF and used it to compare its Joint Density result with the measured Joint Density result, and concluded that by modelling the PDF, we are merely modelling the PDF statistics and the model does not capture the other unique important information regarding packet sequences and source behaviour which the Joint density result provides us with. We then also looked at the method of extracting packet sequence information and used it to calculate the number of traffic sources that generated the statistics. We also observed a very interesting and significant result, when we considered the IPT statistical results of two versions of Markov models with two different packet generation processes. We established

that two different models can have the same PDF but completely different Joint Densities, thus signifying the importance of the Joint Density plots. We then also looked at the symmetrical and asymmetrical nature of Joint Density results and demonstrated that the measured Joint Density results of IP traffic were all symmetrical in nature. Finally we also looked into the impact of introducing random packet loss in to the Markov Monte Carlo simulations of models and drew relevant conclusions on its impact on the Joint density results of IPT.

## Understanding and Modelling Joint Density Results – Periodic Processes

In the previous chapter, we had observed the significance of using the Joint Density results and the additional information they provide on the packet sequences of IP traffic, and we also looked at how Gaussian Markov Models can be used to model the IPT Joint Densities, with focus on packet sequences. There is another important unique property related to traffic sources that is observed in the Joint Density results. This is the presence of the *curve of periodicity* on the logarithmic scale, which hints on the presence of a periodic process or a set of periodic processes in the IPT statistics of measured results. Again this is another feature which is observed from the Joint Density Function (JDF) alone, whereas the PDF does not show such type of information. The *curve of periodicity* was briefly mentioned earlier in Chapter 3, while discussing the Joint Density of measured results. Throughout this chapter we will focus on it, understand it, and explore the use of Markov models to model such a curve in the Joint Density.

In section 5.1, we will review some measured IP traffic results, highlight the presence of periodic events, understand the principle behind the *curve of periodicity* and then discuss the possible applications that generate such type of periodic events from the perspective of IPT statistics. In section 5.2, we will review the setup of the Markov Monte Carlo simulation for modelling the *curve of periodicity* in the Joint Density results. Section 5.3 looks in detail at the various parameters of Markov models involved in modelling the *curve of periodicity*. Section 5.4 looks at the impact of additional periodic processes on the IPT Joint Density results and in section 5.5 we will look at two examples of Markov models that can be used to model the two types of periodic processes observed in the measured results. Section 5.6 looks at the nature of simulation of these two models and finally, section 5.7 summarises the chapter.

## 5.1 Elements of Periodicity in the Joint Density Results

In Chapter 3, we had looked at some examples of measured Joint Density results, when referring to the results in section 3.2 for IPT of traffic associated with an IP telephony client measured at the UDP Port 15010. In section 3.2.3, we had briefly indicated the presence of curves, one of the unique elements we observed in the Joint Density results and which we attributed to the bi-directional periodic nature of traffic for the application or protocol running over the network. Fig 5.1 shows the Joint Density result with one of the *curves of periodicity* highlighted. Fig 5.2 is another example of a Joint density result of IPT based on a UDP traffic stream, showing the *curves of periodicity*.

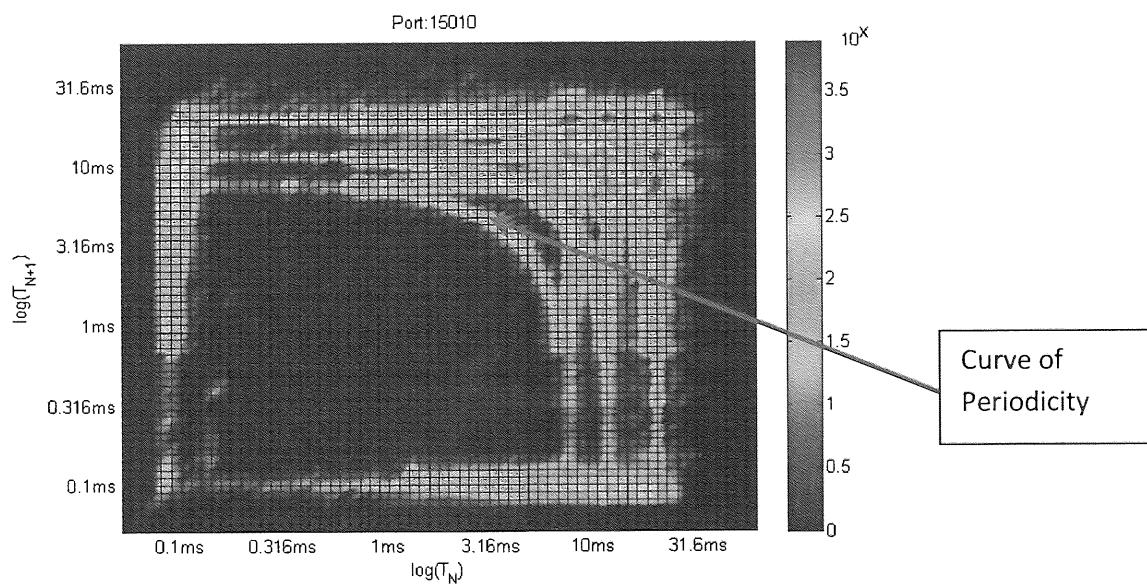


Figure 5.1: Joint Density Result of IPT for IP telephony traffic with a curve of periodicity highlighted

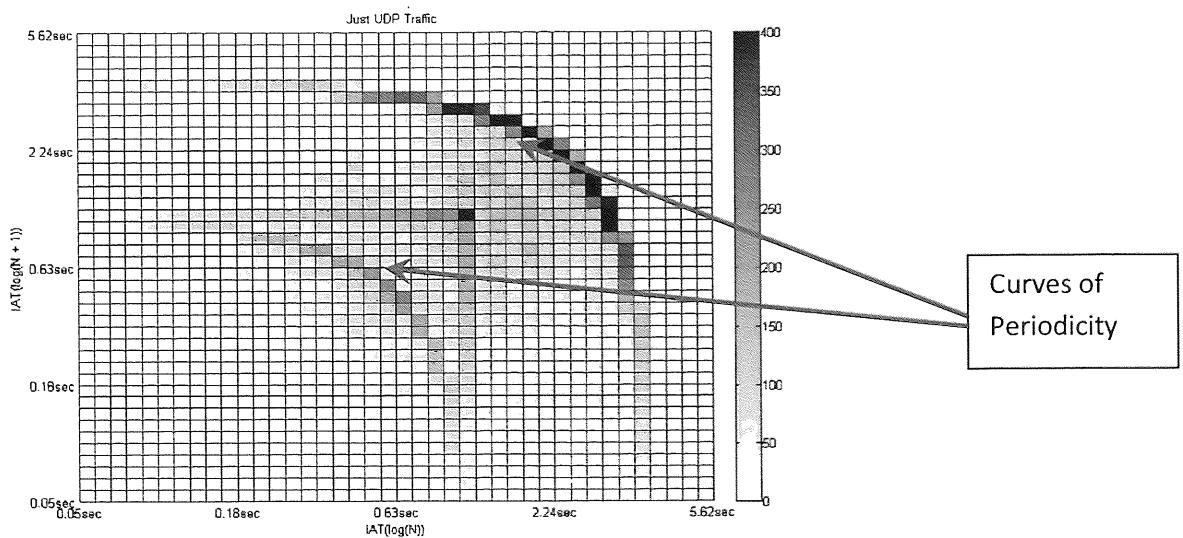


Figure 5.2: Joint Density Result of IPT for a UDP traffic stream with curves of periodicity highlighted.

### 5.1.1 Presence of periodic processes in results associated with real time streaming protocols

These curves of periodicity have been found on numerous occasions involving multimedia traffic, with bi-directional packet streams with packets being sent at regular intervals. This has often been found to be related to traffic associated with real time multimedia and telephony services, which use protocols like Session Initiation Protocol (SIP), Real time transport protocol (RTP) and RTP Control Protocol (RTCP) which are responsible for delivering video and audio content from source to destination in the form of packet streams. We specially observed that these curves of periodicity are exclusive to traffic generated by VoIP and Video Telephony services.

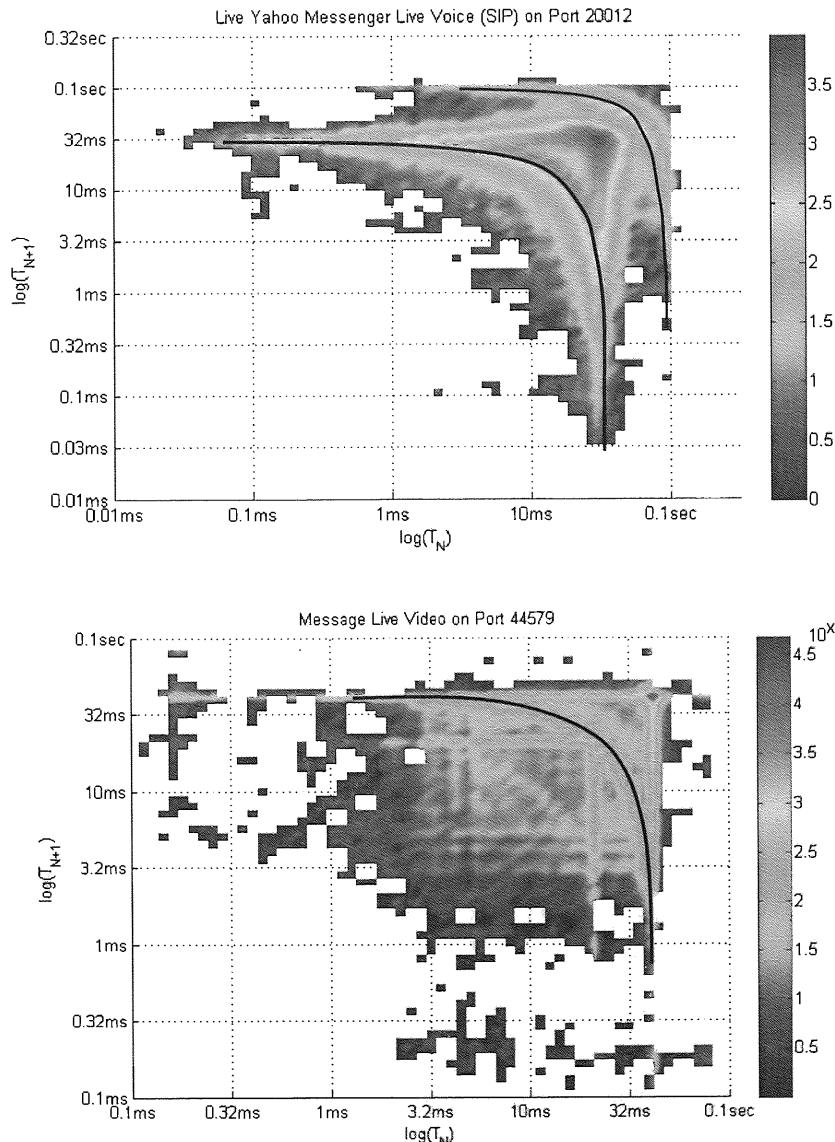


Figure 5.3: Joint Density Results of IPT for traffic of (a) Windows Live Messenger client measured at Port 20012 and (b) Yahoo Messenger client measured at Port 44579, with curve of periodicity indicated by black curves.

These curves have not been found in the Joint Density IPT statistics associated with other types of traffic not using bi-directional streaming, for example, IP TV and IP radio (Fig. 3.4 and 3.6). Fig. 5.3 (a) and (b) show another set of Joint density results for IPT of traffic for Windows Live Messenger and Yahoo Messenger, which are VoIP clients, measured at Ports 20012 and 44579 using the UDP transport protocol. The *curves of periodicity* in the results are clearly indicated both in Fig. 5.3(a) and 5.3(b) by the black curves plotted over the results. In the next section, we will look at the principle behind this *curve of periodicity* in the Joint Density results.

### 5.1.2 Principle behind the presence of the Curve of Periodicity

The curve of periodicity indicates the phenomenon in which an event with time  $t_1$  is always followed by an event with time  $t_2$ . The values of  $t_1$  and  $t_2$  are usually offset by a very small value, which we can label as  $\Delta$ . This value of  $\Delta$  is a finite number. Fig. 5.4 illustrates this phenomenon of the curve of periodicity as observed in the Joint Density results.

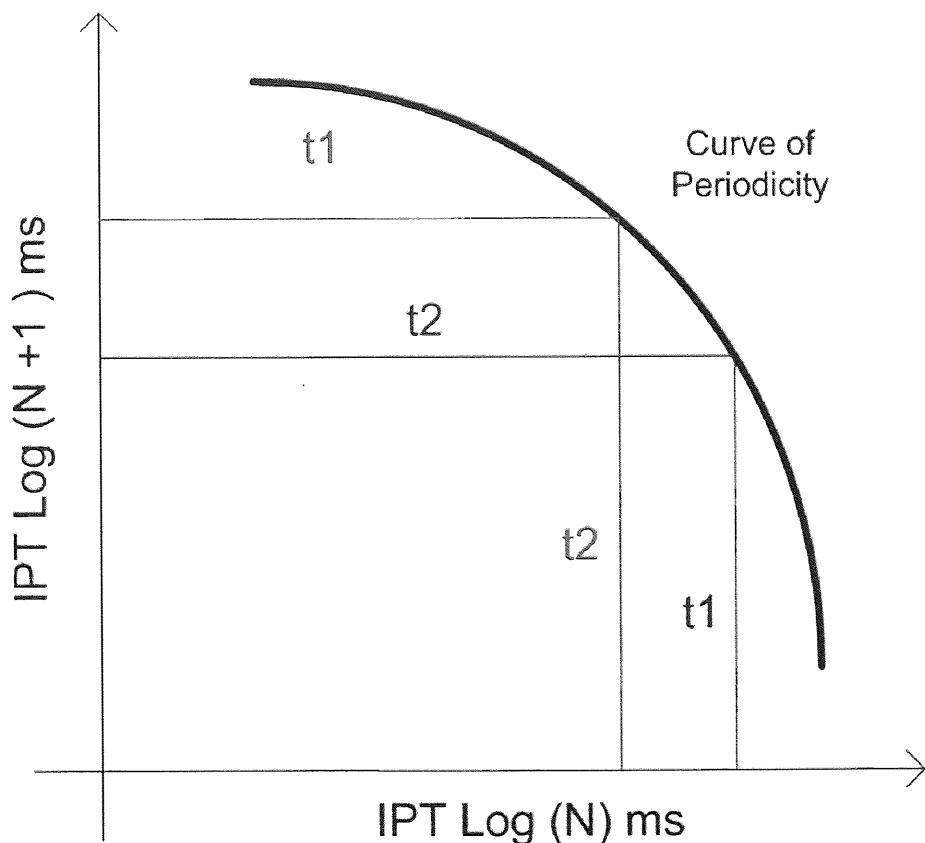


Figure 5.4: Explanation of the Curve of Periodicity in an example of a Joint Density result

When plotted as with all previously shown Joint Density results, the curve shown in Fig 5.4 results from an underlying periodic process with period  $T$  such that  $t_1+t_2 = T$ . The curve is induced by plotting the times logarithmically as we normally have to do because of the large range of timescales involved. A simple example of such a process would be the periodic polling of an email server which responds at a random time. Another way in which such a curve is produced is when two periodic processes, with slightly different periods, are interleaved.

In the real world, this could also be related to a communication link involving two bi-directional periodic streams involving the Real Time Streaming Protocol (RTSP), typically in a scenario of an IP telephony call, where two RTSP or RTCP streams involving codecs of fixed time units operate in both directions with times slightly offset from each other. A second possible scenario is that the machines on two sides of a bi-directional periodic process, log the times involving the same periods, but because the system clock of one of the machines measuring the time stamps of events is out of synchronisation, there is a fixed amount of delay, labelled as  $\Delta$ , which is the time difference between the periods of the two processes.

## 5.2 Using Markov Models to model the periodic elements

Since we now have a strong understanding that the curve in the Joint Density refers to periodic processes, the next step is to look at how we can use Markov models to model the periodic processes.

### 5.2.1 Presence of periodic processes in results associated with real time streaming protocols

The idea of using Markov Models to represent Periodic Processes may seem tricky to some extent, because in its pure sense, the Markov Models follow the Markov Property which states that events of the present and the future are independent of the past events. In other words, they are memory less systems. However, as shown in Fig 5.5, we could represent each of the two periodic sources with two

simple two-state Markov Models ‘C’ and ‘D’ (they have been named C and D to avoid naming conflicts with Markov models used in the previous chapter), that are used as a set, where only States 1 of each Markov Model emit packets A and B respectively. Each model is similar to the two-state model with non-Poisson statistics studied in Chapter 2, with the only exception that the model considered here spends a negligible amount of time in State 2, which can be ignored, and also the state holding time distribution parameters are defined by a Gaussian distribution.

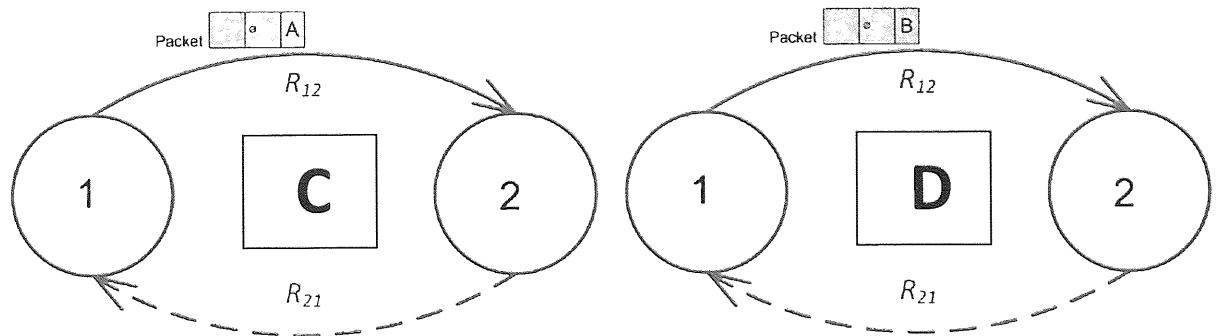


Figure 5.5: Markov Models ‘C’ and ‘D’ for modeling Periodic Processes.

### 5.2.2 Setting up the Markov Monte Carlo Simulation Process for modelling Periodic events

The setup of Markov Monte Carlo simulations for the modelling of Periodic events is slightly different than our earlier approaches as it involves the calculation of absolute time stamps as well. States 1 of Markov Models ‘C’ and ‘D’ can be associated with Gaussian Distributions with means  $\mu_1 = t_1$  (Packet A) and  $\mu_2 = t_2 + \Delta$  (Packet B) respectively, where  $\Delta$  can be small or large as compared to the period T. The third key parameter  $\sigma$ , the common standard deviation of the two Gaussian distributions, can also be varied according to need. The two Markov Models ‘C’ and ‘D’ emit packets A and B at absolute times  $ABTS_C$  and  $ABTS_D$  respectively. To set the simulation with correct Inter Packet times (IPT) for the two periodic processes, the maximum absolute arrival time ( $ABTS_D$ ) of the Markov Model ‘D’ is limited to the maximum absolute arrival time ( $ABTS_C$ ) of Markov Model ‘C’ (shown in Fig. 5.6). The absolute times  $ABTS_C$  and  $ABTS_D$  are merged and finally the Inter-packet times are calculated and the Probability density and Joint density results are plotted as before.

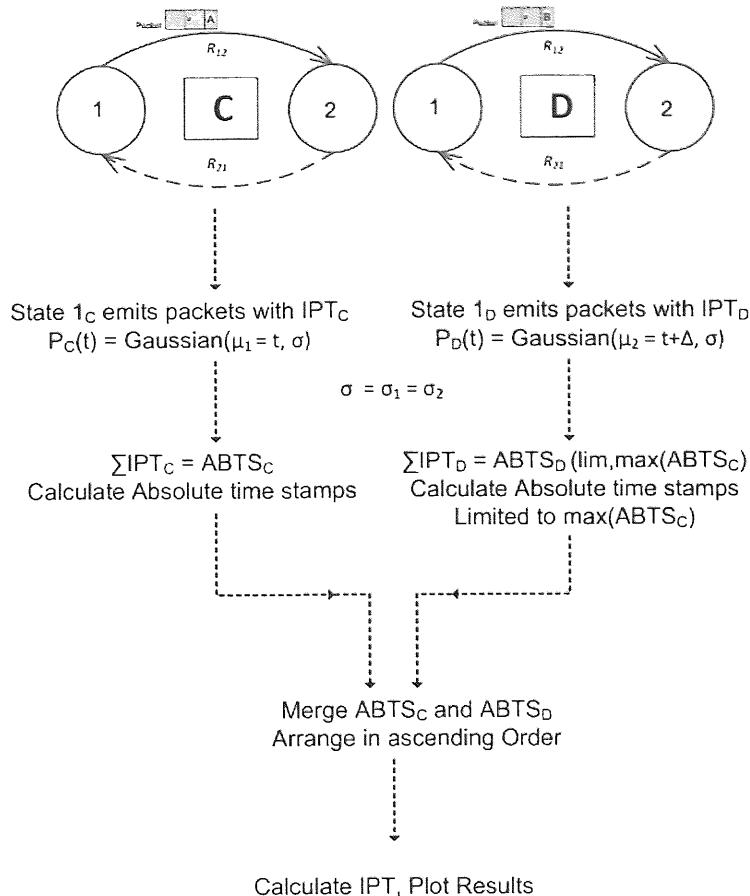


Figure 5.6: Markov Monte Carlo Setup for Modelling Periodic Events

Fig. 5.6 illustrates the Monte Carlo process for two Markov Models 'C' and 'D' modelling two periodic processes using models 'C' and 'D' emitting packets 'A' and 'B' respectively. If additional periodic processes are required to be simulated, a chain of the required number of additional two state Markov models representing each periodic process is added to the simulation and their absolute times are merged together and then the inter-packet times are calculated as described earlier.

### 5.3 Analysis of results of Monte Carlo Simulation of Markov models representing periodic processes

In this section we will look at the results of the Markov Monte Carlo simulation for a variety of cases where we change the three key parameters associated with the two Markov Models in consideration i.e.  $(\mu_1, \mu_2, \sigma)$  and  $\Delta$ .

### 5.3.1 Case 1: Small values of $\mu$ , with varying $\Delta$ and fixed $\sigma$

In the first case of analysis, we will consider small values of  $\mu_1$  and  $\mu_2$  which are closer to the origin, fixing the standard deviation  $\sigma = \sigma_1 = \sigma_2$  to 0.645 for all our initial analysis. We first consider lower values of  $\Delta$ , and observe the curve of periodicity, and then increase the value of  $\Delta$  and notice its impact.

**Case 1.a:  $\mu_1 = 2.5$ ,  $\mu_2 = 2.7$ ,  $\sigma = 0.645$  and  $\Delta = 0.2$**

Setting the values of  $\mu_1 = 2.5$ ,  $\mu_2 = 2.7$  for the Markov models C and D respectively, we run the Markov Monte Carlo simulation and obtain the following PDF and JDF results:

Let us first look at the PDF of the IPTs for this case as shown in Fig. 5.7.

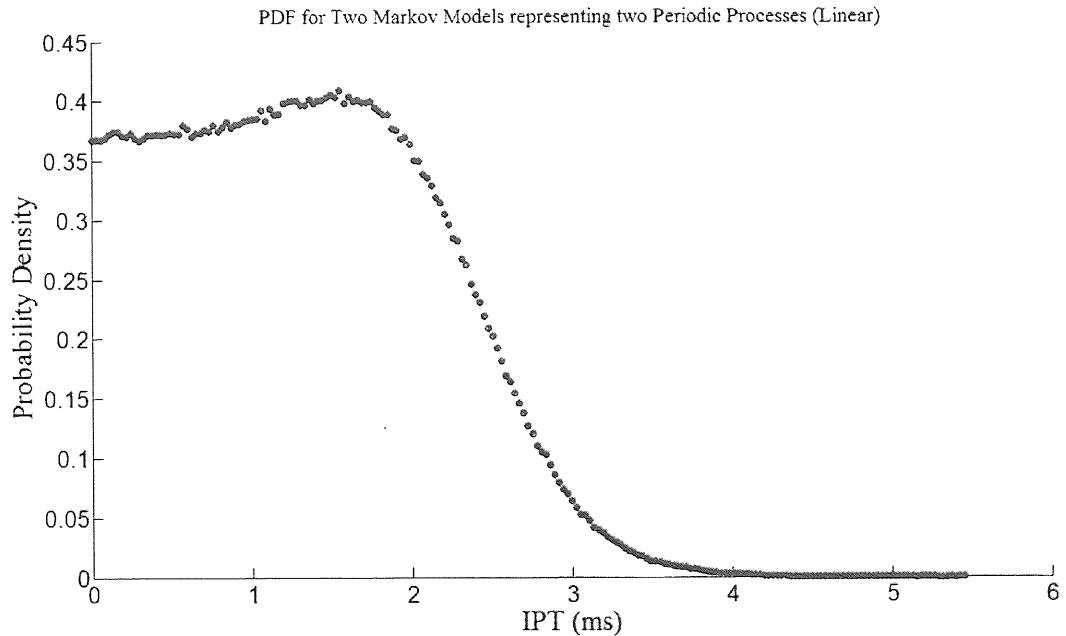


Figure 5.7: PDF for IPT, of  $\mu_1 = 2.5$ ,  $\mu_2 = 2.7$ ,  $\sigma = 0.645$  and  $\Delta = 0.2$

The PDF shows the presence of IPTs in the range 0-3 ms but gives no information about any underlying periodicity. So we will move on to the Joint density result as shown in Fig. 5.8, which shows a rim like structure on the linear scale. Notice how when viewed from any of the axes in Fig. 5.8, one can see the projected PDF shown in Fig. 5.7.

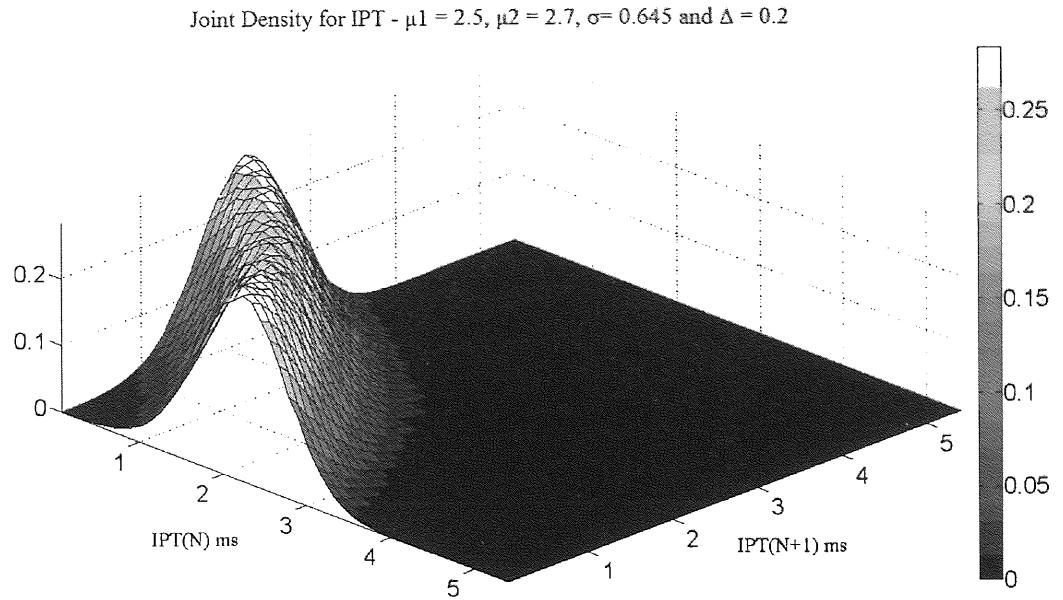


Figure 5.8: JDF (Linear version) for IPT, of  $\mu_1 = 2.5, \mu_2 = 2.7, \sigma = 0.645$  and  $\Delta = 0.2$

From the Joint Density result (Fig. 5.8) in the three dimensional form (linear scales), we can see a rim at times of  $\mu_1 = 2.5$ . The curve of periodicity seen in the two dimensional Joint Density plot shown in Fig. 5.9, represents the top view of the rim structure that we see in Fig. 5.8, but on logarithmic scales.

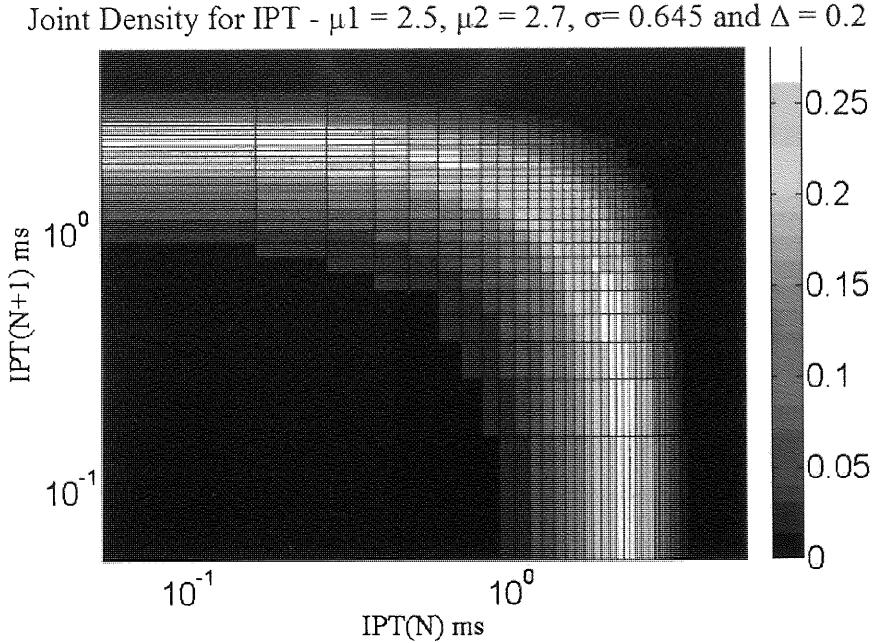


Figure 5.9: JDF (Logarithmic version) for IPT, of  $\mu_1 = 2.5, \mu_2 = 2.7, \sigma = 0.645$  and  $\Delta = 0.2$

Fig. 5.9 clearly shows that time  $t_1$  of packet is always followed by time  $t_2 + \Delta$  of another packet in the network and with this result we have demonstrated that we can achieve the curve of periodicity with the Markov Monte Carlo simulation of the two Markov models 'C' and 'D' respectively. It is useful to

note that the result when  $\Delta = 0$ , looks very similar to the result above. For the next case, we will increase the value of  $\Delta$  by 1.

### Case 1.b: $\mu_1 = 2.5$ , $\mu_2 = 3.7$ , $\sigma = 0.645$ and $\Delta = 1.2$

The PDF and JDF results for the increase in  $\Delta$  by 1 are shown in Fig. 5.10 and 5.11 respectively.

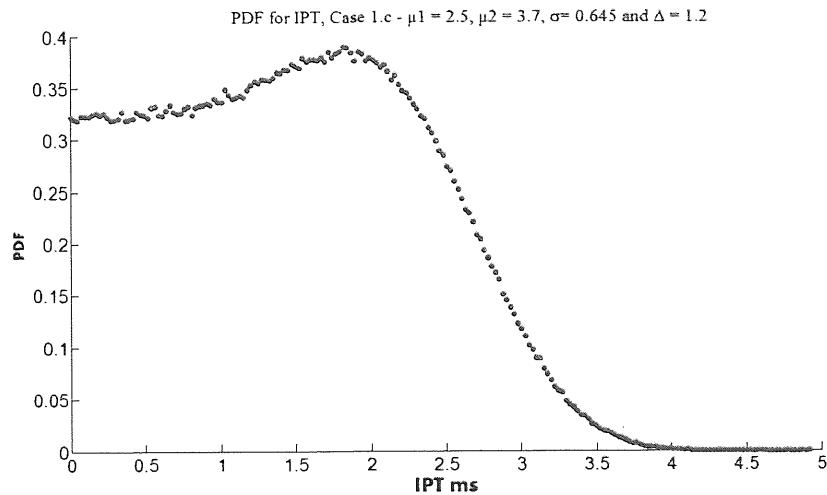


Figure 5.10: PDF for IPT, of  $\mu_1 = 2.5, \mu_2 = 3.7, \sigma = 0.645$  and  $\Delta = 1.2$

JDF for IPT, Case 1.b:  $\mu_1 = 2.5, \mu_2 = 3.7, \sigma = 0.645$  and  $\Delta = 1.2$

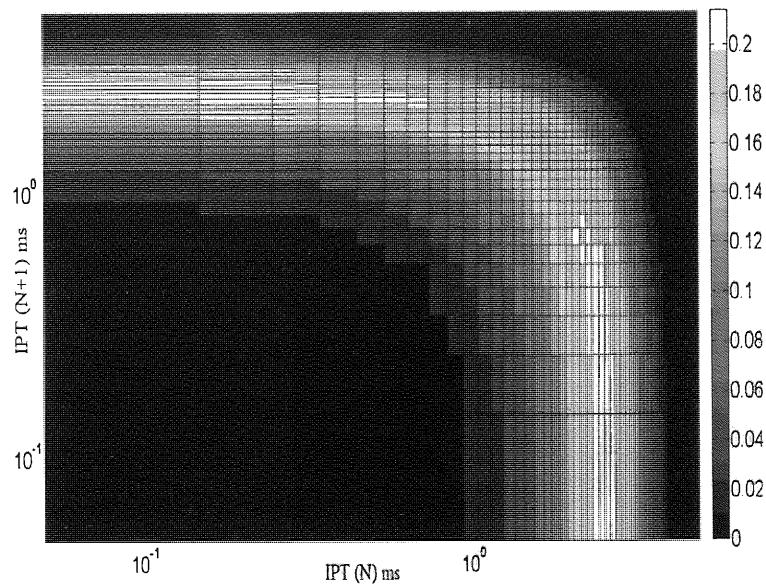


Figure 5.11: JDF for IPT, of  $\mu_1 = 2.5, \mu_2 = 3.7, \sigma = 0.645$  and  $\Delta = 1.2$

From these figures, we see very minor differences in the result as the curve largely remains intact from the previous result. However, we see that the peak at 2.5 ms begins to broaden out as we increase ' $\Delta$ ', as will be evident from further results. In the next case we will increase the  $\Delta$  to 7.2.

**Case 1.c:  $\mu_1 = 2.5$ ,  $\mu_2 = 9.7$ ,  $\sigma = 0.645$  and  $\Delta = 7.2$** 

The PDF and JDF results for IPT for  $\Delta = 7.2$  are as shown in Fig 5.12 and 5.13 respectively.

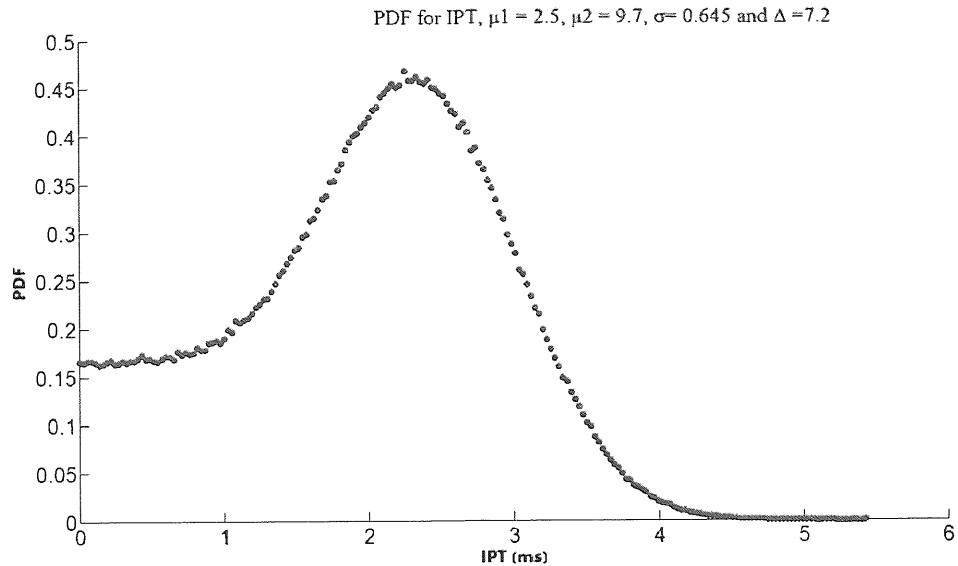


Figure 5.12: PDF for IPT, of  $\mu_1 = 2.5$ ,  $\mu_2 = 9.7$ ,  $\sigma = 0.645$  and  $\Delta = 7.2$

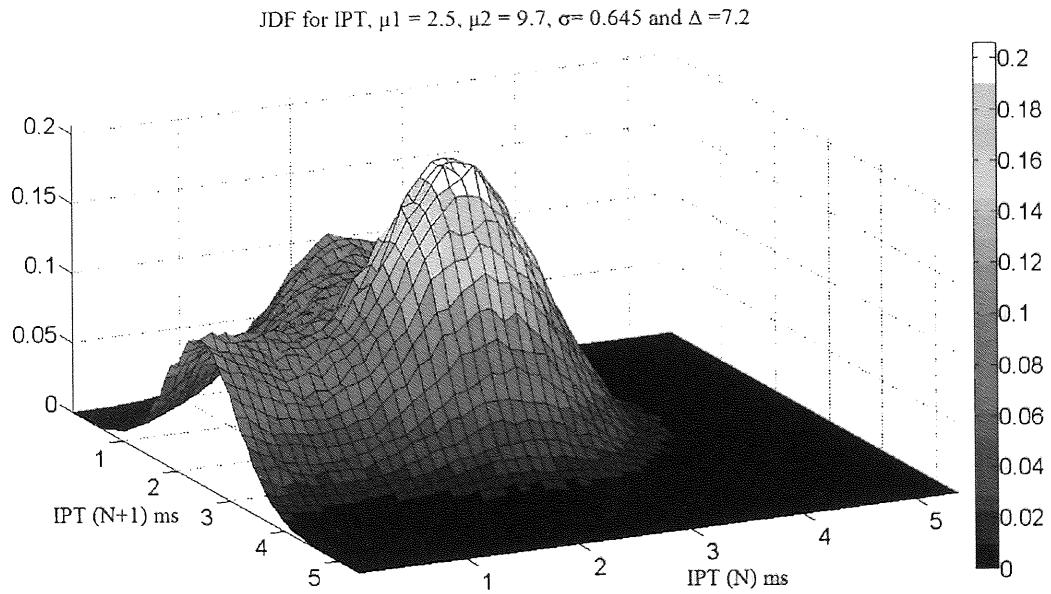


Figure 5.13: JDF (Three dimensional Linear) for IPT, of  $\mu_1 = 2.5$ ,  $\mu_2 = 9.7$ ,  $\sigma = 0.645$  and  $\Delta = 7.2$

We see that the peak at times 2.5 ms vs. 2.5 ms moves out further out of the curved rim surface as the value of  $\Delta$  is increased to 7.2.

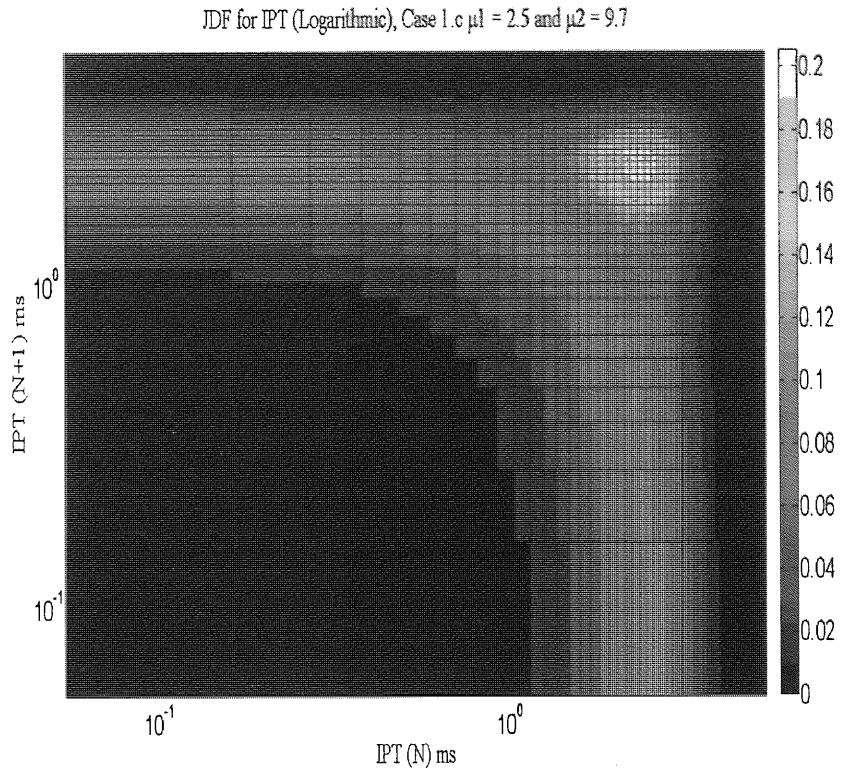


Figure 5.14: JDF (Logarithmic) IPT, of  $\mu_1 = 2.5$ ,  $\mu_2 = 9.7$ ,  $\sigma = 0.645$  and  $\Delta = 7.2$

From these three trials, we have learnt that as the value of  $\Delta$  is increased we see a peak protruding out of the curve as witnessed from the figures of this section so far.

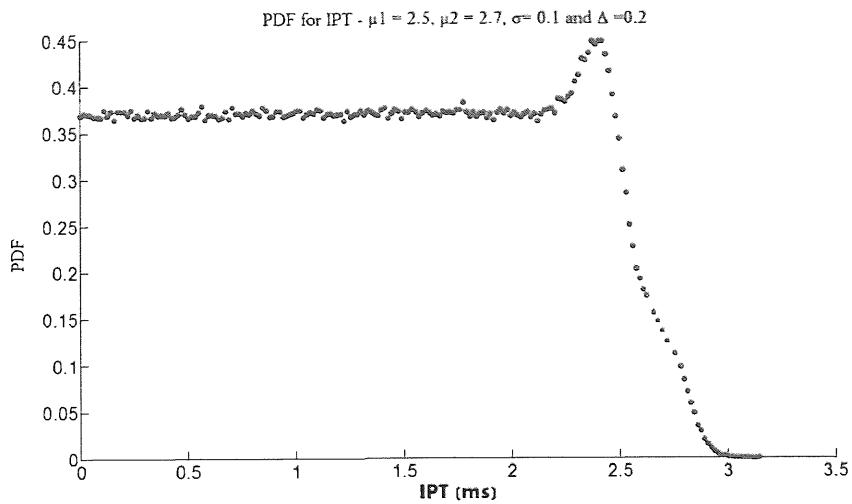
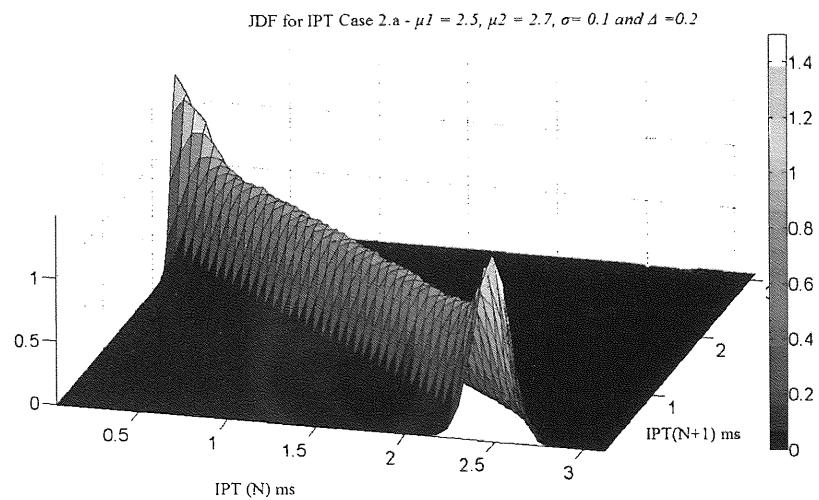
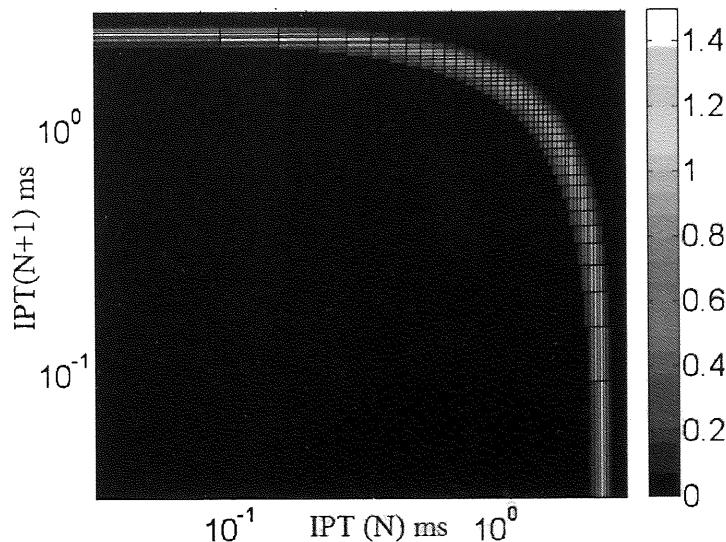
### 5.3.2 Case 2: Impact of Varying $\sigma$ of Models ‘C’ and ‘D’

Next, we will study the impact of varying the value of  $\sigma$ , the common standard deviation parameter of the two Gaussian distributions associated with the two periodic processes.

**Case 2.a:  $\mu_1 = 2.5$ ,  $\mu_2 = 2.7$ ,  $\sigma = 0.1$  and  $\Delta = 0.2$**

We will once again consider the model parameters used in Case 1.c, with the exception of the value of  $\sigma$ . For this case 2.a, we will decrease the value of  $\sigma$  from 0.645 to 0.1.

Fig 5.15, 5.16 and 5.17 show the Monte Carlo simulated results for the PDF, Linear JDF (three dimensional) and Logarithmic JDF (Two dimensional) results respectively.

Figure 5.15: PDF for IPT, of  $\mu_1 = 2.5, \mu_2 = 2.7, \sigma = 0.1$  and  $\Delta = 0.2$ Figure 5.16: JDF (Linear X and Y axes, three dimensional) for IPT, of  $\mu_1 = 2.5, \mu_2 = 2.7, \sigma = 0.1$  and  $\Delta = 0.2$ Figure 5.17: JDF (Logarithmic X and Y axes, two dimensional) for IPT, of  $\mu_1 = 2.5, \mu_2 = 2.7, \sigma = 0.1, \Delta = 0.2$

We observe from Fig. 5.16, that the curvature of the rim structure has broadened and as seen in Fig. 5.17 we also find that the curve's thickness has reduced as a result of lowering the  $\sigma$  value.

### **Case 2.b: $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 0.01$ and $\Delta = 0.2$**

In this next example, we assign a very low value of  $\sigma$  keeping the other parameters of the model intact. Executing Markov Monte Carlo simulation for these parameters again we obtain the Joint Density result as shown in Fig. 5.18. The plot has been plotted uniquely on a white plane to show the presence of the thin curve of periodicity (not visible in conventional plots) obtaining a better perspective of the result.

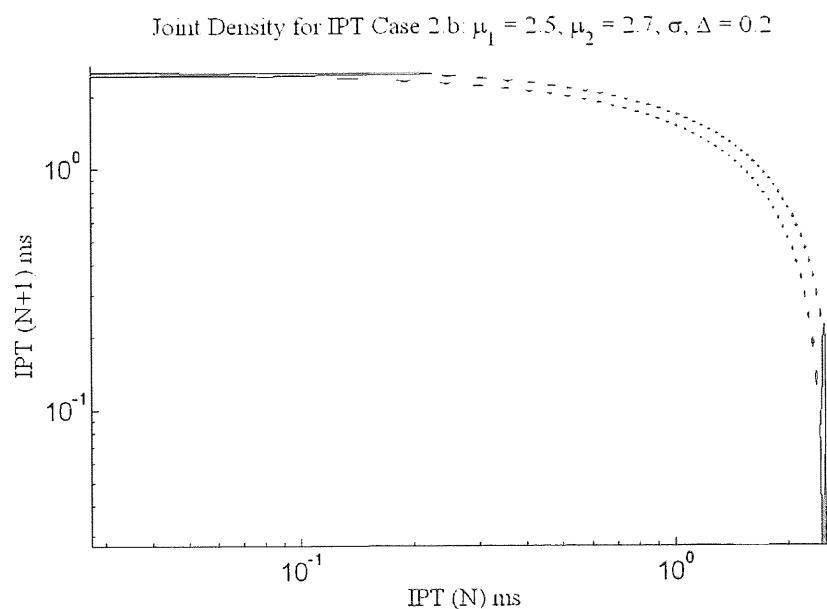


Figure 5.18: JDF (Logarithmic X and Y axes, two dimensional) for IPT, of  $\mu_1 = 2.5$ ,  $\mu_2 = 2.7$ ,  $\sigma = 0.01$ ,  $\Delta = 0.2$

We learn from these results so far, that the lower the value of  $\sigma$  for the two Gaussian distributions, the narrower the width of the curve of periodicity. Therefore we now know that the width of the curve can be controlled by the  $\sigma$  of the two associated distributions.

### **Case 2.c: $\mu_1 = 2.5$ , $\mu_2 = 2.7$ , $\sigma = 1$ and $\Delta = 0.2$**

For the final case of varying  $\sigma$ , we increase the  $\sigma$  to a fairly large value. The results for this case would of course be Poisson-like as seen in Fig. 5.19. This is because, as  $\sigma$  is very large, the distributions associated with the periodic process almost completely overlap each other leading to Poisson like statistics.

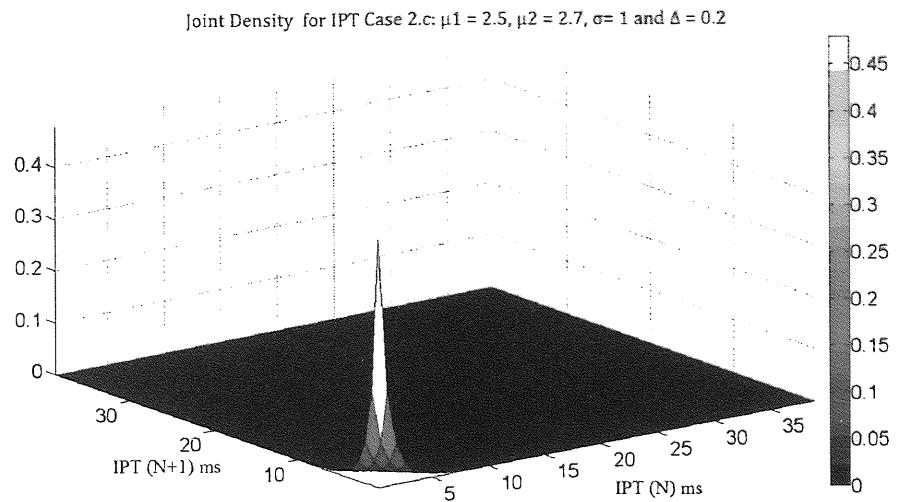


Figure 5.19: Joint Density Function (Three dimensional) – Poisson like statistics for large  $\sigma$

### 5.3.3 Case 3: Choosing $\mu_1$ and $\mu_2$ far away from the origin

Now taking our analysis with these Markov Models further, we will now look at the impact of pushing the means  $\mu_1$  and  $\mu_2$  further away from the origin along with varying the value of  $\Delta$ .

**Case 3.a:  $\mu_1 = 9.5$ ,  $\mu_2 = 9.7$ ,  $\sigma = 0.645$  and  $\Delta = 0.2$**

Let us consider an example similar to that of Case 1.a, but with larger values of  $\mu$  as detailed above.

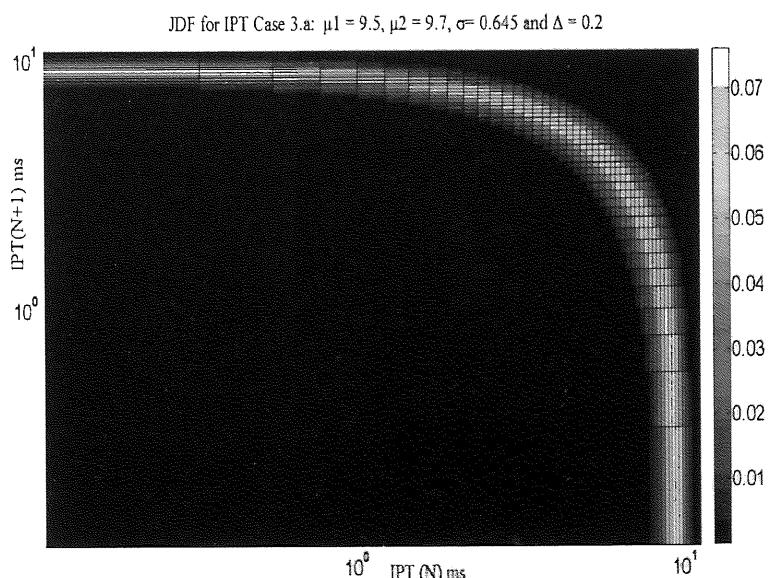


Figure 5.20: JDF for IPT, of  $\mu_1 = 9.5$ ,  $\mu_2 = 9.7$ ,  $\sigma = 0.645$  and  $\Delta = 0.2$

We observe from Fig. 5.20, that the results look very similar to that of case 1a where  $\Delta$  was small. So, we will now look at increasing  $\Delta$  to a larger value.

### **Case 3.b: $\mu_1 = 9.5$ , $\mu_2 = 19.7$ , $\sigma = 0.645$ and $\Delta = 10.2$**

With this increase in the value of  $\Delta$ , as seen in Fig. 5.21, we begin to see an additional peak forming out of the curve as was also demonstrated in section 5.3.1. The curve structure also has protruding arms, culminating into an additional peak which has now moved further away from the curve.

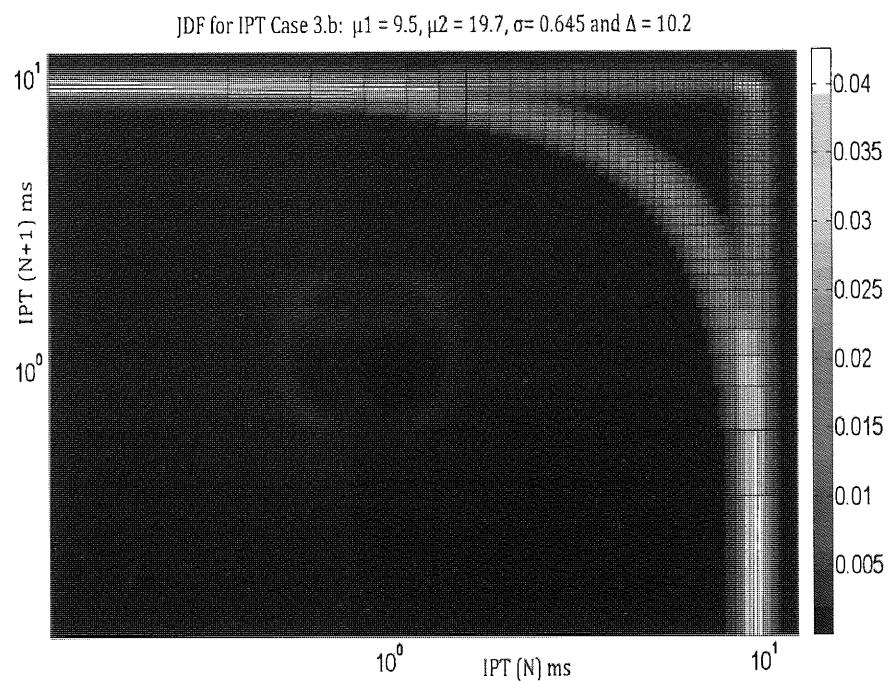


Figure 5.21: JDF for IPT, of  $\mu_1 = 9.5$ ,  $\mu_2 = 19.7$ ,  $\sigma = 0.645$  and  $\Delta = 10.2$

### **Case 3.c: $\mu_1 = 20.5$ , $\mu_2 = 100.7$ , $\sigma = 0.645$ and $\Delta = 80.2$**

By further increasing  $\Delta$  for the above mentioned large values of  $\mu_1$  and  $\mu_2$  in the Markov Monte Carlo simulation of models C and D, we get the following Joint Density result as shown in Fig. 5.22.

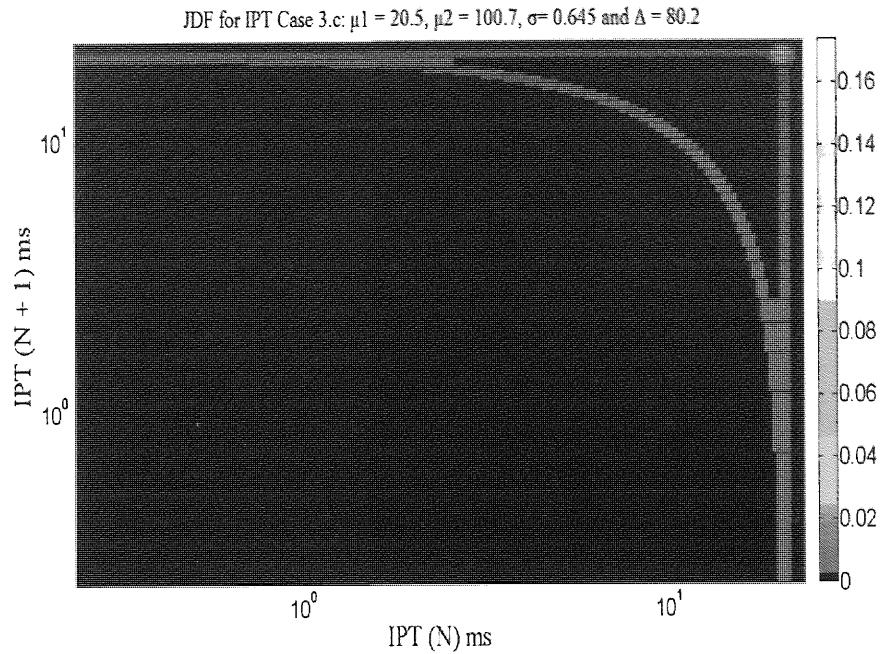


Figure 5.22: JDF for IPT, of  $\mu_1 = 20.5$ ,  $\mu_2 = 100.7$ ,  $\sigma = 0.645$  and  $\Delta = 80.2$

From the plot, we observe that with a larger  $\Delta$  value the curve reduces in density growing thinner, while the peak which moves yet further away from the curve, connected by the protruding tangent arms, increases in density. This kind of result is very similar to the features we observed in our measured results as shall be confirmed in the section 5.5. The PDF of IPT for this case is shown in Fig.5.23.

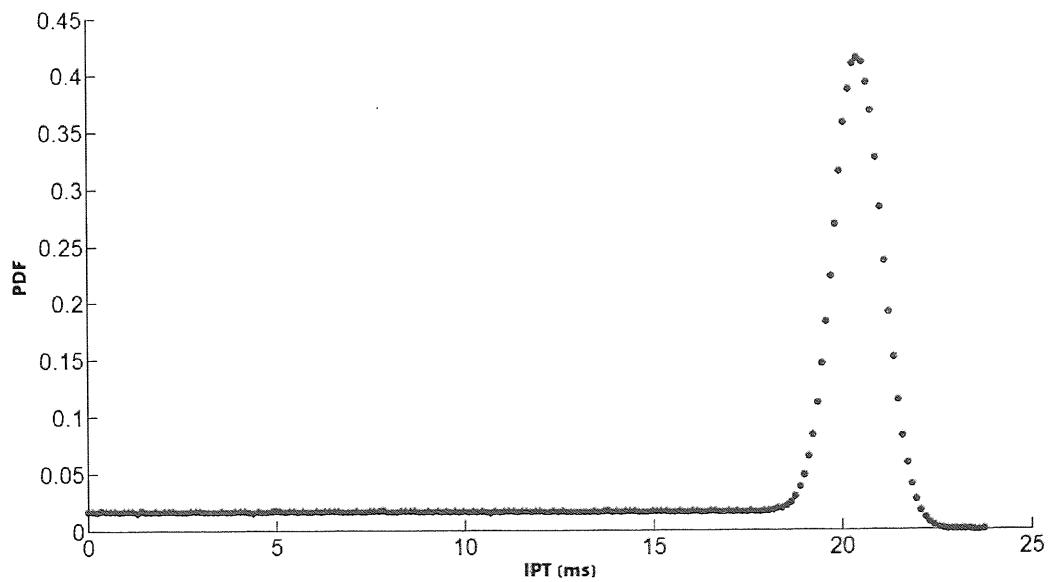


Figure 5.23: PDF for IPT, of  $\mu_1 = 20.5$ ,  $\mu_2 = 100.7$ ,  $\sigma = 0.645$  and  $\Delta = 80.2$

### 5.3.4 Summary of Analysis of Results of Monte Carlo simulations

To summarize our observations made in all the cases of Markov Monte Carlo simulations covered in this section, we have demonstrated that the curve representing two periodic processes can be modelled by two of the two state Markov Models ‘C’ and ‘D’ with only one state of each emitting packets with Gaussian distributions, the means of which differ by a value of  $\Delta$ . The general conclusion from the results is that two Markov models with limited or no difference between their mean value (low value of  $\Delta$ ), can clearly produce the periodicity curve we seek to model. As the value of  $\Delta$  is increased, we start to see a peak at times  $t_2$  which moves out of the curve. At larger values of  $\Delta$ , this peak is connected with perpendicular tangent arms reaching the curve as observed in cases 3.b, and 3.c. This structure has been found in various measured plots and is of significant importance again as shall be revealed in the section 5.5. As  $\Delta$  increases, the curve diminishes in density and the peak has a higher density. In addition to all of this, we know that varying the common standard deviation  $\sigma$ , by a marginally small value, controls the width of the curve, whereas a large value of  $\sigma$ , results in an overlap of the two distributions leading to Poisson statistics.

## 5.4 Confirming the presence of only two periodic processes associated with a curve of periodicity

Throughout the previous sections, dealing with analysis of results, we have considered and assumed the presence of only two periodic processes. In this section, we will evaluate the presence of additional processes that can be associated with the curve of periodicity observed in the Joint Density result, by introducing an additional number of periodic processes in the Markov Monte Carlo simulation process explained in section 5.2.2. We will consider two test cases where we add one and two additional periodic processes to the Markov Monte Carlo simulation respectively.

### 5.4.1 Test Case 1: Introducing one additional periodic process to the Monte Carlo Simulation.

To the previous approach as highlighted in section 5.2.2, we will add one more periodic process by including an additional Markov Model E, similar to models C and D, with mean  $\mu_3$  for its Gaussian distribution. For this first test case we will consider that the Markov Models emits packets with  $\mu_1 = 2.5$ ,  $\mu_2 = 3.0$  and  $\mu_3 = 3.5$ , hence  $\Delta = 0.5$ , and the common standard deviation  $\sigma$  is 0.645 as before. Looking at the PDF for IPT for this case, it clearly shows Poisson statistics (Fig 5.24):

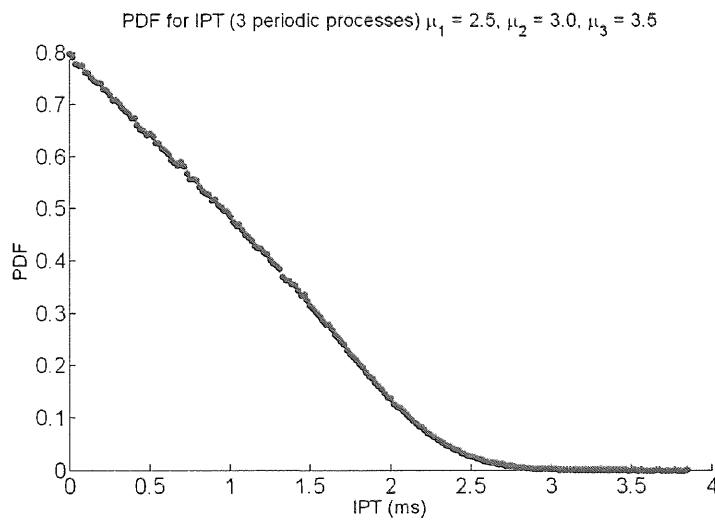


Figure 5.24: PDF for 3 periodic Processes with  $\mu_1 = 2.5$ ,  $\mu_2 = 3.0$ ,  $\mu_3 = 3.5$ ,  $\sigma = 0.645$  and  $\Delta = 0.5$

This PDF is a projection of the Joint Density structure shown in Fig. 5.25.

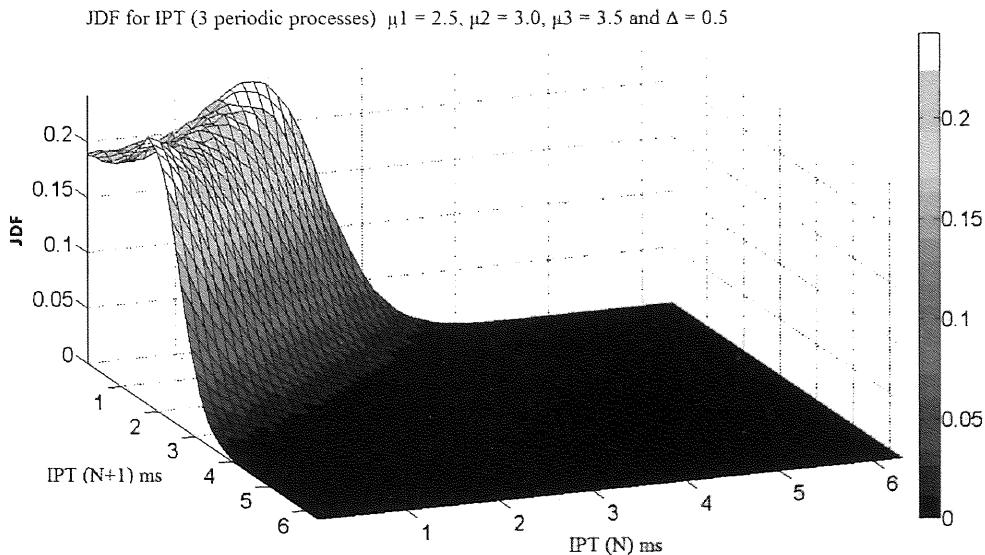


Figure 5.25: JDF (3d Linear) for 3 periodic Processes,  $\mu_1 = 2.5$ ,  $\mu_2 = 3.0$  and  $\mu_3 = 3.5$ ,  $\Delta = 0.5$

We now look at the Joint Density results on the logarithmic X and Y Axes shown in Fig. 5.26, and we see that the structure associated with periodicity is not curve but rather a symmetrical solid structure.

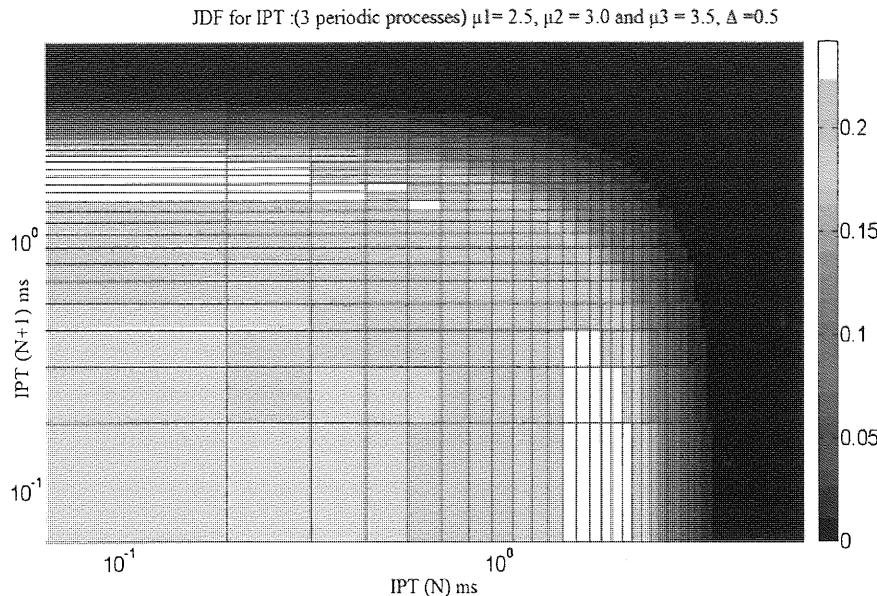


Figure 5.26: JDF (2d Logarithmic) for 3 periodic Processes,  $\mu_1 = 2.5$ ,  $\mu_2 = 3.0$  and  $\mu_3 = 3.5$ ,  $\Delta = 0.5$

It is clear from the above Joint Density plots, that with the addition of a third periodic process, the curve statistics are reduced to Poisson statistics with the inner area of the curve filled (solid surface), to look like a quadrant of a hemisphere.

### 5.4.2 Test Case 2: Including two additional periodic processes to the Monte Carlo Simulation

In this second test case, we will include two additional periodic processes. Four Markov models will emit packets in the Markov Monte Carlo simulation, with the following parameters  $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5$ , such that  $\Delta = 0$ , and  $\sigma = 0.645$  as before. For a change, we have considered the same value of means for all the four Markov models and hence  $\Delta = 0$ . The PDF for IPT for the four periodic processes is as shown in Fig. 5.27. We observe similar Poisson statistics again, with the difference that the PDF curve is now linear because all the four periodic processes have the same mean parameters.

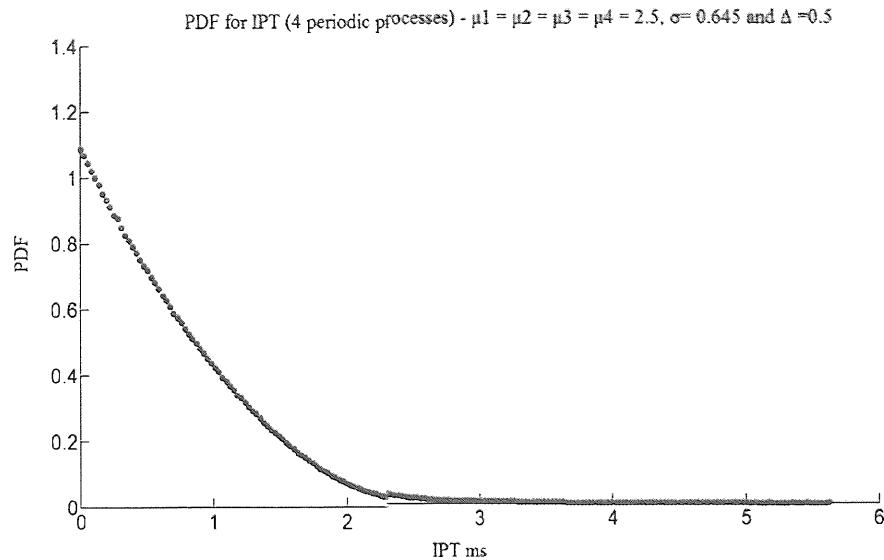


Figure 5.27: PDF for 4 periodic Processes with  $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5, \sigma = 0.645$  and  $\Delta = 0.5$

With the Joint Density result shown in Fig 5.28, we observe a structure that arises from a two dimensional Poisson process.

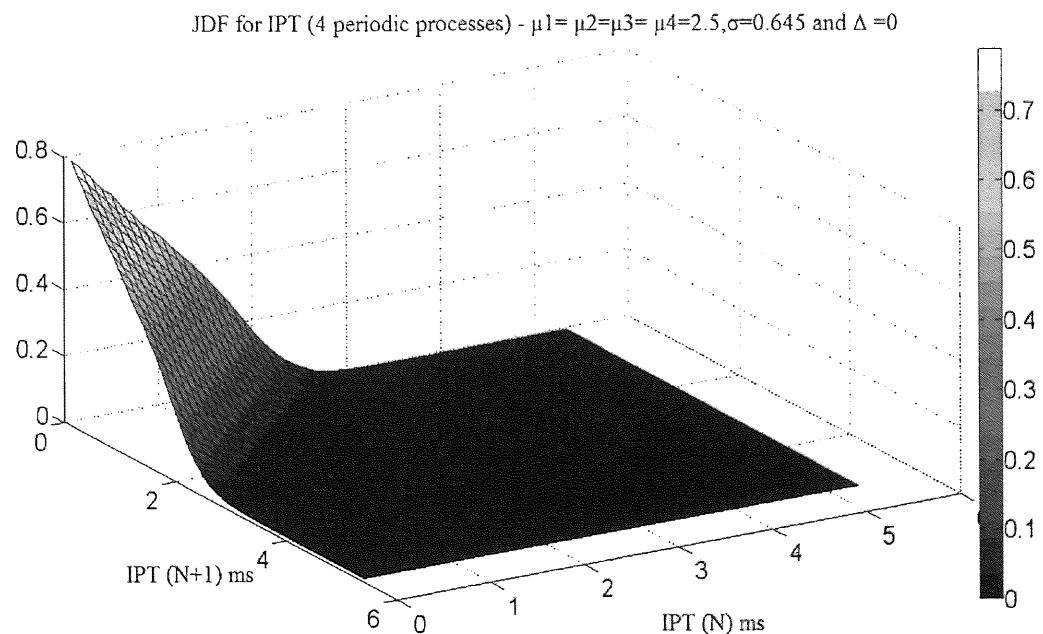


Figure 5.28: JDF (Linear scale) for 4 periodic Processes with  $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5, \sigma = 0.645$  and  $\Delta = 0.5$

As with the previous case, plotting the results on the logarithmic X and Y axes as shown in Fig. 5.29, reveals a solid surface and not the curve of periodicity observed in the case of presence of only two periodic processes.

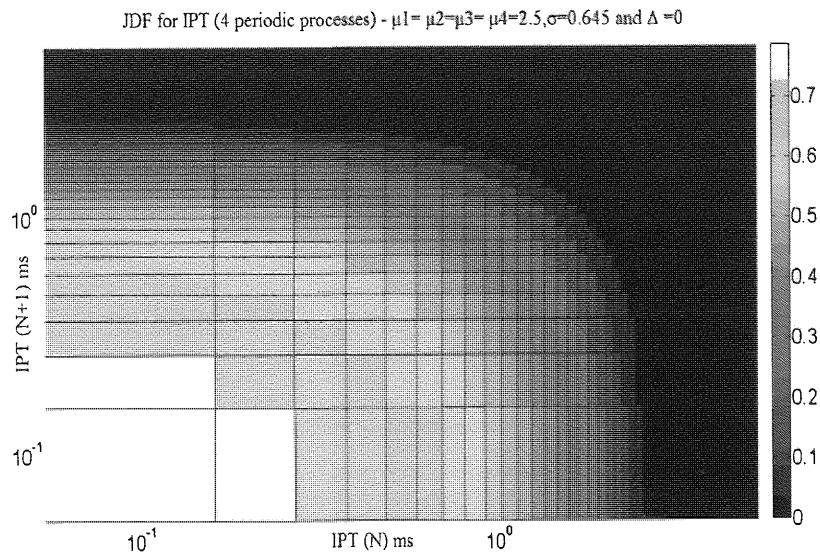


Figure 5.29: JDF (2d Logarithmic) for 4 periodic Processes,  $\mu_1 = \mu_2 = \mu_3 = \mu_4 = 2.5$

### 5.4.3 Conclusion on Test results

The results shown in section 5.4.1 and 5.4.2 confirm analytically and numerically, that there can be only two periodic processes which represent the curve indicating the presence of periodic elements in the Joint Density, as clearly the addition of one or more periodic processes, changes the dynamics of the results and reduces the results to Poisson statistics as observed.

The fact that the Joint Density results (on logarithmic X and Y axes) for the two cases shown in Fig. 5.26 and 5.29 show Poisson statistics can be confirmed by comparing these results to that of the purely Poisson Joint density result (Fig. 2.6) of the symmetric two-state Markov model considered in section 2.2.1.1 of Chapter 2. Fig. 2.6 showed a cut-section of the Poisson Joint Density result with linear X and Y axes. If that result (Fig.2.6) were to be plotted on a logarithmic X and Y axes, with a linear density, the result would look like as shown in Fig. 5.30.

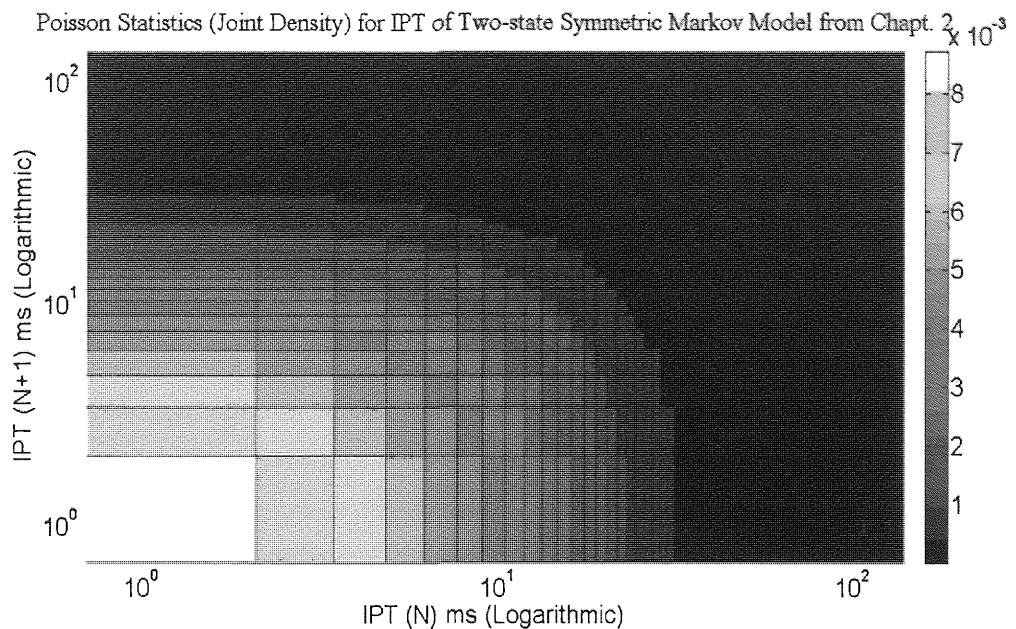


Figure 5.30: Poisson statistics (Joint density) for IPT of Two-state Symmetric Markov Model from Chapter 2 (Fig.2.6)

Comparing the purely Poisson result of the two-state symmetric Markov model shown in Fig. 5.30, with the results shown in Figs. 5.26 and 5.29 respectively, we observe that they look similar in structure, hence confirming Poisson statistics for cases which feature more than two periodic processes in the Monte-Carlo simulation. This also establishes that the curve of periodicity occurs only when two periodic processes are involved.

## 5.5 Relating the modelled Joint Density results to the measured Joint Density results

Finally, we come to the most important part of this chapter, where we use the results of models ‘C’ and ‘D’ to relate to the features observed in a few of the measured results which indicated the presence of periodic elements as shown in the earlier sections of this chapter. To proceed further, we will consider the two unique types of observed results of IPT Joint Density, which consist of the result with small  $\Delta$  giving us a curve (as in Fig. 5.17), and the result with large value of  $\Delta$ , giving us a peak with perpendicular tangent arms to the curve (as in Fig. 5.21). We will use these results and designate them as types ‘ $S\Delta$ ’ and ‘ $L\Delta$ ’, as shown in fig. 5.31 (a) and (b) respectively, to use them to indicate similar unique features and elements in the measured Joint Density IPT results.

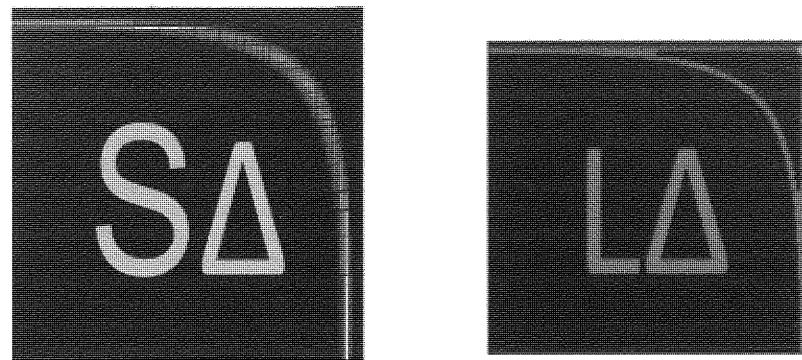


Figure 5.31: (a) The  $S\Delta$  Model and (b) The  $L\Delta$  Model

### 5.5.1 Example 1: UDP Streaming Traffic Results

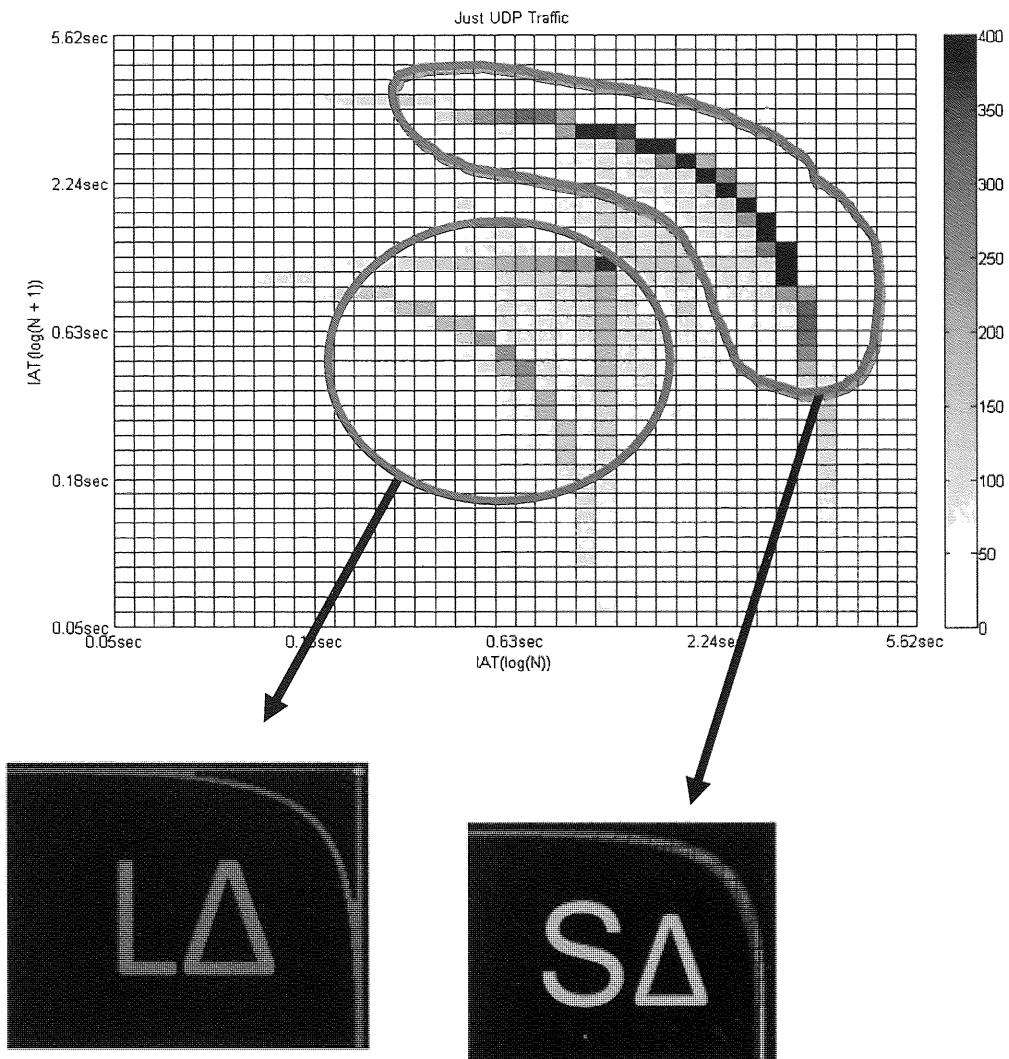


Figure 5.32: Example Result showing how two statistical features of the Joint Density of IPT have been captured by the Markov Monte Carlo simulation of the  $L\Delta$  and the  $S\Delta$  models

The example result in Fig 5.32 shows how two sets of periodic processes represented within the measured Joint Density results recalled from section 5.1 can be modelled by the two sets of Markov Models. The feature highlighted by the ‘ $S\Delta$ ’ model can be modelled using two periodic processes with a large  $\Delta$  and the feature highlighted by the ‘ $L\Delta$ ’ model can be modelled using two periodic processes with a very small  $\Delta$ . This confirms that the above statistical result of UDP traffic, demonstrates the presence of two sets of periodic processes, with large and small  $\Delta$  respectively.

### 5.5.2 Example 2: Voice/Video streaming over IM clients

We will now look at two more examples based on traffic from Voice and Video telephony conversation from Windows Live Messenger and Yahoo Messenger Instant Messaging Clients. Please refer Fig. 5.3 (a) and 5.3 (b) for the actual Joint Density measured results.

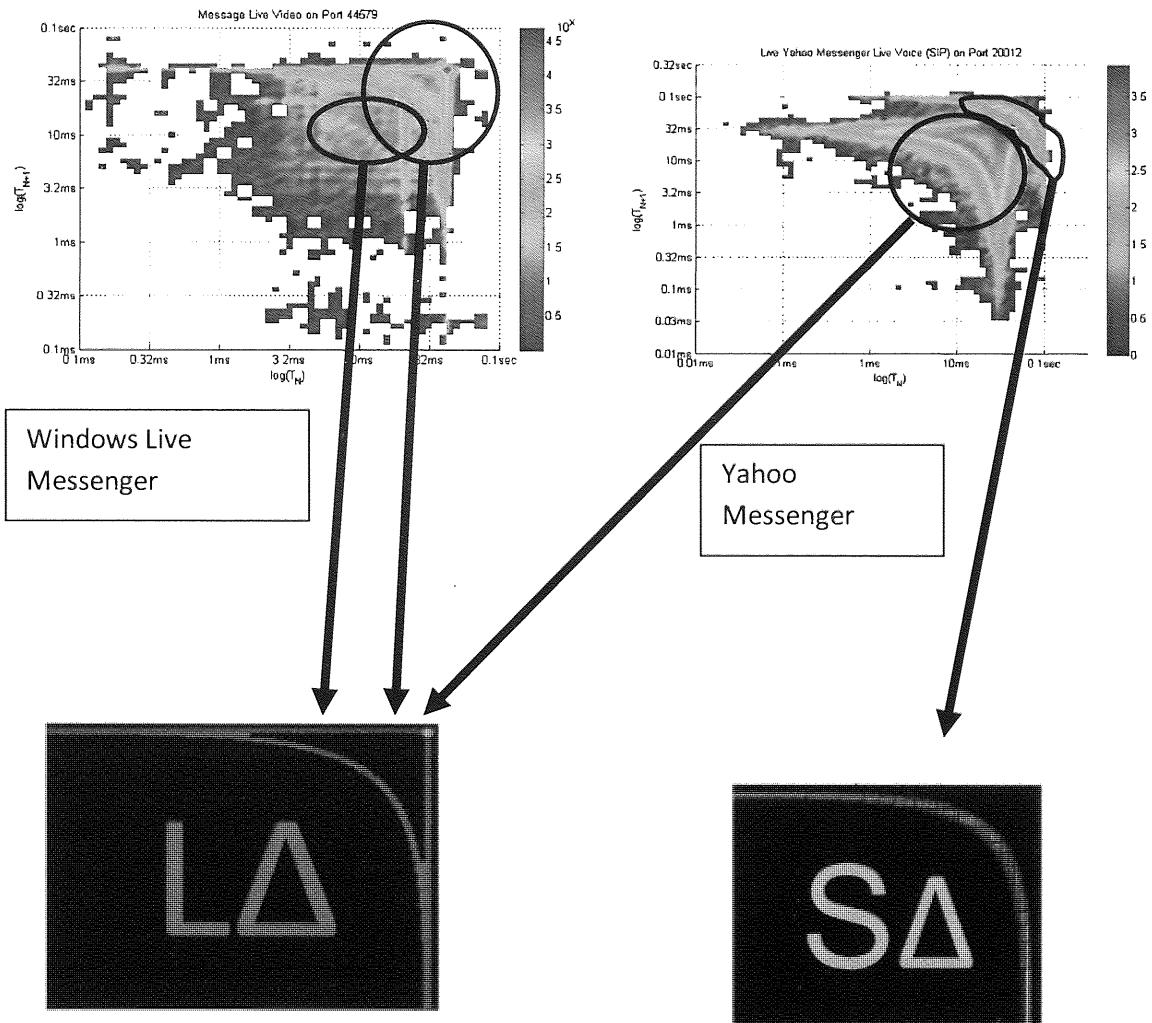


Figure 5.33: Example Results of Joint Density statistics of traffic on Windows Live Messenger and Yahoo Messenger IM clients showing how their statistical features have been captured by the Markov Model sets ‘ $L\Delta$ ’ and ‘ $S\Delta$ ’ representing periodic processes with large and small  $\Delta$  respectively.

The example results in Fig. 5.33 show through the aid of the Markov Model sets ‘ $L\Delta$ ’ and ‘ $S\Delta$ ’, that the Joint Density result of IPT of the Windows live Messenger client has two similar sets of periodic processes both of which have a large  $\Delta$  (‘ $L\Delta$ ’). Whereas, the Joint Density result of IPT of the Yahoo Messenger Client shows elements of two different sets of periodic processes, one with large  $\Delta$  (‘ $L\Delta$ ’) and the other one with small  $\Delta$  respectively (‘ $S\Delta$ ’).

## 5.6 Phased nature of IP Traffic Transmission

In the previous section through the examples shown in Fig. 5.32 and 5.33, we have demonstrated how the Joint Density results featuring the curve of periodicity can be modelled with a combination of the ‘ $L\Delta$ ’ and the ‘ $S\Delta$ ’ models respectively. However, it is important to note that the Joint Density result contains two sets of periodic processes at two different sets of times, and in order for such a result to be simulated the Monte-Carlo simulation featuring a combination of the ‘ $L\Delta$ ’ and ‘ $S\Delta$ ’ models should be executed in such a way such that the individual models are never ‘On’ at the same time. In other words, the generation of IPT samples from the two models can only be in phases. If the results were generated simultaneously from both the models, it will result in a structure different to the one being modelled. This important property implies that the traffic being modelled occurs in two phases, where the first phase of the traffic is described by characteristics of one model and the second phase of the traffic is described by the characteristics of the second model.

This implication can be demonstrated by considering the measured Joint Density result of UDP traffic shown in Fig. 5.2. Plotting the Joint Density results of the IPT for only the first 30% of the captured packets, results in Fig. 5.33.

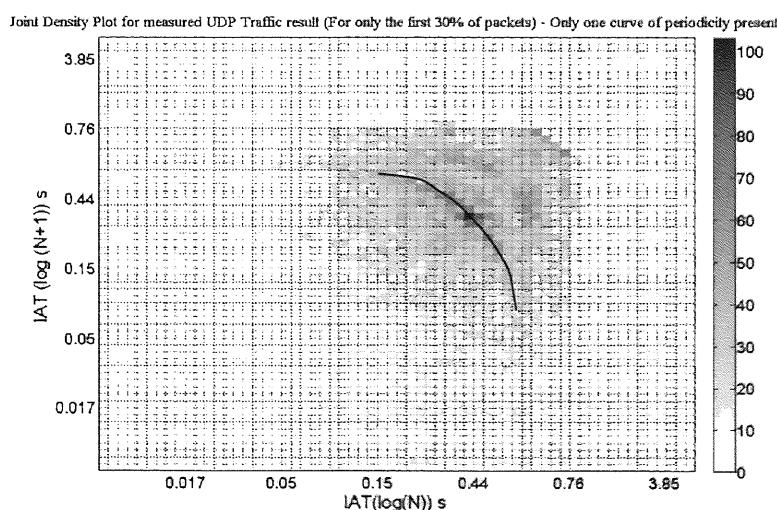


Figure 5.34: Joint Density Plot of UDP Traffic for only the first 30% of the captured packets

The Joint Density result for the first 30% of the captured IP packets reveals the presence of only the bottom curve of periodicity (along with slightly visible arms) from Fig. 5.2. This is highlighted by a black curve over Fig. 5.34. This observation demonstrates the fact that the traffic transmission, for which the statistics were plotted, occurred in phases. Comparing Fig. 5.2 and Fig. 5.34, we find that the first 30% of the captured packets characterise only one curve of periodicity and there is no sign of the other periodic process, as clearly stressed by a comparison of the time values on the IPT axes of both the plots. It can be observed in Fig. 5.34, that the curve of periodicity in concern, has not completely formed (in density structure) for the first 30% of the packets, which implies that the phases of traffic transmission can occur in any form of sequences which add up coherently to the complete Joint Density result for all the captured packets. Thus, using the Joint density results and Markov models, we have not only demonstrated the periodic nature of traffic but also the phased nature of IP traffic transmission as proved by the experimental test on a UDP traffic capture.

## 5.7 Summary

This chapter began with a detailed insight into the study of a unique feature of the measured Joint density results of IPT, which revealed the presence of periodic processes, represented by a curve of periodicity that occurs at logarithmic scales of the X and Y axes. We have also seen some examples of the measured Joint Density results featuring the curves of periodicity followed by a section which explained the setup of the proposed set of 2-state Markov models which modelled two periodic processes. We have seen the detailed procedure for setting up the Markov Monte Carlo simulation for generating IPT samples based on the two periodic processes. A detailed analysis on the Markov Monte Carlo simulation results were also covered which included assessing the impact on the Joint Density results with changes in the various parameters associated with the Gaussian distributions of the two models. Further, upon tests with additional periodic processes, it was proved that the curve of periodicity can result from a pair of periodic processes only. Finally, we singled out two types of unique results from the analysis, one with a small difference in mean, which we designated as ‘ $S\Delta$ ’ and the other with a large difference in the means for the periodic processes, which we designated as ‘ $L\Delta$ ’, and showed how these two sets of models can be used to model all the measured Joint Density results which showed signs of periodicity of packet events. We have also used these models to demonstrate the phased nature of traffic transmission. With this chapter, our work on the IP traffic modelling with the aid of Markov Monte Carlo simulations and Probability and Joint density plots concludes and the next two chapters of this thesis deal with the analysis of ON and OFF period characteristics of Voice traffic over IP networks.

## **PART – II**

### **Analysing and modelling the ON-OFF characteristics of VoIP traffic**

# Chapter Six

## Analysis of Voice over IP traffic from a multi-lingual VoIP call database

The approach used for modelling voice traffic over IP networks is slightly different to the approaches used for conventional IP traffic modelling, as it involves the use of models which are based on conversational speech. Voice packets based on VoIP traffic using *silence suppression*, are generated only during the talk spurts of the conversation, which correspond to the ON periods of a VoIP traffic source. No packets are generated during the OFF periods of the traffic source, corresponding to the silence durations of conversational speech. This chapter looks in detail, at the characteristics and analysis of voice traffic over IP networks, based on the statistics of ON and OFF periods of a voice traffic source. These statistics are based on the measurements of ON and OFF periods taken from a significantly large voice call database consisting of native speakers of both genders speaking in some of the world's most widely spoken languages. The chapter first starts with a brief background to voice traffic modelling in section 6.1, where we also look at some related work carried out on the subject. Section 6.2 discusses the silence detection process used to measure the ON and OFF periods of voice traffic from the voice call database. Section 6.3 covers the statistical analysis of the ON and OFF periods and also investigates the impact of differences in the linguistic and prosodic structure of speech and speaker behaviour on the statistics, drawing relevant conclusions. Finally, section 6.4 summarises the work covered in this chapter.

### 6.1 Background on Voice Traffic Modelling

The recent explosive growth of Voice over IP (VoIP) solutions has triggered a phenomenal increase in the amount of voice traffic carried over the IP network. Unlike the traditional circuit-switched telephony system, the packet-switched VoIP technology has many advantages to offer such as cost efficiency, portability and flexibility, thus attracting a large number of users and gradually reducing the share of the traditional circuit-switched telephony system in the market. However, VoIP does have a few disadvantages with respect to the IP network. VoIP traffic is very sensitive to network

end-to-end delay, jitter and packet loss. Moreover, the best effort nature of the IP network does not provide a guarantee for the reliable quality of a VoIP call. These potential issues call for an improvement in the design and performance of the IP networks and technology that carry the voice traffic. This can be done with the aid of efficient voice traffic models that characterise the voice traffic flows as accurately as possible.

Voice traffic has often and traditionally been modelled with the use of ON-OFF models which were briefly described in Chapter 1. This is because, a VoIP conversation can be modelled as an alternating process of periods of voice burst or *talk spurts* separated by gaps or *silence lengths*. These talk spurts and silence lengths can be represented as  $T_{ON}$  and  $T_{OFF}$  periods corresponding to the ON and OFF periods respectively for a voice traffic source. The source starts emitting voice packets during the  $T_{ON}$  periods and stops emitting voice packets during the  $T_{OFF}$  periods. This process of turning off the traffic sources during periods of silence is called *silence suppression*. The majority of voice codecs used in VoIP solutions today support silence suppression. The key advantage of using silence suppression is that, it reduces the bandwidth used on the network, by only transmitting voice packets when necessary. Thus the  $T_{OFF}$  periods of a voice traffic source, corresponding to the silence lengths in the conversation of a VoIP call, are extremely significant and it is important that the OFF period characteristics of the voice traffic model are captured as accurately as possible.

Several ON-OFF models for voice traffic have been proposed in the past [72]. The conventional and the most prominent of these models assumed an exponential distribution for the holding times for the ON and OFF states [73] [74] [75] [76] [77] [78] [79] [80] [81]. In other words, they assumed that the talk spurts and silence lengths were exponentially distributed. One such widely referenced model is the 8 - state model, used in our simulation work for comparison purposes, by Stern et al [79] [80], shown in Fig. 6.1. This model is an intelligent modification of the original 6-state ON-OFF model proposed by P.T.Brady in his pioneering work based on conversational speech [74] [75]. Stern et al. [79], defined three types of silence lengths with relevance to the OFF periods of a voice traffic source: a) *Listening or Long silence lengths*: Silences which occur when listening, b) *Medium silence lengths*: Silences which occur in between sentences or long phrases and c) *Short silence lengths*: These are silences which occur in between the syllables of a word when the user is speaking and are below 200 ms in duration. For example, the small gap of silence that can be observed between the letters ‘s’ and ‘t’ when one pronounces the word ‘disaster’ is an example of a short silence length.

Stern et al. [79] established the significance of these critical short silence lengths in their work and suggested that by including these short silence gaps (< 200 ms) which other models, including [74], had ignored, the total duration of  $T_{OFF}$  periods could be increased significantly, thus bringing an increase in bandwidth efficiency of the network links carrying the voice traffic. The study of these short silence lengths was also supported by the work of [81], who measured silence lengths below 200 ms. However, the results of the conventional models are based on old and outdated measurements of talk spurts and silence lengths, from calls of very limited duration which were recorded with the aid of analogue recording equipment. In addition, the calls were based on English speakers only. This provides a good scope for the use of a new multi-lingual database based on packetized voice, to explore the statistics of the  $T_{ON}$  and  $T_{OFF}$  periods, using an efficient silence detection mechanism.

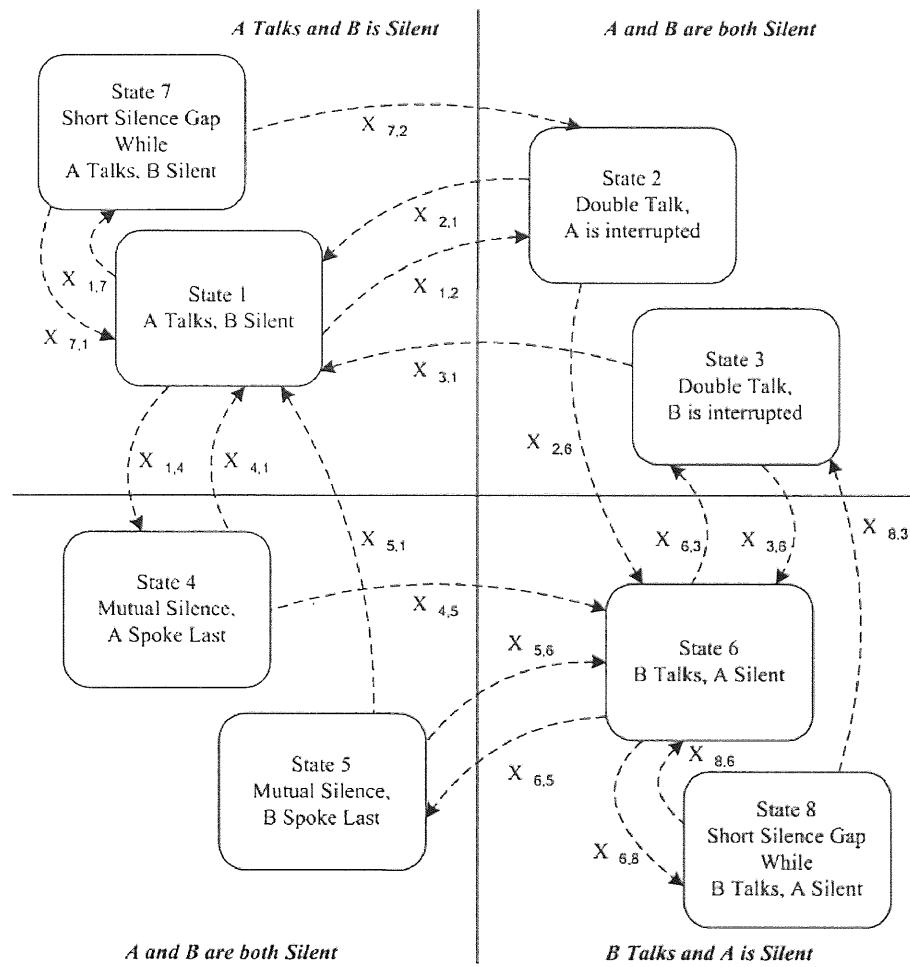


Figure 6.1: The 8-state ON-OFF model proposed by Stern et al.

As pointed out earlier, the vast majority of the voice traffic models used an exponential distribution or purely Markov processes for their models. However, recent work on packetized voice traffic has suggested that the  $T_{ON}$  and  $T_{OFF}$  periods actually follow a heavy tailed distribution instead of an

exponential distribution [82] [83] [84] [85]. While each of these models proposed a unique heavy tailed distribution for the  $T_{ON}$  and  $T_{OFF}$  periods from simulated VoIP traffic flows, they completely ignored the short  $T_{OFF}$  periods ( $< 200$  ms), the significance of which Stern et al. [79] had established in their work. Thus, they observed much longer  $T_{ON}$  and  $T_{OFF}$  periods on average [85]. In our work, as detailed in this chapter, we present new and accurate measurements of the  $T_{ON}$  and  $T_{OFF}$  periods of VoIP conversational activity from a very large multilingual database of VoIP calls and monolog speech recordings, of sufficiently long durations so as to achieve good quality statistical information. We then investigate the impact of the language and speaker behaviour on the  $T_{ON}$  and  $T_{OFF}$  periods before proposing our own ON-OFF model in Chapter 7.

## 6.2 Measuring the $T_{ON}$ and $T_{OFF}$ periods for VoIP traffic

### 6.2.1 The VoIP Call database

For the measurement of the  $T_{ON}$  and  $T_{OFF}$  periods of VoIP traffic, we considered a large database of two-way voice call recordings retrieved from the ‘Callfriend’ section of TalkBank.org [86]. This database consisted of over 150 hours of voice call recordings involving both male and female native speakers of six of the world’s most widely spoken languages which include Mandarin Chinese, Spanish, Japanese, German, French and English. Each of the call samples typically lasted from 30 to 50 minutes and were based on the G.711 codec with a sampling frequency of 8 KHz and a bit rate of 64 kb/s. Apart from linguistic differences, the call samples featured a varying degree of random dynamics of conversation and presence of non-grammatical silences as one would observe in a typical telephone conversation. In addition to the large VoIP call database, a large selection of multilingual monolog data for the above mentioned languages was also considered for the purpose of analysis of short silence lengths, based on the monolog bursts of a speaker in a VoIP call. This monolog data was retrieved from the Open Speech Repository [87].

### 6.2.2 The Silence Detection Process

#### 6.2.2.1 Extracting the speech signal

The first step involved in the silence detection process was to split both sides of the conversation in the stereo PCM based VoIP call recording file to get two separate mono-channel PCM files containing

the individual talk and silence sequences of the two speakers A and B involved in the conversation, respectively. This was achieved by a process of left and right audio panning of the stereo file with the aid of the popular audio editing software, *Audacity* [88]. The speech signal  $S(t)$  is then extracted from the monaural speech file using an appropriate tool such as *Sox* [89]. This is done for the entire number of speech samples in the file and these values when plotted against time gives a speech wave pattern as shown in Fig. 6.2.

It is useful to note that for the silence detection process, we only consider the ‘conversation part’ for statistical analysis and exclude the ‘ringing part’ of the call recording as described in [85]. Fig. 6.1 shows a small extract of a random speech signal  $S(t)$  for a monolog burst of a caller conversing in English. Here we clearly see periods of talk activity separated by periods of silence when the speaker is listening. Accordingly, the  $T_{ON}$  and  $T_{OFF}$  periods with respect to a voice source model have also been marked approximately in the figure.

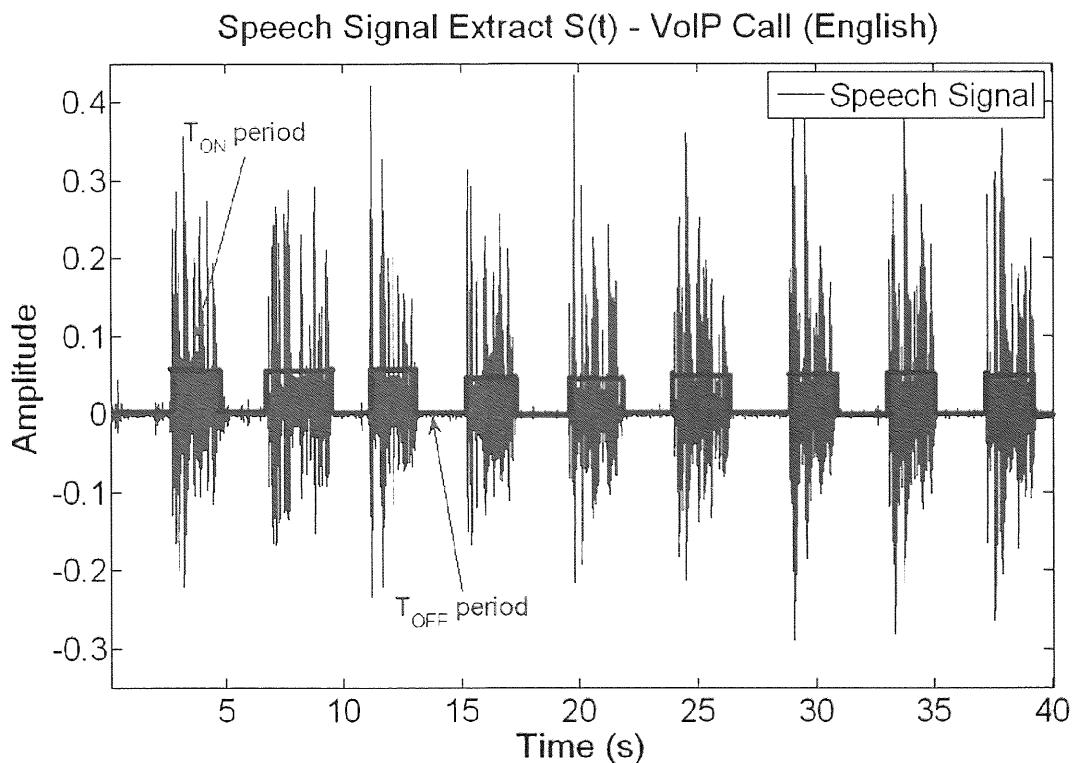


Figure 6.2: Example of waveform of extract of a speech signal  $S(t)$  with the  $T_{ON}$  and  $T_{OFF}$  periods shown

When a very small section of the speech signal is expanded, say for example, up to 100 ms accuracy, one can visually observe the small  $T_{OFF}$  periods which occur between words and syllables, as shown in Fig. 6.3, where we notice a small  $T_{OFF}$  period of 125 ms between two words.

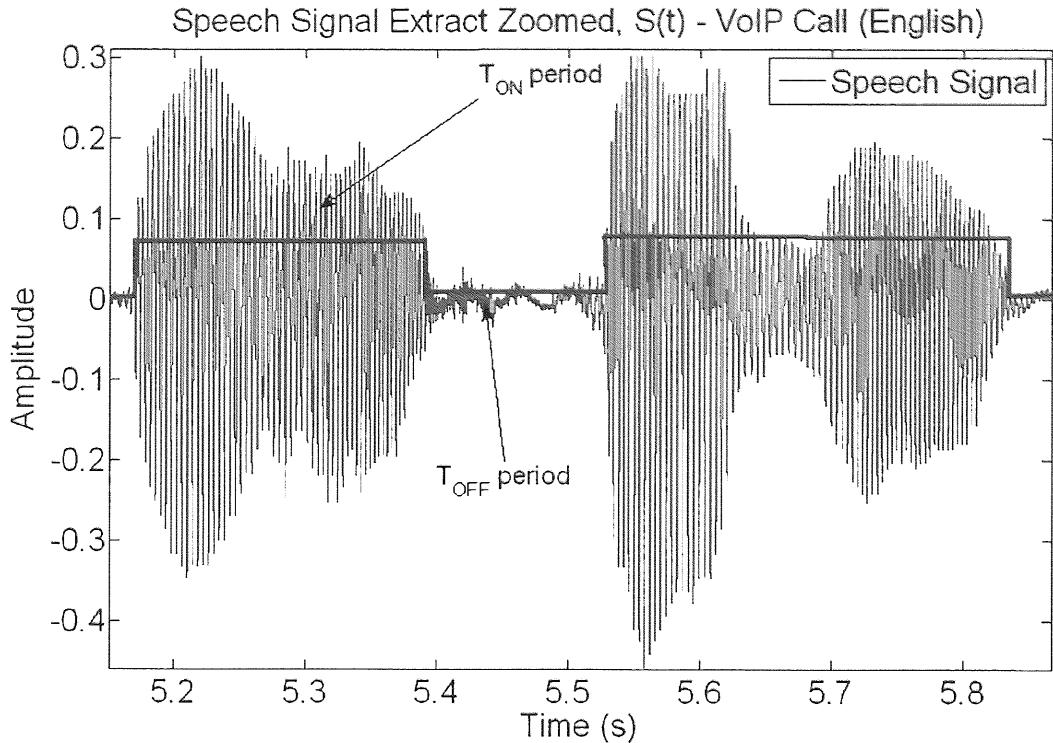


Figure 6.3: Zoomed section of a waveform of extract of a speech signal  $S(t)$  showing a small  $T_{OFF}$  period

### 6.2.2.2 Computing the acoustic envelope using an FFT Low pass filter

The next step involved in the silence detection process, is to compute the smooth acoustic envelope of the speech signal  $S(t)$  in order to aid the decision making process in the silence discrimination of the speech extract of the VoIP call. This involves the use of an efficient FFT based low pass filtering technique to obtain a filtered speech signal  $\hat{S}(t)$ , and then computing the acoustic envelope  $m(t)$ .

The FFT based low pass filter uses the following frequency response function:

$$H(f) = \begin{cases} 0, & |f| > f_c \\ 1, & |f| < f_c \end{cases} \quad (6.1)$$

where  $f_c$  is the cut-off frequency of the low pass filter. Since the speech signals associated with the VoIP call recordings consist of very large datasets, there would be memory issues if a standard FFT approach was taken. Therefore, in order to avoid this, we make use of the Overlap-Add method [90], where the FFT in the algorithm was implemented with the use of the efficient *FFTW* library [91] for fast FFT computations. The principle behind the Overlap Add method is that it performs the low pass filtering of the speech signal segment-by-segment and then reconstructs the filtered speech signal. Fig.6.4 illustrates the Overlap Add method.

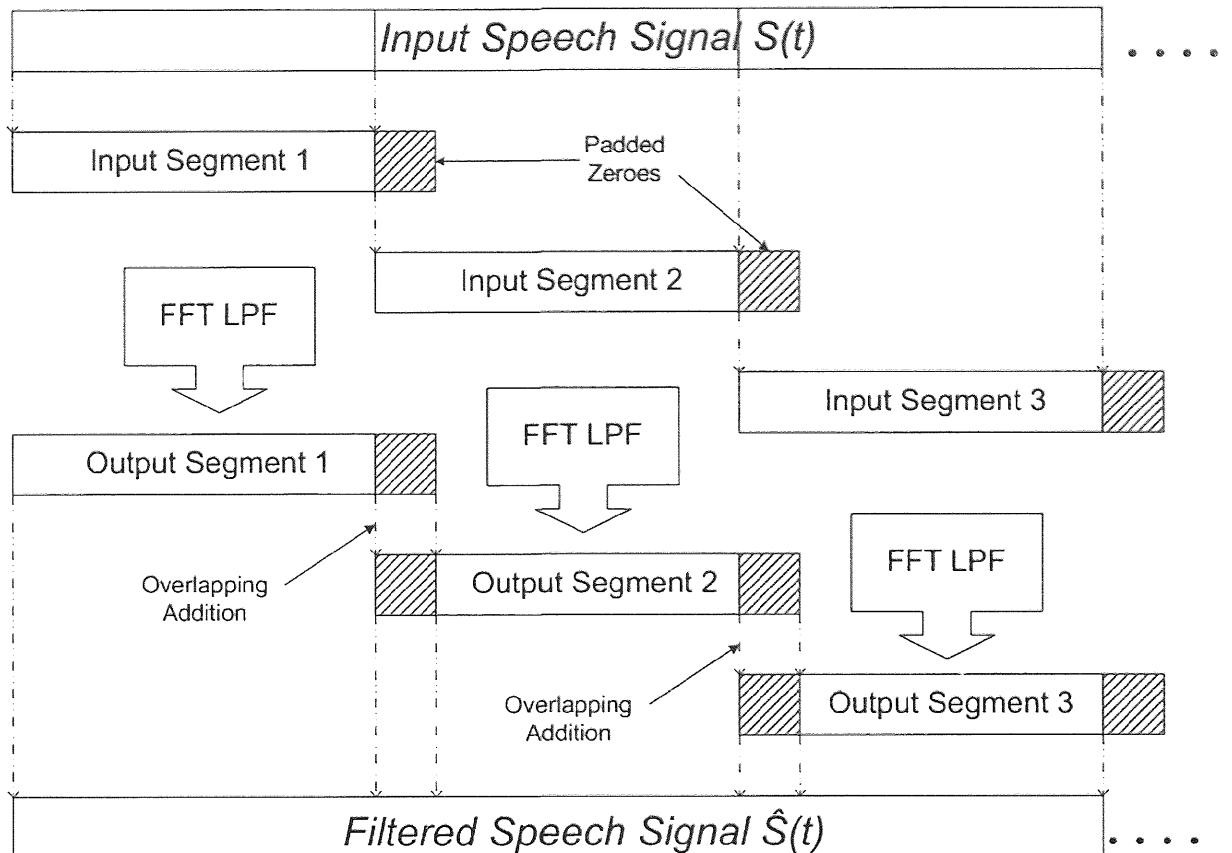
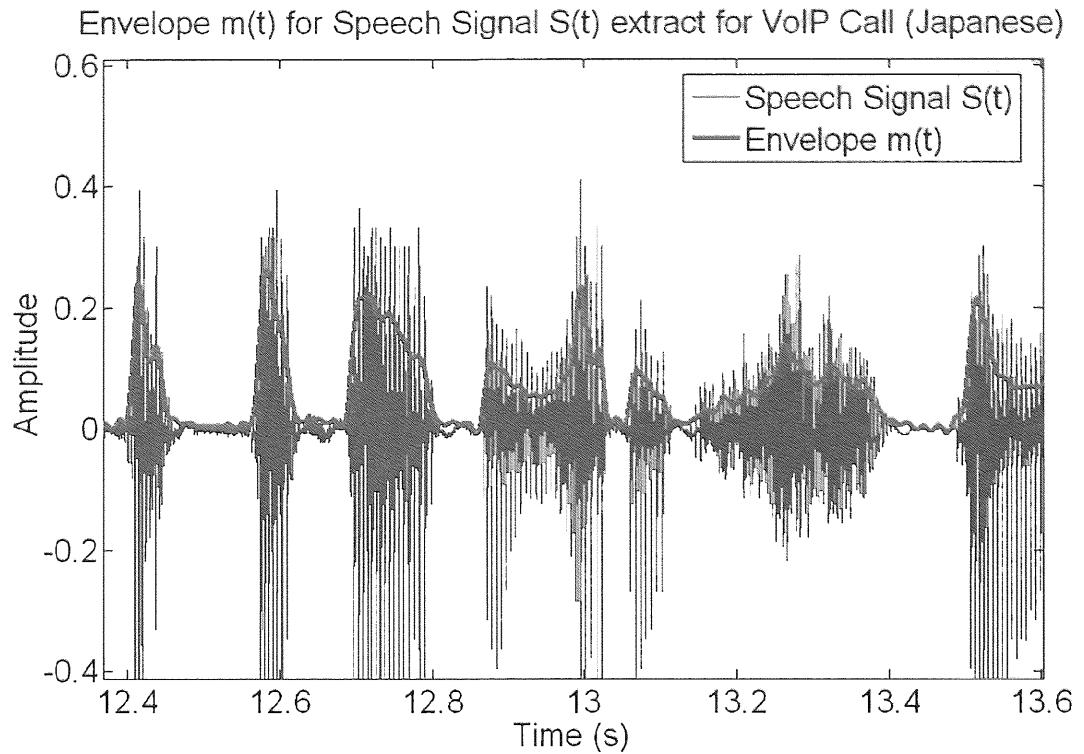


Figure 6.4: An illustration of the Overlap Add method used for Low pass filtering the speech signal

The signal is first broken down into several identically sized segments or blocks. For each block the FFT low pass filtering process is performed by first computing the Forward Fast Fourier transform (FFFT), followed by the introduction of the cut off frequency in the frequency domain, and then obtaining the filtered segment using the Inverse Fast Fourier Transform (IFFT). The complete filtered signal is then obtained by overlapping and adding the individual filtered segments with the use of zero padding to prevent wrap around effects as shown in Fig. 6.4. The overlap here arises due to the fact that a linear convolution always ends up with a sample segment longer than the original segment. Finally once the complete filtered speech signal,  $\hat{S}(t)$  is obtained, the desired acoustic envelope  $m(t)$  is derived by computing the modulus of the filtered signal  $\hat{S}(t)$ , as shown below:

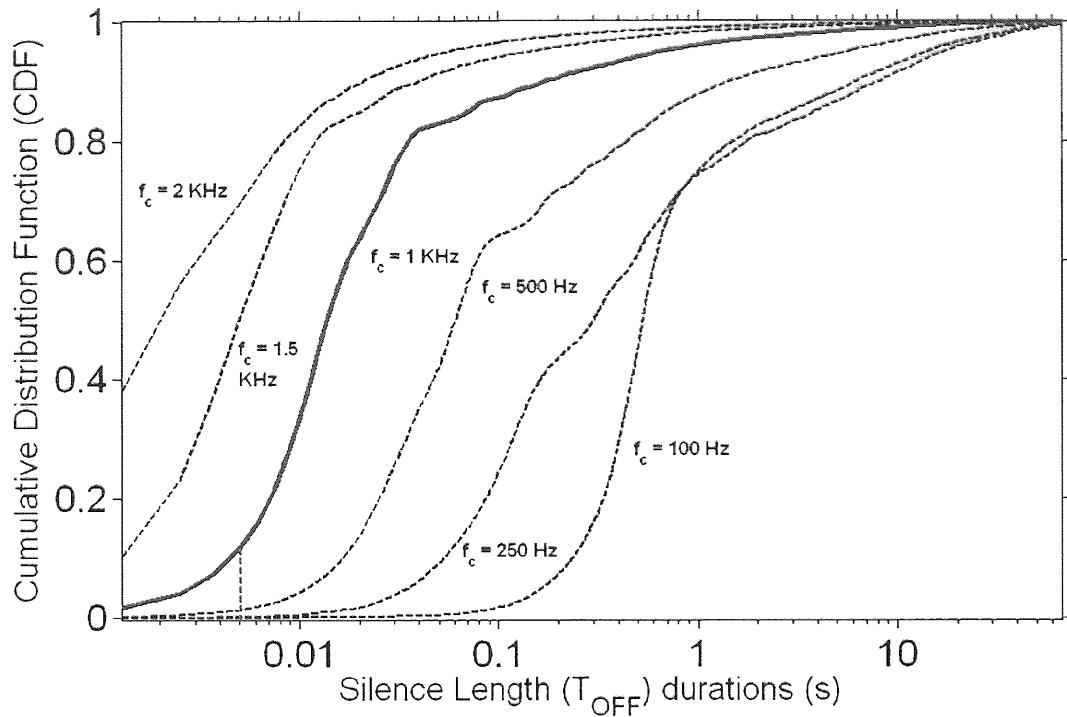
$$m(t) = |\hat{S}(t)| \quad (6.2)$$

Fig. 6.5 shows the acoustic envelope  $m(t)$ , of a speech signal  $S(t)$  for one of the samples of the VoIP call recordings.

Figure 6.5: The Acoustic Envelope  $m(t)$  of a speech signal  $S(t)$ 

### 6.2.2.3 Calibrating the choice of the low-pass filter cut off frequency $f_c$

Since we are only interested in obtaining the envelope of the speech signal and not reconstructing the speech signal with the fundamental carrier components, the choice of  $f_c$  is carefully chosen with the aim of tracking all short silence lengths ( $T_{OFF}$  periods) as accurately as possible. Fig. 6.6 shows the cumulative distribution of  $T_{OFF}$  periods and the impact of varying the values of  $f_c$  on it, for an arbitrary call sample. We observe that for low values of  $f_c$ , we see fewer amount of short silences. For example, for a value of  $f_c = 100$  Hz, we observe that the results closely match those by [74] where silences below 200 ms were not observed. For a large value of  $f_c = 2$  KHz, we see that there is noise set in, thus completely damaging the statistics. Research has suggested that the smallest silence gap that can be detected by an adult in speech perception is 5 ms [92]. Moreover the choice of frame sizes used in the current generation voice codecs vary from 5 ms to 30 ms [93] [94]. Thus the value of  $f_c$  was calibrated so that we record as many  $T_{OFF}$  periods close to 5 ms as possible, limiting the presence of  $T_{OFF}$  periods shorter than 1 ms assuming those to be noise. As shown in Fig. 6.6, upon such a calibration, a cut off frequency  $f_c$  of 1 KHz proved to be the most appropriate choice for our requirements for the low pass filtering process.

Figure 6.6: Impact of change in value of  $f_c$  on the  $T_{OFF}$  periods

#### 6.2.2.4 Discriminating the silences ( $T_{OFF}$ periods) from talk spurts ( $T_{ON}$ periods)

Once the acoustic envelope of the speech signal is obtained, it is then subjected to a silence discrimination process where the talk spurts ( $T_{ON}$  periods) are separated from the silence periods ( $T_{OFF}$  periods) by the use of a Threshold Crossing Value (TCV) which is set comfortably above the estimated noise floor for each speech signal. The noise floor of course varies with each speech signal, and the choice of the TCV also varies with each VoIP call recording file, but as a general rule of thumb in each case, the TCV is chosen midway between the noise floor and the smallest talk spurt spike that can be detected which can stand out from the noise floor, which is roughly about 1.5 – 2.0 % of the maximum amplitude range of the speech signal. This value is then used in a decision making process whereby the duration for which the envelope  $m(t)$  lies below the TCV, is marked as a silence ( $T_{OFF}$  period) and the duration for which the envelope lies above the TCV, is marked as a talk spurt ( $T_{ON}$  period). Once the values of the  $T_{ON}$  and  $T_{OFF}$  periods are accumulated for the entire speech signal, we finally compute the probability density and cumulative density plots for statistical analysis. The impact of the choice of the TCV on the statistics of the talk spurts and silence lengths is consistent with previous work [73]. For lower threshold values we record a larger number of shorter silence lengths and talk spurts, whereas for higher threshold values, we record fewer and longer silence

lengths and talk spurts. However we found the difference between using the two extremes of the values of the threshold to be marginally small compared to those observed by [73] and that the distributions of  $T_{OFF}$  and  $T_{ON}$  periods are less affected as long as the TCV is small enough to capture as much of the speech component of the call recording as possible. Fig. 6.7 shows the impact of the change in TCV values on the cumulative statistics of the  $T_{OFF}$  periods. The TCVs are represented as percentages of the maximum positive amplitudes of the speech signal  $S(t)$ .

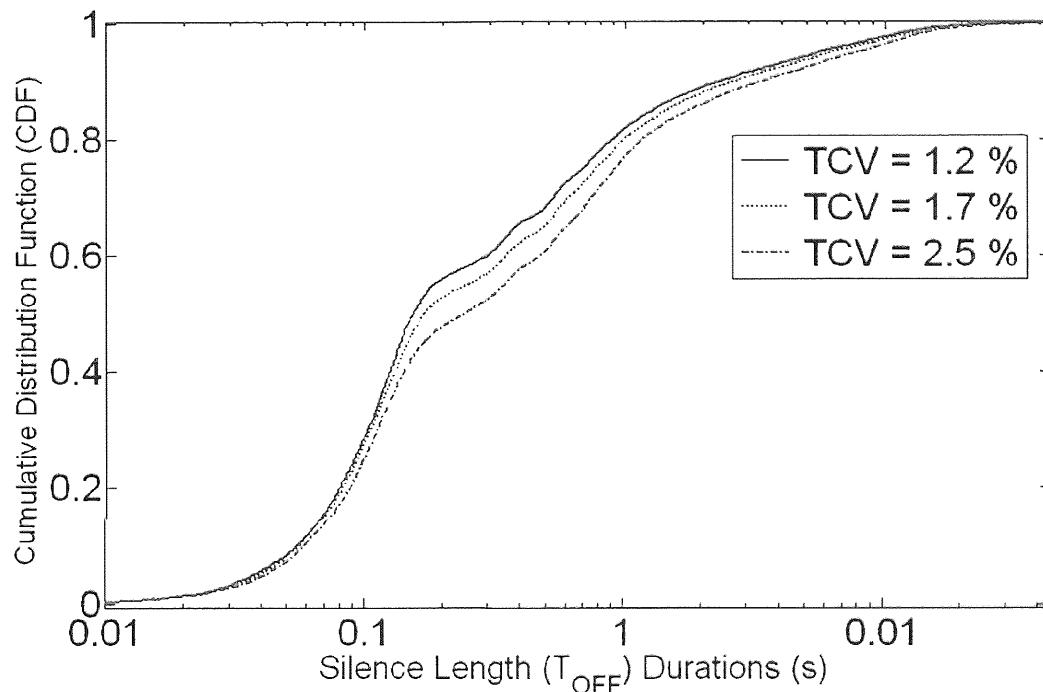


Figure 6.7: Impact of change in value of TCV on the  $T_{OFF}$  periods

### 6.3 Analysis of $T_{ON}$ and $T_{OFF}$ periods and the impact of language, speech rate and caller behaviour

Since most VoIP traffic models, including the 8-state model [79] are based on monolog data, we first study the characteristics of the  $T_{OFF}$  periods based on the monolog data from [9]. Plotting the average statistics for the monolog samples of each group of languages, it was observed that the incidence of short silence lengths or  $T_{OFF}$  periods is unique to each language. A good example of this observation is shown in Fig 6.8 which shows the probability densities for the  $T_{OFF}$  periods for three different languages. While we see no major differences in densities for the long and medium silence lengths, we see that Mandarin Chinese has the highest number of short silence lengths (<200 ms) , followed

by French and then English. This is because the short silence lengths in speech vary with the size of the prosodic units and speech rate involved in the language [95]. Thus we see in our results that Chinese and French, being syllable-timed languages, have a larger number of short silence lengths due to the shorter prosodic units as compared to English, a stress-timed language which has relatively longer prosodic units [96]. This is consistently seen with any combination of the various individual speech samples for these languages. In other words, the observation is independent of the samples used.

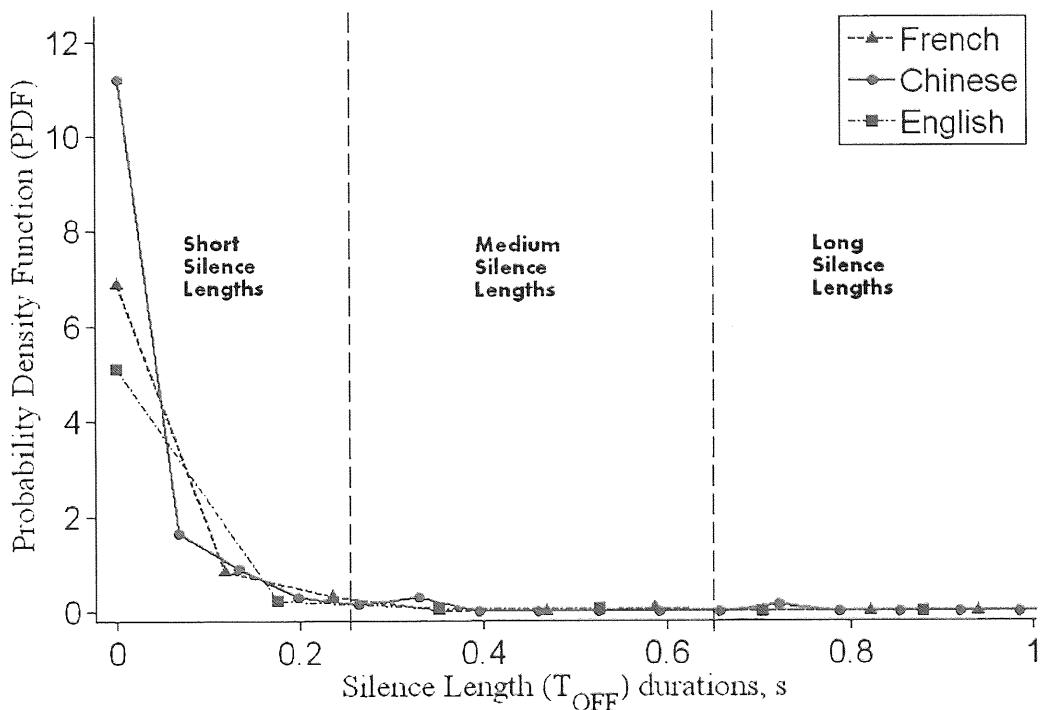


Figure 6.8: Comparison of monolog silence lengths ( $T_{OFF}$  periods) for Chinese, French and English Languages

These differences in the short silence lengths with every language have been proved to be statistically significant [97]. Voice models based on monolog speech should therefore, take the language into consideration as the speech rate varies with every language. While we have noted the dependence of the short silence lengths on the language, we did not observe any critical differences in the  $T_{OFF}$  period distributions between the male and female speakers.

However, in our analysis of silence lengths or  $T_{OFF}$  periods of the two way VoIP Call recordings from [86], we observe that the statistical differences in languages are overshadowed by the random dynamics of the interaction of the speaker. Fig. 6.9 shows the cumulative distribution functions for the entire number of  $T_{OFF}$  periods for calls in different languages. We present the  $T_{OFF}$  periods in

frames of 5 ms as in [79] and [81] to facilitate ease of comparison. The plot shows that we have recorded about 10-30% more short silence lengths (<100 ms) in our measurements when compared to those in [81].

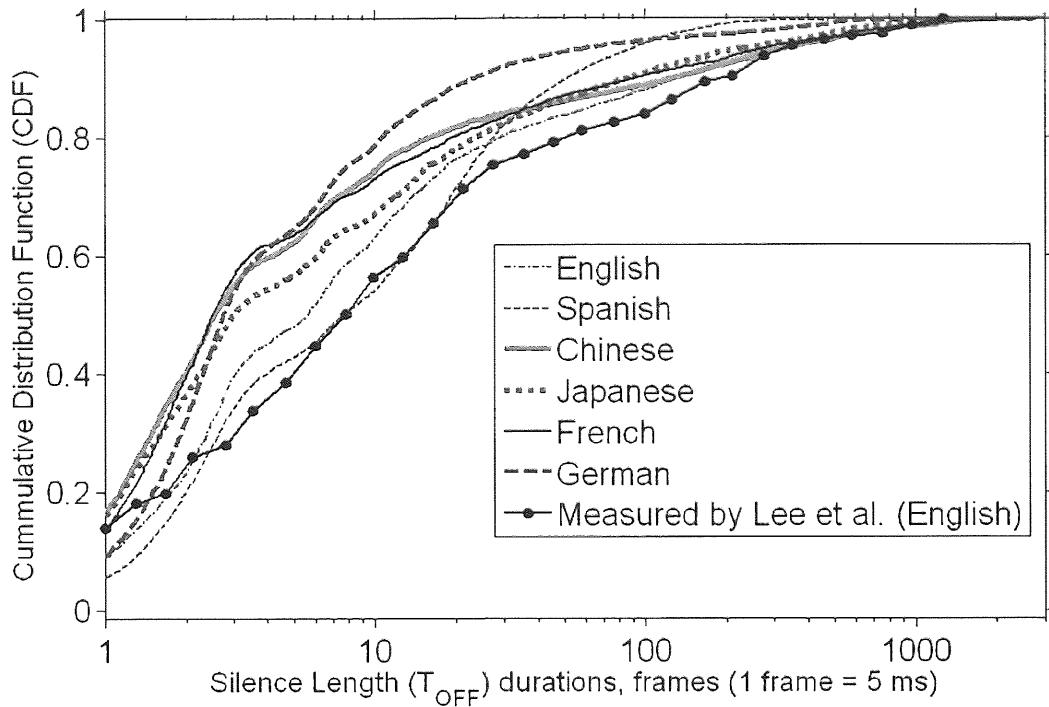


Figure 6.9: Cumulative distribution of silence lengths ( $T_{OFF}$  periods) for various language samples of two-way VoIP calls in contrast to measurement by [81]

We observe from the results that although there are differences in distributions of silence lengths in the voice call samples for each language, the frequent presence of non-grammatical silence lengths such as those related to listening, thinking and hesitation in the voice call samples, yields a complex anomaly which does not present a viable option to model the voice calls silences on a linguistic or speech rate basis alone. This is clearly observed with the German and Spanish samples which have a large presence of prolonged silences of hesitation and semi-intentional silence lengths, thus yielding higher numbers of medium and long silence lengths compared to the other language samples. In addition, results of the extensive two-sample Kolmogorov-Smirnov tests [98] performed for a large selection of the samples, revealed that the distance between the cumulative distributions of  $T_{OFF}$  periods for samples of calls in different languages were nearly the same as those of the same language, thus ruling out the possibility of modelling the  $T_{OFF}$  periods by a unique distribution for each individual language. This is clearly apparent from the results in Fig. 6.10 which show the cumulative distributions of  $T_{OFF}$  periods for eight different call samples for the Spanish language. As

can be seen, this plot with the distributions looks similar to the one shown in Fig. 6.9 involving the  $T_{OFF}$  periods with different languages.

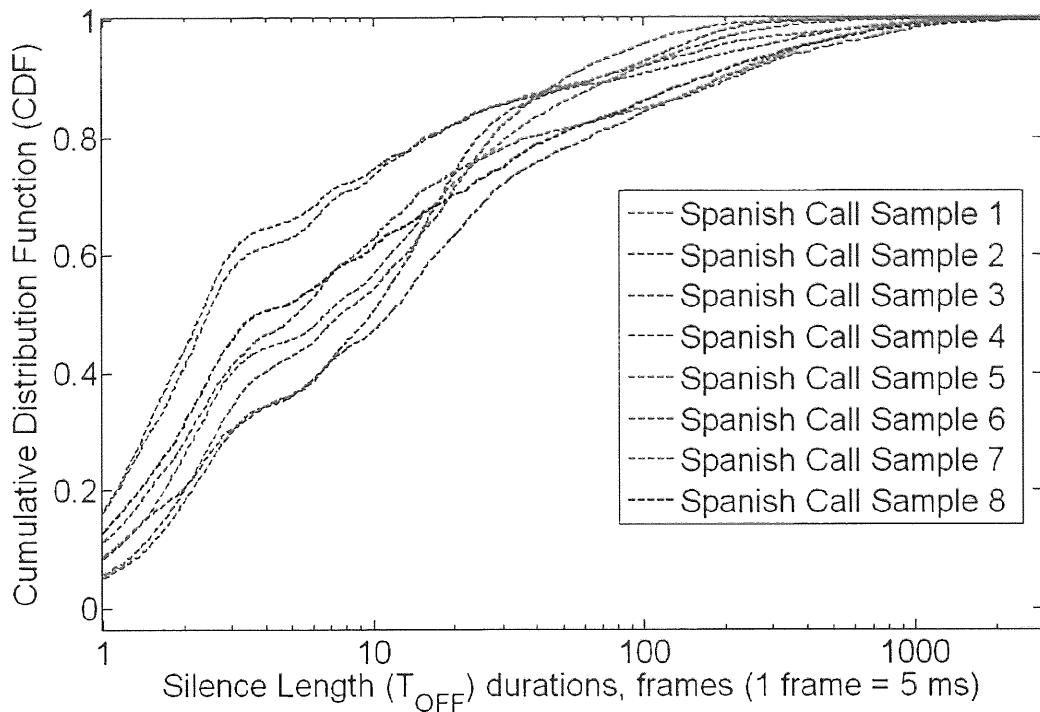


Figure 6.10: Cumulative distribution of silence lengths ( $T_{OFF}$  periods) for eight different calls for the Spanish language

A similar observation has been made with the cumulative statistics of the  $T_{ON}$  periods. Fig. 6.11 shows the cumulative distributions for the  $T_{ON}$  periods for calls in different languages. The differences in the  $T_{OFF}$  period distributions are clearly apparent here as well. We observe that the call samples of Chinese and French languages show a larger number of talk spurts when compared to English and the other languages, as in principle these languages have a short and spontaneous prosodic and temporal structure associated with the words and syllables, since they are syllable timed languages. The call sample for Japanese, a *mora-timed* language [99], where some vowels take twice the time as others, in principle should have a lot lower number of talk spurts when compared to the Chinese and French, but we notice that the percentage of talk spurts for Japanese is not distinctively different from that of Chinese. Thus we conclude that there is a presence of other non-grammatical natured talk spurts which make it almost impossible to characterise the  $T_{ON}$  periods on the basis of languages. The speech-rate with which the caller speaks in a particular language is a more dominant parameter than the prosodic or temporal structure of the language, when it comes to the statistical count of the  $T_{ON}$  periods. This speech rate changes from one call to another as a direct result of the varying nature of the emotions and psychological behaviour of the speaker. Hence there is an absence of uniformity

observed in the characteristics of  $T_{ON}$  periods for various call samples of the same language as was observed in the case of the  $T_{OFF}$  periods as well. Another example of this kind of inconsistency can be observed in the Spanish sample shown in Fig. 6.11, where we see that the call sample has a very limited number of occasions when the speakers were talking, as there was the presence of several long silence periods related to hesitation. This is likely to be due to the speakers involved in the conversation being very conscious about the recording process and in a state of hesitation in relation to their talk activity. Hence it can be seen from the Spanish Call sample, that there is a much smaller number of  $T_{ON}$  periods when compared to other languages. In addition to the factors already mentioned, it is significant to note that the frequent presence of non-speech expressions such as coughing, laughing, anger, also get included into the  $T_{ON}$  periods thus making the characteristics of the  $T_{ON}$  and  $T_{OFF}$  even more complex and independent of language.

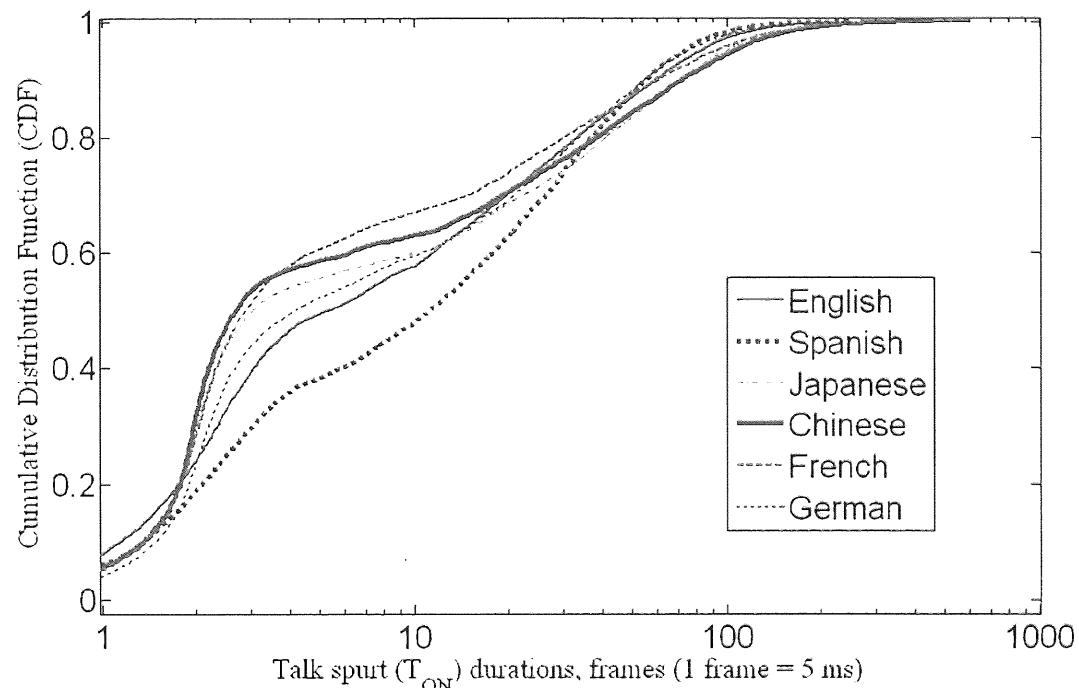


Figure 6.11: Cumulative distribution of silence lengths ( $T_{ON}$  periods) for call samples for different languages

The study of the impact of languages on the OFF period characteristics reveals that, while there are distinguishable differences observed with monolog speech, a more complex anomaly of  $T_{OFF}$  and  $T_{ON}$  periods is observed with the two-way call conversations due to the presence of random conversational silences and associated dynamic speech rates of the speakers. Therefore, in order to propose a good model it is imperative that it is based on the composite statistics of the  $T_{ON}$  and  $T_{OFF}$  periods which average out all the complex statistical differences.

## 6.4 Summary

In this chapter, we looked in great detail, at the background to voice traffic modelling and highlighted the advantages and limitations of some popular voice traffic models. We then looked at the detailed procedure of the measurement of the  $T_{ON}$  and  $T_{OFF}$  periods from a very large multi-lingual database of VoIP call recordings. We then made an in-depth assessment of the impact of varying languages on the characteristics of the  $T_{ON}$  and  $T_{OFF}$  periods and concluded that the random conversational dynamics of the speakers dominate the differences between the call samples and hence it is not possible to model the  $T_{ON}$  and  $T_{OFF}$  periods based on individual languages, but rather use of composite statistics is required for modelling of  $T_{ON}$  and  $T_{OFF}$  periods accurately so that a generic model can be proposed based on averaged statistics of language and conversational behaviour.

# Chapter Seven

## An ON-OFF VoIP Source Model with Lognormal Transitions

In this Chapter, an ON-OFF VoIP traffic Model is proposed based on the composite statistics of the  $T_{ON}$  and  $T_{OFF}$  periods. Section 7.1 looks at the use of the composite statistical samples in a logarithmic analysis, where multimodal distributions are fitted, the parameters of which are used for an ON-OFF model proposed in section 7.2. Section 7.3 covers the simulated results of the  $T_{ON}$  and  $T_{OFF}$  periods from the Log-normal model and compare them with the results of other popular models which include short silences in their statistics. Finally we summarise the results in section 7.4.

### 7.1 Analysis of Composite Statistical Samples of the $T_{ON}$ and $T_{OFF}$ periods

In the previous chapter, we have seen from the analysis of VoIP calls for various languages, that although there are differences in distributions of silence lengths in the voice call samples for each language, the frequent presence of non-grammatical silence lengths such as those related to listening, thinking and hesitation in the voice call samples yield complex behaviour, inconsistent with modelling the voice call silence ( $T_{OFF}$ ) or talk spurt ( $T_{ON}$ ) periods on a linguistic or speech rate basis alone. Therefore, in order to propose an appropriate model which averages these conversational characteristics and the linguistic differences of all the call samples in the database, it is necessary to produce a model which is based on the consolidated composite statistics of the  $T_{ON}$  and  $T_{OFF}$  periods, for the entire VoIP call database for all speakers of all languages considered.

The composite statistical samples for the  $T_{ON}$  and  $T_{OFF}$  periods were produced by merging all  $T_{ON}$  and  $T_{OFF}$  periods accumulated individually for all available call samples. This composite statistical sample, which averages over all linguistic and conversational differences between each call sample, provides an excellent platform to produce a model the characteristics of which, would closely match with nearly all samples considered. Although considerably large in size, the composite statistical samples

were good enough to be easily used for statistical analysis performed on a Pentium-D dual-core machine without hitting any computing performance bottle necks.

Recent studies on the structure of conversational speech have revealed that the human perception of time with respect to talk spurts and silences are logarithmic in nature. Both [97] and [100] have showed in their recent work that logarithmic fits are more appropriate when modelling the temporal characteristics of speech. Also [101] showed that the voice call holding time distributions for the connected periods are best modelled by a multi-modal distribution on a logarithmic scale. In line with these observations, we follow a similar approach and we produce log-linear plots of the probability densities of the composite statistics for the analysis and modelling of  $T_{ON}$  and  $T_{OFF}$  periods.

### 7.1.1 Logarithmic Analysis and Modelling of $T_{OFF}$ periods

We first start by looking at the log-linear density plot for the composite statistical sample of  $T_{OFF}$  periods shown in Fig. 7.1. We clearly observe from the composite sample (red asterisks), that the  $T_{OFF}$  periods have a tri-modal Gaussian structure on a logarithmic scale. The results suggest that at least 67% of the  $T_{OFF}$  periods or silence lengths are less than 100 ms, a characteristic which many models have failed to capture so far. We find that the three peaks are associated with three different ranges of silence durations. The first peak from the left indicates that a vast majority (56%) of the silence lengths are actually close to 10 ms which correspond to *respiratory pauses* as specified by Zellner [95]. These respiratory pauses are very small breathing gaps between words in spontaneous speech, which has a lot to do with the kind of stress and speech rate of the interlocutor. The next peak describes silence lengths ranging from 80 ms to 100 ms which contribute to 11% of the  $T_{OFF}$  statistics. These kinds of silence lengths most likely correspond to the *Intra-segmental pauses* which are defined as pauses which occur between two occlusions of the vocal tract as defined by [4]. This definition of Intra-segmental pauses closely matches with the general definition of *short silence lengths* which Stern et al used in their work [79], also mentioned earlier in the previous chapter. The third peak corresponds to silences greater than 100 ms which represents the longer silence periods consisting of inter-lexical pauses (occurring between words and sentences [95]), the semi-intentional silences of hesitation, and the listening silences.

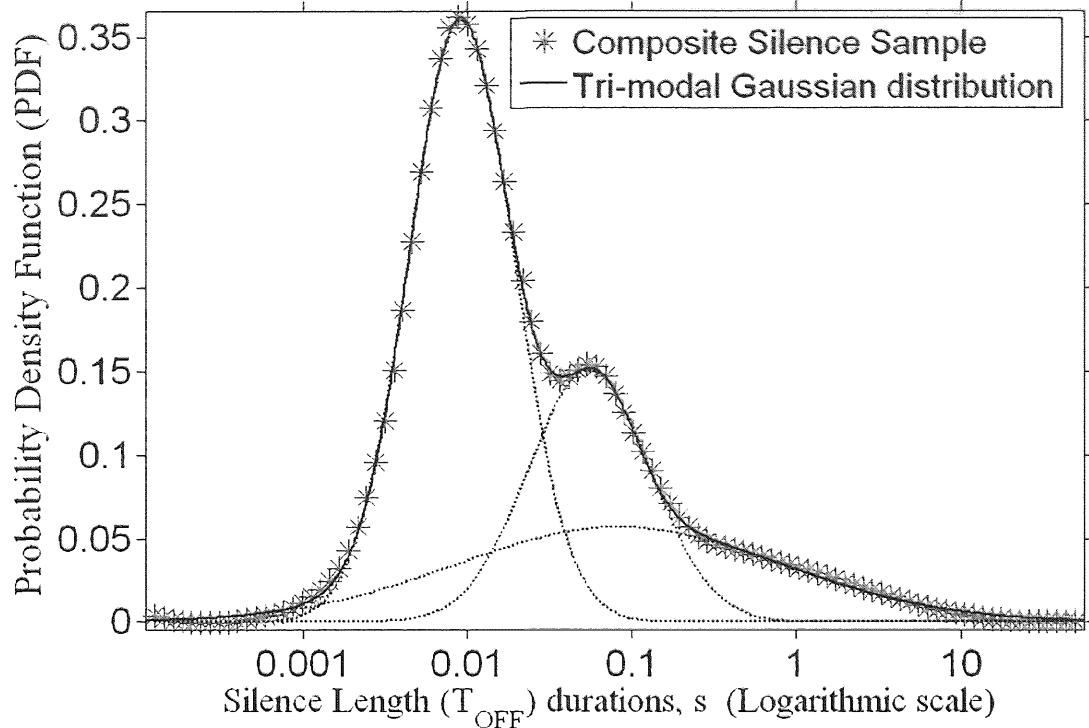


Figure 7.1: Tri-modal Gaussian distribution fit to Probability Density of Composite Sample of  $T_{OFF}$  periods

In order to model the probability density function of the  $T_{OFF}$  periods, with a tri-modal Gaussian distribution, we follow the same procedure of Gaussian component fitting, which we had discussed earlier in section 3.3 of Chapter 3 which was used for fitting simple finite state Gaussian Markov models to the measured PDFs of the IPT of IP packet traffic over the network. The probability density of the composite  $T_{OFF}$  periods in the logarithmic domain can be fitted with a mixture of three Gaussian components as indicated clearly by the black dotted lines in Fig. 7.1. The parameters of this tri-modal Gaussian distribution are derived in the same manner as before using the Expectation-Maximization (EM) algorithm. So, recalling the convex combination equation used in Chapter 3, we have the tri-modal Gaussian mixture defined in equation 7.1.

$$f_{(T_{OFF}, \text{Logarithmic})} = \alpha_1 G_1(\mu_1, \sigma_1) + \alpha_2 G_2(\mu_2, \sigma_2) + \alpha_3 G_3(\mu_3, \sigma_3) \quad (7.1)$$

Where each Gaussian component is given by the equation (7.2):

$$G_i(\mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(t-\mu_i)^2}{2\sigma_i^2}}, i = 1, 2, 3 \quad (7.2)$$

where  $\alpha_i$  is the mixture probability ratio for  $i = 1$  to 3 such that  $\alpha_1 + \alpha_2 + \alpha_3 = 1$ .  $\mu_i$  is the mean and  $\sigma_i$  is the standard deviation of each Gaussian component for the  $T_{OFF}$  periods on the logarithmic scale. With reference to statistical theory, it is obvious that a tri-modal Gaussian mixture distribution on the logarithmic scale can be represented by a tri-modal Log-normal distribution on the linear scale. This is because the Log-normal distribution is defined as the probability distribution of a random variable whose logarithm is Gaussian distributed.

The Log-normal distribution has a PDF given by equation 7.3:

$$P(t)_{(T_{OFF}, \text{Log-normal})} = \frac{1}{t\sigma\sqrt{2\pi}} e^{\frac{-(\log_{10}(t)-\mu)^2}{2\sigma^2}} \quad (7.3)$$

Where,  $\mu$  and  $\sigma$  are the mean and standard deviation of the constituent Gaussian component respectively. Therefore using this principle, equation 7.1 representing the fitted PDF for  $T_{OFF}$  periods on a logarithmic scale, can be reformulated into equation 7.4, which represents the fitted PDF for  $T_{OFF}$  periods on a linear scale.

$$f_{(T_{OFF})}(t) = \sum_{i=1}^3 \alpha_i \frac{1}{t\sigma_i\sqrt{2\pi}} \exp\left(\frac{-(\log_{10}(t) - \mu_i)^2}{2\sigma_i^2}\right) \quad (7.4)$$

The parameters estimated for the above equation using the EM algorithm are tabulated in Table 7.1.

Duration	Parameters	Value		
		$i = 1$	$i = 2$	$i = 3$
$T_{OFF}$ Periods (States 3, 4 and 5)	A	0.56	0.11	0.33
	$\mu$	-2.05	-1.29	-1.08
	$\sigma$	0.29	0.23	0.99

Table 7.1: Estimated values of parameters for Lognormal Densities for  $T_{OFF}$  periods

We see from Fig. 7.1, that the tri-modal Log-normal distribution fits the composite statistic well. We had mentioned earlier in section 6.1 of Chapter 6 that recent work has suggested that the  $T_{OFF}$  periods have been found to be heavy tailed. We can check this by plotting the composite statistic along with the fitted log-normal distribution and the two-term exponential distribution proposed by

Stern et al. in their work [79] on a logarithmic density scale, see Fig. 7.2. The figure which shows the Complementary Cumulative Density Function (CCDF) (i.e. 1-CDF) confirms the heavy-tailed nature of the  $T_{OFF}$  periods, as the composite data clearly tracks the heavy-tailed Lognormal distribution well rather than the two-term exponential distribution proposed by [79]

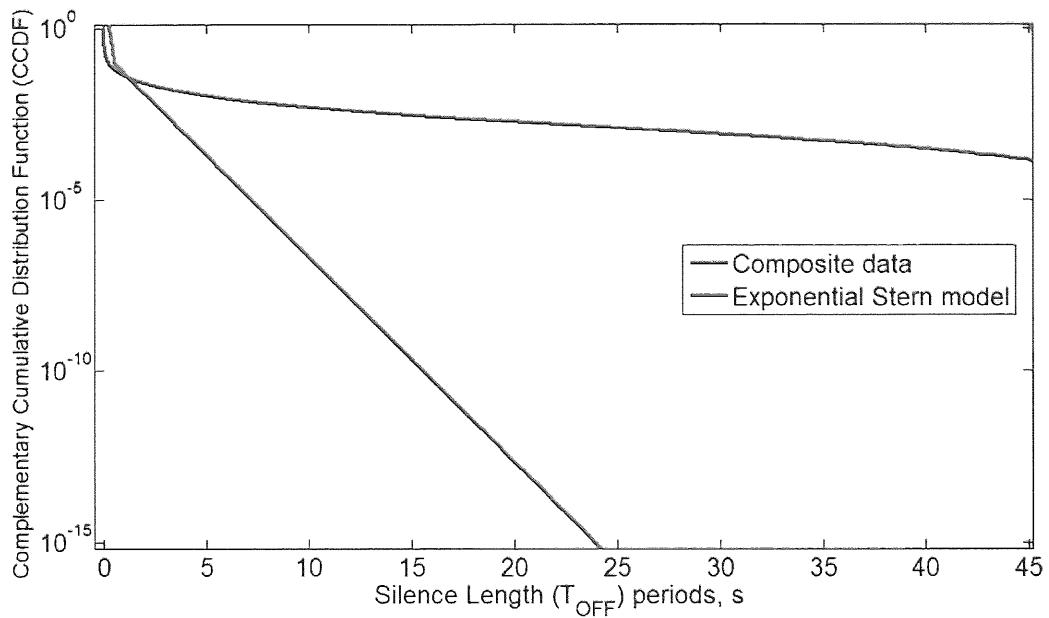


Figure 7.2: Confirming the heavy-tailed nature of the  $T_{OFF}$  periods

### 7.1.2 Logarithmic Analysis and Modelling of $T_{ON}$ periods

Moving on to the logarithmic analysis of the  $T_{ON}$  periods, we repeat a similar approach for analysis to that used with the  $T_{OFF}$  periods. Plotting the probability density of the  $T_{ON}$  periods on a log-linear plot, shown in Fig. 7.3, we observe a bi-modal profile for the density of the  $T_{ON}$  periods, given by the blue asterisks. These results show that we have been able to accurately track talk spurts below 100 ms, which [81] and [79] were not able to track and classify. We see that a significant amount of talk spurts which we have tracked, lie in the region of 6 ms to 10 ms in contrast to those longer than 100 ms. While these spurts may randomly represent elements of background noise recorded during the conversation of the VoIP call, we believe that, given that we had  $T_{OFF}$  periods in the same time region, these small  $T_{ON}$  periods detected are small vowel segments or prosodic units as pointed out in [95].

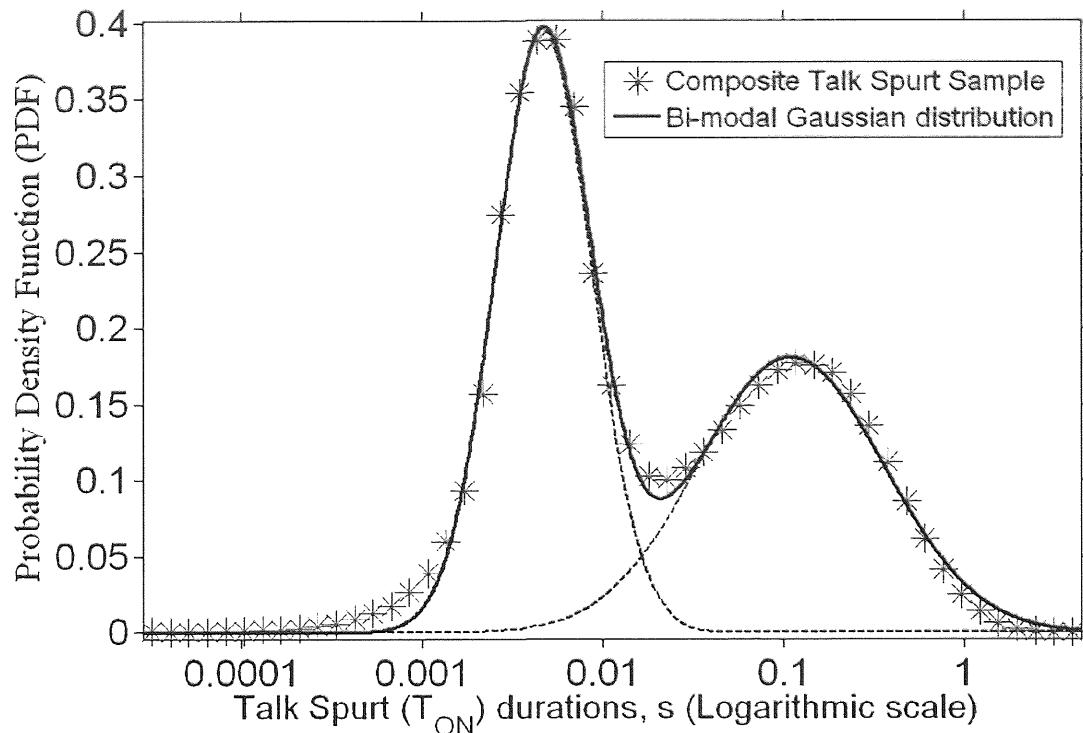


Figure 7.3: Bi-modal Gaussian distribution fit to Probability Density of Composite Sample of  $T_{ON}$  periods

To this bi-modal structure, we fit a bi-modal Gaussian distribution with the two peaks represented by the black dotted lines clearly shown in Fig. 7.3. Now the equation for the bi-modal distribution fit in the logarithmic domain is given by equation 7.5.

$$f_{(T_{ON}, \text{Logarithmic})} = \beta_1 \text{Gaussian}(\lambda_1, \varsigma_1) + \beta_2 \text{Gaussian}(\lambda_2, \varsigma_2) \quad (7.5)$$

where  $\beta_i$  is the mixture probability ratio such that  $\beta_1 + \beta_2 = 1$ .  $\lambda_i$  is the mean and  $\varsigma_i$  is the standard deviation of each Gaussian component for the  $T_{ON}$  periods on the logarithmic scale, and  $i = 1$  to 2. Again as explained before, equation 7.5 which represents the fitted PDF for  $T_{ON}$  periods on a logarithmic scale can be reformulated into equation 7.6, which represents the fitted PDF for  $T_{ON}$  periods on a linear scale.

$$f_{(T_{ON})}(t) = \sum_{i=1}^2 \beta_i \frac{1}{t \varsigma_i \sqrt{2\pi}} \exp\left(\frac{-(\log_{10}(t) - \lambda_i)^2}{2\varsigma_i^2}\right) \quad (7.6)$$

The parameter values for equation 7.6, estimated using the EM algorithm, are tabulated in table 7.2.

Duration	Parameters	Value	
		i = 1	i = 2
T <sub>ON</sub> Periods (States 1,2)	$\beta$	0.53	0.47
	$\lambda$	-4.653	-1.919
	$\varsigma$	0.531	1.02

Table 7.2: Estimated values of parameters for Lognormal Densities for T<sub>ON</sub> periods

## 7.2 The ON-OFF Source Model

Based on the fitted log-normal distributions for the T<sub>ON</sub> and T<sub>OFF</sub> periods given by equations 7.4 and 7.6, a simple five state Semi-Markov ON-OFF source model can be constructed as shown in Fig. 7.4. It consists of two ON states and three OFF states. The three modes of the tri-modal log-normal distribution of the T<sub>OFF</sub> given by equation 7.4 are represented by the states 3, 4 and 5 respectively and the two modes of the bi-modal lognormal distribution of the T<sub>ON</sub> periods given by equation 7.6 are represented by the states 1 and 2 respectively.

The model transits between the ON and OFF states alternately such that when simulated, a T<sub>ON</sub> period is always followed by a T<sub>OFF</sub> period as a talk spurt is always followed by a period of silence. The state transition probabilities (probability of visiting a state) are represented by the normalising weights  $\alpha_i$  and  $\beta_i$  of equations 7.4 and 7.6, shown in brackets in Fig. 7.4. The simple state transition probability matrix  $P_{ij}$ , for the Monte Carlo Simulation is then given by the equation 7.7.

$$P_{ij} = \begin{pmatrix} 0 & 0 & 0.56 & 0.11 & 0.33 \\ 0 & 0 & 0.56 & 0.11 & 0.33 \\ 0.53 & 0.47 & 0 & 0 & 0 \\ 0.53 & 0.47 & 0 & 0 & 0 \\ 0.53 & 0.47 & 0 & 0 & 0 \end{pmatrix} \quad (7.7)$$

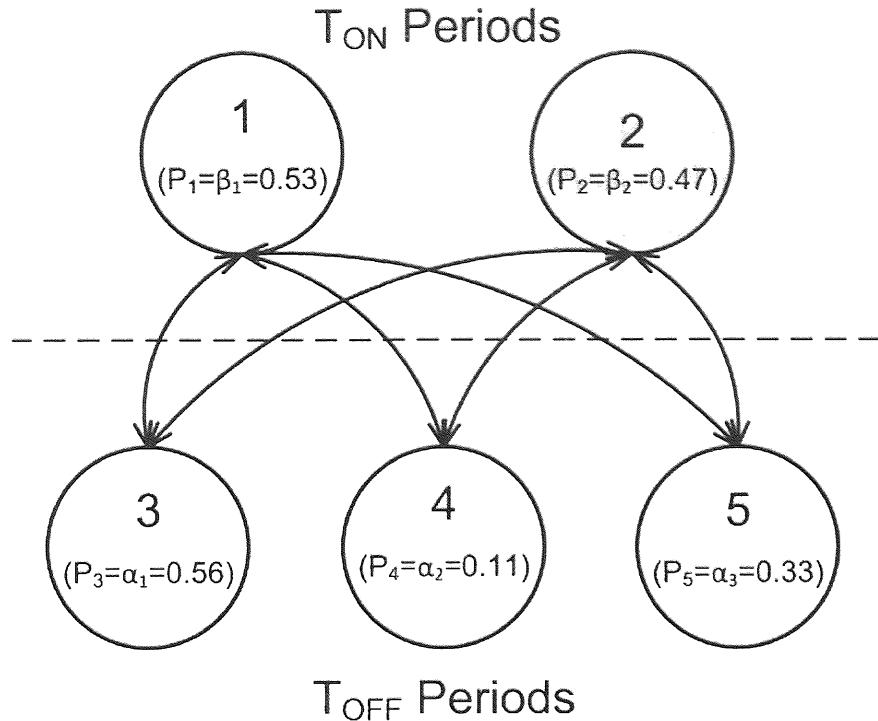


Figure 7.4: The ON-OFF VoIP source Markov model with Log-normal transitions

The model spends a log-normally distributed amount of time in each state before transiting to the next state. In other words, the transitions of the model are log-normal. The parameters used for the log-normal state holding time distributions of the ON and OFF states are given in tables 7.2 and 7.1 respectively.

## 7.3 Simulated Results and Analysis

The proposed five state ON-OFF source model was used in a Markov Monte Carlo simulation, for simulating the  $T_{ON}$  and  $T_{OFF}$  periods. For the simulation setup, the same procedure as the one detailed in section 2.1.4 of Chapter 2 was used. 2 million samples in total were generated, so that there are a million samples of  $T_{ON}$  periods and a million samples of  $T_{OFF}$  periods respectively. In the analysis that follows, we will refer to the proposed five state ON-OFF model as the Log-normal Model and compare its results to the widely referenced measured results of Lee et al [81] and to that of the 8-state model [79] both of which considered short silences, and make relevant conclusions. The results of the 8-state model were simulated using the transition parameters mentioned in [79].

### 7.3.1 Analysis of simulated results of $T_{OFF}$ periods

Fig. 7.5 shows the cumulative distribution functions (CDFs) for our measured (composite) and the simulated  $T_{OFF}$  periods of the Log-normal model and we compare these results to the simulated and measured results of [81] and [79] respectively. As before, to facilitate ease of comparison, the CDFs for the  $T_{OFF}$  are represented in frames of 5 ms as in [81] [79]. The figure clearly shows that our simulated results for the  $T_{OFF}$  periods clearly track our measured (composite) result accurately. With our results, we record an average of 30% more short silences than the English language based, limited time measurements of [80] [79] based on a frame by frame decision process for the silence detection. The Stern model attempted to model the Lee data closely, but this thesis work recorded and modelled higher number of silences than the Stern and Lee models and the figure highlights the inaccuracy of the popular 8-state Stern model in tracking our measured  $T_{OFF}$  periods, especially for the smaller  $T_{OFF}$  periods.

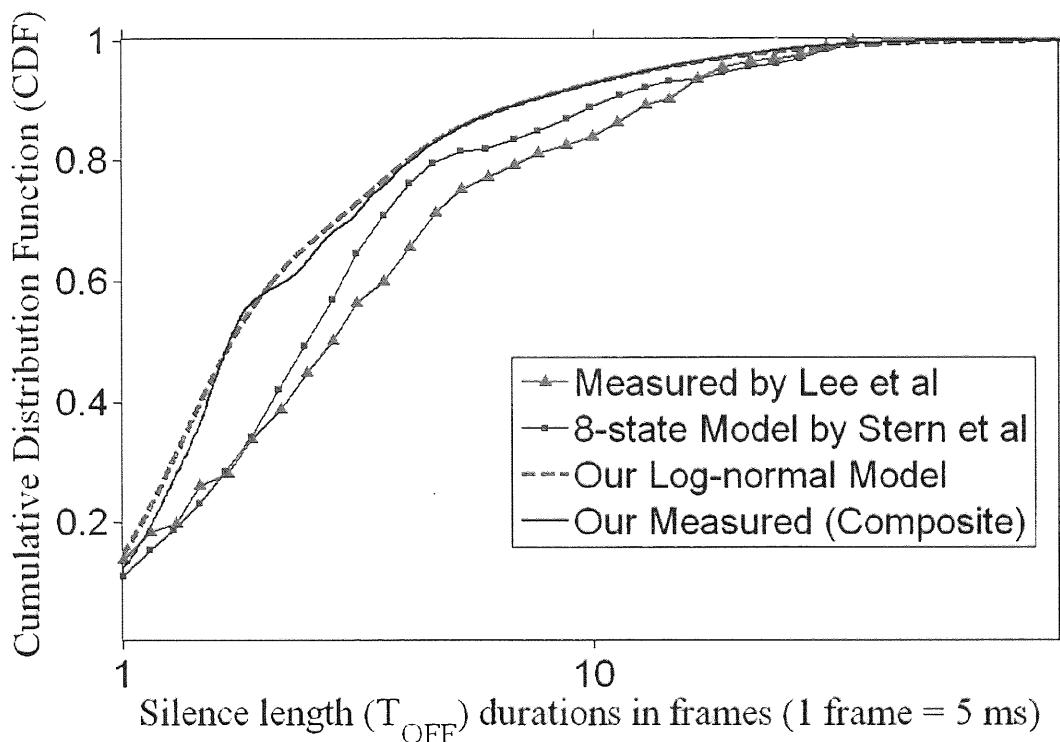


Figure 7.5 : Cumulative distributions for  $T_{OFF}$  periods - Measured values (Composite) vs. Our Log-normal model and results from [79] and [81]

### 7.3.2 Analysis of simulated results of $T_{ON}$ periods

Fig. 7.6 shows the cumulative distributions for the measured (composite) and simulated results of the  $T_{ON}$  periods along with the results by [81] and [79]. Again we observe that the simulated results [187]

of our Log-normal model almost accurately fit our measured (composite) results. In contrast, we see that the results of the 8-state model widely differ with our results as both [79] and [81] state in their work that they could not accurately track talk spurts shorter than 100 ms. Also in our results we find that around 60% of the talk spurts of the multilingual database are below 50 ms in comparison to the 40% observed by [79].

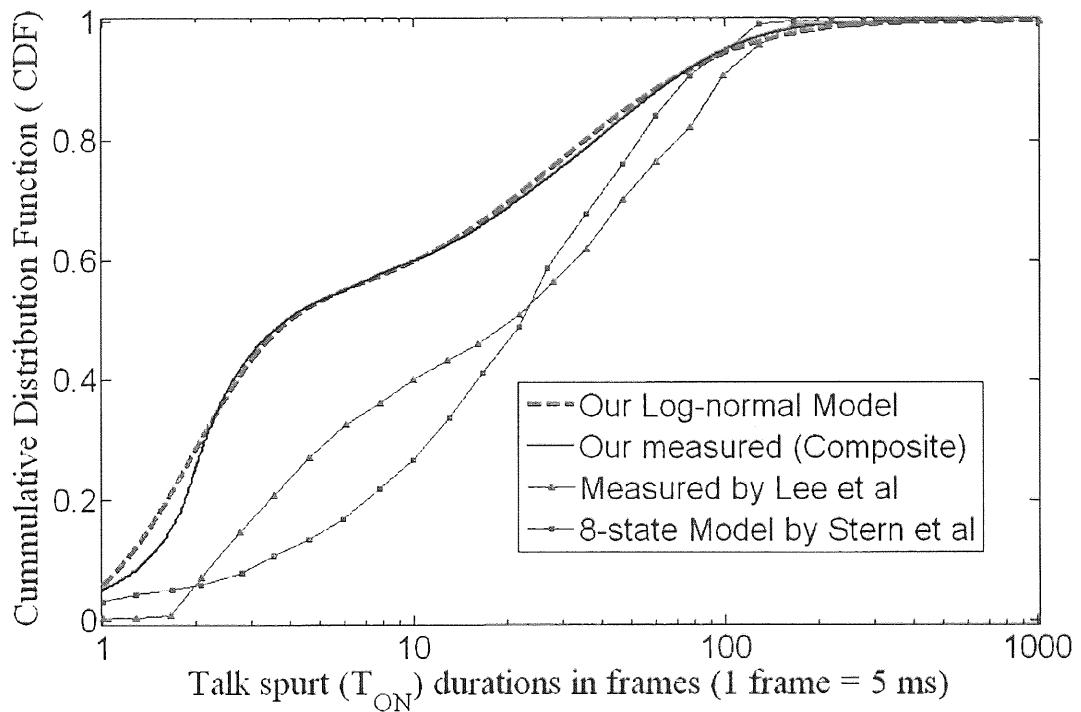


Figure 7.6: Cumulative distributions for  $T_{ON}$  periods - Measured values (Composite) vs. Our Log-normal model and results from [81] and [79]

While it is also apparent from Fig. 7.6 that around 25% of the talk spurts are below 10 ms, further analysis revealed that if we filter out all talk spurts below 10 ms from our composite measured sample, our result would look very similar to that of [79] as shown in Fig. 7.7. This highlights the inefficiency of the detecting methodology used by [79] to track the shorter talk spurts, thus leading to the differences observed in our measurements.

### 7.3.3 Conclusion on the measured and simulated results of the $T_{ON}$ and $T_{OFF}$ periods

From the measured and the simulated results of the  $T_{OFF}$  periods, we have observed higher percentage of shorter silences when compared to the previous studies. This suggests that these detected short silences can be well exploited by a sophisticated Voice Activity Detector(VAD) in [188]

terms of silence suppression with a codec like G.729a which has a frame rate of 10 ms, to further enhance system capacity by increasing the efficiency of the traffic sources with increased  $T_{OFF}$  times. This is particularly crucial for wireless networks where spectral economy is vital. Our proposed Log-normal ON-OFF source model clearly tracks our measured results for  $T_{ON}$  and  $T_{OFF}$  very well and can be easily implemented. Conventional Markov source models with exponential OFF periods, still widely used due to ease of implementation, fail to accurately model our measurements and we conclude that the proposed ON-OFF source model should be used for simulation of  $T_{ON}$  and  $T_{OFF}$  periods of VoIP traffic instead.

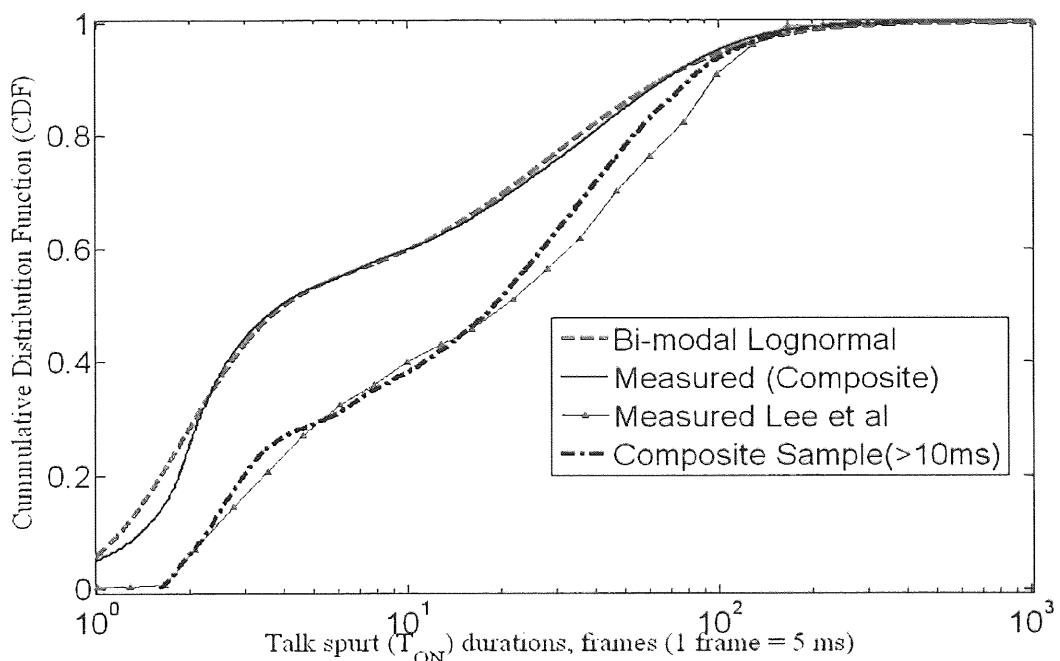


Figure 7.7: Cumulative distributions for  $T_{ON}$  periods – with a composite sample excluding silences less than 10 ms

## 7.4 Summary

In this chapter, we have looked at the composite statistics of the  $T_{ON}$  and  $T_{OFF}$  periods of VoIP traffic, and modelled the  $T_{ON}$  and  $T_{OFF}$  period densities with bi-modal and tri-modal log-normal distributions respectively confirming that these periods are heavy-tailed in nature. We then used the fitted distributions to propose an ON-OFF Semi-Markov source model with Log-normal transitions and studied the results of the simulated  $T_{ON}$  and  $T_{OFF}$  periods and highlighted the differences with results of other popular models which also included short silences in their statistics. Finally we also showed that if the short period detection is artificially cut off at 10 ms, we could reproduce the experimental measurements previously published.

## **PART – III**

### **Conclusions**

# Chapter Eight

## Conclusions and Future Work

In this concluding chapter, we present the conclusions of this thesis in section 8.1 with relevance to the key contributions presented in the preceding chapters, followed by recommendations and directions for future work, in section 8.2.

### 8.1 Conclusions

The following are the conclusions based on the key contributions of this thesis:

#### 8.1.1 Analysis of IPT statistics based on simple Markov traffic models with exponential and non-exponential holding times

In Chapter 2, we had explored the use of simple continuous-time Markov models in Monte Carlo simulations to generate and analyse IPT statistics, adopting a white box approach. In contrast to traditional approaches used with Markov traffic models, where the statistics are based on the packet events executed, we used a slightly different and innovative approach, where we associated the packet emissions with transitions and used the holding times of the states to generate the IPT statistics. In order to understand some of the most primitive features observed in the measured IPT results and their relation to traffic sources, we analysed the IPT statistics generated from some of the simplest and basic examples of the Markov traffic models. We started the analysis with a basic two-state Markov traffic model, moving upwards to n-state models with exponential holding times, confirming our analytical understanding of the models. We demonstrated that the IPT statistics with Poisson characteristics can be modelled by an n-state model, the analytical solution of which is given by the sum of the state holding time distributions weighted by their respective state visiting probabilities. We then extended the work further, by tweaking the models such that they can model the non-Poisson features of the measured IPT statistics. We demonstrated that the peaks observed in the measured network IPT statistics, can be modelled with n-state loop Markov Models where

only one state is emitting packets and that the analytical solution for the characteristic PDF of IPT of such models is given by the convolution of the state holding time distributions. Finally, using the n-state models with non-Poisson statistics as an analytical platform, we introduced the Gaussian Markov models and derived the analytical solution for these models. This work established a strong foundation for the rest of the thesis, contributing to the understanding of how these Markov models, can not only be used for modelling, but also for understanding how the traffic sources behave and contribute to the statistics.

### 8.1.2 Use of Gaussian Markov Models for modelling traffic

In Chapter 3, we looked at how the Gaussian Markov models can be used for modelling the measured IPT statistics. They are not only convenient to be used for modelling the first-order statistics, but they have also been proved to be very instrumental in modelling the features of the Joint Density results as was evident from Chapters 4 and 5. The proposed Gaussian Markov modelling technique follows a white box approach, where each significant part of the IPT statistics can be characterised by a traffic source and its network timing parameters. The strongest implication of these Gaussian Markov models, is that when used for analysing measured IPT statistics of the network, the key statistical components contributing to the statistics can be directly related to a state. This state represents a traffic source and has a Gaussian holding time distribution, because the IPT statistics can be broken down to Gaussian components, where each Gaussian component ideally represents a network's typical packet transmission delay interval parameter which is statistically dispersed around a mean value as is the characteristic property of network delays associated with a traffic source. This work has paved way for the Gaussian Markov models to be used for analysing and estimating a realistic number of traffic sources and their associated packet events that would have contributed to the IPT statistics, thus providing a transparent window to the physical approximation of the entire or part of the network structure under study.

### 8.1.3 Significance of the Joint Density results for understanding and modelling network IPT statistics

In Chapters 3, 4 and 5 we had established the importance of the Joint Density results by observing some unique features in these plots which revealed that there is a lot more information from these

plots which the first order statistics do not provide. One such kind of information was that of packet sequences as covered in Chapter 4. For example, a TCP packet in one direction is always likely to be followed by an ACK packet in the opposite direction and this is clearly not indicated by first order statistics, but can be observed only with the use of the Joint Density plots. Another key advantage with the use of these Joint Density results, as explained in this thesis earlier, is that the simple Gaussian Markov models can be used to model the unique features of the Joint Density results. In Chapter 4, we also proved the significance of the Joint Density results, with relation to packet sequences, by proving how two simple Markov traffic models with different topologies can have exactly the same probability density result and yet have different Joint Density results. We also revealed in Chapter 4, that the first order density results are just a projection of the Joint Density results which carry much more information on the underlying packet sequence of the model. Thus, by modelling traffic based on the first order density, the traffic modelling community merely models the first order statistics only and not the source behaviour of the network.

In Chapter 4, we also demonstrated that the measured IP traffic was symmetric in nature. This symmetric nature of the Joint Density results proved that for those particular traffic captures, the communicating protocol had packet transmissions such that any of the packets associated with the protocol can be followed by any of the packets in random order, and that the protocol did not have a strict sequence for the types of packets it used. This is again significant information about the protocol retrieved from the Joint Density results, and it helps in predicting what protocol was used for generating the traffic. In Chapter 4, we also showed how packet loss affects IPT statistics and proved that a high level of packet loss in the network symmetrises the IPT statistics.

In Chapter 5, we revealed the presence of another important feature in the Joint Density results, the *curve of periodicity*, induced by plotting the IPTs logarithmically and indicating the presence of periodic packet arrivals in the network. We observed its presence in the IPT results for traffic associated with Voice/Video telephony and any traffic consisting of bi-directional periodic packet streams. Using Gaussian Markov models again, we not only modelled and simulated these curve features of the Joint Density results, but also understood that the two types of curves that we observed in the results confirm that there are two sets of traffic sources sending periodic packet traffic streams with small and large differences in mean values of their packet interval distributions respectively. In another unique and significant contribution, we demonstrated with experimental proof, using Joint Density results and simple Markov models representing periodic processes, that certain types of IP traffic transmissions occur in phases, where one phase of traffic is described by the characteristics of one model and the other phase is described by the characteristics of another

model. Again the significance of the Joint Density results is stressed here as these kinds of unique features of Joint Density results give strong hints on the identification of the communication protocol and the application associated with the traffic captured.

#### **8.1.4 Markov modelling with Joint Density results: A powerful tool for analysis and an alternative to Deep Packet Inspection (DPI)**

Deep Packet Inspection (DPI) is a technique often used for network monitoring for collecting information on protocol and source behaviour and also for detecting network intrusion attempts. As the name suggests, DPI refers to a process where the IP packet is deeply inspected for the header and payload to extract, timing protocol and other necessary information for network monitoring. In order to regulate traffic management for networks with bandwidth constraints, network service providers engage in the use of Deep Packet Inspection to analyse and monitor services and user activities such as those related to P2P (Peer to Peer) applications. However this method of network monitoring has a lot of disadvantages. Firstly, they need sophisticated hardware and software for the inspection process. There is a significant amount of overhead in the network due to the inspection process. Finally, the most crucial of all, it invades the privacy of the user. Besides, the current generation, state of the art, encryption systems also provide an opportunity for network users to become immune to DPI in some form or the other. With respect to traffic and network monitoring, while DPI remains a sophisticated and popular option, we know that in terms of traffic analysis the timing structure of the packet traffic is more important than the content of the packet traffic. Large amount of work has been done on classifying traffic without the use of DPI, such as those by [102] [103] [104] [105]. However, the work in this thesis proposes the use of Markov models along with the Joint density plots which forms a very good alternative to study the timing structure of packet traffic without necessarily engaging in privacy intruding techniques of DPI. We pointed out in section 8.1.1, the advantages of using simple Markov models with a white box approach, and how they can be used along with the information from the Joint Density results (section 8.1.3) on factors such as packet sequences and periodic events, to measure and model the network sources. For example, as shown in Chapter 5, we could use simple Markov models and Joint Density results to identify that the results were exclusive to Voice and Video Telephony applications and involved the use of real time transmission streaming protocols. Thus it is far easier to filter out a particular kind of traffic, based on its statistical behaviour, rather than the use of expensive and less efficient approaches such as DPI.

### **8.1.5 Impact of languages, speech rate and prosodic structure on the ON-OFF characteristics of Voice traffic over IP networks**

In Chapter 6, exploiting the advantages of the use of silence suppression in Voice traffic in transmission, we presented accurate measurements of  $T_{ON}$  (On) and  $T_{OFF}$  (Off) periods associated with VoIP traffic measured from a statistically significant and large database of multi-lingual VoIP calls consisting of native speakers speaking in their native languages. A massive increase in the statistics of the  $T_{ON}$  and  $T_{OFF}$  periods was recorded compared to previously published work. A sophisticated and efficient FFT filtering technique was used for accurately measuring the smallest  $T_{ON}$  and  $T_{OFF}$  periods that can be detected and included in statistics. It was observed that while the amount of short  $T_{ON}$  and  $T_{OFF}$  periods were found to vary consistently with the varying prosodic structure of speech from language to language for the monolog samples, the ON-OFF period statistics for the Voice calls were found to be incredibly complex, as the statistics were dominated by the random dynamics of human conversation involving non-grammatical silences and talk spurts. It was confirmed that the statistics of  $T_{ON}$  and  $T_{OFF}$  periods for calls of the same languages differed as much as that of different languages. Also from the call recordings database, it was found that the speech rate of VoIP conversation differed significantly from one person to another, hence making it further difficult to study the exact impact of each language, prosodic structure and the speech rate of the speaker on the VoIP statistics. In the conclusion of this statistical study of ON-OFF periods from the multi-lingual database of VoIP calls, it was proposed that the composite statistical samples for the  $T_{ON}$  and  $T_{OFF}$  periods be used for modelling as they average out the linguistic and conversational behaviour differences. Also, in conclusion, it must be additionally pointed out that if enough attention is not given to the calibration of the cut-off frequency and threshold levels used in the silence discrimination process, this can affect the measured statistics of  $T_{ON}$  and  $T_{OFF}$  periods.

### **8.1.6 An ON-OFF model with Log-normal transitions and its significance**

In Chapter 7 of this thesis, a logarithmic time analysis of the composite statistics of the  $T_{ON}$  and  $T_{OFF}$  periods from the large multi-lingual database was performed which revealed that the composite statistics have a tri-modal aspect to the  $T_{OFF}$  periods and a bi-modal aspect to the  $T_{ON}$  periods. It was also proved that they are heavy tailed in nature. A simple five-state ON-OFF VoIP source model with Log-normal transitions was proposed, based on the accurate measurements of  $T_{ON}$  and  $T_{OFF}$  periods.

Key differences of results of the proposed Log<sub>e</sub>-normal model with the other results were highlighted including the inaccuracy of the detection of short silences and talk spurts by the other models which also considered short silences and talk spurts. As stated earlier, a massive increase in statistics for the analysis of the T<sub>ON</sub> and T<sub>OFF</sub> periods of VoIP activity was produced. The composite data was averaged over 100 hours of call data, with all sorts of anomalies that can be found in a typical call, including hesitational silences, fast and slow speech rates, and non-grammatical silences and talk spurts. Thus, the proposed model will closely fit any call sample as long as silences as low as 5 ms are detected. From the results of the model, an increase of 8% in the T<sub>OFF</sub> periods smaller than 100 ms and a 20% increase of T<sub>ON</sub> periods smaller than 100 ms was recorded when compared to previously published work. It was also recorded that around 67% of the T<sub>OFF</sub> periods are below 10 ms, which is crucial as an implication to the field of design and implementation of VoIP codecs supporting silence suppression. The measured T<sub>OFF</sub> periods statistics are of immediate use for further research on VoIP traffic modelling and on the choice and implementation of voice codecs supporting silence suppression. It is recommended that the high percentage of short T<sub>OFF</sub> periods measured, should be exploited by new generation codecs which have a frame rate as low as 5 ms, thus further enhancing capacity by increasing bandwidth efficiency.

## **8.2 Future Work**

While this thesis has provided a breakthrough for an in-depth analysis of IP traffic IPT statistics using simple Markov Traffic models and first and second-order density plots, there is still some amount of work, that is outside the scope of this thesis, which can be done to extend the work further and make it more interesting. A few directions for future work are given as follows:

### **8.2.1 The fourth dimension: Packet Sizes (PS)**

Traffic modelling involves the use of two main kinds of parameters for modelling and understanding the network behaviour: namely the packet sizes (PS) and the inter-packet times (IPT). Packet Sizes (PS) are a characteristic parameter of the network which defines the sizes of each packet in bytes. These Packet Sizes (PS) vary with the protocol and application being used in the network. In the beginning of the thesis, we had mentioned that the scope of the work for this thesis has permitted the measurement, analysis and modelling of network Inter-packet times (IPT) only. To exploit further the advantages of the simple Markov models and the Joint Density results as seen in Chapters 3-5,

the problem can now be made four dimensional, where the statistics of Packet Sizes (PS) can also be introduced. The vector of simulated or measured Packet Sizes (PS)  $b_n$ , can also be right cyclic shifted to obtain  $b_{n+1}$ , which can be then used for generating a Joint Density plot, by plotting the density of  $b_n$  and  $b_{n+1}$  against each other. This then gives us information that a packet of size ' $b_1$ ' is followed by a packet of size ' $b_2$ '. The work on packet sequences as presented in Chapter 4, can also incorporate Packet Sizes (PS) to determine what is likely the next packet after a given packet. The Gaussian Markov models proposed in this thesis for modelling the first order and second order densities of IPT can also be conveniently modified, such that their state holding time distributions can be represented by the distributions of packet sizes with a mean  $\mu$ . This is of course very practical, because in a real network the packet sizes from a source are not always constant, and vary due to various factors, such as errors on links, payloads and transfer medium. The analysis can be made further interesting by plotting any of the four combinations:  $t_n$ ,  $t_{n+1}$ ,  $b_n$  and  $b_{n+1}$  against each other. This opens up further boundaries for analysis of a variety of other network parameters such as throughput, network load factor, bandwidth efficiency etc.

### 8.2.2 Applying the use of Joint Density results and Markov models to model and understand specific networking protocols

An important reason as to why the measured IPT statistics considered in this thesis, were filtered only by the IP addresses and the Transport Layer protocol with out any emphasis on the Application Layer protocols, is because the generic nature of the IPT results needed to be preserved so that the conclusions on the analysis of these results and associated models, would be independent of the application and protocol. However, given that this thesis has established the significance of the Joint density results and the white box approach Markov models, the analysis of IPT statistics can now be applied to specific Application Layer protocols such as Hypertext Transfer Protocol (HTTP), Session Initiation Protocol (SIP), File Transfer protocol (FTP) and Simple Mail Transfer Protocol (SMTP). For example, when considering the HTTP protocol, the operation and typical timing structure of the packet events between a selected number of sources can be easily studied with reference to RFCs associated with the protocol. This gives us the pre-established knowledge necessary to understand the packet sequences from measured IPT statistics of a traffic capture filtered to a HTTP connection. A network *ping* from the host machine to the server at the time of measurement, gives the typical Round Trip times between the HTTP server and the HTTP client and this can aid in further relating the IPT peaks to the states of the physically approximate model for the HTTP connection which can be

produced by using Gaussian Markov models or Poisson Markov models depending on the nature of the statistics. Each of the states of these models can represent the nodes involved in the HTTP connection and the model can be designated as being specific to a HTTP connection. This can be similarly repeated for other protocols for producing other protocol specific models. The analytical understanding on the operation of these models can be then verified with the IPT statistics of the measured results by comparing the measured and modelled results. Once the models are deemed as realistically accurate, the possibility for associating protocol specific features in statistics of large complex internet connections to these specific models becomes a reality and thus further alleviating the need of complex approaches such as Deep Packet Inspection (DPI).

This thesis ends with a note of recommendation to the traffic modelling community, on the use of Joint Density Function along with the Probability density results for modelling and understanding IP traffic in the most efficient manner, using simple Markov models. A lot of information can be achieved with a combination of these, compared to the other complex traffic analysis methods.

## List of Publications Related to the Thesis

- [1] A. D. Shaikh, M. Eberhard, K. J. Blow and S. A. Fowler, *Statistical Study of Talk Spurts and Silence Lengths with relation to various Languages and Speakers for Advanced Silence Suppression in Next Generation Wireless Communication Devices*, In Proc. SCIENT2008, page P-2, Suita, Osaka, Japan, August 2008, Osaka University
- [2] A. D. Shaikh, Y. Morioka, *Assessing Quality of VoIP Calls in a Wireless LAN Environment (IEEE 802.11b)*, In Proc. AMP/SCIENT 2008, page AMP-15, Suita, Osaka, Japan, September 2008, Osaka University
- [3] A.D.Shaikh, M.Eberhard, K.J.Blow and S.A.Fowler, *Traffic modelling using simplified Markov Traffic Models*, Poster presented at poster session, Next Generation Networks meeting, Apr. 2009, Essex University, Colchester, U.K.
- [4] A. D. Shaikh, M. Eberhard, K. J. Blow and S. A. Fowler, *Significance of Second Order statistics in Markov Internet Traffic Modelling*, In Mathematics for Networks-8, MON-8 Conference, Sept. 2009, Cambridge, U.K.
- [5] A. D. Shaikh, M.A. Eberhard, K. J. Blow and S. A. Fowler, *Characterising OFF periods of a VoIP Traffic source by a Tri-modal Lognormal Distribution*, IET/IEEE Electronics Letters, (in review).
- [6] A. D. Shaikh, M. Eberhard, K. J. Blow and S. A. Fowler, *An ON-OFF VoIP Source Model with Lognormal Transitions*, In IEICE Transactions on Communications, (in review). [A longer elaborate version of this paper is proposed to be submitted to International Journal of Signal Processing]
- [7] A. D. Shaikh, M. Eberhard, K. J. Blow and S. A. Fowler, *Significance of Joint Density results in Markov Internet Traffic Modelling*, IEEE Communications Letters, (under preparation). [Another longer form is being prepared to be submitted to 'Journal of Mathematical and Computer Modelling (Elsevier)']

# List of References

- [1] Advanced Micro Devices, Inc. (2009, Aug.) AMD 50x15 - The World Internet Usage. [Online]. [http://www.50x15.com/en-us/internet\\_usage.aspx](http://www.50x15.com/en-us/internet_usage.aspx)
- [2] J. R. Okin, "Packet Switching: Lifeblood of the Internet," in *The Internet Revolution*. Maine, USA: Ironbound Press, 2005, vol. 1, ch. 4, p. 134.
- [3] J. Postel, "Internet Protocol," RFC 791, 1981.
- [4] International Telecommunications Union, "ISO - Open Systems Interconnection - Basic Reference Model," X.200 Series Recommendations, 1994.
- [5] R. Braden, "Requirements for Internet Hosts -- Communication Layers, RFC 1122," Internet Engineering Task Force RFC 1122, 1989.
- [6] N. Wilkinson, "Packet Switched Technologies," in *Next Generation Network Services - Technologies and Strategies*. Chichester, England: Wiley, 2002, ch. 4, pp. 53-56.
- [7] A. K. Erlang, "The theory of Probabilities and Telephone Conversations," *Nyt Tidsskrift for Matematik*, vol. B, no. 20, pp. 33-39, 1909, Translated in: The Life and Work of A.K.Erlang.
- [8] V. S. Frost and B. Melamed, "Traffic modeling for telecommunications networks," *Communications Magazine, IEEE*, vol. 32, no. 3, pp. 70-81, Mar. 1994.
- [9] M. K. P, V. M. Gadre, and U. B. Desai, "Teletraffic Modeling: A historical Perspective," in *Multifractal Based Network Traffic Modeling*. Kluwer Academic Publishers, 2003, ch. 1, pp. 4-6.
- [10] L. Kleinrock, "Queuing System," in *Computer Applications*. John Wiley & Sons, 1978.
- [11] J. R and R. S, "Packet Trains--Measurements and a New Model for Computer Network Traffic," *Selected Areas in Communications, IEEE Journal on*, vol. 4, no. 6, pp. 986-995, Sep. 1986.
- [12] R. Gusella, "A measurement study of diskless workstation traffic on an Ethernet ,," *Communications, IEEE Transactions on*, vol. 38, no. 9, pp. 1557-1568, Sep. 1990.
- [13] W. Willinger and V. Paxson, "Where Mathematics meets the Internet," *Notices of the American Mathematical Society*, vol. 45, pp. 961-970, 1998.
- [14] V. Paxson and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling," *IEEE/ACM Transaction on Networking*, vol. 3, pp. 226-244, Jun. 1995.
- [15] A. Adas, "Traffic models in broadband networks," *Communications Magazine, IEEE*, vol. 35, no. 7, pp. 82-89, Jul. 1997.
- [16] W. Willinger, V. Paxson, and M. S. Taqqu, "Self-Similarity and Heavy Tails: Structural Modeling of

Network Traffic," in *Statistical Techniques and Applications*. Verlag, pp. 27-53.

- [17] J. Gordon, "Pareto process as a model of self-similar packet traffic," in *Global Telecommunications Conference, 1995. GLOBECOM '95., IEEE*, 1995, pp. 2232-2236.
- [18] C. -N. Chuah and R. H. Katz, "Characterizing Packet Audio Streams from Internet Multimedia Applications," *Proc. IEEE Int. Conf. on Communications (ICC 2002)*, vol. 2, pp. 1199-1203.
- [19] B. Chandrasekaran, "Survey of Network Traffic Models," Washington University in St.Louis, 2006.
- [20] G. Samorodnitsky, "Long Range Dependence," *Foundations and Trends in Stochastic Systems*, vol. 1, no. 3, pp. 163-257, 2006.
- [21] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet Traffic," *Proceedings of ACM SIGCOMM'93*, pp. 183-193, 1993.
- [22] A. Erramilli, O. Narayan, and W. Willinger, "Experimental queueing analysis with long-range dependent packet traffic," *IEEE/ACM Transactions on Networking*, vol. 4, no. 2, Apr. 1994.
- [23] J. Cao, W. S. Cleveland, D. Lin, and D. X. S. Unknown, "Internet traffic tends toward Poisson and independent as the load increases," *Nonlinear Estimation and Classification*, 2002.
- [24] W. Willinger, M. S. Taqqu, R. Sherman, and D.V.Wilson, "Self similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source levele," *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, Feb. 1997.
- [25] M. T. L. e. al, "Statistical Characterization of Wide-area IP traffic," *Proc. of 6th International conference on Computer Communications and Networks*, 1997.
- [26] J. Beran, R. Sherman, and W. Willinger, "Long-range dependence in VBR video traffic," *IEEE Transactions on Communications*, vol. 43, no. 2, pp. 1566-79, 1995.
- [27] L. Yao, M. Agapie, J. Ganbar, and M. Doroslovacki, "Long Range Dependence in Internet Backbone Traffic," *Proceedings of IEEE International Conference on Communications, ICC'03*, vol. 3, pp. 1611-1615, May 2003.
- [28] T. Kariagiannis, M. Molle, and M. Faloutsos, "Long-Range Dependence: Ten Years of Internet Traffic Modeling," *IEEE Internet Computing*, vol. 8, no. 5, pp. 57-64, Oct. 2004.
- [29] H. Heffes and D. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance," *Selected Areas in Communications, IEEE Journal on*, vol. 4, no. 6, pp. 856-868, Sep. 1986.
- [30] L. Muscariello, M. Meillia, M. Meo, M. A. Marsan, and R. L. Cigno, "An MMPP-based hierarchical model of Internet traffic," *Communications, 2004 IEEE International Conference on*, vol. 4, pp. 2143-2147, Jun. 2004.

- [31] S. Shah-Heydari and T. Le-Ngoc, "MMPP models for multimedia traffic," *Telecommunication Systems*, vol. 15, no. 4, pp. 273-293, Nov. 2000.
- [32] S. Shah-Heydari and T. Le-Ngoc, "MMPP modeling of aggregated ATM traffic," *Electrical and Computer Engineering, 1998. IEEE Canadian Conference on*, vol. 1, pp. 129-132, May 1998.
- [33] N. Chee-Hock and S. Boon-Hee, *Queuing Modelling Fundamentals - With Applications in Communication Networks*, 2nd ed. John Wiley and Son, 2008.
- [34] H. Kobayashi and B. L. Mark, *System Modeling and Analysis: Foundations of System Performance Evaluation*. Prentice Hall, 2009.
- [35] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains," *The Annals of Mathematical Statistics*, vol. 41, no. 1, pp. 164-171, Feb. 1970.
- [36] D. Bruckner, B. Sallans, and G. Russ, "Hidden Markov Models for Traffic Observation," *Industrial Informatics, 2007 5th IEEE International Conference on*, vol. 2, pp. 989-994, Jun. 2007.
- [37] A. Dainotti, et al., "QRP07-2: An HMM Approach to Internet Traffic Modeling," in *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, 2006, pp. 1-6.
- [38] C. Wright, F. Monrose, and G. M. Masson, "HMM profiles for network traffic classification," in *Proceedings of the 2004 ACM Workshop on Visualization and Data Mining For Computer Security (Washington DC, USA, October 29 - 29, 2004). VizSEC/DMSEC '04. ACM*, New York, 2004, pp. 9-15.
- [39] N. Hohn, "Measuring, Understanding and Modelling Internet Traffic," University of Melbourne Thesis, 2004.
- [40] M. Menth, A. Binzenhofer, and S. Muhleck, "Source Models for Speech Traffic Revisited," *IEEE/ACM Transactions on Networking*, no. DOI 10.1109/TNET.2008.2006222, Jan. 2009.
- [41] P. Mockapetris, "Telephony's next act," *IEEE Spectrum*, vol. 43, pp. 28-32, Apr. 2006.
- [42] G. P. Basharin, A. N. Langville, and V. A. Naumov, "The life and work of A.A. Markov," *Linear Algebra and its Applications*, vol. 386, no. Special issue on the Conference on the numerical solutions of Markov Chains 2003, pp. 3-26, Jul. 2004.
- [43] M. Baron, "Markov Processes and Markov Chains," in *Probability and Statistics for Computer Scientists*. Chapman & Hall/CRC, 2007, ch. 6, p. 146.
- [44] Leonard Kleinrock, *Queueing Systems: Volume I – Theory*. New York: Wiley Interscience, 1975.
- [45] J. R. Norris, *Markov Chains*. Cambridge: Cambridge University Press, 1998.
- [46] D. Revuz, *Markov Chains*. Elsevier, 1984.

- [47] D. Freeman, *Markov Chains*. Holden-Day, 1971.
- [48] P. Bremaud, *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer, 1999.
- [49] A. B. Johnston, *SIP: understanding the Session Initiation Protocol*, 2nd ed. Artech House, 2004.
- [50] P. Olofsson, *Probability, Statistics and Stochastic Processes*, 1st ed. Wiley, 2005.
- [51] T. Karagiannis, M. Molle, M. Faloutsos, and A. Broido, "A nonstationary Poisson view of Internet traffic," *INFOCOM 2004. Proceedings of Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 3, pp. 1558-1569, Mar. 2004.
- [52] V. Jacobson, C. Leres, and S. Mccane, "The TCPDUMP Manual Page," Lawrence Berkely Laboratory Manual, 1989.
- [53] S. Mccane and V. Jacobson, "The BSD packet filter: A New Architecture for User-level Packet Capture," *Proceedings of the 1993 Usenix Conference, San Diego*, pp. 259-269, 1993.
- [54] W. R. Stevens, "The Tcpdump Program," in *TCP/IP Illustrated Volume 1: The Protocols*. Addison-Wesley, 1994, ch. Appendix A.
- [55] S. B. Shamsuddin and M. E. Woodward, "Modeling Protocol Based Packet Header Anomaly Detector for Network and Host Intrusion Detection Systems , " in *Lecture Notes in Computer Science*. Berlin: Springer, 2007.
- [56] F. Fuentes and K. D.C., "Ethereal vs. Tcpdump: a comparative study on packet sniffing tools for educational purposes," *Journal of Computing Sciences in College*, vol. 2, no. 4, pp. 169-176, 2005.
- [57] M. Allman and V. Paxson, "On estimating end-to-end network path properties., " in *Proceedings of the Conference on Applications, Techonologies, Architectures, and Protocols For Computer Communication* , New York, 1999, pp. 263-274.
- [58] G. Combs. (2009, Aug.) Wireshark - Go deep.. [Online]. <http://www.wireshark.org/>
- [59] A. Dabir and A. Matrawy, "Bottleneck Analysis of Traffic Monitoring using Wireshark," in *Proceedings of 4th International Conference on Innovations in Information Technology*, Dubai, 2007, pp. 158-162.
- [60] G. Munz and G. Carle, "Distributed Network Analysis Using TOPAS and Wireshark," in *Proceedings of Network Operations and Management Symposium Workshops, NOMS 2008*, Salvador do Bahia, 2008, pp. 161-164.
- [61] V. Jacobson, C. Leres, and M. S., "Libpcap: Packet Capture Library," Lawrence Berkely Laboratory Release Statement, 1994.
- [62] S. Mack, *Streaming Media Bible*. Hungry Minds, 2002.

- [63] British Broadcasting Corporation. BBC IPlayer Radio Service. [Online].  
<http://www.bbc.co.uk/iplayer/radio>
- [64] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, pp. 1-38, 1977.
- [65] T. K. Moon, "The expectation maximization algorithm," *Signal Processing Magazine, IEEE*, vol. 13, no. 6, pp. 47-60, 1996.
- [66] C. P. Robert and G. Casella, *Monte Carlo statistical methods*, 2nd ed. Springer, 2004.
- [67] C. T. Calafate, P. Manzoni, J.-C. Cano, and M. P. Malumbres, "Markovian-based traffic modeling for mobile ad hoc networks," *Computer Networks*, vol. 53, no. 15, pp. 2586-2600, Sep. 2009.
- [68] A. Reyes-Lecuona, E. González-Parada, E. Casilar, J. C. Casasola, and A. Díaz-Estrella, "A page-oriented WWW traffic model for wireless system simulations," *Teletraffic science and engineering*, vol. 3, pp. 1271-1280, Jun. 1999.
- [69] Z. Liu, J. Almhana, V. Choulakian, and R. McGorman, "Online EM algorithm for mixture with application to internet traffic modeling," *Computational Statistics & Data Analysis*, vol. 50, no. 4, pp. 1052-1071, Feb. 2006.
- [70] B. Chen, Y. Xu, J. Hu, and Ling Zhang, "Modeling and Analysis Traffic Flows of Peer-to-Peer Application," *Innovative Computing Information and Control, 2008. ICICIC '08. 3rd International Conference on*, pp. 383-383, 2008.
- [71] F. Qian, G.-m. Hu, and X.-m. Yao, "Semi-supervised internet network traffic classification using a Gaussian mixture model," *AEU - International Journal of Electronics and Communications*, vol. 62, no. 7, pp. Pages557-564, Aug. 2008.
- [72] M. Menth, A. Binzenhofer, and S. Muhleck, "Source Models for Speech Traffic Revisited," *IEEE/ACM Transactions on Networking*, vol. 17, no. 4, pp. 1042-1051, Aug. 2009.
- [73] P. T. Brady, "A Technique for investigating On-Off Patterns of Speech," *Bell Systems Technical Journal*, vol. 44, no. 1, pp. 1-22, Jan. 1965.
- [74] P. T. Brady, "A Statistical Analysis of On-Off Speech Patterns in Two-Way Conversations," *Bell Systems Technical Journal*, vol. 47, no. 1, pp. 73-91, Jan. 1968.
- [75] P. T. Brady, "A Model for Generating On-Off Patterns in 16 conversations," *Bell Systems Technical Journal*, vol. 48, no. 9, pp. 2445-2472, Sep. 1969.
- [76] K. Sriram and W. Whitt, "Characterizing Superposition Arrival Processes in Packet Multiplexers for Voice and Data," *IEEE Journal on Selected Areas in Communications*, vol. 4, no. 6, pp. 833-846, Sep. 1986.
- [77] ITU-T, "P.59: Telephone Transmission Quality Objective Measuring Apparatus: Artificial

Conversational Speech," Recommendation, 1993.

- [78] J. G. Gruber, "A Comparison of Measured and Calculated Speech Temporal Parameters Relevant to Speech Activity Detection," *IEEE Transactions on Communications*, vol. 30, no. 4, pp. 728-738, Apr. 1982.
- [79] H. P. Stern, S. A. Mahmoud, and K. .-K. Wong, "A Comprehensive Model for Voice Activity in Conversational Speech - Development and Application to Performance Analysis of New-Generation Wireless Communication Systems," *Wireless Networks*, vol. 2, no. 4, pp. 359-367, Dec. 1996.
- [80] H. P. Stern, S. A. Mahmoud, and K. .-K. Wong, "A Model for Generating On-Off Patterns in Conversational Speech, including short silence gaps and the effects of interaction between parties," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 4, pp. 1094-1100, Nov. 1994.
- [81] H.H.Lee and C. K. Un, "A Study of On-Off Characteristics of Conversational Speech," *IEEE Transactions on Communications*, vol. 34, no. 6, pp. 630-637, Jun. 1986.
- [82] T.D.Dang, B. Sonokoly, and S. Molnar, "Fractal Analysis and modeling of VoIP Traffic," in *Networks 2004*, Vienna, 2004, pp. 123-310.
- [83] C. N. Chuah and R. H. Katz, "Characterizing Packet Audio Streams from Internet Multimedia Applications," in *Proceedings of IEEE ICC 2002*, New York, 2002, pp. 1199-2203.
- [84] H. M. a. F. S. E. Casilar, "Modelling of Voice Traffic Over IP Networks," in *Proceedings of COMMUNICATION SYSTEMS, NETWORKS AND DIGITAL SIGNAL PROCESSING 2002*, Staffordshire, 2002, pp. 411-414.
- [85] P. Pragtong, K. M. Ahmed, and T. J. Erke, "Analysis and Modeling of Voice over IP Traffic in the Real Network," *IEICE Transactions on Information and Systems*, vol. E89-D, no. 12, pp. 2886-2896, Dec. 2006.
- [86] B. Macwhinney. (2007) The TalkBank Project; Creating and Digitizing Language Corpora: Synchronic Databases. [Online]. <http://talkbank.org/media/CABank/CallFriend/>
- [87] T. V. t. :. T. Incorporated. The Open Speech Repository. [Online]. [http://www.voiptroubleshooter.com/open\\_speech/index.html](http://www.voiptroubleshooter.com/open_speech/index.html)
- [88] A. T. F. A. E. a. Recorder. [Online]. <http://audacity.sourceforge.net/>
- [89] Sox - Sound Exchange, Audio Editor. [Online]. <http://sox.sourceforge.net/>
- [90] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. New Jersey: Englewood Cliffs, Prentice-Hall, 1975.
- [91] M. Frigo and S. G. Johnson. (2005) Fastest Fourier Transform in the West - FFTW3. [Online]. <http://www.fftw.org/>

- [92] D. H. Sanes, T. A. Reh, and W. A. Harris, "Behavioral Development," in *Development of the Nervous System*. Academic Press, 2006, ch. 10, p. 303.
- [93] Cisco Systems. (2009) Cisco Systems: Voice Over IP - Per Call Bandwidth Consumption. [Online]. [http://www.cisco.com/en/US/tech/tk652/tk698/technologies\\_tech\\_note09186a0080094ae2.shtml](http://www.cisco.com/en/US/tech/tk652/tk698/technologies_tech_note09186a0080094ae2.shtml)
- [94] B. Goode, "Voice Over Internet Protocol (VoIP)," *Proceedings of the IEEE*, vol. 90, no. 9, pp. 1495-1517, Sep. 2002.
- [95] B. Zellner, "Pauses and the temporal structure of speech," in *Fundamentals of speech synthesis and speech recognition*. Chichester: John Wiley, 1994, pp. 41-62.
- [96] J. Cao, "The Rhythm of Mandarin Chinese," *ournal of Chinese Linguistics, Monograph Series*, vol. 17, 2002.
- [97] E. Campione and J. Veronis, "A Large-Scale Multilingual study of Silent Pause Duration," in *Proc. Speech Prosody*, Aix-en-Provence, France , 2002, pp. 199-202.
- [98] F. J. Massey, "The Kolmogorov-Smirnov Test for Goodness of Fit," *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68-78, 1951.
- [99] C. John, Y. Collin, and F. Janet, *Introduction to Phonetics and Phonology*. Oxford: Blackwells, 2007.
- [100] M. Demol, W. Verhelst, and P. Verhove, "The Duration of Speech Pauses in a Multilingual Environment," in *Proc. Interspeech 2007*, Antwerp, Belgium, 2007, pp. 27-31.
- [101] V. A. Bolotin, "Modeling Call holding time distributions for CCS network design and performance analysis," *IEEE Journal of Selected Areas in Communication*, vol. 12, no. 3, pp. 433-438, Apr. 1994.
- [102] A. W. Moore, "Internet traffic classification using Bayesian analysis techniques," *Proceedings of ACM SIGMETRICS - Performance Evaluation Review*, vol. 33, no. 1, pp. 50-60, 2005.
- [103] T. c. u. c. algorithms, "Jeffrey Erman; Martin Arlitt; Anirban Mahanti," *Proceedings og SIGCOMM 2006*, 2006.
- [104] T. c. o. t. fly, "L. Bernaille; R. Teixeria; I. Akodkenou; Kave Salamatian," *Proceedings ACM SIGCOMM 2006*, vol. 36, no. 2, pp. 23-26, 2006.
- [105] D. Zuev and A. W. Moore, "Traffic classification using a statistical approach," *Passive and Active Network Management*, pp. 321-324, 2005.
- [106] M. R. Crovella and A. Bestavros, "Self-similarity in World Wide Web Traffic: evidence and possible causes," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835-846, Dec. 1997.