Packet Level Measurement over Wireless Access

By

Syeda Samana Naqvi

SUBMITTED FOR THE DEGREE OF MASTER OF PHILOSOPHY

Supervised by Dr. John Schormans

Department of Electronic Engineering

Queen Mary University of London

February 2011

Dedicated to

My Family

Authors Declaration

This is to certify that

1. The work presented in this thesis is the author's own.
2. Due acknowledgements has been made in the text to all other material used.
3. The thesis is less than 60,000 words in length (including footnotes), exclusive of tables, maps, bibliographies and appendices.

Syeda Samana Naqvi

Department of Electronic Engineering

Queen Mary University of London

February 2011

Abstract

Performance Measurement of the IP packet networks mainly comprise of monitoring the network performance in terms of packet losses and delays. If used appropriately, these network parameters (i.e. delay, loss and bandwidth etc) can indicate the performance status of the network and they can be used in fault and performance monitoring, network provisioning, and traffic engineering. Globally, there is a growing need for accurate network measurement to support the commercial use of IP networks. In wireless networks, transmission losses and communication delays strongly affect the performance of the network. Compared to wired networks, wireless networks experience higher levels of data dropouts, and corruption due to issues of channel fading, noise, interference and mobility. Performance monitoring is a vital element in the commercial future of broadband packet networking and the ability to guarantee quality of service in such networks is implicit in Service Level Agreements.

Active measurements are performed by injecting probes, and this is widely used to determine the end to end performance. End to end delay in wired networks has been extensively investigated, and in this thesis we report on the accuracy achieved by probing for end to end delay over a wireless scenario. We have compared two probing techniques i.e. Periodic and Poisson probing, and estimated the absolute error for both. The simulations have been performed for single hop and multi- hop wireless networks.

In addition to end to end latency, Active measurements have also been performed for packet loss rate. The simulation based analysis has been tried under different traffic scenarios using Poisson Traffic Models. We have sampled the user traffic using Periodic probing at different rates for single hop and multiple hop wireless scenarios.

Active probing becomes critical at higher values of load forcing the network to saturation much earlier. We have evaluated the impact of monitoring overheads on the user traffic, and show that even small amount of probing overhead in a wireless medium can cause large degradation in network performance. Although probing at high rate provides a good estimation of delay distribution of user traffic with large variance yet there is a critical tradeoff between the accuracy of measurement and the packet probing overhead. Our results suggest that active probing is highly affected by probe size, rate, pattern, traffic load, and nature of shared medium, available bandwidth and the burstiness of the traffic.

Acknowledgements

First of all I would like to thanks Allah Almighty for giving me strength and courage to complete this thesis. Along with the challenges of this independent research besides motivation and dedication, there are some key people in both my academic and personal life that had leaded the completion of this work.

My sincerest "Thanks" is to my supervisor, Prof John Schormans for his understanding and encouragement. I would like to thanks everybody in Electronics Department at Queen Mary University of London for creating a nice friendly working environment.

I would like to thank my parents who have been very supported and encouraging. I appreciate them for their never ending prayers and blessings for me. To my brothers Ali, Hassan, and Touseef who always helped me to keep my morale high. A Big thanks to my best friends Moneeza and Sadaf who have been in contact with me all this time.

The last but not least, my special thank-you to my dearest husband, Dr Qalab e Abbas, though he was far away in china for studies, still helped me through the difficult times and gives me his unconditional love and support all this time. Thanks Skype to make me talk to him every day☺.

Syeda Samana Naqvi

Table of Contents

List of Figures

List of Tables

Glossary

ACK	Acknowledgement
ATM	Asynchronous Transfer Mode
CAC	Connection Admission Control
CMIP	Common Management Internet Protocol
CSMA/CA	Carrier Sense Multiple Access protocol with Collision Avoidance
CTS	Clear to Send
CW	Contention Window
DCF	Distributed Coordination Function
DIFS	DCF Inter Frame Space
DR	Decay Rate
DSDV	Destination-Sequenced Distance-Vector Routing
DSR	Dynamic Source Routing
DSSS	Direct Sequence Spread Spectrum
EIFS	Extended inter frame space
FIFO	First In First Out
FHSS	Frequency Hopping Spread Spectrum
ICMP	Internet Control Message Protocol
IETF	Internet Engineering Task Force
IFS	Inter Frame Space
IFQ	Interface Queue
IGMP	Internet Group Management Protocol
IP	Internet Protocol
IPPM	Internet Protocol Performance Metrics
ITU-T	Internet Telecommunication Unit - Telecommunication
LRD	Long Range Dependent
LL	Link Layer
MAC	Medium Access Control
MBAC	Measurement Based Admission Control
MIB	Management Information Base

MIMO	Multiple Input Multiple Output
MMPP	Markov Modulated Poisson Process
MPLS	Multi Protocol Label Switching
NAM	Network Animator
NAV	Network Allocation Vector
NS	Network Simulator
NSP	Network Service Provider
OFDM	Orthogonal Frequency Division Multiplexing
OWAMP	One Way Active Measurement Protocol
PASTA	Poisson Arrival See Time Averages
PCF	Point Coordination Function
PGM	Probe Gap Model
PRM	Probe Rate Model
PLP	Packet Loss Probability
PMA	Passive Measurement and Analysis
QoS	Quality of Service
RTD	Round Trip Delay
RTP	Real Time Protocol
RTT	Round Trip Time
RTS	Request to Send
SAA	Service Assurance Agent
SLA	Service Level Agreement
SNMP	Simple Network Management Protocol
SIFS	Short Inter Frame Space
SRD	Short Range Dependent
TORA	Temporally Ordered Routing Algorithm
TTL	Time to Live
TWAMP	Two Way Active Measurement Protocol
UDP	User Datagram Protocol
VoIP	Voice over Internet Protocol
WLAN	Wireless Local Area Network

1. Introduction

1.1 Research Motivation

The fast increase in both the volume of traffic and variety of applications available on the internet, in combination with the severe competition among network service providers, is making traffic performance monitoring increasingly essential to network management.

Network monitoring also provides crucial information that can be used for network redesign and network management including capacity planning. Examples of network management include the characterization of current utilization for explicit admission control and the configuration of a scheduler to divide link-capacity among competing traffic classes. Without measurements, these management techniques have relied upon the accurate characterization of traffic — without accurate traffic characterization, network resources may be under or over utilized [4]. Using measurements can remove the need for complicated and incomplete models of network traffic. Measurement-based schemes can be used where no model is available.

End-to- end delay is a dominant network "health indicator" in any network. It captures service degradation due to congestion, link failure and routing anomalies. Accurate delay statistics are of huge importance to both customers and ISP's. Delay has been a key parameter in Service Level Agreements (SLAs) between Internet Service Provider (ISP) and customers [1].

Offering and maintaining delay guarantees is a difficult task in wireless networks and more generally in multi-hop wireless networks. Recently, the work on quality of service (QoS) in ad-hoc network has attracted a lot of attention. Most of research e.g. [8] [9] assumes that the core technology used in these networks is the IEEE 802.11, as this technology is widely available, less expensive, and provides distributed radio medium access that can be easily employed in ad-hoc networks. The random radio medium access provided by IEEE 802.11 makes the radio medium sharing difficult in a Multihop context [2]. Much of the work focus on QoS of IEEE 802.11 based ad-hoc a network is aimed at providing either throughput guarantees or delay guarantees or both. While monitoring such networks questions arise like how often delay statistics should be measured? What metric(s) capture the network delay as well as loss performance in a meaningful way? How do we implement these metrics to ensure limited impact on the performance of the network?

In general, methods for measuring network performance are divided into two types: active and passive. Passive measurement is a means of tracking the performance and behaviour of the user traffic without creating additional traffic or modifying existing traffic. It is implemented by integrating into the network devices to enable identification and recording of the characteristics of user traffic. This method is popular but it requires access to network resources which might not be possible. In active probing, probe packets are sent and measurements of performance is based on the data e.g. delay or loss ratio of those probe packets. Active measurement is easy to deploy, e.g. in the world-wide deployment of monitoring host for the internet end-to-end Performance project [3].

Authors in [4] [5] [6] have published research findings that aim to bound the accuracy achievable when probing for delays and loss. They have concluded that probing is prone to considerable error. [7] Reports research to measure packet loss probability (PLP) across the network within a predefined measurement error. All these researches have been carried out for wired scenarios. This motivated us to investigate the applicability of active probing over wireless access. Therefore, in this thesis we aim to study the limitation of probing over a wireless access to measure the delay distribution.

1.2 Objectives of Thesis

The overall aim is to investigate the accuracy of probing for delay distribution and packet loss.

The objectives of my research are:

- The main objective of this report is to analyze active measurement methodology for delay performance evaluation. The effectiveness of the measurement scheme is examined through extensive simulation of single and multi hops scenarios.

- To develop code in NS2 to capture the end to end delay distribution for data as well as probe traffic over single as well as multiple hops. We validated the accuracy of measurement by making use of sampling theory, and calculated absolute error for each probing example.

- To investigate the effect of data traffic (packet) size, probe size, probe pattern, probe rate and traffic type on active measurement.

- To explore how effectively we monitor packet loss rate using active probing.

1.3 Research Contribution

The work reported in this thesis is novel. The main contributions are:

- Using active probing technique at different probing rates, we evaluated the accuracy of measuring end to end delay distribution and packet loss rate of user traffic.

- For different probing rate employed, we observed that probing rate of 5 probes per second gave us more accurate results.

- The packet size of the probes must be of same size of user traffic.

- Our research suggests that active probing provides effective results for single hop wireless network. The absolute error in measurement increases to unacceptable extent, when number of hops increases, mainly due to interference caused by neighbouring nodes and wireless medium contention.

1.4 Thesis Layout

Chapter 2 discusses the importance of network monitoring, different measurement techniques, features of measurement methods and their applications. The two basic measurement schemes: active and passive measurement and their variants found in industry are also studied in this chapter. The merits and shortcomings of the two schemes are discussed.

Chapter 3 provides a literature review, and highlights related work in measurement specifically for delays over Ad hoc networks. We have also overviewed IEEE 802.11 standard, MAC layer legacy, end to end measurement issues, components of delay and the challenges involved in performing active monitoring over wireless access.

Chapter 4 presents the techniques used in research to simulate packet networks. It briefly explains simulation models, i.e. static and dynamic simulation models, deterministic and stochastic simulation models, continuous and discrete simulation models. We have report on different types of traffic models used in network simulations e.g. IPP, MMPP, ON/OFF models. We discuss the simulation tool "Network Simulator 2" (NS2) used in our research. We highlight the queuing behavior for a buffer with multiplexed Markovian traffic. In the end of chapter 4 we present a simulation based comparison for queue length and decay rates for wired and wireless scenarios.

Chapter 5 discusses the active probing mechanism, types of sampling techniques e.g. simple random sampling, stratified random sampling, systematic/periodic sampling, and the effect of data packet size and probe packet size on our monitoring technique. We access the accuracy of active probing for determining the end-to-end delay distribution of a single hop wireless network and extended our investigation to multi-hop networks. We compare the results obtained from both probing techniques used in our in simulations i.e. Periodic and Poisson probing, and estimate the accuracy of these measurements by making use of sampling theory.

Chapter 6 we used active probing to calculate the packet loss rate over wireless access. We compared the results for different probing rates and calculate the absolute error in the measurements. We also investigate user traffic driven by a Power law distribution and conclude that measurability of packet loss rate worsens when network traffic becomes burstier.

Chapter 7 consists of a summary of research and conclusive remarks. In addition future work directions related to this research are addressed.

2. Network Measurement Techniques

Managed IP networks are coming to lead the way information is brought to the users on a worldwide basis. As the numbers of customers, the range of users and the heterogeneity of application increases, so does the demand of providing guaranteed QoS [16]. The SLAs enumerate the network performance, typically mean delay, delay jitter and PLP. Measurement and monitoring of the network performance is vital to ensure guarantees are being met.

This chapter discusses the network measurement in different aspects e.g. importance and motivation, measurement types and comparison between the measurement techniques and their applications.

2.1 *Importance of Network Measurement*

Network measurement is an essential tool to collect information for the following:

1. Network performance evaluation - QoS performance metrics e.g. throughput, packet delays and losses.
2. Study network properties -e.g. traffic, path, and link characteristics.
3. Report generation – SLA validation.
4. Assisting the network operators e.g. dimensioning, capacity planning, network monitoring and fault detection.
5. Input for decision making schemes -e.g. routing, Connection Admission Control (CAC).

Performance degradation in a network gives rise to questions such as: Why is an application running slow? Where do delays come from? How can we improve the performance of a network? Statistics must be collected in order to answer these questions. Without measurements, there will be no objective record or benchmark of how a network behaves.

Measurements show whether changes improve or degrade the network's performance, and by how much [13].

Measurement for capacity planning provides the information for network operators to calculate necessary capacity, in order to avoid congestion in the network. The degradation of performance may cause long packet delay time, jitter or larger packet loss ratio. For VoIP the voice quality is not acceptable if the packet loss probability is more than 10^{-3}[8] [9] [10]. Figure 2-1 shows some common network measurement metrics. These metrics are availability, packet loss, delay and throughput. We have performed One Way Delay Measurements (OWD) as it provides a better characterization of parameters. Several measurements of OWD have been made before, in a much closed environments with low precision hardware or using short time intervals [40]. However, we have followed the simulation based approach by simulated different wireless scenarios using probing technique to infer OWD.

There are two basic techniques for network measurement *passive* and *active*.

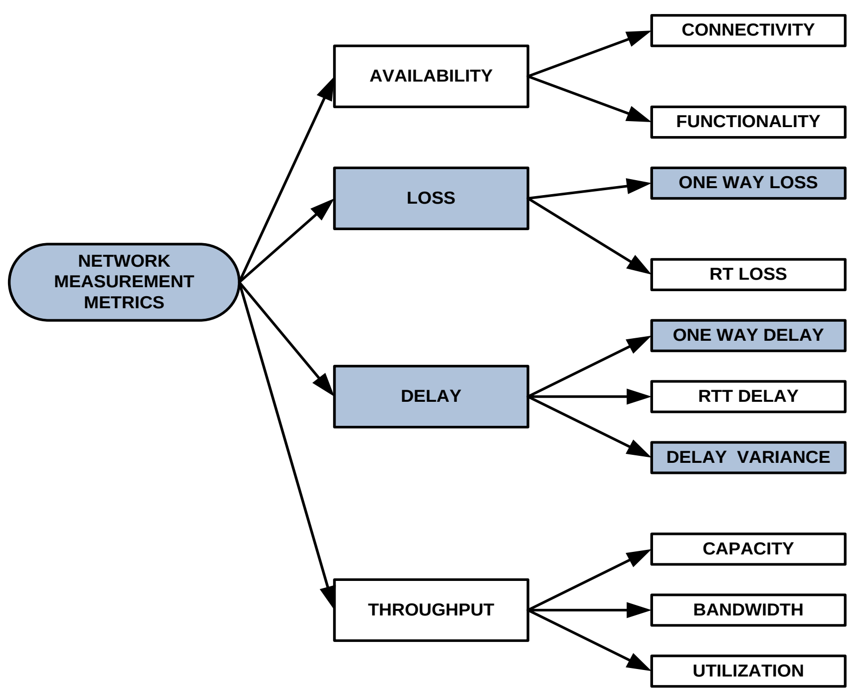

Figure 2-1 Network Measurement Metrics

2.2 Passive Measurement

Reading the buffer queue statistics from the router's buffers can give the packet loss probability directly. It is also possible to infer end to end delay distribution from convolution assuming link independence. Passive measurements can be implemented by incorporating some additional intelligence into network devices to enable them to identify and record the characteristics and quantity of the packets that flow through them. The packet statistics can then be retrieved without an additional flow of traffic. The following types of information can be obtained using passive monitoring: [17]

- Bit or packet rates
- Packet timings
- Queue levels in buffers
- Traffic / protocol mixes

Passive measurement tools include packet sniffing (hardware- based passive measurement), Tcpdump (software-based passive measurement), polling MIB.

2.3 Active Measurement

Active monitoring injects '*Probe*' traffic into the stream and uses the associated time stamps to get the delay distribution. Occasional loss of probe packets allows the packet loss probability to be measured.

Active Probing Infrastructure

The basic components of an end-to-end active measurement infrastructure are shown in Figure2-2. In an experimental setup of active probing the sender creates and transmits a probe stream which passes through the network and ends at the receiver, a sink. The information of probe sequence number along with packet arrival and departure time is retrievable from the packet payload. These data are recorded by sender and receiver monitors. Probe characteristics are important as they constitute the experimental data which is used for analysis.

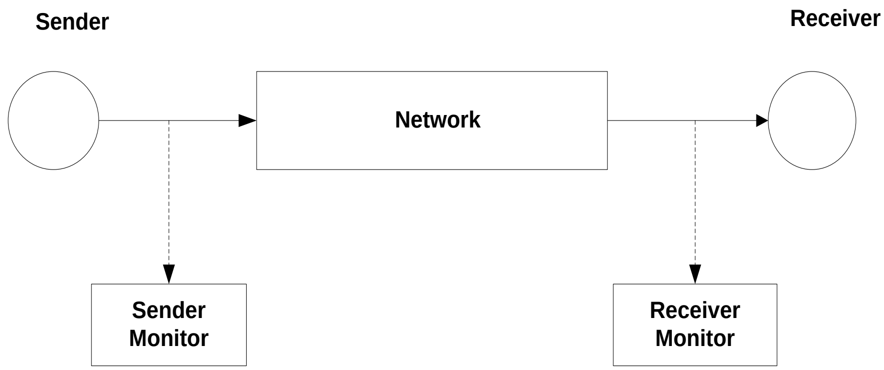

Figure 2-2 Basic components of Active Probing Experiment

Each component in Figure2-2 can be a potential source of error involving timing in some form. Further errors include those due to reference source timing, and synchronization issues [23]. Some of the active measurement techniques are:

2.3.1 Ping

The most widely used method to determine network delay is for a measurement host to construct and transmit an ICMP echo request packet to an echo host. As the packet is sent, the sender starts the timer. The target simply reverses the ICMP headers and sends the packets back to the sender as an ICMP echo reply. When the packet arrives at the original sender's system, the timer is halted and the elapsed time noted, i.e. the Round Trip Time (RTT) is calculated by the difference between the time the echo request is sent and the time a matching response is received by the PING application [10]. Figure 2-3 depicts an ICMP PING implementation.

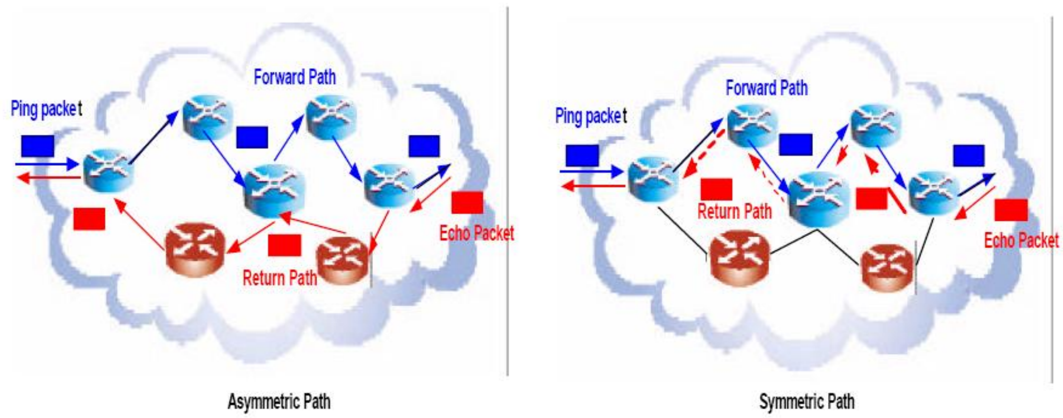

Figure 2-3 ICMP Ping implementation

The basic applications of PING are [23]:

- To test the availability of a path between the sources and destination end.
- To measure the RTT

2.3.2 Trace route

The traceroute technique allows a measurement host to deduce the forward path destination. Traceroute makes use of the TTL field in the IP packet and ICMP control messages. When the packet is dropped due to an expired TTL, the information from the intermediate node will be returned to the measurement host. A Traceroute example is shown in Figure 2-4.

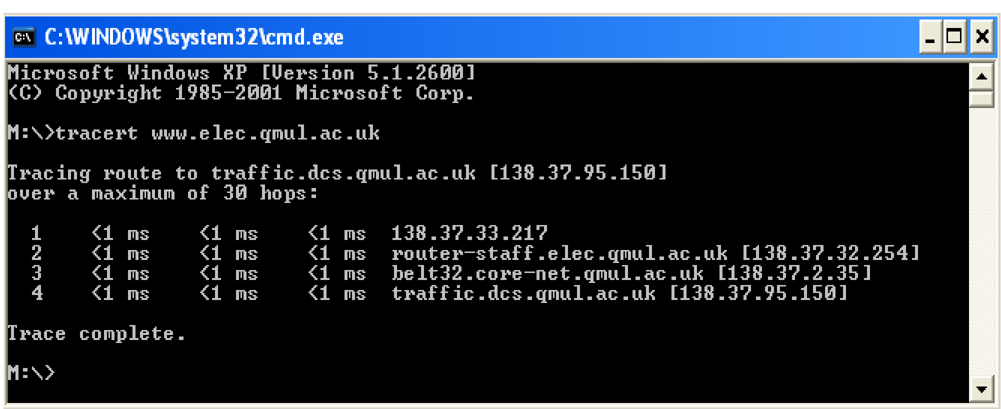

Figure 2-4 A Trace route Example

2.3.3 One- way Active Measurement Protocol

One-way Active Measurement Protocol (OWAMP) was proposed by the Internet Engineering Task Force's (IETE) Internet Protocol Performance Metrics (IPPM) [26]. The Active Monitoring Tools are intentionally placed at two measuring points: sending and receiving ends [25] as shown in the Figure 2-5 [19].

The most common method used to infer network performance is the well-developed approach of one way active probing. In this approach, probe packets are sent into the network with precisely controlled departure times, and their arrival times elsewhere in the network are measurement. Such probing requires installation of probe equipment into the network, but this equipment is typically fairly cheap, and it does not require special access to a network. These factors have led to active probing being the most widely deployed form of IP performance measurement [17].

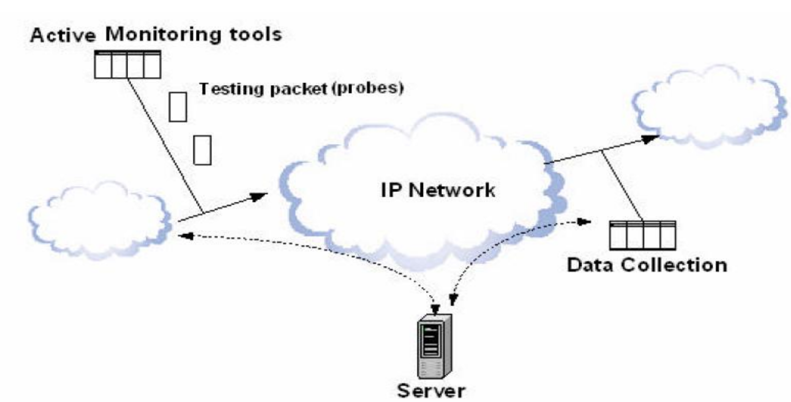

Figure 2-5 One way Active Measurement

The benefit of OWAMP is that it can run almost anywhere in the network and can easily estimate the end to end network performance.

2.4 *Sampling Methods for Active Measurement*

Poisson and *Periodic sampling* are the two most widely used sampling techniques for active measurement. Periodic sampling is still widely used as it is simple and easy to use. Poisson sampling is a kind of random additive sampling, i.e. the intervals between consecutive sampling instances are independent random variables which have the same exponential distribution. Poisson sampling is a special case of irregular sampling. There are two main advantages which make Poisson sampling a natural candidate for network measurement; the first is the characteristics of Poisson Arrival See Time Averages (PASTA), which ensure that, the average state observed by probes will converge to a true average of the system (network) under observation. Poisson sampling is also unbiased for the state that is sampled. The second is that it can avoid potential network synchronization as the next arriving packet cannot be anticipated, and so such measurement becomes harder to manipulate [12].

[25] Has introduced the concept of *Adaptive sampling* into active measurement, and investigated how adaptive techniques can be use to adjust the sampling rate of each parameter monitored according to the availability of source statistics. One benefit of this method is that it can eliminate the bias caused by synchronization.

2.5 Active Measurement Errors

While performing Active measurement on finer time scales, several milliseconds to seconds, periodic as well as more sophisticated and refined probe streams are employed to measure bottleneck as well as available bandwidth, and to explore the detailed statistical structure of delay and loss. However, at such time scales timing problems in common measurement infrastructures can result in measurement errors of the order of the time intervals one is trying to measure, and the inter-departure times of probe packets one is trying to control [18]. Even for the less demanding low resolution measurements, errors in the individual measurements of delay are no less significant, even if critical.

Some of the common errors encountered while performing active measurements are listed below:

2.5.1 Sampling Error

It is known from sampling theory that the greater the variability of sampled data the more samples are needed for accurate estimation, even for the mean value of the distribution only [8]. In the case of actively probing packet networks the variability is highly dependent on two factors: the load on the network, and the type of the traffic being carried.

High probing rates can lead to inaccuracies in the measurements. Therefore keeping the probe rate to low values helps to minimize inaccuracies due to probe overhead. However, high probing rates affects the number of measurement samples, hence the measurement accuracy too. It has also been discussed in [13] that at higher loads the sample error increases. This is because the variance of the queue size and hence delay is larger, relative to mean queue size, at higher loads. This error is worse if the traffic is highly bursty. This effect can be clearly observed in the Figure 2-6 [14] as the load is increasing, the bandwidth requirement for the probes in suddenly increased. We can observe from the Figure 2-6 that, with limited probing bandwidth standard error in the measurement becomes unacceptably high. On the other hand as the bandwidth taken by the probes is increased to a certain point (i.e. x with reference to Figure 2-6), there is a sudden decrease in error. Where, x is a certain threshold limit beyond with the error in measurements becomes unacceptable. The value of x can be different for different applications (e.g. voice, video etc). Based on the fact that it is practically impossible to get an infinite bandwidth, we cannot ignore this effect, especially when probing a wireless medium that is already constrained in terms available bandwidth.

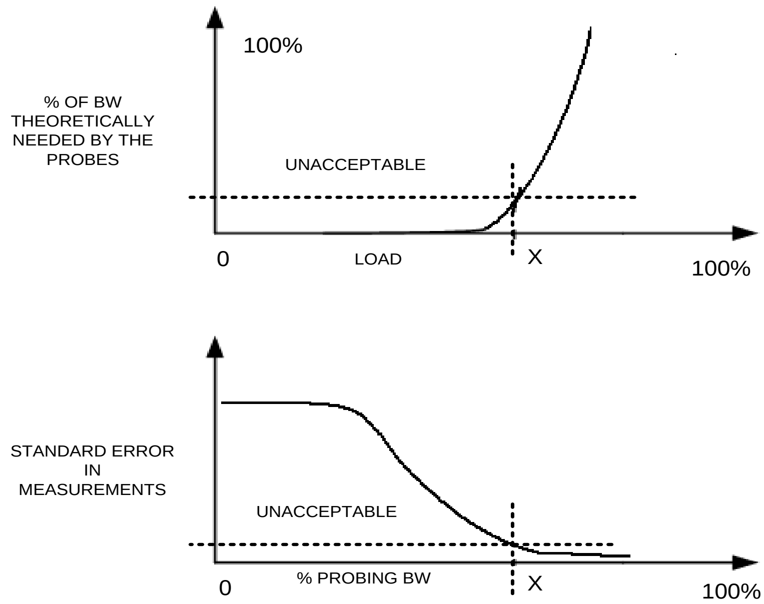

Figure 2-6 Behavior of Active Probing

Research in [27] has shown that highly bursty traffic is frequently found in packet networks and this results in very large variances associated with the number of packets in queues, and hence the packet delays [28] [29].

2.5.2 Self interference

Probe packets can themselves act as a source of packet-level interference. This is more likely to happen when the traffic is bursty and buffers are running out of space. There are methods to reduce self interference like minimizing the number of probes or reducing the size of the probes. By reducing the size of the probes we cannot judge that the loss pattern of probes is same as that of data packets, this fact has been proved in chapter 5 by our simulation results.

2.6 *Active and Passive Measurement Comparison*

In Table 2-1 [9] the features of Active Measurement and Passive Measurement have been outlined. As the main difference between Active Measurement and Passive Measurement is the presence or absence of the extra testing traffic. In Active Measurement, the testing packets pass through the network elements and the end-to-end path properties are measured. Without the probe packets, it is difficult for Passive Measurement to be employed in the end-to-end QoS related measurement. As end-to-end communication is often provided over a number of separately owned networks, there will be little chance of getting access to all the equipment's internal measurements [43]. So, intrusive measurement using probe packets is often the only option available.

	Active Measurement	Passive Measurement
Configuration	Multi-point	Single or multi-point
Data size	Small	Large
Network Overhead	additional traffic	- Device Overhead - NO overhead if splitter or copper network is used
Purpose	Delay, packet loss, availability	Throughput, traffic pattern
CPU Requirement	Low to Moderate	High – traffic capture Low – polling MIB (like queue length monitoring)
Intrusion	***YES***	***NO***

Table 2-1 Comparison between active and passive measurement approaches

The most important limitation of active monitoring turns out to be that, under certain circumstances, the bandwidth available for probing may be so small that there are not enough probes available per measured hour to provide a level of accuracy that makes monitoring worthwhile;. Active measurement is more challenging on the access links (typically operated to high utilization) than the core network [43].

2.7 Network Measurement Based Applications

In the previous sections we discussed the network measurement techniques and comparison between them. This section is organized to highlight some of important applications of network measurement.

2.7.1 Network Tomography

Optimizing communication network routing and service strategies requires knowledge of the delay (congestion) at different points in the network. However, as it is practically impossible to measure packet delay at each and every router. End-to-end delay is comparatively easy [18]. *Network Tomography* is a technique employed to infer the internal network elements characteristics such as loss characteristics and delay distribution of an intermediate node.

Network Tomography methods are addressed in [19], which is based on end to end measurement using active, multicast probing in order to infer packet correlation statistics and therefore gathering information about internal network characteristics. In addition to that multicast- based method may not provide precise characterization of network behavior as routers treat multi-cast packets differently to Unicast packets. Unicast network Tomography is easily carried out on most networks and is scalable. Unicast methods provide more accurate estimate of network characteristics as most of the networks are predominately Unicast in nature [25].

2.7.2 Available Bandwidth Estimation and Adjustment

Due to the statistical complexity of traffic in broadband networks, CAC algorithms often use estimates of required bandwidth. As a result, bandwidth adjustment is a methodology used to fine-tune the allocated bandwidth in order to achieve the optimum network operation point.

Available bandwidth is of great importance for various network applications as the knowledge of available bandwidth on end-to-end path can improve admission control, congestion control, adjusting codec rates for streaming applications, configuring overlay routes in overlay networks etc.

Present schemes for available bandwidth estimation fall into two broad classes. The first class of schemes is based on statistical cross-traffic model, are known as *Probe Gap Model* (PGM) such as Delphi and *Initial Gap Increasing* (IGI) [20][21].

PGM scheme makes use of the information in the time gap between the arrivals of two consecutive probes at the receiver. A probe pair is sent with a time gap Δ_{in} , and reaches the receiver with a time gap Δ_{out} . Assuming a single bottleneck and that the queue does not become empty between the departure of the first probe in the pair and the arrival of the second probe, then Δ_{out} is the time taken by the bottleneck to transmit the second probe in the pair and the cross traffic that arrived during Δ_{in} , as shown in the Figure 2-7 [18]. Thus, the time to transmit the cross traffic is $\Delta_{out} - \Delta_{in}$, and the rate of the cross-traffic is $\frac{\Delta_{out} - \Delta_{in}}{\Delta_{in}}.C$ [24]
Where C is the capacity of the bottleneck, the available bandwidth is:

$$\left(1 - \frac{\Delta_{out} - \Delta_{in}}{\Delta_{in}}\right).C$$

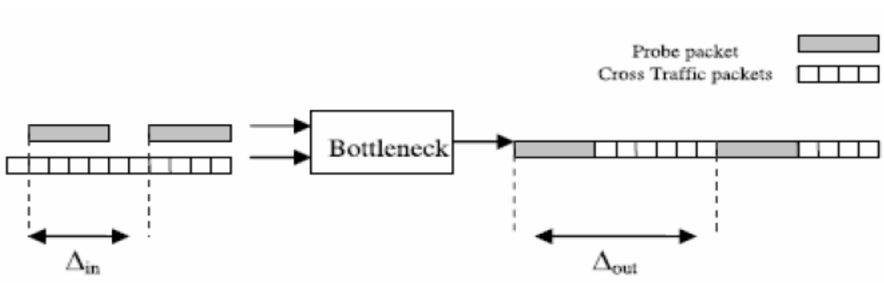

Figure 2-7 PGM for estimating the available bandwidth

The second class of scheme is based on the idea of self-induced congestion, also known as the *Probe Rate Model* (PRM). If the probes are sent at a lower rate than the available bandwidth along the path, then the arrival rate of the probes at the receiver end will match the rate at the sender. On the other hand, if the probe traffic is sent a rate higher than the available bandwidth, then queues will start growing inside the network and the probe traffic will be delayed. As a result the probe rate at receiving end will be less than at the sending rate. Consequently, one can measure the available bandwidth by searching for the turning point at which probe sending and receiving rates start matching. However, this cannot be deemed to be the best method to determine the available bandwidth as it causes congestion which affects the measurement process [25].

2.7.3 Measurement Based Admission Control

The role of an admission control algorithm is to ensure that the admission of a new connection into a resource constrained network does not violate service commitments made by the network to the admitted flows, at the same time as maintaining good network utilization. One of the approaches to admission control is known as parameter- based, which computes the amount of network resources required to support a set of flows given a priori flow characteristics. The new connection will not be accepted if the network resource is not sufficient [25].
Measurement-based admission control (MBAC) is an important technique for providing statistical services guarantees and many MBAC algorithms have been proposed in the literature [22][23].

Figure 2-8 [16] shows the structure of MBAC. It reveals that an MBAC algorithm for a network typically consists of three elements:

1. Admission decision algorithm
2. Traffic estimator
3. Resource estimator

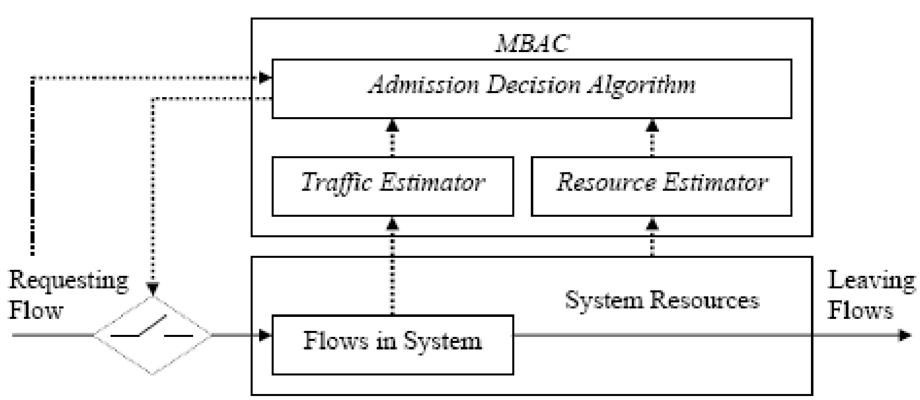

Figure 2-8 Structure of MBAC

The algorithm keeps on measuring the traffic in the network and the remaining network resources such as available bandwidth and buffer size. Based on measurement results, the traffic estimator estimates how much traffic there is in the network, what the characteristics of the network are and how many flows are there in the network; whereas, the resource estimator estimates how many resources remain in the system. When there is a request by a new flow to enter the system, the MBAC algorithm is used to decide if this flow can be admitted. Moreover, the decision also depends on some information from the requesting flow, which typically includes its QoS requirement and its traffic description [22].

[23] Has proposed an end-to-end MBAC mechanism for VoIP networks. In [23], active measurement is used for call admission control to maintain packet loss rate of any voice flow below a pre-set value under different network conditions. This method, in contrast to passive measurement, is very efficient when employed in large-scale VPN networks (where the VPN is carried over a mixture of packet networks that are not owned by single organization).

2.8 Summary

This chapter has discussed the significance of network measurement, different network measurement approaches, features of measurement methods and their applications. Performance measurement of IP networks involves: monitoring network performance, including availability, packet loss and packet delay statistics. A variety of network measurement techniques have been proposed and developed, but they are mainly classified into two main classes' i.e. active measurement and passive measurement. Active measurement is intrusive in nature whereas, passive measurement is not. Examples of active measurements tools are Ping, Traceroute whereas, packet sniffing, tcpdump and polling MIB are classic examples of passive measurement tools.

3. IEEE 802.11 Standard & End-to end Measurement Issues

During last few years a large amount of research has been carried out on end–to-end measurement of wired IP networks. End-to-end methodologies for wireless networks present problems that are not considered by the traditional wired network measurement techniques. In this chapter we will discuss some background of a wireless standard, i.e. IEEE 802.11, the issues involved in measuring end-to-end delays in such networks, the network delay components and the challenges involved in performing active measurements over wireless access networks.

3.1 Overview of IEEE 802.11 Standard Family

IEEE 802.11 denotes a set of Wireless LAN (WLAN) standards developed by working group 11 of the IEEE LAN/MAN Standards Committee (IEEE 802). The IEEE 802.11 standard defines the media access control (MAC) and physical (PHY) layer for a local area network (LAN) with wireless connectivity [30]. It addresses local area networking where the connected devices communicate over the air interface to other devices that are within close proximity to each other using carrier sense multiple access protocol with collision avoidance (CSMA/CA) mechanism.

The 802.11 standard family includes six over-the-air modulation techniques that use the same basic MAC protocol. The first version of the 802.11 standard was released in 1997. It specified three PHY layer options: infrared (IR), frequency hopping spread spectrum (FHSS) and direct sequence spread spectrum (DSSS). FHSS and DSSS operate at the industrial, scientific and medical (ISM) band at 2.4 GHz, and IR uses near-visible light in the 850 nm to 950 nm range for signaling [45].

The three basic options did not become widely used mostly because of the low data rate they can provide, maximum 2 Mbps. Then it was rapidly supplemented by 802.11b in 1999.

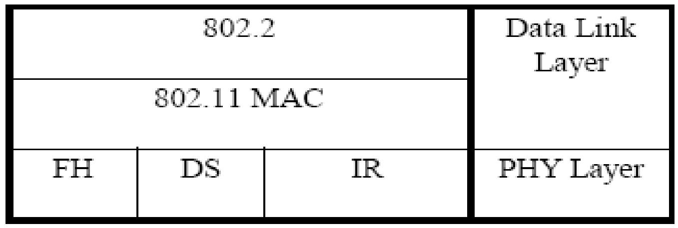

Figure 3-1 IEEE 802.11 Layers Description

The 802.11 protocol, covers the MAC and Physical Layer, The Standard currently defines single MAC which interacts with three PHYs as shown in Figure 3-1.

While in the same year, i.e. 1997, another amendment to the 802.11 legacy, 802.11a was also released. It operates in 5 GHz band to avoid the crowded 2.4 GHz band. Also, 802.11a uses the OFDM modulation scheme and thus gives a much higher data rate of 54 Mbps.

The first widely used amendment to the 802.11 standard was 802.11b. It operates at 2.4 GHz and supports the maximum data rate of 11 Mbps. 802.11b directly uses the DSSS modulation technique of the original 802.11 legacy standard and largely improves its data rate. It can satisfy most of the network applications, thus not long after its release the 802.11b products began to appear on the market and then became very widely used. Because of this, new 802.11 amendments have tended to choose to be backward compatible to 802.11b.

To further improve the data rate, in June 2003, 802.11g was released. It uses OFDM (Orthogonal Frequency Division Multiplexing) as the modulation scheme, and supports a maximum data rate of 54 Mbps, which is the same as 802.11a. It operates at 2.4 GHz and can be viewed as an enhancement of 802.11b.

The latest generation to improve the data rate is 802.11n. This new amendment to the 802.11 standard is estimated to reach a theoretical data rate of 540 Mbit/s (which may require an even higher raw data rate at the physical layer), which is up to 50 times faster than 802.11b, and up to 10 times faster than 802.11a or 802.11g. 802.11n builds upon previous 802.11 standards by adding MIMO (multiple input multiple output). MIMO is the use of multiple antennas at both the transmitter and receiver to improve communication performance [45].

There are some other important amendments that provide new functionalities to the original 802.11 legacy. 802.11i is an amendment to provide secure network access. 802.11b the core of this thesis, we have used for wireless access and performed monitoring over it.

3.2 MAC Layer of IEEE 802.11 Legacy

The original 802.11 standard [1] is also referred to as "802.11 legacy". The 802.11 standard specifies a common medium access control (MAC) layer, which provides a variety of functions that support the operation of 802.11-based WLANs. The IEEE 802.11 MAC layer covers three functional areas: reliable data delivery, security and media access control.

As with a wireless network, a wireless LAN using the IEEE 802.11 physical and MAC layer is subject to considerable unreliability. Noise, interference, and other propagation effects results in the loss of significant numbers of frames. Even with error correction codes, a number of MAC frames may not successfully be received. For this purpose, IEEE 802.11 includes a frame exchange protocol. When a station receives a data frame from another station, it returns an acknowledgment (ACK) frame to the source station.

3.3 Medium Access Control

The 802.11 working group considered two types of proposals for the MAC algorithm: Distributed Access protocol, which is like Ethernet, distributes the decision to transmit over all the nodes using a carrier-sense mechanism; and a Centralize Access Protocol which involves regulation of transmission by a centralized decision maker. A Distributed Access Protocol makes sense for an Ad-hoc network peer work stations and may also be attractive in other wireless LAN configurations that consist primarily of bursty traffic [46].

The original 802.11 standard has defined two medium access mechanisms: the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). The DCF is the basis for the PCF and is used for best effort contention services. Since the DCF is a distributed access method, it can be used not only in infrastructure network configurations, but also in Ad-hoc network configurations. DCF uses a contention algorithm to provide access to all traffic. Ordinary asynchronous traffic directly uses DCF. The PCF on the other hand, is build on the top of DCF to assure access for its users and is required for contention-free services. Furthermore, it is only usable in infrastructure network configurations.

3.3.1 Distributed Coordination Function (DCF)

DCF is a method used by the stations contending for access to the network resources and attempting to send frames when there is no other station transmitting. *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) is the basic multiple access method used by DCF. Since it is contention based, it can be used in both infrastructure and Ad -hoc networks.

CSMA/CA

In CSMA/CA, a station wishing to transmit has to listen to the channel for a predetermined amount of time so as to check for any activity on the channel, which is a radio medium that stations share.

If the channel is sensed "*idle*" then the station is permitted to transmit. If the channel is sensed as "*busy*" the station has to defer its transmission. This is the essence of the "*collision avoidance*" part of the protocol.

Inter Frame Space

To ensure the smooth and fair functioning of this algorithm, DCF includes a set of delays that amounts to a priority scheme. To control the waiting time before medium access, the DCF use the following parameters called *inter frame spaces* (IFS): short inter frame space (SIFS), DCF inter frame space (DIFS), and extended inter frame space (EIFS). The different lengths of the IFSs are defined to provide priority levels for access to the wireless media. SIFS has the shortest waiting time, and is thus used to give the highest priority for medium access. It is used by short control messages, such as ACK (Acknowledgement) and CTS (Clear to Send) frames. DIFS is a longer waiting time than SIFS and thus has lower priority for medium access. It is used to transmit data or management frames by the stations operating DCF. EIFS has the longest waiting time and is only used when a transmission failure occurs. A station that receives an incorrect frame must wait for EIFS before starting its transmission in order to give other stations enough time to acknowledge the frame that the station received incorrectly.

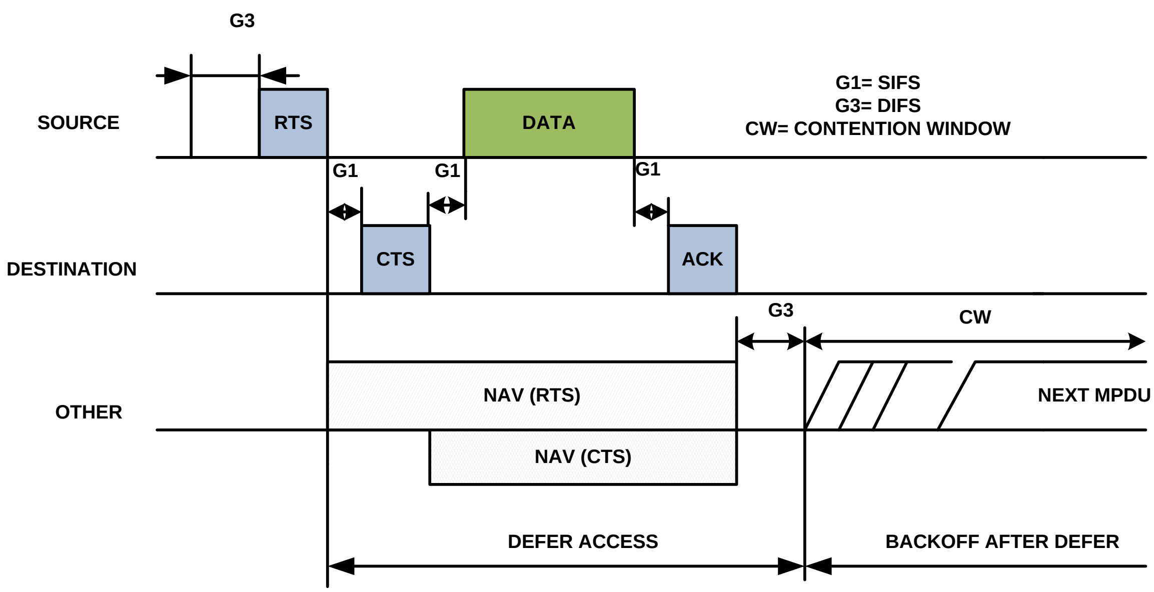

Figure 3-2 Schematics of Access Mechanism CSMA/CA

3.3.2 DCF Access Procedure

Before transmitting, each station in the network has to sense the medium. There are two carrier sense mechanisms in PHY layer and in NAV which is virtually provided by the MAC layer. *Network Allocation Vector* is a counter maintained by each station with amount of time that must elapse until the medium becomes free. It contains the time that station, currently require for transmitting its frame. Stations cannot transmit until the NAV is zero.

As shown in the Figure 3-3, when the station determines that the medium is sensed to be idle for greater than or equal to a DIFS period, it can initiate a transmission immediately. Otherwise, if the medium is sensed to be busy, or if it is sensed to be idle but becomes busy before it reaches the end of DIFS period, the station must defer the pending transmission by itself, wait until the end of the current ongoing transmission which occupies the medium, and start to sense the medium to be idle for a DIFS period without interruption. Then a random back-off process is invoked to reduce the possibility of collision, since in this situation (after the medium becomes idle following a busy medium) several stations may be waiting for the medium to become idle and there is a high probability of collision.

It means when a station senses the medium to be idle again for a DIFS period, then it begins a back-off period. During the back-off time, if the medium is sensed to be idle for a whole slot, the generated random number is decremented by one, and if the medium is sensed to be busy, the back-off timer becomes suspended until the medium is sensed to be idle for the duration of DIFS and then the back-off procedure resumes. The station can start to transmit the pending frame if the back-off timer reaches zero. This procedure ensures that the station selecting the smallest back-off time using the random function will win the contention.

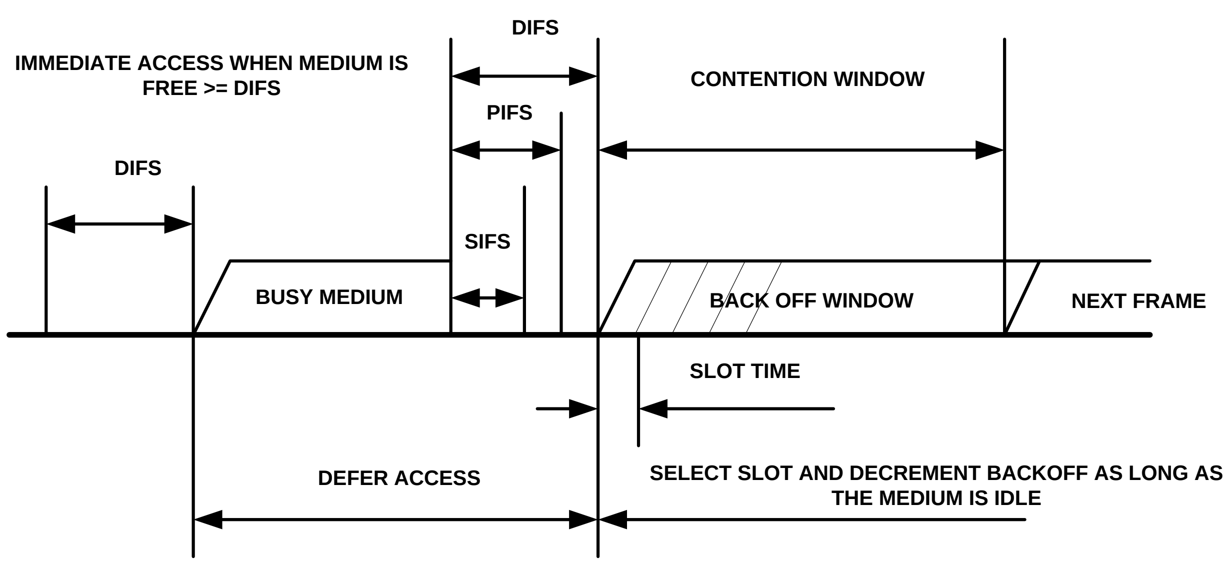

Figure 3-3 Schematics of Access Mechanism with Back-off Window

The value of CW varies between CW_{min} and CW_{max}, which are PHY-dependent. It is possible that two stations which choose the same random number finish the back-off procedure and start to transmit simultaneously, which causes collision again. To reduce the possibility of collision in this situation, each station maintains a CW window with a different length. Initially CW is set to CW_{min} and its value doubles if a collision occurs until CW_{max} is reached. If the transmission is successful, the CW value reverts to CW_{min} before a new random back-off interval is chosen. Upon receiving a frame, the destination station waits for the duration of a SIFS and responds with an ACK frame to notify the sender of a successful reception. Figure 3-4 depicts the access mechanism with back-off window whereas Figure 3-5 illustrates the flow chart representation of MAC access mechanism.

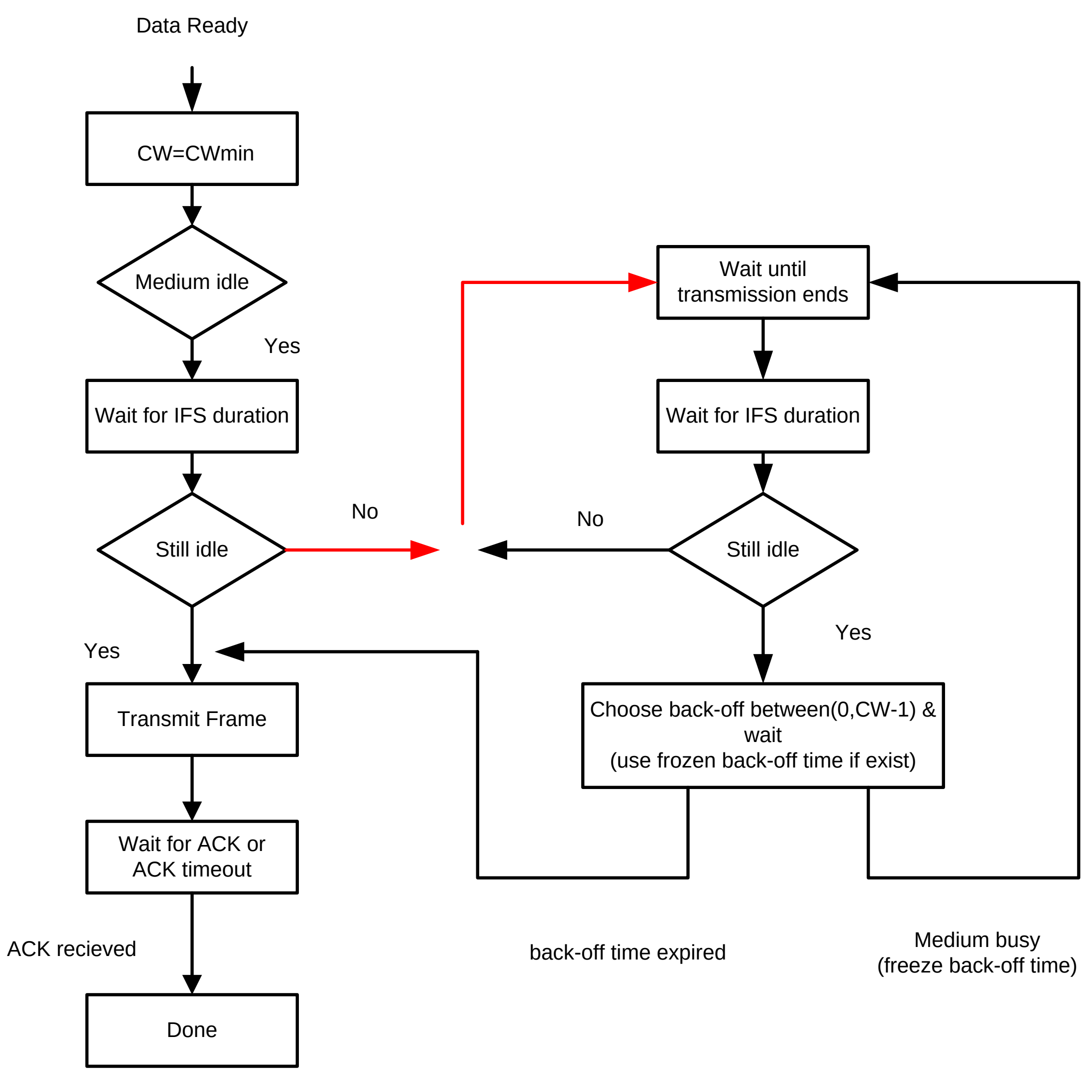

Figure 3-4 IEEE 802.11 MAC Logic

3.3.3 RTS/CTS

One of the problems is that it is not possible to listen while sending; therefore collision detection is not possible. Moreover, the DCF suffers from the hidden station problem that exists in contention based protocols: Two stations are hidden from each other and if they are out of signal range and thus cannot hear each other. Figure 3-6 shows the hidden node problem i.e. a sends to B, C cannot listen to A. C wants to send to B, C senses a free medium. Collision at B not detected by A: C is hidden from A. In this situation the carrier senses mechanism will not work properly and the hidden station may both sense the medium free and start transmitting at the same time to a common receiver causing a collision.

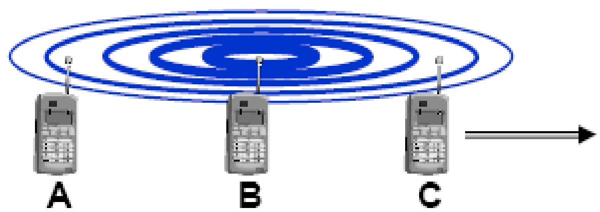

Figure 3-5 Hidden Node Problem

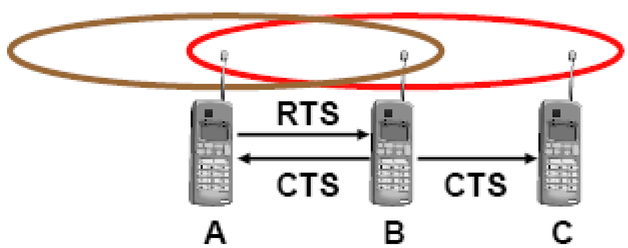

Figure 3-6 RTS/CTS

Therefore *a Request to Send* and *Clear to Send* (RTS/CTS) exchange mechanism can be optionally adopted by CSMA/CA to counter the hidden station problem, and to avoid collision by announcing the impending transmitting status to the neighbor stations. Before sending a data or a management frame, a station can transmit an RTS frame and await a CTS frame. Figure 3-7 shows the RTS/CTS mechanism where, CTS from B to A is also heard by C, which then waits.

The RTS and CTS frames contain a Duration/ID field that defines the period of time that the medium is to be reserved to transmit the actual data frame and the returning ACK frame. All stations within the reception range shall update their network allocation vectors (NAVs) so that they consider the medium busy until the end of the transmission. Thus, collisions occur only on RTS frames and are detected by the absence of a CTS frame. However, RTS/CTS mechanism introduces a large overhead especially for short data frames.

Although, the MAC access mechanism is responsible for reliable delivery of frames the time spend in counting down back-off intervals is a part of the MAC overhead. Choosing a large CW can lead to large back-off intervals and can result in larger overhead, whereas selecting a small CW leads to large number of collision (when two nodes count down to 0 simultaneously).

3.4 Back off Algorithm in Wireless Multihop Networks

A multiple hop wireless network is a network where a node can communicate with far-way nodes through packet relying by intermediate nodes. All the nodes are equal and share the communication medium with any centralized control. Therefore, network performance in terms of packet delays and losses may significantly degrade as the number of competing nodes increases which will bring to large interference. If the medium is busy, a back off algorithm (discussed earlier in section 3.3) is used to avoid repeated collisions. So the beck off algorithm is crucial to the performance of multiple hop wireless networks [63]. Unlike wired networks, problems such as mobility of nodes, shared broadcast channel, hidden and exposed terminal problems, and constraint on resources, such as bandwidth and battery power, limit the application of ad hoc networks.

Recently, the renewed interest in ad hoc wireless networks have centered on using IEEE 802.11 MAC mechanism. In [64], the authors came up with the question: Can the IEEE 802.11 works well in wireless Multihop networks, although as discussed earlier in this chapter that 802.11 MAC can support Multihop mobile networks, in which connectivity is the most prominent aspect. The performance of IEEE 802.11 MAC protocol is determined by the contention window control scheme, RTS/CTS mechanism, transmission range, etc. The metrics for the performance constitute throughput, packet delays, losses, jitter, etc will be highly affected by binary back off contention window algorithm. Measuring such a critical network using probing can proved to be highly challenging. Some of these challenges are discussed in section 3.6 in detail.

The following section discusses the types of delay components in end-to –end measurement.

3.5 End-to-End Delay Measurement and Packet Delay Components in Wireless Networks

Heterogeneous data networks carry different types of applications. Some of them are delay sensitive i.e. the quality of service is affected by delay. Some examples of delay sensitive applications are listed in Table 3-1 [36].

	Application	Minimum user requirement
Inelastic applications	Video	5 frames per second
	Audio	Delay < 400ms
	Interactive real-time multimedia	Delay < 200ms Jitter < 200ms
Elastic applications	Web page access	Latency < 11 seconds

Table 3-1 Minimum latency thresholds for acceptable QoS

Delay in wireless network can therefore summed in three categories i.e. *Intra node delay*, *Inter node delay* and *Response time* as discussed in [65]. Inter node delay constitute time spend in packet header processing, route table look up, queuing time. Theoretically, this delay is equivalent to passing the packet through bottom three layers (i.e. Physical layer, Data Link Layer and Network Layer). Inter node delay involved the air time i.e. propagation delay we have measured the end-to-end delay by determining the difference between the timestamps at the packet's departure from the source and the arrival at the destination. Response Time is the delay incurs at the destination node where the packets is send towards the source as acknowledgement packets.

In the following section we have split end-to end delay into stochastic and deterministic components as follows:

3.5.1 Processing Delay

Processing delay is the time needed to process a packet at each node in preparation for transmission. This delay varies with the voice coder used and processor speed.

During processing of a packet, routers may check for bit-level errors in the packet that occurred during transmission as well as finding where the packet's next destination is. In the past, the processing delay has been ignored as insignificant compared to the other forms of network delay. However, in some systems the processing delay can be quite large particularly where routers are doing complex encryption algorithms and investigating or modifying packet content. Routers that perform network address translation normally have higher processing delays as they need to examine and modify both incoming and outgoing packets [37].

3.5.2 Transmission Delay

ln a network based on packet switching, **transmission delay** (or **store-and-forward delay**) is the amount of time required to push all of the packet's bits into the wire. In other words, this is the delay caused by the data-rate of the link. Transmission delay is a function of the packet's length and has nothing to do with the distance between the two nodes. This delay is proportional to the packet's length in bits,

It is given by the following formula:

$D_T = N / R$

Where

D_T is the transmission delay in seconds

N is the number of bits in the packet, and

R is the rate of transmission (in bits per second) [37]

3.5.3 802.11 MAC inter- transmission delay

In a wireless access system like IEEE 802.11, there is a significant amount of delay introduced by the MAC layer in between the transmission times at a node in a busy network. 802.11 MAC makes use of CSMA-CA technique coupled with Virtual Carrier Sense Mechanism to achieve collision free transmission under contention in a shared channel. The CSMA mechanism has an intense influence on the transmission delays at MAC.

Investigations has shown that in contrast to wired networks, the inter-arrival time distribution at a busy node in 802.11 network shows evidence of pacing introduced by the MAC protocol [38]. Moreover, the inter-arrival times have a multi-modal distribution.

3.5.4 Propagation Delay

The time needed to propagate a bit through the communication link. The propagation delay is deterministic, for example, the travel time of an electromagnetic wave through the physical channel of a communication path like fiber optic, or by electrical signal speed via a cable. It is independent of actual traffic on the link. The propagation delay can be important as, for instance, in trans-Atlantic links or in a satellite link. In satellite communications, the one-way propagation delays are around 110 – 150ms for medium earth orbit system (MEO). However, if the satellite moves with respect to transmitter or receiver then the propagation delay might not be fixed.

3.5.5 Queuing Delay

In telecommunication, the **queuing delay** is the time a packet waits in a queue until it can be served. It is a key component of network delay. This term is mainly used in reference to routers. Packets are processed and transmitted once they arrive at a router. If packets arrive faster than the router can process them (such as in a burst transmission) the router stores them into the queue (also called the buffer) until it can get around to transmitting them as shown in Figure 3-8.

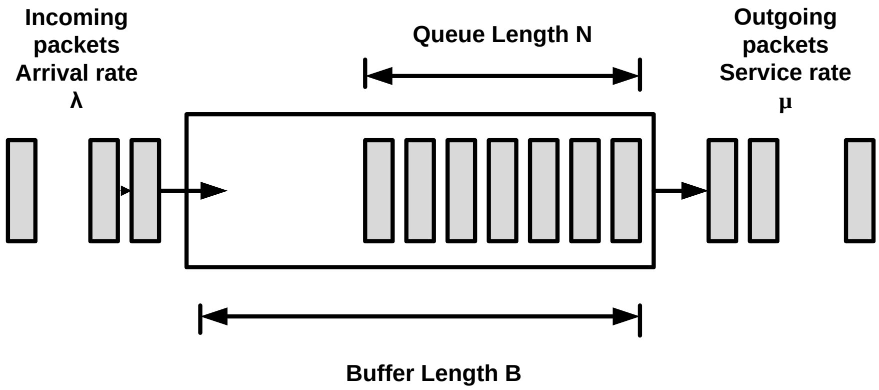

Figure 3-7 Queuing System

The maximum queuing delay is proportional to buffer size in FIFO (First in First Out) queues. The longer the line of packets waiting to be transmitted, the longer the average waiting time is. However, this is often much preferable to a shorter buffer, which would result in dropped packets, which in turn would result in much longer overall transmission times (due to re-transmission) [37].

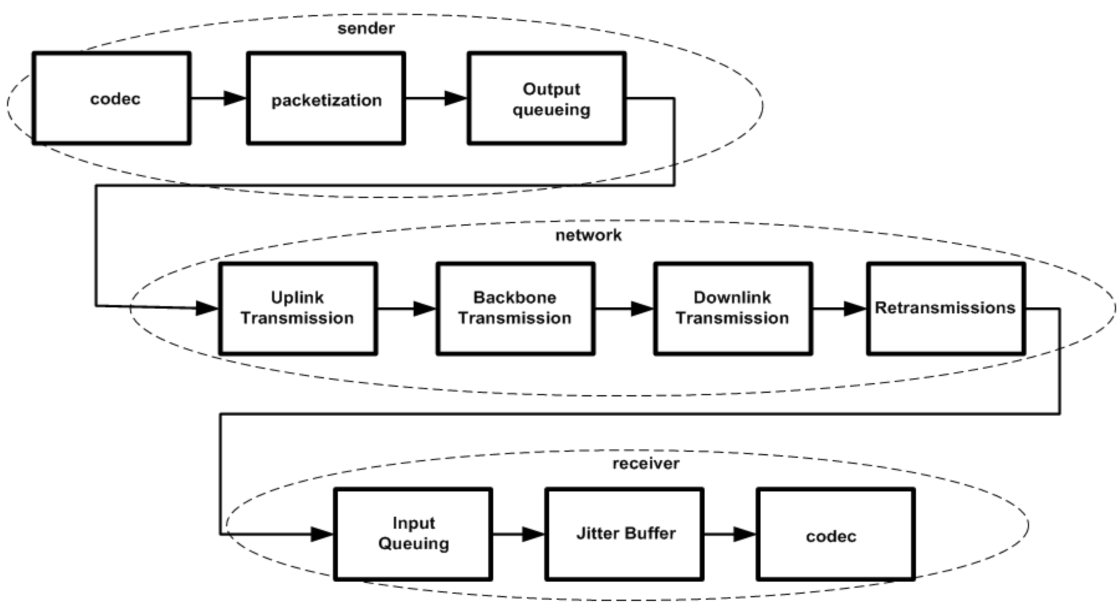

Figure 3-8 Components of Delay in a Wireless Network

3.6 *Challenges Involved in Performing Active Measurement over Wireless Access*

The significance of working with Multihop wireless networks cannot be denied. Wireless Multihop networks e.g. Ad hoc networks are required where a fixed communication, wired or wireless does not exist or has been destroyed. In multi-hop wireless networks, nodes communicate with each other using multi-hop wireless links and there is no centralized stationary base station. Each node in the network also acts as a router, to forward the packet to the other node [41]. Each node in such a network acts as a router, forwarding date packets to other nodes. In our research, we focus on a static multi hop network. We noticed the problems caused by multiple hop wireless links due to channel sharing and binary back off algorithms particularly when there are large numbers of nodes trying to communicate.

Most of the wireless networks are based on an 802.11 standard. IEEE 802.11 uses a contention-based media access control (MAC), 802.11 is incapable of allocating bandwidth and scheduling traffic practically when the load on a network is high. However, Research has been carried out due to increasing popularity of 802.11 based multi-hop wireless networks [63] [64] [65], a lot of research has been done on fabricating efficient monitoring of such networks in order to provide resource management, fault management and in these networks. This objective requires an efficient monitoring infrastructure in the wireless network, which can provide a maximum amount of information, while utilizing the least network resources. However, network monitoring introduces overheads which can affect performance in the network [44].

The characteristic of the wireless medium makes wireless network measurement fundamentally different from wired networks. Specifically, as indicated in [42], the wireless medium has neither absolute nor readily available observable boundaries outside of which stations are known to be unable to receive network frames; the channel is unprotected from outer signals; the wireless medium is significantly less reliable than wired media; the channel has time-varying and asymmetric propagation properties. In ad-hoc networks that rely on a carrier sensing random access protocols, like the IEEE 802.11 the wireless medium characteristics are susceptible to complex phenomena such as the hidden node and the exposed node problem

[30]. For these reasons, when monitoring a wireless access network delays and losses are already potentially high.

IEEE 802.11 encapsulate protocol overheads (including MAC, IP, UDP, RTP, and routing overheads in case of multi-hops) which results in a decrease in available bandwidth for transmission. Active probing of a network requires accessing the wireless network and transmitting probe packets. This means that probes become a part of competing traffic for accessing medium. In the presence of an addition source of traffic, which is also candidate to gain access to the medium, active probing becomes even more challenging than in wired networks.

Early simulation experience in wireless networks, and specifically wireless ad-hoc networks, suggests that their capacity can be surprisingly low. This is due to the bandwidth sharing and that nodes forward each other packets [40]. As the utilization of the medium increases, the contention between the data flow increases. With the increase in the contention associated delays increases too, and the throughput of the network decreases.

Reference [43] indicates that the bandwidth required by the probes (to achieve the desired level of accuracy) increases significantly as the overall load increases above 50%. This clearly means that the performance falls dramatically as the available bandwidth decreases, as load increases and traffic burstiness increases.

The impact of monitoring wireless network overheads on data traffic has been overlooked in most current research. It remains unclear as to how to parameterize such as the number of monitoring probes, the frequency of monitoring probes, their sizes and patterns, so that they have minimal effect on the performance of wireless networks. A small amount of overheads can result in a large degradation in network performance.

3.7 Summary

In this chapter we have described a brief over view of IEEE 802.11 Standard and the DCF access mechanism and CSMA/CA protocol and how it work. We have described different components of delay e.g. processing delay, access delay, queueing delay, transmission delay and retransmissions delay. Lastly we have highlighted challenges typically involved while performing measurements over wireless access.

4. Packet Network Simulation

This chapter presents the techniques used in research to simulate packet networks. It explains briefly explains the types of the simulation approaches, traffic models and the simulation tools used for network simulations.

4.1 Simulation based Network Performance Analysis

The advantage of simulation is that any complex network, whether it is a small or large scale, can be simulated. Simulation is more flexible than analysis and simulation may support the generation of random numbers with practically any distribution [49]. On the other hand, simulation based modelling has its drawbacks, such as possible excessive memory requirement, and a lot of time can be used in order to reach steady state.

4.2 Simulation Models

In this section we will discuss different types of simulation models used while simulating communication networks.

4.2.1 Static and Dynamic Simulation Models

Static simulation modelling (sometimes called *Monte Carlo simulation model*), represents a system in which the parameters do not change with time. In these simulations, the elements and the parameters used in the simulation do not alter during the simulation run. The values of the parameters are set at the start of simulation and they remain unchanged till the end of simulation run. Therefore, they can be set to represent "busy hour" values [49].

Dynamic simulation modelling represents time varying behaviour of the system. These can be run in real time in order to get the virtual response close to the actual system. In such a simulation some or most of the elements of the model change their values during the simulation.

4.2.2 Deterministic and Stochastic Simulation Models

Simulation models that contain no random numbers or probabilistic elements are known as *Deterministic* Models. Such models have a known set of inputs which will result in a unique set of outputs. In a deterministic simulation, the output is determined when the set of input elements and the attributes of the models have been specified.

A *Stochastic* simulation model may contain one or more random variables as inputs that lead to random outputs. In most simulation models randomness is significant to mimic a real scenario. In a stochastic simulation, the output must be treated as statistical estimate of the true characteristics of the system [60].

4.2.3 Continuous and Discrete Simulation Models

A *Continuous* simulation model is used to represent a system in which the state of certain variables changes continuously with time. Continuous simulation models usually use differential equations that give rates of change of the state of the variables with respect to time.

A *Discrete* simulation models a system in which state variables change only at a countable number of points in time. These points are the only ones where events occur. An event is known as an instantaneous occurrence that may change the state of the system. For this type of simulation modelling it is desirable to have a system clock that jumps from instant to instant and only does the things that need doing at those times. Although discrete-event simulation can be done by hand calculations, the amount of data that must be stored and process for most real world system is so large that it is best done on a computer [49].

4.3 Traffic Models

Traffic models have long been used to imitate real network traffic. In this section some of the most useful traffic models that are used for representing real queueing behaviour of networks are discussed. Selection of an accurate traffic model can help in network performance analysis either through analysis or computer simulation. An appropriate traffic model should be able to capture the key characteristics of the actual traffic in an effective way. If the traffic is not represented accurately, the simulation may under estimate or over estimate the network performance [25].

Traffic models can be divided into many different types but in this thesis we will mostly focus on exponential On/Off for short-range dependence (SRD), and a similar model with a Pareto distributed sojourn times to depict long-range dependence (LRD). These are briefly discussed in the next section.

4.3.1 Short- Range Dependent Traffic Models

The first performance models for telecommunication systems were based on the assumption that an arrival pattern of traffic is usually well modelled by a Poisson process [25]. Poisson processes are well-behaved and because they are stateless and the peak loading are not constant, so the queues fill slower [50]. Later, Markov Modulated processes were introduced for demonstrating a variety of traffic scenarios. In the following section, we will review some traffic models that have been used most commonly.

4.3.1.1 Interrupted Poisson Process (IPP) Models

Figure 4-1 shows an IPP traffic model. IPP traffic model is a two state model i.e. ON and OFF. Every interrupted process may be either in ON state (active) or OFF state (idle). The time in ON and OFF states is exponentially distributed with mean $1/\alpha$ and $1/\beta$ respectively. Packets are generated only occur during the ON state and follow a Poisson distribution with the rate λ.

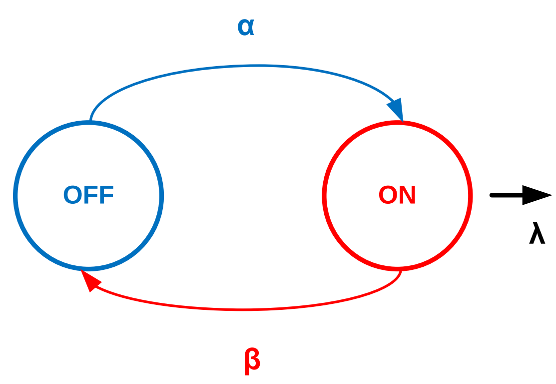

Figure 4-1 IPP Model

4.3.1.2 Markov Modulated Poisson Process (MMPP)

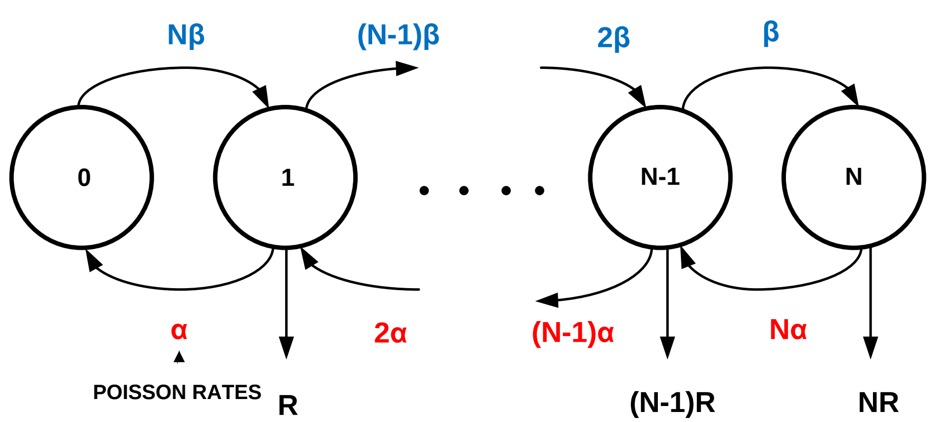

Figure 4-2 Superposition of N sources

Figure 4-2 shows an MMPP model for multiplexed sources. MMPP model has been extensively used to analyze teletraffic models. The MMPP is a doubly stochastic [51] [52] [53] Poisson process as, in which the current rate is determined or modulated by a continuous time Markov Chain. This process has more than two states in which the current state of the underlying process controls the rate of generating traffic [50].

The MMPP model has a finite number of states M, and operates as a Poisson process with the state- dependent rate λ, where 1 ≤ i ≥ M.

4.3.1.3 ON/OFF Models

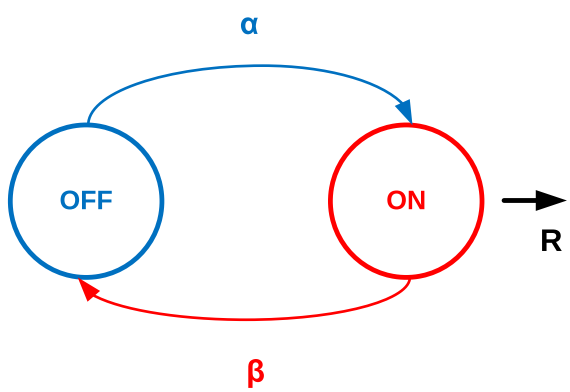

Figure 4-3 ON/OFF Model

Figure 4-3 shows an ON/OFF source model. This model is widely used for modeling voice traffic [2] [12]. These ON/OFF models have only two states i.e. ON state and OFF state as shown in the Figure 4-3. In this model packets are only generated during talk spurts (ON time) and have fixed inter-arrival times. The time spend in ON and OFF states is exponentially distributed with the mean 1/α and 1/β respectively [51].

It has been discovered [61] that Nx2 state process can be replaces by a single 2 state process accurately. Therefore, using traffic aggregation we can multiplex sources and can replace by a single equivalent ON/OFF source as shown in Figure 4-4 b [49].

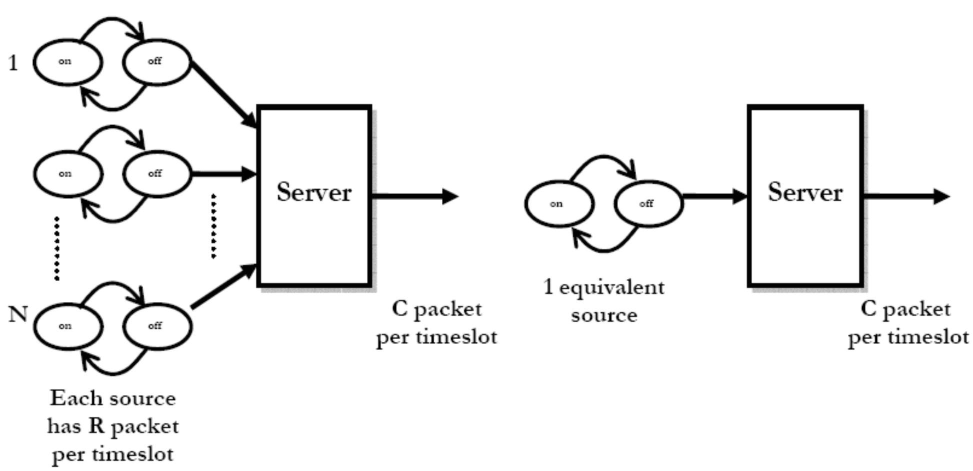

Figure 4-4 a. Conventional N ON/OFF sources b. Traffic Aggregation

4.3.2 Long-Range Dependent Traffic Models

The models discussed in section 4.3.1 have autocorrelation functions that decay exponentially. It is suggested in [55] [56] [57] that traffic measurement of Ethernet Local Area Network LAN traffic, Web traffic and ATM- encoded variable bit rate video are well modelled by process whose autocorrelation function (tail probabilities) decays as a Power law, rather than exponentially. Therefore, a single communication process with heavy tailed sojourn times shows long-range dependence and the aggregation of LRD sources exhibits high variability or infinite variance (also known as *Noah effect*), results in self- similar traffic [50].

[58] In the authors find that LRD is present in IEEE 802.11b wireless LAN traffic at all the levels of load. Under heavy load conditions, a higher degree of LRD is present. They conclude that as the aggregation level of traffic increases, the degree of long-range dependence increases as well. Furthermore, an investigation in [58] confirms that higher degrees of LRD are observed with shorter sampling intervals. This may be due to the fact that a shorter sampling interval reflects the characteristics of the actual (original) process better.

4.4 Simulation Model and Simulation Tool used

This section describes the simulation model used in this thesis and the simulating tool employed

4.4.1 Traffic Model Used: ON/OFF Model

In this thesis, a superposition of homogenous ON/OFF sources is chosen to produce the SRD and LRD traffic. In order to generate SRD traffic, the ON and OFF periods are exponentially distributed. Whereas, ON periods governed by the Pareto distribution are used to generate LRD traffic. The advantages of using ON/OFF models are [50]:

- The model is simple and can be defined using few parameters
- It is a basic prototype for a bursty traffic source
- It is an effective model for representing packet traffic

As compared to other traffic models like MMPP, traffic modelled using ON/OFF sources only requires a few parameters. The parameters needed to define ON/OFF model include the packet transmission rate, mean ON time, mean OFF time, and the number of aggregated sources.An ON/OFF source is the basic prototype for bursty source and has been extensively used in tele-traffic modelling [50].

4.4.2 Simulation Tool used: Network Simulator (Ns2)

The Simulator used in this research is Network Simulator 2 (Ns2). This is the most extensively used event driven simulator and is open source.

Ns2 is a discrete event simulator aimed at networking research. Ns2 provides considerable support for simulation of UDP, TCP, MAC, Ad-hoc routing, Sensor networks, multicast protocol over wired and wireless (local and satellite) networks.

Ns2 consists of the following components:

- Ns – The Simulator
- Nam – Network Animator for Visual demonstration of Ns output
- Pre-processing – This included either handwritten TCL scripts or Topology generator
- Post-processing – Trace analysis using Perl/TCL/AWK/MATLAB
- From the user's perspective, Ns2 is an OTcl interpreter that takes OTcl scripts as an input and generates a trace file as an output. Figure 4-5 shows the components of Ns2.

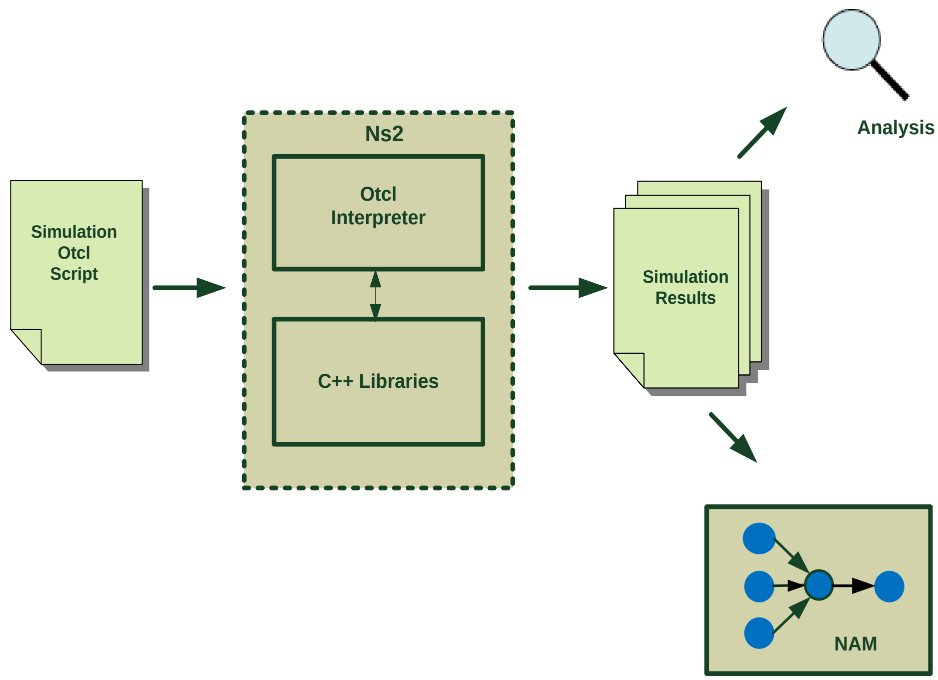

Figure 4-5 Structure of Ns2

In this thesis we have simulated wireless scenarios using Ns2. We begin by creating TCL scripts for wireless simulation. The main component of wireless simulation in Ns2 is a *mobile node*. A mobile node of a network consists of components like Link Layer (LL), Interface Queue (IFQ), MAC Layer. We need to define the type of each of these networks components. Additionally, we need to define other parameters like the type of the antenna, the radio propagation model, the type of Ad-hoc routing protocol (e.g. DSDV, DSR, TORA) used by mobile nodes. Figure 4-6 [58] depicts components of a mobile node in Ns2. We have defined an extra packet header by altering code of NS2 in order to Time-Stamp user traffic as well as probe traffic to collect delay distribution statistics.

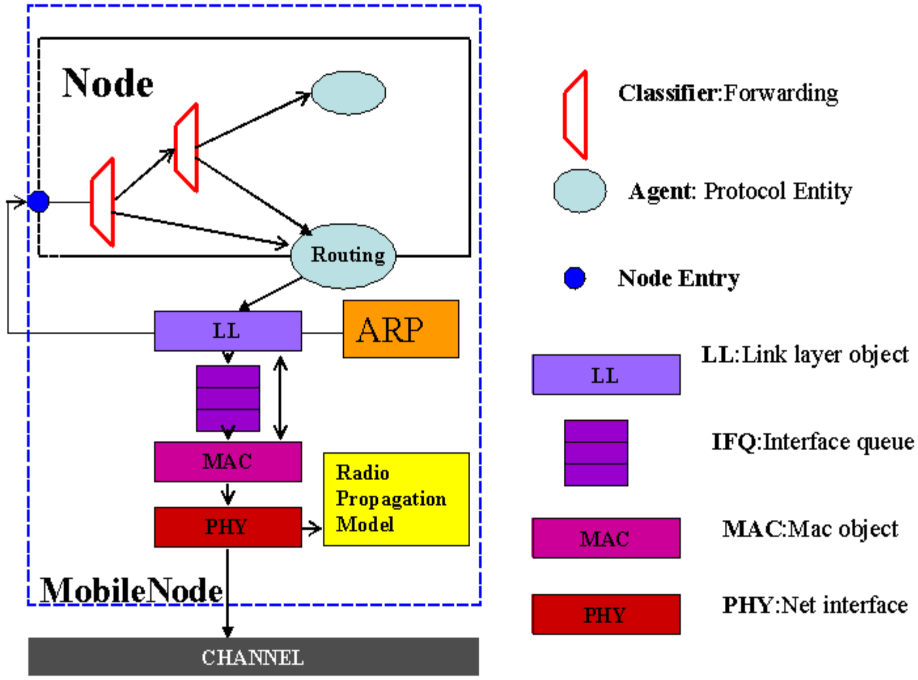

Figure 4-6 General Architecture of mobile node in Ns2

As a result of simulating a wireless scenario, a wireless trace files is generated at the end of simulation run. All the events related to particular simulations are dumped into it. For example stated below is a single event taken from a trace file that looks like:

s 80.000000000 _99_ AGT --- 1800 cbr 64 [0 0 0 0] ------- [98:0 0:0 22 0]

So we can interpret the above trace as:

" Application 0 (port number) on node 99 sent a CBR packet whose ID is 1800 and size is 64 bytes, at the time 80.0 second, to application)on the node 0 with TTL is 22 hops. The next hop is not decided yet".

One of the major disadvantages of NS2 is limited documentation, when compared with other commercial products. However, most of the problems can be solved by mailing lists and discussion groups online. This second shortcoming is that Ns2 consumes a large amount of memory and CPU resources by dumping information into very large files [59].

4.5 *Queueing Behaviour for a FIFO Buffer with Multiplexed ON/OFF Traffic*

This section investigates the behaviour of bursty traffic, modelled by multiplexing ON-OFF sources sharing a buffer. Before investigation we want to give a brief introduction about the two different types of queue behaviour based on the queue length distribution.

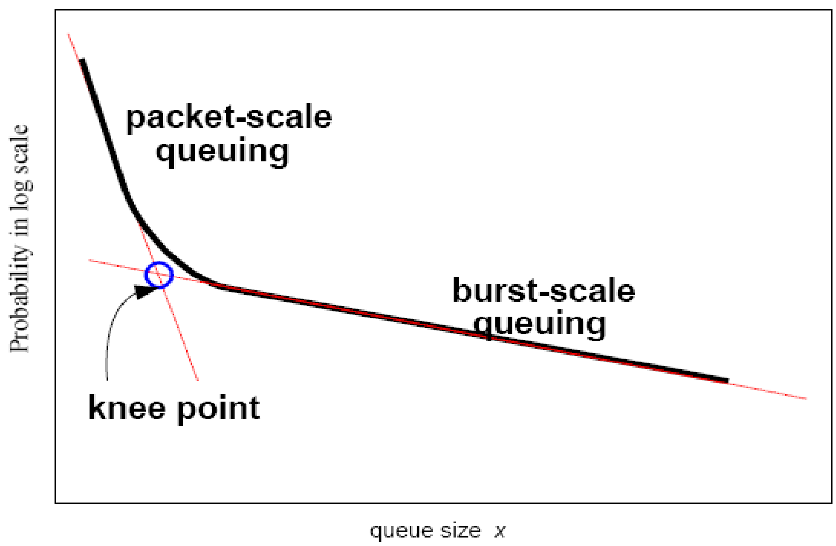

Figure 4-7 Queue length distributions for a FIFO multiplexing Markovian traffic

Figure 4-7 [39] depicts the typical queue length distribution for a FIFO queue multiplexing ON/OFF traffic. With reference to Figure 4-7, the queue length distribution consists of two regions, often referred to as packet-scale region and burst-scale region [28].The queue length distribution starts with the packet scale component with a steeper slope. The distribution then gradually changes to a second branch with a smaller slope, the burst-scale component. The intersection of the tangents to both branches is often designated as the "knee" point [1]. The qualitative explanation of this behavior is given in the following sections:

4.5.1 Packet-Scale Queuing

To account for the packet-scale queuing, it is easier to view the aggregate traffic produced by a group of individual connections. Each connection switches between ON and OFF models. Packets are generated only in the ON state. When the number of connections in ON states is relatively small, the equivalent packet arrival rate is smaller than the queue service rate. Then, the queue is most likely to be empty. Nevertheless, a smaller queue may happen due to the random arrival of the packets. This is called packet-scale queuing as illustrated in Figure 4-8 [28].

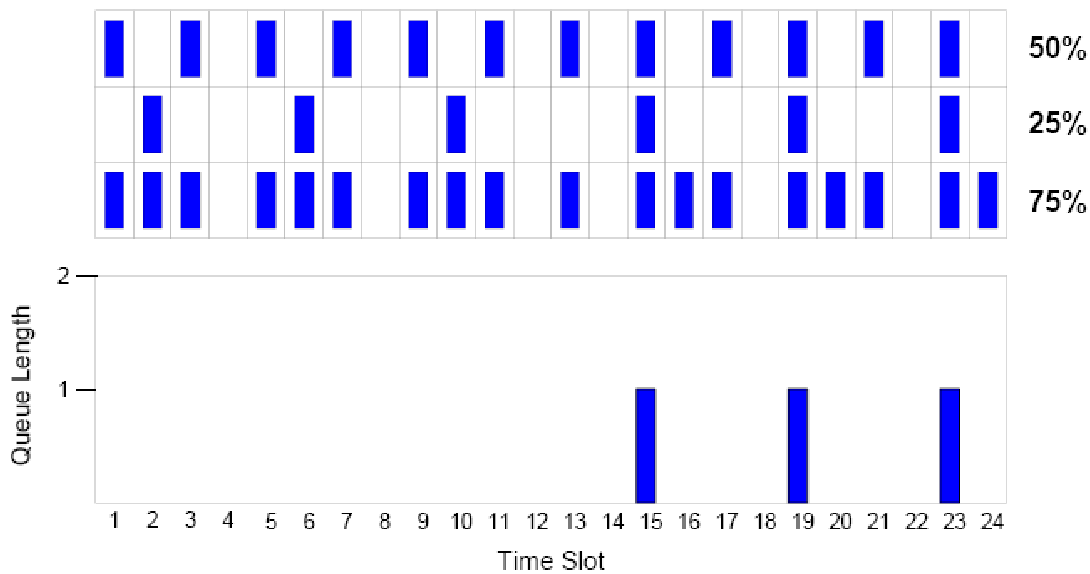

Figure 4-8 Packet-scale queuing behavior

In the example as shown in Figure 4-8, there are three connections. The activity factor of the first, second and third connection are 0.5, 0.25 and 0.75 respectively. Two packets can be removed from the queue during a time slot. Under this configuration, the load is equal to 0.75.

From the diagram, it can be noted that only a short queue exists. In other words, only packet-scale queuing behavior occurs. The small queue is established owing to the pattern of the incoming packets [28].

4.5.2 Burst-Scale Queuing

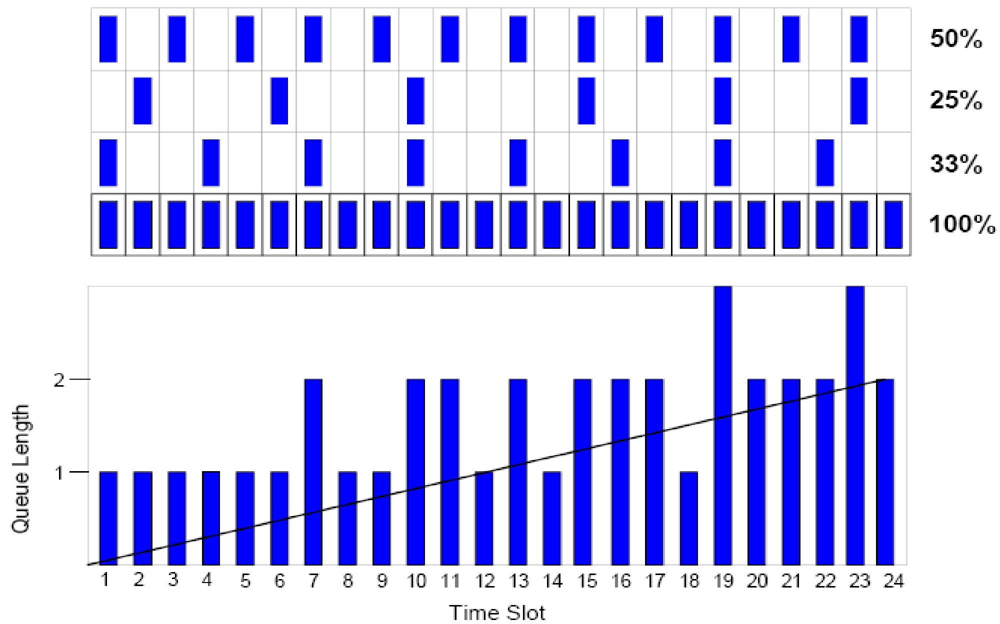

Figure 4-9 Burst Scale Queuing Behavior

Figure 4-9 [28] illustrates the burst-scale queuing effect. Similar to the previous example, two packets can be served in each time slot. In this example, the third connection has the activity factor equal to 33% and an extra connection is added with a unity activity factor, so the load is 1.041. The equivalent arrival rate is greater than the service rate. With reference to Figure 4-9, it can be noted that the queue length is gradually built up as indicated by the black solid line, although some fluctuation is shown.

The queue length distribution can be analyzed with Excess Rate Analysis. The Excess-rate (ER) packets refer to "packets which must be queued as they represent an excess of (instantaneous) arrival rate over the service rate" [29].

As discussed above, the packet-scale queuing behavior is caused by the accidental coincidence of packet arrivals of two or more connections when the simultaneous transmission rate is overall smaller than the service rate. When several connections emit bursts of packets simultaneously, and the transmission rate is greater than the service rate, when it lasts for a considerable longer period than a time-unit, the queue begins to fill and burst-scale queuing happens. The more connections are in a burst period simultaneously, the larger the growth in queue length. But as it is unlikely that all connections send bursts at the same time, the very large queue lengths have a small probability [28].

The knee position, in general, is where just enough sources are active as to overload fill the queue capacity. There are some general trends in this behavior. When the number of connections or the peak rate increases, the knee moves upward, reducing the packet-scale queuing component, and the burst scale component becomes flatter. Increasing the burst length while keeping the load fixed leaves the knee point more or less unchanged but makes the burst component flatter [47].

4.6 Comparing Decay Rate Analysis for Wired and Wireless Access

The simulation experiment conducted in this section is to check if we can get agreement on the decay rate (for queue length) for wired and wireless scenario (if the headers effect is cancelled out for the wireless scenario).

4.6.1 Simulation Topology and Model

The topology used is the same for both wireless and wired scenarios except that for the wired topology we have a buffered link as shown in the Figure 4-10 and Figure 4-11. We have two nodes with one edge node that transmits the packets to the destination node. For the wired case it consists of a single link, and for the wireless case both nodes are in hearing range[1] of each other. The traffic sources used are ON/OFF with exponentially distributed ON and OFF periods.

The arrival process of any single source is deterministic and the sources only produce packets when 'ON'. In the experiment, the queue size distribution is obtained at different values of high load conditions (in order to observe the burst scale decay rate) by adjusting the arrival rate of traffic and the number of ON/OFF sources. The values for parameters are typical values used for digitized voice taken from [28].

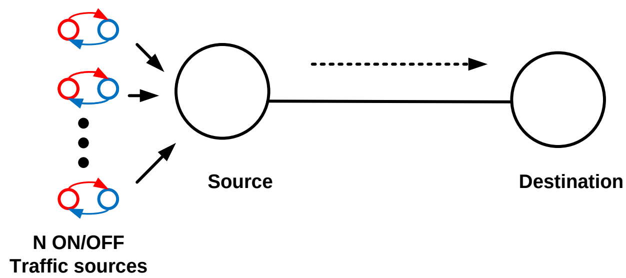

Figure 4-10 Wired Scenario Topology

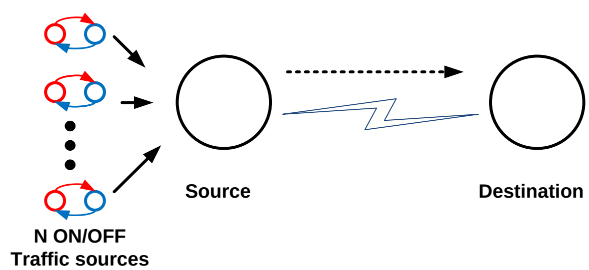

Figure 4-11 Wireless Scenario Topology

The Following parameters have been used

Packet Size = 1000 bytes

T_{on} = 0.34

T_{off} = 0.67

R= 85600bps

C = 1Mbps

Where,

T_{on} = mean ON time of a single ON-OFF source (sec)

T_{off} = mean OFF time of a single ON-OFF source (sec)

R= arrival rate of a single on-off source during the ON period (packets/sec)

α = activity factor i.e. $T_{on}/(T_{on} + T_{off})$

C= capacity of transmission link

4.6.2 Simulation Results

According to the Figure 4-12 and Figure 4-13 the burst scale decay rate for wireless and wired scenarios seems to be very close. Whereas, in reality it is not like that because there are overheads of extra headers in wireless simulations due to which the burst scale decay is observed at a much lower value of the load. In this set of simulations, we have balanced these overheads and noticed that in reality 10% of the wireless link bandwidth is used by the packet headers. Here we want to mention that, in the wireless scenario, only two nodes are communicating with each other and there is no competition for medium access, as one is source and the other one is destination. The medium access effect (medium contention) on measurements (for delays and losses) is considered in chapter 5 and chapter 6. Figure 4-12 reveals the queue state probability for a wired buffered link and Figure 4-13 illustrated the burst scale decay trend at load values of 0.75, 0.8. 0.85 and 0.9 respectively in wireless scenarios.

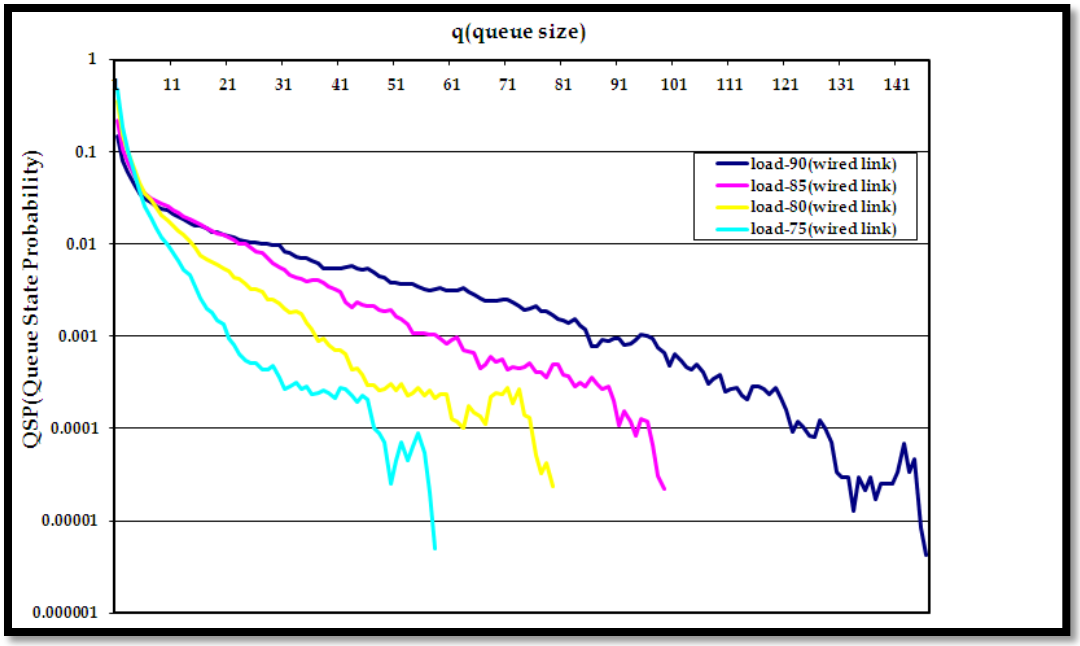

Figure 4-12 Queue State Probability of a buffered wired link under different Load values

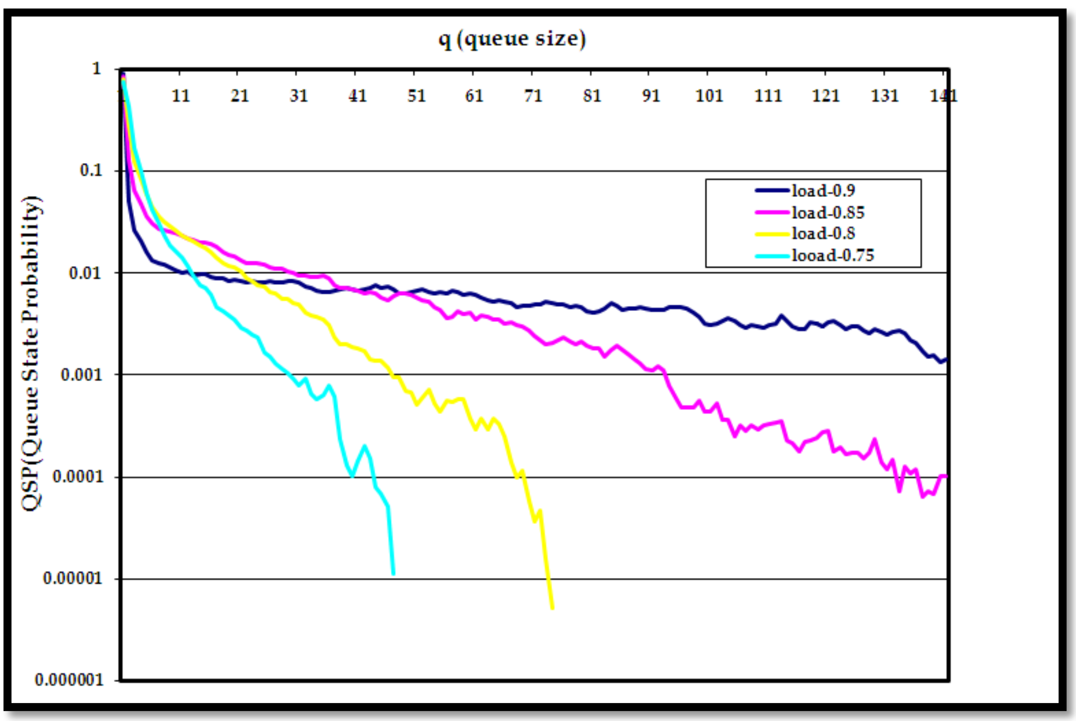

Figure 4-13 Queue State Probability of a wireless node under different Load values

This set of simulation experiments has been performed to analyze the wireless queue length distribution and compare it to that of wired. We know that queueing occurs between Link Layer and MAC Layer in wireless setup (can be seen from Figure 4-7) whereas; in a wired scenario a buffered link is used to monitor queue length. Our evaluation reveals that if all the conditions are kept ideal for wireless medium i.e. minimum interference, the nodes are in hearing range and no contention effects and subtracting the header effects the queue length distribution obtained is quite similar to a wired case. These experiments gave us a base for future comparison as measurement may get more critical over wireless medium because of above mentioned factors, the results shown in Table 4-1 seem quite closer as after cancelling the extra wireless header effects. We have also compared our decay rate values with the analytical formula presented in [48]. Appendix A describes the analytical formula derivation for DR analysis.

Load	**DR (wireless)**	**DR (wire)**	**Vindya's DR formula(modified)**	**Vindya 's DR formula(original)**
0.9	0.9704	0.9647	0.978	0.994
0.85	0.9389	0.9292	0.937	0.986
0.8	0.8885	0.8857	0.867	0.963
0.75	0.7899	0.7715	0.763	0.957

Table 4-1 Decay Rate Comparison

Load	Queue -Length(wireless)	Queue Length (wire)
0.9	23	19.21
0.85	12	7.13
0.8	4.03	3.15
0.75	1.87	1.51

Table 4-2 Queue Length Comparison

From Table 4-2 it can be seen that the values of queue length for both wired scenario and wireless scenario are quite closed when the wireless packets are experiencing no MAC header effect. Figure 4-14 shows the queue length comparison graphically.

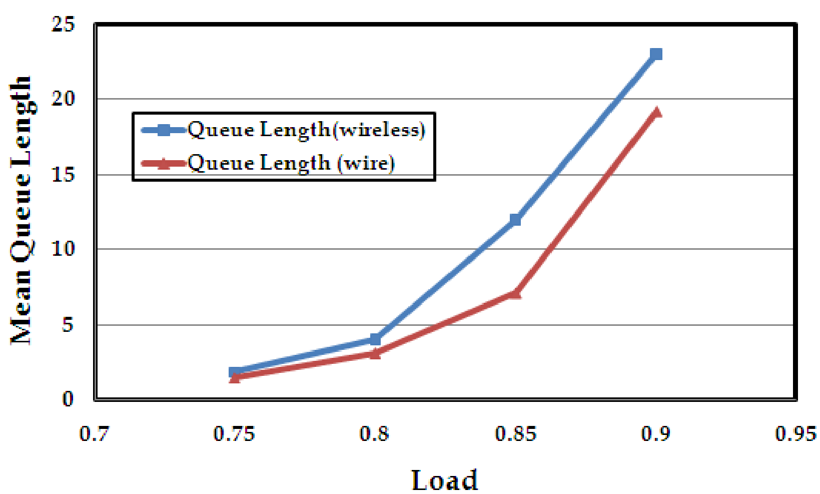

Figure 4-14 Queue Length Comparison graph

The formula used to compare the simulation results proposed by [48] for Burst Scale decay rate analysis

$$R \rightarrow \left[\frac{1 - \left[\frac{\ln\left(h/C\right)}{\ln(\rho)} + \frac{\left(h^2 T_{on} \rho\right)}{\left(C(1-\rho)^2\right)} \right]^{-1}}{1 - \left[\rho(1-\rho)^2 \Big/ (h/C) \cdot T_{on}.[(1-\rho)C + h \cdot \rho)] \right]} \right]$$

Equation 4-1

Where,

h = arrival rate of a single on-off source during the ON period (packets/sec)

C = capacity of transmission link (Mbps)

R = decay rate

ρ = value for load

Above stated is the original formula but it works well with short packet sizes if we remove $\frac{\ln\left(h/C\right)}{\ln(\rho)}$ factor which is for short packet size correction. We get more accurate results as stated in Table 4-1.

4.7 Summary

In this chapter, we have described different types of simulation models used in network simulation. We also give an account of the different types of traffic models relevant to this research i.e. short range dependent and long range dependent traffic. We explain the choice of model used in the research: the ON/OFF model and the simulator employed for conducting the simulation experiments: the Ns2 Simulator. We highlight the queuing behaviour for a buffer with multiplexed Markovian traffic where queue length can be split into two regions: the packet scale region and burst scale region. Finally, we give a simulation based comparison for queue length and decay rates for wired and wireless scenarios and compared the results with some recent research presented by [48].

5. Measurement of Delay Distribution over Wireless Access using Active Probing

In this chapter the main motive is to assess the accuracy of active probing for determining the end-to-end delay distribution of a single hop and well as multiple hops wireless networks. The current approach of end-to-end delay performance assessment in most practical networks is mainly based on active probing. In this section we will perform the end to end delay measurement using active probing. In the following section we briefly discuss the active measurement mechanism.

5.1 Active Probing Mechanism

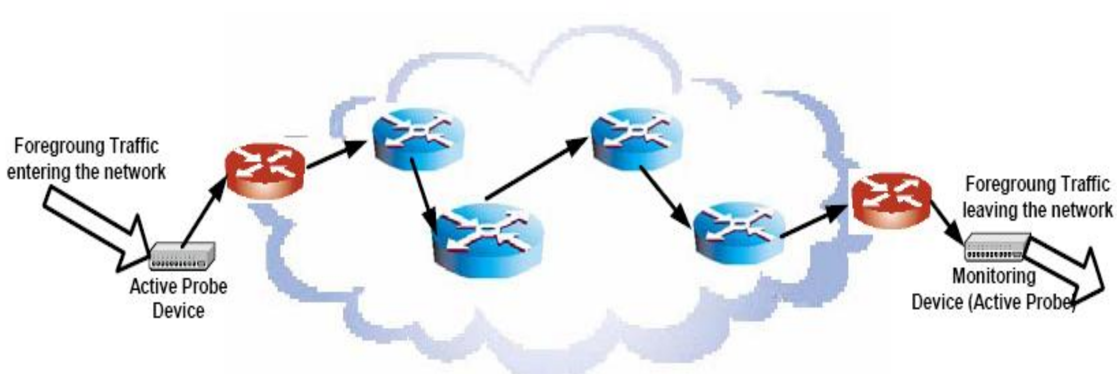

Figure 5-1 Active Probing

Figure 5-1 depicts the active probing framework .The traffic of interest is referred to as the *foreground traffic*. There exist active probing devices at both ends of the path which the foreground traffic traverses. At the sending side the active probing device is responsible for generating the probe packets. These probe packets are time stamped and received at the receiving end by the monitoring device. In our simulation we have connected a separate sink for collecting the probe packets and to measure the end-to-end delay. Both the user traffic and probe traffic are 'Time-Stamped' in order to get the delay distribution statistics for both traffics. These statistics are used for comparison. The whole experimental setup is designed using network simulator NS2.

Active probing is based on sampling. The testing packets are a small fraction of background measured traffic to prevent overloading of the available resources.

5.2 *Sampling Techniques*

The purpose of the testing packets is to sample the end-to-end packet delay data from its parent population (in our case parent population is the background traffic). In [39], the authors investigated various sampling techniques for QoS performance measurement, simple random, stratified random and system sampling. These are described in more detail in the following sections.

5.2.1 Simple Random Sampling

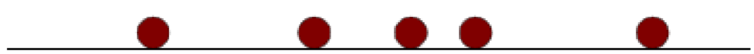

Figure 5-2 Random Sampling

Figure 5-2 shows simple Random Sampling, The sampling takes place randomly during the measurement period. Random Sampling alleviates any problem associated with periodicity. [62], reports that the duration between the sample times may have an exponential distribution in order to achieve memory-less effect so that the correlation between the samples can be minimized.

5.2.2 Stratified Random Sampling

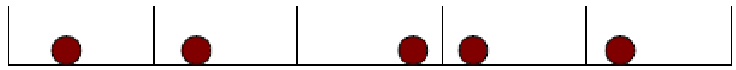

Figure 5-3 Stratified Random Sampling

Stratified Random Sampling splits the population into *N* sub-populations. These populations do not overlap and they cover the whole of the original population. The sampling events take place randomly within the sub-population. As for Simple Random Sampling, Stratified Sampling alleviates any problems associated with periodicity [62].

5.2.3 Systematic/Periodic Sampling

Figure 5-4 Systematic Sampling

Systematic Sampling takes place every *T* time units as illustrated in Figure 6.4. This sampling method is employed in our experiments for the following reasons.

Generally, the larger the sampling size, the greater the accuracy of the measurement result will be obtained. To study the performance of active probing, we would like to make the sampling size a fixed portion of the foreground traffic to limit the measurement cost as discussed above. Since the amount of the foreground traffic during the measurement period is usually unknown in advance, therefore, Systematic Sampling is easier to be implemented compared with Random Sampling if we want to limit the sampling size to be a fixed portion of the foreground traffic.

Stratified Random Sampling is also able to ensure the probing traffic volume is a fixed portion of its parent traffic. Stratified Random Sampling provides better sampling results than Systematic Sampling provided that a prior knowledge of the correlation structure of the parent population is available [4]. The splitting size as shown in Figure 5-3 becomes the size having small correlation in the parent population. Nevertheless, this prior knowledge is not normally available during measurement. However, we have also used Simple Random Sampling and compared our results with those of Periodic Sampling.

5.3 *Probe Requirements*

The probing technique used in this thesis will need to report on the following network characteristics

- Delay distribution
- Packet loss ratio

The probes will measure these metrics by flowing alongside the actual traffic. The characteristics of these probes will be recorded separately from background traffic. We have considered three attributes of probes while simulating in order to minimize probes interference with traffic and capture the actual characteristics of the traffic. These three probe attributes are:

- Probe size
- Probe rate
- Probe type

The probe type used is periodic and random as discussed in section 5.2.3. The probe size varies between the smallest possible packet size which is 40 bytes and the largest packet size 1000bytes (In the literature largest the packet size used is 1500 bytes, but we have used 1000 bytes, as NS2 splits the packet into two if its greater than 1000 bytes and this will result in large amounts of overhead, distorting the simulation results). The rate of probes depends on the size of the probes as higher probing rates coupled with a large size will result in severe disruption of the network traffic. However, higher probing rates provide more accurate results, as there will be more samples.

5.4 The Simulation Setup

We have evaluated the delay distribution of foreground (probe traffic) and background traffic (data traffic) for three different scenarios under varying load conditions.

I. Single Hop
II. Two Hops
III. Three Hops

The above three scenarios are probed for three different probing rates

I. 1 probe every 10 second (periodic probing)
II. 1 probe per second (periodic probing)
III. 10 probes per second (periodic probing)
IV. 5 probes per second (periodic probing)

The objective of these set of simulation results is to investigate firstly how accurately the probes measure of delay in traffic. Secondly by increasing the probe rate we discovered how much the interference caused by the probe traffic affects the actual delay distribution. Figure 5-5 lists all the simulation scenarios

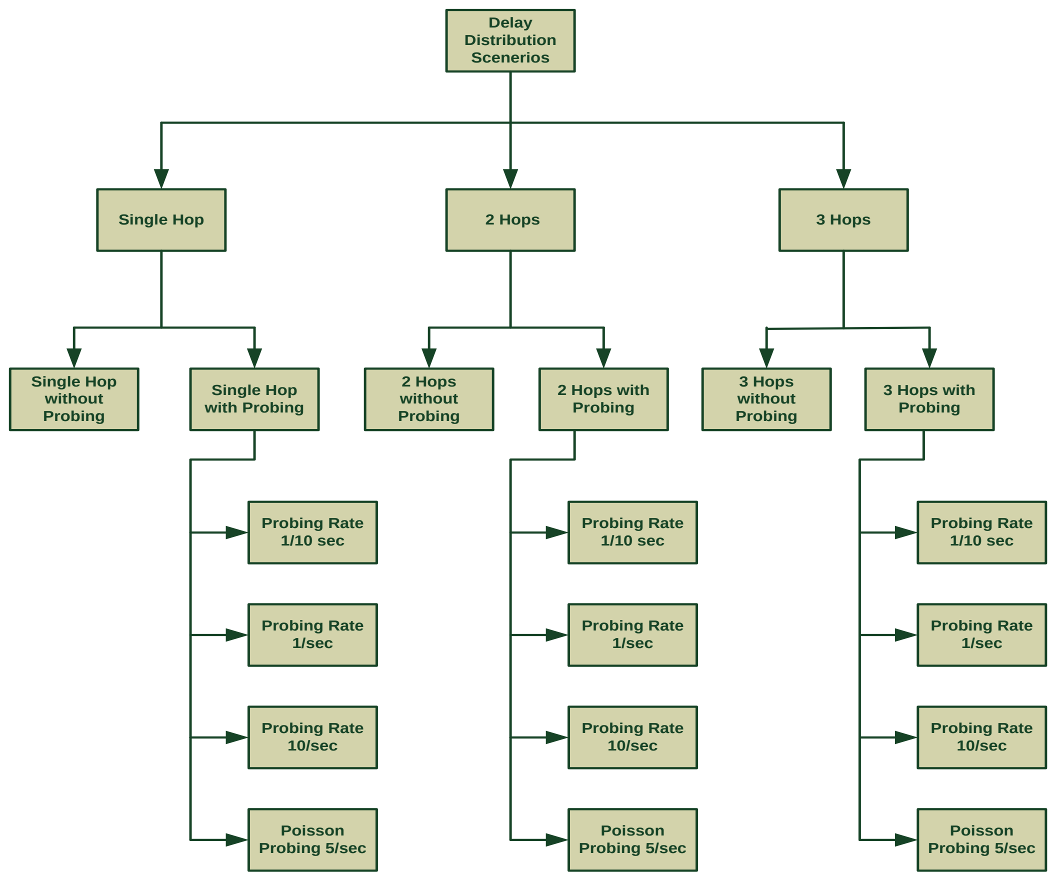

Figure 5-5 Simulation Scenarios

5.4.1 Effect of Packet Size

Before we describe the simulation topology and experimental analysis we will justify why we used the packet size of 1000 bytes for all simulation experiments. To justify this we performed a simple experiment simulating two different packet sizes for delay and packet delivery ratio as well as throughput. From the Figures 5-6, 5-7 and 5-8 it is obvious that when we used a small packet size i.e. 200bytes for transmission over a wireless access, delay suffered by the traffic is far greater than that suffered by a bigger size packet i.e. 1000bytes. The reason behind is, although the packets are small but the headers attach to them are same size to that of bigger packets which becomes an extra overhead leading to increase in delay. Therefore, the bigger packet size is used. In this way we can save some bandwidth, and packet delivery ratio can be improved. Research in [65] also support this finding stating " packet size optimization could facilitate efficient usage of wireless network resources, improving service provided to all end users sharing the network".

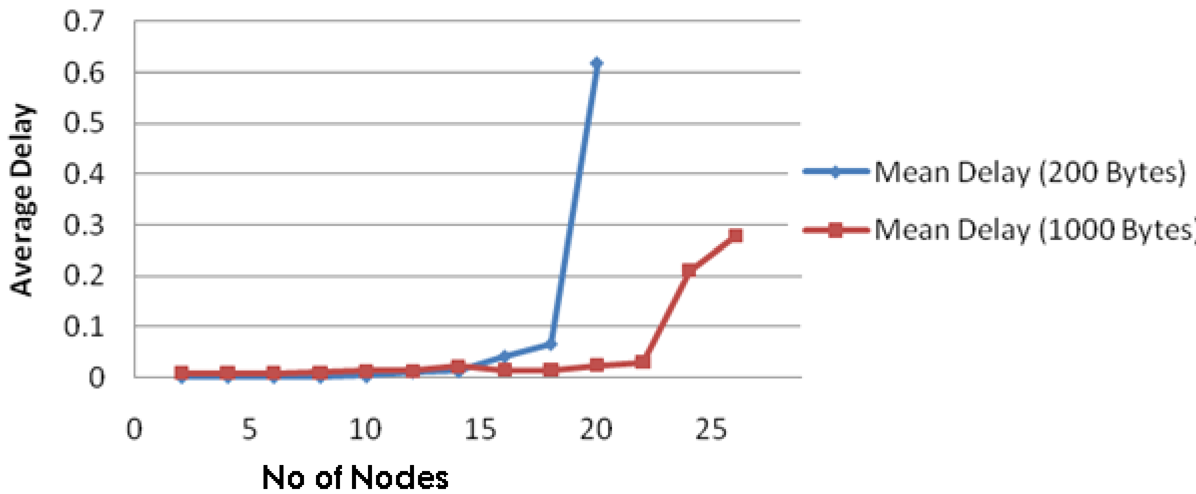

Figure 5-6 Delay comparison for different packet sizes

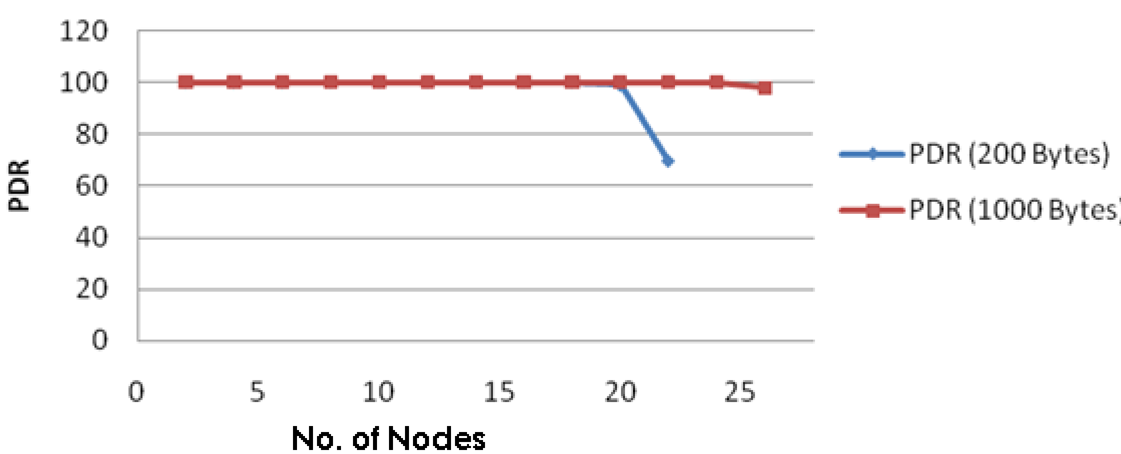

Figure 5-7 Packet Delivery Ratio Comparison for different packet sizes

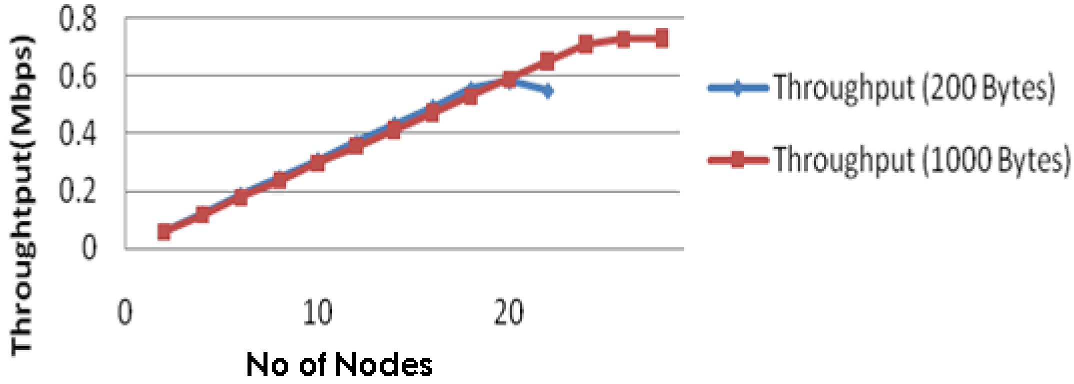

Figure 5-8 Throughput comparison

5.4.2 Effect of Probe Packet Size

The packet size for both probes (foreground) and user (background) traffic is the same for our investigations, because if we use the small size of probe packet in order to minimize the probing overhead it will lead to a larger difference in delay values when compared to actual traffic delays. This is confirmed by the results shown in Figure 5-9 which provides a comparison of average delay in which the size of probe packet is 40 bytes (with probing frequency of 1 probe per sec) whereas the packet size of user traffic is 1000 bytes.

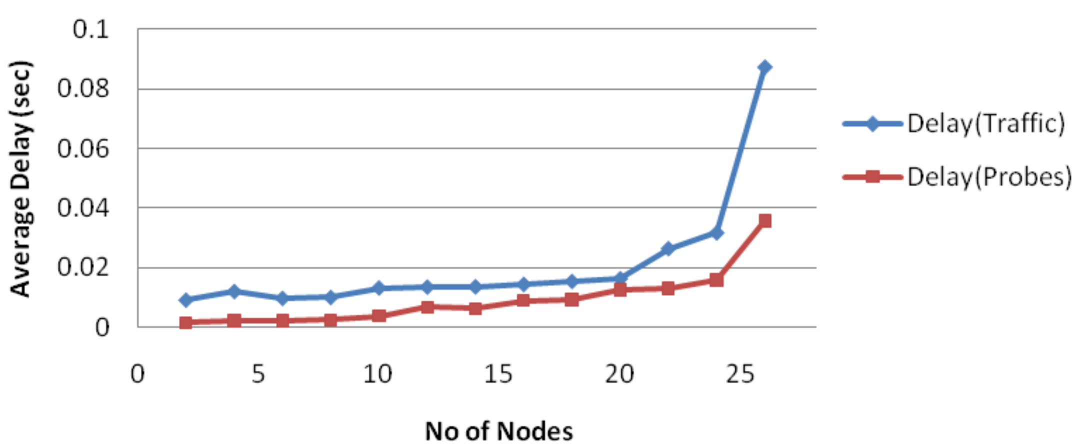

Figure 5-9 Comparing average delay for different traffic packet size(1000 bytes) and probe packet size (40bytes)

5.5 Delay Distribution for Single Hop Scenario

5.5.1 Simulation Topology and parameters

In our single hop simulation scenario the traffic is monitored by an active probing device that is responsible for generating testing packets (i.e. probes). These testing packets are time stamped and received at the receiving end by the probe sink that measures the end to end delay. However, data traffic is received at a traffic sink. Figure 5-10 describes the simulation topology. The wireless nodes are within the hearing range of each other. The simulated traffic parameters used are listed in the Table 5-1. The simulation parameters used for simulating all scenarios are listed in Table 5-2.

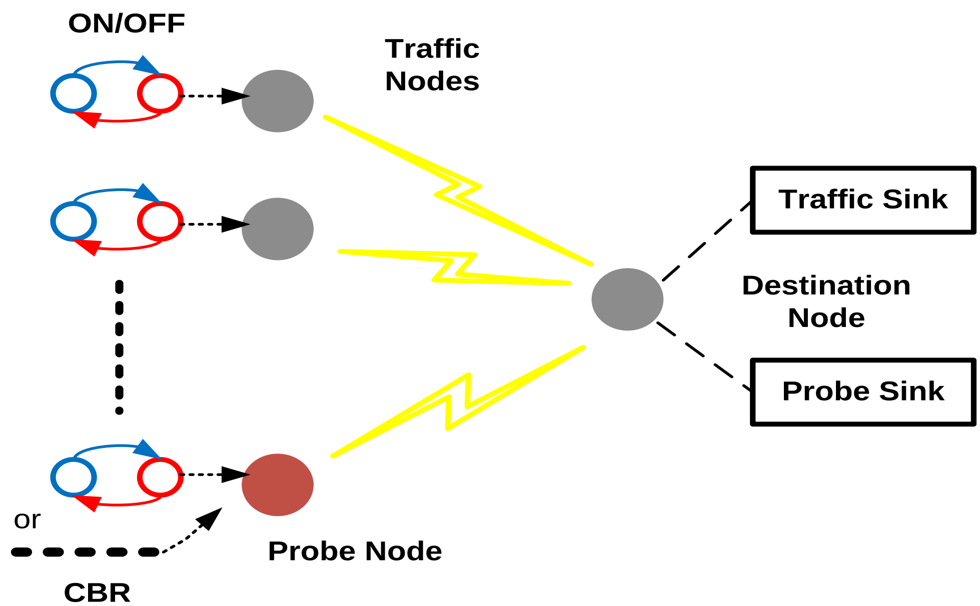

Figure 5-10 Simulation Topology for Single Hop Scenario

Traffic type	Exponential ON/OFF
Application Type	UDP
Probe Traffic Type	CBR
Packet Size (data traffic)	1000 bytes
Packet Size (probes traffic)	1000 bytes
Mean ON time of data traffic (Ton)	0.35 sec
Mean OFF time of data traffic (Toff)	0.67 sec
Input data traffic rate (R)	85.6kbps
Probes rate (Periodic probing)	10/sec, 1/sec, 1/10sec
Probes rate (Poisson probing)	5/sec

Table 5-1 Traffic Parameters

Simulator	**NS version 2.31**	
	Network Size	1000x500 m
	No of Nodes	Min=2, Max=28
	Simulation Duration	3600 sec
Physical Layer	**Signal Propagation Model**	Two Ray Ground
	Max Transmission Range	250m
	Antenna Type	Omi directional
Mac Layer	**MAC Protocol**	802.11
	Link Bandwidth (data rate)	1Mbps
	Interface Queue Type	Drop Tail/Pri-Queue
	Interface Queue Size	8000 packets

Table 5-2 NS2 Parameters used in Simulation

5.5.2 Simulation Results

In this section we provide graphs of the delay probability experienced by the user traffic and the probes. All the delay distribution graphs given in Figure 5-11 to Figure 5-19 are for sampling rates listed in Table 5-1. In the graphs "Background traffic" stands for *user traffic* and "Foreground traffic" stands for *probes*.

It can be clearly seen from the Figure 5-11 that when there are just two nodes sending traffic to a destination node (along with a probe generating node). The mean delay experienced by the user traffic and probes is almost the same, and so is the variance in delay distribution which can be seen from the values given in the Table 5-3. This results in a very small, almost negligible, sampling error especially for the case when we are probing with the probing frequency of 10 probes per second (periodic probing) and 5 probes per second (random probing).

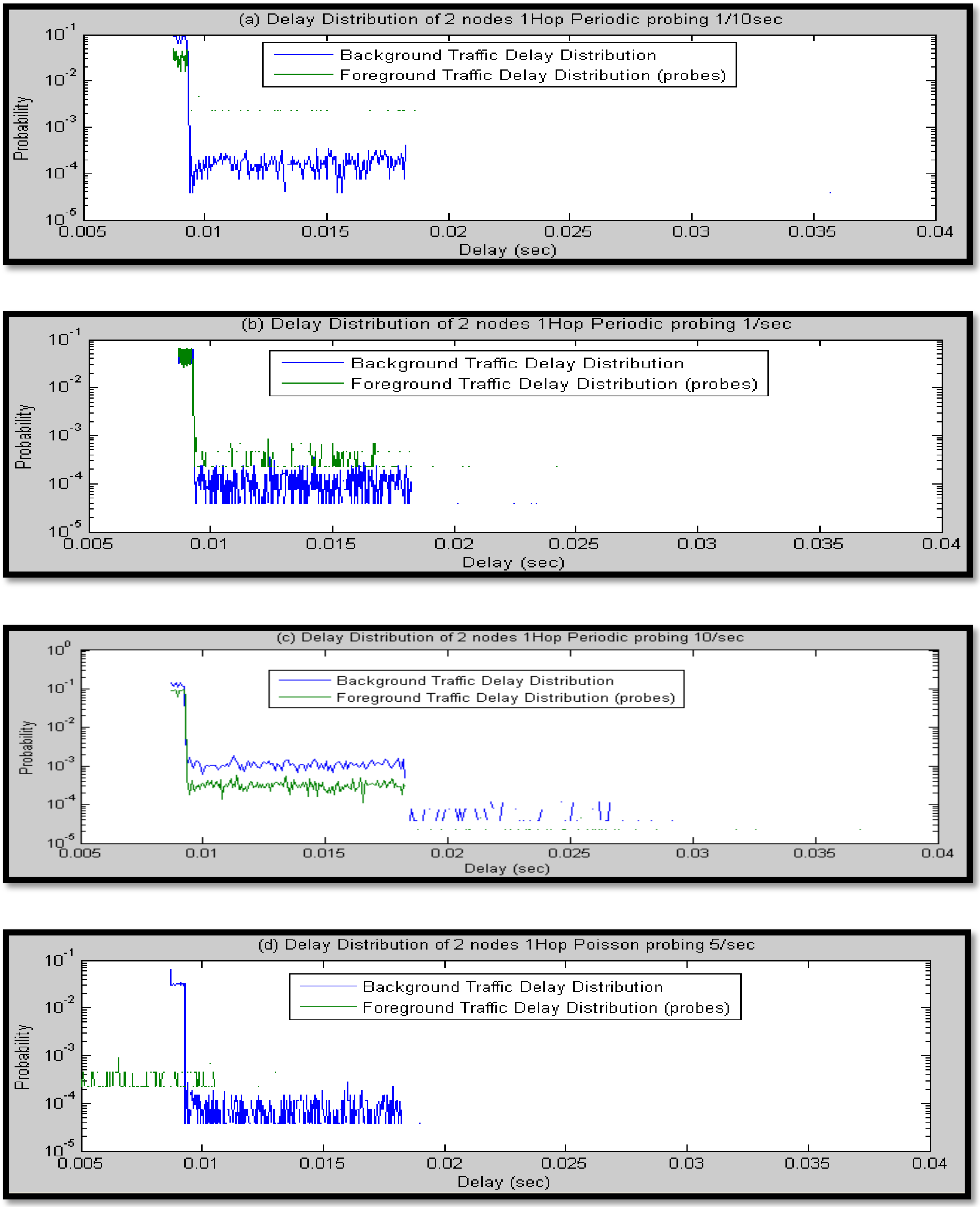

Figure 5-11 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 2 nodes scenario at different probing rates

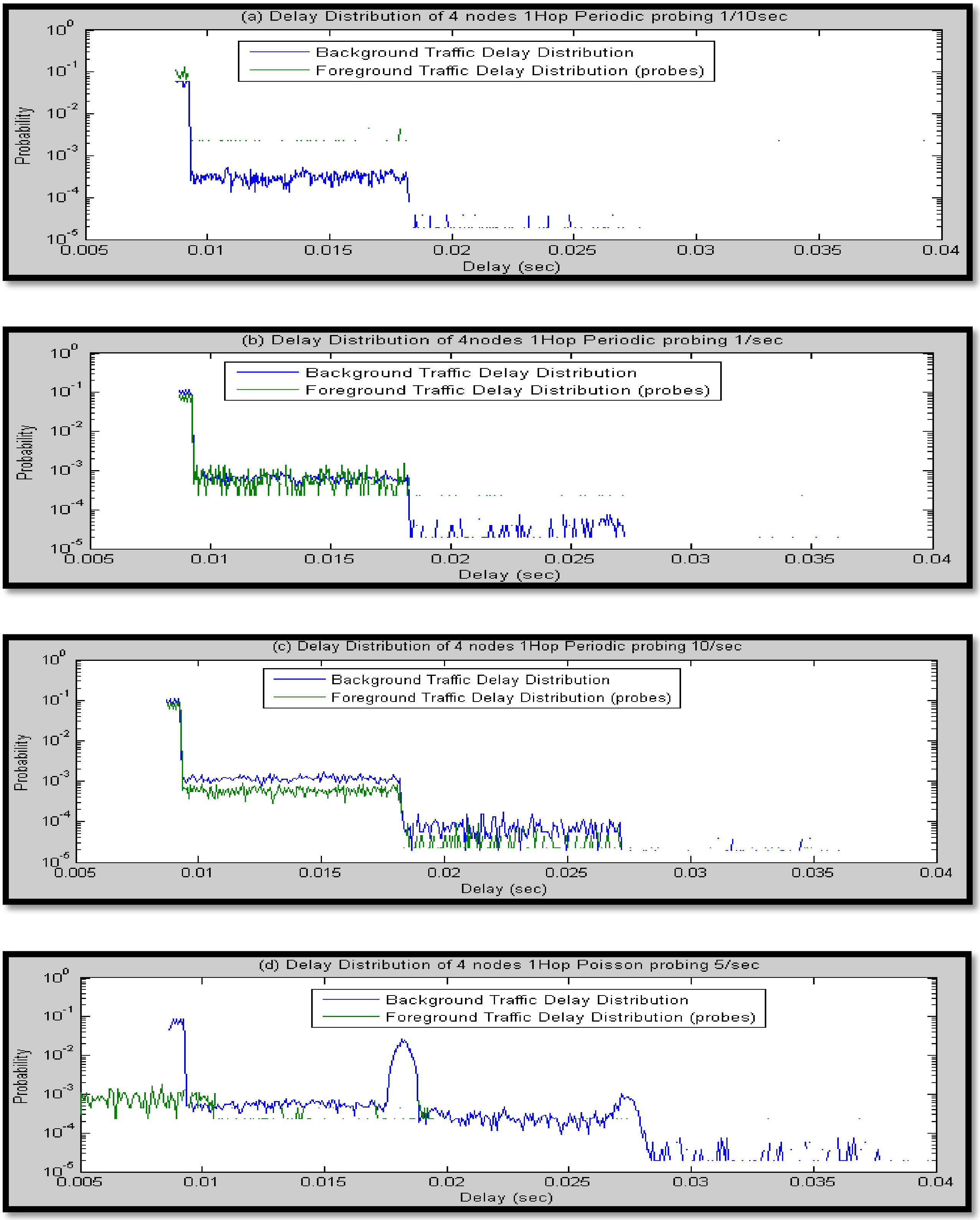

Figure 5-12 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 4 nodes scenario at different probing rates

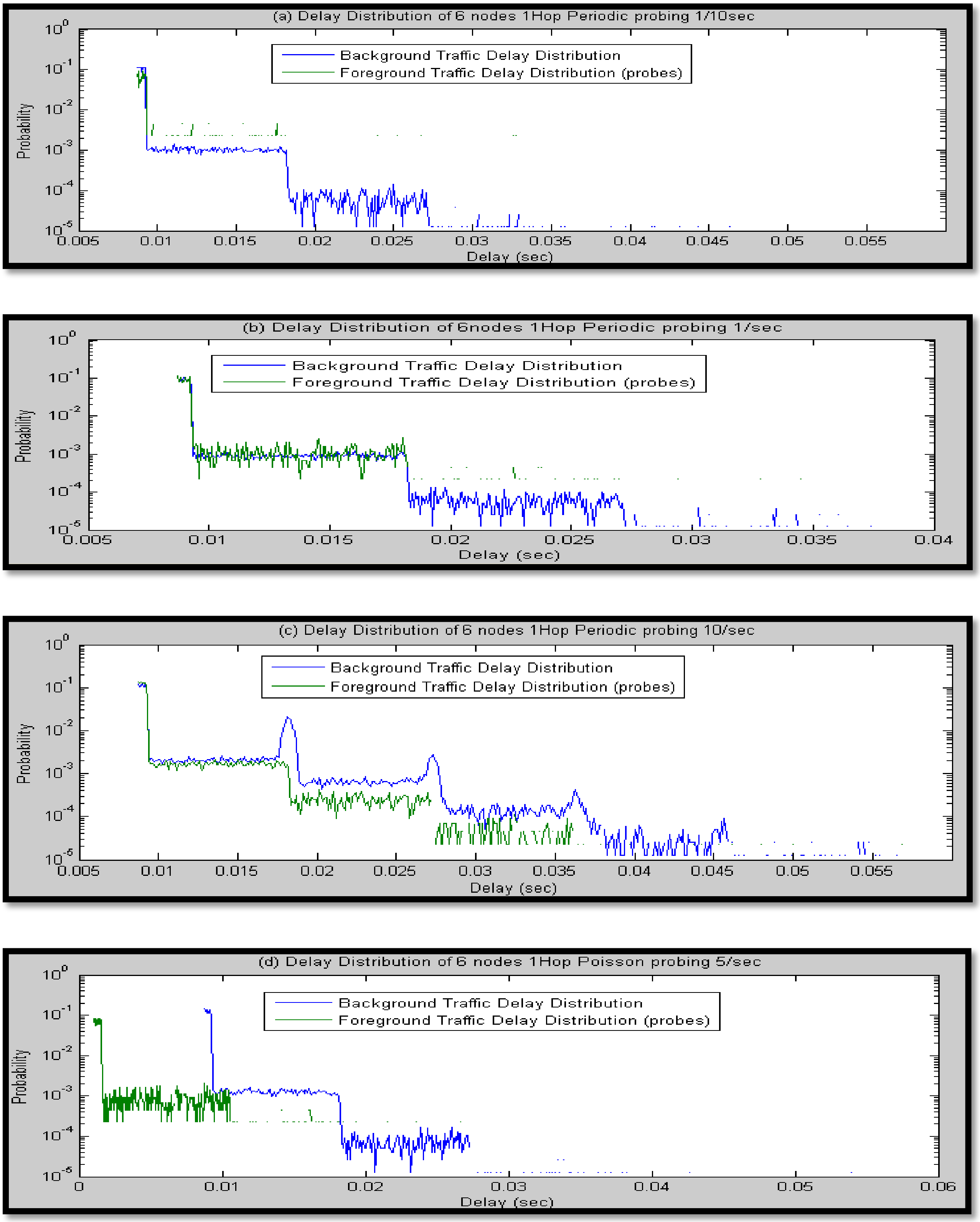

Figure 5-13 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 6 nodes scenario at different probing rates

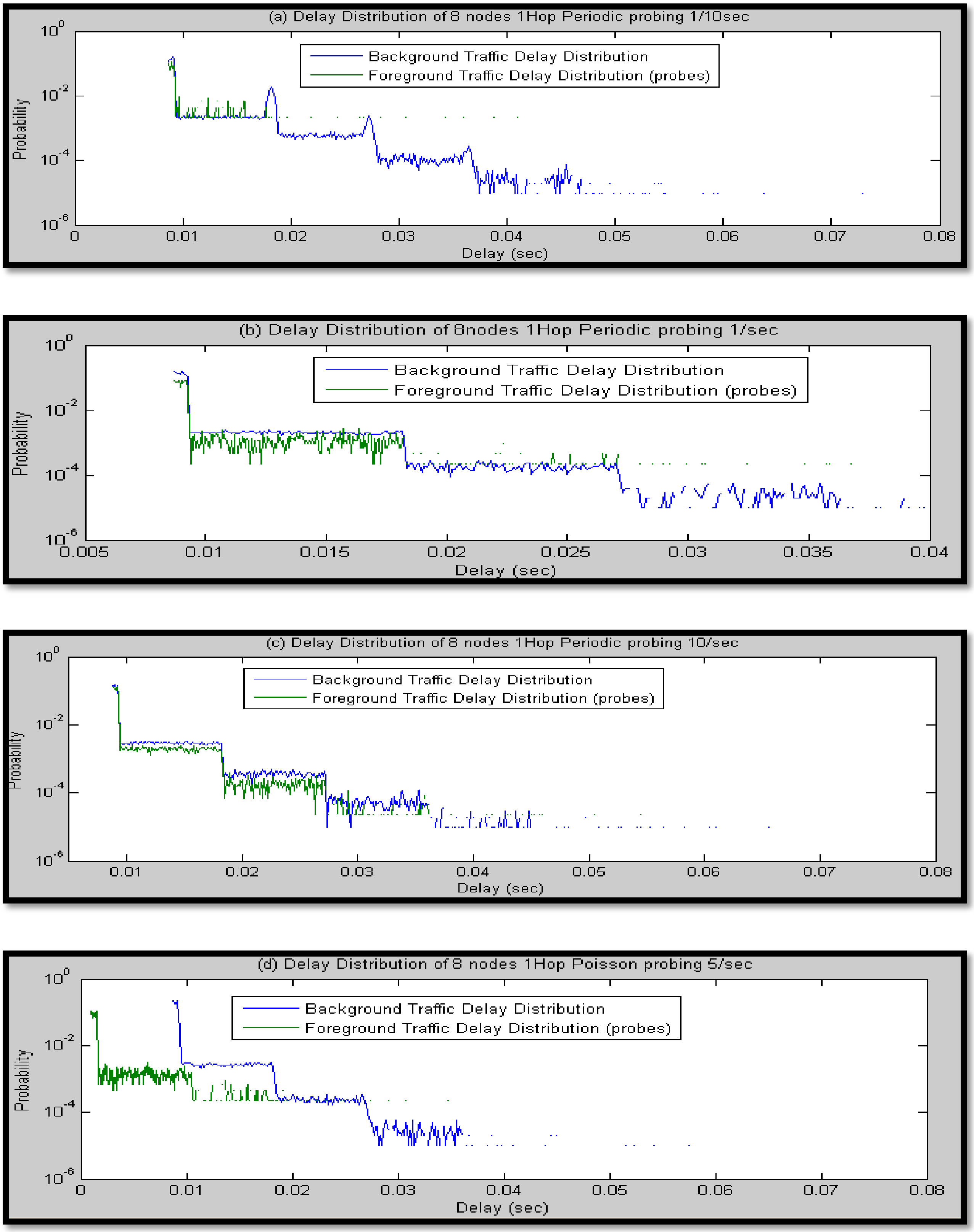

Figure 5-14 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 8 nodes scenario at different probing rates

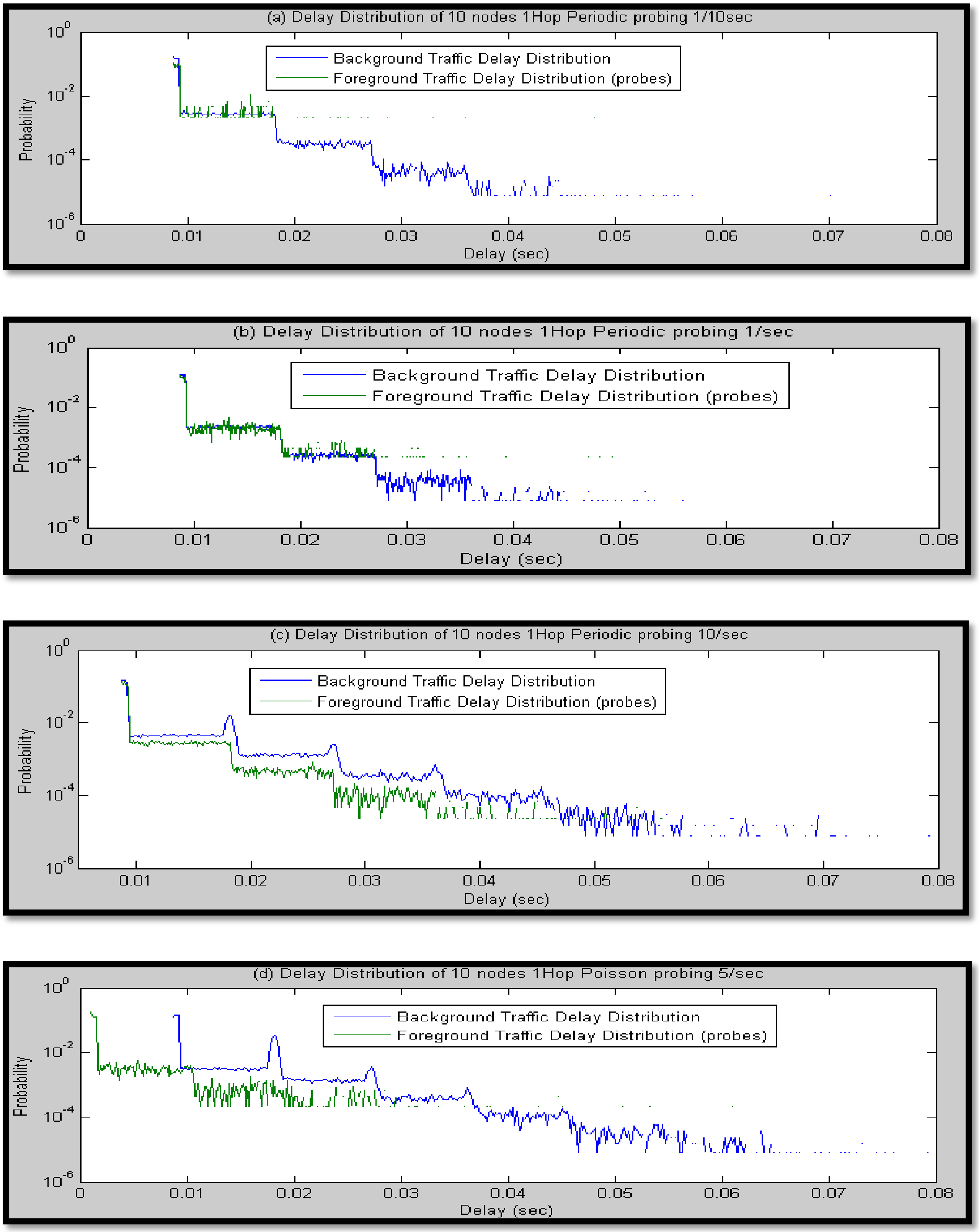

Figure 5-15 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 10 nodes scenario at different probing rates

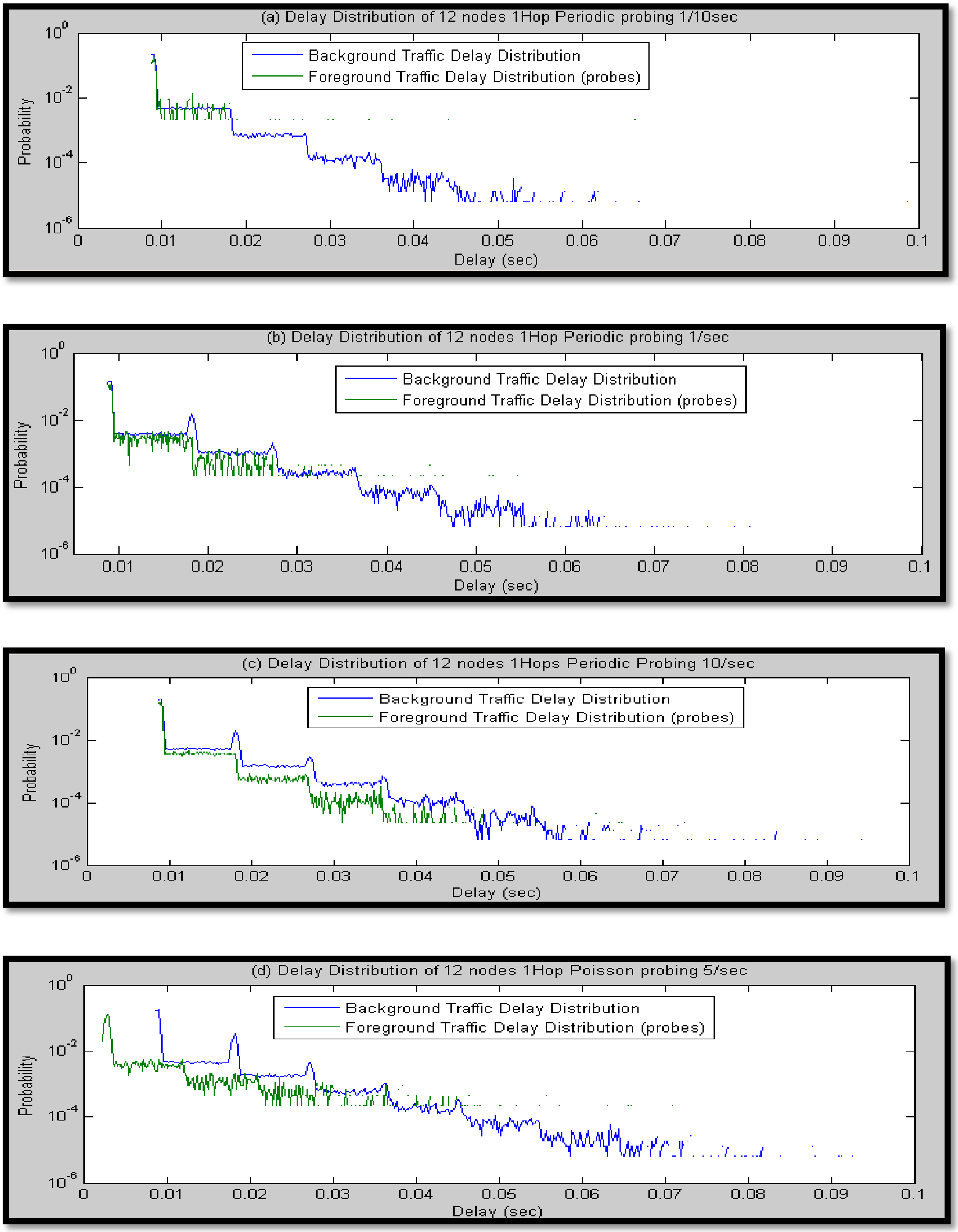

Figure 5-16 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 12 nodes scenario at different probing rates

In our simulations, when the number of wireless nodes contending for the medium increases, we observe a step like graph as shown in Figures 5-12, 5-13, 5-14, 5-15, 5-16. The reason behind this is the back-off window explained in chapter 3; section 3.2.2, i.e. when a collision is encounter and the size of the contention window is increased exponentially in order to reduce the collision probability. As a result the access delay in the network increases. When the number of collisions increases as a result of increase in number of wireless nodes contending for the medium, the size of the window increases as well until it reaches the maximum.

Currently the 802.11 DCF resolves collisions through multiple levels of CWs and back-off stages. In the initial back-off stage the value of CW has the minimal value CW_{min}. After each collision, the CW will be doubled until reaching the maximum CW_{max}. After each successful transmission, the back-off will resume at the initial stage and the CW will be reset to *CW*min regardless of the network condition or the number of competing nodes. By resetting the CW to *CW*min, DCF increases the probability of collision and frequent retransmissions remain high until the CW attains appropriate values(The back off value is random between 0 and CW_{max}).This is obviously not optimal since a high collision rate in the network means poor network exploitation. On the other hand, the intrinsic back-off randomness makes it difficult to instantaneously absorb an increasing number of flows. The back-off process is basically intended to reduce the collision rate when using a higher contention stage.

When the window size fluctuates more frequently, the graph becomes smoother and starts showing the burst-scale decay, which can be clearly observed from Figure 5-17 to 5-19.

The distribution is presented in Figure 5-18 for 20 wireless nodes sending data to the destination node. In Figure 5-18 (d) represents the Poisson Probing comparison plot for delay distribution showing that shape of the delay distribution of user traffic i.e. referred as background traffic in graph and probes traffic i.e. Foreground traffic are not in accordance with each other. However in this case a large number of probing packets were injected leading to increased delay as compared to Figure 5-18 (a) and (b). In (a) the sampling frequency is 1/10 sec i.e. 1 probe sent after every 10 sec. It can be seen that the distribution measured by the active probing is "floored" at about $2.4x10^{-3}$. This is because the number of probing packets generated during the measurement period is nearly 416, whereas in case (b) when probing frequency was 1 probe per second nearly 4400 were generated (this is obvious from the graph where the floor of the active probing distribution is lying at approximately 2.27×10^{-4}). (c) It is the obvious case of sampling traffic at high rate (i.e. in our case 10/sec) which results in overloading the network by injecting a large number of probes. This leads to higher delay values yet it is not able to measure the traffic delay distribution fully.

Figure 5-19 (a) shows that the probing rate of 1 probe every 10 second is not sufficient to capture the delay distribution of the user traffic as there are not enough probes generated. In Figure 5-19 (b) although the probes are injected every second yet they are not sufficient, whereas in Figure 5-19 (c), when we increase the sampling rate to 10 probes per second, the network becomes overloaded leading to much higher values of delays than actual user traffic delay, which is not desirable. We then repeated our investigation using Poisson probing with the probing rate of 5 probes every second. In this case, although sufficient probes were generated to map actual traffic distribution, but delay is increased to larger values and the tail of distribution still remained poorly resolved, as shown in Figure 5-19 (d)

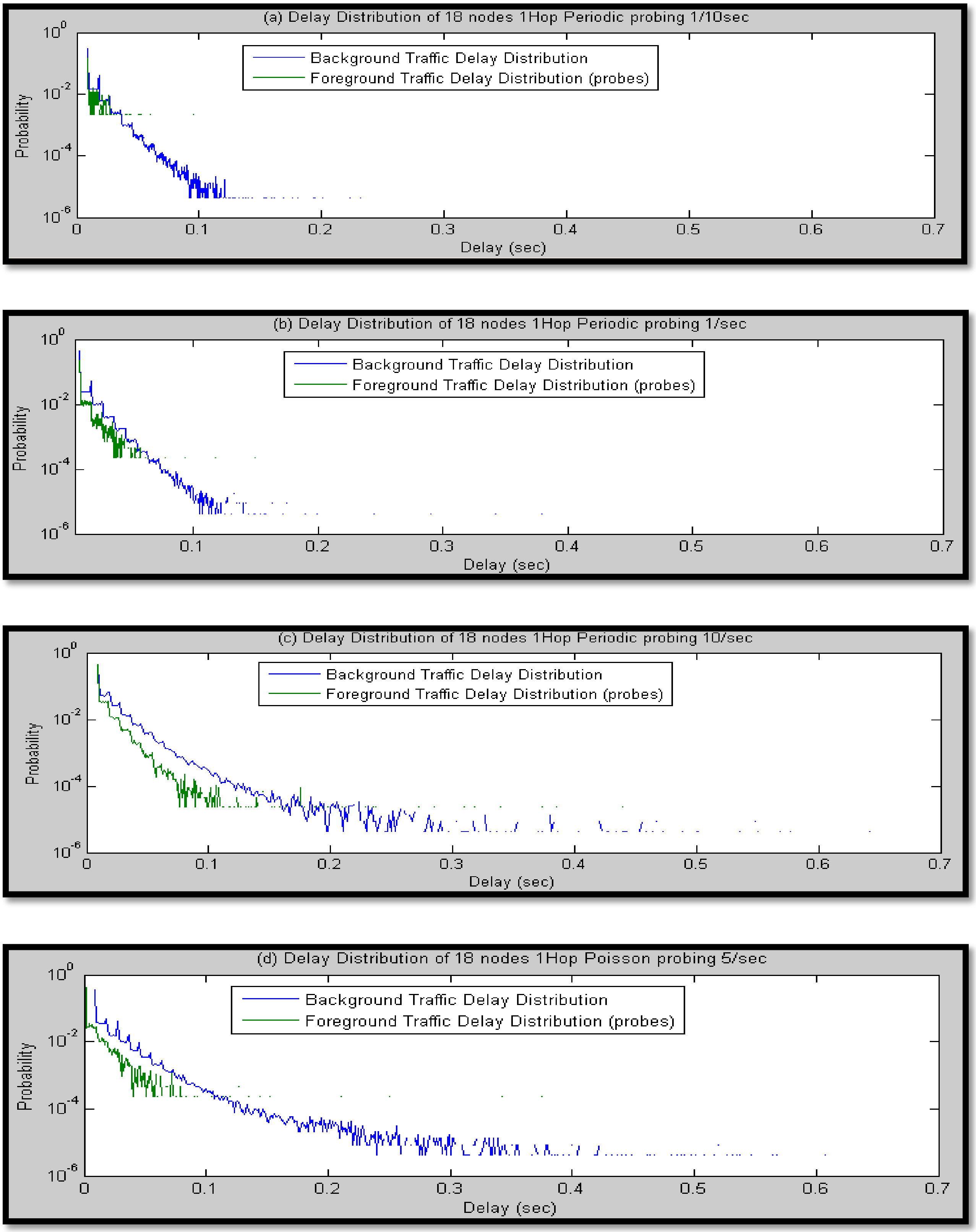

Figure 5-17 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 18 nodes scenario at different probing rates

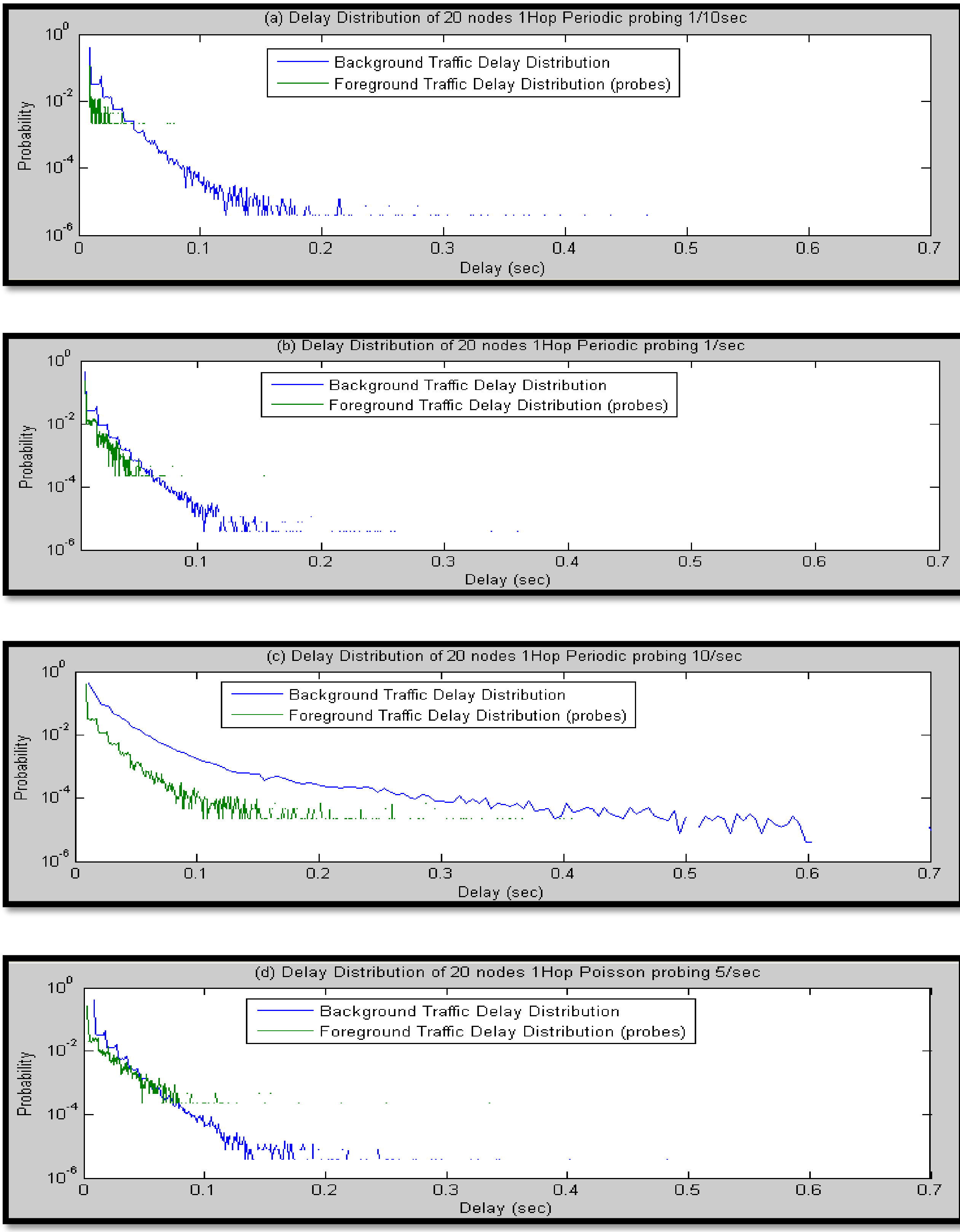

Figure 5-18 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 20 nodes scenario at different probing rates

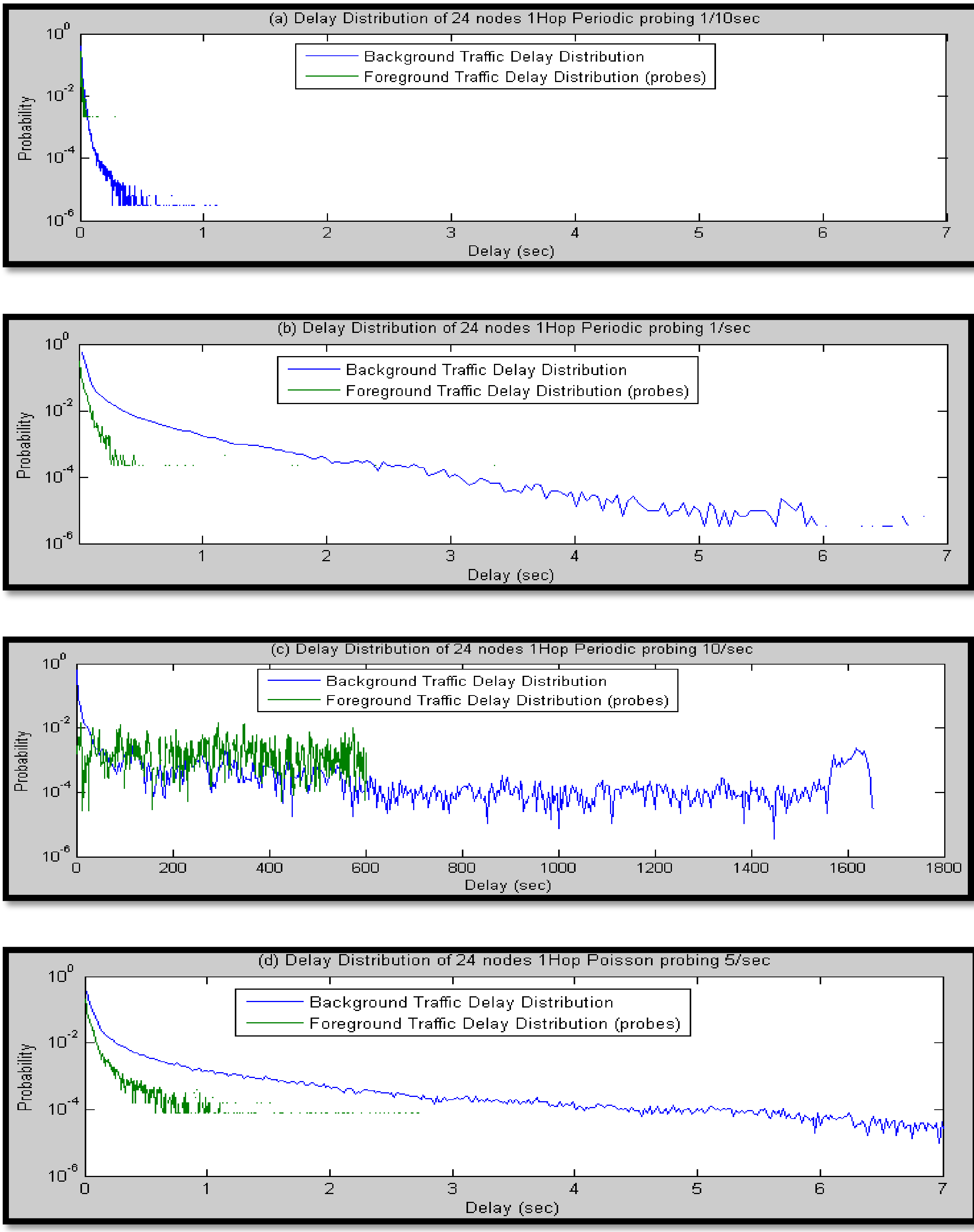

Figure 5-19 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in single hop, 24 nodes scenario at different probing rates

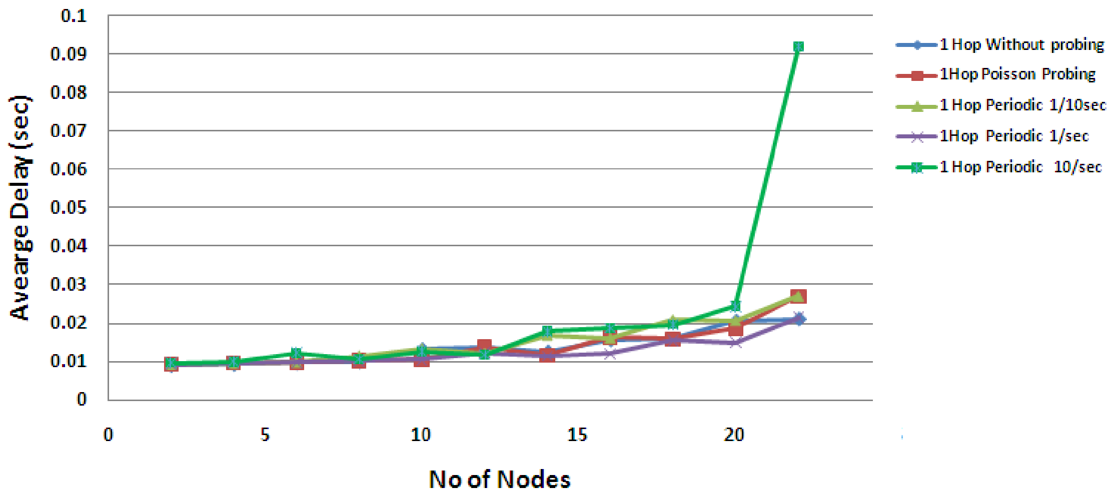

Figure 5-20 Single Hop Delay Comparison for various probing rates

Figure 5-20 shows average delay results for each probing pattern. It can be seen by the graph that a probe rate of one packet per second represents a good compromise for this particular scenario as it gives a good accuracy across the load range. However, periodic probing with 10 probes per second shows a larger amount of probe overhead is resulting in huge amount of extra delay and making active probing almost useless.

Now the question arises how accurate is the measurement? As we know that only if it is sufficiently accurate, and doesn't require excessive bandwidth overhead, can active measurement be used with confidence in measuring packet delay distribution over wireless access. The simulation study reveals that active probing gives a very good degree of approximation for delay distribution of a network if it's not heavily loaded, and requires minimal bandwidth overhead. In our particular network scenario for single hop wireless network, as the number of nodes exceeds 20, accuracy in measurement decreases due to unavoidable probing overhead.

5.5.3 Sampling error in delay distribution measurement

Sampling theory provides methods for calculating the probability that any delays differ from reality by more than a certain amount. In order to determine how much our average probe delay deviates from the mean value, define

N = the number of the probe packets

$t_{N-1,1-\frac{\alpha}{2}}$ = the Student t-distribution value for N-1 degrees of freedom (i.e. sample size=N), and

(100-α) % confidence interval for the estimate of the mean delay (L)

δ = absolute error in the measurements

L_N = mean delay, estimated from N measurements

S_N = standard deviation of the measurements

S^2_N = variance of the measurements

We have got mean delay values from simulation results, we know the numbers of probes generated, by calculation the standard deviation values from mean delay we can evaluate the absolute error using Equation 5-1

$$L_N \pm t_{N-1,1-\alpha/2} \cdot \frac{S_N}{\sqrt{N}}$$

Equation 5-1 [14]

Figure 5-21 shows the absolute error obtained by using different probing patterns using confidence coefficient =0.95 where we observed the average delay of the packets through the network by sending probe packets through the network and measuring their delay. Each probe packet samples the delay distribution, and so we take their sample mean, and use this as a measure of the average network delay. The values of mean delay and variance used to compute absolute error are given in Table 5-3.

Nodes	1Hop Average Delay (sec)	1Hop delay distribution variance of traffic	1Hop Poisson Probing variance	Sampling Error Poisson Probing	1 Hop Periodic 1/10sec variance	Sampling Error 1/10 sec	1Hop Periodic 1/sec variance	Sampling Error 1/sec	1 Hop Periodic 10/sec variance	Sampling Error 10/sec
2	0.0091	7.55E-07	1.6386E-06	2.50895E-05	1.9291E-06	1.30E-04	1.41E-06	3.51E-05	1.72E-06	1.22398E-05
4	0.0094	2.54E-06	3.2901E-06	3.55517E-05	0.000005613	2.21E-04	3.59E-06	5.60E-05	3.65E-06	1.78476E-05
6	0.0097	4.76E-06	5.4837E-06	4.58979E-05	7.36E-06	2.53E-04	6.00E-06	7.24E-05	1.00E-05	2.95418E-05
8	0.01	7.65E-06	8.9955E-06	5.87853E-05	1.59E-05	3.73E-04	8.90E-06	8.82E-05	9.05E-06	2.81041E-05
10	0.0132	4.40E-05	1.3507E-06	2.2779E-05	1.70E-05	3.86E-04	1.25E-05	1.05E-04	4.24E-05	6.08239E-05
12	0.0135	5.03E-05	0.000031642	0.000110252	2.43E-05	4.61E-04	2.28E-05	1.41E-04	1.83E-05	3.99785E-05
14	0.0126	3.76E-04	0.000023465	9.49437E-05	4.18E-05	6.04E-04	2.73E-05	1.54E-04	7.48E-05	8.0815E-05
16	0.0157	9.94E-05	0.000061897	0.000154202	8.68E-05	8.70E-04	3.62E-05	1.78E-04	7.15E-05	7.90002E-05
18	0.0159	9.26E-05	0.00008334	0.00017893	7.31E-05	7.99E-04	6.99E-05	2.47E-04	1.39E-04	0.000109989
20	0.02046	0.0027	0.00062505	0.00049002	8.60E-05	8.67E-04	7.36E-05	2.54E-04	0.063	0.002345309
22	0.021	0.0071	0.0459	0.00419916	0.0014	3.50E-03	1.96E-04	4.13E-04	0.0326	0.001687091
24	0.1368	0.1115	0.0286	0.003314661	0.0286	1.58E-02	9.10E-03	2.82E-03	2.85E+04	1.57729913
26	0.2261	0.4166	2.22E+02	0.292069333	0.0044	6.20E-03	0.0013	1.07E-03	2.47E+04	1.46791974
28	165.8351	1.98E+05	3.43E+05	11.47847356	0.113	3.14E-02	0.2155	1.37E-02	4.72E+05	6.416364136

Table 5-3 Variance and Sampling Error for single hop wireless scenario

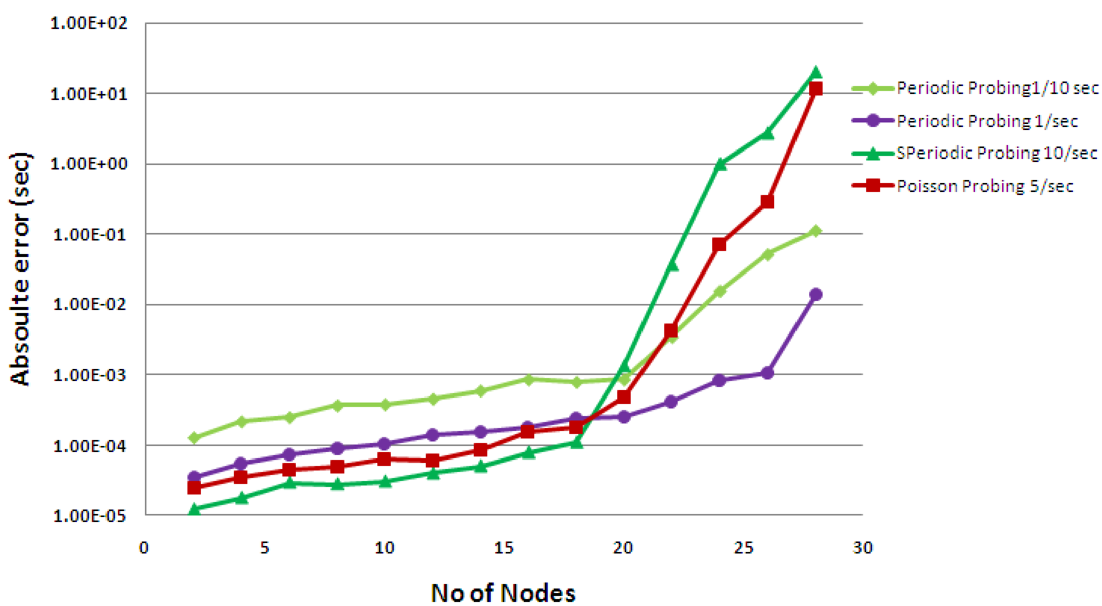

Figure 5-21 Sampling error for different probing patterns and rates for single hop scenario

5.6 Delay Distribution for Two Hops Scenario

5.6.1 Simulation Topology

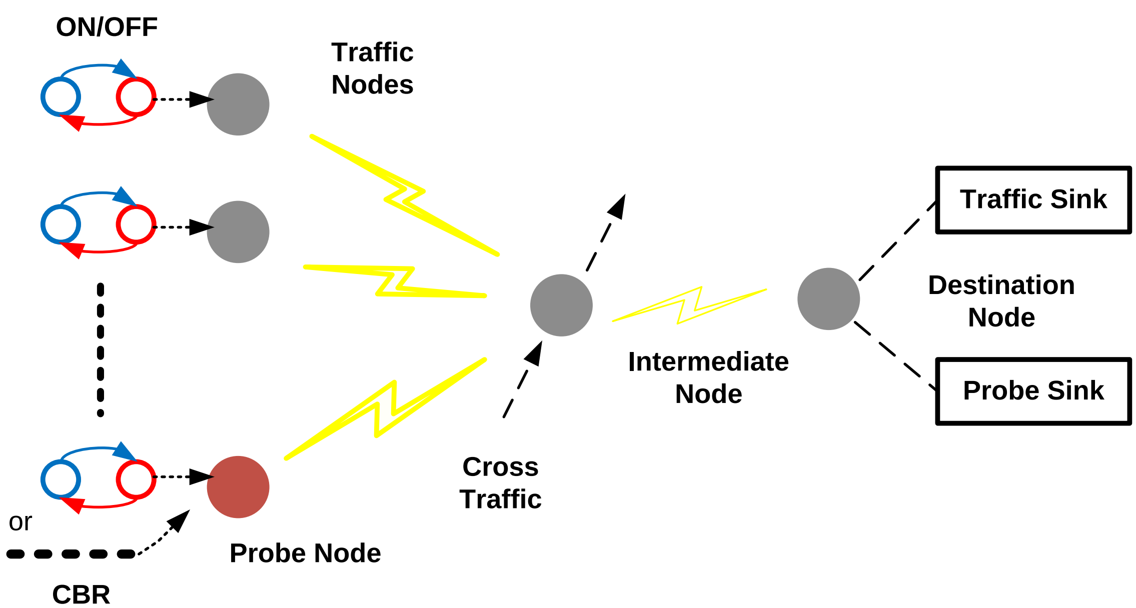

Figure 5-22 Simulation Topology for Two Hops Scenario

The experimental setup for the two hops wireless network is the same as the single hop, just with the addition of an intermediate node between source nodes and destination node. The simulation parameters used for traffic are the same as given in Table 5-1, and the simulator parameters are the same as listed in Table 5-2. The number of traffic sending nodes will be increased in order to increase the load on the network and create contention for channel access. The middle node acts as a forwarding node. We have extended the network to examine the effectiveness of active measurement over multi-hop wireless access which includes hidden node and exposed node issues.

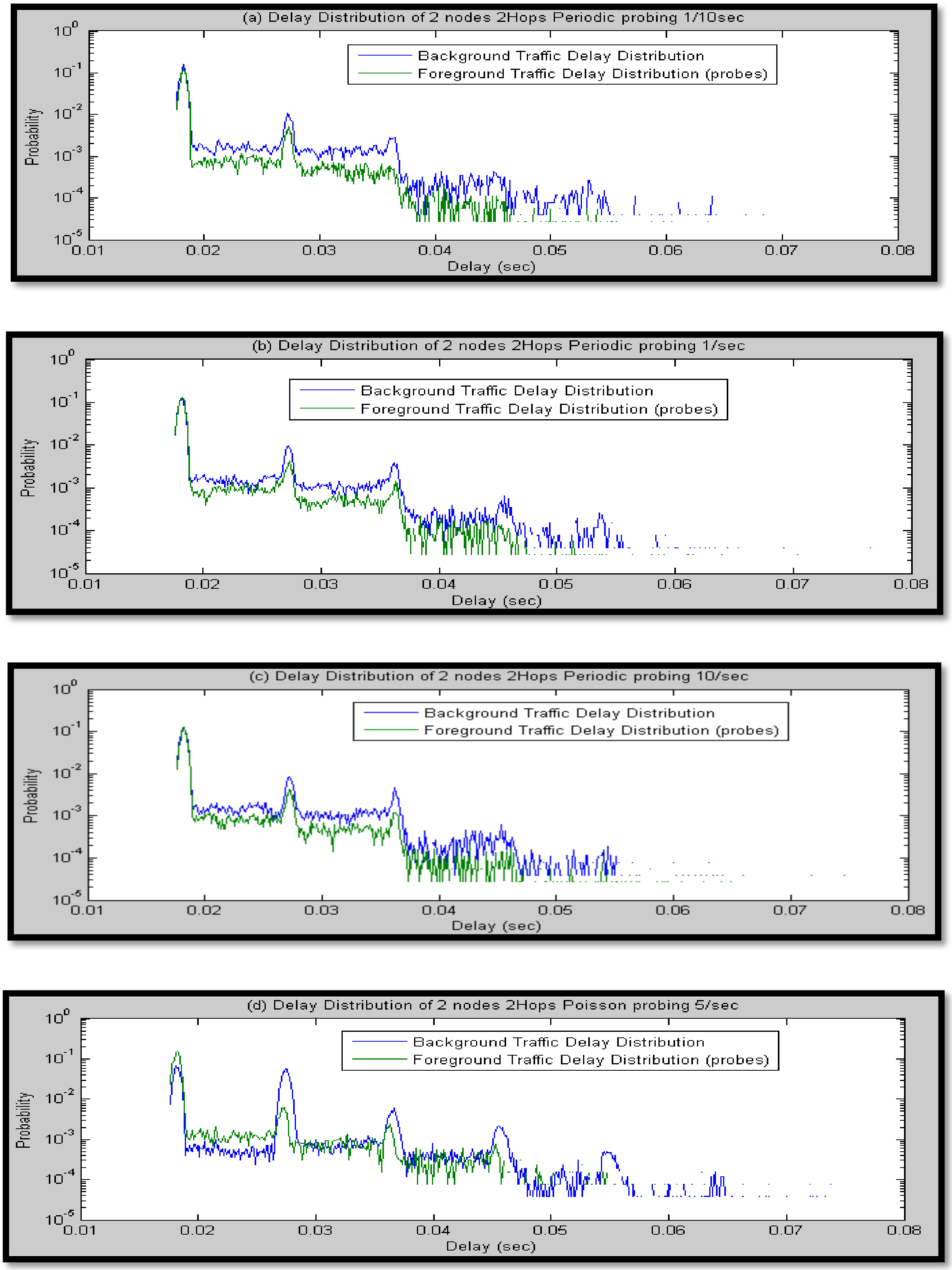

Figure 5-23 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a two hops, 2 nodes scenario at different probing rates

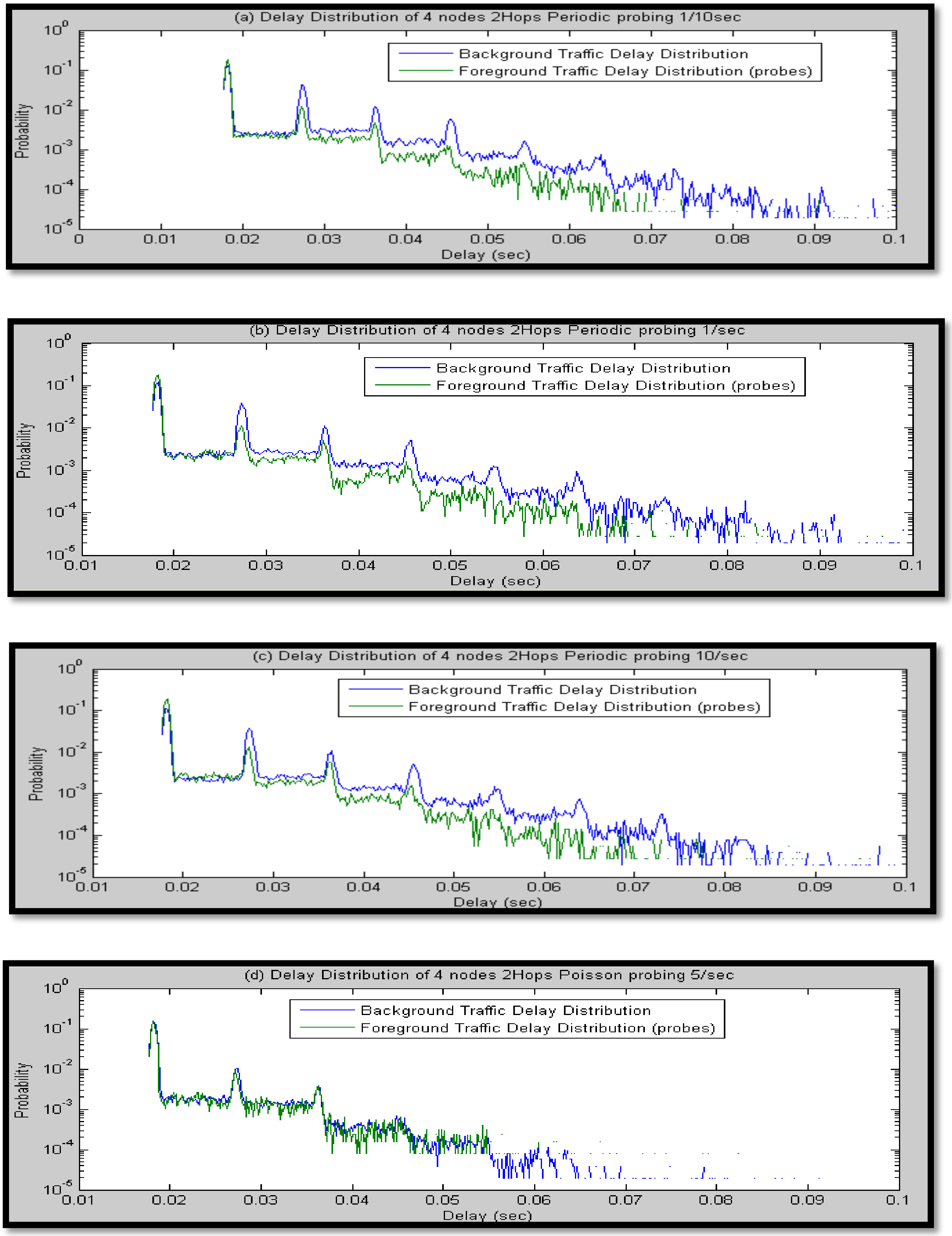

Figure 5-24 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a two hops, 4 nodes scenario at different probing rates

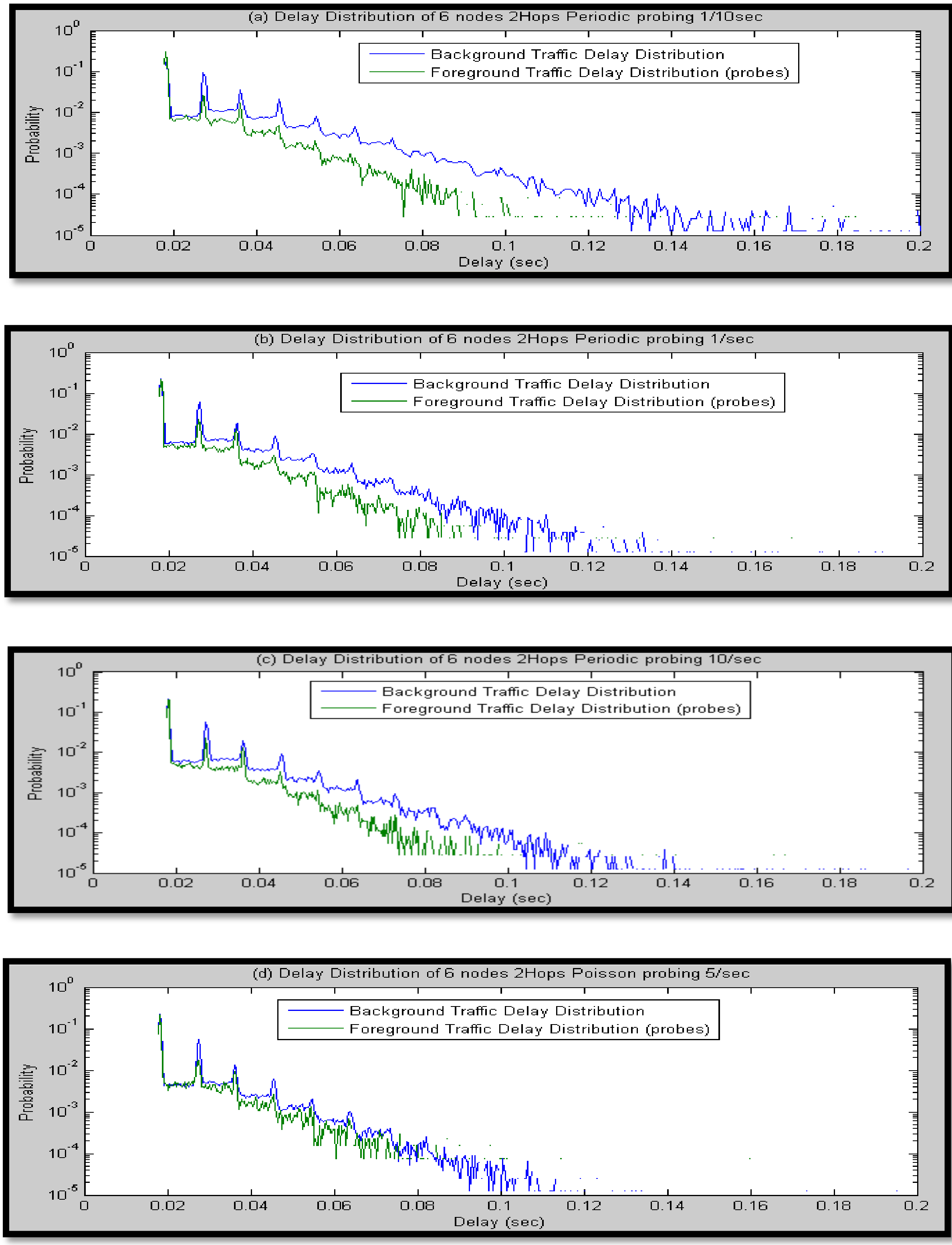

Figure 5-25 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a two hops, 6 nodes scenario at different probing rates

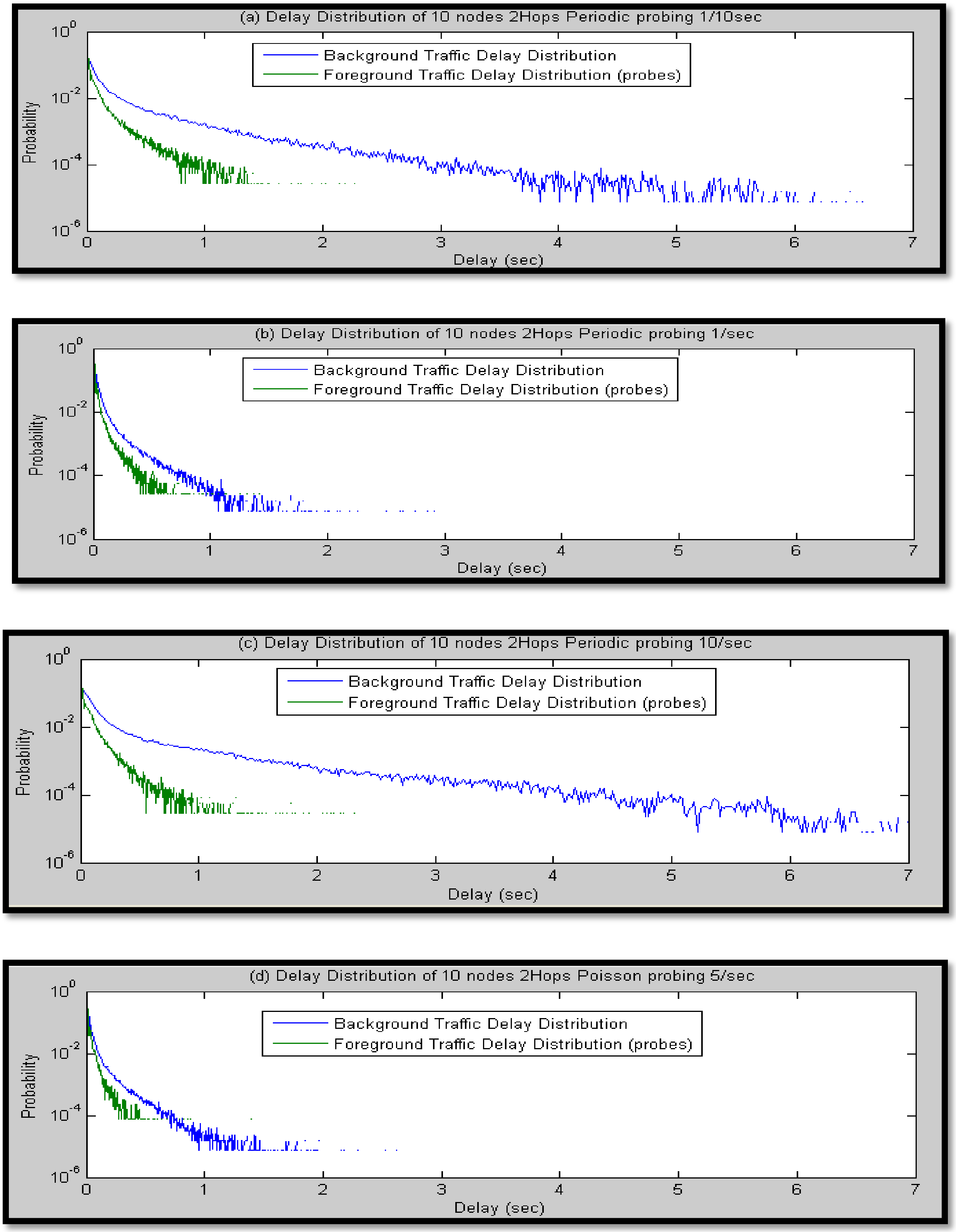

Figure 5-26 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a two hops, 10 nodes scenario at different probing rates

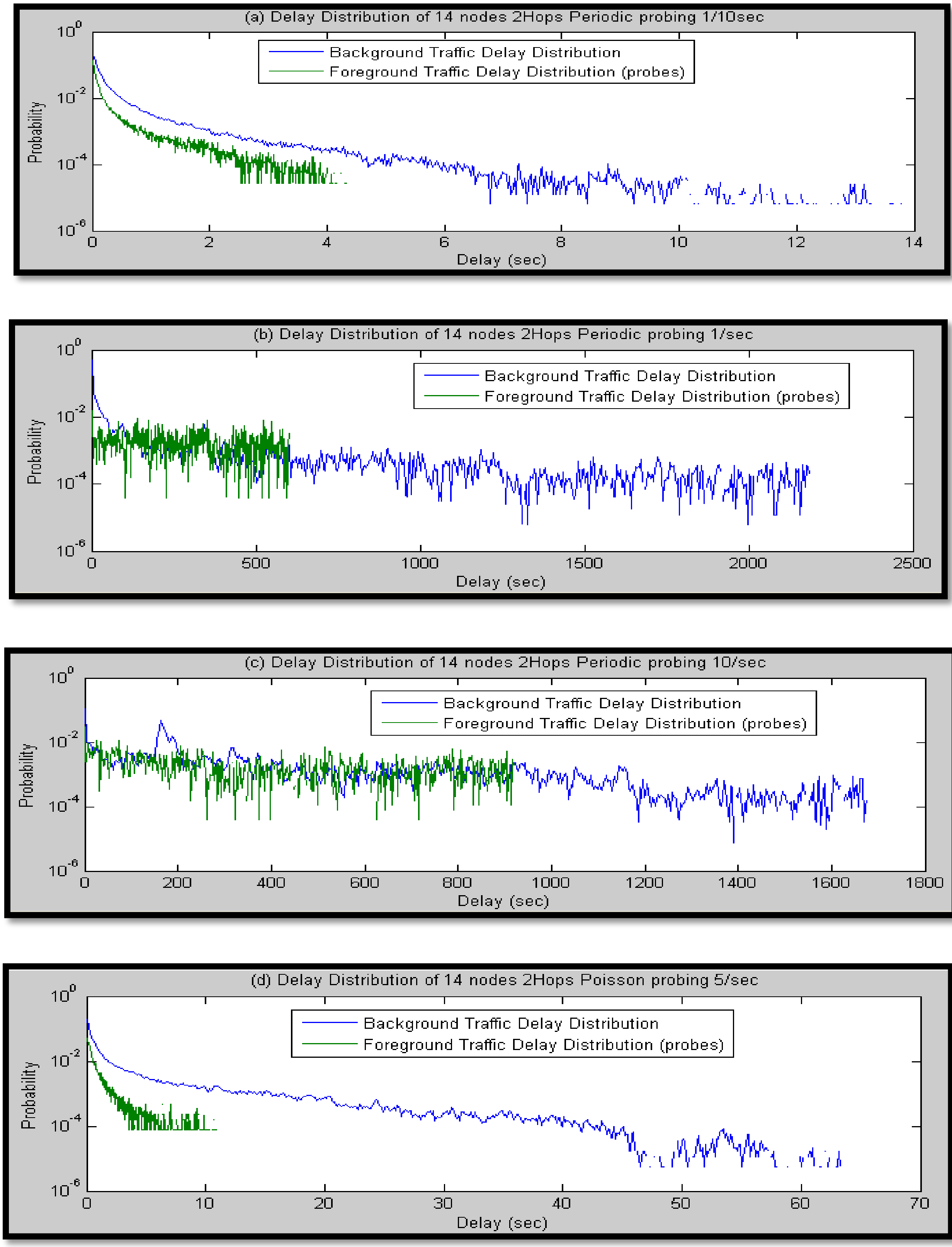

Figure 5-27 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a two hops, 14 nodes scenario at different probing rates

5.6.2 Simulation Results

The simulation results for two hops are very much in accordance with the single hop scenario, i.e. initially we can see the same step-like decay response for probability of delay. It is interesting to note that the delay results we got for 8 nodes in a single hop network are the same as 8 nodes in two hops. This clearly indicates that the bandwidth available for transmission for 8 nodes in a single hop case is nearly equal to the bandwidth available for 2 nodes traffic transmission in a two hops network scenario. Therefore, as indicated from results, probing multi-hop networks becomes even more challenging.

In two hops, after 14 nodes, the network saturates and the delays jump to a very large value. Figure 5-27 indicates that when we probe 14 nodes two hop wireless network with a probing frequency of 1 probe every 10 second the results are closer to the delay values of user traffic but doesn’t closely follow the delay probabilities of user traffic. Therefore when we increase the probing rate to 1 per second the network almost saturates, we are not able to see the packet scale delay as it is shifted much higher. Probing at 10 probes per second appears to be totally unacceptable as it leads to huge delays and network saturation. When we employed random probing with 5 per second though the network doesn’t saturate, the tail of the delay distribution remained unresolved.

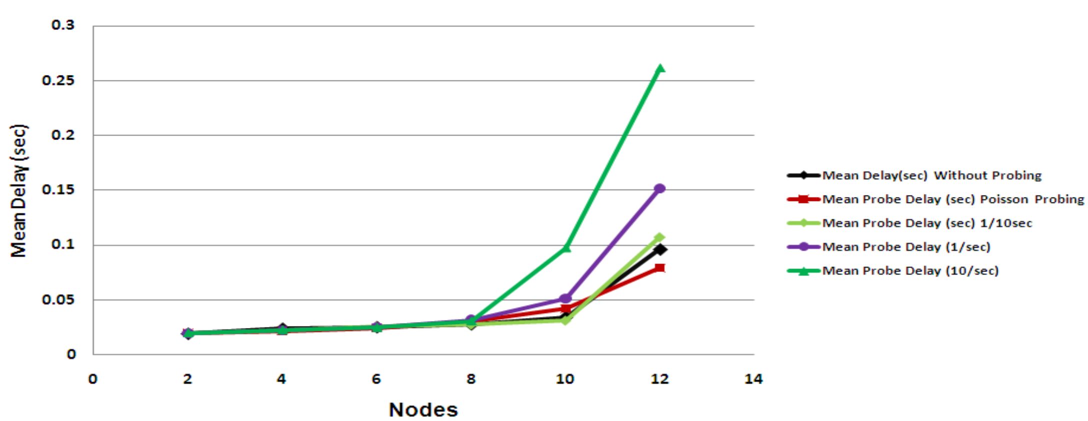

Figure 5-28 Two Hops Delay Comparison for various probing rates

Nodes	2Hops Average Delay (sec)	2Hops delay distribution variance of traffic	2Hops Poisson Probing variance	2 Hop Sampling Error Poisson Probing	2 Hop Periodic 1/10sec variance	2 Hops Sampling Error 1/10 sec	2 Hops Periodic 1/sec variance	2 Hops Sampling Error 1/sec	2 Hops Periodic 10/sec variance	2 Hops Sampling Error 10/sec
2	0.0192	1.08E-05	2.02E-05	2.20E-06	1.67E-05	1.82E-05	1.81E-05	1.90E-06	1.87E-05	1.9283E-07
4	0.0238	6.96E-05	4.97E-05	3.45E-06	6.45E-05	3.58E-05	6.57E-05	3.61E-06	6.58E-05	3.6138E-07
6	0.0252	1.12E-04	1.05E-04	5.02E-06	1.41E-04	5.28E-05	1.12E-04	4.71E-06	1.10E-04	4.6807E-07
8	0.0279	2.14E-04	3.96E-04	9.75E-06	2.56E-04	7.13E-05	5.52E-04	1.05E-05	4.58E-04	9.537E-07
10	0.0343	7.28E-04	0.0022	2.30E-05	0.022	6.61E-04	0.0037	2.71E-05	0.0147	5.4008E-06
12	0.0961	2.28E-02	0.0121	5.39E-05	0.2969	2.43E-03	2.9756	7.68E-04	0.2742	2.3326E-05
14	73.3609	1.39E+04	0.1823	2.09E-04	2.92E+04	7.61E-01	3.03E+04	7.75E-02	8.38E+04	0.01289827
16	91.0486	4.80E+04	8.4635	1.43E-03	6.31E+03	3.54E-01	3.74E+04	8.61E-02	8.55E+03	0.00411933
18	322.6987	1.83E+05	5.15E+04	1.11E-01	6.10E+04	1.10E+00	4.63E+04	9.59E-02	6.10E+04	0.01100543

Table 5-4 Variance and Sampling Error for two hop wireless scenario

Table 5-4 refers to the variance values and probing error values obtained for 2 hops wireless scenario. Figure 5-28 shows the amount of delay experienced by different probing patterns in comparison with user network delay without probing.

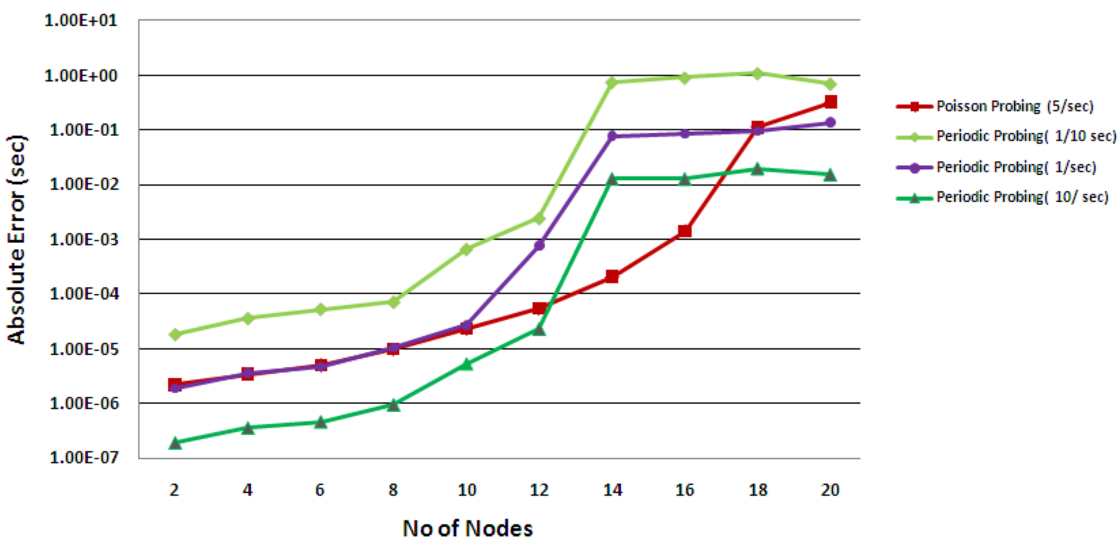

Figure 5-29 Sampling error for different probing patterns for two hops scenario

Figure 5-29 shows the absolute error with Poisson probing (5 probes per second) and Periodic probing (10 probes per second). It initially shows that those work well, but the accuracy in measurement degrades as the variability in the measured delay distribution increases. Extra load caused by adding probe packets becomes suddenly extreme. This is seen at high loads an i.e. value of load for which performance monitoring is most essential.

5.7 Delay Distribution for Three Hops Scenario

5.7.1 Simulation Topology

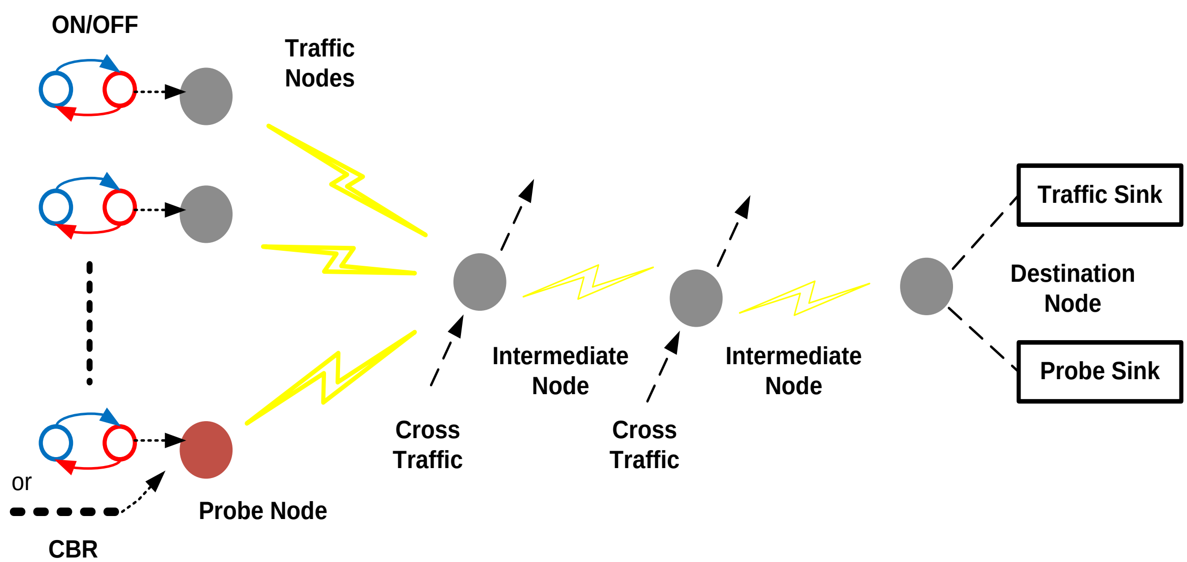

Figure 5-30 Simulation Topology for Three Hops Scenario

In this section, the investigation is further extended to three hops. We introduce two intermediate nodes between source nodes and destination. We also called these nodes 'Forwarding nodes', as their main purpose is to rely or forward the user packets to the destination node. The parameters are the same as listed in Table 5-1 and Table 5-2. Cross traffic has been introduced (within the hearing range of competing nodes i.e. 250 meters) to make the network scenario more realistic. Here cross traffic is also contending for radio medium access so the interference level can be much higher as compared to previous cases. The parameters for cross traffic are same as background user traffic.

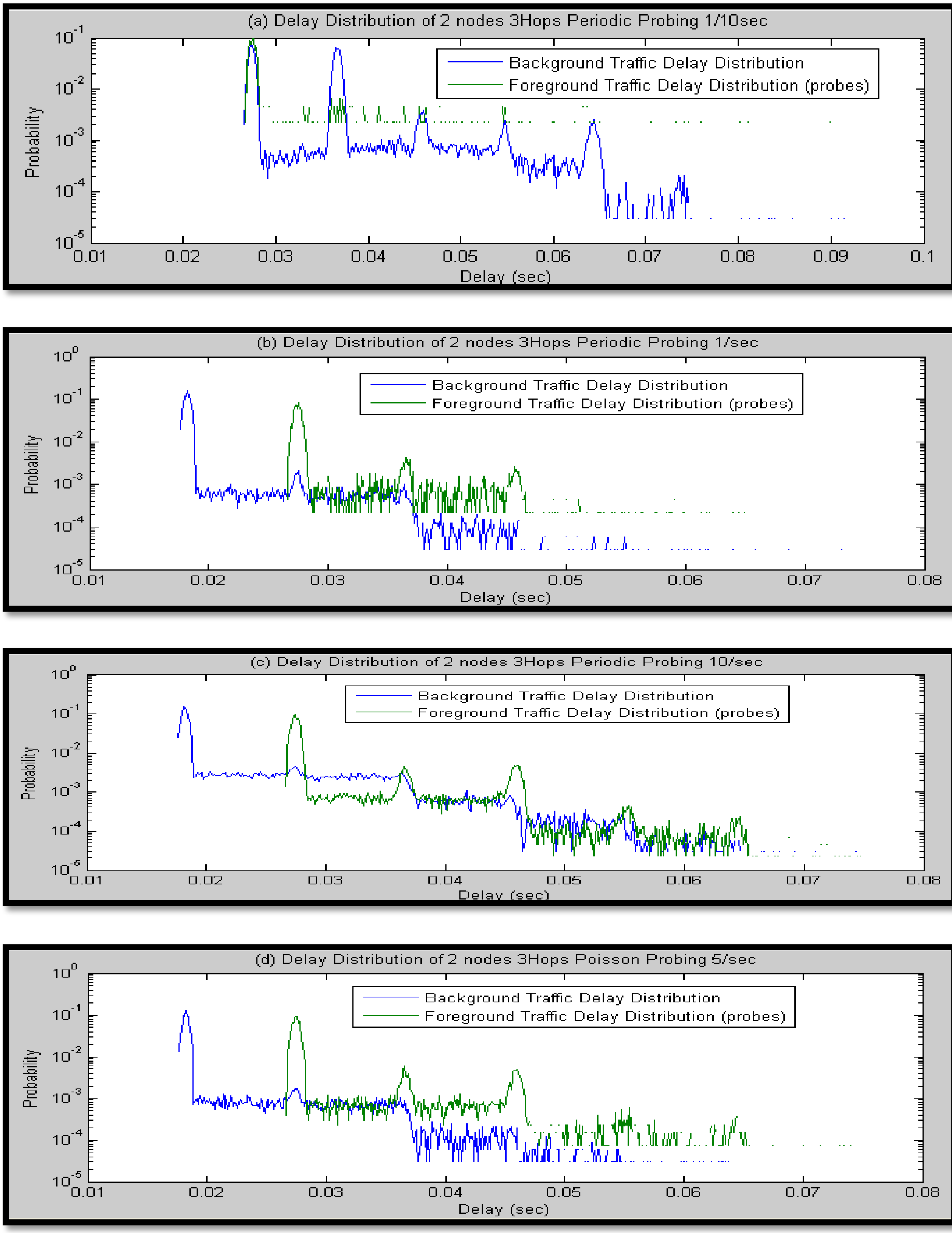

Figure 5-31 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a three hops, 2 nodes scenario at different probing rates

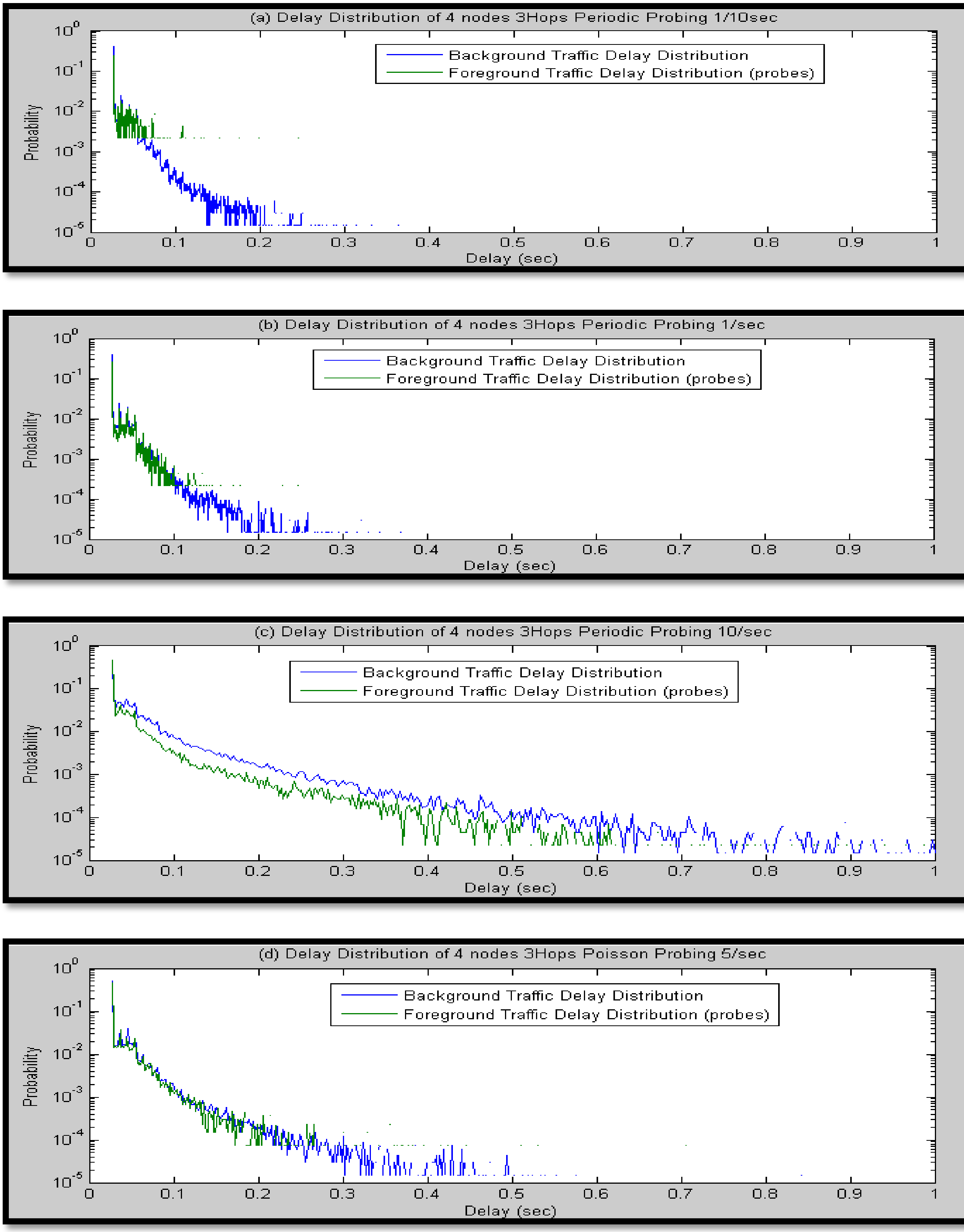

Figure 5-32 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a three hops, 4 nodes scenario at different probing rates

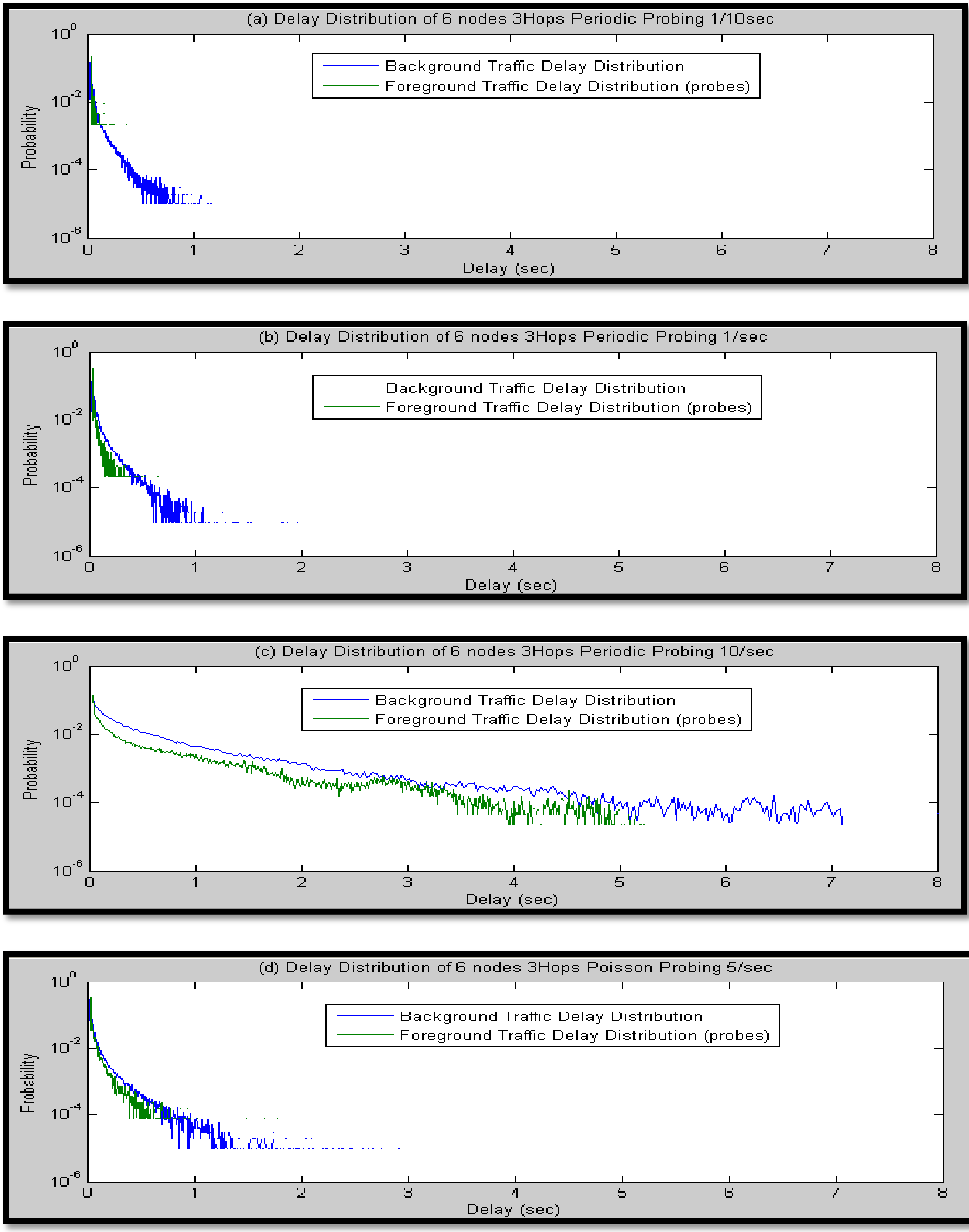

Figure 5-33 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a three hops, 6 nodes scenario at different probing rates

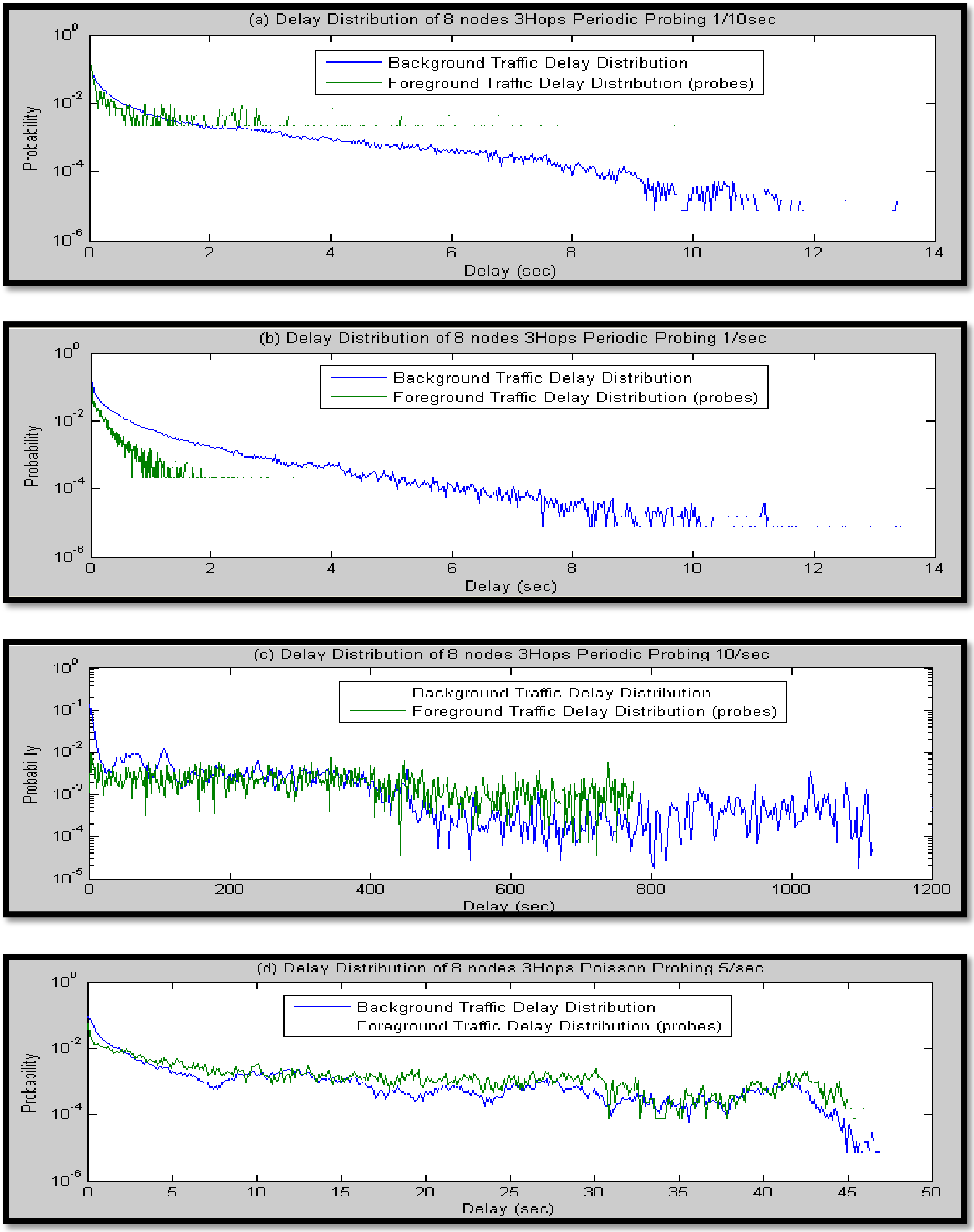

Figure 5-34 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a three hops, 8 nodes scenario at different probing rates

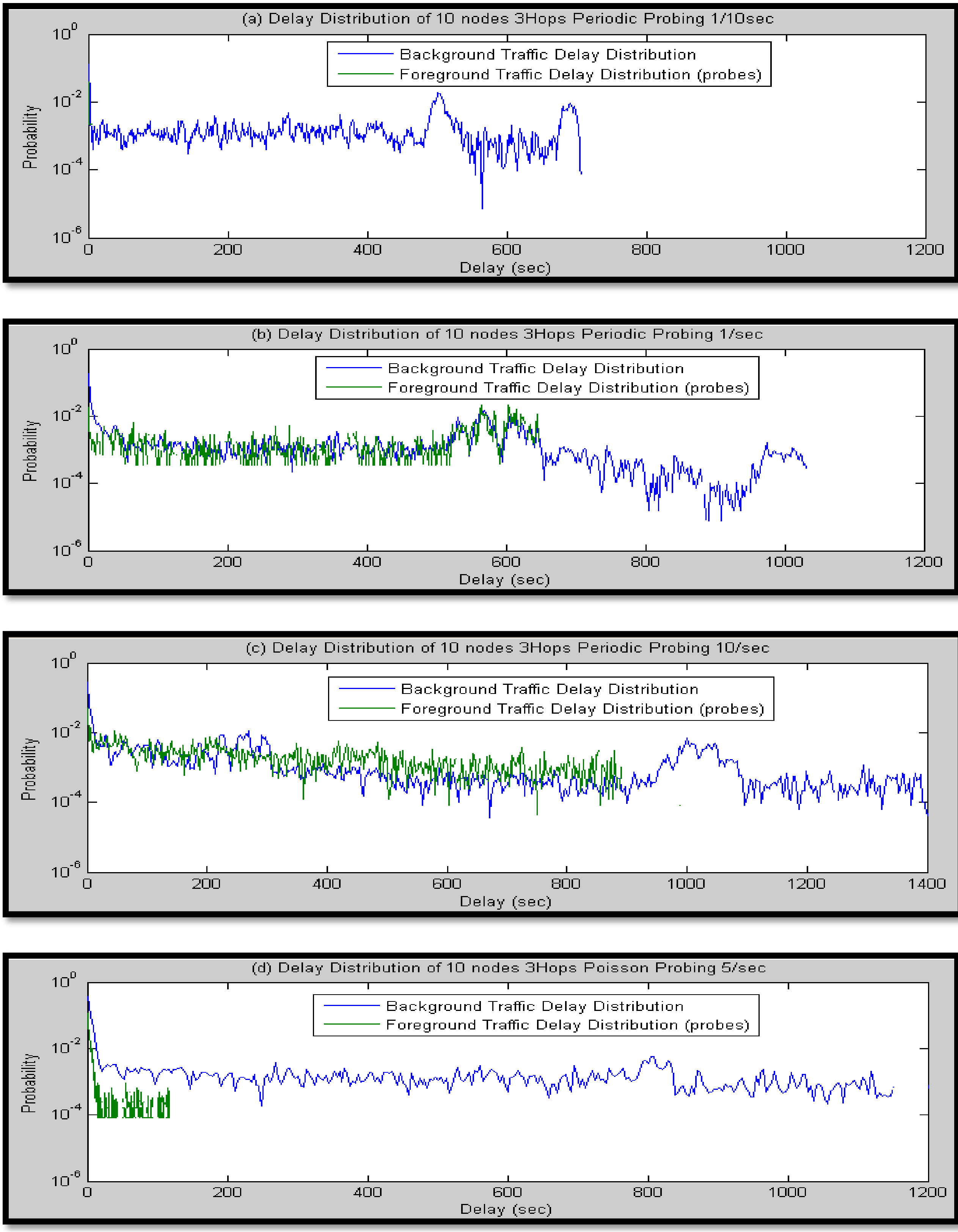

Figure 5-35 Delay probability plots for probes (foreground traffic) and user traffic (background traffic) in a three hops, 10 nodes scenario at different probing rates

5.7.2 Simulation Results

Figure 5-31 shows delay distribution for a two nodes case. (a) Indicates that the numbers of probes are insufficient to capture delay distribution, therefore in (b) and (c) we increased the probing rate to 1/sec and 10/sec. (c) shows that probes map the variability in delay distribution quite nicely. (d) Used Poisson probing the results for the probe delay was quite similar to that of user traffic (see Figure 5-36).

We note that in Figure 5-32 (d) and Figure 5-33 (d) there are sufficient probes to resolve the delay distribution, but it lead to an increase in delay. The network goes into saturation only with 10 nodes in the network. This can be seen from Figure 5-35 and probing in this case is useless. Table 5-5 refers to the variance values and probing error values obtained for 3 hops wireless scenario.

Figure 5-36 gives the average delay comparison for all probing patterns in 3 hop wireless network, and Figure 5-37 illustrates the absolute error for all the probing patterns and with different probing rates used for investigation for applicability of active probing over 3 hops wireless network scenario.

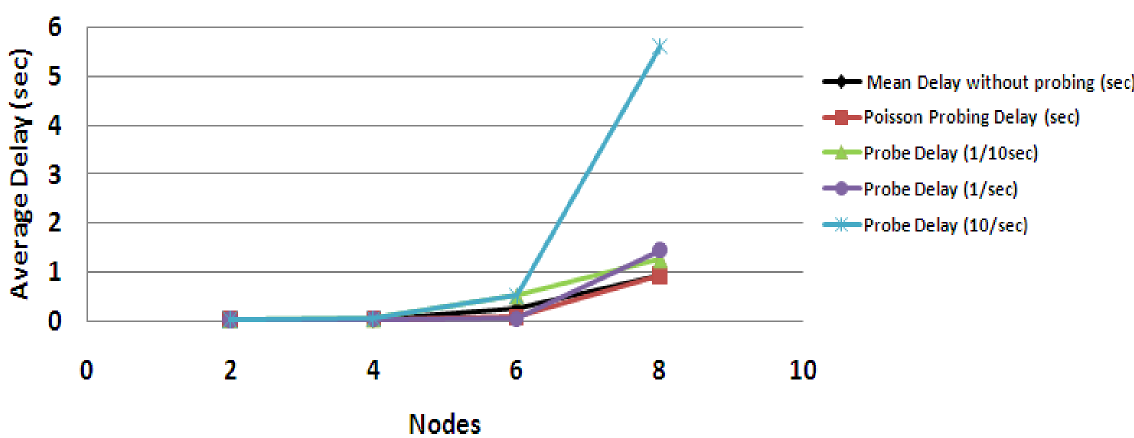

Figure 5-36 Three Hops Delay Comparison

Nodes	3Hops Average Delay (sec)	3Hops delay distribution variance of traffic	3Hops Poisson Probing variance	3 Hop Sampling Error Poisson Probing	3 Hop Periodic 1/10sec variance	3 Hops Sampling Error 1/10 sec	3 Hops Periodic 1/sec variance	3 Hops Sampling Error 1/sec	3 Hops Periodic 10/sec variance	3 Hops Sampling Error 10/sec
2	0.0245	4.19E-05	4.79E-05	3.39234E-06	1.13E-04	4.08114E-05	2.87E-05	2.05785E-06	4.19E-05	2.48552E-07
4	0.0333	1.96E-04	3.97E-04	9.76503E-06	5.50E-04	9.01084E-05	2.86E-04	6.49159E-06	0.0036	2.30455E-06
6	0.0376	2.73E-04	0.0051	3.4993E-05	0.0016	0.000153636	0.0023	1.84204E-05	0.5597	2.8735E-05
8	0.1045	0.0189	1.4705	0.000594194	2.8487	0.006482722	0.1005	0.000121763	4.39E+04	0.00804824
10	0.3159	0.2934	3.6137	0.000931477	0.1238	0.001351432	4.65E+03	0.082849753	6.16E+04	0.009533273
12	261.5264	7.11E+04	6.55E+04	0.125378735	0.1436	0.001455497	0.0963	0.000119192	7.85E+04	0.01076339

Table 5-5 Variance and Sampling Error for two hop wireless scenario

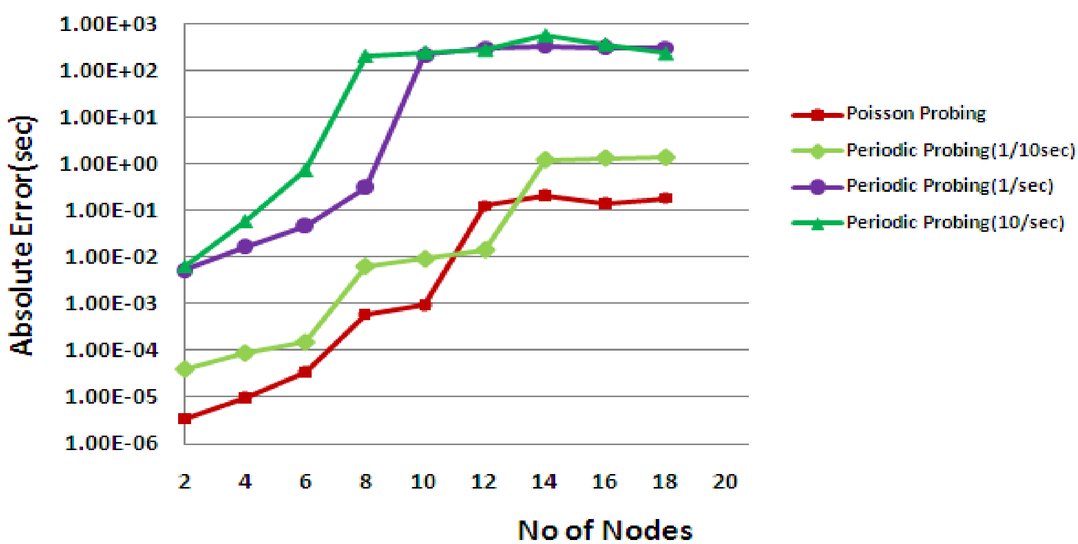

Figure 5-37 Sampling error Comparison

5.7.3 Discussion

In this chapter we have quantitively explored how active measurement works when it is used to evaluate the packet delay distribution over small scale wireless ad-hoc networks. We have observed that, as the load increases, active measurement systematically underestimates the mean delay of the user traffic, no matter whether Poisson probing or Periodic probing is employed. The absolute error is not significantly different for either probing technique. The possible reasons might be that Poisson and Periodic sampling are evenly distributed with time, however user traffic is bursty.

We have observed that, with a greater number of nodes in the network, the variance in the delay distribution obtained is larger as well. In order to measure this variability in delay distribution, a large number of samples are needed which leads to severe interference with traffic packets, skewing the delay results and increasing the inaccuracy to a greater extent. This has been clearly observed in the simulation results. Therefore, there is a critical trade-off between accuracy of measurement and packet probing overhead in wireless ad-hoc networks, with shared medium and limited bandwidth. Figure 5-38 shows how the packet delivery ratio falls dramatically with increase in hop count in a wireless ad-hoc network. When we perform active measurements for multiple hops, we observed that as the chain length increases, i.e. number of hops, available capacity for transmission decreases, this leads to low throughput (as shown in Figure 5-39) and the network becomes overloaded much earlier.

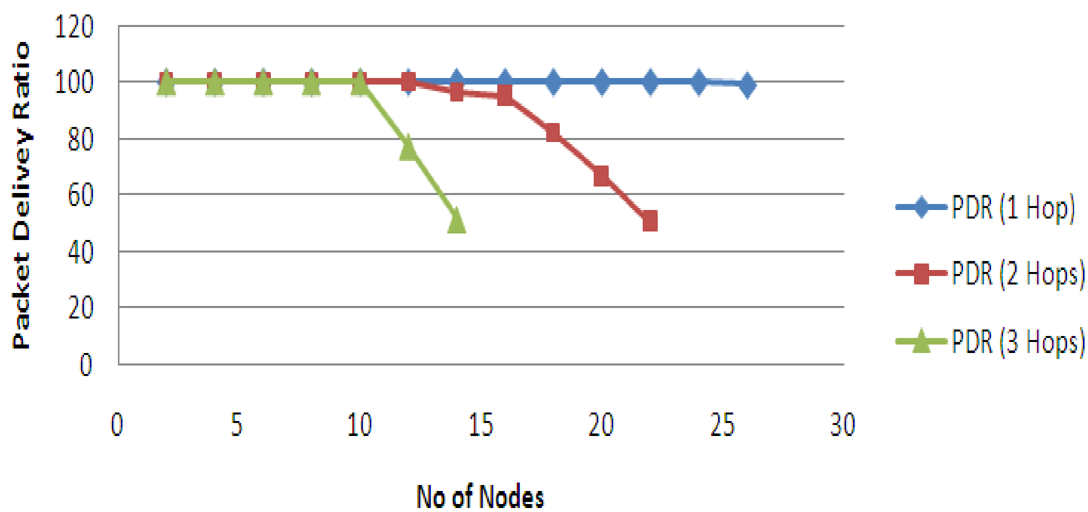

Figure 5-38 Packet Delivery Ratio Comparison

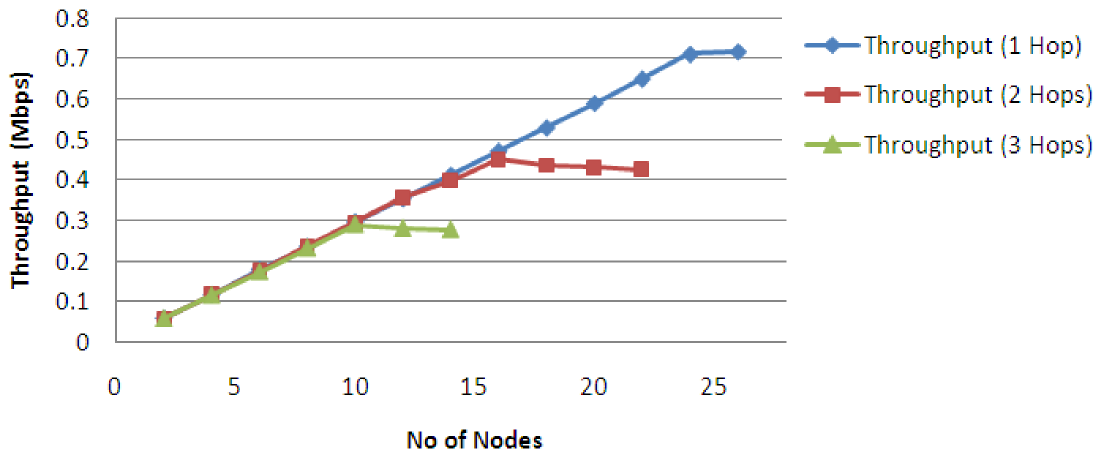

Figure 5-39 Throughput Comparison

5.8 Summary

In this chapter we discussed the active probing accuracy and different types of active probing techniques. Our results suggested that active probing is highly affected by the probe size, pattern and traffic load. We have compared two probing techniques i.e. Periodic and Poisson probing, and estimated the absolute error by using sampling theory. The simulations have been performed for single hop and multi-hop wireless networks to perform end-to-end latency measurement via probing.

6. Packet Loss Probing over Wireless Access

In this chapter we have measured the packet loss rate of probes in comparison with user traffic. We have three simulated three different scenarios i.e. Single hop, two hops and three hops wireless networks.

6.1 *Single Hop Scenario*

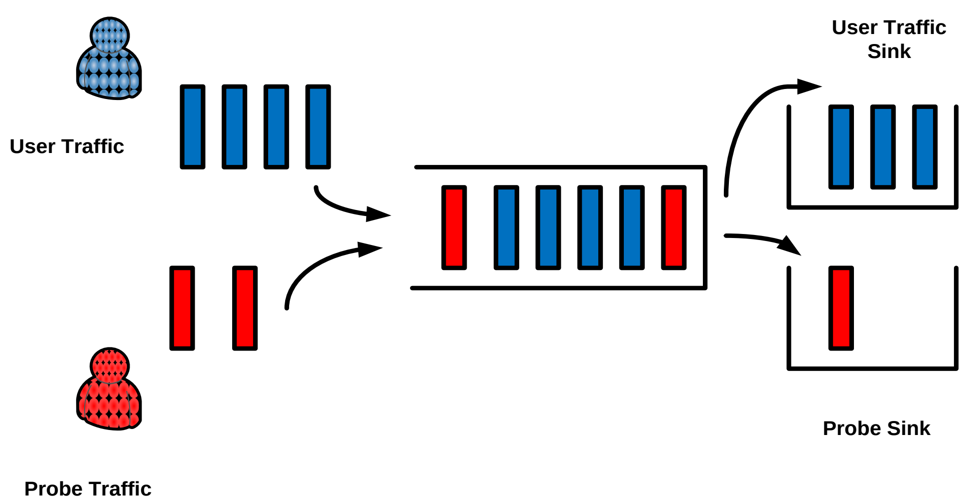

Figure 6-1 Schematic of Single Hop Loss Probing Scenario

The network used for the first set of experiments contains one destination node and the rest of the nodes are source node. The topology is the same as described in chapter 5 section 5.5.1. Two Sinks are attached to the destination node in order to collect the probe traffic and user traffic. Probes are generated at three different rates: 1/sec, 5/sec and 10/sec for Periodic probing. The rest of the parameters used for simulation are listed in Table 6-1, and the simulator parameters used are given in Table 6-2. Initially the simulation was run without probe traffic and the numbers of packets sent and received at the destination node were collected by step wise increasing the number of the nodes in the network until the network saturates.

Saturation happened with 28 wireless nodes in the network; results from chapter 5 also suggested the same for single hop simulation. The Packet loss ratio was calculated as:

Packet loss ratio = number of packets lost/ number of packets received

Traffic type	Exponential ON/OFF
Application Type	UDP
Probe Traffic Type	CBR
Packet Size (data traffic)	1000 bytes
Packet Size (probes traffic)	1000 bytes
Mean ON time of data traffic (Ton)	0.35 sec
Mean OFF time of data traffic (Toff)	0.67 sec
Input data traffic rate (R)	85.6kbps
Probing rates	1/sec, 5/sec, 10/sec

Table 6-1 Traffic Parameters

Simulator	**NS version 2.31**	
	Network Size	1000x500 m
	No of Nodes	Min=14, Max=28
	Simulation Duration	3600 sec
Physical Layer	**Signal Propagation Model**	Two Ray Ground

	Max Transmission Range	250m
	Antenna Type	Omi directional
Mac Layer	**MAC Protocol**	802.11
	Link Bandwidth (data rate)	1Mbps
	Interface Queue Type	Drop Tail/Pri-Queue
	Interface Queue Size	8000 packets

Table 6-2 Key NS2 Parameters used in Simulation

As stated in previous paragraph, the simulation scenarios are run twice; once with the probes switched off and once with probe switched on. As shown in Figure 6-2, the rates at which the probes are sent are of vital importance. Too few and the results are not accurate and if there are too many the results are disturbed. When probing at the rate of 1 probe per second we observed that packets are lost for user traffic, but there is no loss observed in probe traffic. At higher values of load, although probes show some packet loss they still are not match-able with the user traffic packet loss rate values. While, probing at the rate of 5 probes per sec, the loss ratio of probes and user traffic was closer to each other. Probing at 10 probes per second forced the network to go in to saturation much earlier.

There are two things that need to be considered when assessing how well the probing is performing; firstly how accurate the probed results are be and secondly how much the traffic is being affected by the probes. We have discussed the accuracy of probes in section 6.4.

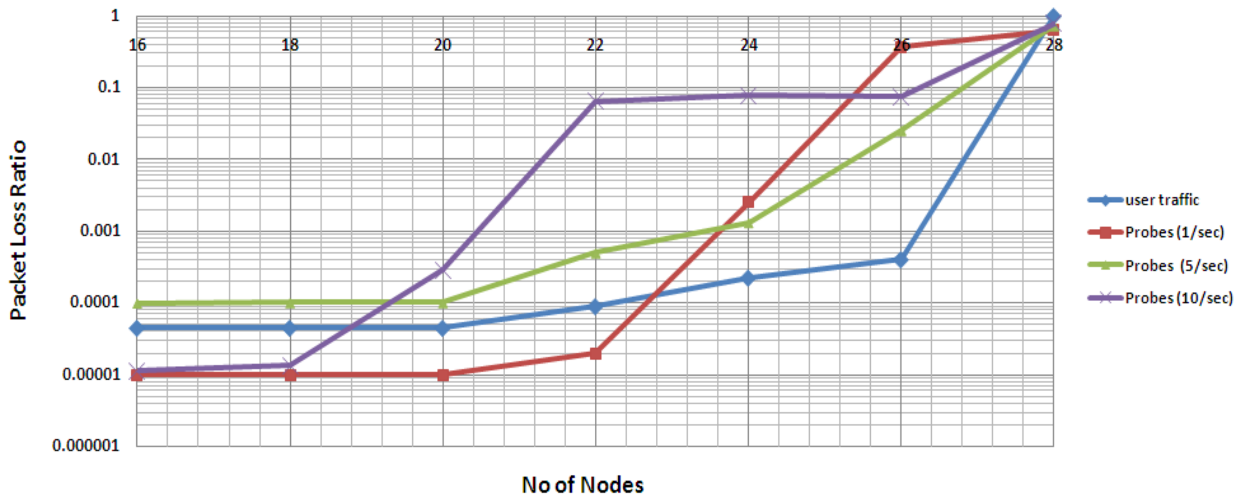

Figure 6-2 Packet loss ratio comparison of probes traffic and user traffic

6.2 *Two Hops Scenario*

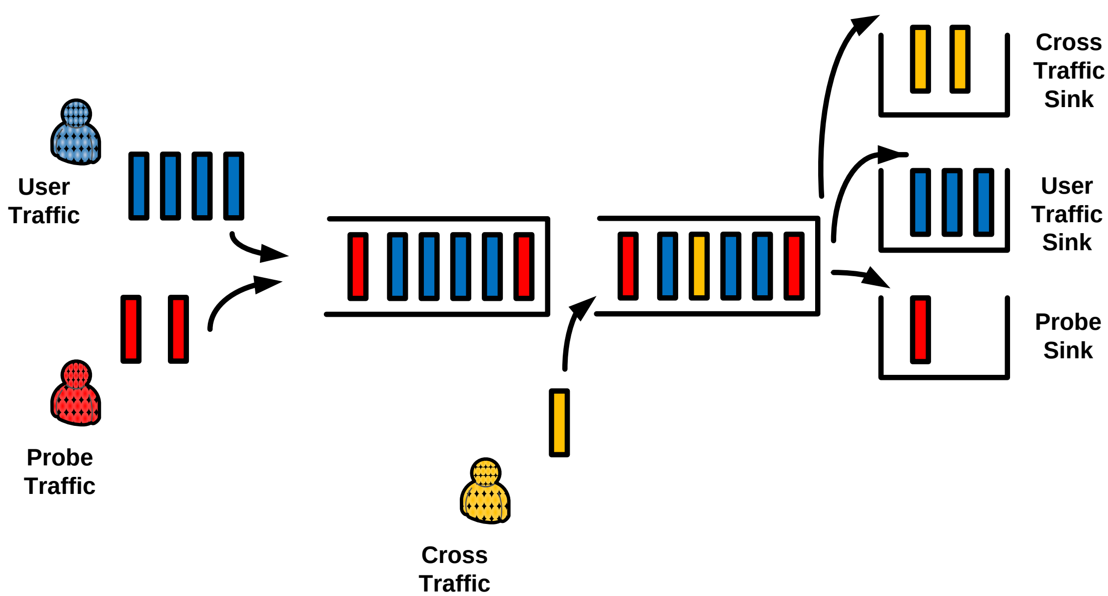

Figure 6-3 Schematic of Two Hop Loss Probing Scenario

The experimental setup for the two hops wireless network is the same as the single hop, just with the addition of an intermediate node between source nodes and destination node as shown in Figure 6-3. The simulation parameters used for traffic are the same as given in Table 6-1, and the simulator parameters are the same as listed in Table 6-2. The number of traffic sending nodes was increased in order to increase the load on the network and create contention for channel access. The middle node acts as a forwarding node. In order to make this scenario more realistic cross traffic is introduced on the forwarding node. However, care was taken to ensure that the load caused by cross traffic is not so high, that it destroys our simulation purpose. The cross traffic parameters are stated in Table 6-3.

Cross Traffic type	Exponential ON/OFF
Application Type	UDP
Packet Size (cross traffic)	1000 bytes
Mean ON time of cross traffic (Ton)	0.35 sec
Mean OFF time of cross traffic (Toff)	0.67 sec
Input data traffic rate (R)	64 kbps

Table 6-3 Cross Traffic Parameters

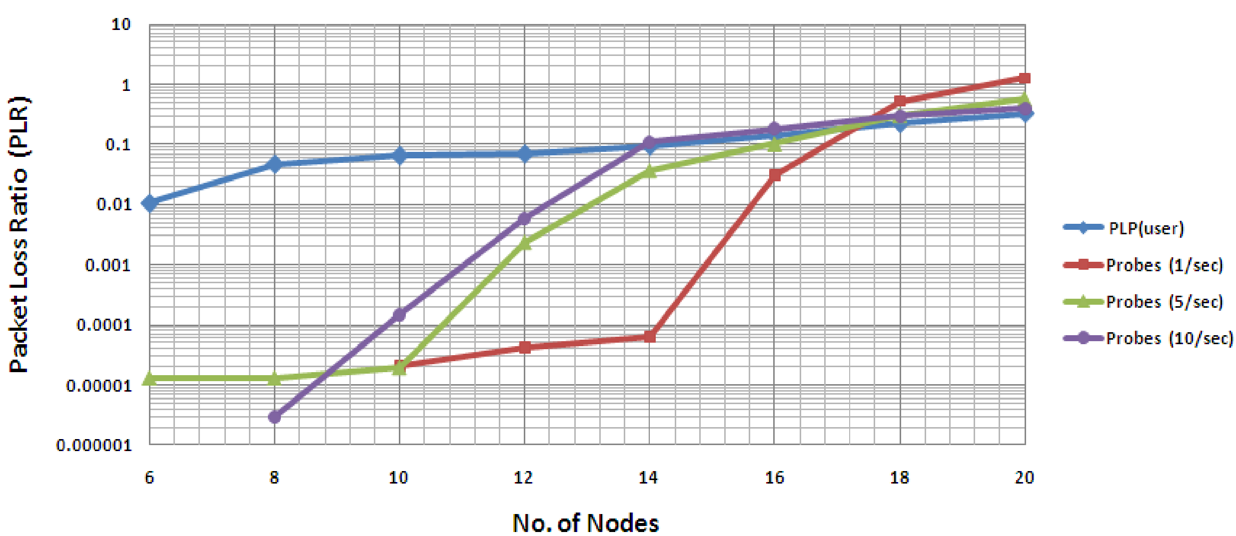

Figure 6-4 Packet loss ratio comparison of probes traffic and user traffic

From Figure 6-4 it can be observed that the packet loss ratio of the user traffic is far more than that of the probes initially. Results start getting better for probes when the number of nodes go beyond 12 nodes. Although the loss ratio estimated by 1 probe per second, is not quite comparable to that of the user traffic, probing at 5/sec resulted in very close values to that of user traffic at high load values. The results for probe's packet loss ratio are very much in accordance with user packet loss ratio, however it has been observed it is not practically acceptable, as user traffic delay increase to large values as shown in chapter 5.

6.3 *Three Hops Scenario*

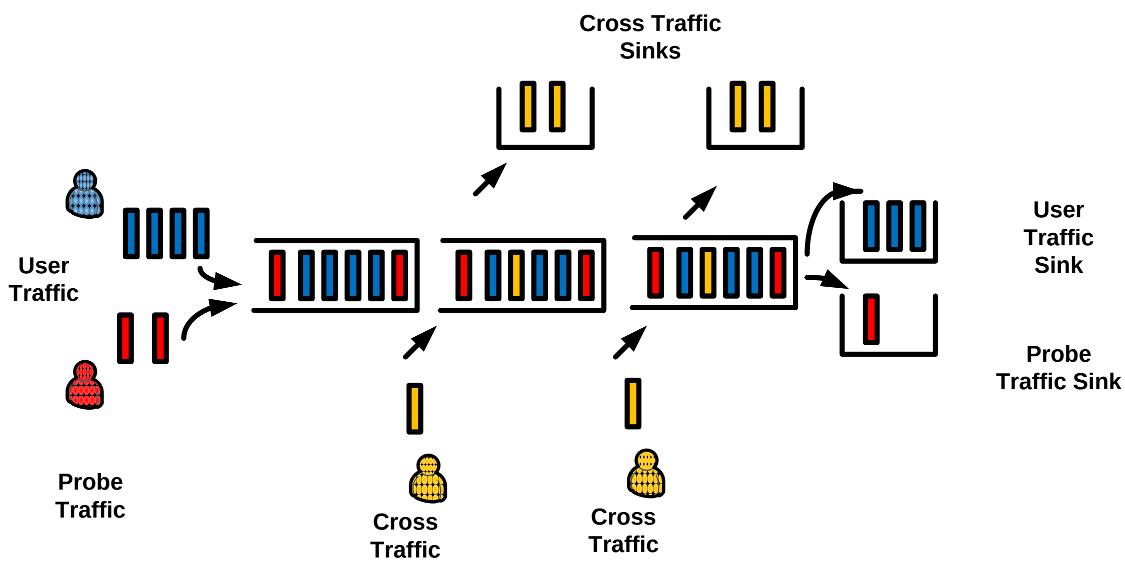

Figure 6-5 Schematic for Three Hops Loss Probing Scenario

In this section, the investigation is further extended to three hops as shown in Figure 6-5. We introduce two intermediate nodes between source nodes and destination. We also called these nodes 'Forwarding nodes', as their main purpose is to rely or forward the user packets to the destination node. The parameters are the same as listed in Table 6-1 and Table 6-2. The parameters for cross traffic are same as listed in Table 6-3.

It has been observed that as the number of hops increases, the accuracy achieved by the probes in measuring packet loss rate decreases, especially when probing at a low rate. It can be seen in Figure 6-6, that the packet loss ratio obtained by probes at the rate 1 probe/sec is not close to the user traffic. One of the key points noticed during the simulation is that, in a Multihop wireless network, it is not necessary that the probes and user traffic packets traverse the same path to reach the destination node. The distance of the probing node and user traffic nodes to the destination nodes in terms of hops also matters.

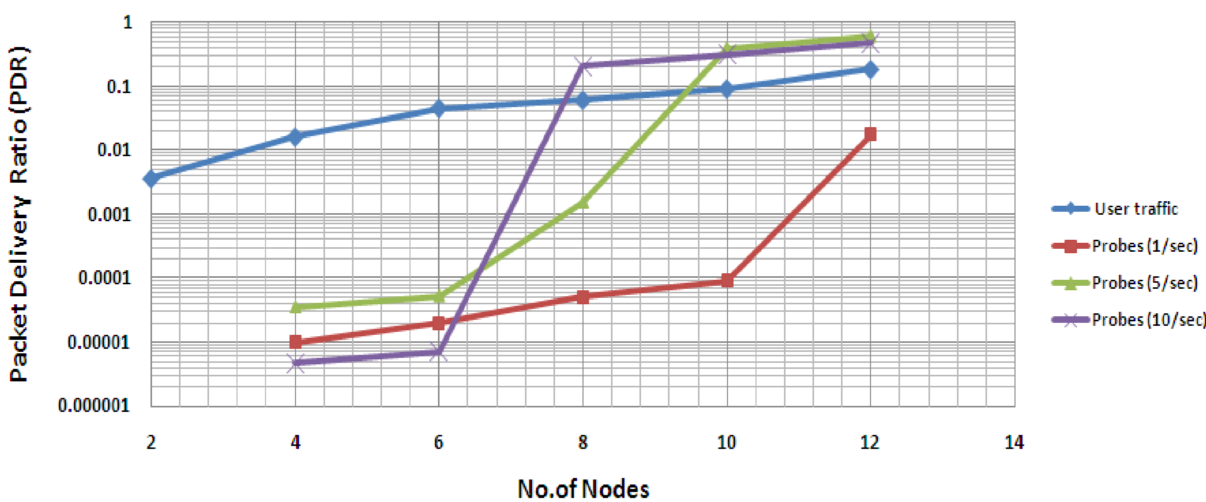

Figure 6-6 Packet loss ratio comparison of probes traffic and user traffic

We have observed from all of the simulation scenarios that active probing can provide a good estimate of user traffic characteristics for single hop when probing at 5/sec, therefore it can be used to investigate the losses and delays occurring in last hop wireless networks. As the number of hops and alternative paths increases, the issue of hidden node, exposed node as well as path selection cannot be ignored. From our experimental results 5/sec probing rate observed the closest correlation with user traffic packet loss ratio at higher values of load.

6.4 Accuracy

In this section we compare the absolute error for the three probing rates, and it is observed that a probing rate of 1/sec results in a higher value of standard deviation, thus the absolute error is high as well and provides the least accurate results as compared to the other two probing rates (i.e. 5/sec and 10/second). The three graph 6-7, 6-8 and 6-9 gives a comparison of accuracy for the three probing rate for single hop, two hops and three hops respectively.

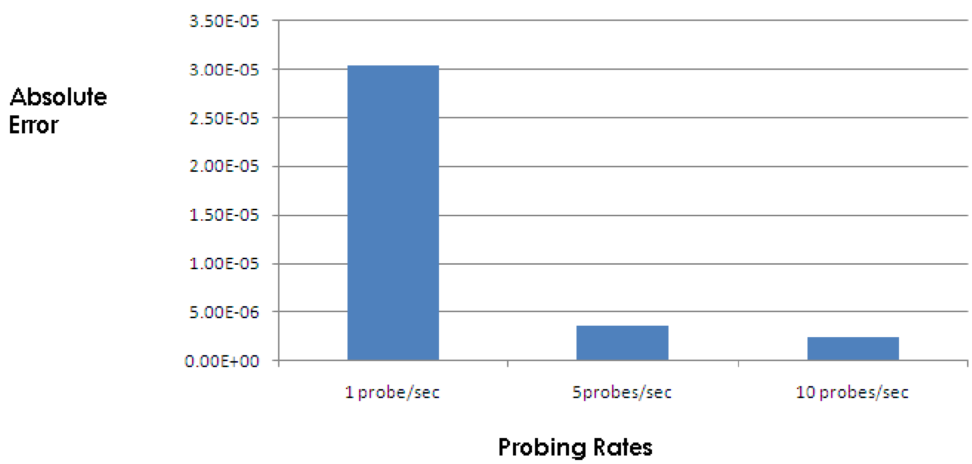

Figure 6-7 Absolute error in PLP for 1 hop with 28 nodes

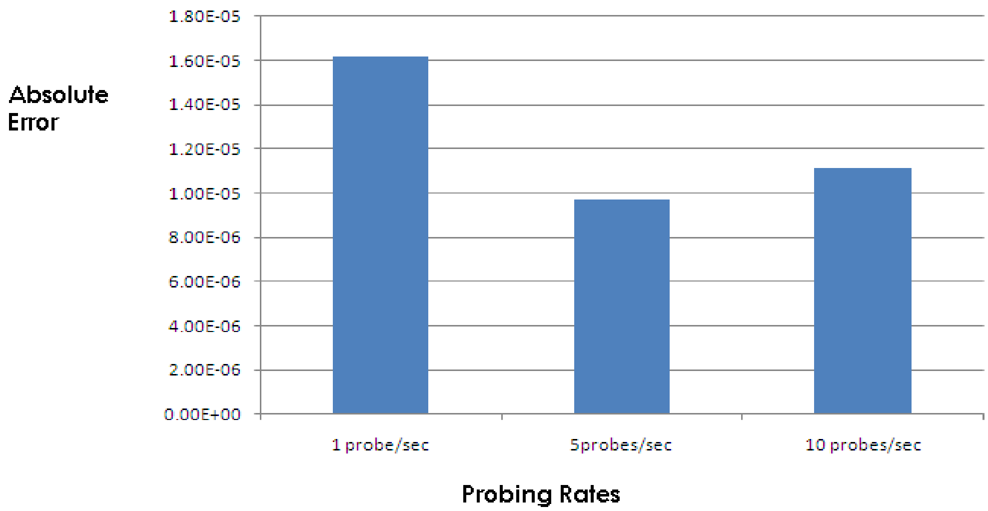

Figure 6-8 Absolute error in PLP for 2 hops with 20 nodes

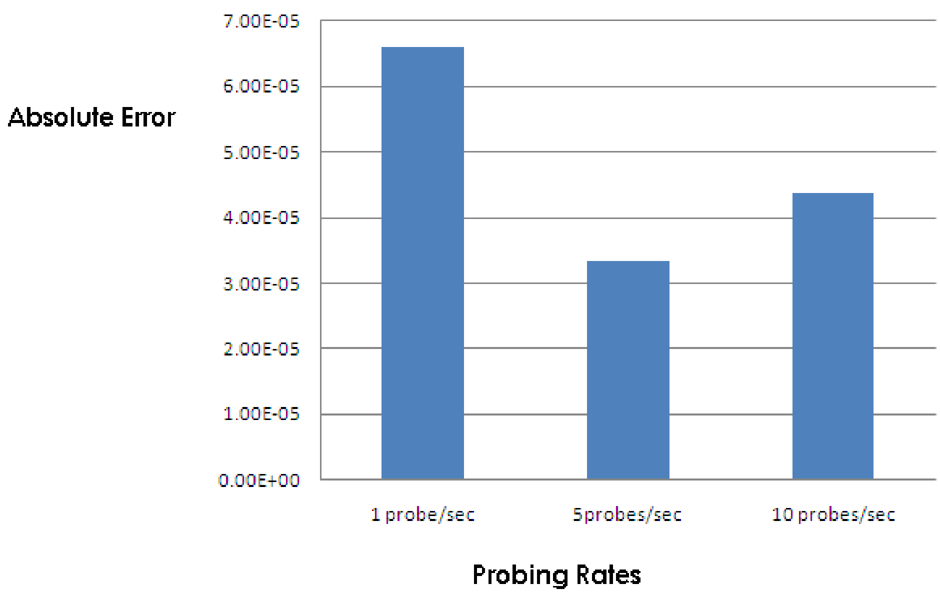

Figure 6-9 Absolute error in PLP for 3hops with 12 nodes

It can be seen from Figure 6-7 that absolute error was least for 10 probes per second but as the hop count increased to 2, the probing error first decreased when probing at 5 probes per second and then increased again for 10 probes per second. It was observed that as the number of hops increased the packet loss ratio of probes were higher as compared to other user nodes as they found alternative path to reach destination and probes are more frequently lost as compared to user traffic.

6.5 Effect of Probing on User Traffic

Figure 6-10 is a clear example of probe effect on user traffic. Without probes the maximum achievable throughput of a single hop wireless network was 0.78Mbps which fell to 0.63Mbps with probing rate of 5/sec, and to 0.51 Mbps for 10 probes per second probing frequency. This implies that around 15% of total bandwidth is consumed by probes as probing overhead when probing at the rate of 5 probes per second and around 26% of total bandwidth is eaten up by probes when probing at the rate of 10/sec.

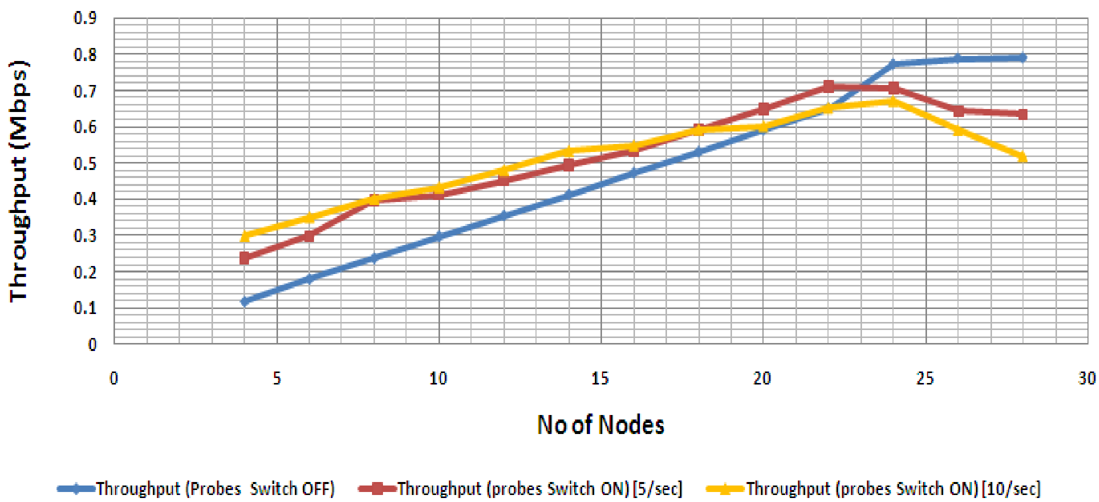

Figure 6-10 Throughput Comparison for Single Hop

6.6 Summary

In this chapter we have probed the network for the packet loss rate. We have employed periodic probing for three different probing rates. We observed that when we probe the network with the rate of 1 probe every second there is almost no association in packet loss rate results of user traffic and probe results.

A probe could only measure loss if the buffer/queue inside the mobile node is already full of user traffic. If the queue is full when a probe enters the queue, it is likely that the user traffic has already lost some packets. As the probes would not encounter full queues all the time (non-bursty period as said in queuing literature) their loss would be less than the traffic loss; also probes are sent as single packets were as user traffic is often bursty at higher values of load. A single packet is more likely to be stored in buffers than a whole burst of packets. Probes gave quite accurate results for higher load values when r probing frequency of 5/sec was used.

7. Conclusions and Future Works

7.1 Conclusions

The aim of the research was to find the accuracy of active probing over wireless access. We have used NS2 simulator to simulate network scenarios in order monitor the accuracy of the probes and affects of probes on user traffic. In this research we have employed periodic and poisson probing techniques to evaluate the effectiveness of measuring end to end delay distribution and packet loss rate of user traffic.

We have probed single hop and multiple hop static Ad-hoc networks, which rely on carrier sensing random access protocol. Simulation experience revealed that wireless networks specifically Ad-hoc wireless networks, have surprisingly low capacity, due to bandwidth sharing and nodes forwarding packets o each other. Such characteristics of medium made active probing even more challenging.

We experimented with different probe packet sizes and concluded that when a small probe packet size was used in order to minimize probing overheads, it led to larger difference in delay values as compared to actual traffic delay (smaller sized packets experience smaller delay values). Therefore, packet size of probes should be equal to the actual user traffic data so that what probe measure is as close to the target as possible. For this reason we used same size for probes and user traffic packets i.e. 1000bytes for investigation purposes.

Four different probing rates were employed to probe user traffic, i.e. 1 probe every 10 sec, 1 probe per second, 5 probes per second and 10 probes per second. Out of all we observed that probing rate of 5/sec gave us more accurate results. Probe rate of 1/10sec failed to capture the delay distribution; 1/sec proved accurate results for lightly loaded network but failed for higher values of load whereas, probing rate of 10/sec forced network to early saturation due to huge amount of probing overhead and skewing results to unacceptable extend.

An Important finding during active probing the wireless access for delay distribution was, unlike wired networks when we get “packet scale decay” i.e. a slope in log plot of probability, we observed “step like or stair response” the size of these stairs gradually gets smaller with increase in nodes (in the network) and at higher values of load it further smoothes down to a slope, and where we can observe both “packet scale decay” and “burst scale delay”. The possible reason behind these steps like response is the fluctuation in the size contention window due to collision.

We observed that when we extended our investigation for multiple hops, the trend of results for delay distribution were more or less same with the difference that network goes in saturation more earlier. This was because the throughput and packet delivery ratio decreases with the increase in hops between source and destination nodes.

Our Research suggests that active probing provides effective results specifically for a single hop wireless network. The absolute error increase when the number of hops increases because it is not necessary that probe traffic and user traffic are traversing the same routing path to reach destination, if there are large amount of nodes in the network. This situation can get more critical with mobile nodes. However, for static networks (with dedicated paths or while monitoring a bottleneck wireless link) or last hop wireless network, active probing can provide a good approximation of user traffic characteristics.

7.2 Future Works

Suggestions for future work falls into two categories: expanding the range of simulation for active probes and trying out different other types of sampling techniques for probing.

This research assumed deterministic i.e. periodic probing and poisson probing; however other monitoring strategy that involve release of probes as batched, rather singly. Other probing techniques can be designed e.g. adaptive optimal probing by sensing wireless channel and adjusting the rate and size of probe packets accordingly in order to minimize the probing overhead and increasing the accuracy in measurement. Active measurement and Passive measurements can be combined in order to get more accurate results. Moreover, researchers can work on schemes that will be scalable and require minimal overhead so that they can be used to optimally manage measurements and monitoring aspect of wireless networks This will be important for wireless network e.g. mobile and 3G and beyond as they suffer from a major constraint of limited bandwidth.

In this research, only ON/OFF static model was used, however measurement can be performed for more electric traffic like TCP.

The active probing simulation could be extended to investigate the measurement accuracy in last hop wireless network, mesh networks. In addition to this the effect of mobility in wireless network can be taken in consideration while probing.

8. Author's Publications

1. Naqvi, S.S, Schormans, J.A.: "*Simulation based Analysis on Applicability of Active Measurement over Wireless Access*" PGNET Conference, 2010, Liverpool John Moore University.
2. Naqvi, S.S, Schormans, J.A.: "*Assessing Active Measurement for Delay Distribution of Single & Multi-hop Ad-Hoc Wireless Network* to be submitted for publication to IET Communications.

9. References

[1] Choi, B.K.; Moon, S.; Cruz, R.; Zhang, Z.L.; *Quantile Sampling for Practical Delay Monitoring in Internet Backbone Networks,* Computer Networks: The International Journal of Computer and Telecommunications Networking, Vol. 51, Issue 10, July, 2007.

[2] Sarr, C.; Lassous, I.G.: *Estimating Average End–to-End Delays in IEEE 802.11 Multihop Wireless Networks.*

[3] Ishibashi, K.; Kanazawa, T.; Aida, M.: *Active/passive combination-type Performance Measurement Method using change of measure Framework,* IEEE GLOBECOM '02, issue 17-21, pp. 2538-2542, vol.3, March, 2002.

[4] Timotijevic, T. and Schormans, J.A.: "*Bandwidth overhead of probe technology guaranteeing QoS in packet networks*". IEE Electronics Letters, 15th Vol. 39, No. 10, pp 816-818. May, 2003.

[5] Roughan, M.: "*Fundamental bounds on the accuracy of Network Performance Measurements*" Joint International Conference on Measurement and Modeling of Computer Systems archive Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems , pp. 253-264 , Banff, Alberta, Canada 2005.

[6] Sommers, J.; Barford, P.; Duffield, N.; Ron, A: "*Improving accuracy in end-to-end packet loss measurement*" Proceedings of ACM SIGCOMM, August, 2005.

[7] Leung, C.M.; Schormans, J.A.; "*Measurement-based Traffic Characteristic Estimation for QoS Oriented IP Networks"* In 41st European Telecommunications Congress, Genoa, pp. 4-7, Sept. 2002.

[8] Giuseppe, A; Borgia, E; Conti, M; Gregori, E: "Wi-Fi in Ad-Hoc Mode: A Measurement Study", Pervasive Computing and Communications, pp. 145-154, March 2004.

[9] Meky, M.M.; ,Saadawi, T.N,; *Degradation Effect of Cell Loss on Speech Quality over ATM Networks*", in Broadband Communication Ed. Mason & Canaca, pub. By Chapman & Hall, 1996.

[10] Oodan,A.P.; Ward,K.E.; Savolaine,C.G, Daneshmand,M.; Hoath,P.; *Telecommunications Quality of Service Management from legacy to emerging services services* IEE Telecommunications Series 48, 2003.

[11] http://wand.cs.waikato.ac.nz/wand/publications/jamie_420/final/node9.html

[12] Wang, K.; Li, Z.C.; Yang, F.; Wu, Q.; Bi, J.P.; *Experiment and Analysis of Active Measurement for Packet Delay Dynamics,* ICCNMC 2005, pp.1063-1072,2005.

[13] Leung, C.M,; *Non-Intrusive Measurement in Packet Networks and its Applications*, PhD Thesis, University of London, 2004

[14] Chien, M.W.; Su, H.K.; Yang, W.H.; "Efficent Backoff Algorithm in Wireless Multihop Ad Hoc Networks" International Journal of Advancements in Computing Technology, Vol. 3, No. 1, Feb. 2011

[15] Hong, J.; *Traffic Measurement for High-Speed Internet*, 2003 ICAT Workshop, Seoul, Korea, Apr. 2003

[16] Moon. B.S.; Measurement *and Analysis of End-to-End Delay and Loss in the internet*, PhD Thesis, University of Massachusetts Amherst 2000.

[17] Pasztor, A.; Veitch, D.; *High Precision Active Probing for Internet Measurement* Proceedings of INET 01, April 2009.

[18] Coates, M.; Nowak, R.; *Network Tomography for Internal Delay Estimation*, IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2001.

[19] NONPARAMETRIC INTERNET TOMOGRAPHY: http://www.ngisupernet org/NGI-PI-2001/Towsley.pdf

[20] Ribeiro,V.; Coate, M., Riedi, R.; Sarvotham, S.; Hendricks,B.; and Baraniuk,R.; *Multifractal cross -traffic estimation*, Proc. of ITC Specialist Seminar on IP Traffic Measurement, Sept. 2000.

[21] Hu, N.; Steenkiste, P., *Evaluation and Characterization of Available Bandwidth Probing Techniques,* IEEE JSAC Special Issue in Internet and WWW Measurement, Mapping, and Modeling, 2003.

[22] Jiang, Y. ; et al., *Measurement-Based Admission Control: A Revisit*, Center for Quantifiable Quality of Service in Communication Systems Department of Telematics. Norway, 2004.

[23] Mase, K.; Toyama, Y., *End-to-end measurement based admission control for VoIP network,* Communications, 2002. ICC 2002. IEEE International Conference on Volume 2, 28 April-2 May 2002, pp.1194 - 1198.

[24] Strauss, J.; Katabi, D.; Kaashoek, F., *A Measurement study of available bandwidth estimation tools,"* IMC 2003.

[25] Hasib, M.; *Analysis of Packet Loss Probing in Packet Networks*, PhD thesis, Queen Mary, University of London, Jun. 2006

[26] http://e2epi.internet2.edu/owamp/

[27] Willinger, W.; Taqqu, M.S; Sherman, R.; Wilson, D.V.; *Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at traffic at the source level,* IEEE/ACM Transaction on Networking, vol.5, no.1, pp. 71-86, 1997

[28] Pitts, J.M.; Schormans, J.A.; *Introduction to IP and ATM design and performance*, Wiley, 2000. Part 3 : IP Performance and Traffic Measurement.

[29] Pitts, J.M.; Schormans, J.A.; *Introduction to ATM design and performance*, Wiley, 1996.

[30] ANSI/IEEE STD 802.11, \Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", 1999

[31] Ishibashi, K.; Kanazawa, T.; Aida, M., *Active/passive Combination-type Performance Measurement Method using Change-of-measure Framework"* Global Telecommunications Conference, GLOBECOM '02. IEEE ;vol.3, pp.2538- 2542, 2002.

[32] Keshave, S., *A Control-Theoretic Approach to Flow Control,* Proc. ACM SIGCOMM '91, 1991.

[33] Caceres, R.; Duffield, N.G.; Horowitz, J.; Towsley, D.F.; Bu, T., *Multicast-Based Inference of Network-Internal Characteristics: Accuracy of Packet Loss Estimation*, Proc. IEEE INFOCOM, 1999.

[34] *Internet End-to-end Performance Monitoring,* URL: http://www-iepm.slac.stanford.edu/

[35] Tariq, M.M.B.; Dhamdhere, A.; Dovrolis, C.; Ammar, M., *Poisson versus Periodic Path Probing,* ACM Internet Measurement Conf., Berkeley, CA, pp. 119-124, Oct 2005.

[36] DeMeer, H.; Sasse, H.; *Of Packets and People A User Centered Approach to QoS*, IEEE/IFIP International Workshop on QoS, 2000.

[37] http://en.wikipedia.org/wiki/Processing_delay

[38] Tickoo, O.; Sikdar, B., *Traffic Characteristics and Heavy-Tails in 802.11 MAC Based Wireless Networks*, Globecom, Nov. 2000 .

[39] Leung, C.M., "*Non-Intrusive Measurement in Packet Networks and its Applications*", PhD Thesis, University of London, 2004

[40] Hernandez, A.; Mangana, E.; "One-Way Delay and Characterization" Third international conference on Networking and Services (ICNS'07), 2007.

[41] Macker, J.P.; Corson, M.S.; *Mobile Ad Hoc Networking and the ISTF, ACM Mobile Computing and Communications Review*", Vol.2, No. 1, Jan 1998.

[42] Xu, K.; Gerla, M.; Sang, B., *How Effective is the IEEE 802.11 RTS/CTS Handshake in Ad Hoc Networks* in Proc. Of IEEE GLOBECOM' 02, Vol. 1, pp. 72-76, Nov. 2002.

[43] Schormans, J.A.; Leung, C.M., *Measurement for Guaranteeing QoS in Broadband Multiservice Networks*. Invited chapter (8) in *Mobile and Wireless Systems Beyond 3G: Managing New Business Opportunities* , ISBN 159140544-0. Prof. Pagani editor. 2005.

[44] Gupta.D; Mohapatra.P; Chen, N.C, *Efficient Monitoring in Wireless Mesh Networks: Overheads and Accuracy Trades –offs*, 5th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems, pp. no : 13-23, 2008.

[45] Larsson, T.; Liu, Y.; *A Study of EDCA and DCF in Multihop Ad Hoc Networks*, MSc. Thesis, March 2008.

[46] Yu, C.; *Mobile Computing* : MSc. Lectures, Cleveland State University.

[47] Michiel, H.; Laevens, K.; *Teletraffic engineering in a broad-band era*, Proceedings of the IEEE, vol. 85, pp. 2007-2003, Dec. 1997

[48] Amaradasa,V.; Schormans, J.A.; Pitts, J.M,; Leung C.M.; *Evaluating Overflow Probability For VoIP Buffer Dimensioning*, Journal of Communication IET, vol. 3, pp.636-643, Apr. 2009

[49] Affirrin, S.H.S.; *Accelerated Simulation for Packet Buffer with a Non FIFO Scheduler*, PhD thesis, Queen Mary, University of London, Mar. 2006

[50] http://www.cs.kent.ac.uk/people/staff/pfl/presentations/longrange/sl003.htm

[51] Adas, A.; Traffic Models in Broadband Networks, Georgia Institute of Technology, IEEE Communications Magazine ,July 1997.

[52] Li, H.; Muskulus, M.; Wolters, L.; *Job Arrivals in a Data-Intensive Grid",* Leiden University, The Netherlands.

[53] Singh, L.N.; Dattatreya, G.R.; *A Novel Approach to Parameter Estimation in Markov modulated Poisson processes*, IEEE Emerging Technologies Conference (ETC), Richardson Texas, Oct 2004.

[54] Heffes, H.; Lucantoni, D.; *A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance*, IEEE JSAC, pp. 856-868 Sept. 1986.

[55] Leland, W.; *et al.*; *On the Self-Similar Nature of Ethernet Traffic(Extended Version)*, IEEE/ACM Trans. Networking, pp. 1-15, Feb. 1994.

[56] Klivanski, S.; Mukherjee, A.; Song, C.; *On Long-Range Dependence in NSFNET Traffic*, Technical Report GIT-CC-94-61, Georgia Tech., 1994.

[57] Crovella, A.; Bestavros, A.; *Explaining World Wide Web Traffic Self-Similarity*, Technical Report TR-95-015, 1995.

[58] www.isi.edu/nsnam/ns/ns-tutorial/wireless.ppt

[59] Amaradasa,V.; *Performance Evaluation by Aggregation of ON-OFF Traffic in Processor Sharing and Related Systems,* PhD thesis, Queen Mary, University of London, Aug. 2008

[60] Gibb, Allan.; St-Jacques, J.C.; *A Comparison of Deterministic vs. Stochastic Simulation Models for Accessing Adaptive Management Techniques over Disadvantaged Tactical Communication Networks.*

[61] Schormans, J.A.; Pitts, J.M.; Scharf, E.M.; Pearmain, A.J.; and Phillips, C.I.; *Buffer Overflow Probability for Multiplexed ON-OFF VoIP Sources*, IEE Electronic Letters, pp. 523-524, 36, (6), March 2000.

[62] Hill, J.; *Assessing the Accuracy of Active Probes for Determining Network Delay, Jitter and Loss*, MSc Thesis in High Performance Computing, The University of Edinburgh, 2002.

[63] Zeng, J.L.; Su, Q.; "*Performance Investigation of backoff Algorithms in Multihop Wireless Networks*" 9th International Conference for Young Computer Scientists, 2008.

[64] Xu, S.; Saadawi, T.; "*Does IEEE 802.11 MAC Protocol Work Well in Multihop Wireless Ad Hoc Networks?*" IEEE Communication Magazine, pp.130-137, 2001

[65] Korhonen, J.; Wang, Y.; " Effect of Packet Size on Loss Rate and delay in Wireless Links"IEEE Communication Society WCNC 2005,Los Angeles, pp. 13-17, Mar. 2005.

10. Appendix A: Decay Rate Analysis

The duration of the ON periods are exponentially distributed. Define first for each of the individual sources:

N = No of ON/OFF traffic sources

T_{on} = Mean ON Time of a traffic source

T_{off} = Mean OFF Time of traffic source

h = ON Rate of single source

A = the mean applied load in data unit time

C= the output rate

The rate, at which these active periods arrive, from the population of N packet sources, is

$$F = \frac{N}{T_{on} + T_{off}}$$ **Equation 10-1**

Therefore, we can find the overall mean load, A_p, and the offered traffic, A, in erlangs

$$A_p = F.T_{on}.h$$ **Equation 10-2**

$$A = F.T_{on}$$ **Equation 10-3**

The maximum number of the sources that can be served simultaneously, without exceeding the buffer's service rate is

$$N_o = \frac{C}{h}$$

Blocking Probability is given by

$$B = \frac{\frac{A^{N_o}}{N_o!}}{\sum_{r=0}^{N_o} \frac{A_r}{r!}}$$

$$D = \frac{N_o B}{N_o - A + A.B}$$ **Equation 10-4**

$$R_{on} = C + h.\frac{A_p}{C - A_p}$$ **Equation 10-5**

$$R_{off} = \frac{A_p - D.R_{on}}{1 - D}$$ **Equation 10-6**

$$T_{(on)} = \frac{h.T_{on}}{C - A_p}$$ **Equation 10-7**

$$T_{(off)} = T_{(on)}.\frac{1 - D}{D}$$ **Equation 10-8**

The reduced 2-state model, has the following properties: In the ON state when the total input rate exceeds the service rate of the buffer, and the buffer fills (rate of increase = Ron - C); in the OFF state the total input rate is less than the output rate of the buffer, so allowing the buffer to empty (reduction rate = C - Roff). When not less than C/h sources are active at any time the aggregate process is in the ON state, otherwise it is in the OFF state [25].

Now we can calculate the parameters, a and s, and hence the decay rate

$$a = 1 - \frac{1}{T_{(on)}.(R_{on} - C)}$$ **Equation 10-9**

$$s = 1 - \frac{1}{T_{(off)}.(C - R_{off})}$$ **Equation 10-10**

$$\textit{Decay Rate} = \frac{a}{s}$$ **Equation 10-11**

11. Appendix B: Delay Distribution Measurement of Power Law Traffic using Poisson Probing

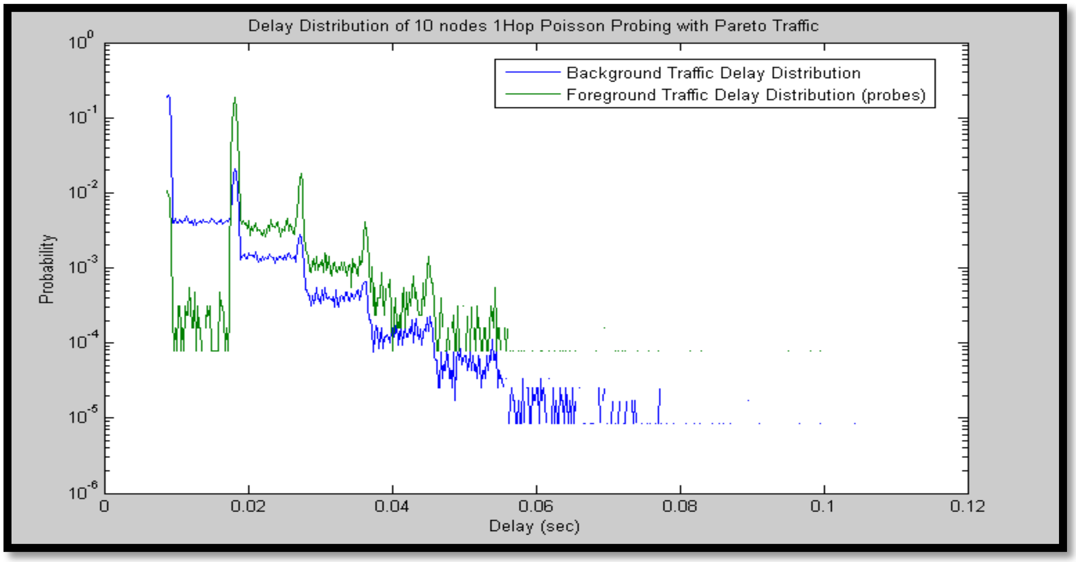

Figure 11-1 Delay probability plots for probes(foreground traffic) and user traffic(background traffic) in a single hop, 10 nodes secenario with probing rates of 5/sec

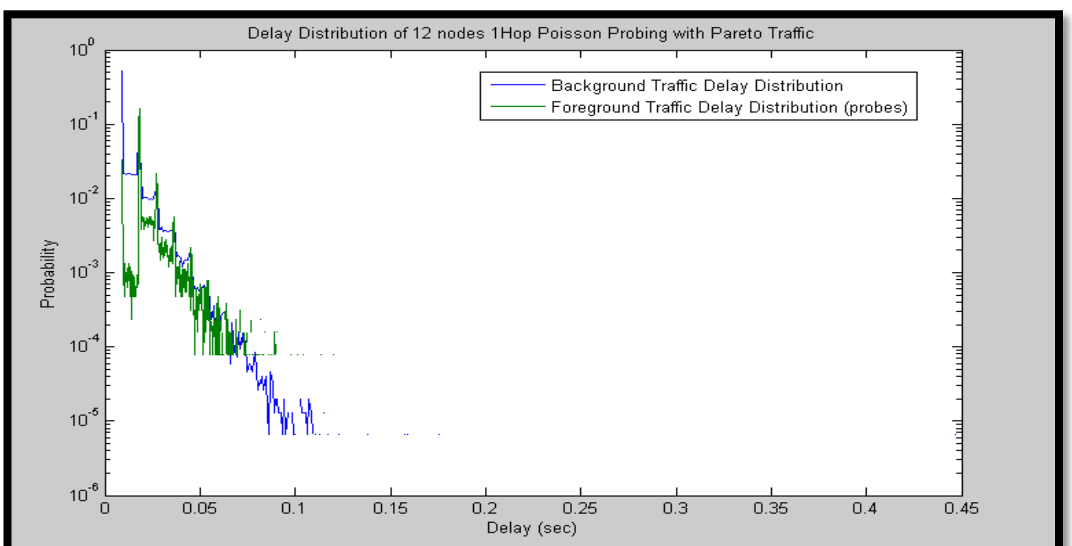

Figure 11-2 Delay probability plots for probes(foreground traffic) and user traffic(background traffic) in a single hop, 12 nodes secenario with probing rates of 5/sec

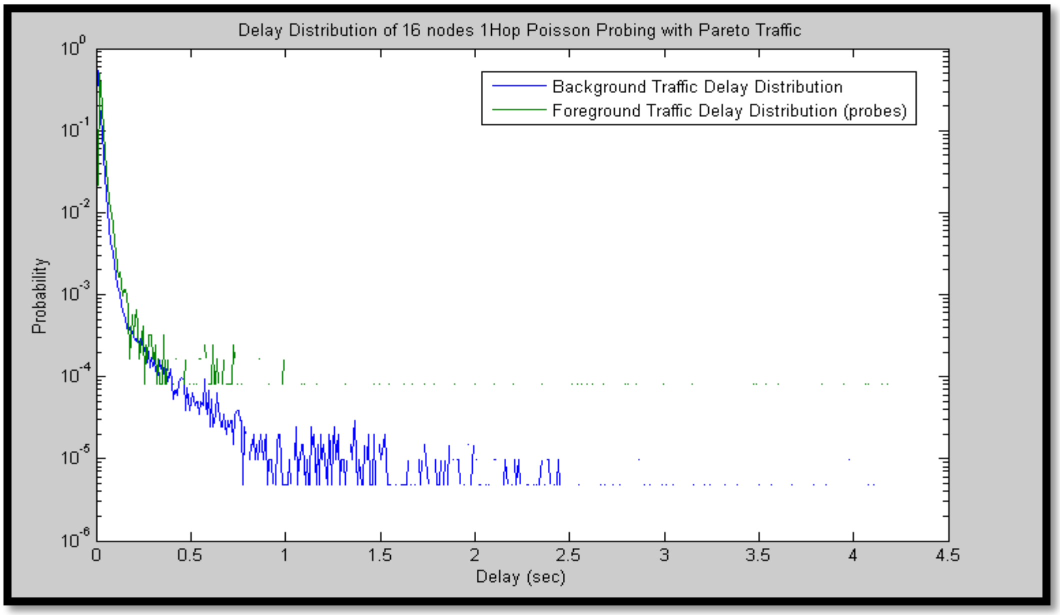

Figure 11-3 Delay probability plots for probes(foreground traffic) and user traffic(background traffic) in a single hop, 16 nodes secenario with probing rates of 5/sec

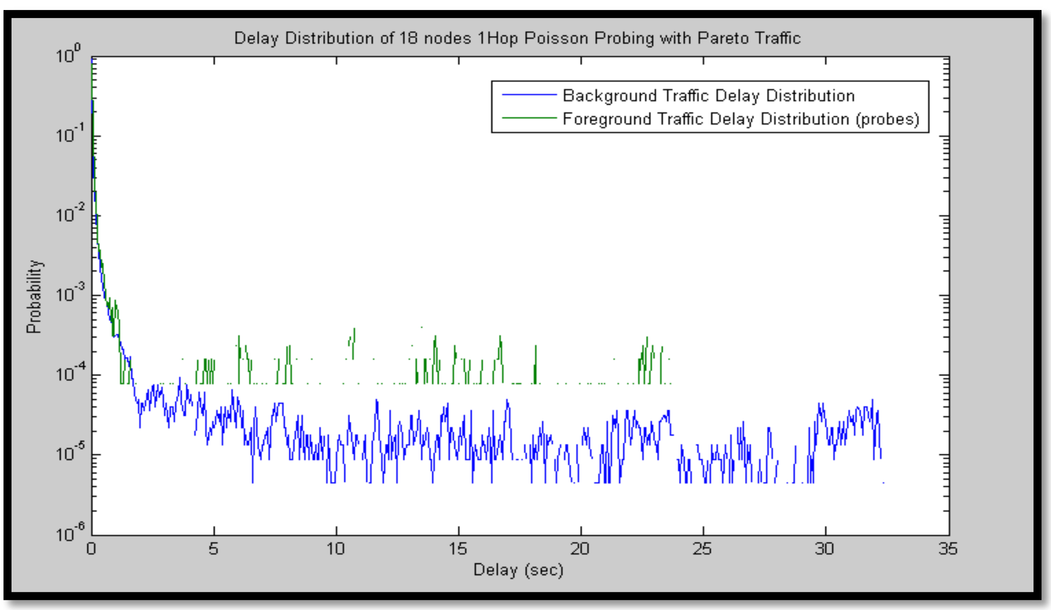

Figure 11-4 Delay probability plots for probes(foreground traffic) and user traffic(background traffic) in a single hop, 18 nodes secenario with probing rates of 5/sec

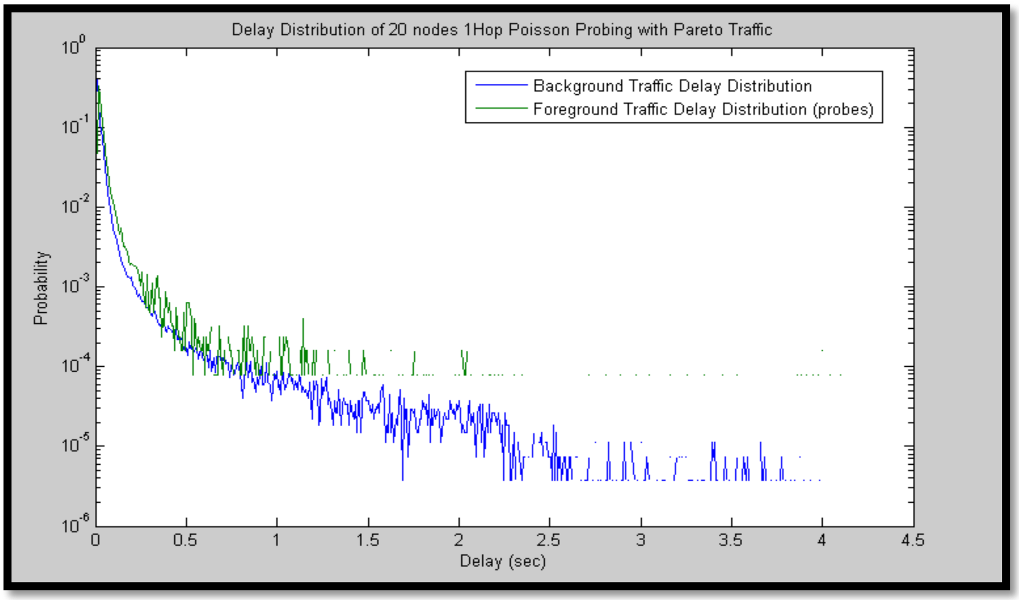

Figure 11-5 Delay probability plots for probes(foreground traffic) and user traffic(background traffic) in a single hop, 20 nodes secenario with probing rates of 5/sec

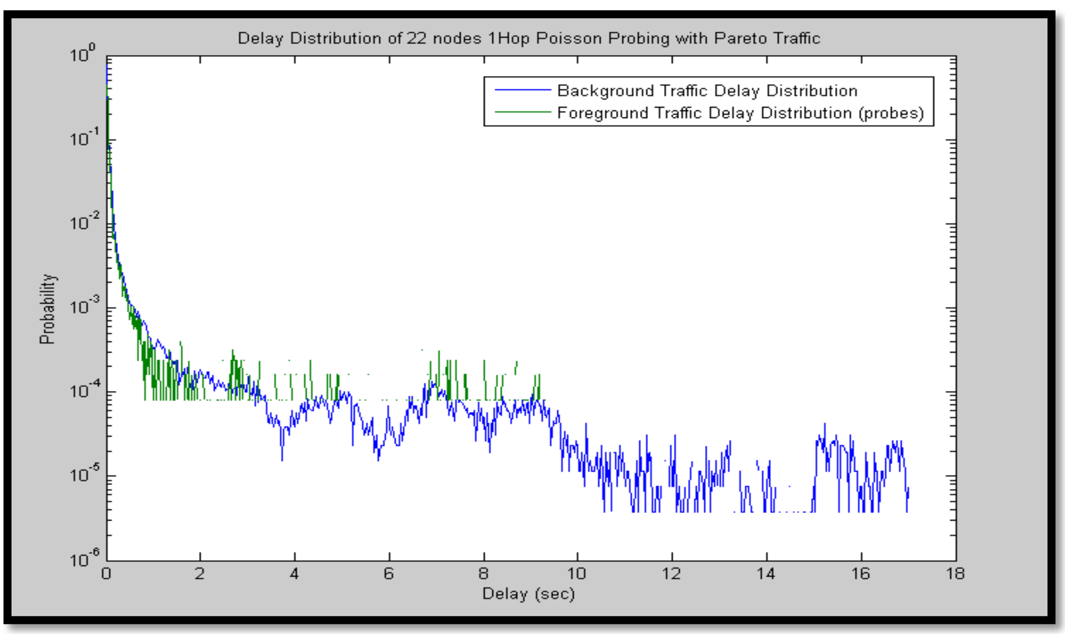

Figure 11-6 Delay probability plots for probes(foreground traffic) and user traffic(background traffic) in a single hop, 22 nodes secenario with probing rates of 5/sec