IMPACT OF WIRELESS LOSSES ON THE PREDICTABILITY OF END-TO-END FLOW CHARACTERISTICS IN MOBILE IP NETWORKS

A Thesis

by

SAMEER BHOITE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2004

Major Subject: Mechanical Engineering

IMPACT OF WIRELESS LOSSES ON THE PREDICTABILITY OF END-TO-END FLOW CHARACTERISTICS IN MOBILE IP NETWORKS

A Thesis

by

SAMEER BHOITE

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Alexander G. Parlos (Chair of Committee)	Suhada Jayasuriya (Member)
Madhav Pappu (Member)	Dennis L. O'Neal (Head of Department)

December 2004

Major Subject: Mechanical Engineering

ABSTRACT

Impact of Wireless Losses on the Predictability of End-to-End Flow Characteristics in Mobile IP Networks. (December 2004)

Sameer Bhoite, B. E., Visvesvaraya Regional College of Engineering

Chair of Advisory Committee: Dr. Alexander G. Parlos

Technological advancements have led to an increase in the number of wireless and mobile devices such as PDAs, laptops and smart phones. This has resulted in an ever-increasing demand for wireless access to the Internet. Hence, wireless mobile traffic is expected to form a significant fraction of Internet traffic in the near future, over the so-called Mobile Internet Protocol (MIP) networks. For real-time applications, such as voice, video and process monitoring and control, deployed over standard IP networks, network resources must be properly allocated so that the mobile end-user is guaranteed a certain Quality of Service (QoS). As with the wired and fixed IP networks, MIP networks do not offer any QoS guarantees. Such networks have been designed for non-real-time applications. In attempts to deploy real-time applications in such networks without requiring major network infrastructure modifications, the end-points must provide some level of QoS guarantees. Such QoS guarantees or QoS control, requires ability of predictive capabilities of the end-to-end flow characteristics.

In this research network flow accumulation is used as a measure of end-to-end network congestion. Careful analysis and study of the flow accumulation signal shows that it has long-term dependencies and it is very noisy, thus making it very difficult to predict. Hence, this work predicts the moving average of the flow accumulation signal. Both single-step and multi-step predictors are developed using linear system identification techniques. A multi-step prediction error of up to 17% is achieved for prediction horizon of up to 0.5sec.

The main thrust of this research is on the impact of wireless losses on the ability to predict end-to-end flow accumulation. As opposed to wired, congestion related packet losses, the losses occurring in a wireless channel are to a large extent random, making the prediction of flow accumulation more challenging. Flow accumulation prediction studies in this research demonstrate that, if an accurate predictor is employed, the increase in prediction error is up to 170% when wireless loss reaches as high as 15% , as compared to the case of no wireless loss. As the predictor accuracy in the case of no wireless loss deteriorates, the impact of wireless losses on the flow accumulation prediction error decreases.

To my parents and sister

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Alexander G. Parlos for his sound advice, valuable guidance, kind cooperation and constant encouragement throughout my research under him. I am also thankful to all my friends who were with me whenever I needed their support. I would like to thank Dr. Jayasuriya and Dr. Pappu for being on my committee, and Dr. Langari and Dr. Zoghi for their support.

Lastly, I thank my family for all their sacrifices that made me what I am today.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

TABLE Page

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

The Internet was developed with the primary intention of connecting the different military networks of the United States of America. Since then it has come a long way and is now a primary source of information exchange. The current technological advancement has led to the increase in the number of portable, wireless devices such as personal digital assistants (PDAs), laptops and smart phones. This has caused an ever increasing demand for instant access to the Internet. Consequently, in the near future a major portion of the traffic would be wireless and mobile traffic.

A. Motivation

Increase in the number of Internet users has made it necessary to allocate network resources efficiently so that the end-users are guaranteed certain Quality of Service (QoS). Subsequently, QoS has become an important topic for researchers. This carries greater importance for real-time audio-video applications with mobile end-users. For such real-time applications one must provide the desired QoS. The principal objectives of this research are to predict network congestion for a certain class of mobile wireless networks and to study the impact of wireless losses on the accuracy of these predictions. There are at least two ways by which network congestion can be predicted. One is to analyze end-to-end delays, while the second is by analyzing network accumulation. Accumulation is defined as the number of packets in transit for the flow under consideration at any given point in time. Some researchers have addressed the problem of network congestion using end-to-end delays. From their research results

---

The journal model is *IEEE Transactions on Automatic Control*.

it can be concluded that it is very difficult to predict end-to-end delays [1]. Hence, feasibility of prediction of end-to-end network accumulation will be explored. Based on the current and past values of accumulation, the future values of accumulation will be predicted. A controller will use these predictions to dynamically adjust the source send-rate. The round trip time (RTT) of the network might be up to few seconds. This implies a controller might have to wait for few seconds before it gets the necessary network information and implement the control. In order to compensate for this dead-time one must perform multi-step-ahead predictions. Different prediction horizons will be explored to investigate the accuracy of the predictors developed.

B. Proposed Network Topology

Researchers are currently trying to take advantage of the cheap wireless local area networks (WLAN), in order to provide Internet access to the mobile end-users. This way, wireless users can stay online even when they are moving and take advantage of the seamless user mobility. Available options for cheap WLAN are IEEE 802.11 [2], Bluetooth [3], Home RF [4] etc. for indoor applications, while for outdoor applications the solution might be wireless coverage provided by existing cellular operators in urban and rural areas. At the same time, attempts are being made by some private Internet Service Providers (ISPs) [5] for providing wireless IP services within US. It is expected that more ISPs will enter into this market.

The methods discussed above have a major disadvantage, that the mobility management is handled by underlying wireless infrastructure. This leads to the problem that users might not be able to move seamlessly across different wireless media.

To overcome these problems new methods have been proposed for handling mobility at the IP layer, and such protocols are being developed by Internet Engineering

Task Force (IETF) [6]. Mobile IP (MIP) is one of the solutions provided by IETF for handling mobility. This concept will be used in this work. With the use of MIP mobile users will be able to change their point of attachment to the wired Internet. In MIP no IP address change is required to allow mobility. Figure 1 shows routing of packets within an MIP network.

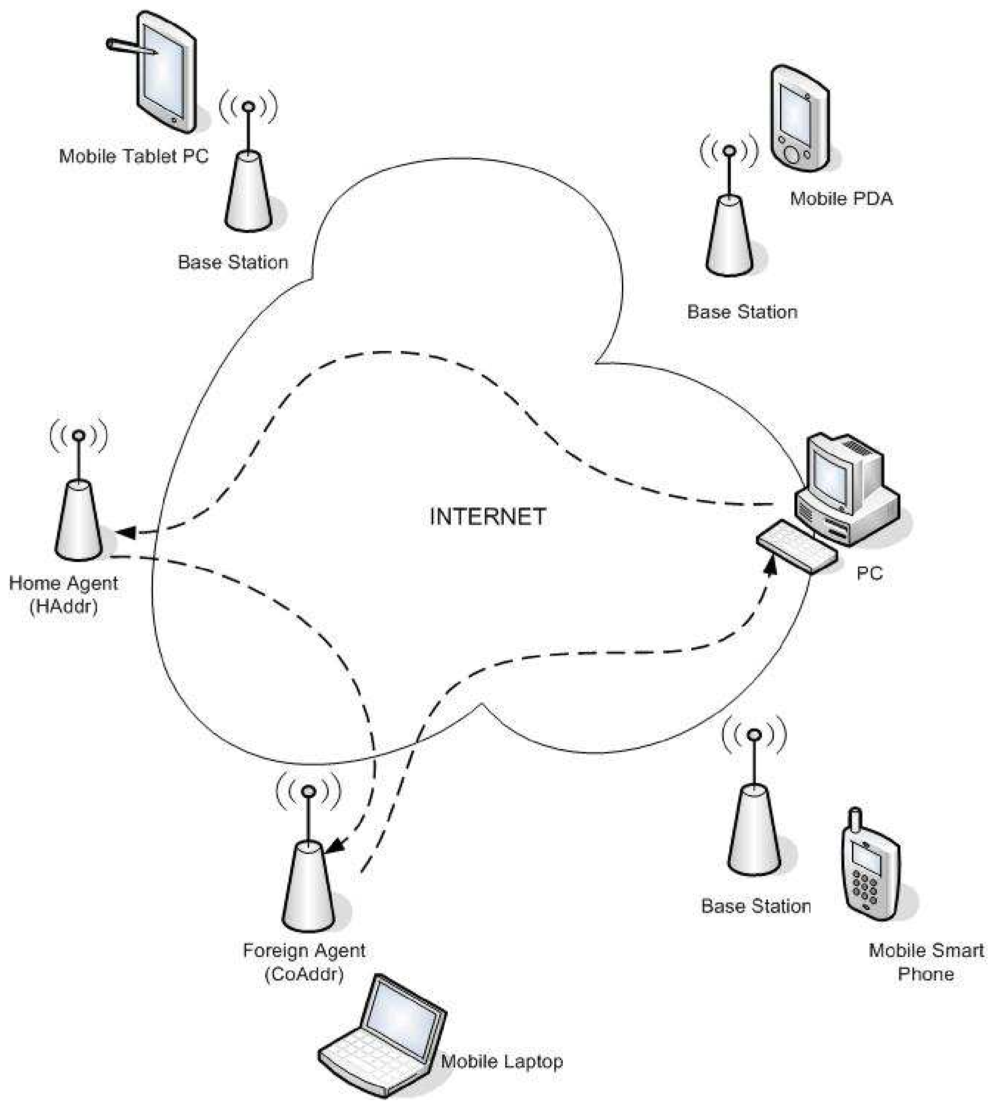

Fig. 1. MIP Network Topology

Each mobile node (MN) is assigned two IP addresses, one is the home address

(HAddr) which never changes, while the other one is the care-of address (COAddr), which is the address given by the visiting subnet to determine the actual position of the mobile node. The COAddr changes as the mobile node changes its Foreign Agent (FA) and is generally the address of the FA, e.g base station. When the MN is away from its Home Agent (HA), HA keeps track of MN and its CoAddr by means of registration procedure. HA forwards the data traffic addressed to MN when it is away. Thus, irrespective of its position in the Internet, mobile node (MN) can communicate using its home address.

C. Problem Statement and Research Objective

As discussed in the previous section, various methodologies have been put forward for communication between mobile end-users through the wired Internet. One of them is the use of MIP networks, which will be used in this work. In these kinds of networks, a base station acts as a gateway between wireless mobile nodes and the Internet infrastructure. The previous section gives detailed information about the functionality of a MIP network. In this work, two mobile nodes will be assumed communicating through a wired network (Internet). The network topology is as shown in Figure 2 and Figure 3. The main thrust of this research is to study the impact of wireless losses on the accuracy of the predictors. Hence in this work, both the nodes are considered stationary while they are communicating.

The main objectives of this research are to design multi-step predictors that can predict the moving average accumulation in MIP networks for a certain prediction horizon and study the impact of wireless losses on the overall performance of the predictors.

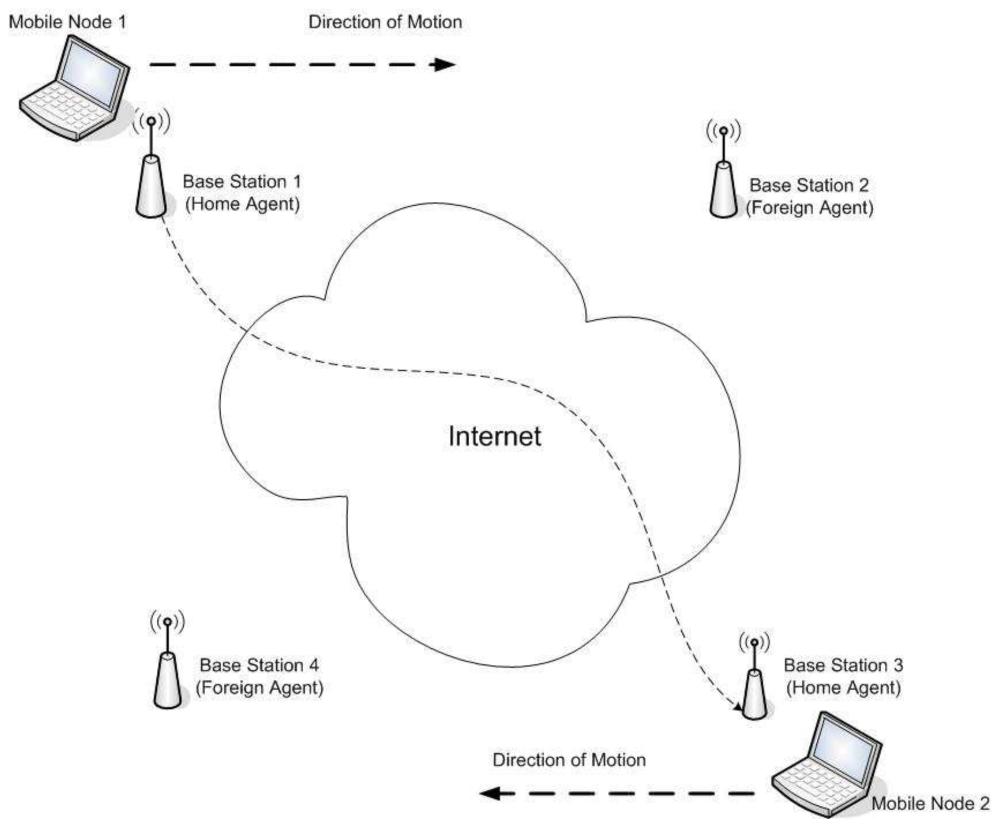

Fig. 2. Initial Position of Mobile Nodes

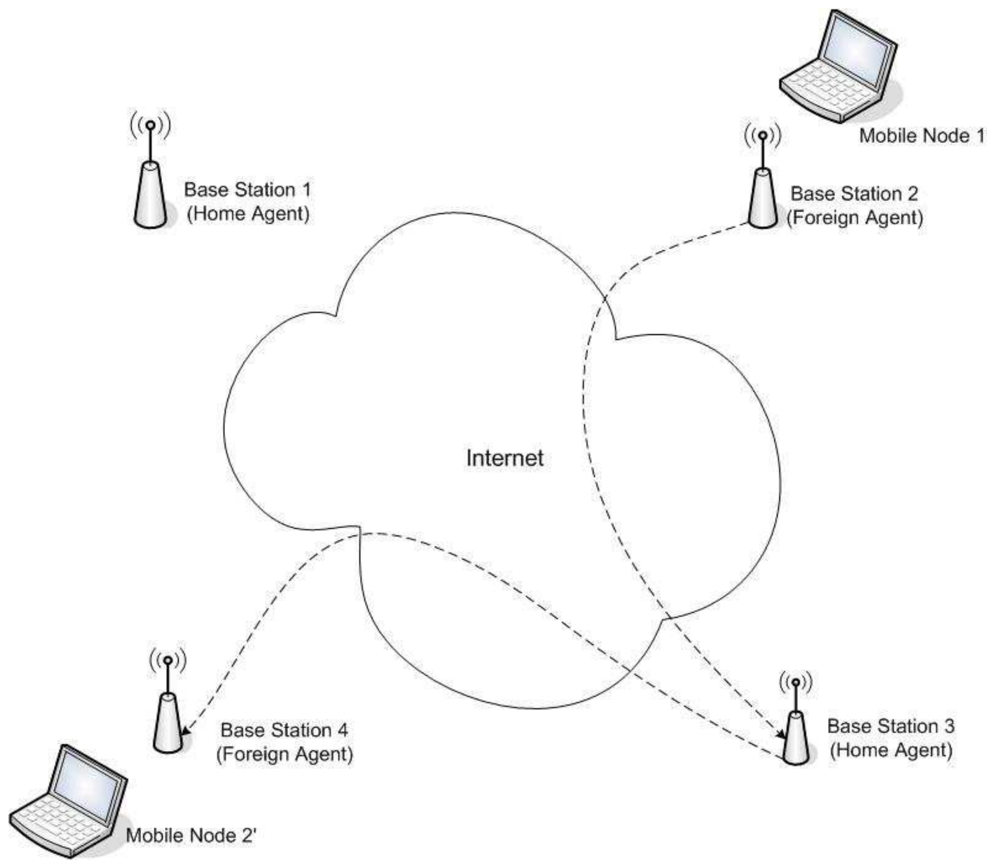

Fig. 3. Final Position of Mobile Nodes

D. Literature Review

1. Review of QoS in Wired Media Applications

Extensive research literature exists on improving QoS in multimedia applications for wired networks. Many applications such as the Web browser and FTP require reliability and "in order" packet delivery. Hence such application use Transmission Control Protocol (TCP) as their transport protocol. Real-time applications are time sensitive and hence use User Datagram Protocol (UDP). UDP does not use any congestion control algorithm nor does it employ retransmission, and hence it is not reliable. Sally Floyd et al. [7] talk about the disadvantages of using unresponsive flows and propose the development of new protocols for best-effort networks. Congestion control and fairness are the key issues in case of multimedia applications over best-effort networks. Some researchers have acknowledged this and suggested adaptive control schemes [8] which promote "TCP-friendly" behavior in order to accomplish their objectives.

Some researchers have attempted to develop TCP-friendly protocols using congestion control algorithms [9, 10]. Implementing such protocols smoothens the bit streams and eliminates use of large buffers. A technique called HPF (Heterogeneous Packet Flows) has also been proposed [11], which provides "in order" delivery.

Methods have been developed for best-effort transmission of audio and video using UDP as the transport protocol [12, 13]. Smith et al. [14] talk about cyclic UDP which uses the notion of rounds and prioritizes packets within rounds. This approach is very useful in case of stored media. Some researchers have tried to use video encoder parameters to develop feedback control [15, 16]. Most of these are receiver-based and control action is taken based on the available bandwidth resources. Some papers talk about sender-based controls, based on queue length and other network parameters [17, 18]. But these control approaches have not been developed for best-effort networks, rather for ATM networks. Sun et al. [19] use prediction of packet loss probability and round-trip-time to develop control schemes. These schemes control source flow rate.

Recently researchers are more inclined towards building systems and architectures that are end-to-end rather than network-centric [20, 21, 22]. The advantage of this method is that analysis and controllers developed for one system can be used by any other system. Yang et al. [23] give an overview of the ways researchers are modeling Internet dynamics and predict the end-to-end delays with emphasis on using system identification techniques. The paper also provides the connection between a control engineer's view-point and a statistician's view-point of addressing the problem.

Ohsaki et al. [24] have revealed the importance of end-to-end delays and the impact on QoS and congestion control. They have used a black-box approach for modeling best-effort networks. This approach is useful for time sensitive applications such as multimedia applications. The above mentioned issues demonstrate the importance of developing methods to accurately measure and predict end-to-end network flow characteristics [25].

2. Review of Wireless Mobile Applications

Internet protocols were not developed for real-time applications nor were they developed for mobile end-users. Recently a lot of research is being conducted on mobile networking which aims at providing real-time services for mobile end-users. Badrinath et al. [26] have put forward the concept of Internet Cellular Phones which is an IP-based backbone network having the capability of delivering packetized voice to moving end-users. Langendoen et al. [27] talk about QoS negotiation frameworks for multimedia applications with mobile end-users.

Of late, network architectures have been put forward that employ multiple wireless hops in route to and from the wired Internet. Gambiroza et al. [28] studied the fairness and end-to-end performance in such multihop wireless backhaul networks. If there are too many users present in a cell, content may experience large delays and QoS may degrade. A dynamic channel allocation scheme assigns more channels to such cells [29].

Bhargava et al. [30] put forward an approach integrating mobile ad-hoc networks with cellular networks. This approach increases the network security as well as throughput.

E. Proposed Approach

In this work traffic data pertaining to wireless losses is obtained from the widely used Network Simulator (NS). Whereas, data pertaining wired network is aggregated using Planet-Lab. Tool UPBAT developed by Yeom is used to generate packet traffic between NIML (TAMU) and different nodes present in Planet-Lab, Princeton. Model replicating real life MIP network is developed by imposing the simulated wireless losses on the real traffic trace. This model is used as input and accumulations are predicted. The accumulation signal has long term dependency which makes it very difficult to predict. Hence, moving average of the signal is calculated. The main purpose of using a moving average is that the large noise content of the raw signal is smoothened out to a certain extent.

The use of linear methods such as Auto-Regressive Exogenous (ARX) in modeling the accumulation in MIP networks has been demonstrated in this thesis. These predictors are used for single and multi-step prediction. In this work, impact of the wireless losses on the accuracy of the predictors developed is studied.

F. Contribution

It can be seen from the literature that the problem of predicting the accumulation in case of MIP networks hasn't been addressed. This is because of the long-term dependencies involved and the unpredictable channel fading errors. This work would be considered as contribution in following ways:

1. Development of single-step and multi-step predictors that can predict the moving average flow accumulation in MIP networks.

2. Study the impact of wireless (fading) losses on the ability to predict end-to-end flow accumulation.

G. Organization of Thesis

Chapter II presents a qualitative discussion on various parameters those can affect QoS in MIP Networks. Chapter III outlines the linear system identification techniques and methods for modeling network accumulation. Measurement and analysis of end-to-end accumulation for MIP networks are presented in Chapter IV. Chapter V demonstrates the development of single-step and multi-step predictors for predicting network accumulation. Impact of wireless losses on the predictors developed is discussed in Chapter VI. Thesis summary and conclusions are presented in Chapter VII. Recommendations for future work is also included in the chapter.

CHAPTER II

OVERVIEW OF QOS IN MOBILE IP NETWORKS

Increasingly, networks are being used as our primary source of information exchange. This implies that users would prefer to have access to this network information even when they are moving. Hence now a days mobility is also an issue while designing networks, network protocols and information services.

Currently researchers are dealing with problem of end-to-end QoS between sender and receiver having static path. But with mobile end-users, the network path might not be static anymore. As the user keeps on moving the end-to-end network path keeps on changing, which will change the end-to-end dynamics.

There are three factors on which QoS in MIP networks depend. One of them is the unpredictable losses occurring in the wireless channel, the second is the congestion level and delay variation in the wired network and the third one is the losses occurring due to congestion in the wired network. The following sections discuss all theses factors in detail.

A. Nature of Losses in Wireless Channel

Key obstacle in offering desired QoS in case of MIP network is the unpredictable losses occurring in the wireless channel. Studying the impact of these losses on the ability to provide end-to-end QoS is the main objective of this work.

In case of wireless mobile communication, packets are not only lost to congestion but also due to errors introduced in the wireless channel (Channel Fading). Performance of wireless channel depends on diffractions around a corner, line of sight (LOS) radiation, reflections from a smooth surface, and scattering caused by an object with dimensions on the order of the wavelength. Hence, estimating the impact of such random losses (on the ability to predict network congestion) is very difficult. These errors do not follow any particular model as in case of wired one, in which the queue-size gives the estimate of occurrence of losses. These are totally random errors and mainly dependent on the surrounding environment and the signals present in it.

Due to this problem, it becomes more difficult to predict the accumulation in MIP networks. Some of the major causes of errors that result in packet loss are:

1. Doppler Shift

Doppler shift occurs due to the relative motion between mobile user and base-station. This causes a frequency shift of the transmitted signal. This frequency shift makes it difficult for base-station to successfully receive the signal without any error. The same effect can be observed as sound pitch varies between a stationary observer and a moving sound source or vice versa.

Doppler effect causes the apparent frequency of the signal to be different than that of the transmitted frequency, if there is a relative motion between the source and the receiver. When there is no relative motion between the source and the receiver the received signal has the same frequency as the transmitted one. When the distance between the source and the receiver increases frequency of the received signal is lower than that of the transmitted. Whereas, if the distance is decreasing the frequency of the received signal will be higher than that of the transmitted one.

In MIP networks as the mobile end-users keep on moving relative to the stationary base-station, the frequency of the received signal might keep on fluctuating depending on the speed of motion and direction of motion. Due to this reason, it is difficult for the base-station to receive signal without any error.

2. Attenuation

Electromagnetic intensity decreases as the distance increases. This results in low signal-to-noise ratio. This problem also exists in case of wired cable. Hence, when it is necessary to transmit signals over large distance repeaters are inserted which increases the strength of the transmitted signal. The distance between repeaters is determined so as to minimize errors due to attenuation.But, in case of wireless transmission one can not have such kind of repeaters, since there might be multiple signals present in the wireless channel.

Higher the frequency of the signal, greater data capability it has. On the other hand, higher frequency signals are more susceptible to attenuation than the lower frequency ones. Generally, line of sight is desired in case of higher frequency signals. Since physical obstructions reduce the signal amplitude to certain extent.

Rain attenuation is the term used to describe the effect of precipitation on the magnitude of the signal. Apart from rain, snow, hail and sleet have a considerable impact on the signal attenuation. Even in the case of a purest day, signal suffers from some degree of attenuation. The only case where signal will not get attenuated is if the signal is traveling in vacuum.

Attenuation is generally expressed in terms of attenuation coefficient, which is defined as the rate of diminution of average power with respect to distance along a transmission path. i.e.

Attenuation Coefficient is given by:

α = power lost per unit length/(2 x power transmitted)

It is generally expressed in dB/mile and attenuation is expressed in dB.

The fading effects due to attenuation is known as large-scale fading.

3. Multipath Fading

Electromagnetic waves get reflected, refracted and scattered from different objects present in the wireless channel. This makes the signal to travel over multiple paths. This can cause fluctuations in received signal's amplitude, phase and angle of arrival. Usually if the wireless medium characteristics are unknown, one assumes it to be free space. It is assumed that there are no objects that can absorb or reflect radio frequency energy that is being transmitted between the mobile-user and the base-station. But in practice it is very rarely realized.

There are three main reasons for Multipath Fading:

1. Reflection: When an electromagnetic wave hits on a smooth surface of large dimensions compared to its wavelength reflection occurs. This may interfere constructively or destructively at the base-station.

2. Diffraction: When electromagnetic wave hits an impenetrable body of large dimension, it gets diffracted. This is also called as shadowing, because the diffracted signal reaches the base-station even if it is shadowed by impenetrable objects.

3. Scattering: When the wireless channel contains objects of dimension on the order of wavelength of the electromagnetic wave, it gets scattered. These kinds of objects deflect the energy from the mobile-user to be transmitted in many directions.

Figure 4 shows phenomenon of reflection, refraction and scattering in urban area. If there is a clear line of sight (LoS) between the mobile-user and base-station the above three mechanisms have minimal impact on the strength of the signal received at the base-station. But if there is no clear LoS then the strength of the signal received

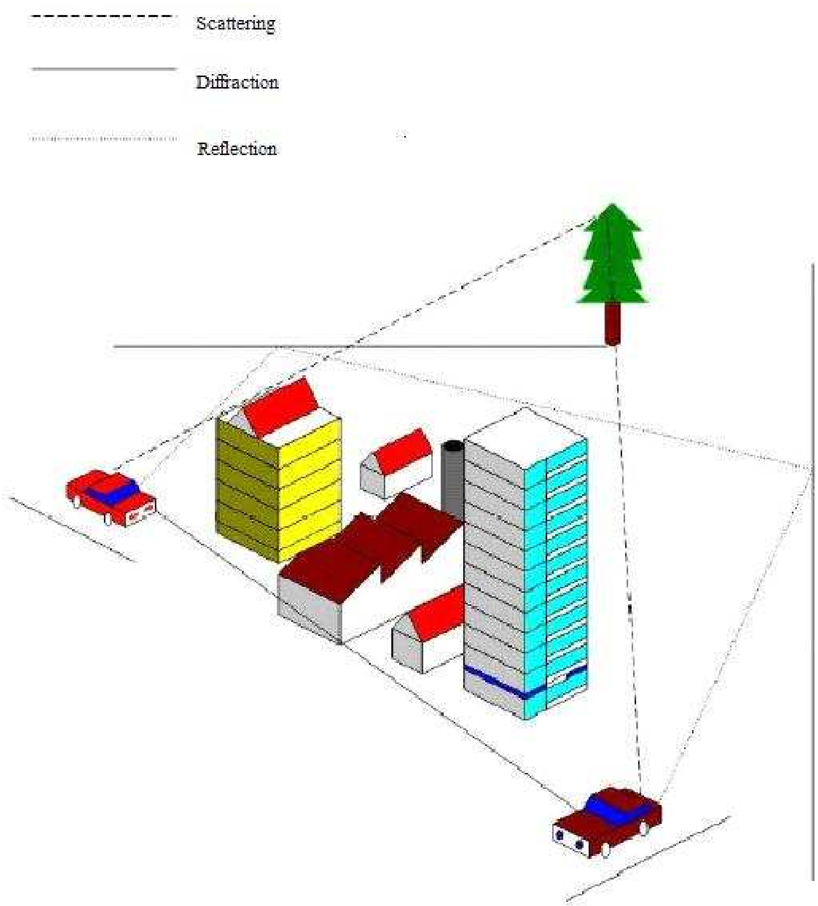

Fig. 4. Reflection Refraction Scattering

will be the sum of the components due to all three of the above. The strength of the signal received depends largely on the relative displacement between mobile-user and base-station, and also on the displacement of other objects present in the medium.

In case of mobile end-users these three factors makes a huge impact on the quality of the received signal.

The fading effects due to multi-path fading is called as small-scale fading.

B. QoS impacted by Wired Network Conditions

QoS in MIP networks not only depend on the wireless losses, but also on the network congestion in the wired Internet. Delay variations (delay jitter) and packet drops are most important parameters impacting QoS for time sensitive real-time applications over the wired Internet. Data sent over a network is split into smaller segments called data packets. These packets are then transmitted to their destination by routers present over different paths. The time taken by a packet to reach the destination application layer, once it leaves the source application layer is called as "delay" of the packet. Variation in the delay between consecutive packets is called as "delay jitter". Various types of delays and the causes of delay variation and packet drops are discussed below.

1. Causes of Delays in Network

End-to-end delay comprises of two components, one is constant while the other one is variable and keeps on changing depending on the network dynamics. The constant part includes delay at the nodes and propagation delay of the links through which the packet travels before reaching the destination. Variable delay includes queueing delay which keeps on changing depending on the cross-traffic and processing time. Following are some of the important factors contributing end-to-end delays [31]

1. Transmission Delay: It is the time taken to transmit a packet into the link. It can also be stated as the time required by a node to push a packet onto the link. For mobile as well as wired networks, it is generally of the order of few microseconds.

2. Propagation Delay: It is the time required by a bit to travel from the beginning to the end of the link. Speed with which a bit travels through a link is called as

propagation speed of the link. It is dependent on the capacity of the physical medium. Mathematically it can be defined as distance between the two nodes divided by the propagation speed of the medium. In case of wired networks it is of the order of milliseconds while in case of wireless channel it is of the order of microseconds.

3. Processing Delay: It is the time required to process the packet to check for bit errors and determine where to route it examining it's header. Generally it is of the order of microseconds.

4. Queueing Delay: It is the time a packet waits in queue before being transmitted onto the link. It depends on the number of packets already present in the queue. This delay varies from few milliseconds to hundreds of milliseconds depending on the cross-traffic present in the network.

2. Causes of Delay Variation in Network

There are several reasons which lead to delay variation known as delay jitter. Some of them are [32, 33]

1. Queues: It is the most prominent cause of delay jitter. If the flow rate of packets onto the link is more than bandwidth of the link, then packets are queued. Queues are built up at routers and switches. Generally, delay jitter happens if there is more than one flow competing for the bandwidth of the link. If two consecutive packets of a flow undergo different waiting time in a queue it leads to delay jitter. If variation in cross-traffic is large and the route to destination is multi-hop, it leads to large delay jitter, since this effect is cumulative.

2. Faulty Clocks: If clocks of transmitting node and the receiving node are not synchronized it leads to time drifts and clock skews. Since it has an impact on the observed state of the network, it is extremely important to deal with clock accuracy.

3. Bursty Traffic: If capacity of the link is less than the incoming burst, then queue starts to build up. If packets of a flow are queued in such a queue, each packet will experience different queueing delay depending on its position in the queue. This leads to a large delay jitter. If there are multiple bursty flows competing for limited bandwidth the jitter can be severe.

4. Route Changes: It happens due to router failures, change in routing algorithm, and sudden recovery of failed routers. Change in route because of any of the above mentioned reasons will lead to jitter, since propagation delays and queueing delays for different routes might be different.

5. Packet Reordering: A packet is reordered if its sequence number is less than any of the packet arrived before it. Reordering of packets happen due to retransmission algorithm present in some reliable protocols like TCP, or due to route changes, or due to multiple buffers at routers.

3. Causes of Packet Loss in Wired Network

In case of wired networks, packets are lost if the queues at the router are full. Routers have finite queue, when the incoming packet comes after the queue has reached its saturation the router drops the packet. Generally, a router has two buffers, the incoming interface buffer and the outgoing interface buffer. If a router is not able to process the incoming packet fast enough, it is dropped at the incoming interface buffer. If the outgoing link of a router is very busy, the packet is dropped at the outgoing interface buffer. Sometimes packets are dropped due to transmission errors, and if the checksum on the packet fails.

C. Chapter Summary

Ensuring QoS in MIP networks is more difficult than wired networks because of the random channel fading errors in the wireless channel. QoS in case of mobile end-users is not only affected by random channel fading losses, but also due to delay jitter and packet losses experienced by the packets in the wired network. This indicates that there is a need for predictive estimation of end-to-end characteristics of MIP networks in order to guarantee a certain level of QoS for mobile end-users.

CHAPTER III

LINEAR EMPIRICAL MODELING TECHNIQUES

A. Introduction

The problem addressed in this research is multi-step-ahead prediction of the moving average of accumulation signal for MIP networks. Prediction can be defined as estimation of a variable of interest at a future point in time given measured data up until and including present time [34]. Two types of models are used in system modeling, first is the physical model whereas second one is empirical model. In physical models, mathematical models are used to describe the relationship between the input system variables and output system variables. Empirical models are derived from the observed data of the system. Empirical models are also known as "black-box" models. The system models developed in this research are empirical.

System identification develops empirical models based on observed data of the system under consideration. System identification is used to develop models in the fields of medicine, process control, computer engineering, and business. As the dynamics of a system becomes increasingly complex and uncertain its mathematical analysis becomes more difficult. In such cases system identification is necessary and very useful for modeling the system. For developing an efficient controller with desired performance, an accurate model representing the system is necessary. Thus, accurately modeling the system is very crucial step of a control problem. The following section gives some information about system identification procedures.

B. System Identification Procedure

Figure 5 shows the sequential steps involved in System Identification (SI). Following are the steps involved in system identification [34]:

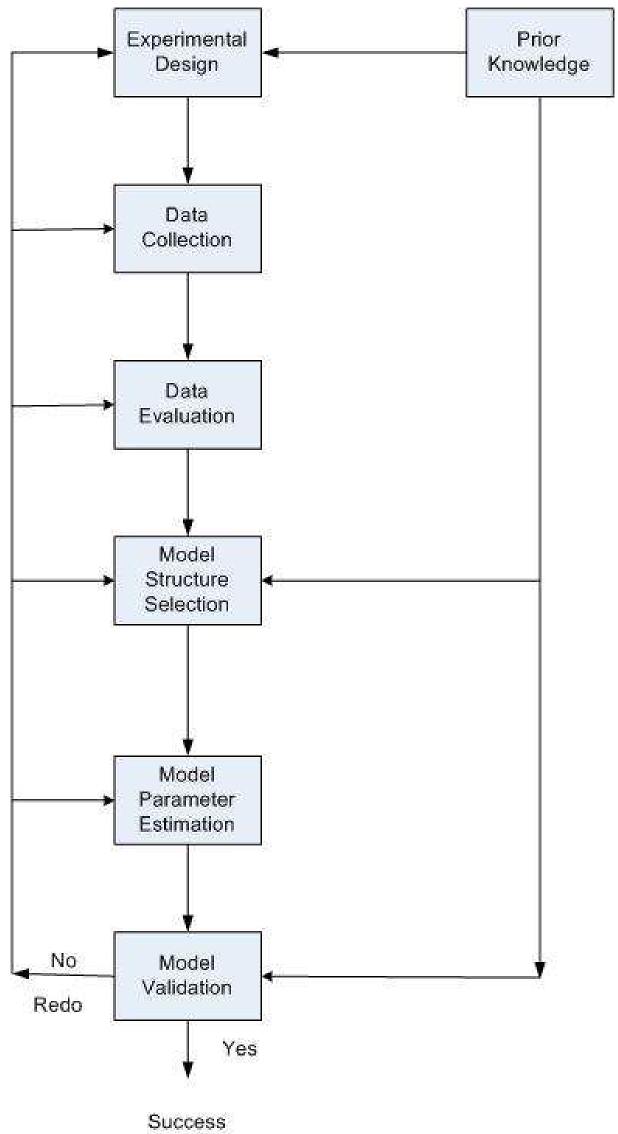

Fig. 5. System Identification Process.

1. Experiment design and data collection: For developing a good model, it is very important to capture the dynamics of the system. Experiments should be conducted in such a way that the input signal must replicate the important dynamics of the system. Hence, the data collected should be analyzed rigorously. Correlation tests should be done to make sure that the data is suitable for system identification. If the data is not replicating the dynamics of the system or not suitable for system identification, experiments should be performed again. Hence, this step is the most important step in system identification process.

2. Model structure selection: After desired data has been collected, the next step is to select an appropriate model. A prior knowledge of the system is very useful for selecting a good model structure. Linear models can be classified into input-output and state-space models. In input-output models, linear regression is used to represent the relation between system inputs and outputs. On the other hand, in state-space models intermediate states are used to represent the system. In this work, input-output models are used for modeling.

3. Model parameter estimation: The third step is to solve an optimization problem using the data collected in the first step and the model selected in the second step. Model parameters are numerical values obtained in this optimization problem which are used to describe the model selected in the previous steps.

4. Model validation: The model developed in the previous three steps must be tested to work in the operating range of interest. For validating a model, completely fresh data are used which are not used in any of the previous steps and the performance of the model is monitored on this data set. If this data set invalidates the model, all of the above steps must be repeated to develop a new model.

Thus mathematically, system identification can be state as follows:

Obtain a free parameter θ of a function *F()* such that one-step-ahead prediction is given by:

$$\hat{y}(t|t-1) = F(u, Y; \theta) \tag{3.1}$$

where, $u = u(1), ..., u(T)$ are finite set of input observations,

$Y = y(1), ..., y(T)$ are the corresponding output observations,

and $\hat{y}(.|..)$ are the past predictions values.

C. Linear System Identification Technique

Linear estimation is based on the assumption that the system being analyzed can be represented as a linear model. Linear models can be classified into input-output systems and state-space system. In input-output systems linear regression is used to present the relation between system inputs and outputs. On the other hand, in state-space models intermediate states are used to represent the system. For modeling, in this work input-output model is used. This is because the accumulation signal is in the form of a time series. Also the system being modeled is single input single output (SISO). The section below gives outline of the model used in this work.

1. Auto-Regressive Exogenous (ARX) Model Structure

This is the simplest of all the linear system modeling techniques. The AR part in ARX denotes the Auto-regressive part while the X part denotes the extra input called exogenous variable. Single-Input Single-Output (SISO) ARX model can be represented by the following equation:

$$\begin{aligned} y(t+1) = & a_1 y(t) + ... + a_{n_y} y(t-n_y+1) \\ & + b_1 u(t-n_d) + ... + b_{n_u} u(t-n_u-n_d+1) + e(t+1) \end{aligned} \tag{3.2}$$

where y(t) is the output of the ARX model, n_y is the number of past outputs commonly known as the lag terms of the model, u(t) is the input to the ARX model, n_u is the number of past input lags used in the model and n_d is the pure time delay (the dead time) in the system. It is assumed that the coefficients $a_1, ..., a_{n_y}$ and $b_1, ..., b_{n_u}$ are known.

From SISO ARX model presented by equation 3.2, single-step-prediction (SSP) of the system output $\hat{y}(t+1|t)$ can be given by following equation:

$$\begin{aligned} \hat{y}(t+1|t) = & a_1 y(t) + ... + a_{n_y} y(t-n_y+1) \\ & + b_1 u(t-n_d+1) + ... + b_{n_u} u(t-n_u-n_d+2) \end{aligned} \tag{3.3}$$

Similarly, multi-step-ahead predictor (MSP) can be written as:

$$\begin{aligned} \hat{y}(t+1|t-p+1) = & a_1 \hat{y}(t|t-p+1) + ... + a_{n_y} \hat{y}(t-n_y+1|t-p+1) \\ & + b_1 u(t-n_d+1) + ... + b_{n_u} u(t-n_u-n_d+2) \end{aligned} \tag{3.4}$$

2. Auto-Regressive Exogenous Parameter Estimation

Previous section described the predictor form of the ARX model. The parameters $a_1, ...a_{n_y}, b_1..., b_{n_u}$ of ARX model are unknown. These parameters must be determined using measurement data, Y and u. The section following discusses estimation of these parameters.

The ARX predictor can also be written as:

$$\hat{y}(t+1|t) = \varphi^T(t+1)\theta \tag{3.5}$$

where,

$\varphi(t+1) = [y(t), ..., y(t-n_y+1), u(t-n_d+1), ..., u(t-n_u-n_d+2)]^T$

$\theta = [a_1, ..., a_{n_y}, b_1, ..., b_{n_u}]^T$

Thus ARX predictor form can be written as a scalar product between the data vector $\varphi(t+1)$ and the parameter vector θ. The equation above is in the form of a linear regression with the parameter vector θ as the regression vector and hence least-squares method can be used to solve for θ [30].

To solve for the parameters of the ARX predictor using least-squares method, the mean-square of the prediction error, $V_N(\theta, Z_N)$ must be defined as:

$$V_N(\theta, Z_N) = \frac{1}{N}\sum_{t=1}^{N}[y(t) - \hat{y}(t|t-1;\theta)]^2 \tag{3.6}$$

where Z^N is the data set of N input-output samples $u(t)$ and $y(t)$ for $t = 1, ..., N$. The above equation can also be interpreted as the objective function of the least-squares problem. The objective function, $V_N(\theta; Z^N)$, is minimized with respect to θ. The solution to this least-squares problem is the value of $\hat{\theta}_N$ that minimizes $V_N(\theta; Z^N)$. This is given by:

$$\hat{\theta}_N = [\sum_{t=1}^{N}\varphi(t)\varphi^T(t)]^{-1}\sum_{t=1}^{N}\varphi(t)y(t). \tag{3.7}$$

Since no exogenous inputs are used, accumulation signal is a time series. In that case model becomes Auto-Regressive (AR) model.

D. Chapter Summary

In this chapter system identification procedure and linear methods for single-step-ahead (SSP) and multi-step-ahead (MSP) are discussed. The linear methods provide accurate linear system models. The next chapter gives an insight on the accumulation data used in this work. Next chapter also highlights the temporal properties of accumulation signal along with its moving average calculations.

CHAPTER IV

MEASUREMENT OF END-TO-END PACKET TRANSPORT: A QUALITATIVE DISCUSSION

A. Introduction

The most important step in modeling end-to-end accumulation is to obtain sufficient data replicating the dynamics of the system. This chapter discusses some important issues regarding the system of interest. A detailed discussion of various end-to-end parameters is done. The term end-to-end measurement relates to the measurements done between the application layer at the source and application layer at the destination, assuming the intermediate Internet as a 'Black-Box'. This chapter gives detailed information regarding data collection procedure used to collect simulated data and real traffic data. The following section discusses about the various possible end-to-end network measurement parameters. This chapter also gives various assumptions made in this work.

B. End-to-End Network Measurement

There are different units of measurements for end-to-end network dynamics. Following section discusses the available units of measurements and their comparison.

1. Signal Measurement at the Source

1. Send Rate: Send rate is defined as number of bytes of data sent by the source per unit time. It is measured in Kb/s

2. Send Flow: Send flow is defined as the total number of packets or bytes sent by the source into the the network at any given instance of time. It is measured in

bytes or number of packets.

2. Signal Measurement at the Destination

1. Arrival Rate: Arrival rate is defined as number of bytes of data received by the destination per unit time. It is measured in Kb/s.

2. Arrival Flow: Arrival flow is defined as the total number of packets or bytes received by the destination at any given instance of time. It is measured in bytes or number of packets.

3. End-to-End Delay: End-to-end delay is the time required by the packet to travel from the application layer of the source to the application layer of the destination. It is given in ms.

3. Packet Measurement and Cumulative Packet Loss

In this work, network packet accumulation is used as an indicator of the end-to-end dynamics of the network. "Accumulation and loss" of a flow is the difference between cumulative send and cumulative arrival flows. Mathematically, Accumulation and losses can be expressed as:

$$AccL(k) = U(k) - Y(k), \tag{4.1}$$

where $Y(k)$ is the arrival flow, $U(k)$ is the send flow, $AccL(k)$ is the accumulation and loss function and k is the discrete time step. The accumulation and loss function has two parts, first is the accumulation and the second one is the cumulative packet loss in the network. Thus, accumulation loss function is:

$$AccL(k) = Acc(k) + L(k), \tag{4.2}$$

where $Acc(k)$ is the true packet accumulation and $L(k)$ is the cumulative packet loss of the flow under consideration at any give time step k. The true accumulation signal can be obtained by removing the trend from the accumulation and loss signal. From now on, the term packet accumulation will denote the true packet accumulation in this work.

4. End-to-End Delay Versus Packet Accumulation

In this work network accumulation is used as an indicator of end-to-end dynamics of the network instead of end-to-end delays because of the following reasons:

1. It is very difficult to distinguish between the packets with very large delays and the packets those are dropped in the network.

2. It is difficult to assign a delay value to the dropped packet.

3. Accumulation signal can be used in case of flow reversal when the packets arrive out of order. On the other hand it is very difficult to account for out of order packet arrival in case of end-to-end delay signal.

C. Major Assumptions

The major assumptions made in this work are as follows:

- The mobile node does not move in a random fashion.

- The fraction of flow under consideration is very less compared to the cross-traffic.

- Losses in wireless channel are Physical Layer losses and there are no Mac layer (collision) losses.

D. Data Collection

As MIP networks are yet to be implemented, there is no data available as of now. Hence, in this work data pertaining to losses in the wireless channel is obtained using popular network simulator (ns-2). Where as, tool UPBAT developed by Yeom is used to generate packet traffic between NIML (TAMU) and different nodes present in Planet-Lab, Princeton. The data generated by the tools discussed above are called as real traffic data. After getting both the data, the wireless losses obtained from ns-2 simulations are imposed on the real data to get the effects of MIP networks. The following sections gives the data collection procedure adopted in detail.

1. Generation of Simulated Wireless Data

This section deals with generation of the wireless loss data using ns-2. The data generated using ns-2 will now be called as simulated data. As described in Chapter II, there are two types of fading effects:

1. Small scale fading: It is caused due to multi-path fading.

2. Large scale fading: It is caused due to signal attenuation.

Shadowing model simulates large scale fading effects in NS-2. But, shadowing model can only be used in case of mobile ad-hoc networks and not in MIP networks.Hence in this work shadowing model is used to get the realistic range of large scale fading effects. Mathematically shadowing model is given by the following equation:

$$[\frac{Pr(d)}{Pr(d_0)}]_{dB} = -10\beta \log(\frac{d}{d_0}) + X_{dB} \tag{4.3}$$

where, $Pr(d_0)$ represents mean power received at a distance d_0, β is called as path loss exponent, and X_{dB} is the Gaussian random variable with zero mean and

Table I. Values of β in various Environmental Conditions

Environment		β
Outdoor	Free Space	2
	Shadowed urban area	2.7 to 5
In Building	Line-of-Sight	1.6 to 1.8
	Obstructed	4 to 6

Table II. Values of σ_{dB} in various Environmental Conditions

Environment	σ_{dB}
Outdoor	4 to 12
Office, hard partition	7
Office, soft partition	9.6
Factory, line-of-sight	3 to 6
Factory, obstructed	6.8

standard deviation σ_{dB}. σ_{dB} is also known as shadowing deviation.

Values of β and σ_{dB} for different environmental conditions are given by Tables I and II [36] respectively.

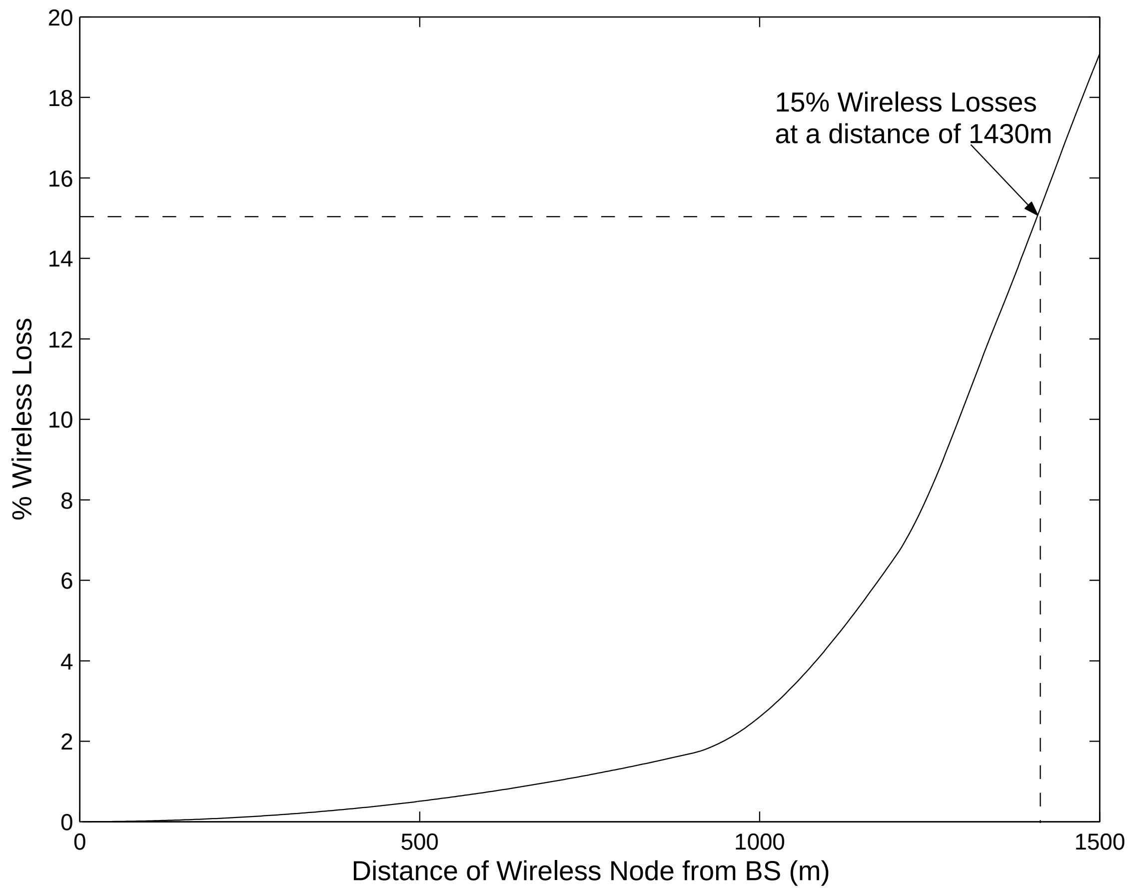

Fig. 6. Wireless loss (%) vs. Distance of Wireless Node from Base-Station

In this work, values of path loss exponent and shadowing deviation are selected for the outdoor conditions. From Figure 6, 15% wireless loss can be observed for a distance of 1430m. In the worst case, wireless losses can be considered as random losses. Hence to model the worst case scenario built in error model in NS was used to generate errors in random numbered packets.

For simulating small scale fading effects commonly used Ricean Fading model is incorporated in ns-2. This model is available online at (http://www.ece.cmu.edu/). Mathematically, Ricean model can be represented as:

$$r = \sqrt{(\sigma x_1 + A)^2 + \sigma x_2^2} \tag{4.4}$$

where, $A = 0$ since it is the line-of-sight component and we are assuming there is no clear line of sight between mobile node and base station,

$\sigma x_1, \sigma x_2$ are in-phase and quadrature phase component respectively, and

r is the magnitude of the fading envelope.

These error models, introduce errors in randomly picked packets and drop them before queuing. For generating simulated data, the flow conditions are maintained the same as in case of wired network. Which means packet size and bit rate of the packets is maintained the same as in case of the wired networks. Figure 7 shows the basic topology used to get the desired simulated data using NS-2.

The MIP network considered has the following components:

1. Mobile Nodes: As shown in the Figure 6 proposed network has two mobile nodes. Initially, the mobile nodes are considered to be stationary, so as to get fading effect due to wireless channel. To have a realistic effect of fading due to mobility after some time both nodes move at 55 mph. In this work, fading effects due to wireless channel only is considered and fading due to mobility is not considered.

2. Base Stations: As shown, the MIP network considered has four base-stations. Each Base station acts as a gateway between wireless mobile node and the Internet infrastructure network. As discussed in chapter I, base-station 1 is home-agent (HA) for mobile-node1 while base-station 3 is the HA for mobile-node 2. Thus initially, base-station 1 acts as a gateway between mobile-node 1 and the Internet. As node 1 starts moving away from its HA, the strength of the signal received by HA weakens. After some time period the mobile

node is unable to connect to any of the base stations and all of its information transmitted gets lost. When mobile node 1 comes in the vicinity of base-station (2), it (base-station 2) becomes its foreign-agent (FA) and routes its data packets through wired Internet. Mobile-node 2 follows the same mobility pattern, except when it reaches its FA (base-station 4), packets addressed to it arrives at its HA first. HA then reroutes those packets to mobile-node 2 via its FA.

3. Wired Link: NS-2 simulations are used in this work to get the data pertaining to wireless and mobile losses. Hence instead of having wired network only a wired link is used, as shown in the network topology. The wired link will be replaced by wired Internet as discussed in the following section.

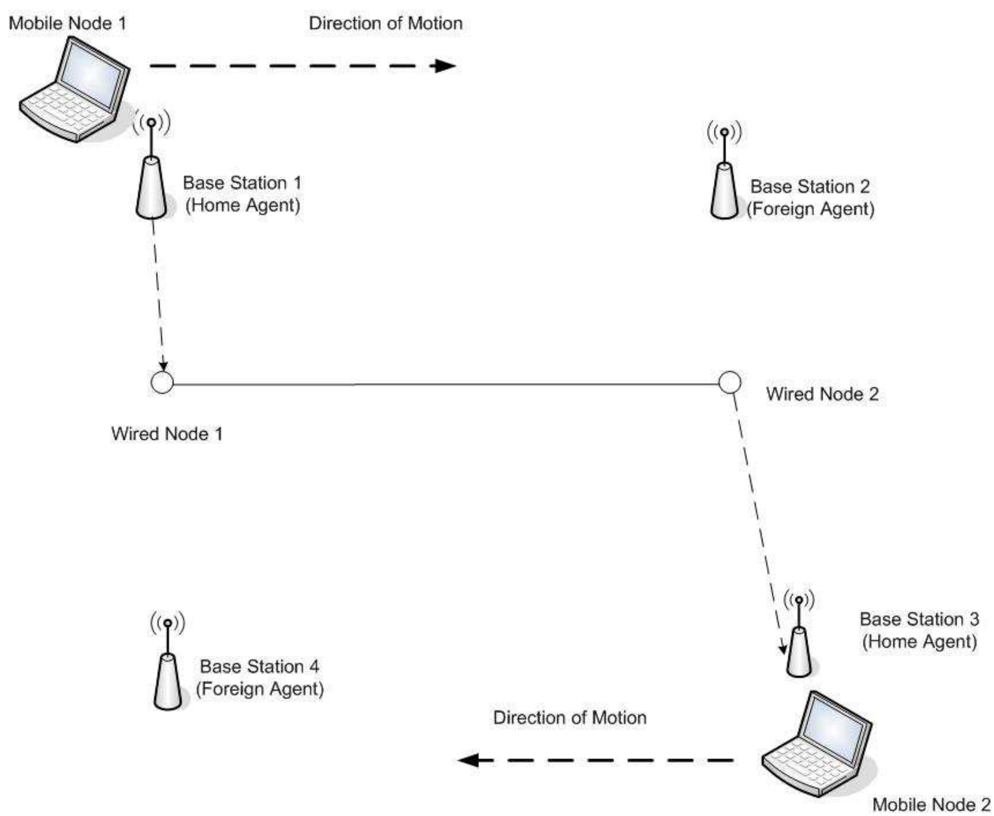

Fig. 7. Network Topology for Simulated Data

2. Collection of Wired Traffic Data

In this work Planet-lab is used for collecting real traffic trace. Planet-lab is open infrastructure for invention of next generation wide area services. It is the foundation on which next generation Internet can emerge. Its a different kind of network test-bed:

- It is not a distributed super-computer.
- It is not a collection of pipes or giga-pops
- It is geographically distributed network services.
- It provides alternative network architectures and protocols.

Planet-lab has more than 430 nodes across 201 sites worldwide which includes Universities, Labs and Internet2. It is an active and growing research community.

Tool UPBAT developed by Yeom [35] are used for generating packet traffic between NIML (TAMU) and different nodes present in Planet-Lab, Princeton. The data generated by these tools will be called as real data. Once the real data is generated, simulated data pertaining to the wireless fading losses is superimposed on the real data to get the replication of MIP networks. The data generated by superimposition will be called as MIP data. MIP data will be used for predictor development and validation. By comparing the validation results of different MIP data, impact of wireless and mobile losses can be studied. Figure 8 gives the topology used for collecting real data using UPBAT tool. The tool requires a source node to send the information packets and the receiver node to receive it.

In this research, real traffic traces between different source node and destination node have been collected. The different source-destination pairs used for data collection are:

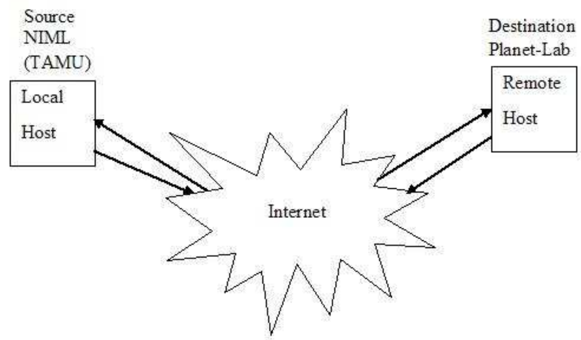

Fig. 8. Network Topology for Actual Wired Traffic Data Collection

1. Source node is NIML (TAMU) and the destination node was pli_pa node present in the planet-lab, Princeton. For this flow protocol used was UDP, packet size was 70 Bytes and bit-rate was 28Kb/s.

2. Source node is NIML (TAMU) and the destination node was gtidsl node present in the planet-lab, Princeton. For this flow protocol used was UDP, packet size was 70 Bytes and bit-rate was 28Kb/s.

3. Source node is gtidsl and the destination node was nbgisp, both these nodes are present in the planet-lab, Princeton. For this flow protocol used was UDP, packet size was 30 Bytes and bit-rate was 9.6Kb/s.

UPBAT uses two threads for sending and receiving data packets. For collecting real traffic data, server program is started on the remote host and client program is started on the local host. Packets are sent from the source host to the destination host. These packets are echoed back to the source with the destination time stamp on it. Thus this tool gives two way delay.

3. Generation of MIP Data

After generating the simulated trace pertaining to wireless, mobile part and real traffic trace, this section discuss about combining these two traces to generate the MIP trace. As discussed earlier, the flow characteristics while generating simulated data are maintained the same as the real traffic. From the simulated data, the information pertaining to the channel fading losses and the mobility losses is derived. After getting this information, those data packets are dropped from the real traffic trace. This will give the pseudo impression of real MIP communication through wired Internet. The data developed by superimposing the wireless and mobile losses on the real traffic data is used throughout this work.

E. Auto-correlation of MIP Data

Autocorrelation of different MIP traces gives us the necessary information about the nature of the data to be predicted. The rate at which the autocorrelation falls with the increase in the number of lags gives some indication about the long-term dependency of the signal. Figures 9 and 10 give autocorrelation coefficient for 20lags. From the figures it can be observed that autocorrelation function does not drop below 0.82 even after 20 lags. The study of autocorrelation function is necessary since it helps in deciding the order of linear predictive model like ARX.

From Figures 9 and 10 it can be seen that the autocorrelation of MIP trace is high even after 20 lags. This shows that the traces has long term dependency. This will make the task of multi-step-ahead prediction very difficult. If the errors in these predictions are very high, the predictor will not be suitable for further development of the source controller. This problem exists because the time series is extremely noisy. In order to smoothen the time series, moving average of the original accumu-

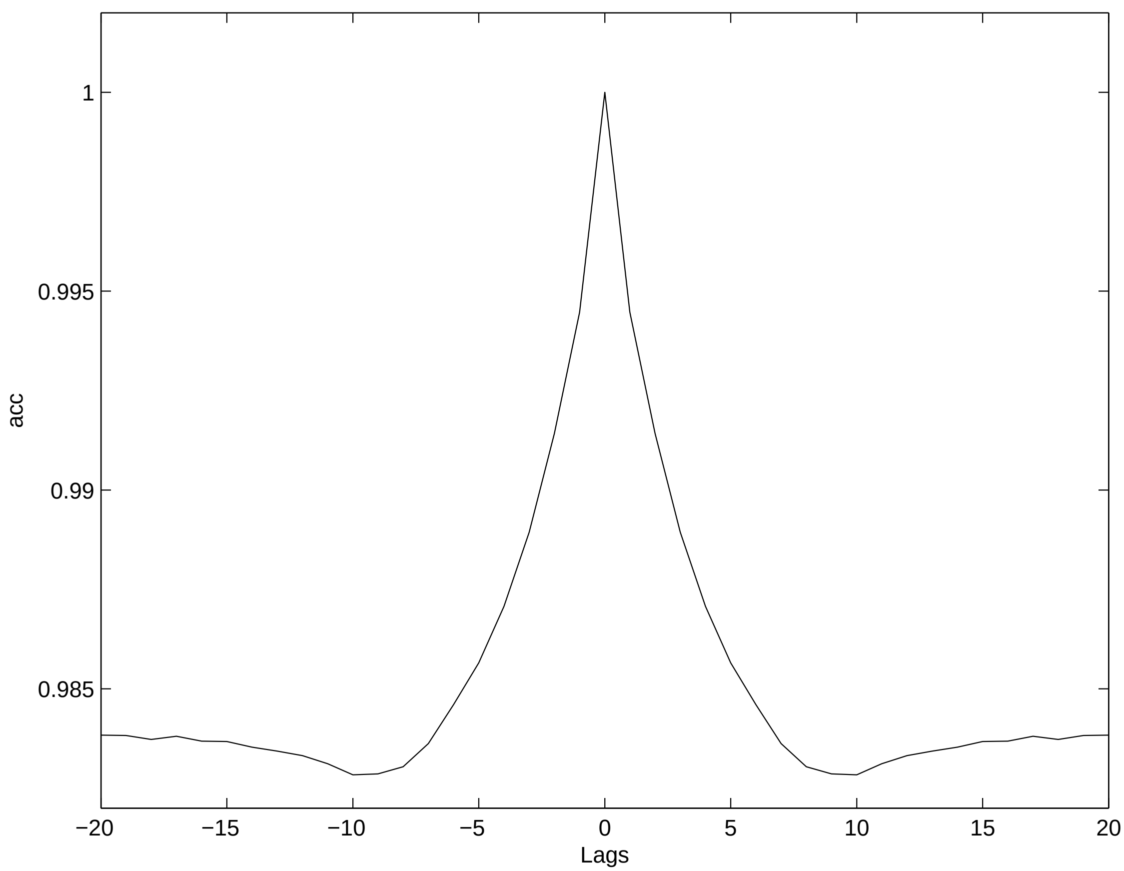

Fig. 9. Auto-Correlation Function of a Typical Accumulation Signal Between TAMU and pli_pa Node in Planet-Lab

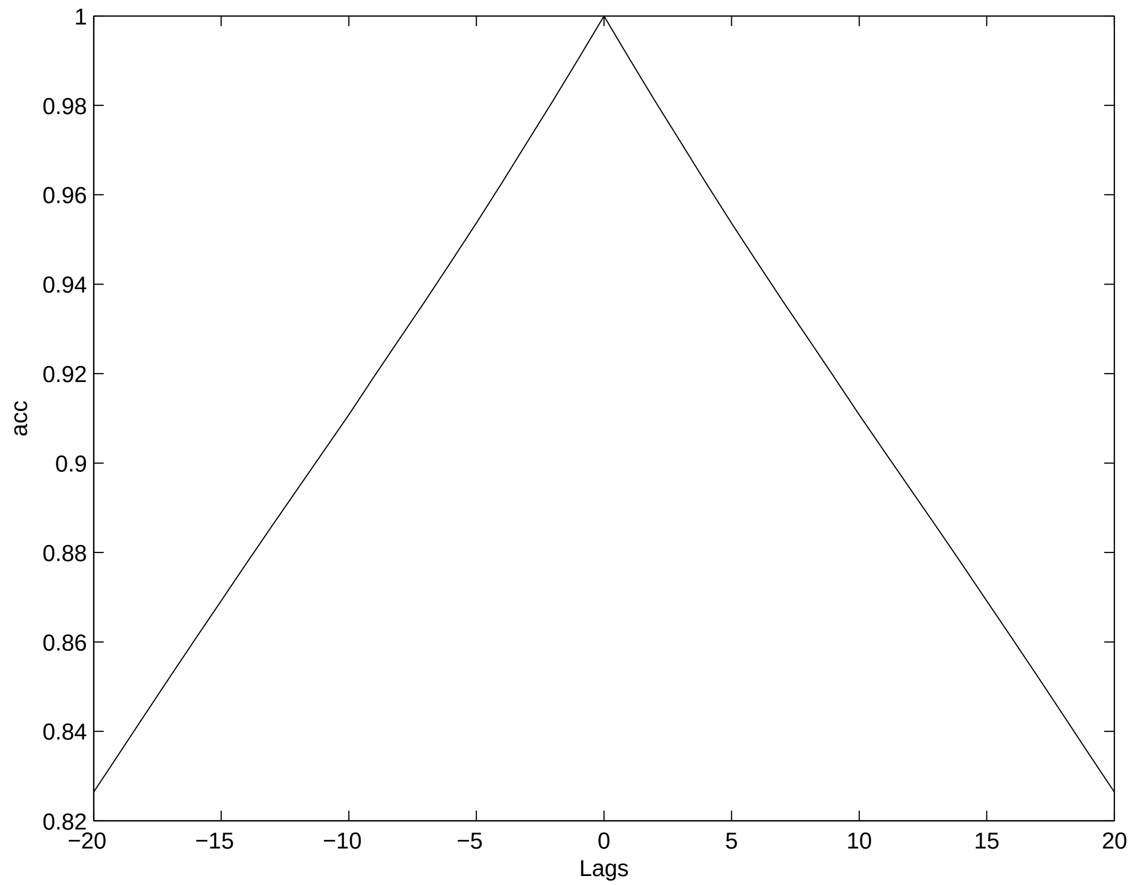

Fig. 10. Auto-Correlation Function of a Typical Accumulation Signal Between TAMU and gtidsl Node in Planet-Lab

lation time-series is taken to generate a mean accumulation time-series. Thus mean accumulation time-series can be given as:

$$Acc_{MA}(k+w-1) = \frac{1}{w}\sum_{j=k}^{k+w-1} Acc(j) \tag{4.5}$$

where $Acc_{MA}(k)$ is k-th moving average of network accumulation and m is the window of the moving average. In this work moving window of 100ms is used to get the mean accumulation time-series. Figures 11 and 12 show the autocorrelation function of the mean accumulation signal.

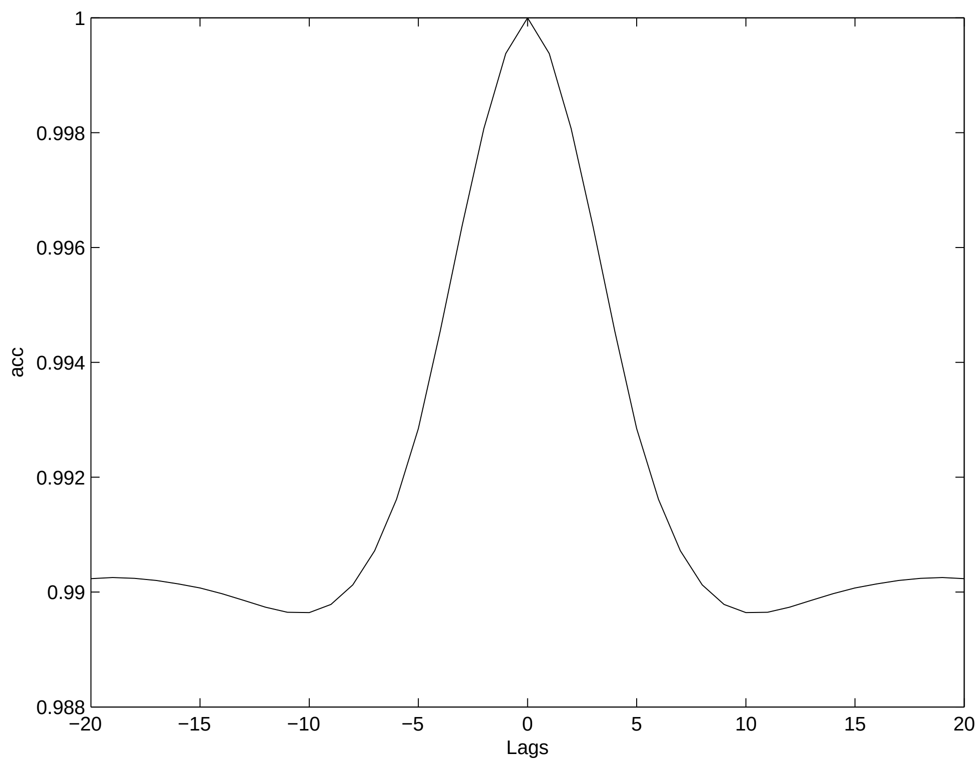

Fig. 11. Auto-Correlation Function of Moving Average Accumulation Signal Between TAMU and pli_pa Node in Planet-Lab

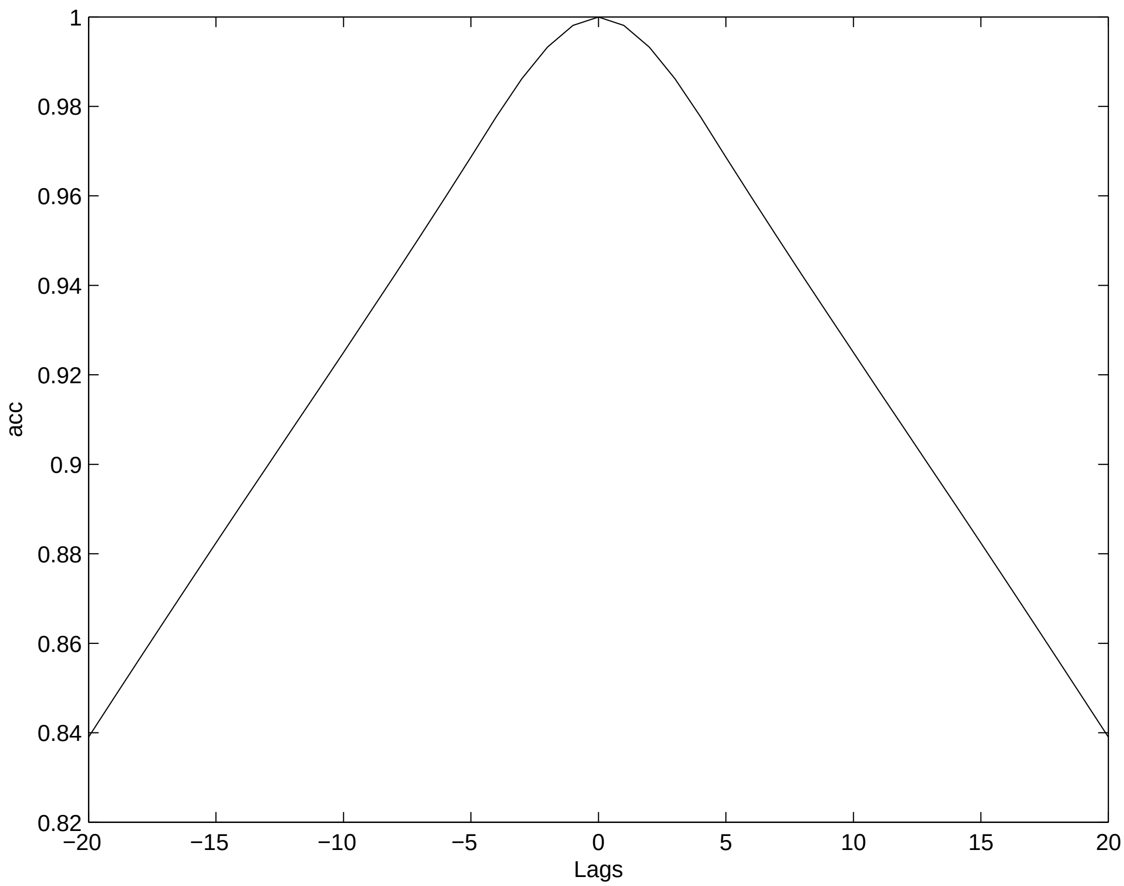

Fig. 12. Auto-Correlation Function of Moving Average Accumulation Signal Between TAMU and gtidsl Node in Planet-Lab

Figures 13 and 14 show the mean accumulation time-series along with the original accumulation time-series.The original accumulation time-series is shown in discrete points. From these plots it can be seen that the original accumulation signal is spiky in nature indicating large noise content of the signal. Moving average accumulation gets rid of these spikes and smoothens the accumulation signal to certain extent.

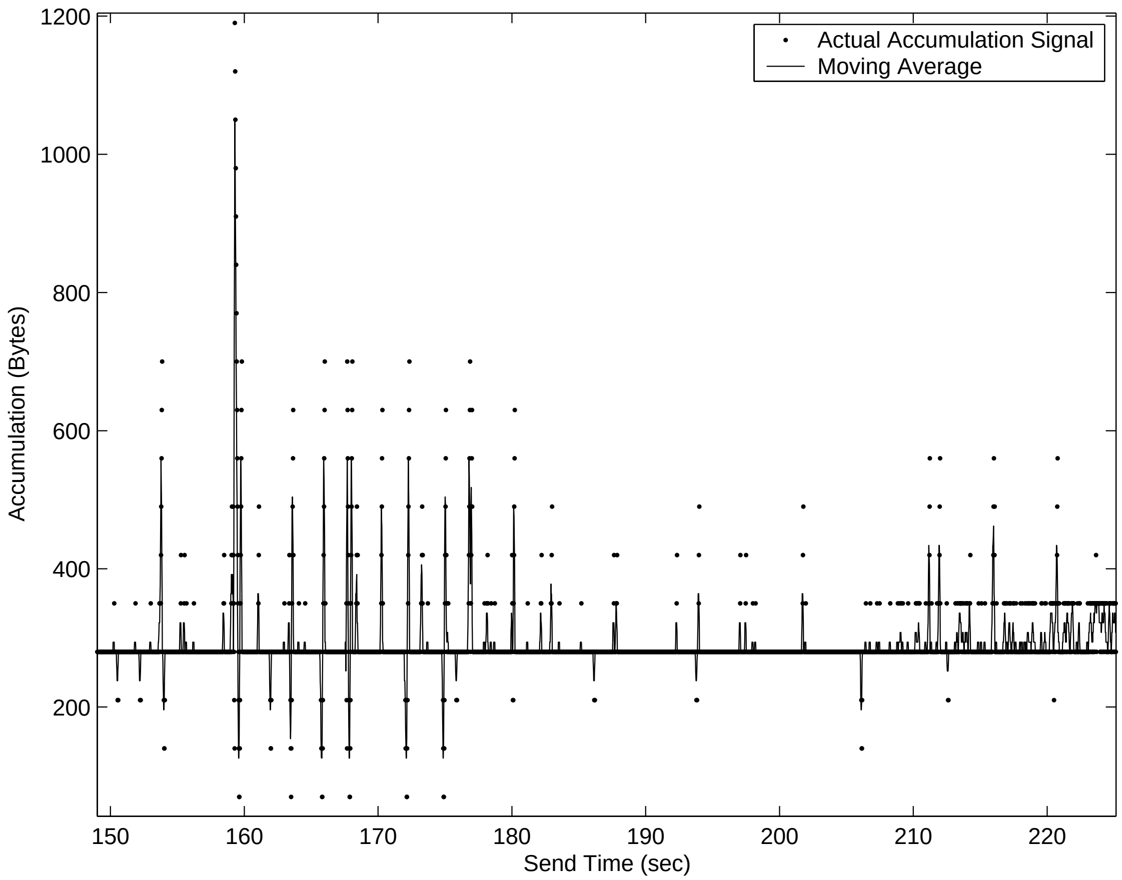

Fig. 13. Moving Average Accumulation Time Series Between TAMU and pli_pa Node in Planet-Lab

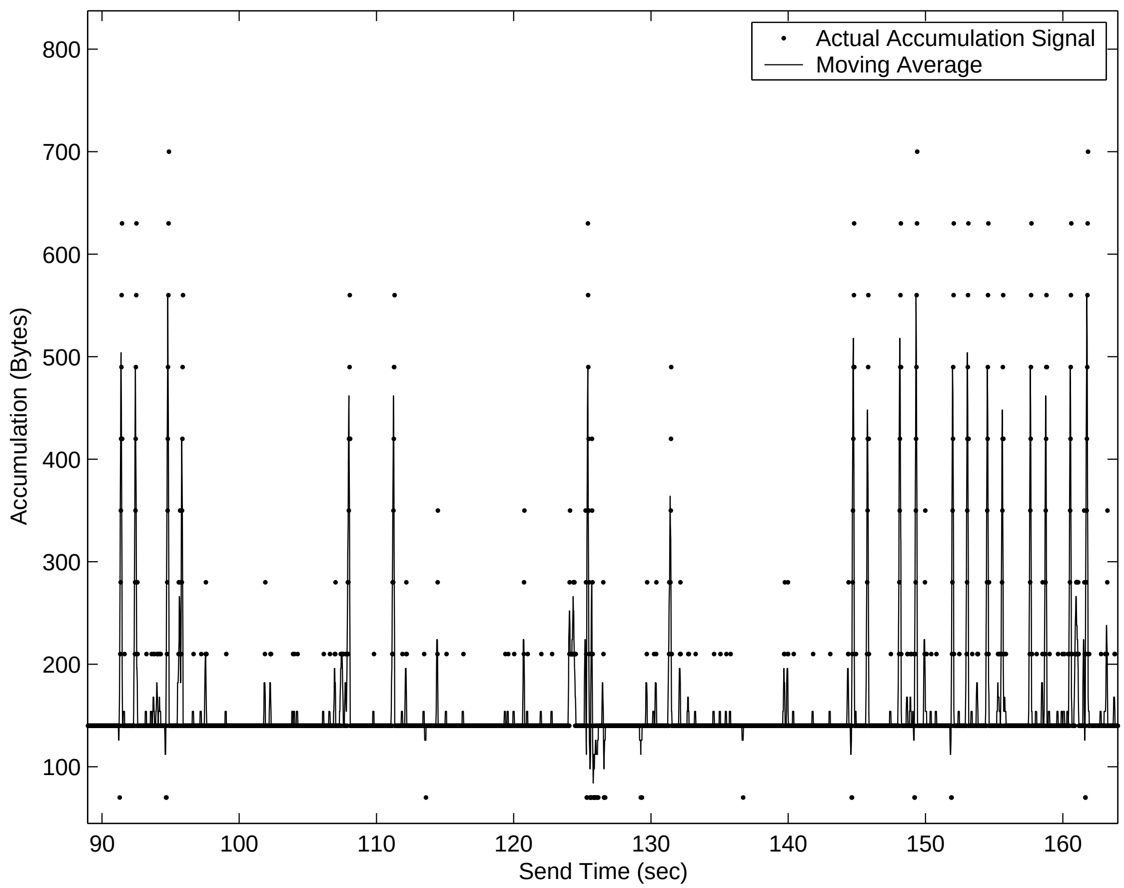

Fig. 14. Moving Average Accumulation Time Series Between TAMU and gtidsl Node in Planet-Lab

F. Chapter Summary

This chapter provides various possible units of end-to-end network measurement, which can be used as an indication for network congestion. A detailed discussion on the data collection methods is done in this chapter.

The accumulation time-series has long term dependency and is extremely noisy. Because of these reasons prediction is very difficult. Hence moving average of accumulation time series is used instead of the original accumulation time-series.

The following chapter deals with the development of predictors and their performance analysis.

CHAPTER V

SINGLE-STEP AND MULTI-STEP PREDICTOR DEVELOPMENT AND TESTING FOR NETWORK ACCUMULATION

A. Introduction

In this chapter, linear predictor models for the moving average accumulation (in MIP networks) are developed and tested. Empirical modeling is done using "black-box" approach and accumulation signal is modeled as time-series. SI technique AR is used for modeling. In this chapter single-step and multi-step predictors are developed.

A detailed discussion on the linear modeling technique AR is done in Chapter III. Next section discusses the performance metrics used while developing the predictors in this work. It will be followed by the development of single-step and multi-step predictors.

B. Performance Metrics

In this work, two types of errors are used for performance evaluation of the predictors developed. The first one is the Mean Square Error(MSE_1). It is the ratio between the sum of the square of the prediction error and the sum of the square of the input data, as as described in Equation 5.1.

$$MSE_1 = \frac{\sum_{k=1}^{N}(x(k) - \hat{x}(k|k-1))^2}{\sum_{k=1}^{N}(x(k))^2} X100 \tag{5.1}$$

where $x(k)$ is the observation, $x(k|k-1)$ is the predicted data and N is the total number of data points. MSE can also be defined as inverse of Signal-to-Noise Ratio (SNR). Hence MSE is considered to be the best performance metric which gives good idea about the quality of predictor.

The second metric used in this work for performance evaluation of the predictors is a variant of mean square error (MSE_2). It is defined as the ratio between the sum of the square of the prediction error and sum of square of input data from which its (input data's) mean has been removed. Mathematically, it can be represented as:

$$MSE_2 = \frac{\sum_{k=1}^{N}(x(k) - \hat{x}(k|k-1))^2}{\sum_{k=1}^{N}(x(k) - \overline{x})^2} X100 \tag{5.2}$$

where, $\overline{x}$ is the arithmetic mean of the observation $x(k)$. In case of signals with large variations, this metric might give better indication of the predictor performance as compared to the MSE_1 discussed earlier.

C. Description on Training, Testing and Validation Data Sets

As described in chapter IV MIP data generated has been used for the predictor development. First moving average of the accumulation time series is generated. As shown in Figure 15, the data obtained is divided into training, testing and validation data.

In this work, different predictors have been developed with data sets with varying wired losses. These predictors are then tested with testing data sets having different wired losses than the one with which it was developed. All the predictors are developed with the ideal wireless channel conditions. Which means that there are no wireless losses in the data with which the predictor is developed. The impact of wireless losses on the accuracy of these predictors is discussed in detail in Chapter VI. The real traffic data is collected with three different source-destination pairs as discussed in section 4.D.2. Description of various data-sets used for predictor testing is as follow:

1. Table III gives the data-sets along with the wired losses present in them. For

Table III. Wired losses(%) present in different data-sets with source node NIML (Texas A&M) and destination node pli_pa (planet-lab)

File Name	Wired Loss(%)
Data-set1	0
Data-set2	0
Data-set3	0.3
Data-set4	0.7
Data-set5	1.4
Data-set6	2.0

these flows source node is NIML (TAMU) and the destination node is pli_pa node present in the planet-lab, Princeton.

2. Table IV gives the data-sets along with the wired losses present in them. For these flows source node is NIML (TAMU) and the destination node is gtidsl node present in the planet-lab, Princeton.

3. Table V gives the data-sets along with the wired losses present in them. For these flows source node is gtidsl and the destination node is nbgisp, both these nodes are present in the planet-lab, Princeton.

These traces are collected at different times and hence have different wired losses.

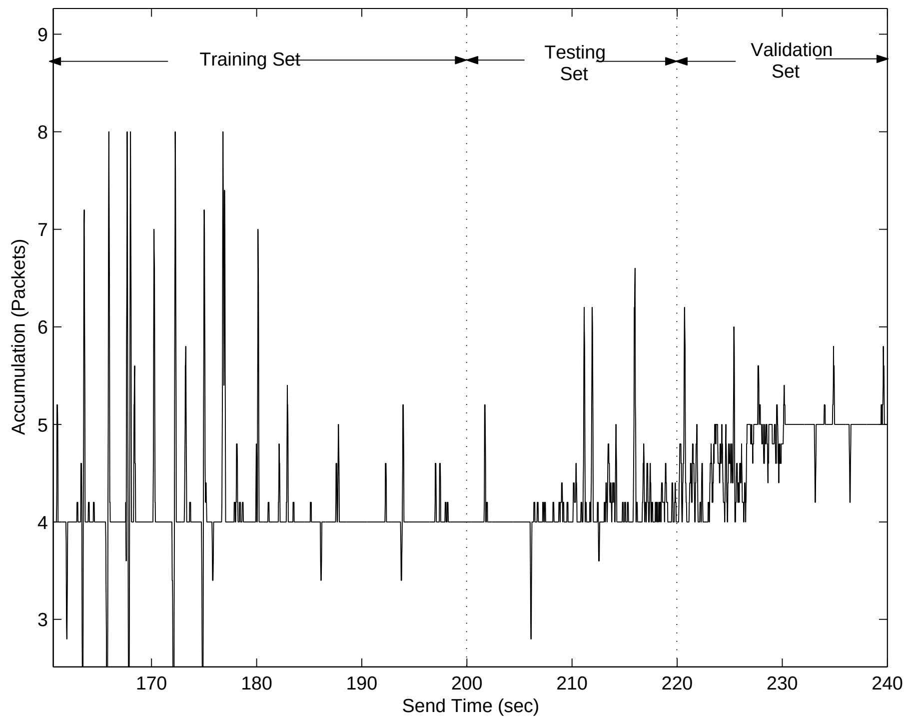

Fig. 15. Representation of Training, Testing and Validation Data Sets

Table IV. Wired losses(%) present in different data-sets with source node NIML (Texas A&M) and destination node gtidsl (planet-lab)

File Name	Wired Loss(%)
Data-set7	0
Data-set8	1.4
Data-set9	2.7
Data-set10	3.6
Data-set11	4.3
Data-set12	5.3
Data-set13	9

Table V. Wired losses(%) present in different data-sets with source node gtidsl (planet-lab) and destination node nbgisp (planet-lab)

File Name	Wired Loss(%)
Data-set14	0
Data-set15	2.6
Data-set16	3.0
Data-set17	5.1
Data-set18	10.4

Table VI. Wired losses(%) present in data-sets used for predictor development with source node NIML (Texas A&M) and destination node pli_pa (planet-lab)

Predictor Name	Wired loss (%) present	Predictor Order
Predictor1	0	7
Predictor2	1	8
Predictor3	2	10

D. Development of SSP

Once the desired data is obtained, the next step is to develop predictor model using system identification techniques. Single-step-ahead prediction in this research means 20msec ahead prediction for the first and second source-destination pair, whereas for the third pair it means 40msec ahead prediction of the moving average accumulation.

For the data-sets with source node NIML (Texas A&M) and destination node pli_pa (planet-lab) three different predictors have been developed. Table VI gives the information about the wired losses present in the traces used for the predictor development.

Table VII gives the prediction errors described in previous section 5.2 for all the SSPs obtained using AR model. Figures 16, 18 and 20 show the performance of the Predictor2 for Dataset1, 2 and 3. The discrete points in the figure presents the predicted value, whereas the continuous line is the actual accumulation signal. Figures 17, 19 and 21 give the prediction error for the same Data-Sets. From the prediction figures it can be seen that predictor can capture the dynamics of the network by predicting moving average of the accumulation signal. From table VII it can be observed that MSE_1 as well as MSE_2 for all the predictors is pretty high for the case of Dataset5. This implies that the predictors developed fails to predict Dataset5 well. It can be seen that the wired losses does not affect the accuracy of the predictors to a great extent. This is because the wired losses follow a trend and are not totally random. For Dataset1 moving average accumulation signal is also plotted. From the plot it can be observed that the predicted values follow closely the moving average signal. This natural since the predictors were developed using moving average accumulation signal.

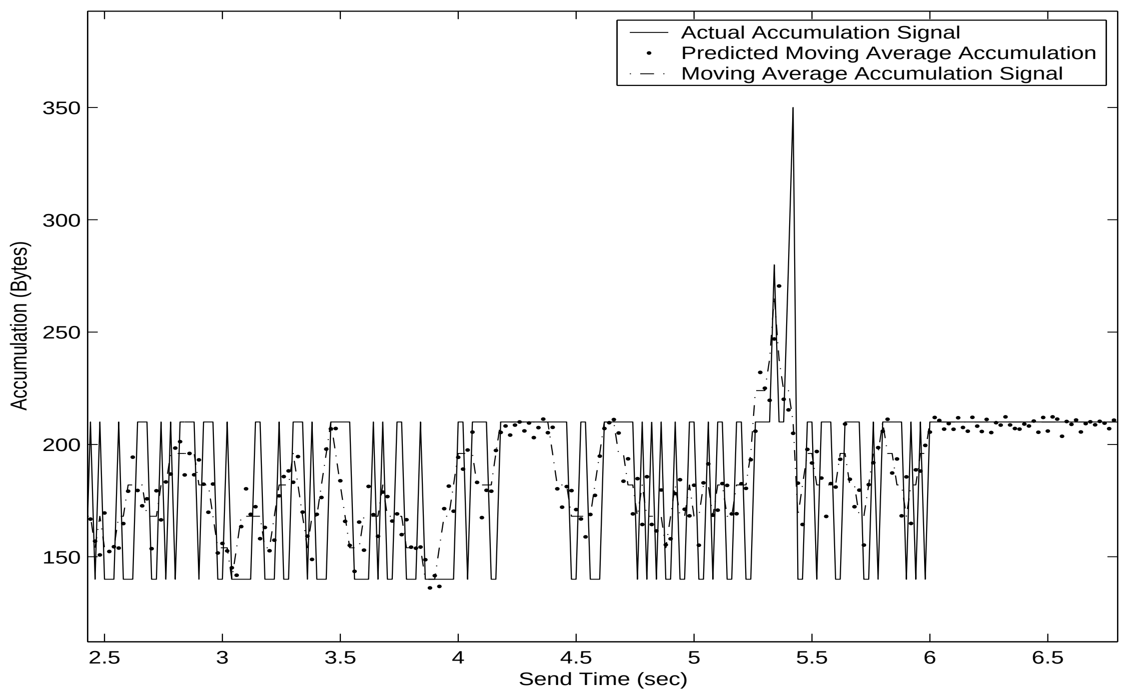

Fig. 16. Single-Step-Ahead Prediction of Data-set1 Using Predictor2

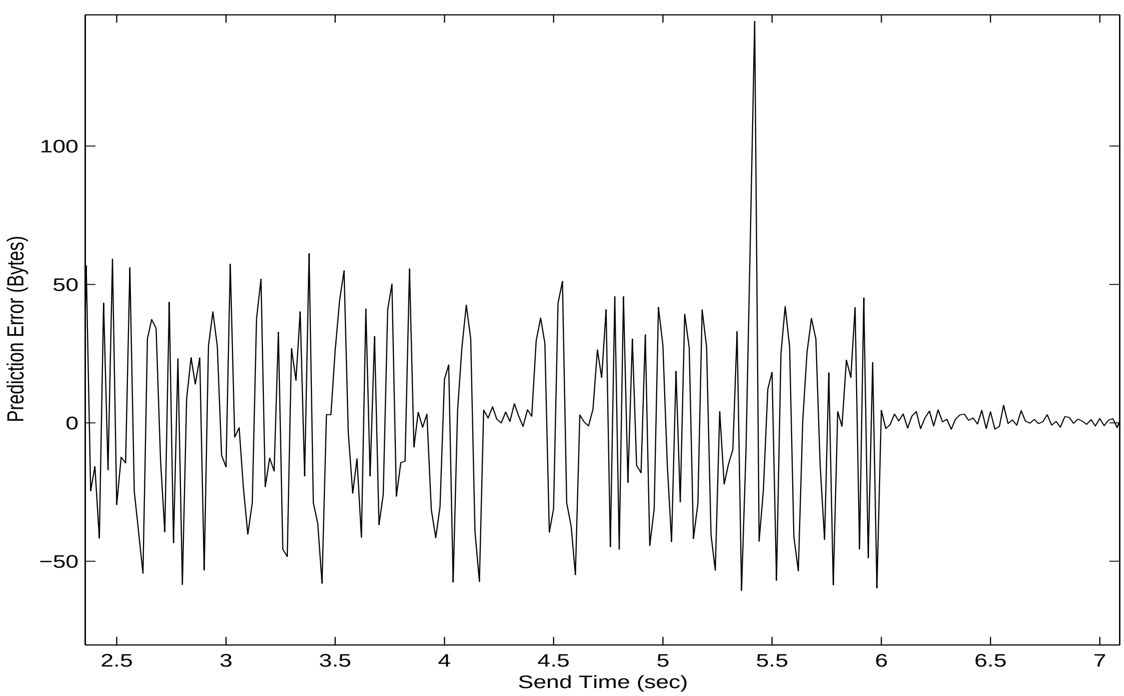

Fig. 17. Single-Step-Ahead Prediction Error of Data-set1 Using Predictor2

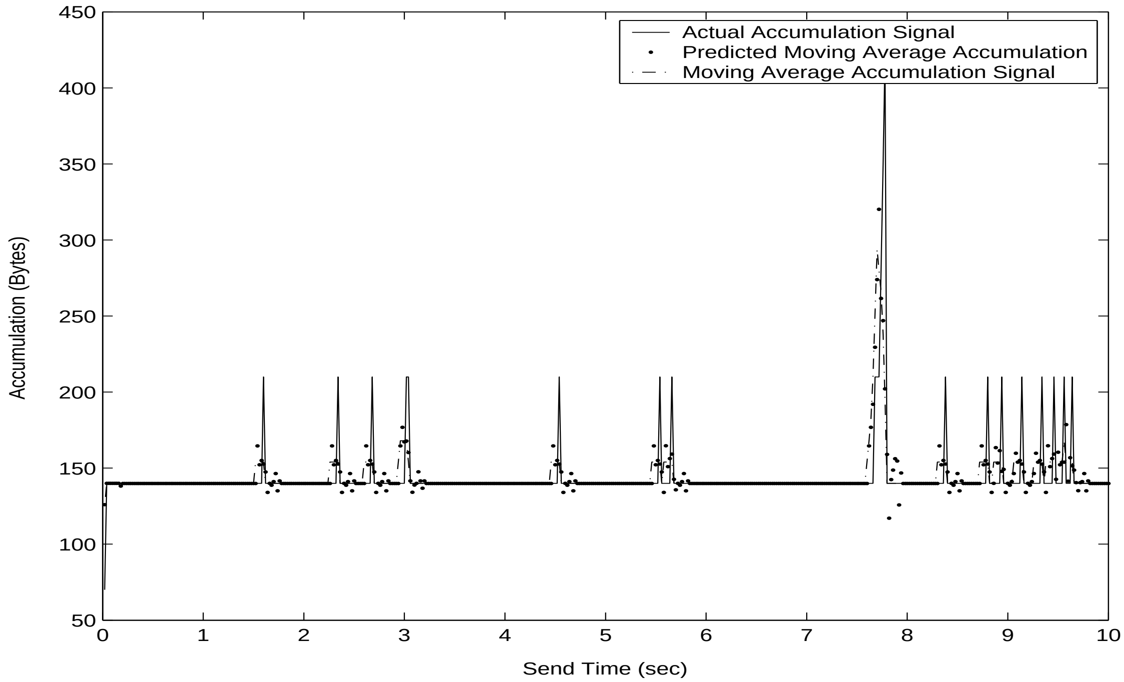

Fig. 18. Single-Step-Ahead Prediction of Data-set2 Using Predictor2

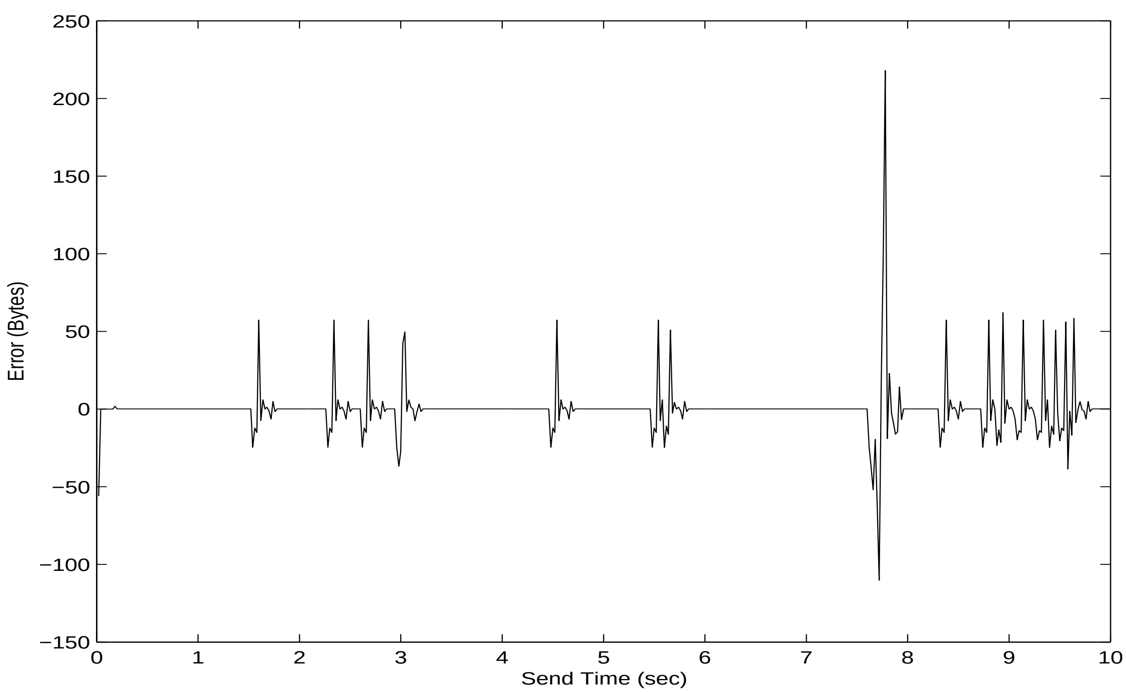

Fig. 19. Single-Step-Ahead Prediction Error of Data-set2 Using Predictor2

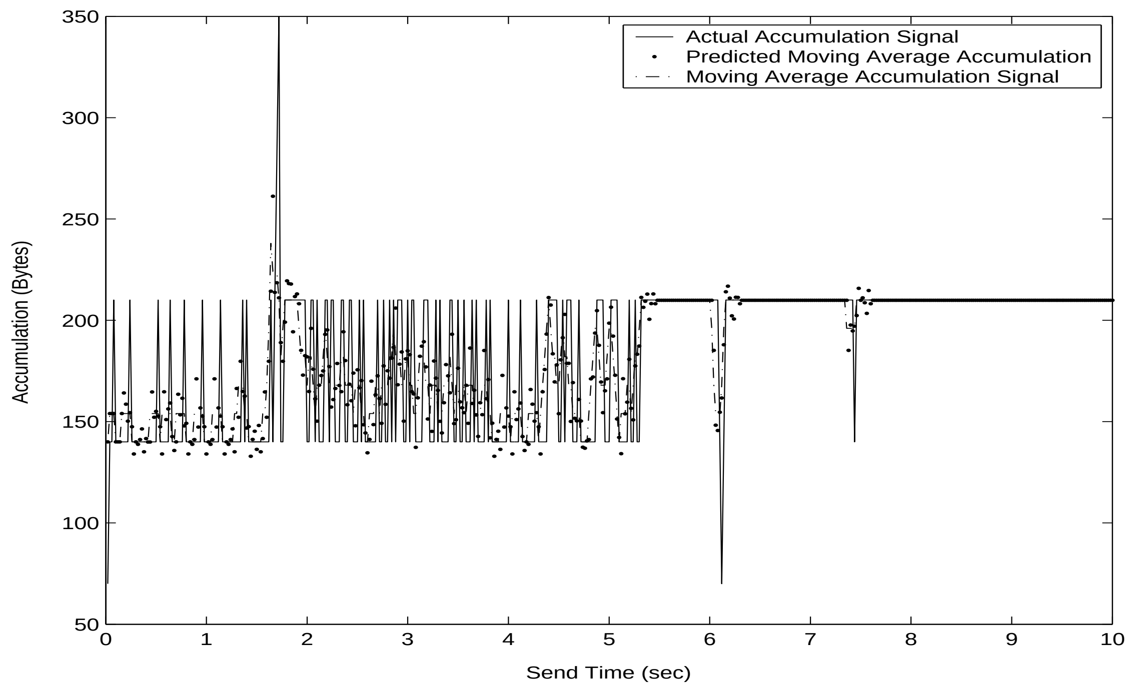

Fig. 20. Single-Step-Ahead Prediction of Data-set3 Using Predictor2

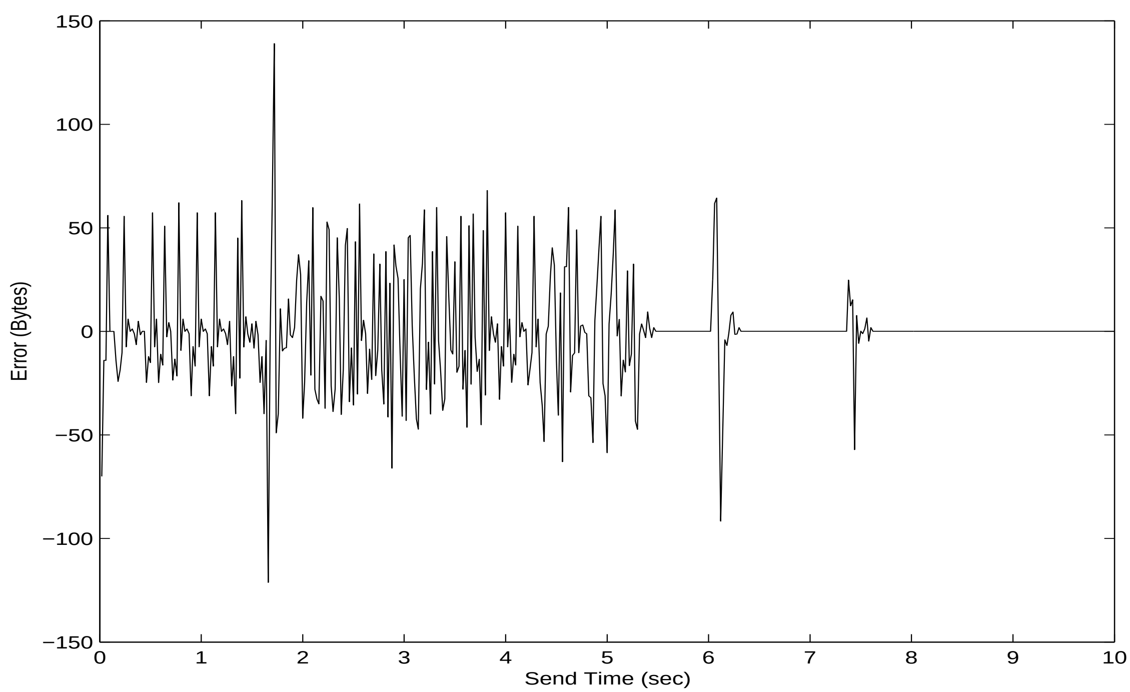

Fig. 21. Single-Step-Ahead Prediction Error of Data-set3 Using Predictor2

Table VII. Single-step-ahead prediction results for Predictor1, Predictor2 and Predictor3

Data-Set	Predictor1		Predictor2		Predictor3	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset1	1.1	43.3	1.3	50.1	1.3	50.5
Dataset2	2.0	49.7	2.2	56.4	2.3	57.6
Dataset3	1.0	36.6	1.1	42.2	1.2	43.2
Dataset4	0.8	35.3	0.9	40.0	0.9	41.3
Dataset5	4.8	40.2	5.8	48.6	6.0	50.5
Dataset6	1.2	37.8	1.6	48.0	1.6	50.2

Table VIII. Wired loss(%) present in the data used for the predictor development with source node NIML (Texas A&M) and destination node gtidsl (planet-lab)

Predictor Name	Wired loss (%) present	Predictor Oder
Predictor4	2.7	9
Predictor5	4.3	5
Predictor6	9	3

For the data-sets with source node NIML (Texas A&M) and destination node gtidsl (planet-lab) three predictors have been developed. Table VIII gives the information about the wired losses present in the traces used for developing these predictors.

Table IX gives information about the prediction errors of the predictors developed on varying wired loss data. Figures 22, 24 and 26 show the performance of the Predictor5 for Datasets7, 8 and 9. Figures 23, 25 and 27 give the error plot of the same traces. Again, it can be seen from table IX wired losses does not impact the predictor accuracy to a great extent. From the figures it can be observed that the predictors developed are able to capture the network dynamics.

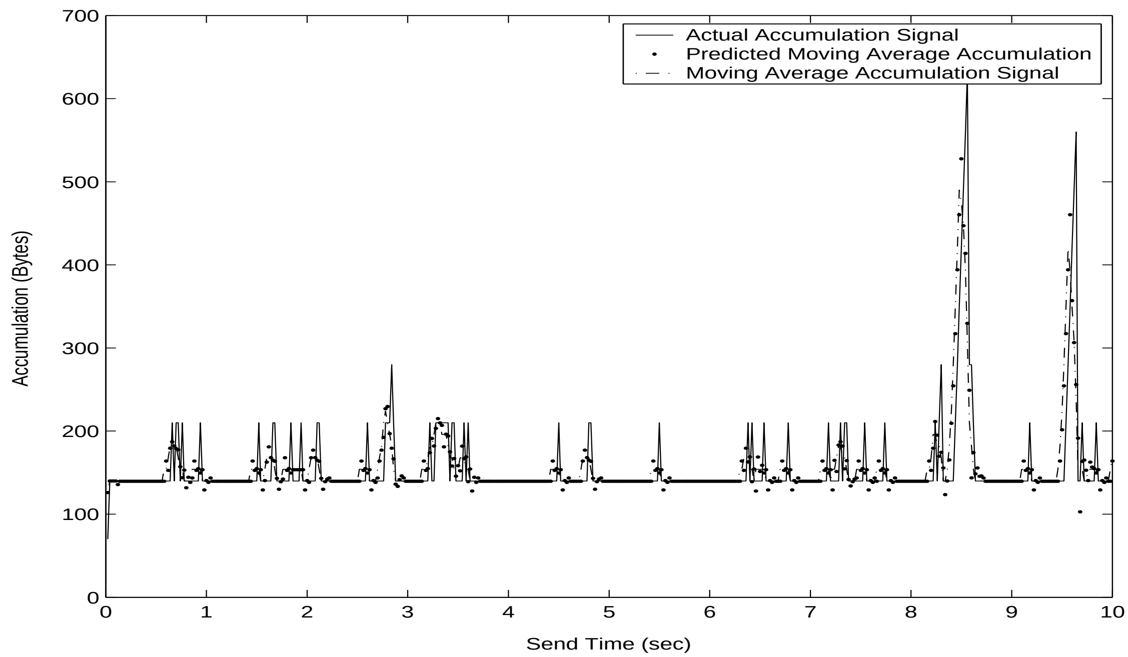

Fig. 22. Single-Step-Ahead Prediction of Data-set7 Using Predictor5

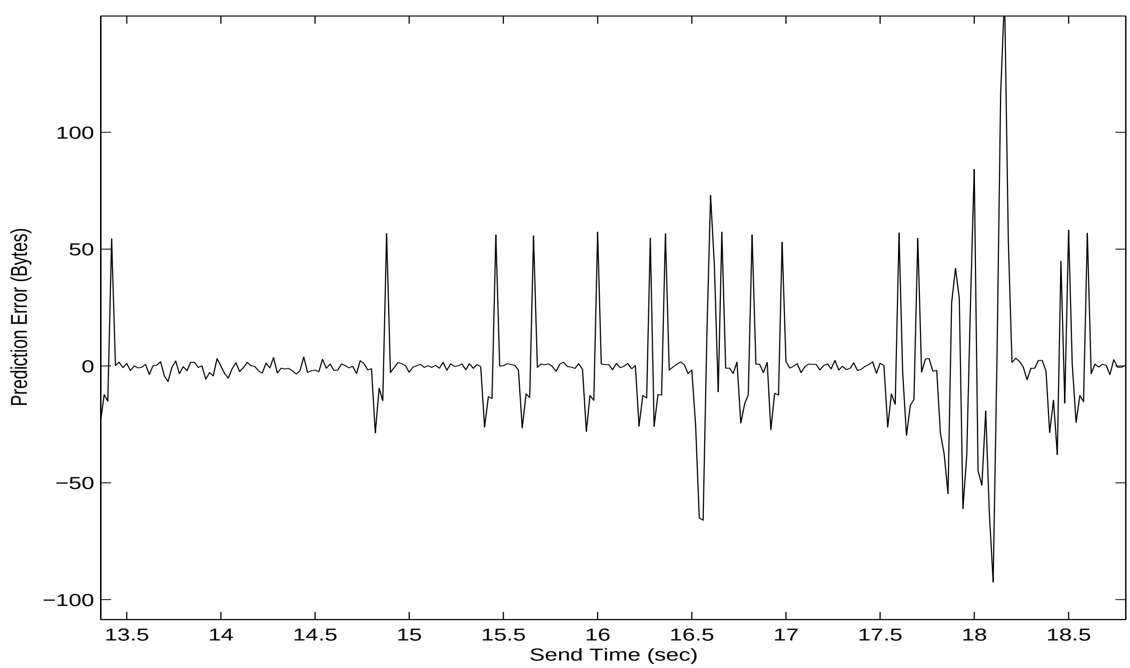

Fig. 23. Single-Step-Ahead Prediction Error of Data-set7 Using Predictor5

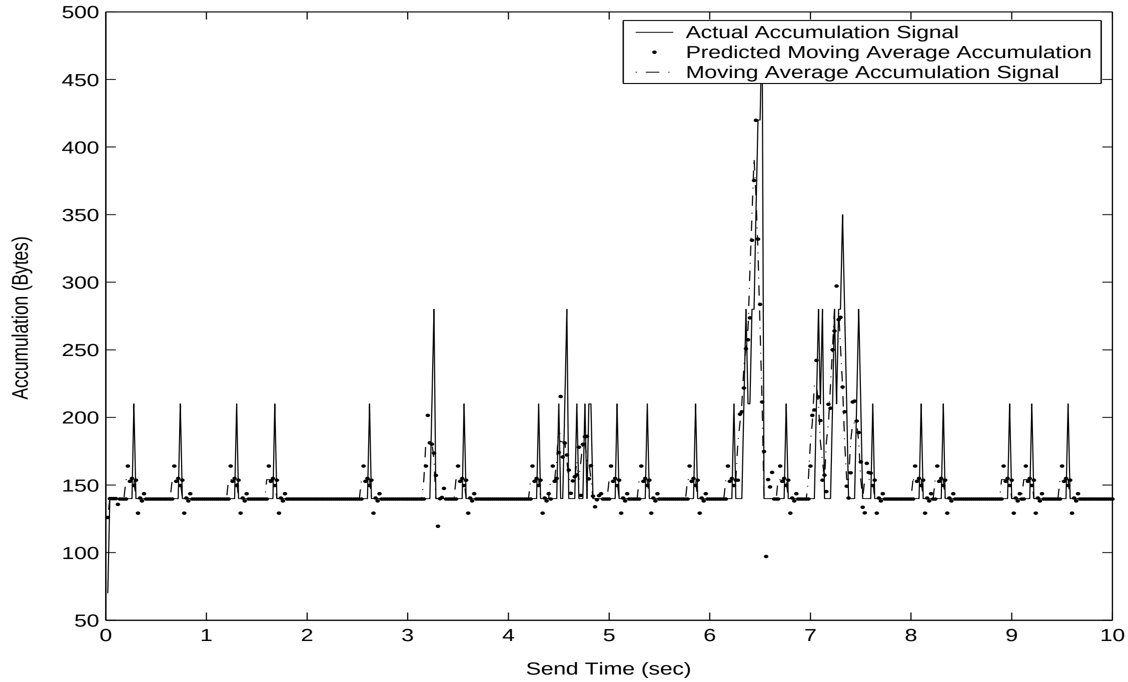

Fig. 24. Single-Step-Ahead Prediction of Data-set8 Using Predictor5

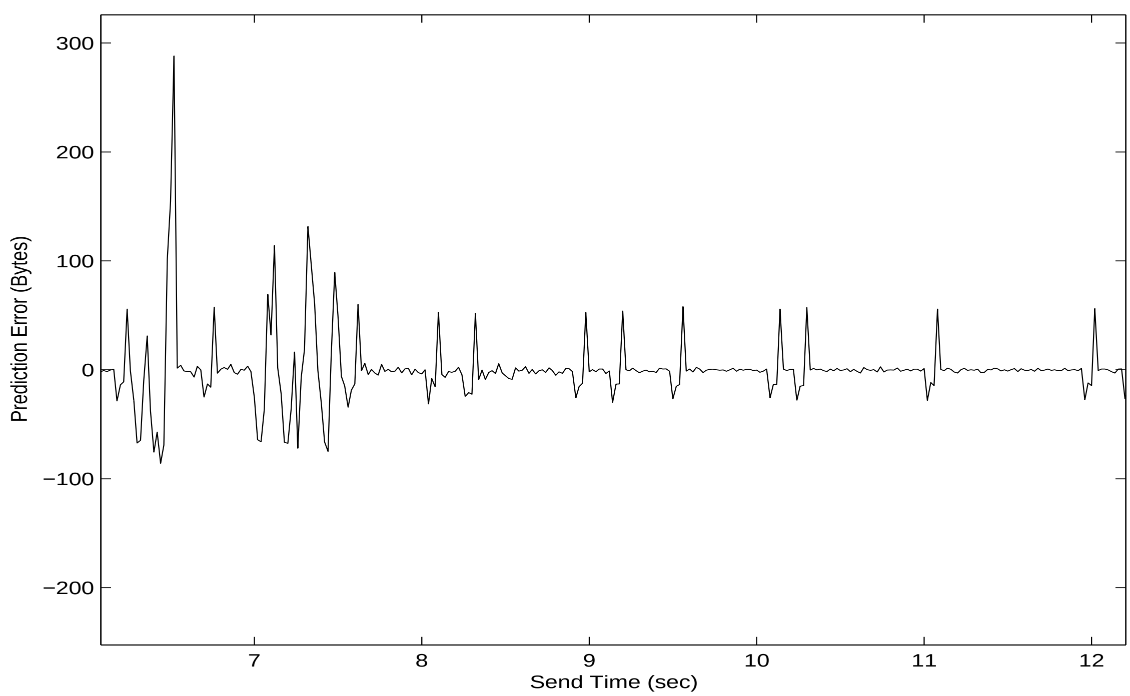

Fig. 25. Single-Step-Ahead Prediction Error of Data-set8 Using Predictor5

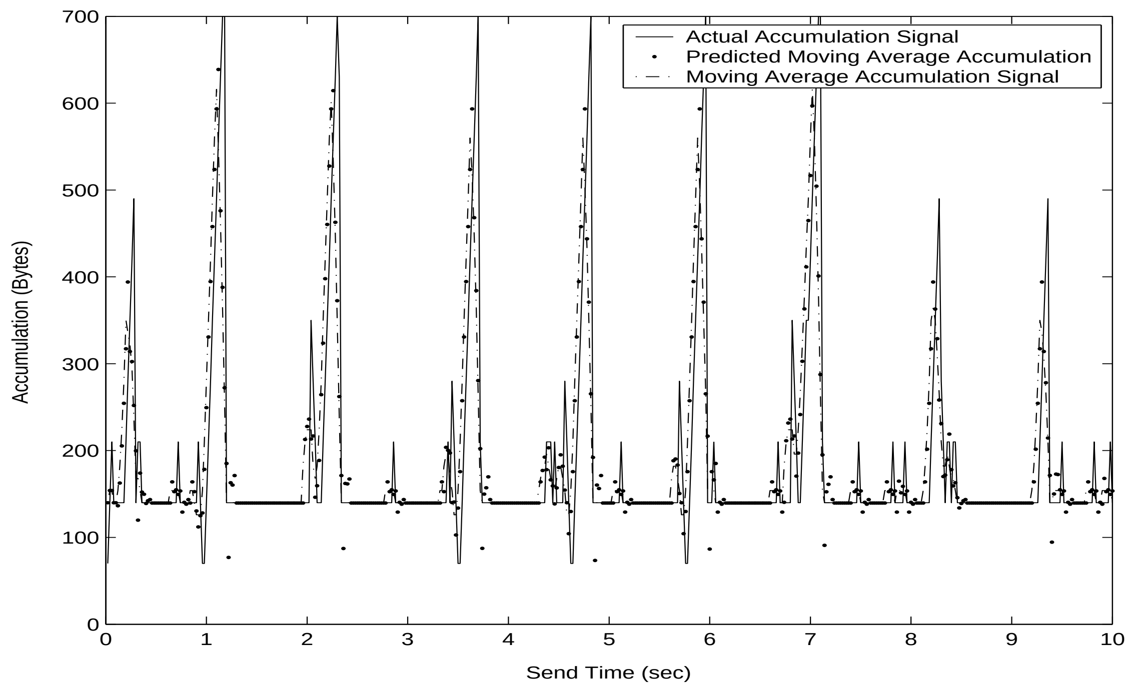

Fig. 26. Single-Step-Ahead Prediction of Data-set9 Using Predictor5

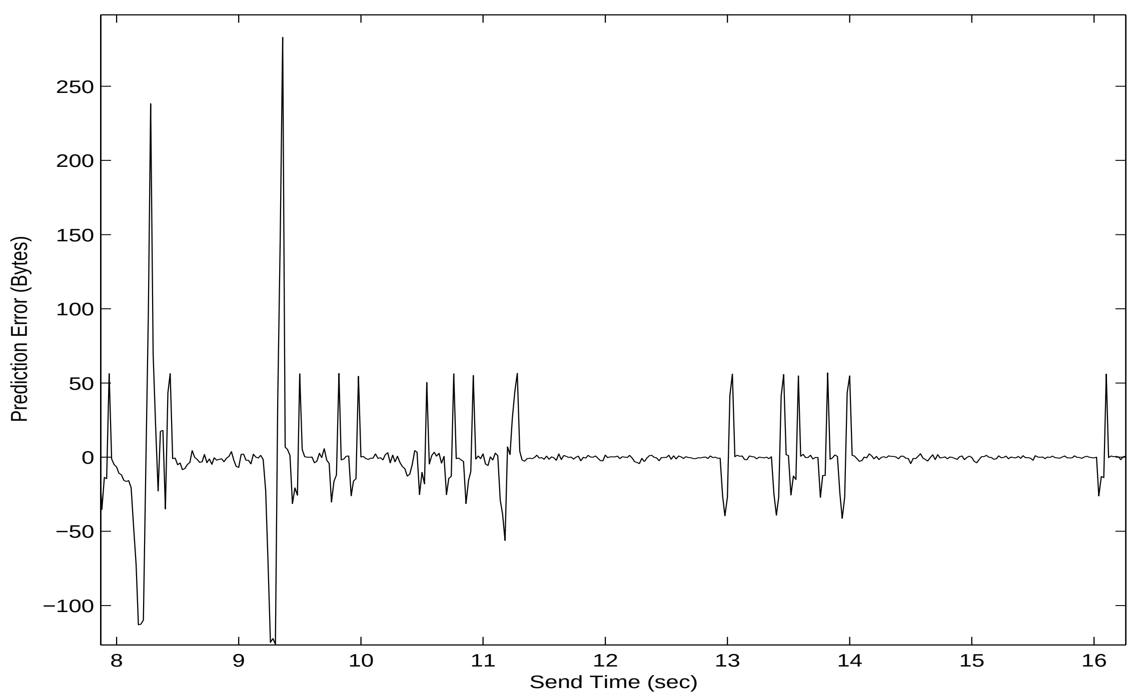

Fig. 27. Single-Step-Ahead Prediction Error of Data-set9 Using Predictor5

Table IX. Single-step-ahead prediction results for Predictor4, Predictor5 and Predictor6

Data-Set	Predictor4		Predictor5		Predictor6	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset7	3.5	50.5	3.5	59.8	2.6	37.3
Dataset8	6.6	38.6	6.3	36.7	4.4	25.6
Dataset9	9.2	41.8	8.5	36.5	5.9	26.6
Dataset10	8.4	41.5	7.9	39.0	5.5	27.3
Dataset11	4.3	56.3	4.1	54.2	3.0	38.9
Dataset12	7.2	31.3	6.8	29.5	4.8	20.9
Dataset13	13.0	31.2	11.2	28.6	8.2	19.8

Table X. Wired losses(%) present in the data-sets used for predictor development with source node gtidsl (planet-lab) and destination node nbgisp (planet-lab)

Predictor Name	Wired loss (%) present	Predictor Order
Predictor7	0	7
Predictor8	2.7	5
Predictor9	6.4	10

For the data-sets with source node gtidsl (planet-lab) and destination node nbgisp (planet-lab) three predictors have been developed. Table X gives the information about the wired losses present in the traces used for developing these predictors.

Table XI gives information about the predictions errors of the predictors developed on varying wired loss data. Figures 28, 30 and 32 show the performance of the Predictor8 for Datasets14, 15 and 16. Figures 29, 31 and 33 give the error plot of the same traces. The predictors performance vary irrespective of the increasing wired losses which was observed earlier also.

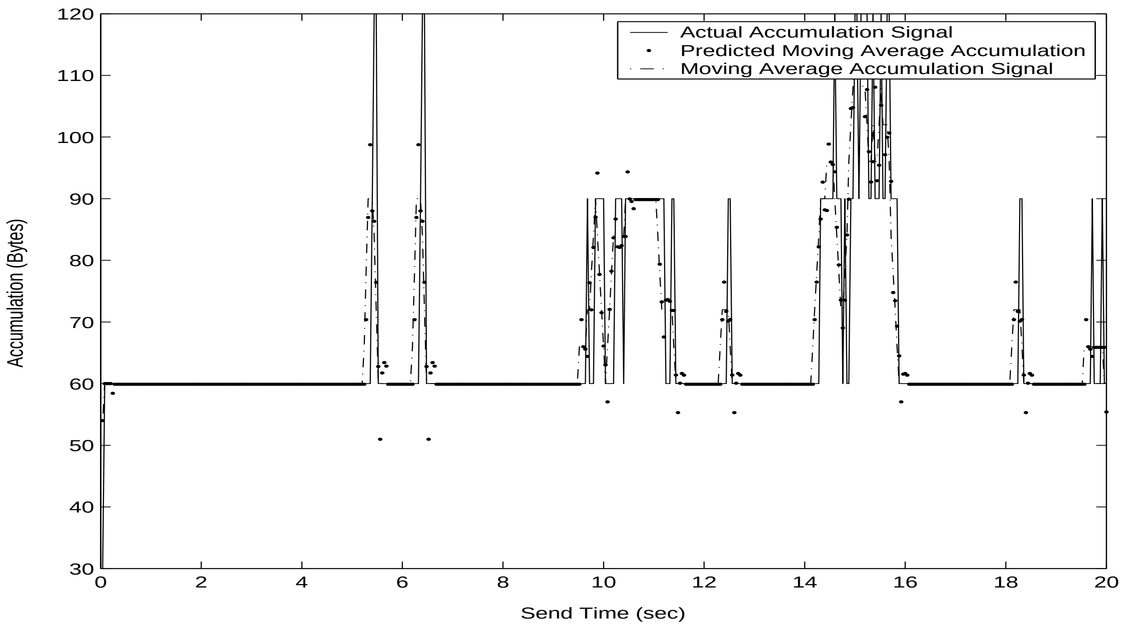

Fig. 28. Single-Step-Ahead Prediction of Data-set14 Using Predictor8

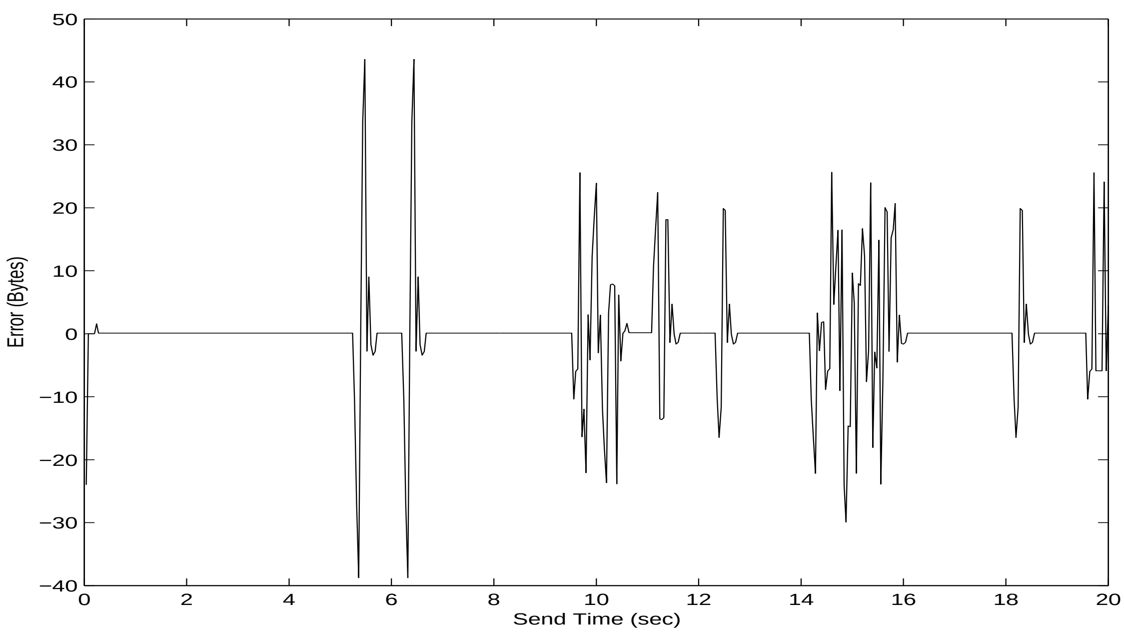

Fig. 29. Single-Step-Ahead Prediction Error of Data-set14 Using Predictor8

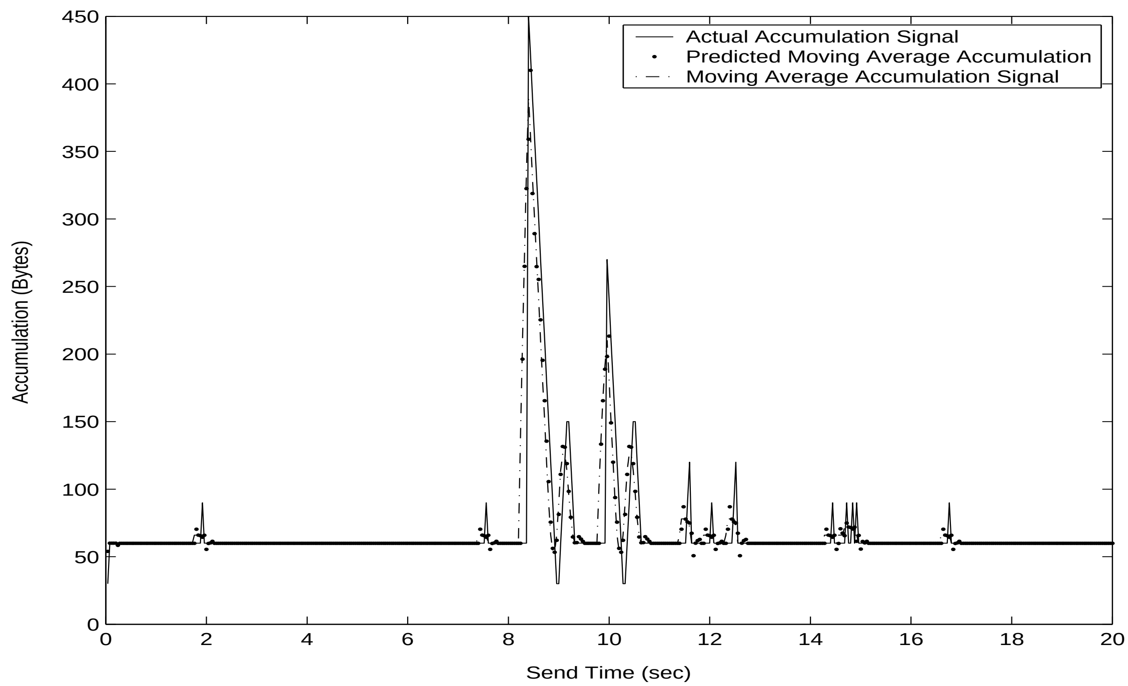

Fig. 30. Single-Step-Ahead Prediction of Data-set15 Using Predictor8

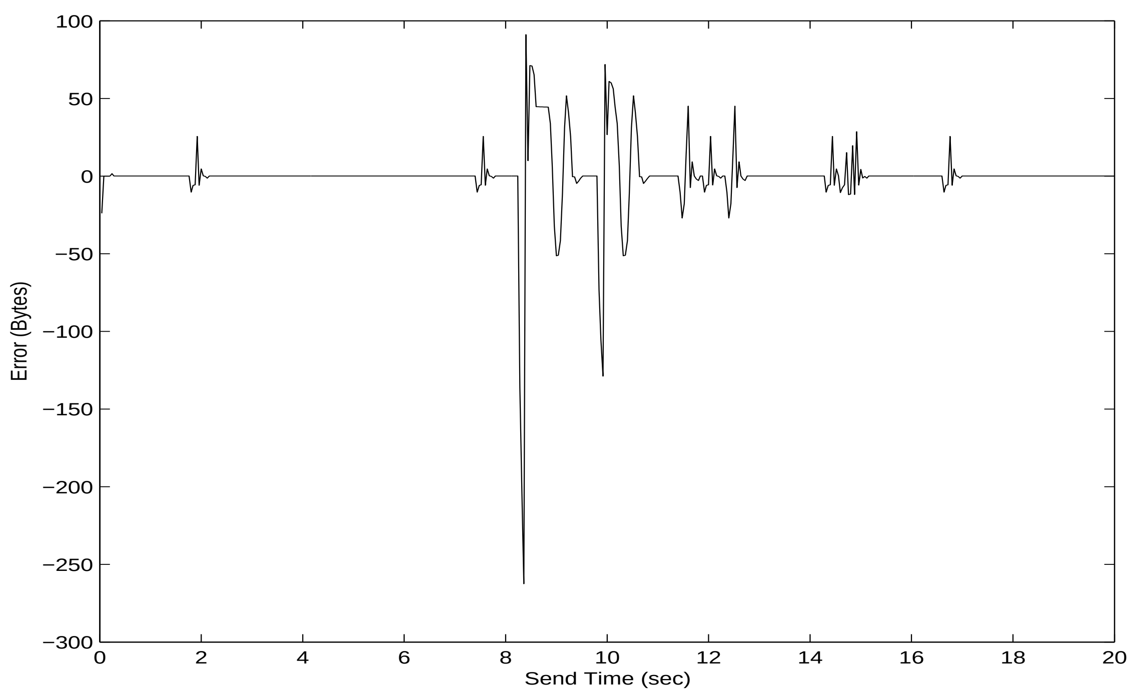

Fig. 31. Single-Step-Ahead Prediction Error of Data-set15 Using Predictor8

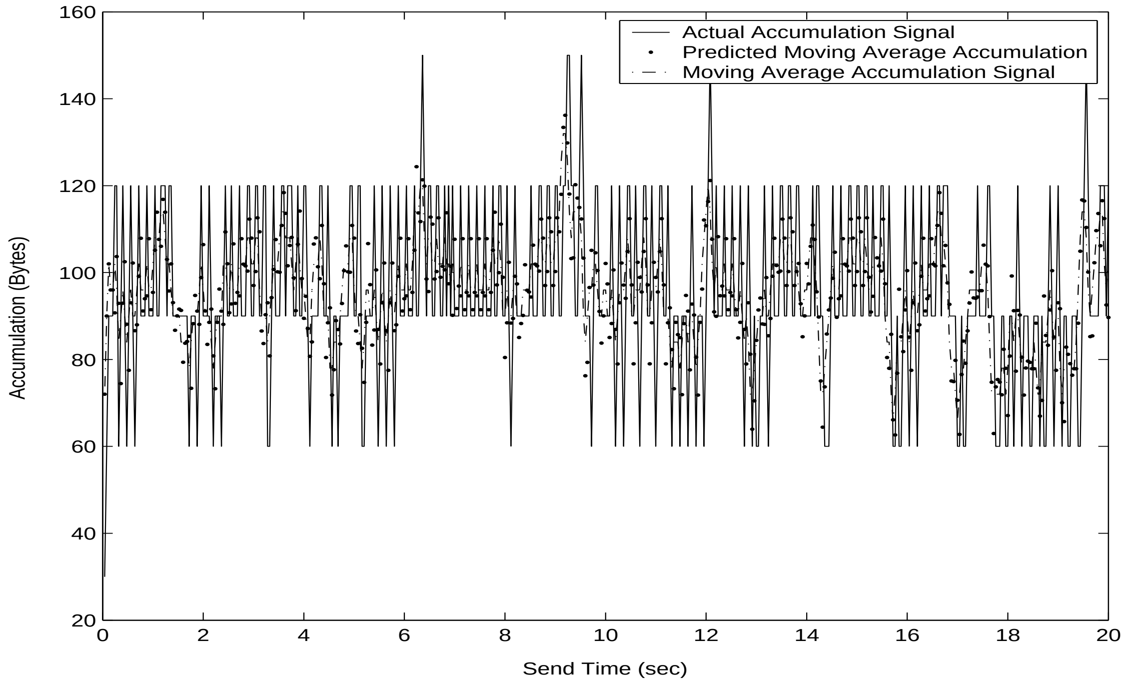

Fig. 32. Single-Step-Ahead Prediction of Data-set16 Using Predictor8

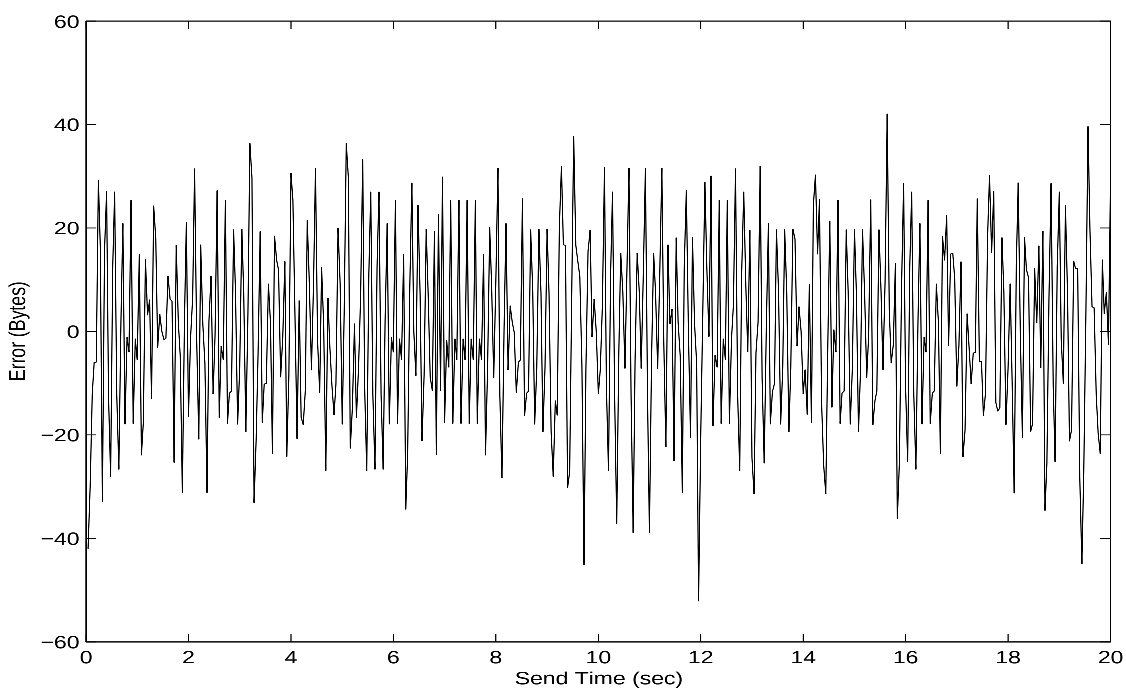

Fig. 33. Single-Step-Ahead Prediction Error of Data-set16 Using Predictor8

Table XI. Single-step-ahead prediction results for Predictor7, Predictor8 and Predictor9

Data-Set	Predictor7		Predictor8		Predictor8	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset14	0.9	31.7	1.0	34.0	0.7	26.6
Dataset15	4.6	28.0	5.2	29.0	3.8	21.4
Dataset16	3.6	78.6	3.3	73.4	2.6	58.1
Dataset17	4.1	68.6	4.1	68.3	3.1	51.6
Dataset18	2.6	3.0	2.9	3.3	2.0	2.4

E. Development of MSP

In this section, procedure followed in developing and evaluating MSP is discussed. Accumulation must be predicted well ahead of time so that it can be used in congestion control/avoidance algorithms. The following sections discuss about the development and performance evaluation of multi-step-ahead predictors.

1. Six-Step-Ahead Prediction

For the flow between gtidsl node and nbgisp node six-step-ahead means 0.24sec ahead prediction of the moving average of accumulation signal. The predictor used for six-step-ahead prediction is the same developed for SSP. The performance of the predictors developed for different data-sets is summarized in table XII. Figures 34, 36 and 38 show the performance of Predictor8 on Datasets14, 15 and 16. Whereas, Figures 35, 37 and 39 show the error plots of the predictions. It can be seen from the Table XII that six-step-ahead prediction errors are higher compared to SSP errors. From the prediction figures and error figures it can be seen that as compared to SSP, the predictor is not so accurate to capture the dynamics of the network ahead of time.

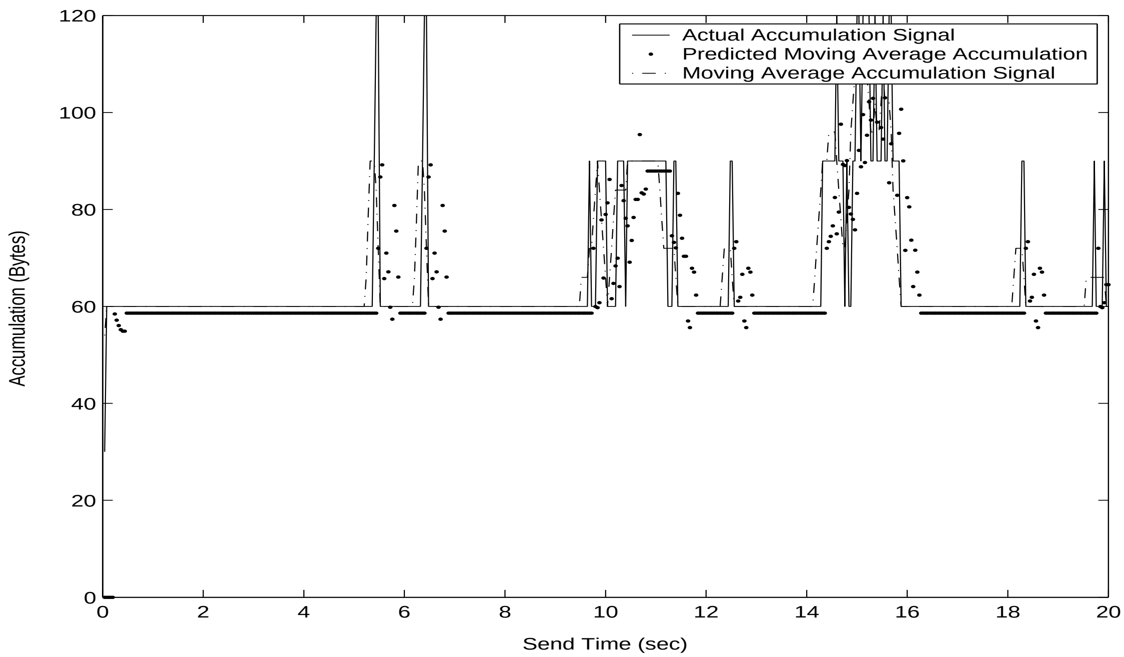

Fig. 34. Six-Step-Ahead Prediction of Data-set14 Using Predictor8

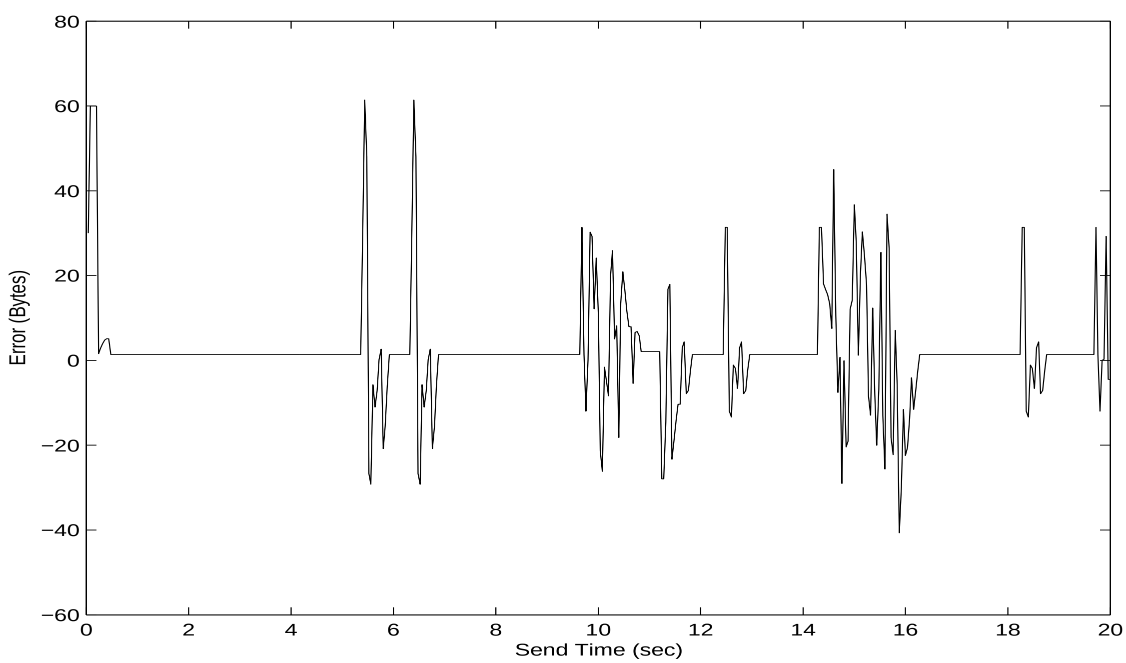

Fig. 35. Six-Step-Ahead Prediction Error of Data-set14 Using Predictor8

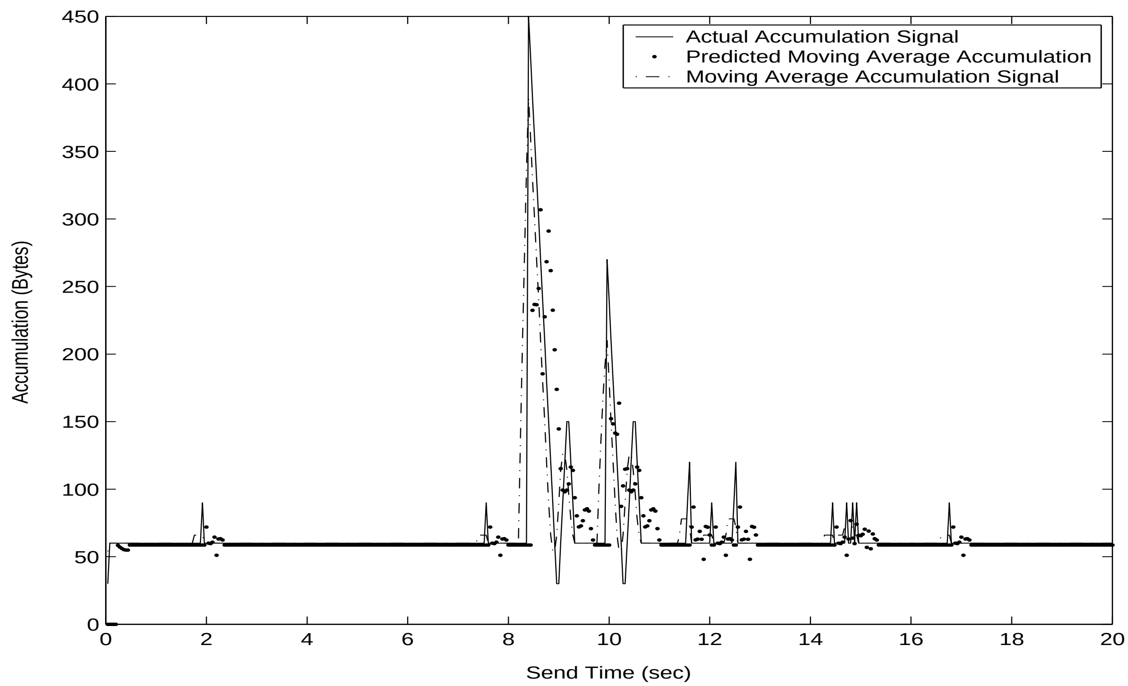

Fig. 36. Six-Step-Ahead Prediction of Data-set15 Using Predictor8

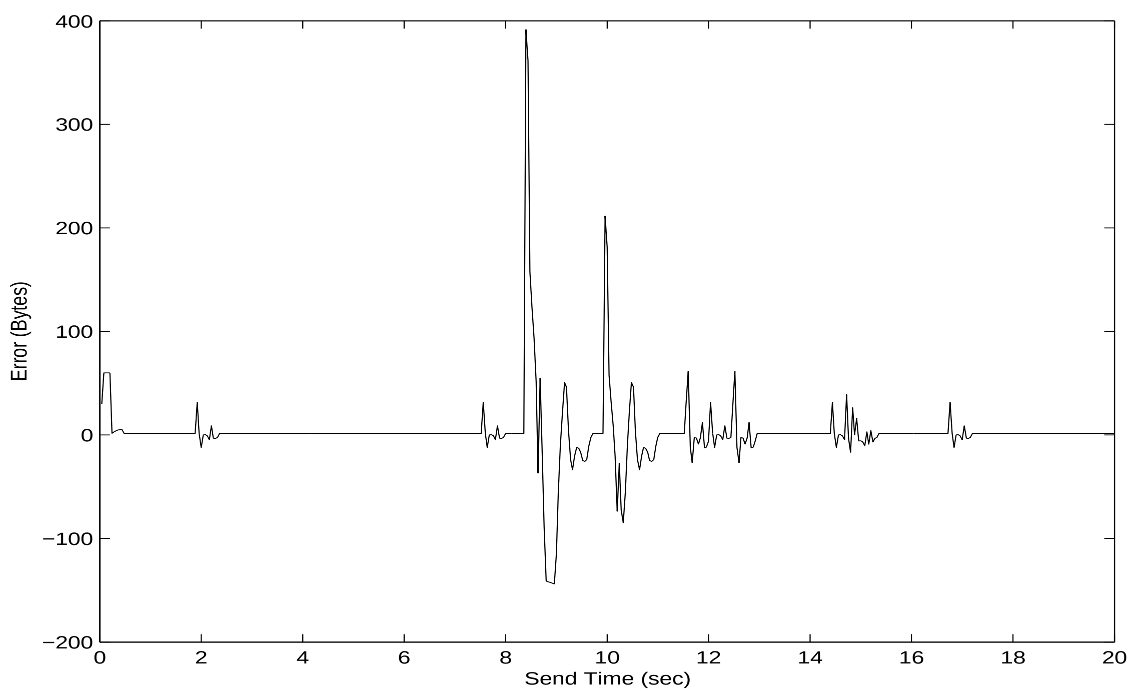

Fig. 37. Six-Step-Ahead Prediction Error of Data-set15 Using Predictor8

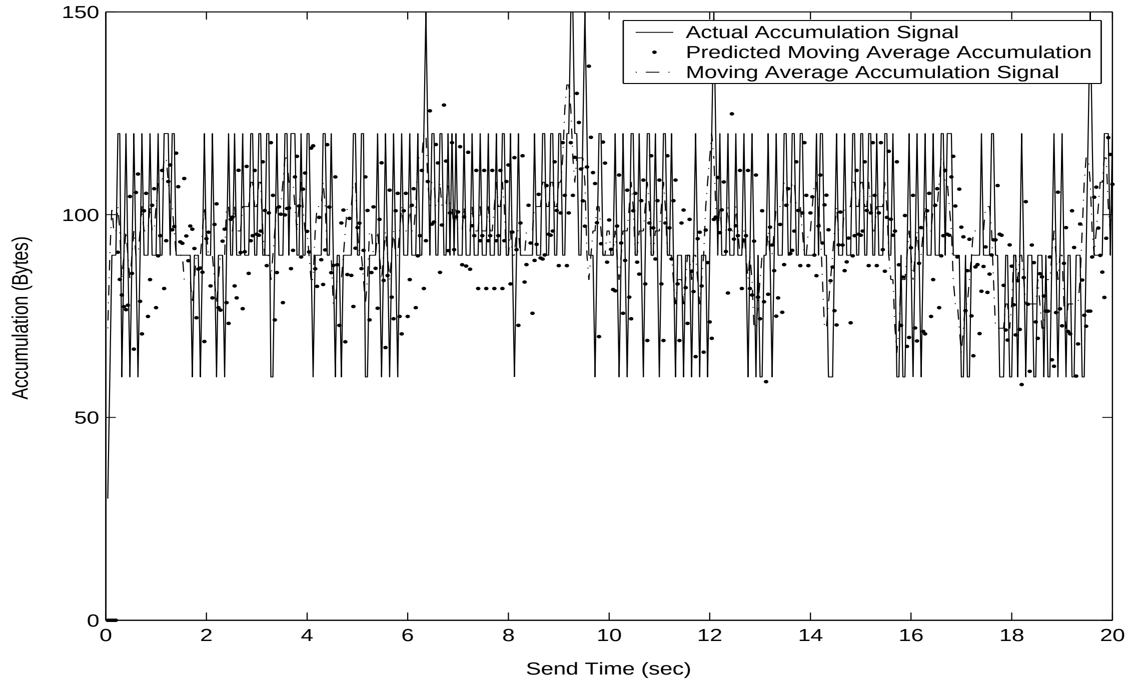

Fig. 38. Six-Step-Ahead Prediction of Data-set16 Using Predictor8

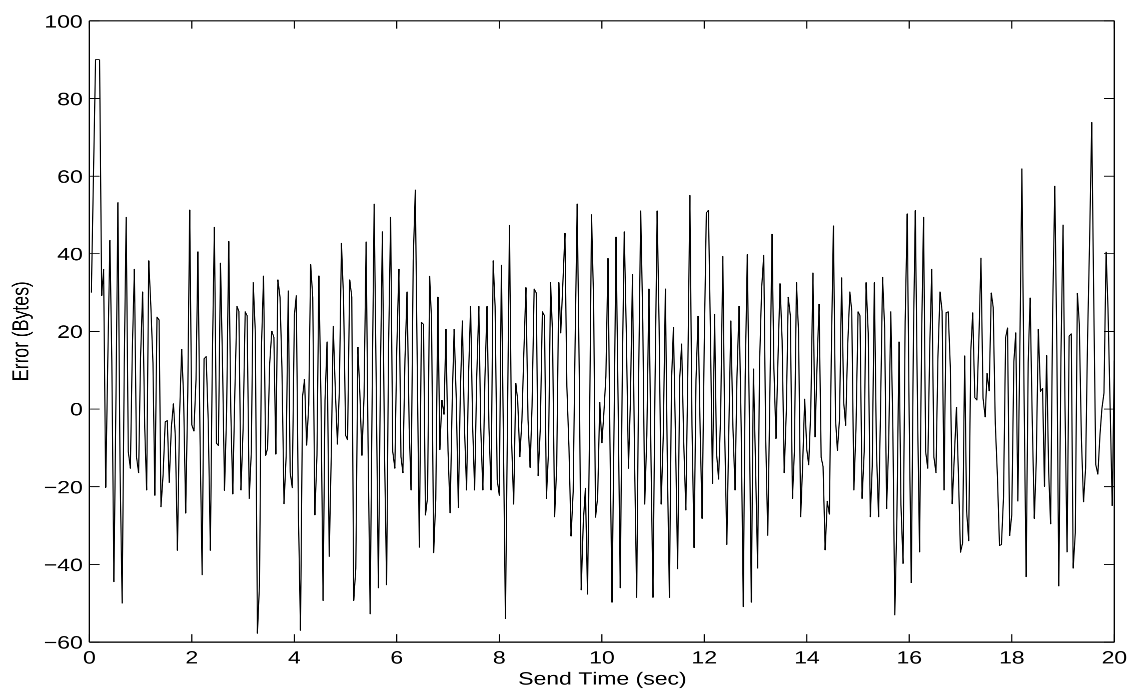

Fig. 39. Six-Step-Ahead Prediction Error of Data-set16 Using Predictor8

Table XII. Six-step-ahead prediction results for Predictor7, Predictor8 and Predictor9

Data-Set	Predictor7		Predictor8		Predictor8	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset14	2.1	75.5	2.2	77.4	2.2	78.9
Dataset15	13.3	73.6	12.4	69.1	17.1	95.0
Dataset16	6.7	148.0	7.3	162.2	4.4	97.7
Dataset17	8.0	133.7	8.5	142.4	6.2	102.9
Dataset18	8.2	9.5	9.0	10.5	11.8	13.8

2. Ten-Step-Ahead Prediction

The AR predictor model used for ten-step-ahead prediction is the same developed for SSP. Ten-step-ahead prediction in this research means 0.2 sec ahead prediction of the moving average accumulation. For the flow between a node in Texas A&M and pli_pa node in planet-lab the prediction errors of the predictors for different data-sets is presented in table XIII. Figures 40, 42 and 44 show the performance of the Predictor2. Figures 41, 43 and 45 give the error plot of the same traces. From the prediction figures it can be observed that predictor gives reasonably good ten-step-ahead predictions. Again, the prediction errors is high compared to the SSP. For Dataset1 moving average accumulation signal is also plotted. From the plot it can be seen that the predicted values follow more closely the moving average signal as expected.

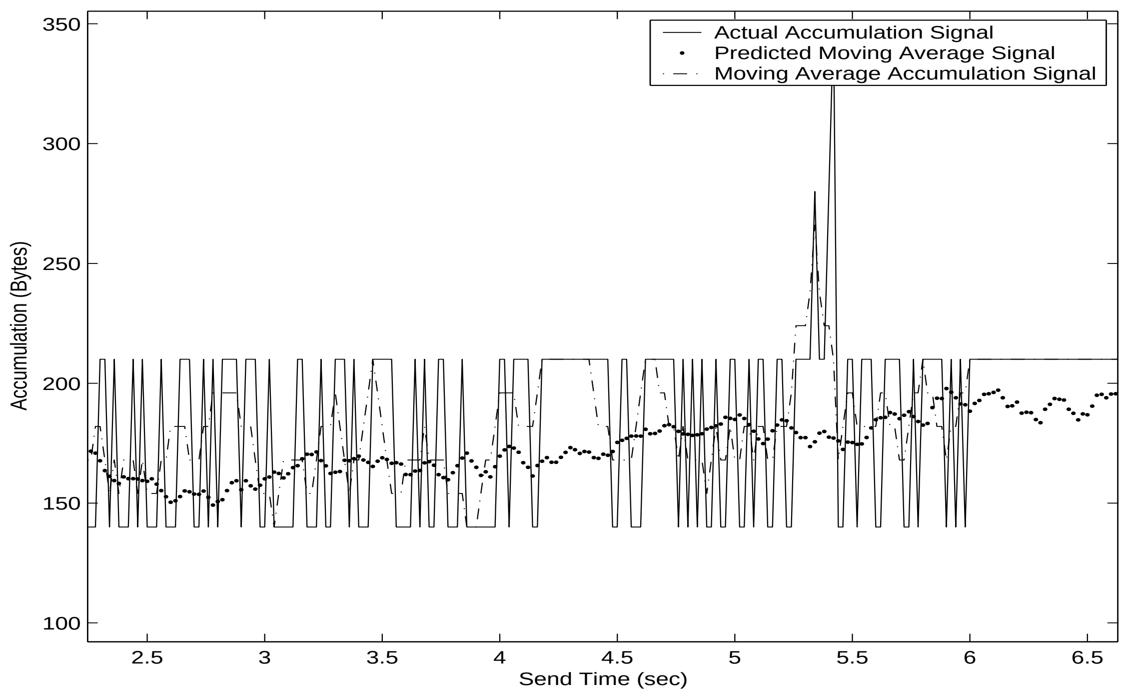

Fig. 40. Ten-Step-Ahead Prediction of Data-set1 Using Predictor2

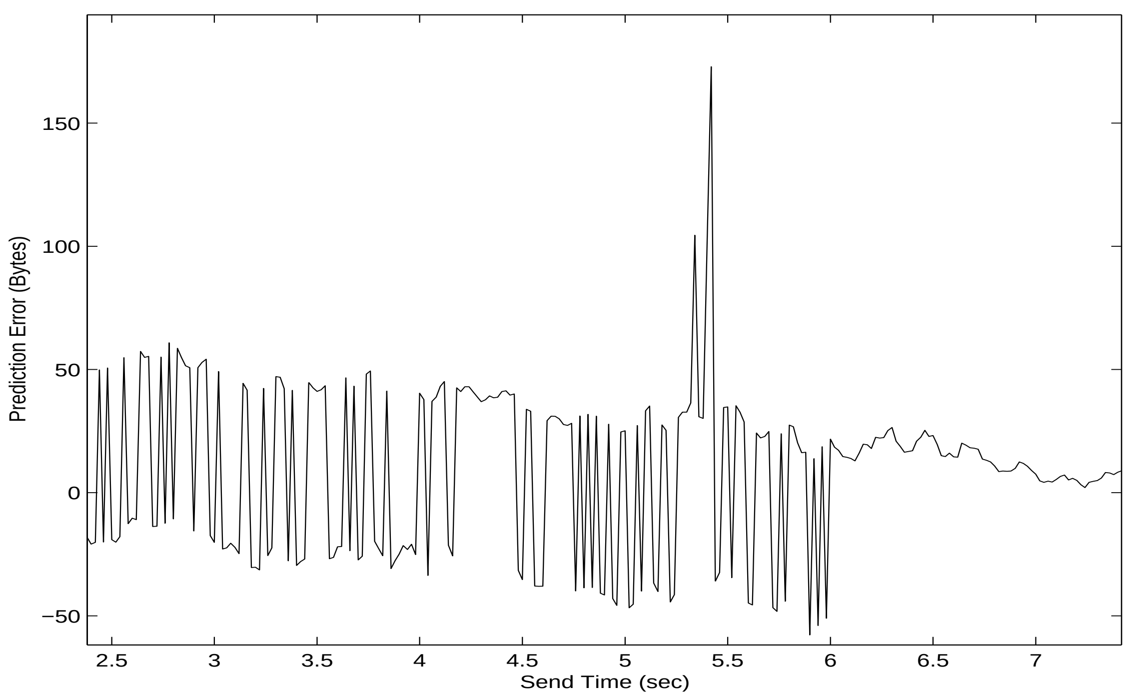

Fig. 41. Ten-Step-Ahead Prediction Error of Data-set1 Using Predictor2

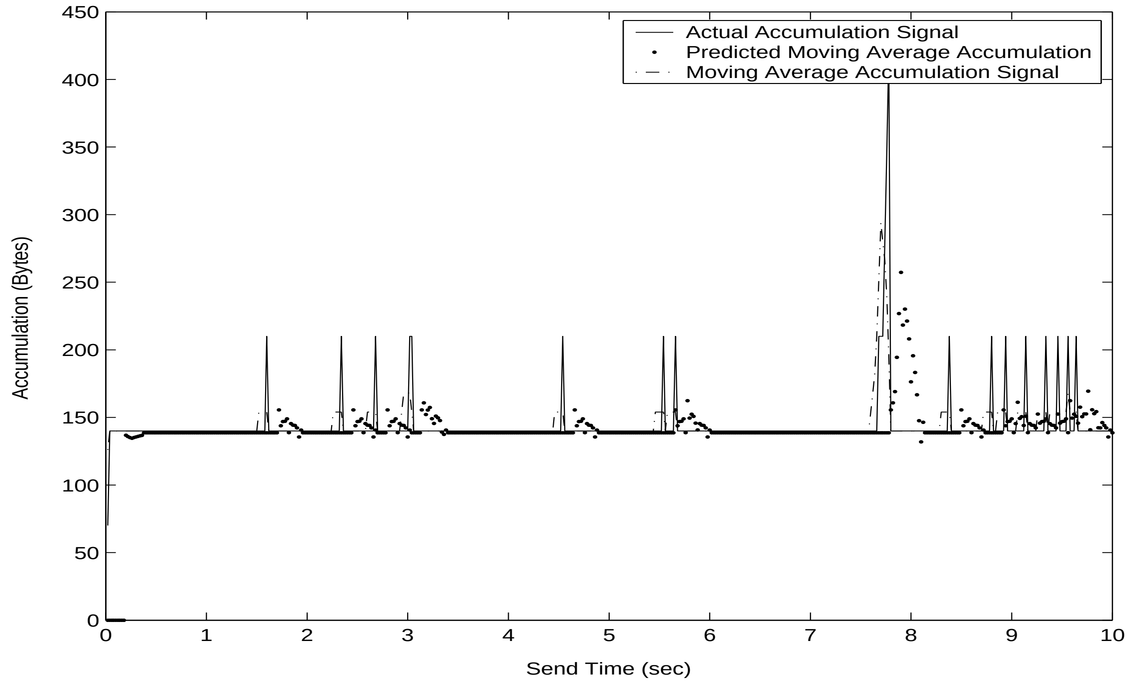

Fig. 42. Ten-Step-Ahead Prediction of Data-set2 Using Predictor2

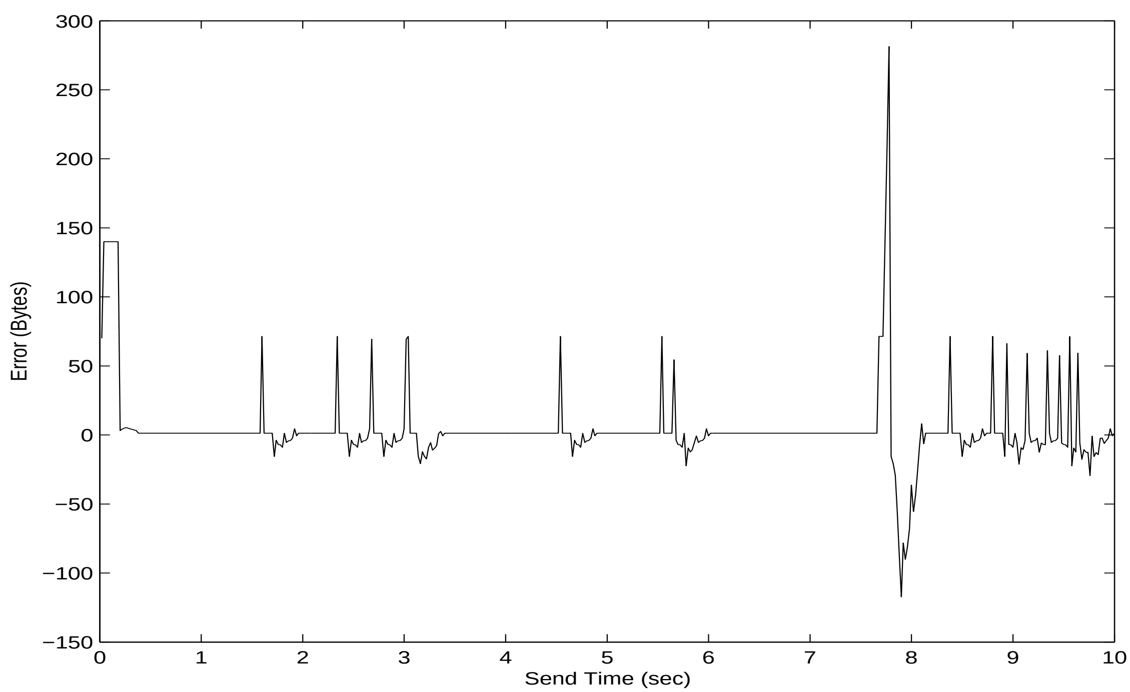

Fig. 43. Ten-Step-Ahead Prediction Error of Data-set2 Using Predictor2

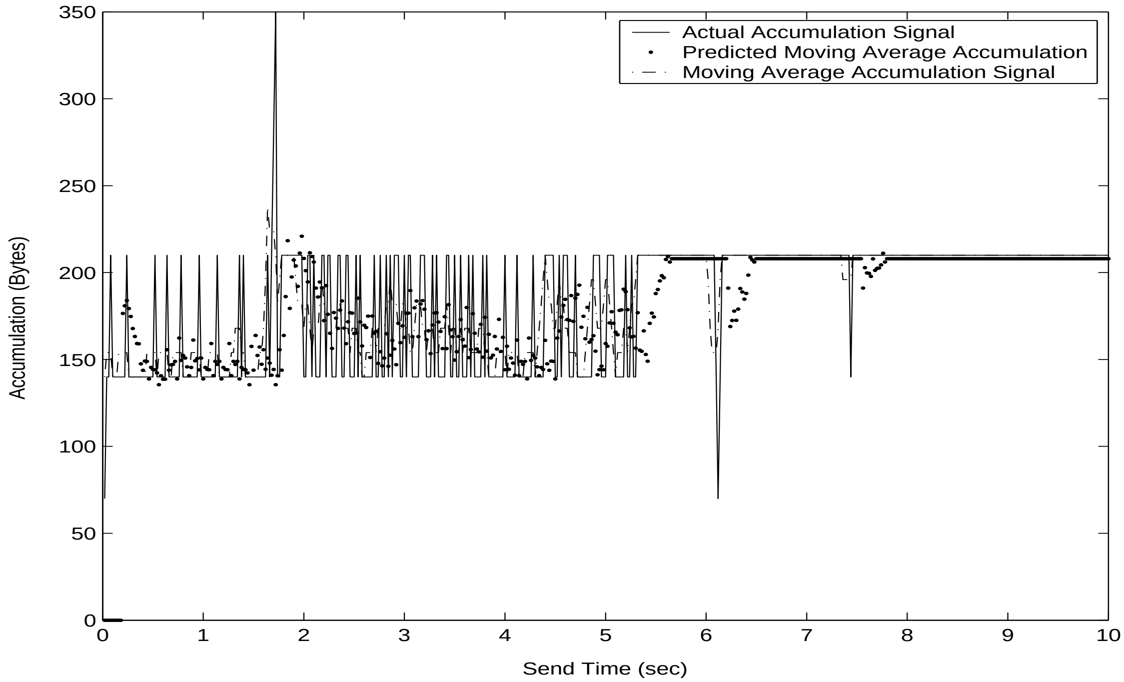

Fig. 44. Ten-Step-Ahead Prediction of Data-set3 Using Predictor2

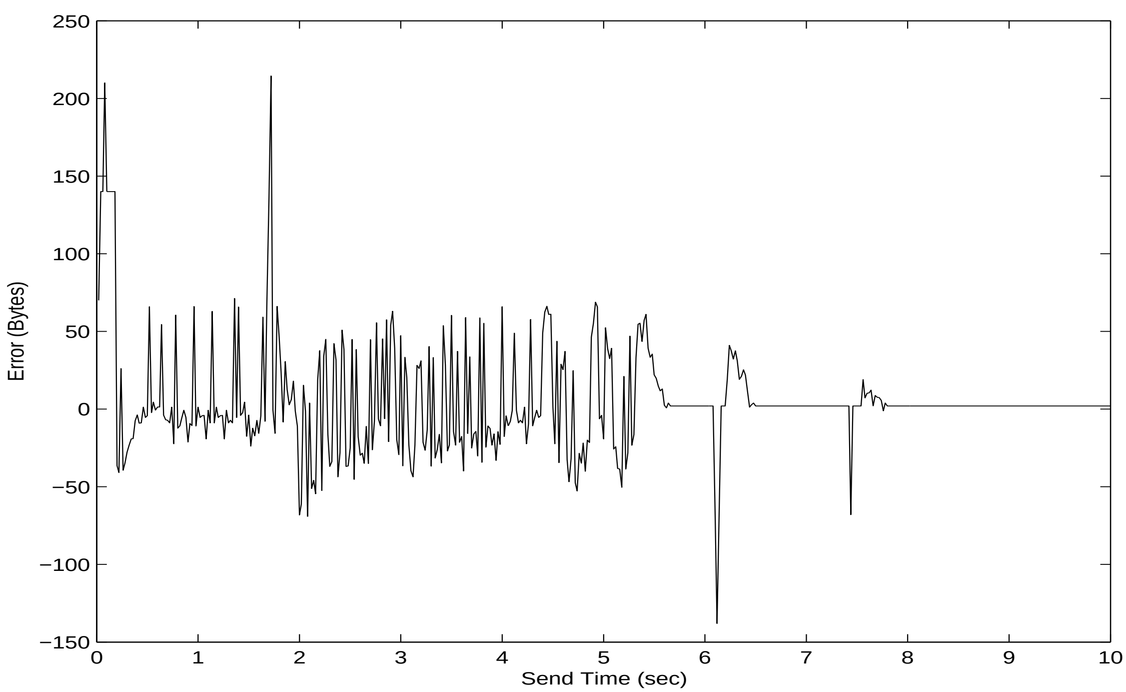

Fig. 45. Ten-Step-Ahead Prediction Error of Data-set3 Using Predictor2

Table XIII. Ten-step-ahead prediction results for Predictor1, Predictor2 and Predictor3

Data-Set	Predictor1		Predictor2		Predictor3	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset1	2.6	98.8	2.6	98.4	2.5	95.8
Dataset2	4.5	114.4	4.5	113.7	4.4	110.0
Dataset3	2.3	83.3	2.3	84.2	2.2	82.4
Dataset4	1.7	79.6	1.7	79.7	1.7	77.8
Dataset5	15.1	126.3	14.6	122.2	13.5	113.0
Dataset6	5.7	175.0	5.4	167.0	4.8	149.5

Similarly, for the flow between a node in Texas A&M and gtidsl node in planetlab the prediction errors of the predictors for different data-sets is presented in Table XIV. Figures 46, 48 and 50 show the performance of the Predictor5 for Datasets7, 8 and 9. Figures 47, 49 and 51 give the error plot of the same traces. Prediction figures show that good ten-step-ahead prediction is achieved. But it can be seen that MSE_1 and MSE_2 has increased compared to SSP.

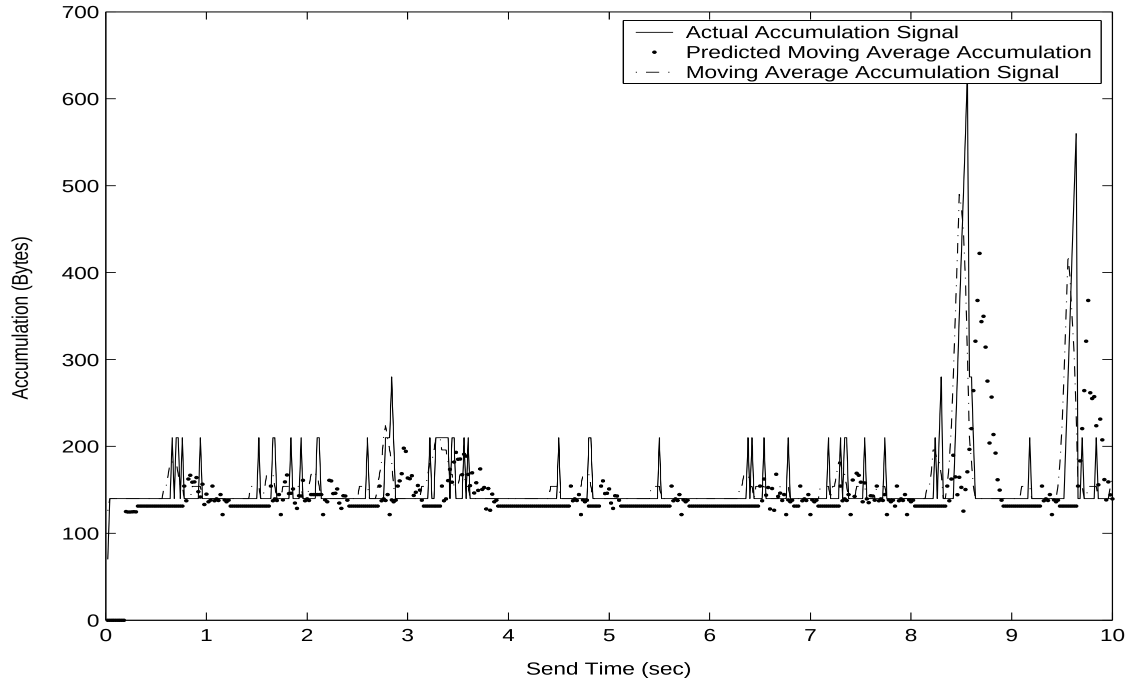

Fig. 46. Ten-Step-Ahead Prediction of Data-set7 Using Predictor5

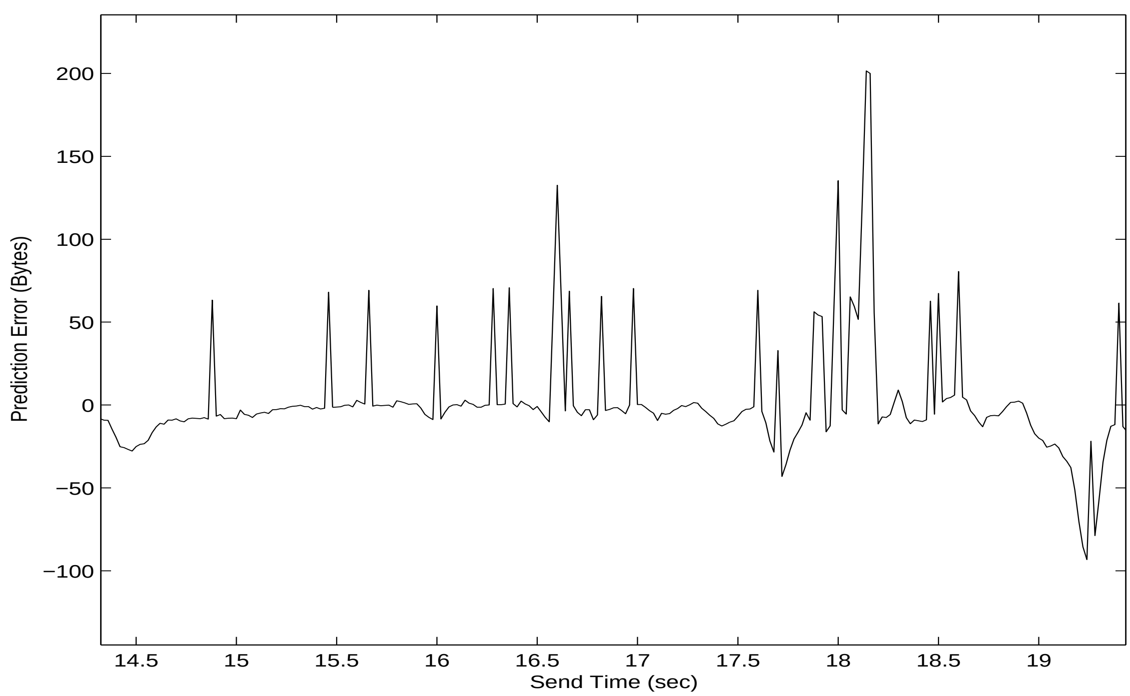

Fig. 47. Ten-Step-Ahead Prediction Error of Data-set7 Using Predictor5

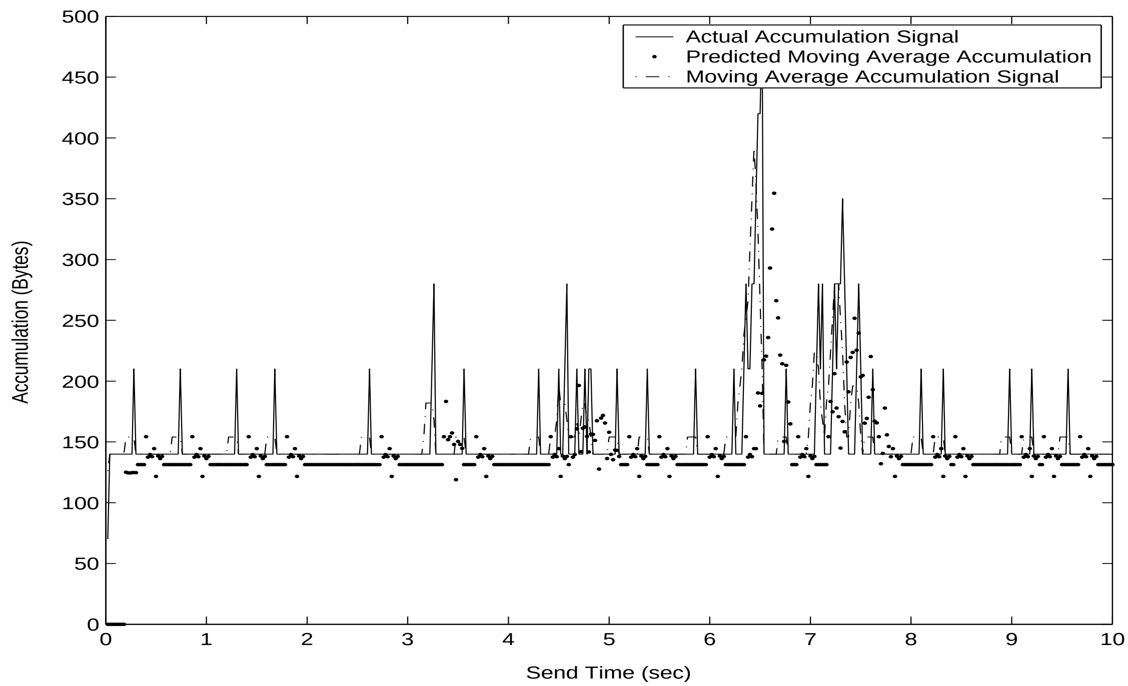

Fig. 48. Ten-Step-Ahead Prediction of Data-set8 Using Predictor5

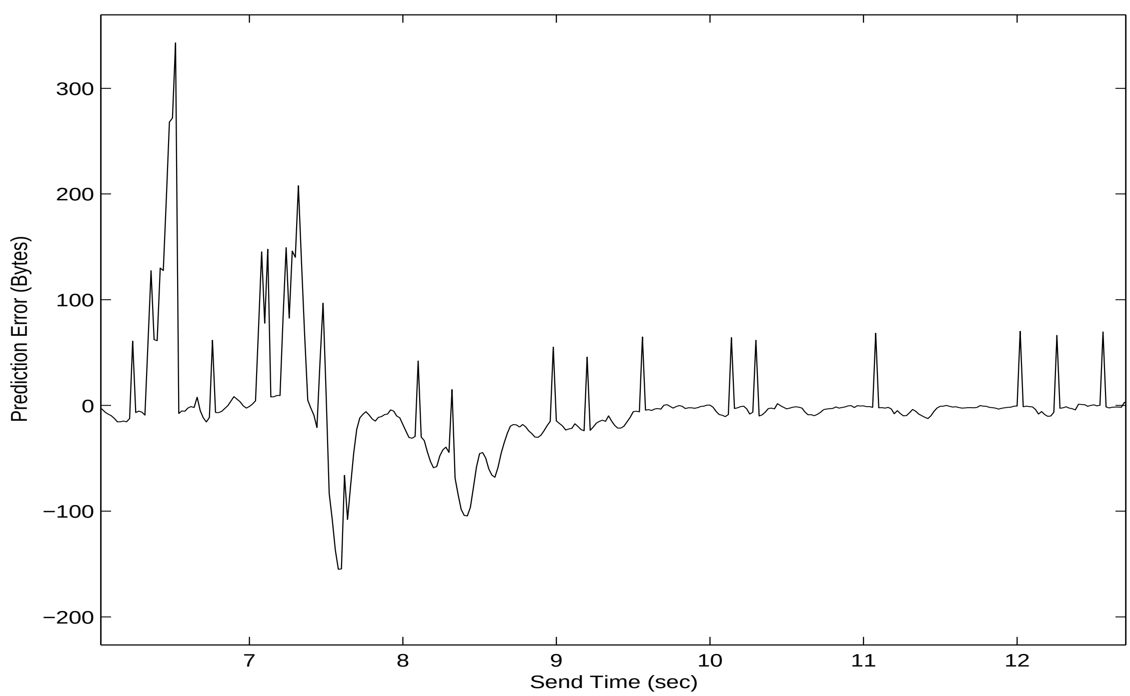

Fig. 49. Ten-Step-Ahead Prediction Error of Data-set8 Using Predictor5

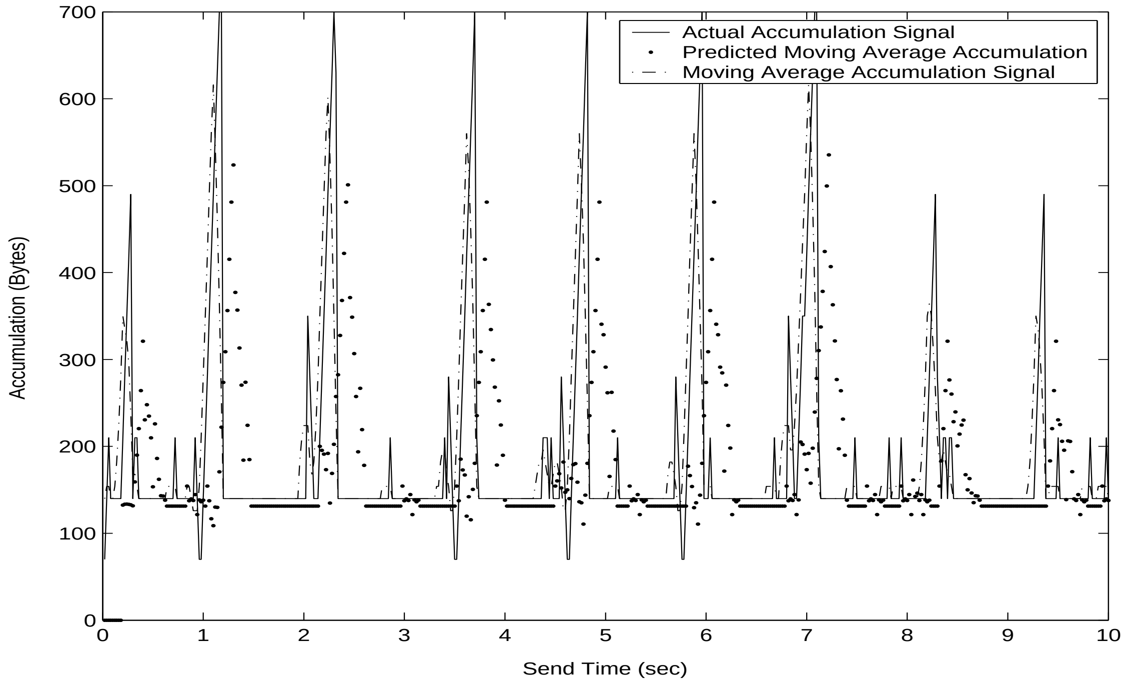

Fig. 50. Ten-Step-Ahead Prediction of Data-set9 Using Predictor5

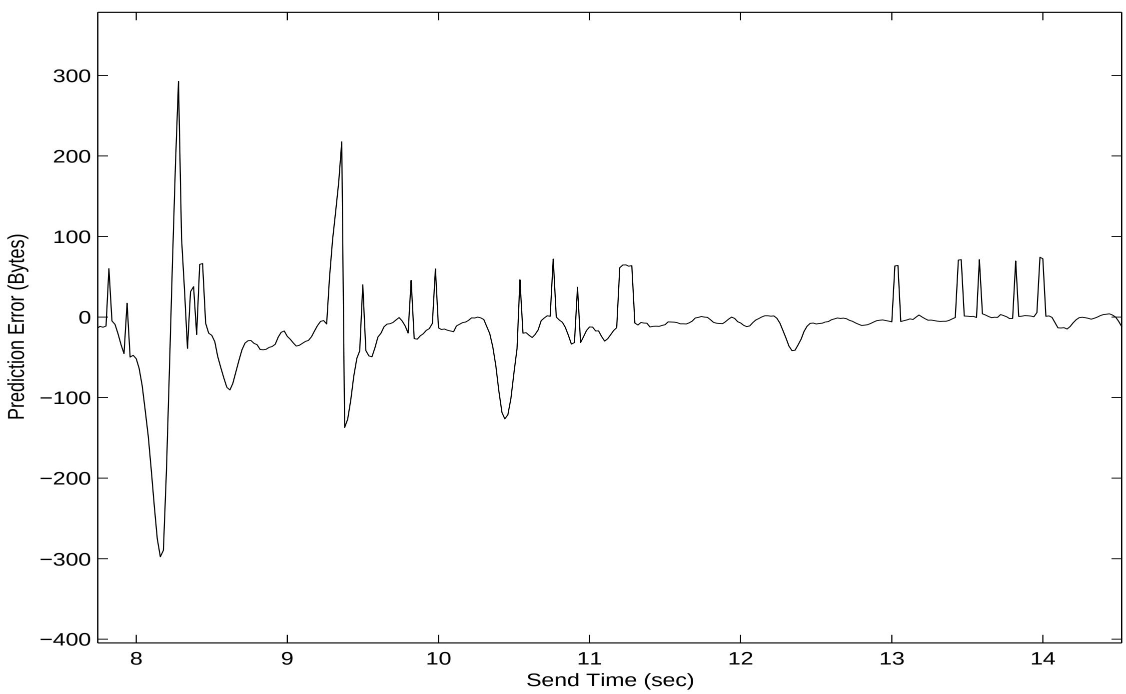

Fig. 51. Ten-Step-Ahead Prediction Error of Data-set9 Using Predictor5

Table XIV. Ten-step-ahead prediction results for Predictor4, Predictor5 and Predictor6

Data-Set	Predictor4		Predictor5		Predictor6	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset7	12.2	176.8	10.5	150.2	11.8	168.8
Dataset8	22.5	131.3	23.4	136.4	28.3	165.0
Dataset9	29.6	134.4	32.0	145.3	38.0	172.8
Dataset10	28.0	138.3	28.9	142.6	33.9	167.5
Dataset11	12.4	163.1	10.4	137.8	12.0	158.8
Dataset12	26.3	115.0	26.4	115.0	29.3	127.4
Dataset13	46.7	112.0	52.0	124.8	62.3	149.4

3. Thirteen-Step-Ahead Prediction

For the flow between gtidsl node and nbgisp node six-step-ahead means 0.52sec ahead prediction of the moving average of accumulation signal. The predictor used for thirteen-step-ahead prediction is the same developed for SSP. The performance of the predictors developed for different data-sets is presented in Table XV. Also Figures 52,54 and 56 show the performance of Predictor8. Whereas, Figures 53,55 and 57 show the error plots of the predictions. From table it can be seen that prediction errors are high as compared to SSP and six-step-ahead prediction. From this result it can be concluded that the predictor performance decreases as the prediction horizon is increased. This means that as we go on increasing the prediction horizon, the predictor's performance will go down.

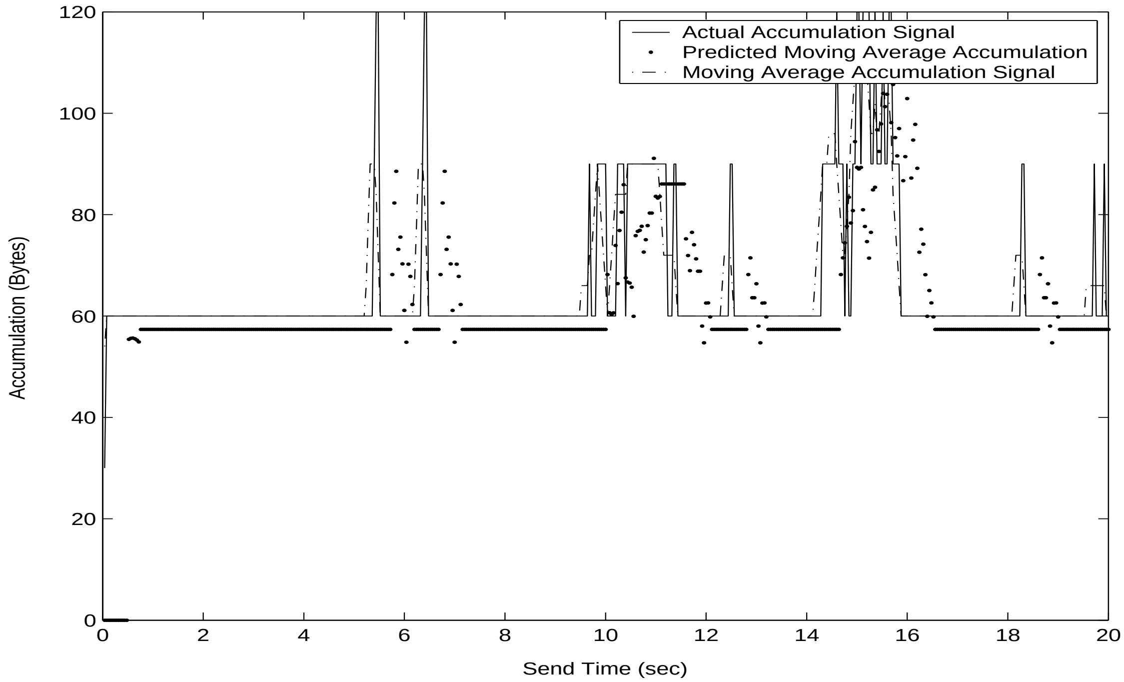

Fig. 52. Thirteen-Step-Ahead Prediction of Data-set14 Using Predictor8

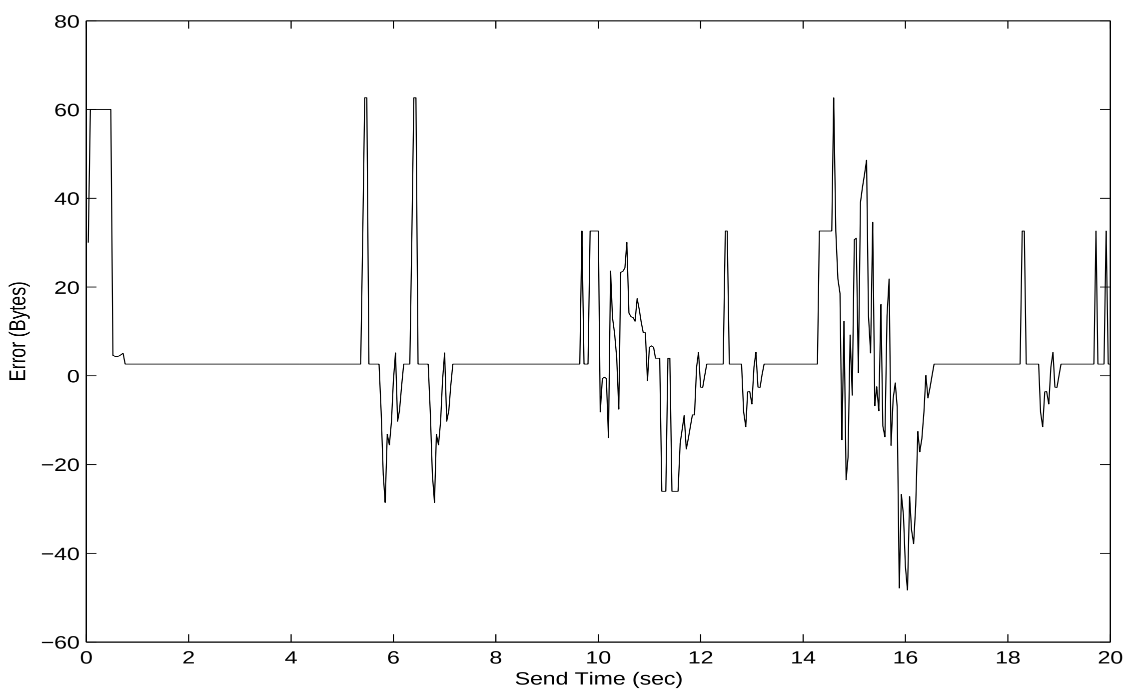

Fig. 53. Thirteen-Step-Ahead Prediction Error of Data-set14 Using Predictor8

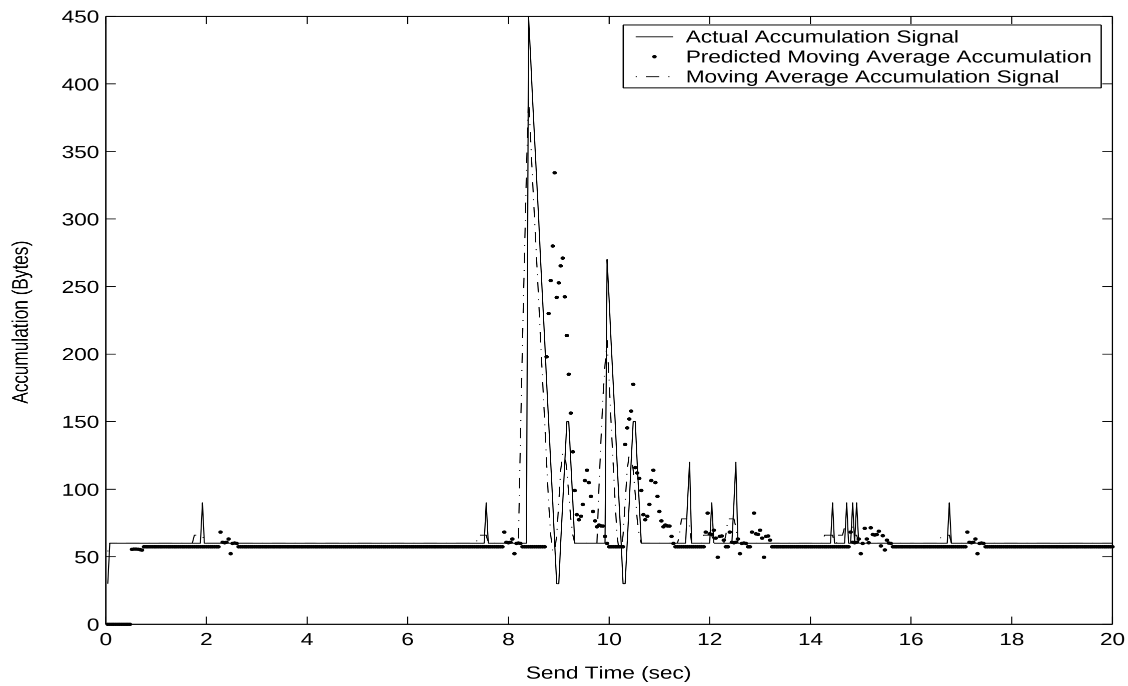

Fig. 54. Thirteen-Step-Ahead Prediction of Data-set15 Using Predictor8

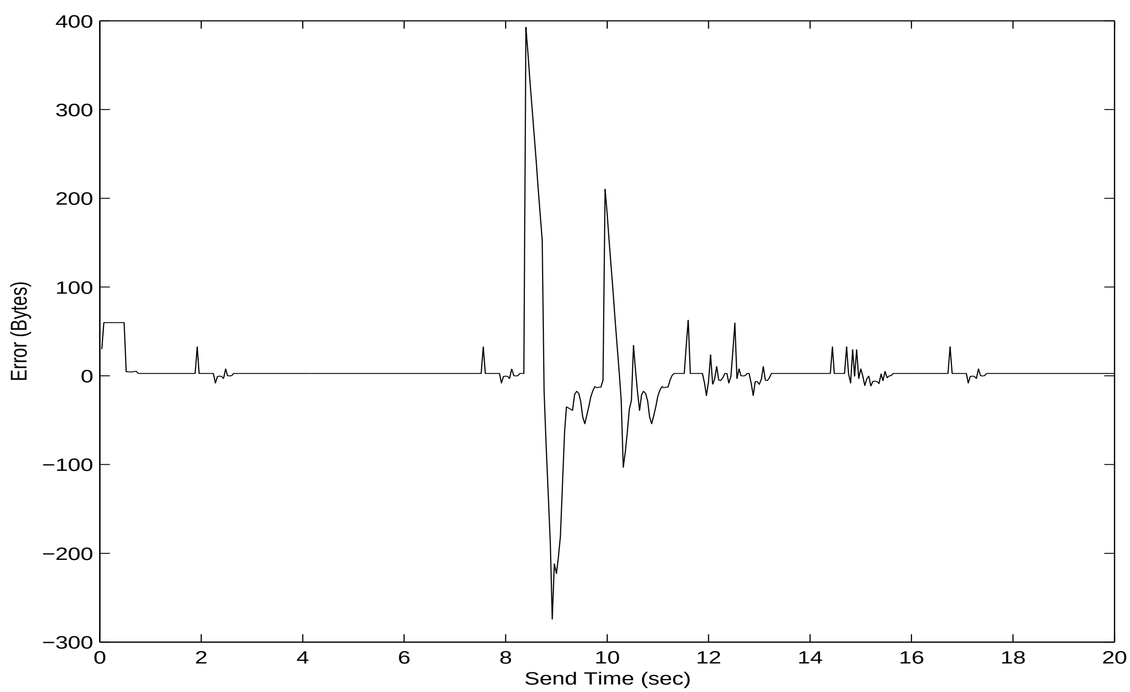

Fig. 55. Thirteen-Step-Ahead Prediction Error of Data-set15 Using Predictor8

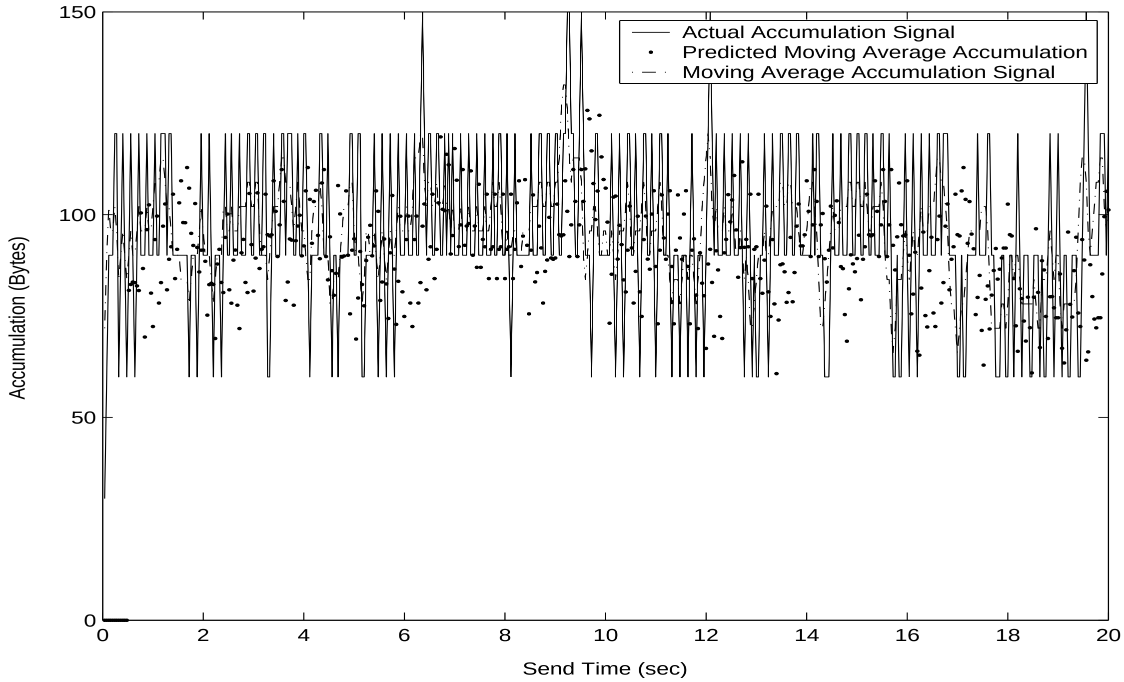

Fig. 56. Thirteen-Step-Ahead Prediction of Data-set16 Using Predictor8

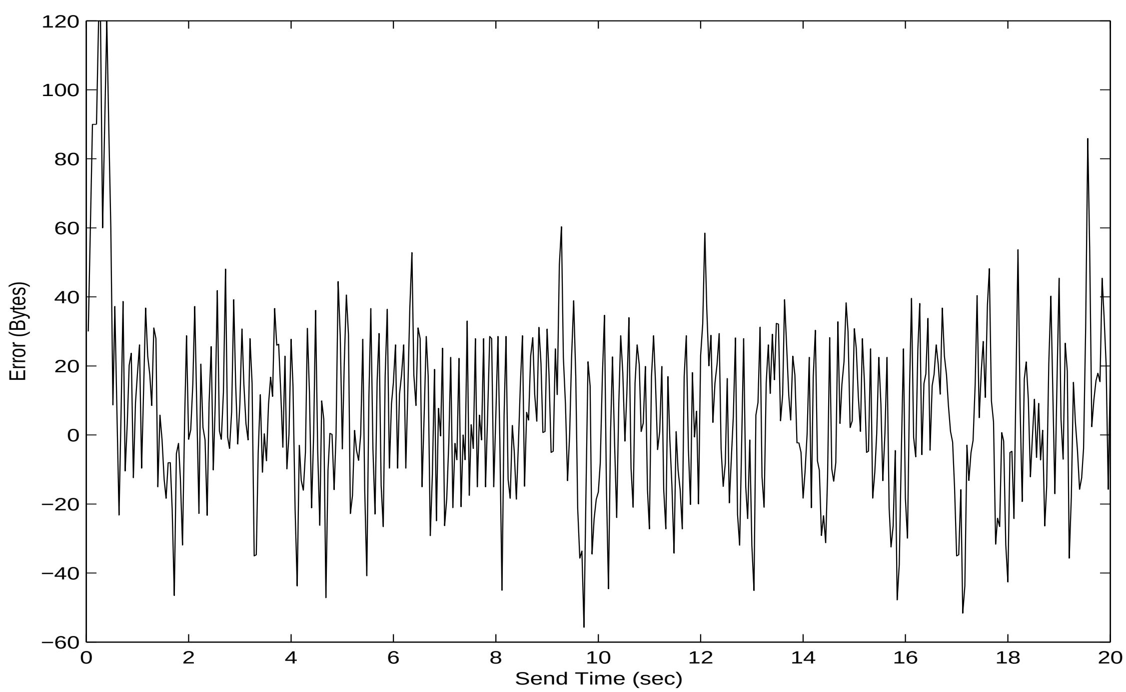

Fig. 57. Thirteen-Step-Ahead Prediction Error of Data-set16 Using Predictor8

Table XV. Thirteen-step-ahead prediction results for Predictor7, Predictor8 and Predictor9

Data-Set	Predictor7		Predictor8		Predictor8	
	MSE_1(%)	MSE_2(%)	MSE_1(%)	MSE_2(%)	MSE_1(%)	MSE_2(%)
Dataset14	3.6	127.9	3.6	130.7	3.5	126.9
Dataset15	28.0	155.8	25.6	142.4	26.1	145.3
Dataset16	6.6	146.5	6.0	124.0	5.8	127.3
Dataset17	8.4	141.5	8.0	133.0	7.2	120.4
Dataset18	59.3	69.4	57.2	66.4	61.5	71.5

4. Twenty-Five-Step-Ahead Prediction

The AR model used for twenty-five-step-ahead prediction is the same developed for SSP. Twenty-five-step-ahead prediction in this research means 0.5 sec ahead prediction of the moving average accumulation. For the flow between a node in Texas A&M and a pli_pa node in planet-lab the prediction errors of the predictors for different files is presented in table XVI. Figures 58,60 and 62 show the performance of the Predictor2 for different traces. Figures 59,61 and 63 give the error plot of the same traces. From the Table XVI it is observed that the MSE_2 in some of the cases goes beyond 200%, which indicates poor predictor performance. But, in most of the cases the predictors give satisfactory prediction. Again, for Dataset1 moving average accumulation signal is plotted. It can be concluded that the predicted values follow more closely the moving average signal as expected.

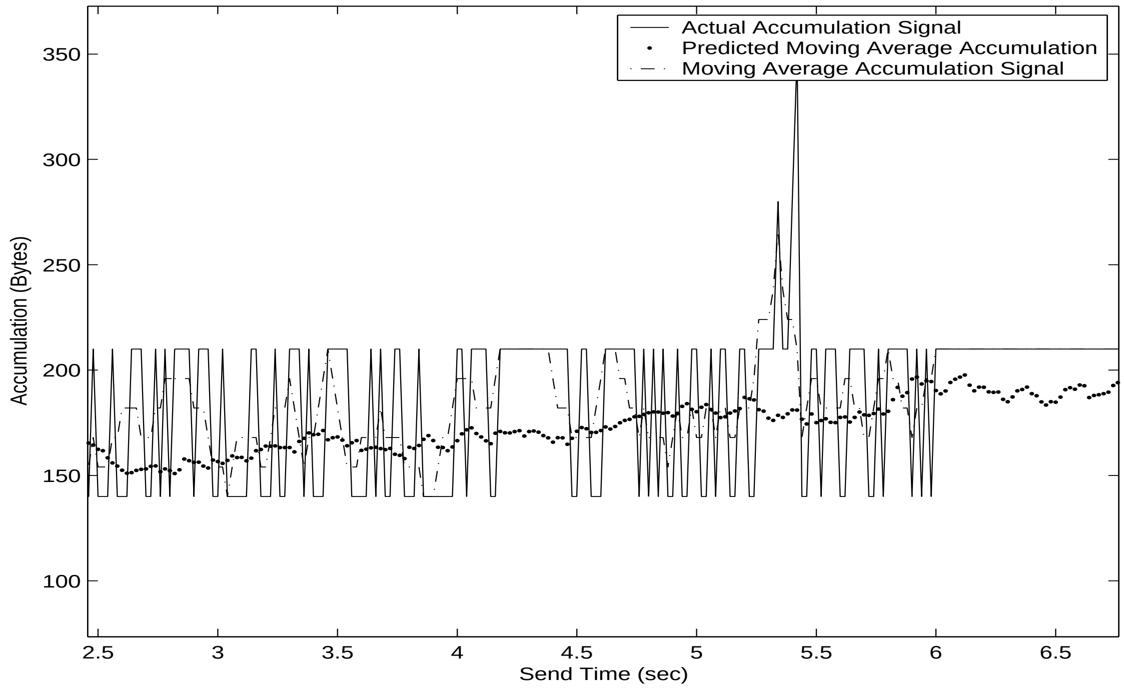

Fig. 58. Twenty-Five-Step-Ahead Prediction of Data-set1 Using Predictor2

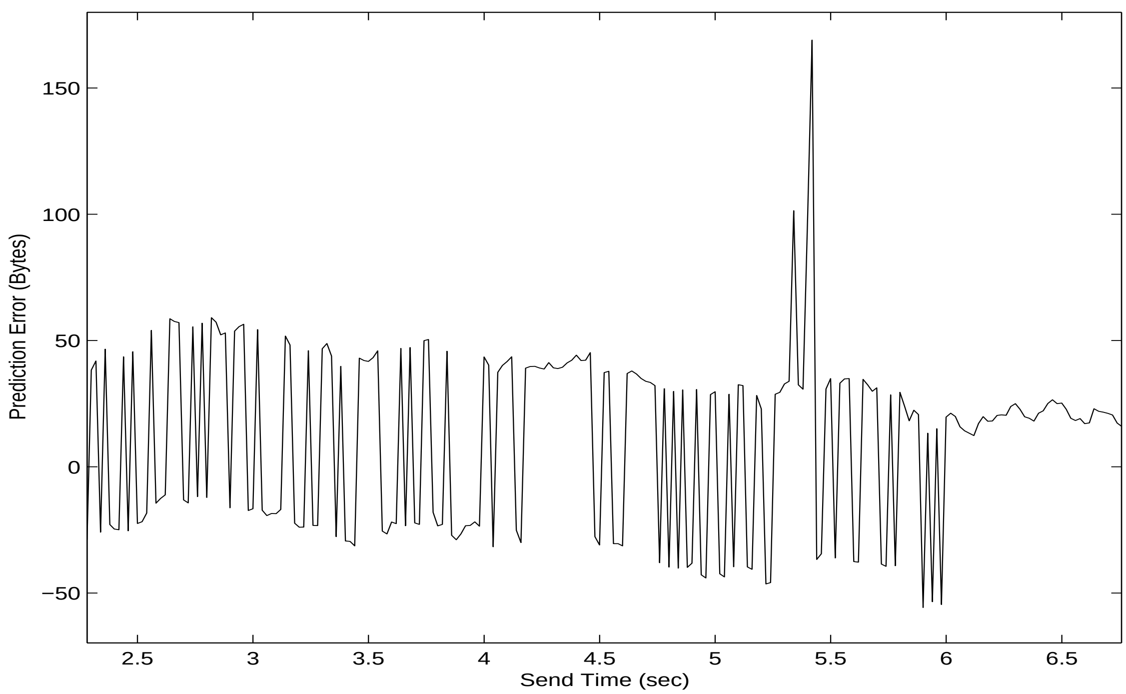

Fig. 59. Twenty-Five-Step-Ahead Prediction Error of Data-set1 Using Predictor2

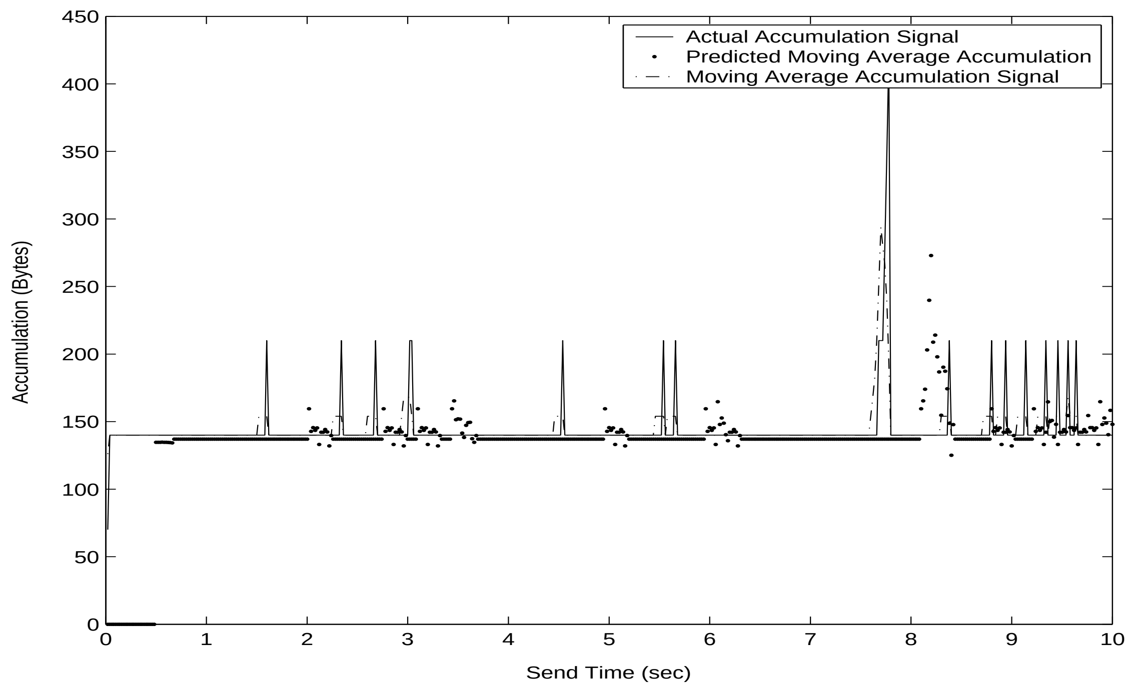

Fig. 60. Twenty-Five-Step-Ahead Prediction of Data-set2 Using Predictor2

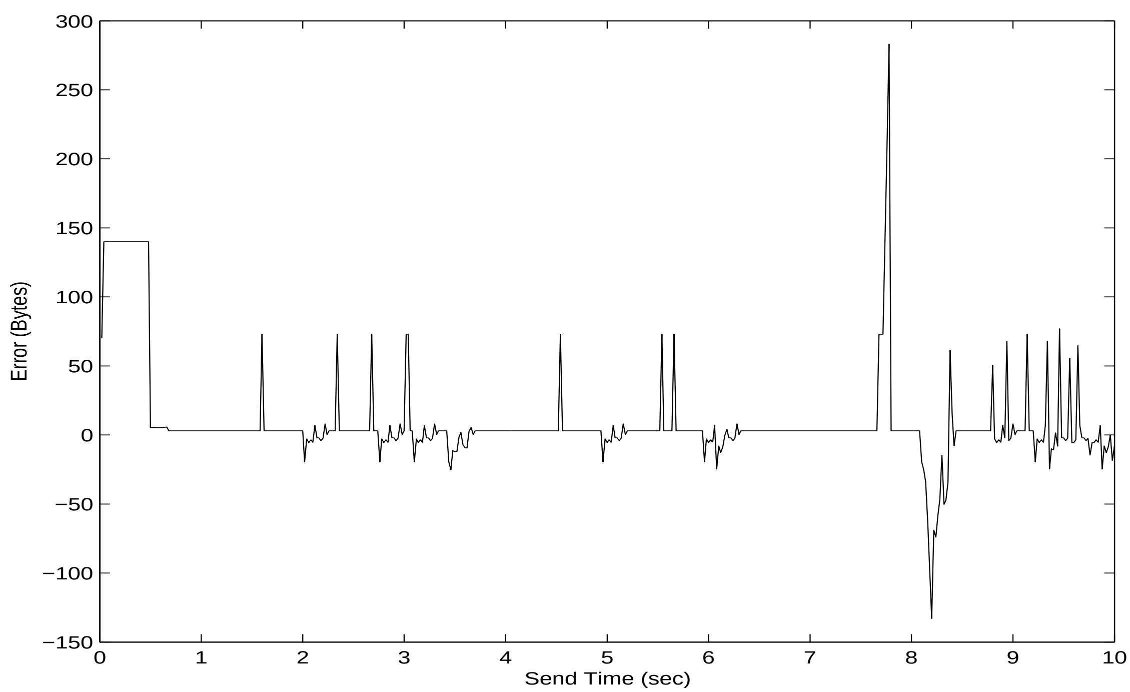

Fig. 61. Twenty-Five-Step-Ahead Prediction Error of Data-set2 Using Predictor2

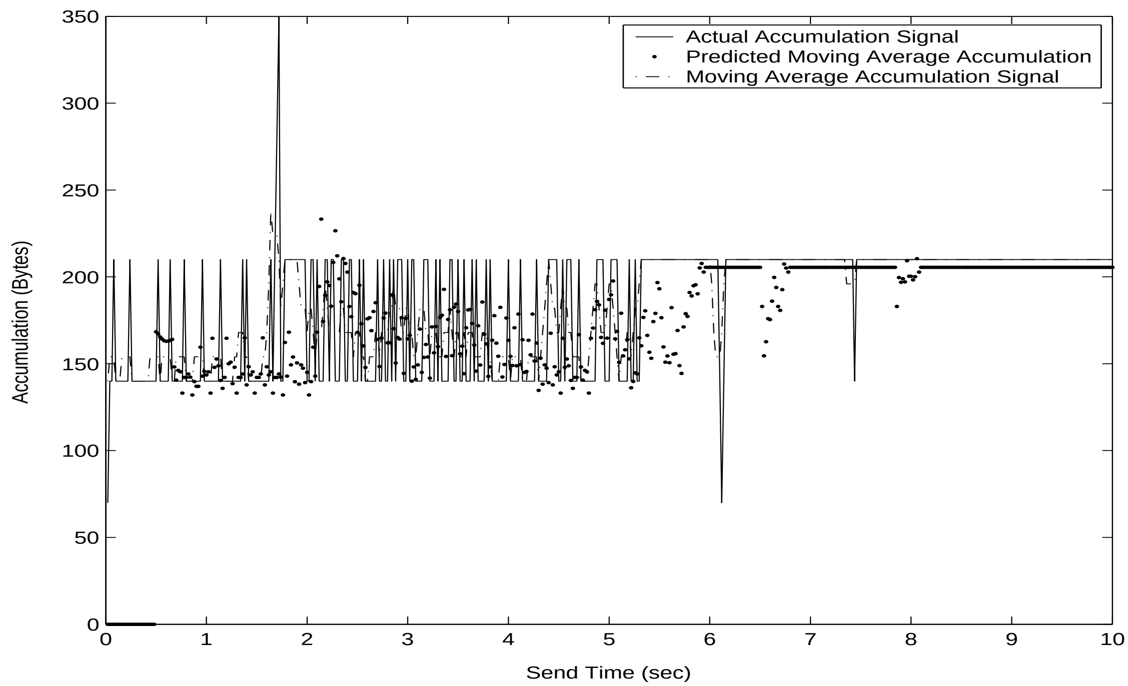

Fig. 62. Twenty-Five-Step-Ahead Prediction of Data-set3 Using Predictor2

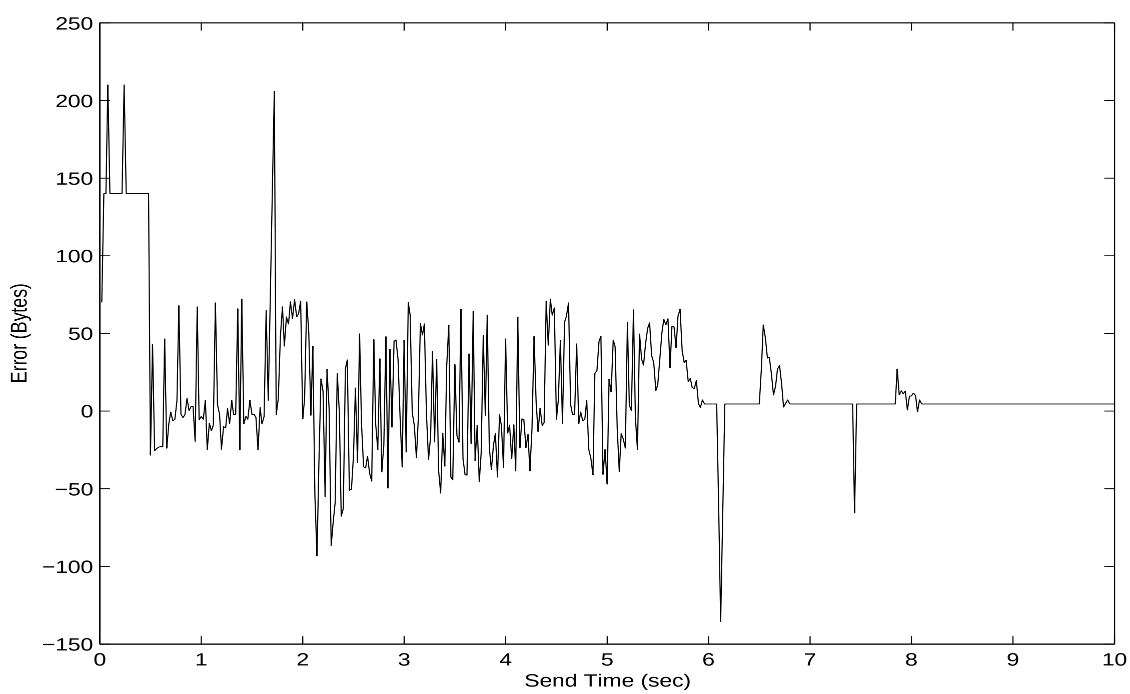

Fig. 63. Twenty-Five-Step-Ahead Prediction Error of Data-set3 Using Predictor2

Table XVI. Twenty-Five-step-ahead prediction results for Predictor1, Predictor2 and Predictor3

Data-Set	Predictor1		Predictor2		Predictor3	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset1	3.7	143.0	3.8	145.3	3.7	142.7
Dataset2	6.0	152.5	6.2	155.1	6.0	150.0
Dataset3	3.2	117.5	3.3	122.0	3.2	119.5
Dataset4	2.5	116.9	2.7	122.5	2.6	120.0
Dataset5	19.2	160.6	19.0	159.1	17.8	149.2
Dataset6	6.3	193.2	6.3	193.9	6.0	186.5

Similarly, for the flow between a node in Texas A&M and gtidsl node in planet-lab the prediction errors of the predictors for different files is presented in table XVII. Figures 64,66 and 68 show the performance of the Predictor5 for different traces. Figure 65,67 and 69 give the error plot of the same traces. From Table XVII we can see that for Data-set6 MSE_2 goes beyond 200%, indicating poor predictor performance in this case. But again in most of the cases predictor gives satisfactory results. By comparing the results of SSP, ten-step-ahead prediction and twenty-five-step-ahead prediction it can be concluded that as the prediction horizon is increased the predictor performance goes down.

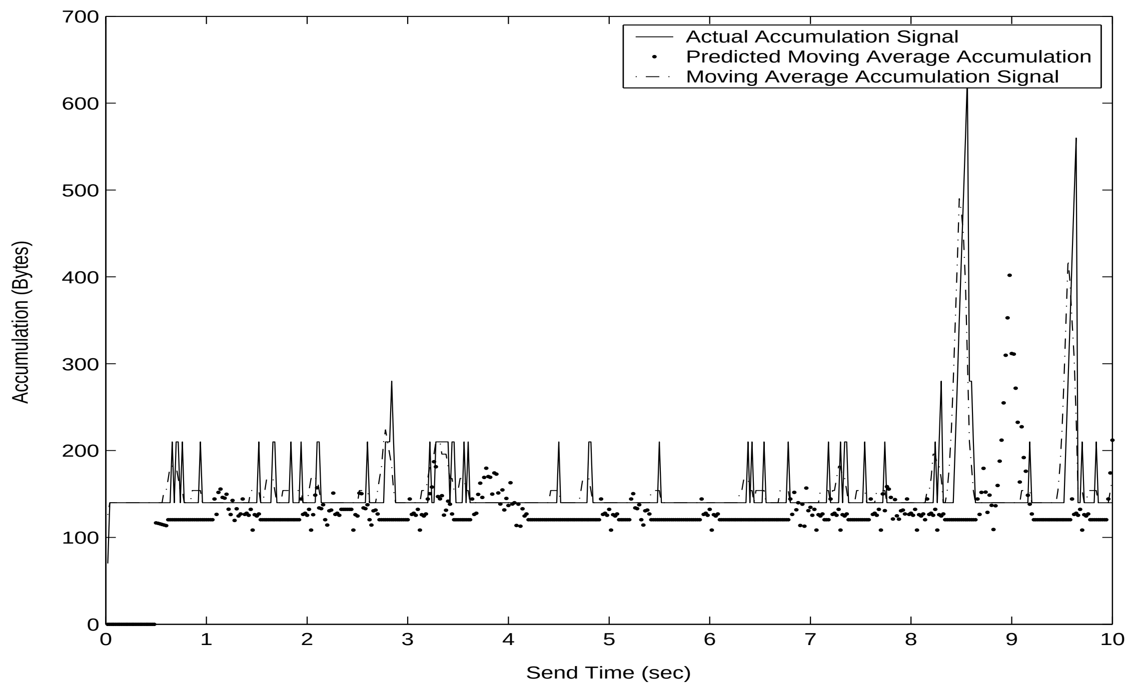

Fig. 64. Twenty-Five-Step-Ahead Prediction of Data-set7 Using Predictor5

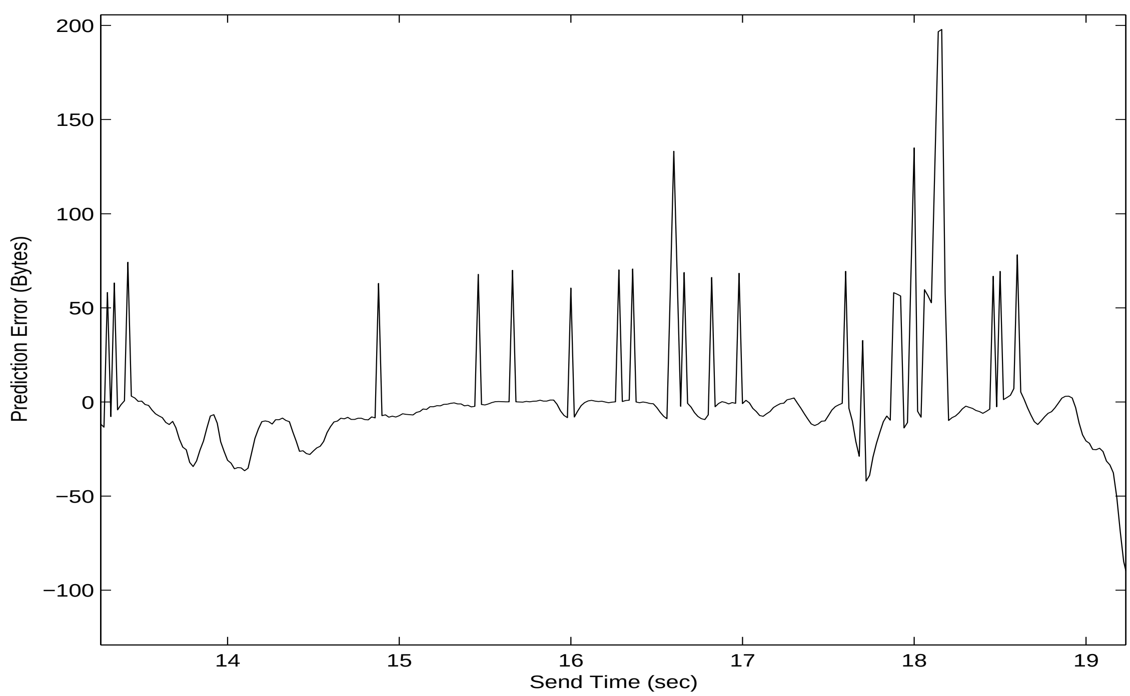

Fig. 65. Twenty-Five-Step-Ahead Prediction Error of Data-set7 Using Predictor5

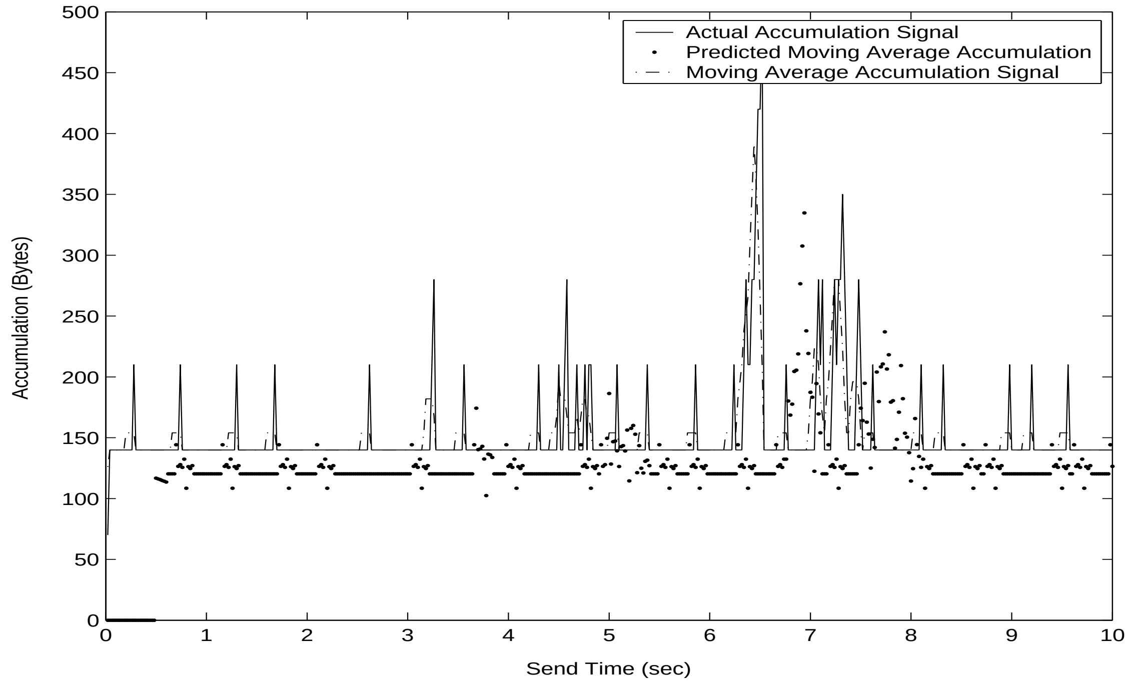

Fig. 66. Twenty-Five-Step-Ahead Prediction of Data-set8 Using Predictor5

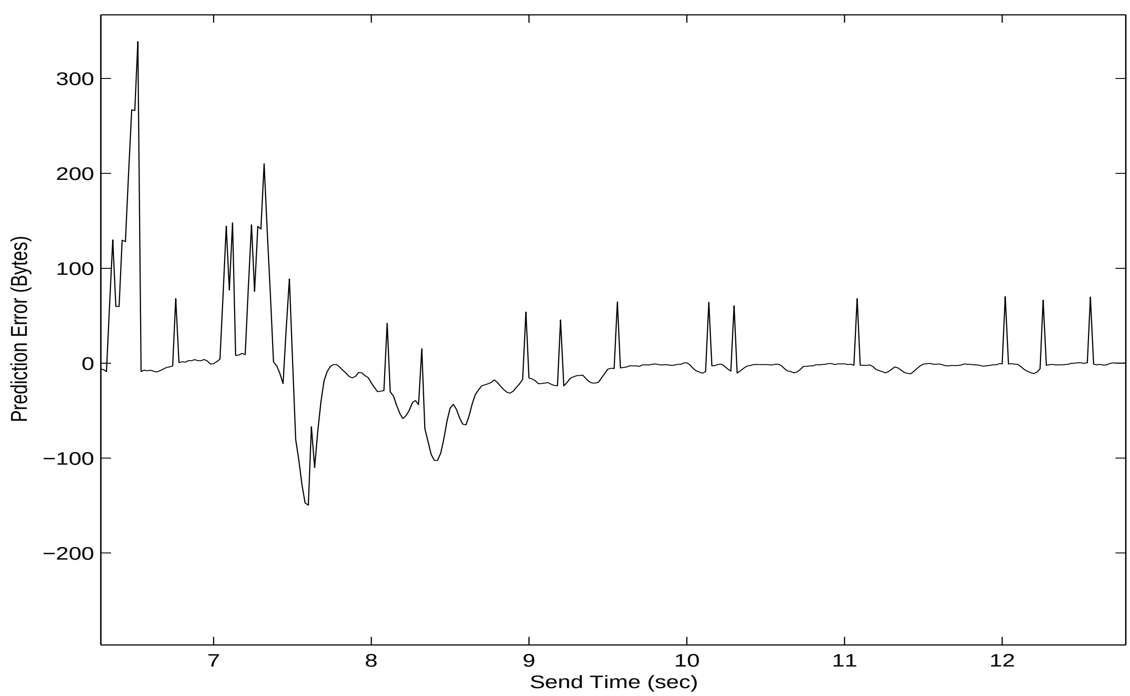

Fig. 67. Ten-Step-Ahead Prediction Error of Data-set8 Using Predictor5

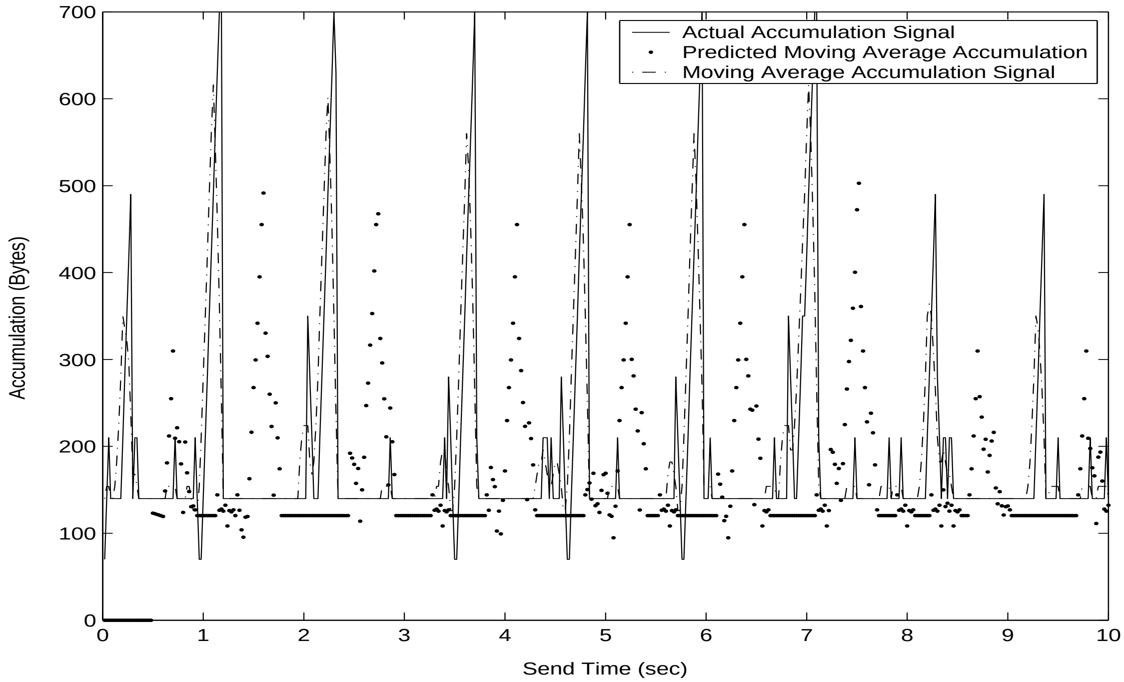

Fig. 68. Ten-Step-Ahead Prediction of Data-set9 Using Predictor5

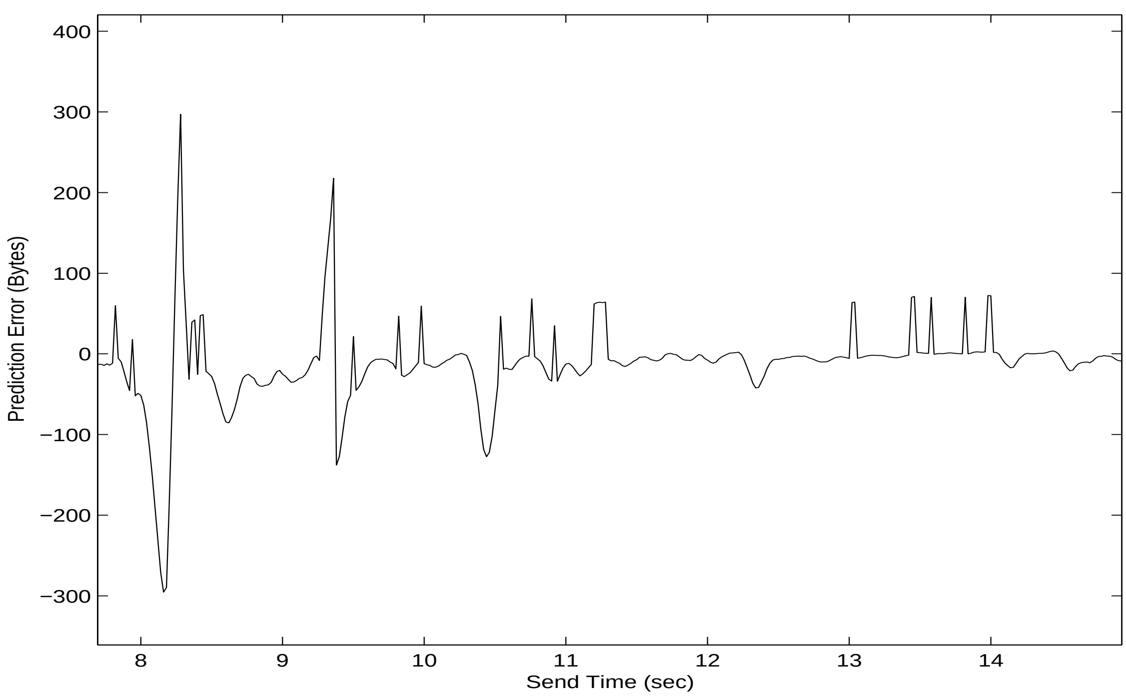

Fig. 69. Twenty-Five-Step-Ahead Prediction Error of Data-set9 Using Predictor5

Table XVII. Twenty-Five-step-ahead prediction results for Predictor4, Predictor5 and Predictor6

Data-Set	Predictor4		Predictor5		Predictor6	
	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$	$MSE_1(\%)$	$MSE_2(\%)$
Dataset7	22.4	256.4	13.4	191.4	13.3	190.2
Dataset8	32.8	191.4	27.0	157.6	30.3	176.6
Dataset9	39.8	181.0	36.8	167.2	42.8	194.5
Dataset10	37.3	184.7	33.6	166.3	38.9	192.6
Dataset11	23.4	180.2	14.7	154.4	16.6	168.3
Dataset12	35.3	153.6	30.8	134.3	35.2	153.4
Dataset13	52.2	125.2	53.8	129.0	63.7	152.8

F. Chapter Summary

This chapter demonstrates training, testing and validation data sets used for predictor development. Performance metrics used for developing the predictors is also discussed. The chapter gives SSP and MSP results of the predictors developed using linear technique. Performance of the predictors developed is discussed and tabulated in terms of the performance metrics discussed. The models gave good prediction on most of the test cases but MSP is not as accurate as SSP and fails on several test cases as the prediction horizon is increased gradually. Acceptable multi-step prediction up to 0.5 sec has been achieved.

CHAPTER VI

IMPACT OF WIRELESS LOSSES ON PREDICTOR ACCURACY

A. Introduction

The previous chapters dealt with the development of SSP and MSP. This chapter discusses about the impact of channel fading losses on the MSP (developed in the previous chapter) accuracy. As discussed in Chapter II, wireless errors are random errors and mostly depends on the surrounding conditions and the signals present in the wireless channel.

B. Impact of Wireless Losses on the Performance of the Predictors Developed

This section discusses the impact of wireless fading losses on the performance of predictors developed. Three predictors were developed for each of the flow discussed in section 4.D.2.2. The section is divided in three subsections, each discussing in detail the impact of wireless losses in each of the above flows. Impact of wireless losses is studied for 0.5sec ahead predictors.

1. Impact of Wireless Losses on the Predictor Performance for the Flow between NIML (TAMU) and pli_pa Node Present in the Planet-lab, Princeton

For each of the predictors developed the wireless loss is increased from 0 to 15% and the impact on the predictor accuracy is studied. Tables XVIII to XXIII give the predictor performance with increasing wireless losses on various data-sets. From the tables it can be seen that the wireless losses has major impact on the MSE_1 while MSE_2 remains nearly the same in all the data-sets. Table XVIII shows more than 100% increase in MSE_1 for all the predictors with 15% increase in wireless loss.

Table XVIII. Impact of wireless losses in data-set1 on the performance of Predictor1, Predictor2 and Predictor3

Wireless loss(%)	Predictor1			Predictor2			Predictor3		
	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	3.7	143.0	0	3.8	145.3	0	3.7	142.4	0
2	4.2	137.2	13.5	4.3	138.7	13.2	4.1	134.2	10.8
3	4.5	136.6	21.6	4.5	137.5	18.4	4.4	133.6	18.9
5	5.0	135.5	35.1	5.1	135.9	34.2	4.9	131.3	32.4
10	6.4	140.6	73.0	6.6	143.6	73.7	6.3	137.7	70.3
15	7.4	132.4	100.0	7.7	136.7	102.6	7.3	130.7	97.3

Note: All values of MSE's are in % and ΔE_1 represents % change in MSE_1 with respect to 0% wireless loss MSE_1.

Similar observation can be made for data-set3 and data-set4 as well. This increase in error can be attributed to the randomness of the wireless errors. In case of data-set2 the increase is nearly 40%. It can be seen that none of the predictors is able to capture the randomness of channel fading errors.

2. Impact of Wireless Losses on the Predictor Performance for the Flow between NIML (TAMU) and gtidsl Node Present in the Planet-lab, Princeton

Wireless losses are increased from 0% to 15% for each of the data-set. The impact of these wireless losses can be seen from Tables XXIV to XXX. Again, it can be observed that MSE_2 does not vary much and its variation is totally random. Nearly 100% increase in MSE_1 is observed in case of data-set11 for Predictor5 and Predictor6 and for Predictor 4 the increase is 45%. This shows that wireless losses has more impact

Table XIX. Impact of wireless losses in data-set2 on the performance of Predictor1, Predictor2 and Predictor3

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	6.0	152.5	0	6.2	155.1	0	6.0	150.0	0
2	6.7	150.5	11.7	6.8	153.6	9.7	6.6	148.5	10.0
3	6.9	158.8	15.0	7.0	161.4	12.9	6.8	156.3	13.3
5	7.4	156.4	23.0	7.5	159.3	21.0	7.3	154.2	21.7
10	8.3	139.5	38.3	8.5	143.0	37.1	8.1	136.2	35.0
15	8.4	120.4	40.0	8.7	125.9	40.3	8.3	120.3	38.3

Table XX. Impact of wireless losses in data-set3 on the performance of Predictor1, Predictor2 and Predictor3

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	3.2	117.5	0	3.3	122.0	0	3.2	119.6	0
2	3.7	119.5	15.6	3.8	123.6	15.1	3.7	120.1	15.6
3	4.0	121.3	25.0	4.2	125.0	27.3	4.0	120.8	25.0
5	4.9	128.9	53.1	5.1	132.9	54.6	4.9	127.6	53.1
10	5.7	118.0	78.1	5.8	121.0	75.8	5.5	115.5	71.9
15	6.8	117.8	112.5	6.9	120.7	109.1	6.6	114.4	106.3

Table XXI. Impact of wireless losses in data-set4 on the performance of Predictor1, Predictor2 and Predictor3

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	2.5	116.9	0	2.7	122.7	0	2.6	120.0	0
2	3.0	113.6	20.0	3.2	119.0	18.4	3.1	116.2	19.2
3	3.2	111.4	28.0	3.4	116.8	25.9	3.3	112.9	26.9
5	3.8	113.5	52.0	3.9	117.5	44.4	3.8	113.6	46.2
10	5.1	124.2	104.0	5.3	128.7	96.3	5.0	122.2	92.3
15	6.3	124.7	152.0	6.4	126.5	137.0	6.1	120.8	134.6

Table XXII. Impact of wireless losses in data-set5 on the performance of Predictor1, Predictor2 and Predictor3

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	19.2	160.6	0	19.0	159.1	0	17.8	149.2	0
2	19.8	159.8	3.1	17.8	149.2	3.2	18.3	148.0	2.8
3	20.4	160.3	6.3	20.2	158.8	6.3	18.9	148.9	6.2
5	21.0	161.3	9.4	20.8	160.2	9.5	19.5	150.2	9.6
10	21.7	158.7	13.0	21.6	157.6	13.7	20.2	147.3	13.5
15	20.6	155.7	7.3	20.4	7.4	6.2	19.3	145.9	8.4

Table XXIII. Impact of wireless losses in data-set6 on the performance of Predictor1, Predictor2 and Predictor3

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	6.3	193.2	0	6.3	193.9	0	6.0	186.5	0
2	6.7	181.2	6.4	6.8	182.8	7.9	6.4	173.9	6.7
3	7.2	180.1	14.3	7.2	180.3	14.3	6.8	171.4	13.3
5	7.0	171.9	11.1	7.1	174.1	12.7	6.8	165.6	13.3
10	9.4	163.7	49.3	9.6	165.8	52.4	9.1	158.1	51.7
15	9.7	156.8	54.0	9.9	160.7	57.1	9.4	152.8	56.7

on the accuracy of a good predictor and less impact if the predictor is less accurate. For data-set7 the increase in error is nearly 25%, if the wireless losses is increased to 15%. From Tables XXVII and XXX it can be concluded that wireless losses has very less impact on the predictor performance, if the predictors developed are poor.

3. Impact of Wireless Losses on the Predictor Performance for the Flow between gtidsl and nbgisp Nodes Present in the Planet-lab, Princeton

Wireless losses are increased from 0% to 15% for each of the data-set. The impact of these wireless losses can be seen from Tables XXXI to XXXV. From Table XXXI it can be observed that increase of 15% wireless losses in data-set14 has lead to increase in nearly 150% error in the prediction. Whereas, this increase in wireless losses has increased the error in prediction to nearly 80% in case of data-set16 and data-set17. Once again, from the predictors performance on data-set18, it can be concluded that if the predictor developed is poor then wireless losses has very less impact on the prediction errors.

Table XXIV. Impact of wireless losses in data-set7 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor4			Predictor5			Predictor6		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	22.4	256.4	0	13.4	191.4	0	13.3	190.2	0
2	22.8	258.2	1.8	14.0	184.8	4.5	14.0	184.4	5.3
3	23.0	257.5	2.7	14.3	184.1	6.7	14.2	183.5	6.8
5	23.6	260.2	5.4	15.0	180.6	11.9	15.0	180.7	12.8
10	25.0	263.6	11.6	16.6	175.3	23.9	16.7	176.2	25.6
15	25.1	258.6	12.1	16.7	171.2	24.6	16.7	170.5	25.6

Table XXV. Impact of wireless losses in data-set8 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor4			Predictor5			Predictor6		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	32.8	191.5	0	27.0	157.6	0	30.3	176.5	0
2	33.0	189.2	0.6	27.2	155.2	0.7	30.3	173.7	0.0
3	33.5	186.9	2.1	27.7	154.5	2.6	31.0	173.2	2.3
5	34.2	184.2	4.3	28.7	154.5	6.3	32.3	173.9	6.6
10	34.3	185.8	4.6	29.2	158.2	8.1	32.9	178.4	8.6
15	35.6	183.7	8.5	30.9	159.6	14.4	34.9	180.2	15.2

Table XXVI. Impact of wireless losses in data-set9 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	39.8	181.0	0	36.8	167.2	0	42.8	194.5	0
2	39.0	182.4	-2.0	35.6	166.5	-3.3	41.2	192.2	-3.7
3	39.8	182.2	0.0	36.8	168.8	0.0	42.8	196.1	0.0
5	39.5	180.9	-0.8	36.3	165.9	-1.4	42.0	192.3	-1.9
10	41.0	182.1	3.0	38.1	169.1	3.5	43.8	194.8	2.3
15	43.2	172.5	8.5	41.1	163.6	11.7	47.7	190.0	11.4

Table XXVII. Impact of wireless losses in data-set10 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	37.7	184.7	0	33.6	166.3	0	38.9	192.6	0
2	37.0	186.7	-0.8	33.2	167.3	-1.2	38.3	193.2	-1.5
3	37.6	181.8	0.8	33.9	164.1	0.9	39.3	189.9	1.0
5	37.0	184.0	-0.8	33.0	164.3	-1.8	38.0	188.9	-2.3
10	38.3	177.4	2.7	34.6	160.6	3.0	39.9	184.6	2.6
15	36.6	182.3	-1.9	32.2	160.4	-4.2	36.3	180.6	-6.7

Table XXVIII. Impact of wireless losses in data-set11 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor4			Predictor5			Predictor6		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	23.3	180.2	0	14.7	154.4	0	14.6	168.3	0
2	24.0	181.2	2.6	15.4	153.7	4.8	15.3	165.9	4.8
3	24.3	179.8	3.9	15.6	157.6	6.1	15.7	166.1	7.5
5	24.8	180.0	6.0	16.3	156.8	10.9	16.5	163.9	13.0
10	33.3	179.5	42.3	27.8	149.0	89.1	30.4	163.0	108.2
15	34.0	178.5	45.3	28.3	148.5	92.5	30.9	162.4	111.6

Table XXIX. Impact of wireless losses in data-set12 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	35.3	153.6	0	30.8	134.3	0	35.2	153.4	0
2	40.3	148.0	14.2	37.6	138.6	22.0	43.3	149.6	23.0
3	35.6	151.6	0.9	31.2	132.6	1.3	35.5	151.2	0.9
5	36.3	150.3	2.8	32.1	133.0	4.2	36.6	151.8	4.0
10	36.6	147.6	3.7	32.4	131.0	5.2	37.2	150.3	5.7
15	38.2	150.0	8.2	34.5	135.8	12.0	39.6	155.7	12.5

Table XXX. Impact of wireless losses in data-set13 on the performance of Predictor4, Predictor5 and Predictor6

Wireless	Predictor4			Predictor5			Predictor6		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	52.2	125.2	0	53.8	129.0	0	63.7	152.8	0
2	52.8	126.5	1.1	54.8	131.3	1.9	65.0	155.5	2.0
3	52.6	125.2	0.8	54.5	129.6	1.3	64.6	153.7	1.4
5	53.8	124.4	3.0	56.1	129.7	4.3	66.4	153.4	4.2
10	56.1	126.7	7.5	59.6	134.8	10.8	71.2	160.9	11.8
15	55.1	123.9	5.6	57.8	130.0	7.4	68.4	153.8	7.2

Table XXXI. Impact of wireless losses in data-set14 on the performance of Predictor7, Predictor8 and Predictor9

Wireless	Predictor7			Predictor8			Predictor9		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	3.6	127.9	0	3.6	130.7	0	3.5	126.9	0
2	6.3	189.3	75.0	6.7	201.2	86.1	6.1	185.4	74.3
3	6.6	187.7	83.3	7.0	199.0	94.4	6.4	182.5	82.9
5	6.8	177.5	88.9	7.2	188.6	100.0	6.5	171.7	85.7
10	7.9	176.4	119.4	8.4	188.0	133.3	7.6	168.4	117.1
15	9.2	166.0	155.6	9.6	173.3	166.7	8.7	156.9	148.6

Table XXXII. Impact of wireless losses in data-set15 on the performance of Predictor7, Predictor8 and Predictor9

Wireless	Predictor4			Predictor5			Predictor6		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	28.0	155.8	0	25.6	142.4	0	26.1	145.3	0
2	32.0	171.8	14.3	29.0	155.9	13.3	30.5	163.9	16.9
3	32.3	170.8	15.4	29.3	155.2	14.5	30.8	163.0	18.0
5	33.1	169.1	18.2	30.0	153.2	17.2	31.5	161.5	20.7
10	37.1	170.8	32.5	33.4	153.5	30.5	35.3	162.1	35.3
15	37.7	166.9	34.6	34.4	152.0	34.4	35.8	158.3	37.2

Table XXXIII. Impact of wireless losses in data-set16 on the performance of Predictor7, Predictor8 and Predictor9

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	6.6	146.5	0	6.0	134.0	0	5.8	127.3	0
2	8.9	186.7	34.9	8.8	184.1	46.7	7.9	165.6	38.9
3	9.3	182.2	40.9	9.1	181.2	51.7	8.2	163.4	43.1
5	9.6	182.0	45.5	9.4	178.9	56.7	8.5	161.2	54.2
10	11.0	168.7	66.7	10.7	163.7	78.3	9.9	151.3	58.3
15	11.4	160.9	72.7	11.2	157.7	86.7	10.2	143.3	79.2

Table XXXIV. Impact of wireless losses in data-set17 on the performance of Predictor7, Predictor8 and Predictor9

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	8.4	145.5	0	8.0	133.8	0	7.2	120.2	0
2	11.2	176.3	33.3	11.2	176.3	40.0	10.0	157.4	38.9
3	11.5	175.2	36.9	11.5	177.8	43.8	10.3	157.4	43.1
5	12.3	176.5	46.4	12.2	175.1	52.5	11.1	159.0	54.2
10	12.6	166.2	50.0	12.7	166.2	58.8	11.4	150.2	58.3
15	14.2	167.1	69.1	14.0	164.4	75.0	12.9	151.4	79.2

Table XXXV. Impact of wireless losses in data-set18 on the performance of Predictor7, Predictor8 and Predictor9

Wireless	Predictor1			Predictor2			Predictor3		
loss(%)	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1	MSE_1	MSE_2	ΔE_1
0	59.3	69.4	0	57.2	66.4	0	61.5	71.4	0
2	59.7	69.5	0.14	57.4	66.8	0.6	61.7	71.9	0.7
3	59.7	69.5	0.14	57.4	66.8	0.6	61.7	71.9	0.7
5	59.7	69.5	0.14	57.4	66.7	0.45	61.8	71.8	0.56
10	59.8	69.4	0.0	57.4	66.6	0.3	61.8	71.7	0.42
15	59.8	69.3	-0.14	57.4	66.6	0.3	61.8	71.7	0.42

C. Chapter Summary

This chapter gave insight about the impact of wireless losses on the performance of the predictors developed. From the results it can be concluded that the unpredictable wireless losses has major impact on the predictors performance. It is observed that in certain cases the wireless losses has increased the error more than 100%. It is also observed that if the predictor developed is poor then wireless losses has very less impact on its performance.

CHAPTER VII

SUMMARY AND CONCLUSION

A. Summary

The objective of this research work is to develop SSP and MSP for network accumulation in MIP networks. Various predictors are developed and tested in this work. Since these networks are highly dynamic in nature the traditional approaches such as queuing theory and other statistical theories are not useful. In this thesis, network accumulation is modeled using linear models like Auto-Regressive eXogenous (AR).

Chapter I gives the literature review done in this work. Most of the literature is related to the QoS related to the best-effort networks and the proposed mechanisms for providing Internet access to mobile end-users. It gives information about the research conducted in end-to-end dynamics measurement of best-effort networks, and use of system identification techniques for modeling network dynamics. It also provides the various techniques researchers have proposed for providing Internet access to the mobile end users. Network topology used in this work is also discussed. This chapter gives the methodology adopted in this work and proposed approach of this research.

Chapter II provides the various factors impacting QoS in MIP networks. It gives in detail, the various types of channel fading losses along with their causes and discusses how these losses can impact the QoS. In this chapter impact of congestion in wired network on the QoS is also discussed.

Chapter III talks about the linear system identification technique AR used in this work. It gives the AR model structure and the mathematical representation of the model. Also discussed in this chapter is the parameter estimation of the models.

Chapter IV gives the data collection techniques and the various end-to-end network measurement parameters. The data collection part is divided in three parts, first collection of the data pertaining to the losses due to channel fading and mobility, the second is the collection of real traffic data and the last part describes how MIP data is obtained by superimposing the simulated channel fading losses on the real traffic data. In this chapter, various end-to-end measurement parameters are discussed which gives the estimate of the congestion level in the network.

In Chapter V, the performance metrics considered for developing and evaluating the SSP and MSP are discussed. The performance of SSP and MSP are presented. SSP results are quite good and as the prediction horizon is increased the prediction of MSP detoriates.

A quantitative discussion on the impact of channel fading losses on the predictors developed in Chapter VI is done in chapter VII. Results shows that the channel fading has major impact on the predictor accuracy.

B. Conclusion and Recommendations

The proposed work has direct impact on the end-to-end network congestion dynamics of the MIP networks. The empirical models developed in this work can be used to develop effective bandwidth allocation and network control strategies for MIP networks. The predicted values can be used to develop controller to control the wireless mobile traffic over internet, which will increase the QoS for the mobile end user. The study of impact of the channel fading losses on the predictor accuracy is useful and it gives us the wireless channel conditions under which predictors will give good results.

Following conclusions can be drawn from this study:

1. Linear system identification techniques such as AR can be used to model MIP network flow accumulation. Real time network measurements can be used to develop models for predicting network congestion.

2. By studying the predictor performance it can be concluded that SSP are more accurate than MSP, but MSP are needed to predict the network accumulation for certain future time horizon.

3. The predictors perform satisfactorily under varying network congestion level for a given sender-receiver pair.

4. Channel fading losses have major impact on the accuracy of the predictors developed. This is because, the channel fading errors are random errors.

5. In the best predictor case 15% increase in wireless losses has increased MSE1 by 170%, while in the worst predictor case it has nearly no impact on the predictor performance.

Following are the recommendations for future work in this area:

1. Impact of Predictor performance needs to be studied as the mobile node moves from one base-station to the other.

2. More accurate fading models representing the actual wireless channel need to be developed and used.

3. Experimental setup with real MIP networks needs to be developed so that real MIP data can be obtained.

REFERENCES

[1] D. Srikar, "Empirical Modeling of End-to-end Delay Dynamics in Best- Effort Networks," M.S. thesis, Texas A&M University, College Station, Texas, 2004.

[2] S. J. Kerry, "The Working Group for Wireless LANs," [Online]. Available: http://grouper.ieee.org/groups/802/11/index.html, Accessed on June 2004.

[3] "Official Bluetooth Info Site," [Online]. Available: http://www.bluetooth.com, Accessed on March 2004.

[4] "HomeRF Resources," [Online]. Available: www.palowireless.com, Accessed on August 2004.

[5] "Wireless IP Services," [Online]. Available: http://www.metricom.com, Accessed on June 2004.

[6] "IETF Home Page," [Online]. Available: http://www.ietf.org, Accessed on July 2004.

[7] S. Floyd and K. Fall, "Promoting the Use of End-to-End Congestion Control in the Internet," *IEEE/ACM Transactions on Networking*, vol. 7, no. 4, pp. 458-472, August 1999.

[8] V. Khariwal, "Adaptive Control of Real Time Media Applications in Best Effort Networks," M. S. Thesis, Texas A&M University, College Station, Texas 2004.

[9] S. Floyd, K. Fall, J. Padhye, and J. Widmer, "Equation-Based Congestion Control for Unicast Applications," *in ACM SIGCOMM '00*, Stockholm, Sweden, August 2000, pp. 43-56.

[10] D. Bansal and H. Balakrishnan, "TCP-Friendly Congestion Control for Real-Time Streaming Applications," *MIT Tech. Report*, MIT-LCS-TR-806, May 2000.

[11] D. Dwyer, S. Ha, J-R. Li, and V. Bhargavan, "An Adaptive Transport Protocol for Multimedia Communication," in *Proceedings of IEEE Multimedia Computing and Systems*, June 1998, Austin, Texas, pp. 23-33.

[12] T. Turletti. and C. Huitema, "Videoconferencing on the Internet," *IEEE Transactions on Networking*, vol. 4, no. 3, pp. 340-350, June 1996.

[13] J. Bolot and T. Turletti, "A Rate Control Mechanism for Packet Video in the Internet," in *Proceedings of the Fourteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, Boston, Massachusetts, April 1994, vol. 3, pp. 1216-1223.

[14] B. Smith, "Cyclic-UDP: A Priority-Driven Best-Effort Protocol," *Tech. Report*, Cornell University, 1995.

[15] S. C. Liew and D. C. Tse, "A Control-Theoretic Approach to Adapting VBR Compressed Video for Transport over a CBR Communications Channel," *IEEE Transactions on Networking*, vol. 6, no. 1, pp. 42-55, February 1998.

[16] A. Kolarov and G. Ramamurthy, "A Control-Theoretic Approach to the Design of an Explicit Rate Controller for ABR Service," *IEEE/ACM Transactions on Networking*, vol. 7, no. 5, pp. 741-753, October 1999.

[17] S. Mascolo, D. Cavendish, and M. Garla, "ATM Rate Based Congestion Control using a Smith Predictor: An EPRCA implementation," in *Proceedings of the Fifteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, San Francisco, California, 1996, vol. 2, pp. 569-576.

[18] L. BenMohamed and Y.T. Wang, "A Control-Theoretic ABR Explicit Rate Algorithm for ATM Switches with Per-VC Queuing," in *Proceedings of the Seventeenth Annual Joint Conference of the IEEE Computer and Communications Societies*, San Francisco, California, 1998, vol. 1, pp. 183-191.

[19] Y.S. Sun, F. Tsou, M. Chang Chen, and Z. Tsai, "A TCP-Friendly Congestion Control Scheme for Real-Time Packet Video using Prediction," in *Proceedings of the 1999 Global Telecommunications Conference*, Rio de Janeireo, Brazil, December 1999, vol. 3, pp. 1818-1822.

[20] D. Wu, Y. T. Hou, W. Zhu, Y. Zhang, and J. M. Peha, "Streaming Video over the Internet: Approaches and Directions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 282-300, March 2001.

[21] D. Wu, Y. T. Hou, W. Zhu, H. Lee, T. Chiang, Y. Zhang, and H. J. Chao, "On End-to-End Architecture for Transporting MPEG-4 Video over the Internet," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 6, pp. 923-941, September 2000.

[22] D. Wu, Y. T. Hou, and Y. Zhang, "Transporting Real-Time Video over the Internet: Challenges and Approaches," *Proceedings of the IEEE*, vol. 88, no. 12, pp. 1855-1875, December 2000.

[23] M. Yang, X. R. Li, H. Chen, and N. S. V. Rao. "Predicting Internet End-to-End Delay: An Overview," *Proc. of 36th IEEE Southeastern Symposium on Systems Theory*, pp. 210-214, Atlanta, Mar. 2004.

[24] H. Ohsaki, M. Morita, and M. Murata, "On Modeling Round-Trip Dynamics of the Internet using System Identification," *IEICE Transactions on Communication*, vol. E85-B, no. 1, pp. 1-9, January 2002.

[25] S. Parekh, N. Gandhi, J. Hellerstein, D. Tilbury, T. Jayram, and J. Bigus, "Using Control Theory to Achieve Service Level Objectives in Performance Management," in *Proceedings of the 2001 IEEE/IFIP International Symposium on Integrated Network Management Proceedings*, May 2001, pp. 841-854.

[26] B. Badrinath, and A. Talukdar "IPv6 + Mobile-IP + MSRVP = Internet Cellular Phones?," [Online] Available: http://www.cs.tcd.ie/Hitesh.Tewari, Accessed on December 2003.

[27] K. Langendoen "QoS Adaptive Services in Best-Effort Mobile Networks: State of the Art and Analysis," Enschede: Telematica Institute, Netherland, 2000.

[28] V. Gambiroza, B. Sadeghi and E. W. Kightly "End-to-end Performance and Fairness in Multihop Wireless Backhaul Networks," in *Proceedings of ACM MobiCom 2004*, Philadelphia, September 2004, pp. 287-301.

[29] X. Wu, S. H. G. Chan and B. Mukharjee "MADF: A Novel Approach to Add an Ad-Hoc Overlay on a Fixed Cellular Infrastructure," in *Proceedings of the IEEE WCNC*, Chicago, Sept. 2000.

[30] B. Bhargava, X. Wu, Y. Lu, and W. Wang "Integrating Heterogeneous Wireless Technologies: A Cellular Aided Mobile Ad hoc Network (CAMA)," *ACM Mobile Network and Applications*, vol. 9, pp. 393-408, 2003.

[31] J. F. Kurose and K. W. Ross, "Computer Networking: a Top Down Approach Featuring the Internet," 1st ed., Addison Wesley, Boston, MA, 2001.

[32] V. Paxson, Measurements and Analysis of End-to-End Internet Dynamics, Ph.D. dissertation, University of California, Berkeley, April 1997.

[33] U. Schwantag, An analysis of the applicability of RSVP, Diploma Thesis, University of Oregon, Oregon and Institute of Telematics, University of Karlsruhe, July 1997.

[34] A. G. Parlos, Neural Networks for Intelligent estimation and control, Class notes for MEEN 689 December 2002.

[35] I. Yeom, ENDE: An End-to-End Network Delay Emulator , M.S. Thesis, Texas A&M University, College Station, Texas, August 1998.

[36] K. Fall, K. Vardhan, "NS Manual," [Online]. Available: http://www.isi.com-nsnam-ns-doc-ns_doc.pdf, Accessed on August 2004.

VITA

Sameer Bhoite was born in Nagpur, Maharashtra, India on the 4th of September, 1979. He received his B.E. degree in mechanical engineering from Visvesvaraya Regional College of Engineering, Nagpur, India, in 2001. In August 2002 he joined the master's degree program in mechanical engineering at Texas A&M University.

Permanent Address:

Sameer Bhoite
"Snehprabha", 16 Balaji Nagar (West),
Nagpur, India-440027
Phone:(091) 712-274-9682
email: sameerb@tamu.edu

The typist for this thesis was Sameer Bhoite.