National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services Branch
Direction des acquisitions et des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Throughput and Delay Performance of Transport User in Congestion Controlled Hybrid ATM/TDMA Networks

Malleswara Rao Talla

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University
Montreal, Quebec, Canada

October, 1995

© M. Talla, 1995

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-10904-6

Canada

ABSTRACT

Throughput and Delay Performance of Transport User in Congestion Controlled Hybrid ATM/TDMA Networks

Malleswara Rao Talla, Ph. D.

Concordia University, 1995

The end-to-end throughput and delay characteristics are analyzed for a virtual circuit (VC) transport user with an underlying ATM global congestion control scheme in a hybrid ATM/TDMA network. First of all, the ARQ transport user is assumed over *M*-node VC to analyze throughput and delay using Norton equivalent queueing model. The service characteristic of the model is obtained from the end-to-end protocol efficiency formulas of go-back-N (GBN) and selective repeat (SR) ARQ schemes. The analysis is applied to the geo-stationary earth orbit (GEO) and the low-earth orbit (LEO) satellite links. The optimum end-to-end ARQ flow control parameters, i.e., packet size, window size, and bandwidth, are also identified for better throughput and low end-to-end delay. Later, a new local congestion control scheme is analyzed by taking into account only the ATM traffic from the local transport users of the ATM multiplexer node. This scheme is based on the leaky bucket and virtual leaky bucket cell discard schemes, and utilizes the interaction between the ATM and higher layers in a hybrid ATM/TDMA network. The transport users are assumed to be generic ATM sources, who modulate their end-to-end flow control parameters, i.e., protocol data unit size in case of video and voice traffic, and window size in case of data traffic, based on the congestion status. Simple analytical

formulas are derived for congestion criteria to represent the required bandwidth to support various classes of service (i.e., video, voice, data, etc.). The ATM multiplexer node buffer is analyzed using a Markov modulated Poisson process queuing model with bulk arrival and bulk service of cells. A new global congestion control algorithm is also designed for the ATM multiplexer node by utilizing *early warning cells*. The global congestion control algorithm prioritizes the transit traffic over local traffic based on the instantaneous global congestion level. It also guarantees the quality of service to all classes of service. The ATM multiplexer node congestion performance criteria, i.e., the mean probabilities of ATM multiplexer node congestion, generating a cell, discarding a cell, buffer content, buffer overflow, and congestion duration, are evaluated with and without the congestion control schemes. Finally, the throughput and delay characteristics are analyzed for the transport user with and without the underlying ATM *global* congestion control.

ACKNOWLEDGMENTS

First, I express my deepest respect and gratitude to my teacher and supervisor Professor A. K. Elhakeem, without whose expert guidance and encouragement the quality and the extent of this research would not have been possible. I would like to express my deepest respect and gratitude to Professor T. Radhakrishnan, without whose encouragement, this research would not have been initiated. I would also like to express my deepest gratitude to Professor M. Kadoch for valuable suggestions during this research work. I would also like to extend my deepest gratitude to the anonymous supervisory committee for the constructive criticism to lead this research to a success.

I would like to express my deepest respect and gratitude to Mr. Dez Andorka, Vice President & General Manager, SITA North America and Caribbean region, Mr. Jean Rech, Director, Network projects, SITA, and Mr. Zafar Khan, Manager, YUL Network projects, SITA, for providing me an opportunity to study for a Ph.D. degree. I would also like to express my deepest respect and gratitude to Mr. Fred McGregor, Manager, Personnel development and Management training, SITA, Mr. Francois Aubin, Manager, Human Resources & Government relations, SITA, Mrs. Patricia Kravec, Human Resources Administrator, SITA, for the constant encouragement during my study. I would also like to thank all my friends for the encouragement during my study.

Finally, I would like to express special thanks to my wife Malleswari, without whose sacrifice of time, this research work would not have been possible.

dedicated to my parents

TABLE OF CONTENTS

List of figures ix

List of symbols xii

List of acronyms and abbreviations xiv

1. INTRODUCTION 1

1.1 Introduction 1

1.2 Research contributions 5

1.3 Scope of the thesis and objectives 6

2. TRANSPORT USER THROUGHPUT AND DELAY 11

2.1 Introduction 11

2.2 Description of the model 12

2.3 Analysis 16

2.4 Computation procedure and results 22

2.5 Conclusions 58

3. ATM NETWORKS 61

3.1 Introduction 61

3.2 Asynchronous Transfer Mode 61

3.3 ATM protocol reference model 64

3.4 ATM traffic models 65

3.5 Congestion in ATM networks 78

4. LOCAL CONGESTION CONTROL 84

4.1 Introduction 84

4.2 Description of the model 87

4.3 Queueing analysis 97

4.4 Congestion performance analysis 103
4.5 Computation procedure and results..... 107
4.6 Conclusions 118

5. GLOBAL CONGESTION CONTROL **120**
5.1 Introduction 120
5.2 Description of the model 123
5.3 Analysis of global congestion control 126
5.4 Global congestion performance results 134
5.5 Transport user performance with underlying ATM congestion control 156
5.6 Conclusions 182

6. CONCLUSIONS **185**
6.1 Summary 185
6.2 Conclusions 185
6.3 Further research work 188

REFERENCES **192**

List of figures

Fig. 2.1 ATM Network

Fig. 2.2 Use of Norton's theorem in queueing networks

Fig. 2.2 (a) Original network

Fig. 2.2 (b) Norton's theorem applied

Fig. 2.2 (c) Equivalent model

Fig. 2.3 Simplified model of sliding window control protocol

Fig. 2.3 (a) Sliding window control, symmetric virtual circuit

Fig. 2.3 (b) Norton equivalent, cyclic queue network

Fig. 2.4 Norton equivalent network

Fig. 2.5 Throughput for Go-Back-N scheme with k=568 bits,n_h=56 bits,R=9600bps

Fig. 2.6 Delay for Go-Back-N scheme with k=568 bits, n_h=56 bits, R=9600 bps

Fig. 2.7 Throughput for SR scheme with k=568 bits, n_h=56 bits, R=9600 bps

Fig. 2.8 Delay for SR scheme with k=568 bits, n_h=56 bits, R=9600 bps

Fig. 2.9 Throughput for Go-Back-N scheme with N=4, n_h=56 bits, R=9600 bps

Fig. 2.10 Delay for Go-Back-N scheme with N=4, n_h=56 bits, R=9600 bps

Fig. 2.11 Throughput for SR scheme with N=4, n_h=56 bits, R=9600 bps

Fig. 2.12 Delay for SR scheme with N=4, n_h=56 bits, R=9600 bps

Fig. 2.13 GBN Throughput vs Transmission speed with N=4, k=568, n_h=56 bits

Fig. 2.14 GBN Delay vs Transmission speed with N=4, k=568, n_h=56 bits

Fig. 2.15 SR Throughput vs Transmission speed with N=4, k=568, n_h=56 bits

Fig. 2.16 SR Delay vs Transmission speed with N=4, k=568, n_h=56 bits

Fig. 3.1 Traffic characteristics of various applications

Fig. 3.2 ATM protocol reference model

Fig. 3.3 Voice source behavior

Fig. 3.4 A statistically packetized voice multiplexer

Fig. 3.5 A three state MMPP

Fig. 3.6 ARMA-ZMNL model

Fig. 3.7 Video cell arrival processes

Fig. 3.8 Birth-death model for *N* sources

Fig. 3.9 Classification of ATM congestion control schemes
Fig. 4.1 Three levels of ATM traffic
Fig. 4.2 Two-phase burst/silence model for ATM call
Fig. 4.3 Cell generation of a bursty call
Fig. 4.4 Queueing model for ATM/TDMA
Fig. 4.5 ATM/TDMA composite frame structure
Fig. 4.6 State transition diagram for the queue length
Fig. 4.7 Congestion control schemes in ATM multiplexer
Fig. 4.8 Computation of performance criteria
Fig. 4.9 Mean probability of congestion
Fig. 4.10 Mean probability of generating a cell
Fig. 4.11 Mean probability of discarding a cell
Fig. 4.12 Mean buffer content
Fig. 4.13 Mean buffer overflow
Fig. 5.1 A cluster of c (c=5) neighboring nodes in ATM Network
Fig. 5.2 Global congestion control state diagram
Fig. 5.3 Mean probability of global congestion
Fig. 5.4 Mean probability of generating a cell
Fig. 5.5 Mean probability of discarding a cell
Fig. 5.6 Mean buffer content
Fig. 5.7 Mean buffer overflow
Fig. 5.8 Mean congestion duration
Fig. 5.9 Transport users over ATM/TDMA network
Fig. 5.10 Go-Back-N end-to-end throughput, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps
Fig. 5.11 Go-Back-N end-to-end delay, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps
Fig. 5.12 SR end-to-end throughput, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps
Fig. 5.13 SR end-to-end delay, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps
Fig. 5.14 Go-Back-N end-to-end throughput, packet=4096 bytes, n_h=7 bytes, R=64 Mbps

Fig. 5.15 Go-Back-N end-to-end delay , packet=4096 bytes, n_h=7 bytes, R=64 Mbps

Fig. 5.16 SR end-to-end throughput , packet=4096 bytes, n_h=7 bytes, R=64 Mbps

Fig. 5.17 SR end-to-end delay , packet=4096 bytes, n_h=7 bytes, R=64 Mbps

List of symbols

a	Propagation constant
τ	Node-to-node propagation delay
λ	Arrival rate
μ	Service rate
N_f	Transmitted packets for one successful packet transmission
n_h	Packet header size in bits
k	Packet size including header
P_b	Probability of bit error
P	Probability of packet error
η	Protocol efficiency
N	Window size
ρ	Normalized applied load, or Traffic intensity
γ	Throughput
K_i	Total number of active transport users of service-class 'i'
K_{av}	Average number of transport users ($K_{av} \geq K_i$)
k_i	Total number of active bursts of service-class 'i'
q_i	Quality of service for service-class 'i'
L_iB_i	Total buffer space of ATM multiplexer node
ll	Number of composite TDMA frames which can be queued
N_{fmax}	Measurement window
n_{ui}	Maximum number of cells, allowed from a transport user of service class 'i'
nn	instantaneous maximum number of cells from 'k_i' active bursts
N_{si}	Number of time slots in a TDMA subframe for a service class 'i'
N_s	Total number of time-slots for all classes of traffic in a TDMA frame
R_{pi}	Peak service-rate for a service class 'i'
S	Total number of service classes
BW	Total bandwidth of ATM multiplexer node
P_a	Probability that an active call generates an active burst

α	Probability of ATM source transition from active to idle state
β	Probability of ATM source transition from idle to active state
θ_l	Probability that 'k_i' active bursts generate 'l' number of cells
$P(k_i)$	Probability that 'K_i' active calls generating 'k_i' active bursts
$\psi(i)$	Probability of ATM multiplexer node buffer state 'i'
p	Instantaneous TPDU size
p_{max}	TPDU size
w	Instantaneous number of TPDUs in the window
w_{max}	Transport user window size, i.e. N
P_a'	Probability of generating active burst based on congestion status
P_c	Probability of congestion
P_{vi}	Probability of violation
ξ	Probability of generating a cell
Δ	Probability of discarding a cell
$P_{overflow}$	Probability of buffer overflow
CD	Congestion duration

List of acronyms and abbreviations

AAL	ATM adaptation layer
ABR	Available bit rate
ADPCM	Adaptive differential pulse code modulation
ARMA	Auto-regressive moving average
ARQ	Automatic repeat request
ATM	Asynchronous transfer mode
BISDN	Broadband integrated services digital network
BPS	Bits per second
BW	Bandwidth
CAC	Connection admission control
CBR	Constant bit rate
CCITT	Consultative committee for international telegraphy and telephony
CD	Congestion duration
ECN	Explicit congestion notification
EWMA	Exponentially weighted moving average
GBN	Go-back-N
GEO	Geo-stationary earth orbit
GFC	Generic flow control
GMDP	Geometrically modulated deterministic process
IP	Internet protocol
IPP	Interrupted Poisson process
ISDN	Integrated services digital network (narrowband)
JW	Jumping window
KBPS	Kilo bits per second (1 Kbps=10^3 bps)
LAN	Local area network
LB	Leaky bucket
LEO	Low earth orbit
MBPS	Mega bits per second (1 Mbps=10^6 bps)
MMPP	Markov modulated Poisson process

mLB	modified leaky bucket
mVLB	modified virtual leaky bucket
MW	Moving window
OAM	Operation and maintenance
OSI	Open systems interconnection
PCM	Pulse code modulation
PCS	Personal communication services
PDU	Protocol data unit
QOS	Quality of service
RM	Resource management
SAD	Speech activity detection
SDU	Service data unit
SPP	Switched Poisson process
SR	Selective repeat
STM	Synchronous transfer mode
TCP	Transmission control protocol
TDMA	Time division multiple access
TJW	Triggered jumping window
TPDU	Transport protocol data unit
UBR	Unspecified bit rate
UDP	User datagram protocol
UPC	Usage parameter control
VBR	Variable bit rate
VC	Virtual circuit
VCI	Virtual circuit identifier
VLB	Virtual leaky bucket
VMTP	Versatile message transaction protocol
VP	Virtual path
VPI	Virtual path identifier
ZMNL	Zero memory non-linearity

CHAPTER 1

INTRODUCTION

1.1 Introduction

Perhaps the most important development in telecommunication networking over the recent years has been the introduction of *frame relay* and *cell relay* capabilities into various telecommunication products. The *frame relay* and *cell relay* technologies take advantages of the high data rates and low error rates of contemporary networking facilities. They are outgrowths of the work on integrated services digital network (ISDN) and broadband ISDN (BISDN) [STA-94]. The BISDN supports multimedia traffic using the high bandwidths supported by new physical media such as optical fibers. The emerging optical fiber technology is still very expensive for transoceanic optical fiber cabling and maintenance, though it provides high bandwidths. So, the communications satellites will greatly be in-use for both regional and international applications [MAH-87]. The low earth orbit (LEO) satellite systems will soon begin offering low-cost worldwide multimedia services [ILA-95]. To take advantages of these new technologies, ATM user oriented LAN-to-LAN interconnected services will probably represent the main service to be provided by broadband networks using satellite links [PRI-89], [INU-92], [CIN-94], [KWA-92], [DEL-95].

Satellite communications are usually realized by geo-stationary earth orbit (GEO) satellites, because the satellite can be treated as a fixed relay station. The GEO is only one orbit, and its capacity (i.e., number of satellites) is obviously limited. Moreover, the

altitude of GEO satellite is so high that the propagation delay for the two-way communication is too large. Big antennas and high power amplifiers are also needed to compensate the serious propagation loss of GEO satellite communication. As a way to prevent difficulties of GEO satellites, LEO satellites smaller than GEO satellites in size are widely becoming popular. The LEO satellites can avoid Van Allen radiation [LEI-92], [DEL-95]. The LEO satellites allow small size antennas and low power transponders due to the low propagation loss of the satellite link. In contrast to GEO satellites, the visible-time of LEO satellites is also limited [HAS-92]. The commercial uses of LEO have been limited mostly to Earth Resources missions. However, other missions notably personal mobile communications (PCS) have recently created a need for low weight, low cost, and high performance LEO satellite systems. It is also expected that LEO-SAT communication systems will expand the possibility of realizing global mobile satellite communications. Moreover, LEO is also being studied for global messaging networks [NOH-93], [KAN-92].

Cell relay is realized in asynchronous transfer mode (ATM). ATM has been adopted by CCITT as the transport and switching technique for BISDN, due to the potential bandwidth efficiency offered by the statistical multiplexing of bursty traffic, and its ability to support bandwidth on-demand. ATM networks should provide guaranteed quality of service (QOS) for a variety of applications ranging from low-speed data to high-speed computer-LAN interconnections, or video image transfer. In such networks, congestion develops rapidly, not only due to the highly bursty traffic generated by its own

traffic sources, but also due to the bursty transit traffic from other nodes. So, it is important to communicate the congestion criteria of an ATM multiplexer node to all neighboring ATM multiplexer nodes periodically. Depending on the congestion status, the ATM multiplexer node may apply reactive and preventive congestion controls to the internal ATM traffic sources. At high transmission speeds, the propagation delay becomes a significant bottleneck even for congestion communication, because the network can be flooded before congestion status is propagated. Two basic traffic control functions, i.e., connection admission control (CAC), and usage parameter control (UPC), are commonly used in ATM networks. The CAC method admits any new connection only when it does not create congestion. Even though some connections exceed their specified traffic rates either maliciously or unintentionally, the UPC method does not deteriorate the QOS of other connections during the information transfer phase [YAM-95], [PAR-95]. Besides the traffic control functions, several new methods were investigated to address the propagation delay problem. These new methods are implemented by means of congestion notification messages, e.g. fast resource management based congestion control using pilot ATM cells for indicating the payload [DA2-93], [DA3-93], fast reservation protocols and explicit congestion notification [WER-92], burst-level feedback control using generic flow control (GFC) field of ATM cell header [RUB-94], backward congestion notification using ATM resource management (RM) cells [NEW-94], reactive congestion control to adapt to traffic load [KAW-95] etc.. Several other methods were also investigated to control active bursts to achieve congestion control in ATM networks [LIU-95], [WID-95], [ABB-94].

Recently lot of research has been conducted to explore new schemes for *call admission control* [SHI-95], [OGI-95], [MUR-91], [RAS-91], *burst admission control* [ST2-94], *flow control schemes (window, credit, and rate based)* [ROB-95], [KUN-95], [RAM-95], [JA2-86], [OHN-88], [DIT-91], *congestion avoidance by adaptive routing* [BAH-94], *congestion notification* [ROU-94], *priority queueing* [HUA-94], [LIN-91] etc., in broadband ATM networks. Hence the *motivation* for this research work is the current industrial need to explore new ATM congestion control schemes for improved end-to-end throughput and delay performance. These congestion control schemes use the existing transport protocol mechanisms and the new ATM technology. The end-to-end throughput and delay characteristics are analyzed for go-back-N (GBN) and selective repeat (SR) automatic repeat request (ARQ) transport users during the first phase of our research. The ARQ transport users are assumed over M-node virtual circuit (VC) with LEO and GEO satellite links in frame relay mode. The optimum protocol parameters of GBN and SR ARQ schemes are also identified. Later, new ATM congestion control schemes are analyzed to ensure quality of service for several transport users in interconnected broadband ATM networks. The ATM multiplexer node buffer is analyzed using Markov modulated Poisson process (MMPP) queueing model with bulk arrival and bulk service of cells. A local congestion control scheme is analyzed with *modified* versions of leaky bucket (LB) and virtual leaky bucket (VLB) schemes. The local congestion control scheme analyzes the congestion created by the traffic generated by the local transport users. The congestion control scheme would not be complete, without the transit

traffic from other ATM multiplexer nodes. So, *early warning cells* are used to communicate burstiness and congestion criteria of an ATM multiplexer node to all adjacent nodes. A global congestion control algorithm is also designed by utilizing *early warning cells* to prioritize transit traffic over local traffic based on the instantaneous global congestion level. Finally, the end-to-end throughput and delay were analyzed with the underlying ATM global congestion control over hybrid asynchronous transfer mode/time division multiple access (ATM/TDMA) network. A hybrid ATM/TDMA network allows its transport users to share the available bandwidth asynchronouly, and services the bandwidth by means of composite TDMA frames.

1.2 Research contributions

* Analysis of end-to-end throughput and delay characteristics of GBN and SR ARQ transport protocols over LEO and GEO satellite links in frame relay mode [TAL-94].

* Optimum end-to-end flow-control parameters, i.e., packet size, window size, and bandwidth in bps for ARQ transport users over frame relay GEO and LEO satellite networks [TA6-95].

* Analysis of a local congestion control scheme for an ATM multiplexer node in hybrid ATM/TDMA networks for congestion criteria, i.e., probabilities of

congestion, generating a cell, discarding a cell, buffer content and buffer overflow [TA2-95], [TA4-95].

* Design of a global congestion control algorithm for ATM multiplexer node for prioritizing transit traffic over local traffic [TA3-95], [TA5-95].

* Quality of service based global congestion control scheme in interconnected ATM multimedia networks [TA5-95].

* A complete analysis for the transport user throughput and delay performance with an underlying ATM global congestion control scheme [TA8-95].

1.3 Scope of the thesis and objectives

The end-to-end throughput and delay characteristics are analyzed for GBN and SR ARQ transport users over LEO/GEO satellite-links in frame-relay mode, during the first phase of research. The objective of this research is to explore new approaches in analyzing the transport user throughput and delay performance. The ARQ protocol efficiency formulas of [SAA-94] are extended to the M-node VC in frame relay environment without any intermediate link level error recovery procedures. As the propagation delay (τ) is generic, the effective end-to-end ARQ protocol efficiency is used to obtain the service characteristic of the resulting M-node queueing system. The transport users are assumed to use GBN and SR ARQ schemes. The optimum flow control parameters are also

identified for these end-to-end ARQ schemes. The effect of VC bandwidth (in bps) over transport user performance is also analyzed.

Later, new ATM congestion control schemes are analyzed for interconnected broadband ATM/TDMA networks. The objective of this research work is to analyze the congestion performance to ensure necessary bandwidth for supporting various classes of service, i.e., video, voice, data, etc., in interconnected broadband hybrid ATM/TDMA networks. In this thesis, two new congestion control schemes are analyzed. They provide local and global congestion controls. The local congestion control scheme regulates the traffic from local transport users, and analyzes the ATM multiplexer congestion based on the local traffic. The global congestion control scheme analyzes both local and transit traffic (by means of *early warning cells* from other ATM multiplexer nodes), to achieve global control by regulating the traffic from the local transport users. The transport users are assumed to be generic ATM sources, who generate active bursts of traffic (bulk arrival of cells) during the connection. The ATM multiplexer node is assumed to provide TDMA service (bulk service of cells) to the generic ATM sources, by means of composite TDMA frames. The ATM layer congestion status is communicated to the transport users by means of *service primitives*. The transport layer user is designed to modulate the active bursts of traffic based on the instantaneous congestion criteria, which is referred as *active burst control* in this thesis. The length of an active burst can be controlled by modulating the transport user protocol data unit (PDU) size based on the congestion criteria in case of video and voice users. As PDU size may not be controlled for data users, the number of

active bursts may be controlled by modulating the transport user window based on the congestion criteria. By controlling the active bursts judiciously at the transport layer, the amount of local traffic can be reduced during congestion periods. Moreover, the video and voice transport users can implement *quality control algorithms* to selectively drop parts of the PDU information during *severe congestion* periods, without significantly sacrificing the quality of the picture and speech information. Hence, the PDU size modulation scheme is a significant improvement over window control mechanisms of [RAT-91]. Furthermore, the transport window control reduces the possible end-to-end re-transmissions that may otherwise be prompted by cell-loss during congestion periods. Hence, the transport window control will enhance the end-to-end throughput for data traffic significantly. The LB and VLB cell discard schemes are proposed to discard the excessive traffic. The analysis of LB scheme is based on discarding all traffic that violates the *quality of service* (QOS) of a service-class. The analysis of VLB scheme is based on admitting all traffic, including that violates the QOS of the service-class, as long as the ATM multiplexer node is not congested. The VLB scheme discards only the traffic that violates the QOS of a service-class during congestion periods. The VLB scheme improves the throughput as long as the transport users do not violate their assigned bandwidths. However, we improved congestion control in the ATM multiplexer node by applying *active burst control* on the top of LB and VLB schemes, to derive *modified* LB (mLB) and *modified* VLB (mVLB), respectively. The mLB and mVLB schemes communicate instantaneous congestion criteria to the transport users by means of *service primitives*, while performing the cell discard functions of LB and VLB schemes, respectively. Hence, the mLB and

mVLB schemes shall control congestion more effectively than the respective LB and VLB schemes. Both mLB and mVLB schemes enhance the system performance by reducing the ATM multiplexer node congestion and the violation of QOS of the service-class. The admitted cells in all these four schemes, namely LB, VLB, mLB and mVLB, are added to the ATM multiplexer buffer. The ATM multiplexer node buffer is analyzed using a MMPP [AKI-94] queuing model with bulk arrival and bulk service of cells. The ATM multiplexer node is assumed to communicate the congestion criteria by transmitting *early warning cells* to all neighboring ATM multiplexer nodes. The early warning cells are used to compute global congestion criteria by means of a new global congestion algorithm. The global congestion algorithm analyzes the congestion of the ATM multiplexer by taking into account both local and transit traffic. It divides the global congestion into different *levels or states*, based on the congestion-propagation among the cluster of adjacent ATM multiplexer nodes. A simplified approach for congestion duration is also presented based on the global congestion algorithm. The global congestion criteria are communicated to the local transport users by means of *service primitives*. The local transport users modulate their active bursts of traffic based on the instantaneous global congestion criteria, to eventually achieve global congestion control. Simple analytical formulas for the ATM multiplexer node congestion criteria, i.e., the mean probabilities of congestion (P_c), generating a cell which gets admitted for transmission (ξ), discarding a cell (Δ), buffer content (i.e., the number of admitted cells) (Ψ), buffer overflow (i.e., the number of admitted cells during peak buffer occupancy, which may be lost when there is no room in the queue), and congestion duration are also derived. The ATM multiplexer-node

congestion performance of mLB and mVLB are compared with LB and VLB schemes, respectively. Finally, the transport user throughput and delay characteristics are analyzed with the underlying ATM global congestion control in a broadband hybrid ATM/TDMA network.

Chapter 2 details the queueing analysis, end-to-end throughput and delay characteristics of transport users over frame relay LEO and GEO satellite links. Chapter 3 provides a brief introduction to ATM networks, state-of-art traffic models and congestion control schemes. Chapter 4 details a local congestion control scheme for hybrid ATM/TDMA networks. Chapter 5 details global congestion control schemes for hybrid ATM/TDMA networks. It also analyzes end-to-end throughput and delay characteristics of a transport user with the underlying ATM global congestion control scheme over broadband hybrid ATM/TDMA network. Finally, chapter 6 provides a brief summary, conclusions, and a set of possible extensions of this research work.

CHAPTER 2

TRANSPORT USER THROUGHPUT AND DELAY CHARACTERISTICS

2.1 Introduction

The low-earth orbit (LEO) satellite communication reduces transmission path loss and delay when compared to geo-stationary earth orbit (GEO) satellites [SUN-94]. To improve performance of the transport protocols, simplified switching processes such as frame relay [RAH-91] may be used by moving link-level flow-control and error-recovery procedures to end systems. Unlike video and voice applications, data applications can not tolerate any loss of data. So, the traditional automatic repeat request (ARQ) schemes, i.e., go-back-N (GBN), selective-repeat (SR), etc., are still popular for data applications. However, several improved versions of ARQ schemes were investigated over the past 20 years [MIL-81], [NET-90], [FAN-90], [WAR-95]. Recent work on performance analysis of ARQ protocols includes queueing theory. In general, throughput and delay are the main performance measures [ZHO-90], [WEL-82].

In this chapter, the end-to-end transport users are assumed over a virtual circuit (VC) with M nodes, to analyze end-to-end throughput and delay, using Norton equivalent queueing model in frame relay service. The transport users are assumed to use GBN, or SR ARQ scheme. The analysis is general, and hence it is applicable to any type of transmission media. However, we used GEO and LEO satellite transmission media to identify optimum end-to-end ARQ flow control parameters such as packet and window

sizes for better throughput and low end-to-end delay. As the link transmission speed (bandwidth) can be allocated dynamically in multimedia environment, the effect of transport user bandwidth over throughput and delay is also analyzed.

2.2 Description of the model

The source and destination transport users are assumed over a VC with *M-1* satellite links, in a sliding window closed queueing model as shown in Fig. 2.1. The source and destination are linked through an added artificial queue labeled *M+1*, whose service rate is equal to λ, i.e., precisely the input rate to the *M*-node VC. A fixed number of packets (*N*), which is equal to end-to-end window of the VC, circulate through the model. All satellites over (*M-1*) links are assumed to exist at the same altitude, leading to same propagation delay (τ) over each satellite link. The transport users at the source and destination nodes are assumed to communicate over the VC in frame relay mode, i.e., without using any intermediate link-level error recovery procedures. The end-to-end GBN or SR ARQ scheme is used over the virtual circuit. According to the Norton theorem [SCH-92] for queueing network of product form (i.e., the state probability of the entire network is given by product of state probabilities in each queue), a single state dependent queue representing the network may be found as shown in Fig. 2.2. Assuming homogeneous VC with a uniform service rate of μ, the model can be further simplified to a single queue as shown in Fig. 2.3. This queue allows *n* packets between source and destination, whereas a total of *N* packets circulate through the closed network. The resultant short-circuit throughput would be *u(n)* packets/sec. Then the probabilities of

state of the lower queue are identical in both Figs 2.1 and 2.3. If the upper queue is in state n, the lower queue is in state N-n in Fig. 2.3 (b). It thus suffices to find the probabilities of state of the upper queue. Note that the equations of state for the upper queue are just those of generalized birth-death process with birth (arrival) rate of λ, and state dependent death (service) rate of $u(n)$, as shown in Fig. 2.4. All satellites over M-1 links are assumed at the same altitude, leading to same propagation delay (τ) over all links. The transport users are assumed to use GBN or SR ARQ scheme in frame relay service, i.e., without any link-level error recovery procedures. The ARQ protocol efficiency formulas of [SAA-94] are extended to the M-node VC using uniform τ. For simplicity, only τ and M are used from the LEO or GEO satellite environment at appropriate altitudes. The throughput and delay are then analyzed using the service characteristic (μ), which is obtained from the ARQ protocol.

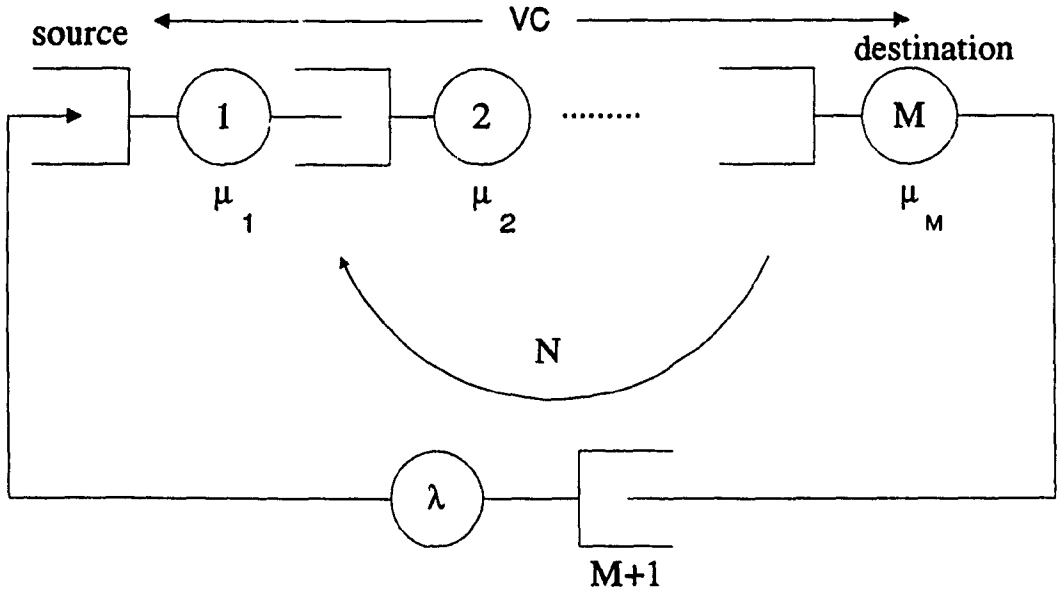

Fig. 2.1. Sliding window control model.

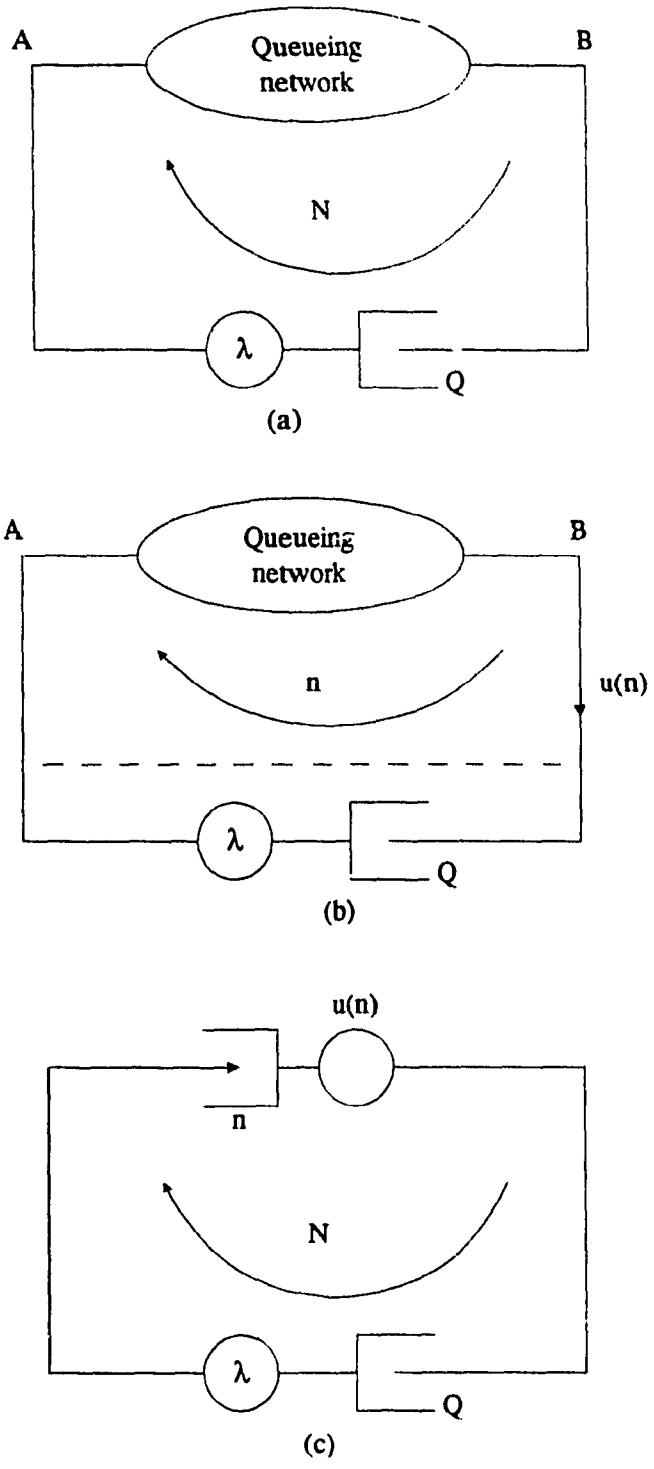

Fig. 2.2. Use of Norton's theorem in queuing networks.

(a) Original network; (b) Norton's theorem applied; (c) Equivalent model.

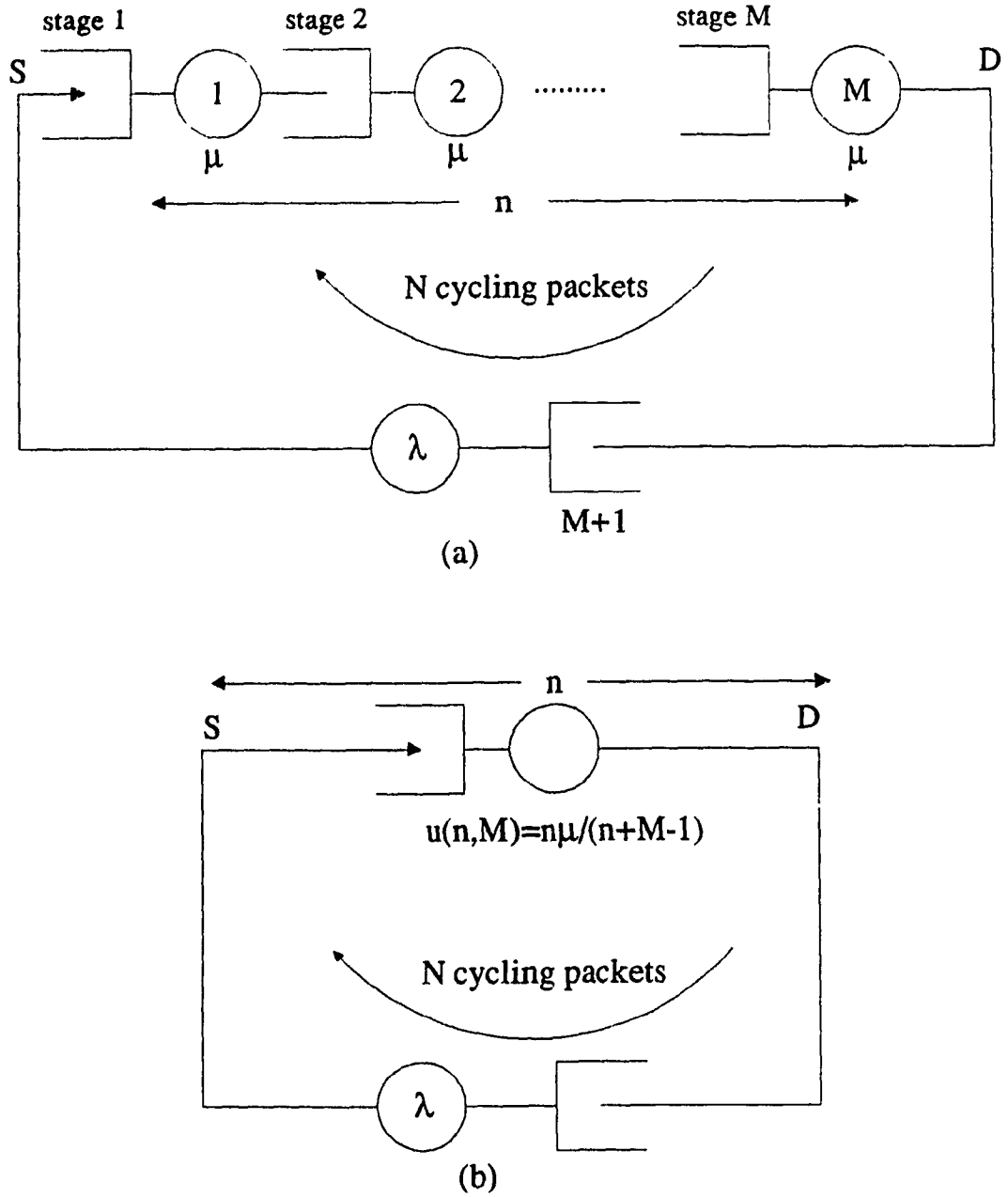

Fig. 2.3. Simplified model of sliding window control protocol.

(a) Sliding window control, symmetric virtual circuit;

(b) Norton equivalent, cyclic queue network.

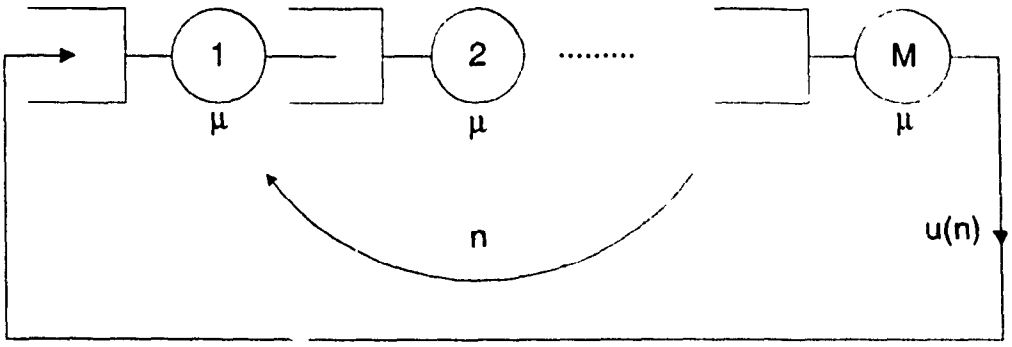

Fig. 2.4. Norton equivalent network.

2.3 Analysis

The probability that the upper queue in Fig. 2.3, is in state n, can be written as

$$p_n = \frac{p_0 \lambda^n}{\prod_{i=1}^{n} u(i)} \tag{2.1}$$

with probability normalization condition

$$\sum_{n=0}^{N} p_n = 1 \, . \tag{2.2}$$

The Norton equivalent for this symmetric network will have a state-dependent service characteristic as

$$u(n) = \frac{n\mu}{n+(M-1)} \tag{2.3}$$

where μ is homogeneous service characteristic, and n is the number of packets distributed over M nodes. Hence, the throughput $u(n)$ with n packets distributed among the M queues is always less than or equal to μ. It is given by

$$u(n) = \mu.\text{Probability}\,(queue - is - nonempty)\ . \tag{2.4}$$

Using equations (2.1)-(2.3), it can be shown that

$$\frac{p_n}{p_0} = \frac{\lambda^n}{\mu^n \prod_{i=1}^{n} \frac{i}{i+M-1}}\ . \tag{2.5}$$

The normalized applied load to the queueing system, is given by

$$\rho = \lambda / \mu \qquad . \tag{2.6}$$

Using equations (2.5) and (2.6), it can be shown that

$$\frac{p_n}{p_0} = \rho^n \begin{pmatrix} M-1+n \\ n \end{pmatrix}. \tag{2.7}$$

The throughput γ of the window controlled VC is given by averaging the service characteristic over all N possible service rates, and hence one can obtain

$$\gamma = \sum_{n=1}^{N} u(n) p_n \tag{2.8}$$

where $u(n)$, p_n are defined in equations (2.3) and (2.7), respectively. Hence, the throughput of the queueing system is

$$\gamma = \sum_{n=1}^{N} \frac{\left\{\frac{n\mu}{n+M-1}\right\} \rho^n \begin{pmatrix} n+M-1 \\ n \end{pmatrix}}{\sum_{i=0}^{N} \left\{\rho^i \begin{pmatrix} i+M-1 \\ i \end{pmatrix}\right\}}\ . \tag{2.9}$$

By Little's formula, the end-to-end delay $E(T)$ over the VC is just the ratio of average number of packets $E(n)$ along the VC to the throughput γ, and hence

$$E(T) = \frac{E(n)}{\gamma} = \frac{\sum_{n=1}^{N} np_n}{\gamma}. \tag{2.10}$$

Equations (2.9) and (2.10) provide throughput and delay characteristics of the equivalent queueing system of M node VC. However, the ARQ schemes are generally used to provide reliable data communication between the peer end-to-end users. When frame relay mode is assumed, the link-level flow-control and error-recovery procedures are moved to end users. Hence the intermediate nodes are allowed to drop erroneous packets. We assume that bandwidth (i.e., data transmission speed R for a VC) is negotiated during the VC establishment, and is ensured throughout the data transfer phase of the VC.

Applying ARQ protocol [SAA-94] environment over the resulting queueing system of Fig. 2.3, the ARQ protocol efficiency of an error-free communication over the VC, is given by

$$\eta = \begin{cases} 1 & \text{when } N \geq 1+2a \\ \dfrac{N}{1+2a} & \text{otherwise} \end{cases} \tag{2.11}$$

where N is the end-to-end window size, and

$$a = \frac{end-to-end \quad packet \quad propagation \quad delay}{packet \quad transmission \quad time}.$$

Hence, the constant 'a' for M-node VC can be defined as

$$a = \frac{\tau(M-1)R}{k} \tag{2.12}$$

where k is total transmitter bits per packet (i.e., packet size + packet header size), R is data transmission speed of the VC, τ is propagation delay per link, M is the total number of nodes, and $(M\text{-}1)$ is the total number of links over end-to-end VC.

Assuming that a packet is transmitted N_f times to ensure reliable communication between the transport users using ARQ protocol scheme, the protocol efficiency is

$$\eta = \begin{cases} \dfrac{1}{N_f} & when \quad N \geq 1+2a \\ \dfrac{N}{N_f(1+2a)} & otherwise. \end{cases} \tag{2.13}$$

When data packet size is chosen to be relatively shorter for supporting multimedia traffic, the packet header size can become a significant overhead. Hence the packet header can not be ignored for protocol efficiency computation. So, equation (2.13) becomes

$$\eta = \begin{cases} \dfrac{k-n_h}{kN_f} & when \quad N \geq 1+2a \\ \dfrac{(k-n_h)N}{kN_f(1+2a)} & otherwise \end{cases} \tag{2.14}$$

where n_h is the packet header size in bits. The value of N_f can be evaluated from the ARQ protocol scheme between the transport users, assuming no queueing delay.

Go-Back-N ARQ:

The average number of transmissions N_f, which are needed for one frame to be successful, is given by

$$N_f\big|_{GBN} = 1+\frac{NP}{1-P} \tag{2.15}$$

where P is the probability of packet error over end-to-end VC.

The acknowledgment (ACK) or NAK (negative ACK) is received only after an end-to-end round-trip propagation delay. The number of packets transmitted during round trip propagation delay, are $1+2a$, which leads to

$$N_f\big|_{GBN} = \frac{1+2aP}{1-P} \quad . \tag{2.16}$$

From equations (2.14), (2.15) and (2.16), the end-to-end GBN ARQ protocol efficiency can be obtained as

$$\eta|_{GBN} = \begin{cases} \dfrac{(k-n_h)(1-P)}{k(1+2aP)} & when \quad N \geq 1+2a \\ \dfrac{(k-n_h)N(1-P)}{k(1+2a)(1-P+NP)} & otherwise. \end{cases} \tag{2.17}$$

The resulting queueing system service rate μ for M-node VC with end-to-end GBN ARQ, can be obtained as

$$\mu|_{GBN} = \begin{cases} \dfrac{(k-n_h)(1-P_b)^{k(M-1)}}{k\left\{1+2a\left[1-(1-P_b)^{k(M-1)}\right]\right\}} & when \quad N \geq 2a+1 \\ \dfrac{(k-n_h)N(1-P_b)^{k(M-1)}}{k(1+2a)\left\{N+(1-N)(1-P_b)^{k(M-1)}\right\}} & otherwise \end{cases} \tag{2.18}$$

where P_b is probability of bit error over one link.

Selective Repeat ARQ:

The improvement of SR ARQ scheme over GBN ARQ scheme, is one packet re-transmission for each erroneous packet. Hence the value N_f for SR ARQ scheme can be obtained as

$$N_f|_{SR} = \frac{1}{1-P}. \tag{2.19}$$

Using equations (2.14) and (2.19), the SR ARQ efficiency can be written as

$$\eta|_{SR} = \begin{cases} \dfrac{(k-n_h)(1-P)}{k} & when\ N \geq 2a+1 \\ \dfrac{N(k-n_h)(1-P)}{k(1+2a)} & otherwise. \end{cases} \tag{2.20}$$

Hence, the service rate (μ) of the queueing system, which represents M-node VC with end-to-end SR ARQ, can be written as

$$\mu|_{SR} = \begin{cases} \dfrac{(k-n_h)(1-P_b)^{k(M-1)}}{k} & when\ N \geq 2a+1 \\ \dfrac{N(k-n_h)(1-P_b)^{k(M-1)}}{k(1+2a)} & otherwise. \end{cases} \tag{2.21}$$

The effective throughput and delay characteristics of the VC transport user can be obtained from equations (2.9) and (2.10), with the service characteristic of equations (2.18) and (2.21) for GBN and SR ARQ schemes, respectively.

2.4 Computation procedure and results

The end-to-end throughput and delay characteristics were obtained using the system of equations (2.1)-(2.21). Once the static system parameters such as k, n_h, R, τ, P_b, ρ, etc., are assumed, the Norton equivalent queueing system service rate μ was computed using equations (2.18) and (2.21) for GBN and SR ARQ schemes, respectively. The Norton equivalent queueing system of Fig. 2.3 uses the service rate μ for a given arrival rate λ and normalized applied load ρ as per equation (2.6). Using equations (2.9) and

(2.10), the end-to-end throughput and end-to-end delay characteristics were obtained with the service characteristic found from equations (2.18) and (2.21) for GBN and SR schemes, respectively. As smaller packet sizes are preferred for multimedia traffic, packet data size of 64 bytes (512 bits) and header size (n_h) of 7 bytes (56 bits) (i.e., k=568 bits) are initially chosen in frame relay mode. The data transmission speed over the VC is assumed to be 9600 bps for simplicity, which may be a portion of the transmitter speed. The propagation delay (τ) per link is 0.27 seconds for GEO at 36000 KM, and 0.02 seconds for LEO at 3000 KM. The normalized applied load ρ is assumed to be 0.99, 0.5, and 0.1 for depicting fully, marginally, and less loaded conditions, respectively. As satellite links generally have higher bit errors, the probability of bit error (P_b) is assumed to be 10^{-3} and 10^{-6} to reflect high and low bit error conditions, respectively. However, the same value of P_b is used for both LEO and GEO links for comparison purpose. It can be noted that due to low altitudes and low media losses, the LEO satellite links may provide lower probability of bit error compared to GEO satellite links. The results were obtained for VCs with 3 and 5 nodes over LEO satellite links. The 3 or 5 node VC uses 2 or 4 LEO satellites respectively. The results were obtained for a VC with 1 GEO satellite (M=2) to cover the same distance of 2 LEO satellite (M=3), or 4 LEO satellite (M=5) VCs.

The end-to-end throughput characteristics were obtained as a function of window size (N) for GBN ARQ scheme in Fig. 2.5. These results were obtained for 9600 bps VC, with a packet data size of 64 bytes and packet header size of 7 bytes. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99,

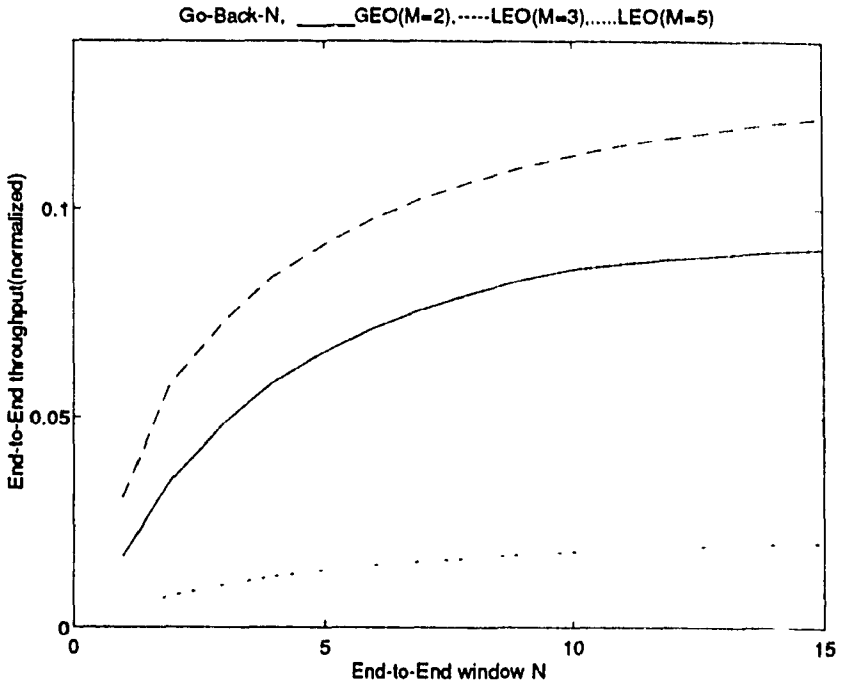

Fig. 2.5 (a) GBN scheme, P_b=10^{-3}, ρ=0.99;

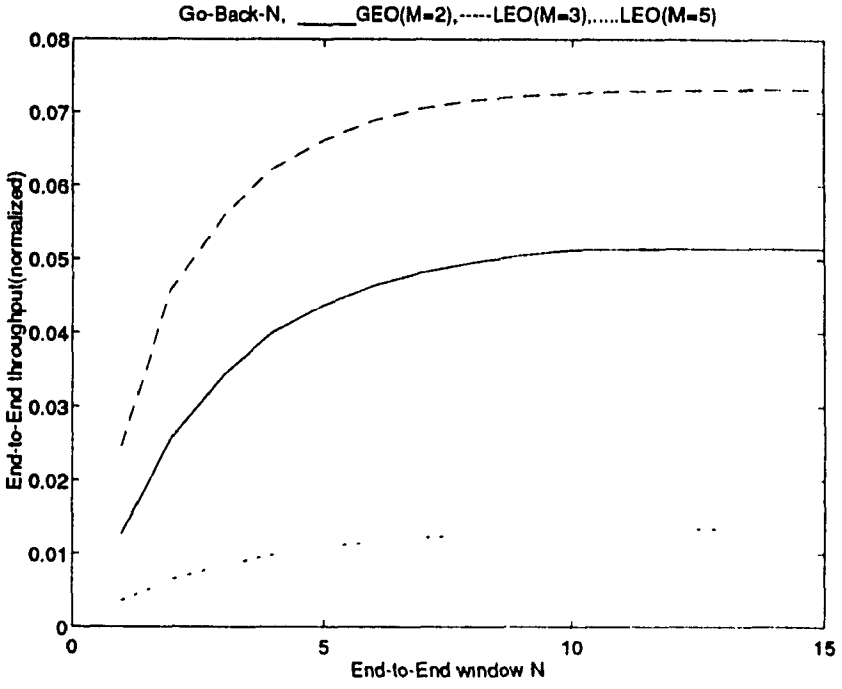

Fig. 2.5 (b) GBN scheme, P_b=10^{-3}, ρ=0.5;

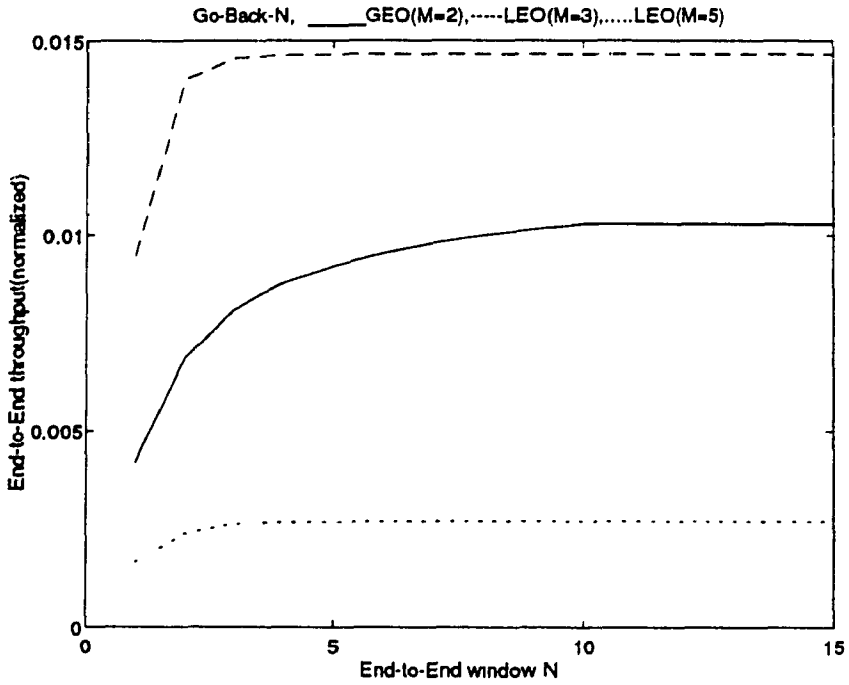

Fig. 2.5 (c) GBN scheme, P_b=10^{-3}, ρ=0.1;

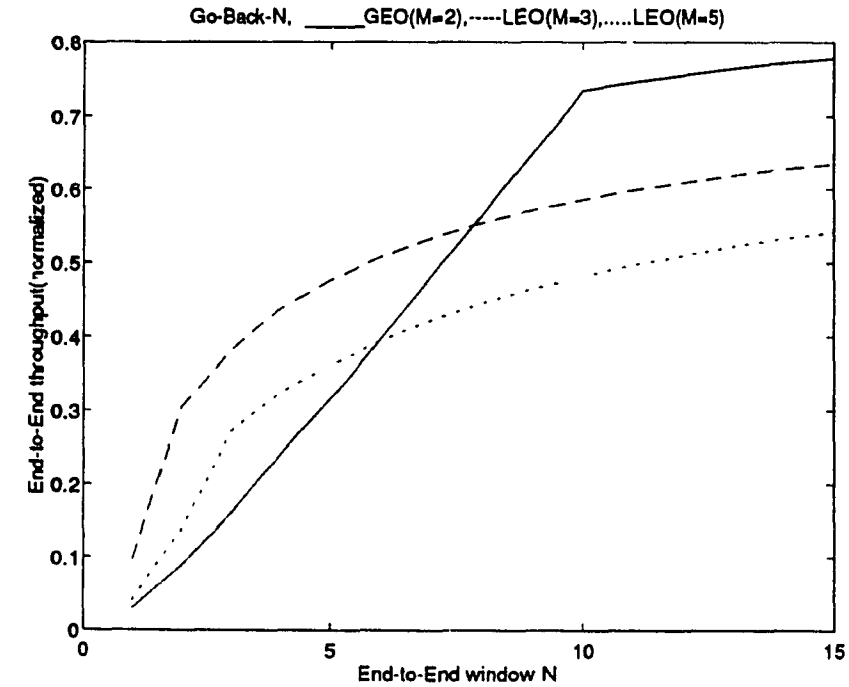

Fig. 2.5 (d) GBN scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.5. Throughput for Go-Back-N scheme with k=568 bits, n_h=56 bits, R=9600 bps.

0.5 and 0.1 in Figs 2.5 (a), (b), and (c), respectively. The results in Fig. 2.5 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. The results clearly show that the end-to-end throughput increases as the window size is increased as long as the window transmission time is within the round-trip propagation delay (i.e., $N < 2a+1$). The end-to end throughput almost saturated after a certain window size (i.e., $N \geq 2a+1$), as the acknowledgment can be received within the window transmission and propagation delays, which results maximum end-to-end throughput. When $N \geq 2a+1$, further improvement in the results can only be observed by reducing P_b and by increasing the packet size at the ARQ user. It can be observed from Figs 2.5 (a) and (d) that the end-to-end throughput significantly improved when the P_b is decreased. It can also be observed from Figs 2.5 (a), (b), and (c) that the 2 LEO satellite configuration (M=3) provides better throughput results compared to 1 GEO satellite configuration (M=2), or 4 LEO satellite configuration (M=5).

The end-to-end delay characteristics were obtained as a function of window size (N) for GBN ARQ scheme in Fig. 2.6. These results were obtained for 9600 bps VC, with a packet data size of 64 bytes and packet header size of 7 bytes. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.6 (a), (b), and (c), respectively. The results in Fig. 2.6 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for

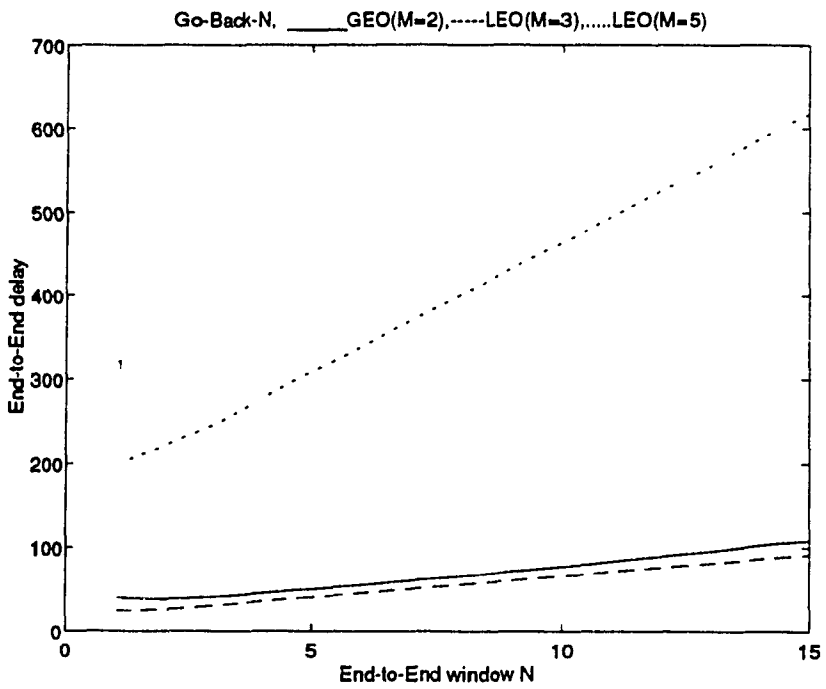

Fig. 2.6 (a) GBN scheme, P_b=10^{-3}, ρ=0.99;

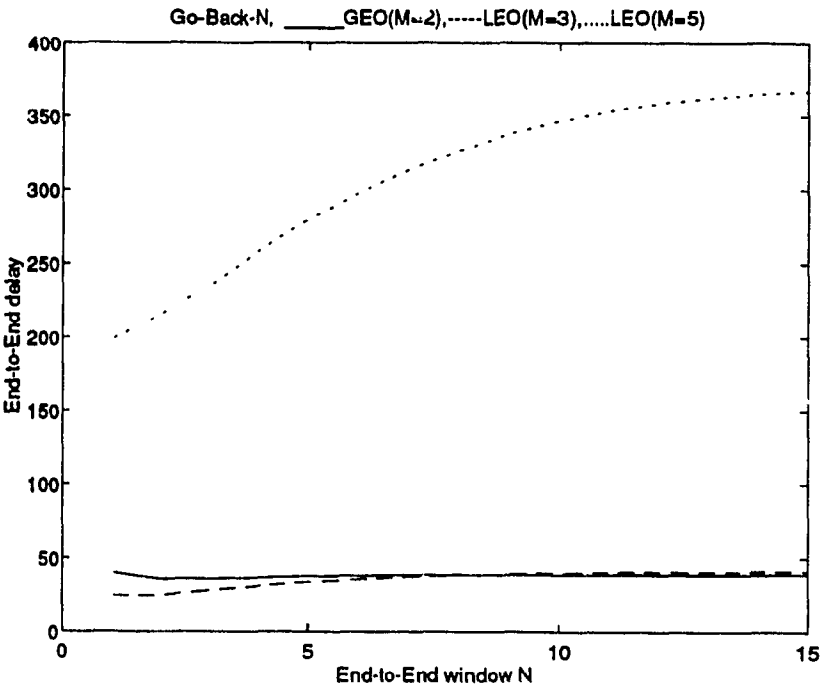

Fig. 2.6 (b) GBN scheme, P_b=10^{-3}, ρ=0.5;

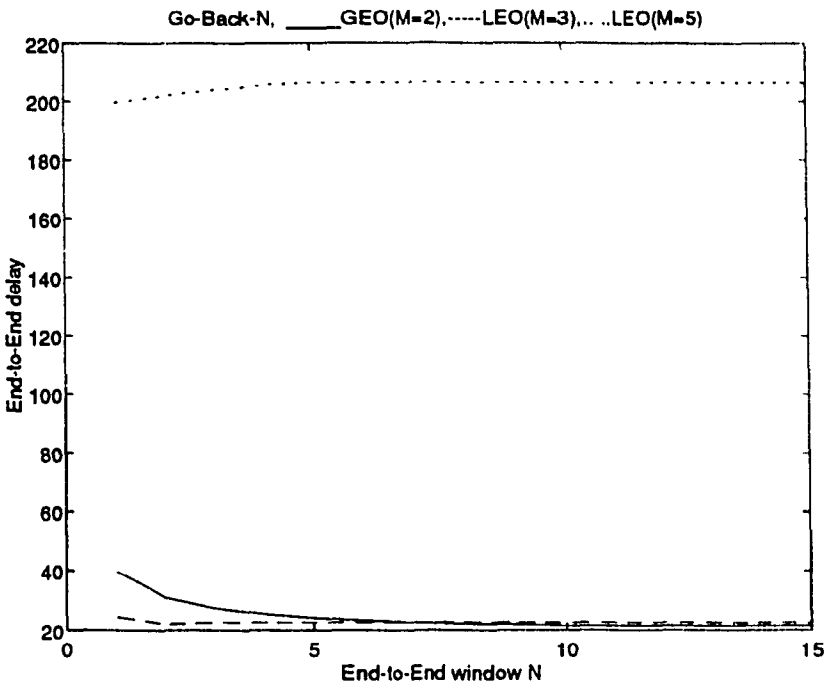

Fig. 2.6 (c) GBN scheme, P_b=10^{-3}, ρ=0.1;

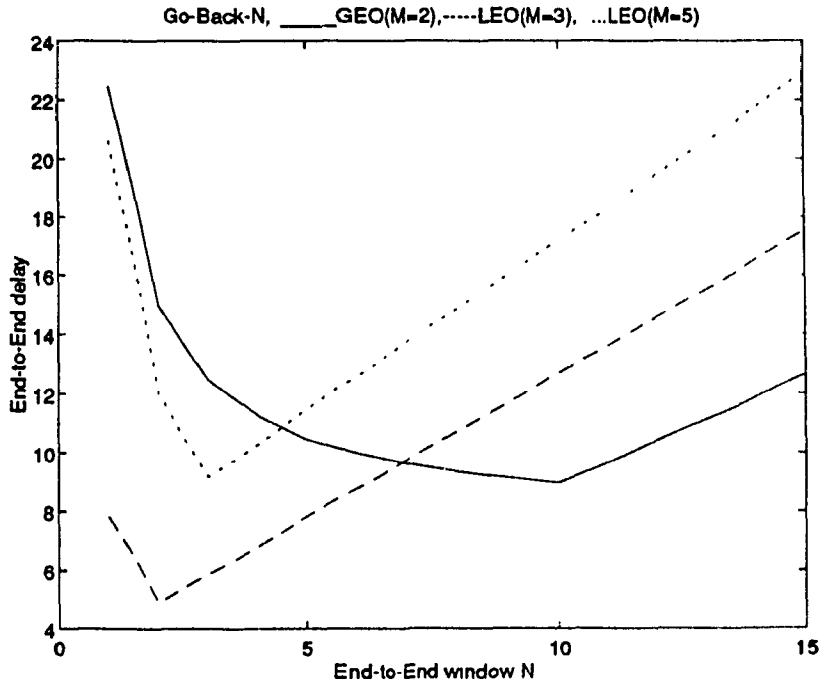

Fig. 2.6 (d) GBN scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.6. Delay for Go-Back-N scheme with k=568 bits, n_h=56 bits, R=9600 bps.

comparison. The end-to-end delay results in Fig. 2.6 correspond to the same set of the protocol parameters that resulted the end-to-end throughput characteristics in Fig. 2.5. These end-to-end delay results clearly show that the delay decreases as long as the window transmission time is within the round-trip propagation delay (i.e., $N < 2a+1$), otherwise the delay increases (when $N \geq 2a+1$), where N is the window size. The lowest end-to-end delay can be observed when the window (N) is equal to $2a+1$. Moreover, the end-to-end delay decreases as the window (N) reaches the value $2a+1$, and almost saturated after a certain window size (i.e., $N \geq 2a+1$), as the acknowledgment can be received within the window transmission and propagation delays. It can be observed from Figs 2.6 (a) and (d) that the end-to-end delay significantly improved when the P_b is decreased, which is due to less number of end-to-end packet retransmissions. It can also be observed from Fig. 2.6 that the 2 LEO satellite configuration (M=3) provides better throughput results compared to 1 GEO satellite configuration (M=2), or 4 LEO satellite configuration (M=5), for lower window sizes. The 2 LEO satellite (M=3), and 1 GEO satellite (M=2) configurations provide better end-to-end delay characteristics for obvious reasons, i.e., less number of hops, compared to 4 LEO satellite configuration (M=5).

The end-to-end throughput characteristics were obtained as a function of window size (N) for SR ARQ scheme in Fig. 2.7. These results were obtained for 9600 bps VC, with a packet data size of 64 bytes and packet header size of 7 bytes. These results correspond to the same link level configuration parameters of Fig. 2.5 for comparison

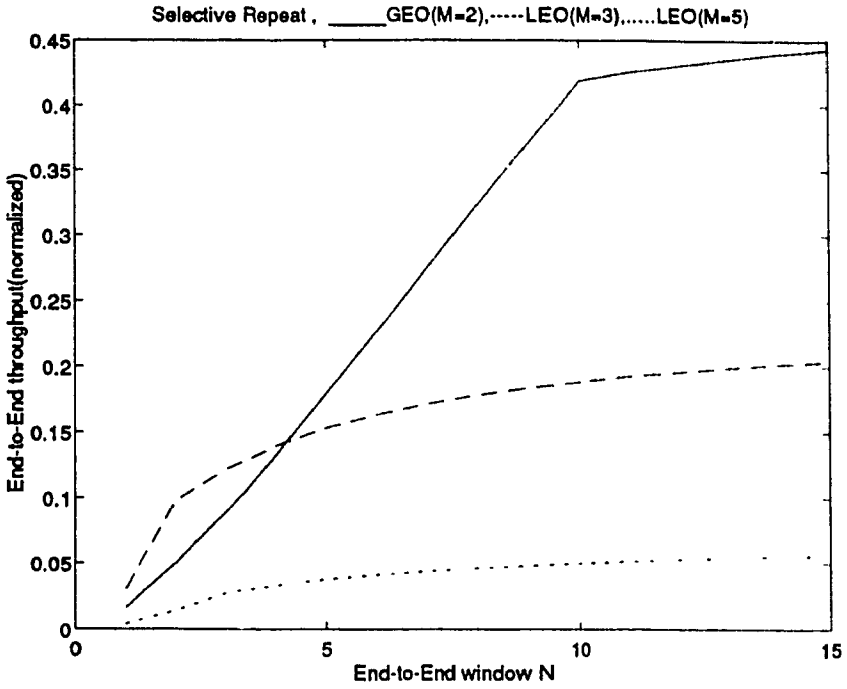

Fig. 2.7 (a) SR scheme, $P_b=10^{-3}$, ρ=0.99;

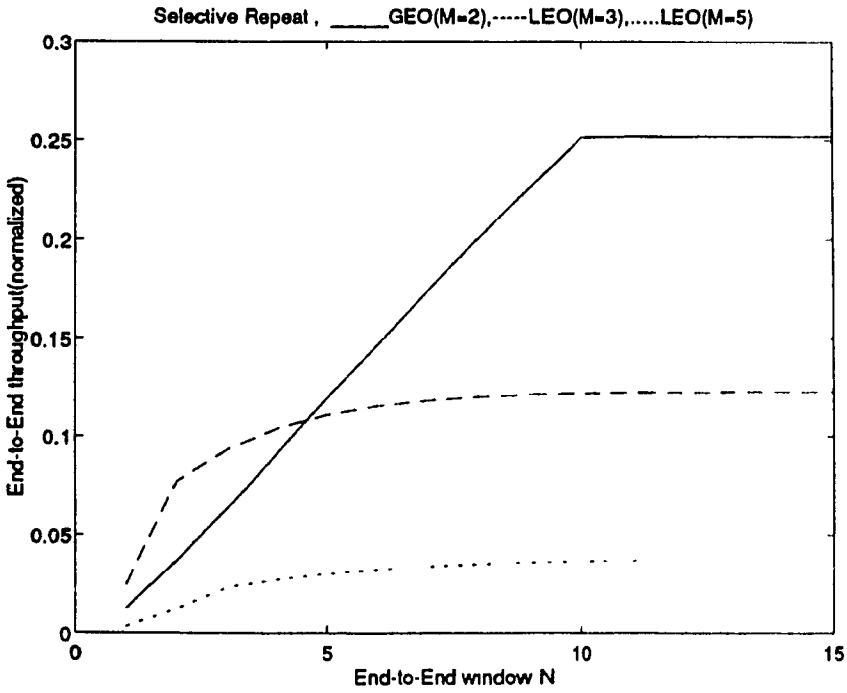

Fig. 2.7 (b) SR scheme, $P_b=10^{-3}$, ρ=0.5;

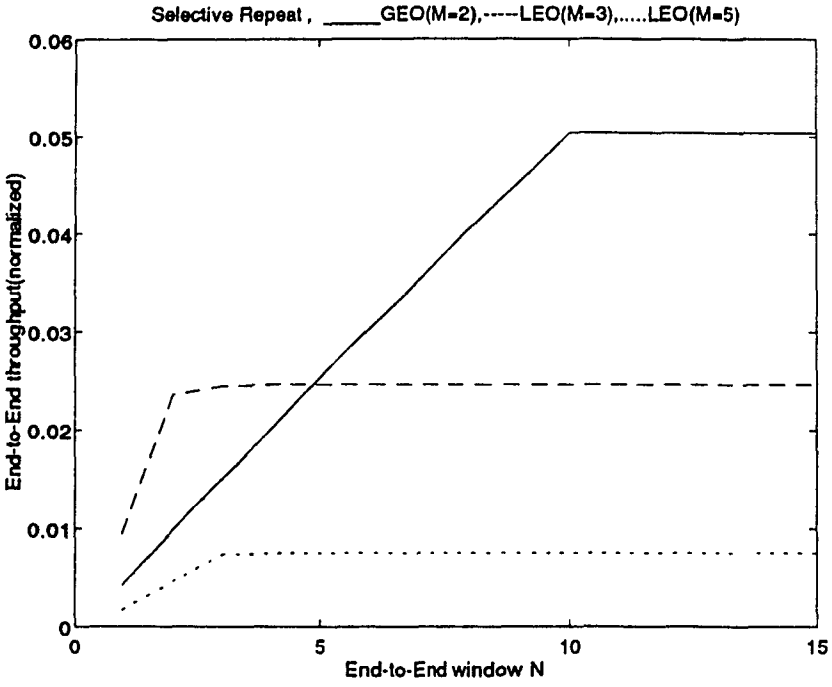

Fig. 2.7 (c) SR scheme, P_b=10^{-3}, ρ=0.1;

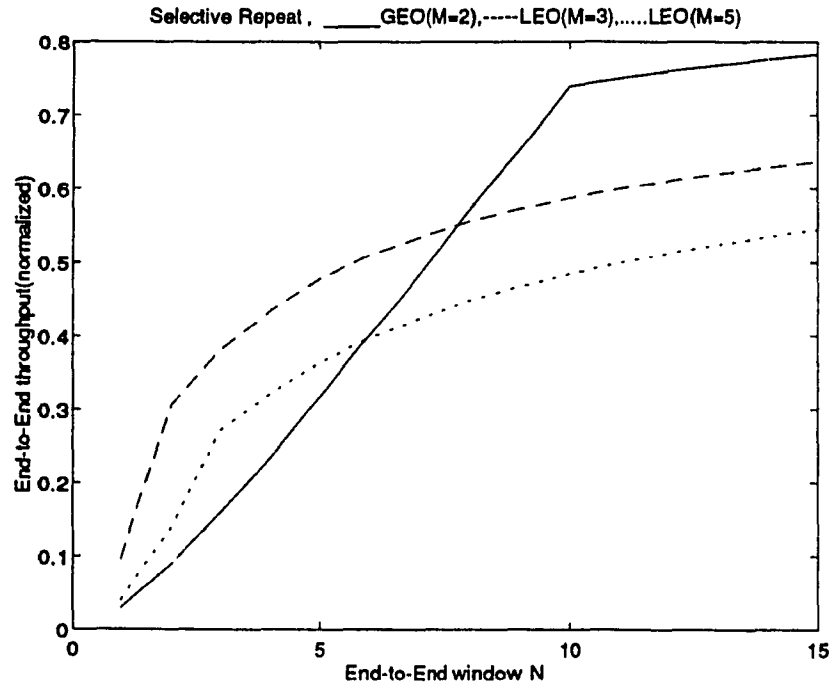

Fig. 2.7 (d) SR scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.7. Throughput for SR scheme with k=568 bits, n_h=56 bits, R=9600 bps.

purpose. The results in Figs 2.7 (a), (b), and (c) were obtained with a probability of bit error (P_b) of 10^{-3} for traffic intensities, i.e., 0.99, 0.5 and 0.1 respectively. The results in Fig. 2.7 (d) were obtained with a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. The results clearly show that the end-to-end throughput increases as the window size is increased as long as the window transmission time is within the round-trip propagation delay (i.e., $N < 2a+1$). The end-to-end throughput almost saturated after a certain window size (i.e., $N \geq 2a+1$), as the acknowledgment can be received within the window transmission and propagation delays. When $N \geq 2a+1$, further improvement in the results can only be observed by decreasing P_b, and by increasing the packet size. The LEO altitude range is 300-3000 K.M. By avoiding Van Allen radiation (which occurs at 1100-19000 K.M.), and by choosing proper orbit altitude, the transmission path loss and delay can be minimized, which can improve P_b. Furthermore, high quality transponders, transmitters, and receivers may also help to reduce the overall P_b. It can be observed from Figs 2.7 (a) and (d) that the end-to-end throughput significantly improved when the P_b is decreased. It is consistently observed from Fig. 2.7 that the 2 LEO satellite configuration (M=3) provides better throughput results compared to 1 GEO satellite configuration (M=2), or 4 LEO satellite configuration (M=5), at lower window sizes. When P_b is as high as 10^{-3}, the SR ARQ scheme provides better performance than GBN scheme as observed in Fig. 2.5 and Fig. 2.7. Both the GBN and SR ARQ schemes resulted identical throughput characteristics, when P_b is as low as 10^{-6}.

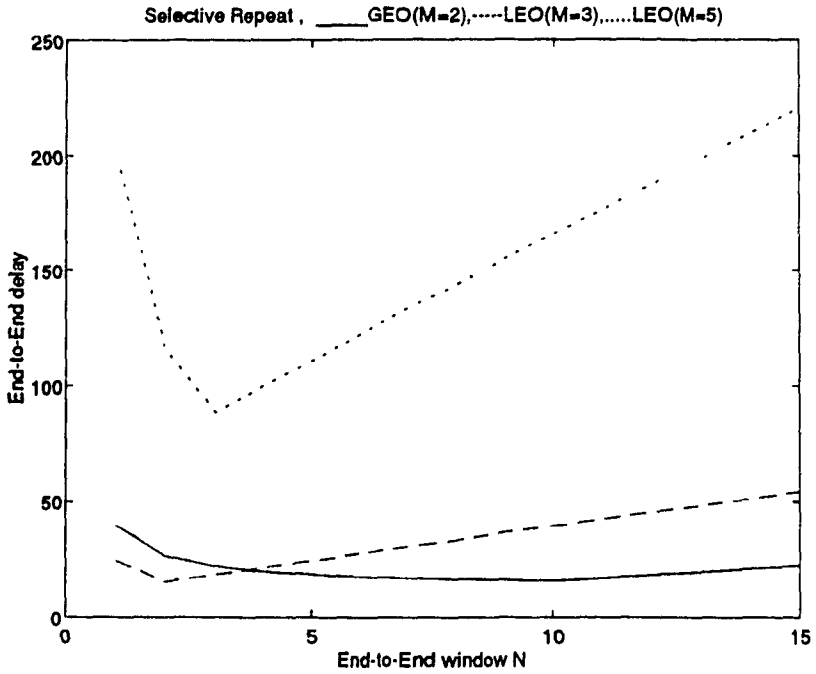

Fig. 2.8 (a) SR scheme, P_b=10^{-3}, ρ=0.99;

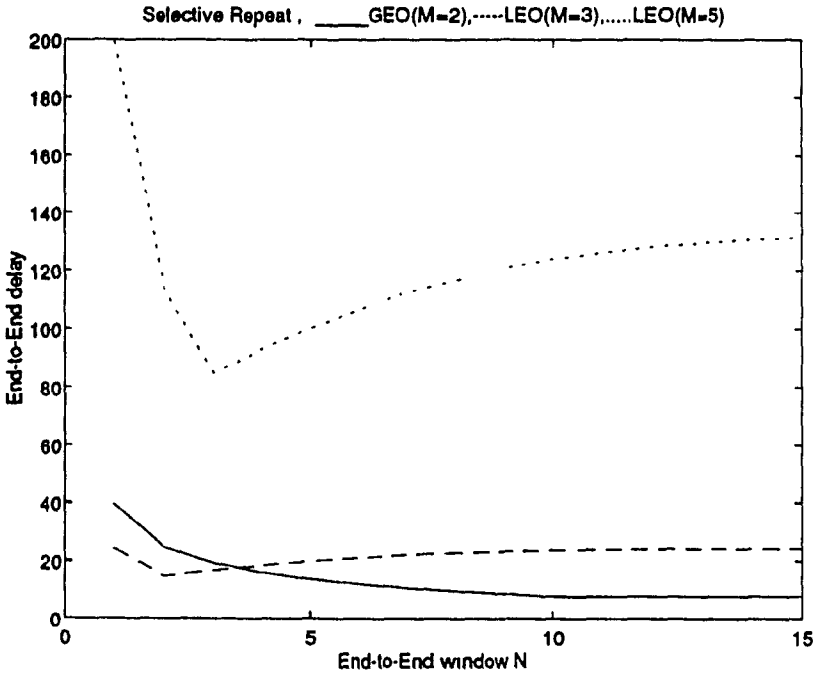

Fig. 2.8 (b) SR scheme, P_b=10^{-3}, ρ=0.5;

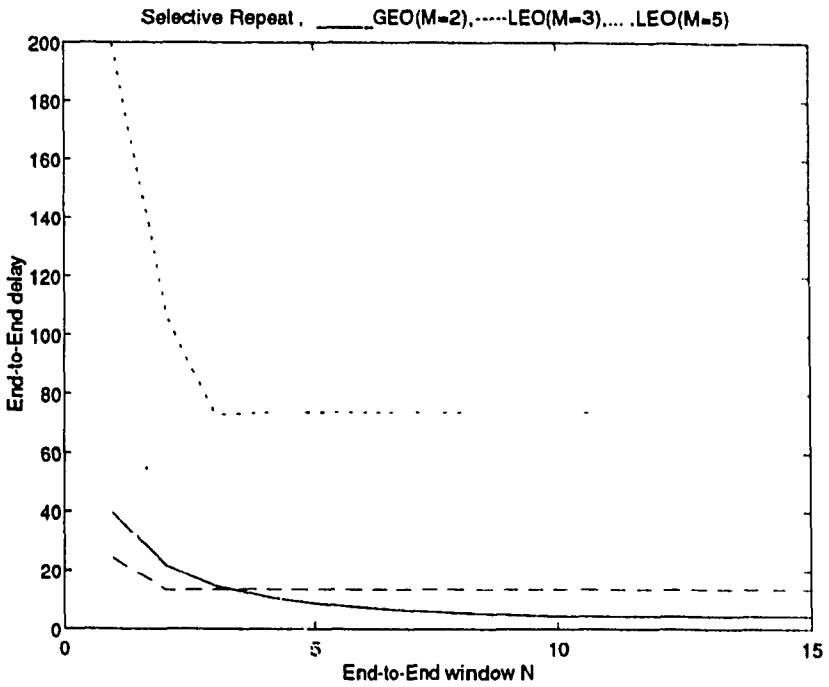

Fig. 2.8 (c) SR scheme, P_b=10^{-3}, ρ=0.1;

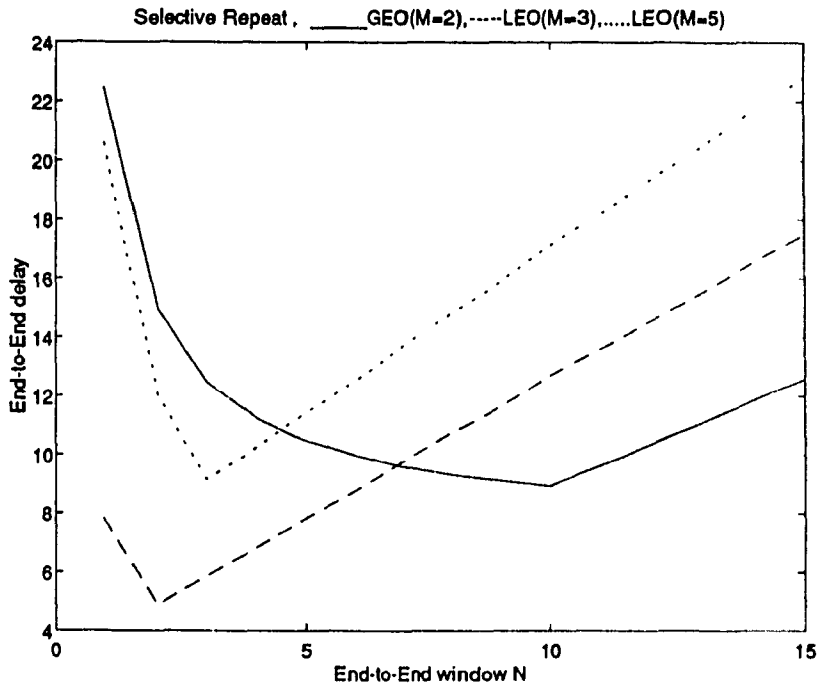

Fig. 2.8 (d) SR scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.8. Delay for SR scheme with k=568 bits, n_h=56 bits, R=9600 bps.

The end-to-end delay characteristics were obtained as a function of window size (N) for SR ARQ scheme in Fig. 2.8. These results were obtained for 9600 bps VC, with a packet data size of 64 bytes and packet header size of 7 bytes. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5, and 0.1 in Figs 2.8 (a), (b), and (c), respectively. The results in Fig. 2.8 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. The end-to-end delay results in Fig. 2.8 correspond to the same set of the protocol parameters that resulted the end-to-end throughput characteristics in Fig. 2.7 for SR ARQ. Moreover, these results were obtained for the same link configurations of the GBN ARQ scheme, for which the end-to-end delay characteristics were obtained in Fig. 2.6. These end-to-end delay results clearly show that the delay decreases when $N < 2a+1$, and increases when $N \geq 2a+1$, where N is the window size. The lowest end-to-end delay can be observed when the window (N) is equal to $2a+1$. Moreover, the end-to-end delay decreases as the window (N) reaches the value $2a+1$, and almost saturated after a certain window size (i.e., $N \geq 2a+1$), as the acknowledgment can be received within the window transmission and propagation delays. It can be observed from Figs 2.8 (a) and (d) that the end-to-end delay significantly improved when the P_b is decreased. It can also be observed from Fig. 2.6 that the 2 LEO satellite configuration (M=3) provides better throughput results compared to 1 GEO satellite configuration (M=2), or 4 LEO satellite configuration (M=5), at lower window sizes. The 2 LEO satellite (M=3), and 1 GEO satellite (M=2) configurations provide better end-to-end delay for obvious reasons, i.e., less number of hops, when compared to 4 LEO satellite configuration (M=5).

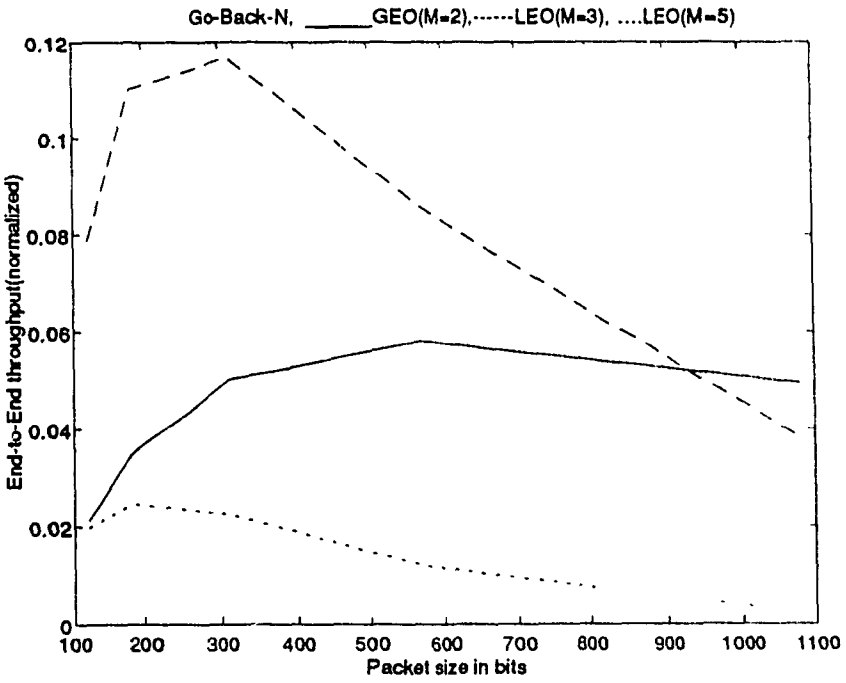

Fig. 2.9 (a) GBN scheme, P_b=10^{-3}, ρ=0.99;

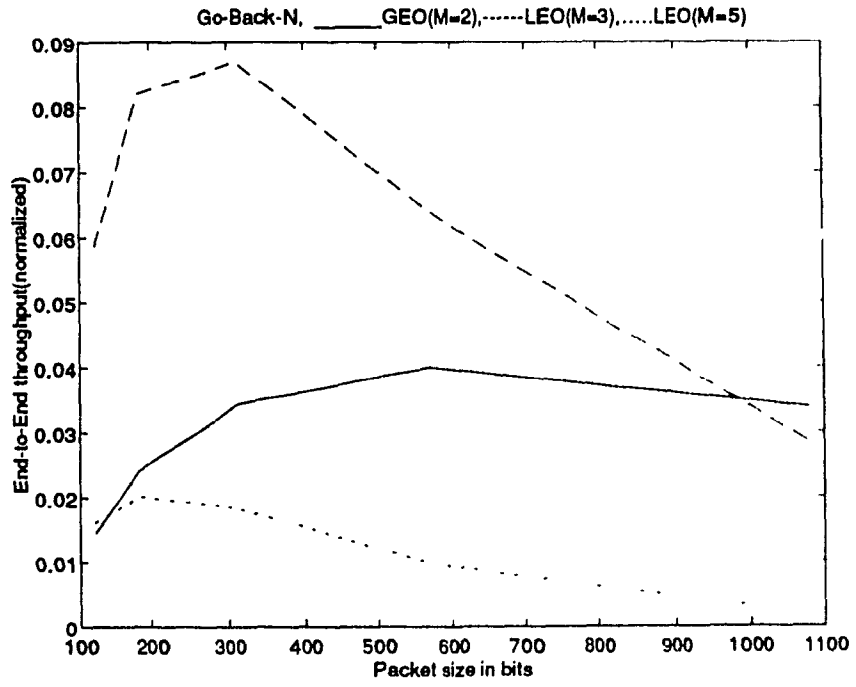

Fig. 2.9 (b) GBN scheme, P_b=10^{-3}, ρ=0.5;

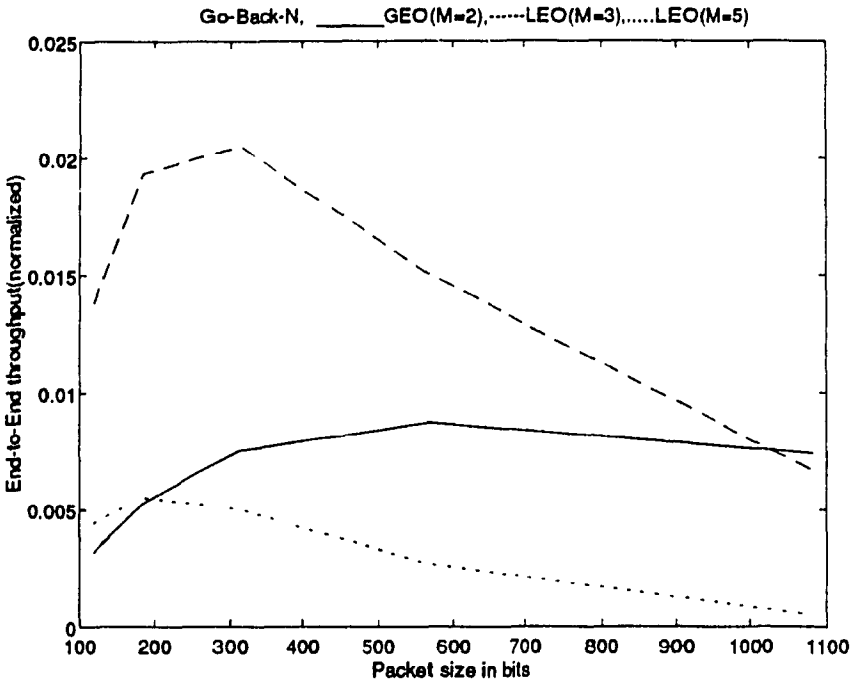

Fig. 2.9 (c) GBN scheme, P_b=10^{-3}, ρ=0.1;

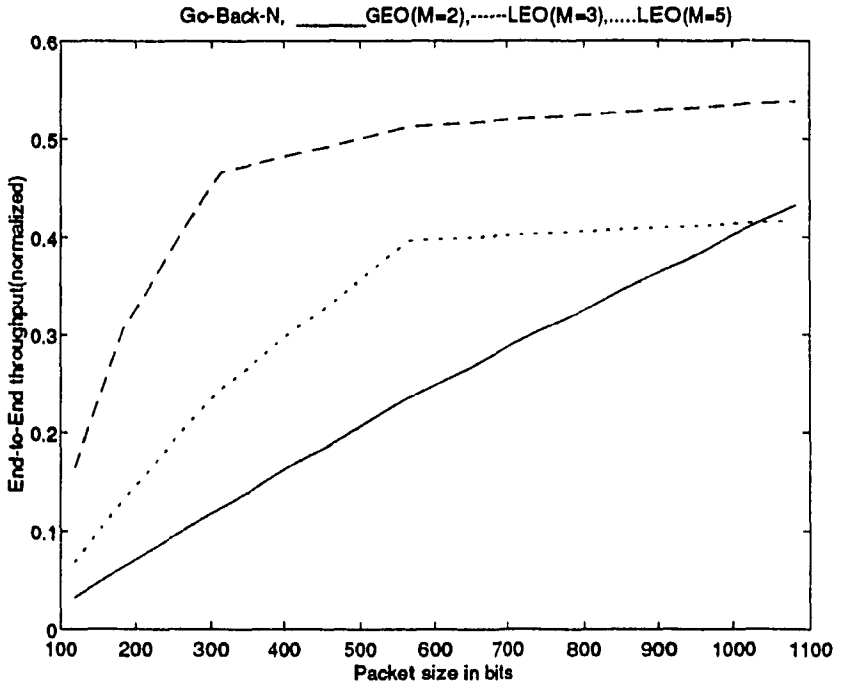

Fig. 2.9 (d) GBN scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.9. Throughput for Go-Back-N scheme with N=4, n_h=56 bits, R=9600 bps.

The end-to-end throughput characteristics were obtained as a function of packet size (data + header) for GBN ARQ scheme in Fig. 2.9. These results were obtained for 9600 bps VC, with a window size of 4 and a packet header size of 7 bytes. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.9 (a), (b), and (c), respectively. The results in Fig. 2.9 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. The results clearly show that the end-to-end throughput increases as the packet size is increased up to a certain value (i.e., $N=2a+1$). As the packet size is increased furthermore the probability of packet error increases, and hence the throughput decreases as observed in the results. But the end-to-end throughput saturates after a certain packet size (i.e., $N \geq 2a+1$), as the acknowledgment can be received within the window transmission and propagation delays. In such cases, optimum throughput results can be obtained by decreasing P_b, or by carefully choosing the packet size and window size in such a way that the transmitted packet get acknowledged just before complete transmission of the window, such that the window rotates smoothly. When $N = 2a+1$, the end-to-end window rotates smoothly, by receiving acknowledgment just before complete transmission of the window, contributing to stream traffic over the VC. It can be observed from Figs 2.9 (a) and (d) that the end-to-end throughput significantly improved when the P_b is decreased. It can also be observed from Fig. 2.9 that the 2 LEO satellite (M=3) configuration provides better throughput results compared to 1 GEO satellite (M=2) configuration, or 4 LEO satellite (M=5) configuration.

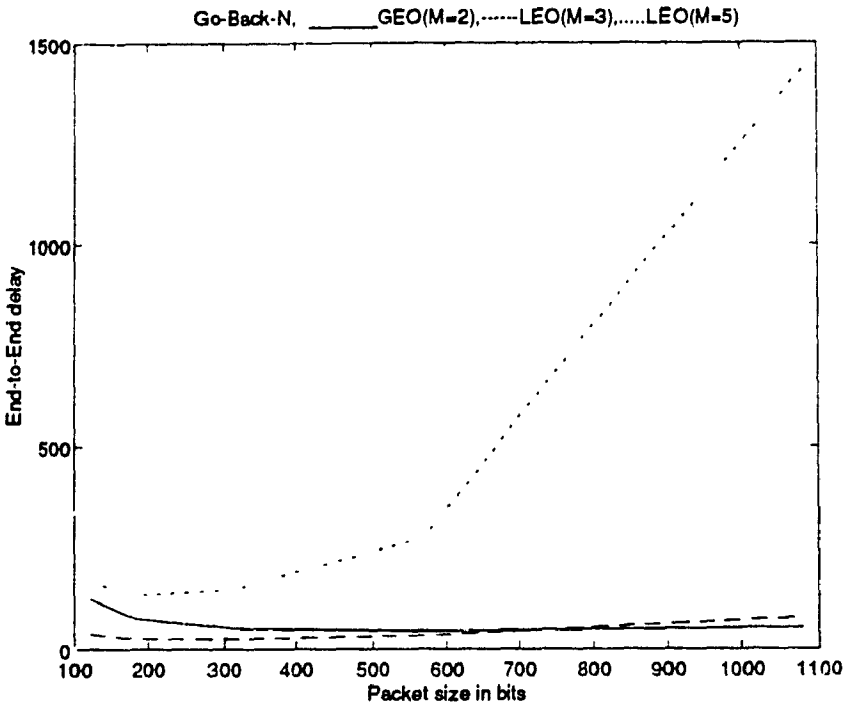

Fig. 2.10 (a) GBN scheme, P_b=10^{-3}, ρ=0.99;

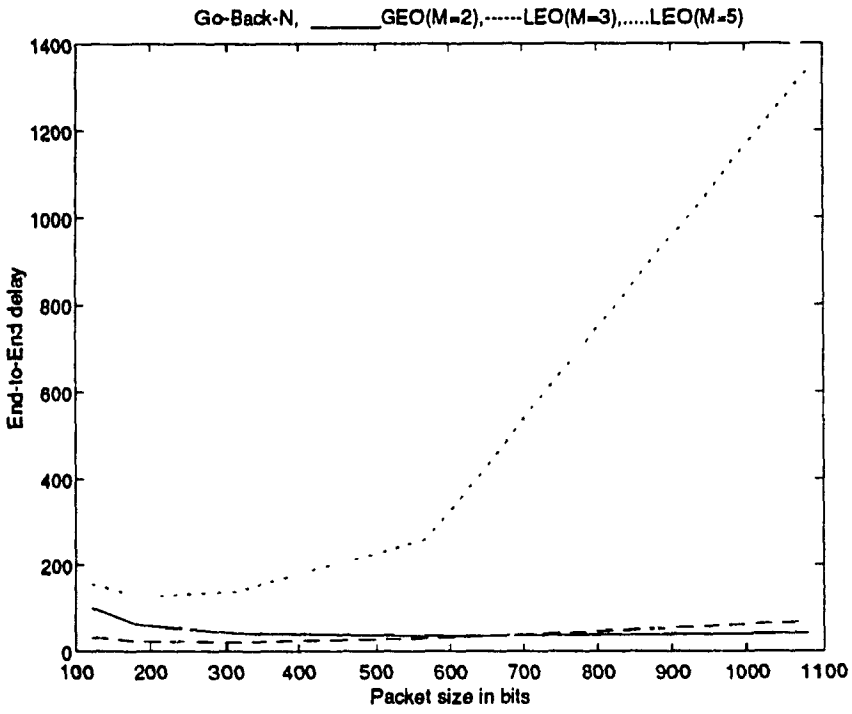

Fig. 2.10 (b) GBN scheme, P_b=10^{-3}, ρ=0.5;

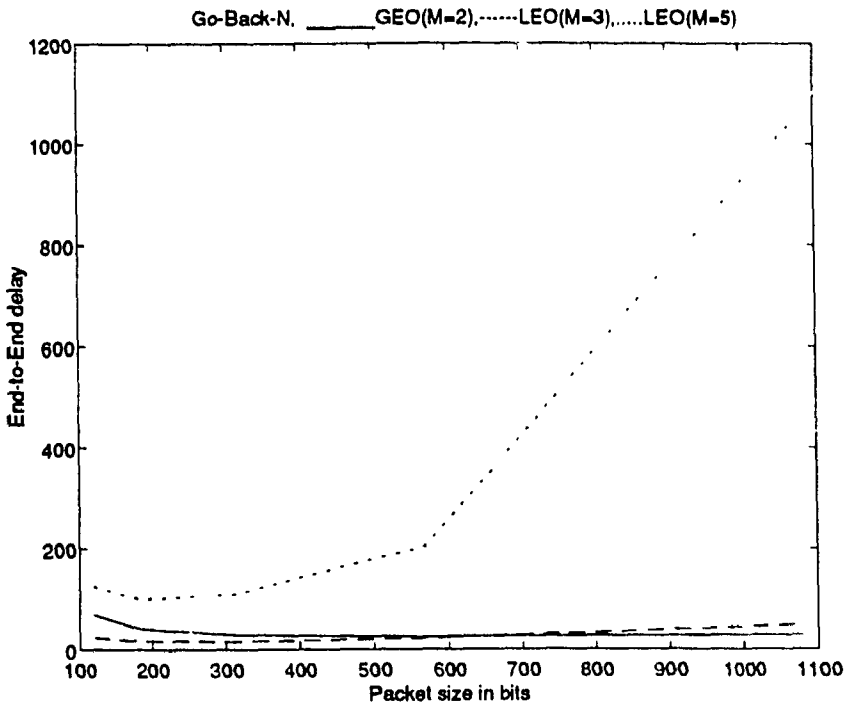

Fig. 2.10 (c) GBN scheme, P_b=10^{-3}, ρ=0.1;

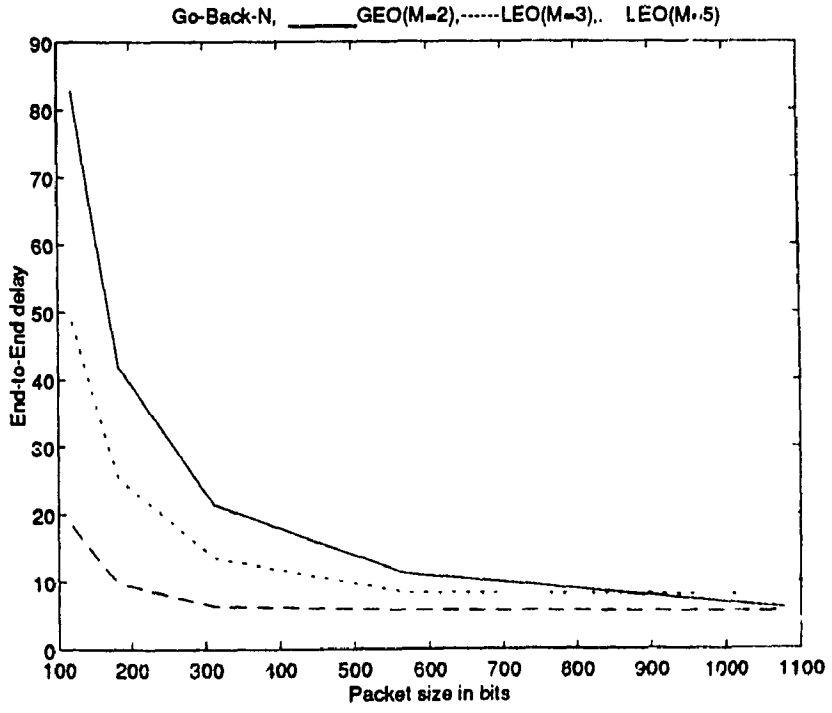

Fig. 2.10 (d) GBN scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.10. Delay for Go-Back-N scheme with N=4, n_h=56 bits, R=9600 bps.

The end-to-end delay characteristics were obtained as a function of packet size (data + header) for GBN ARQ scheme in Fig. 2.10. These results were obtained for 9600 bps VC, with a window size of 4 and a packet header size of 7 bytes. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.10 (a), (b), and (c), respectively. The results in Fig. 2.10 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. It can be noted that the end-to-end delay results in Fig. 2.10 correspond to the same set of the protocol parameters that resulted the end-to-end throughput characteristics in Fig. 2.9. When the probability of bit error is as high as 10^{-3}, the results in Figs 2.10 (a), (b), and (c) clearly show that the end-to-end delay increases as the packet size is increased. However, these results show that 1 GEO satellite (M=2), and 2 LEO satellite (M=3) configurations provide lesser end-to-end delay compared to 4 LEO satellite (M=5) configuration, due to several hops before reaching the destination. As P_b is decreased to 10^{-6} (Fig. 2.10 (d)), the results exhibited lower end-to-end delay as the packet size is increased. Moreover, both 2 LEO satellite configuration and 4 LEO satellite configuration provided lower end-to-end delay when P_b is 10^{-6}. Hence, the probability of bit error (P_b) plays a major role to control end-to-end delay due to multiple end-to-end re-transmissions when P_b is high.

The end-to-end throughput characteristics were obtained as a function of packet size (data + header) for SR ARQ scheme in Fig. 2.11. The results were obtained for 9600 bps VC, with a window size of 4 and a packet header size of 7 bytes. These results were

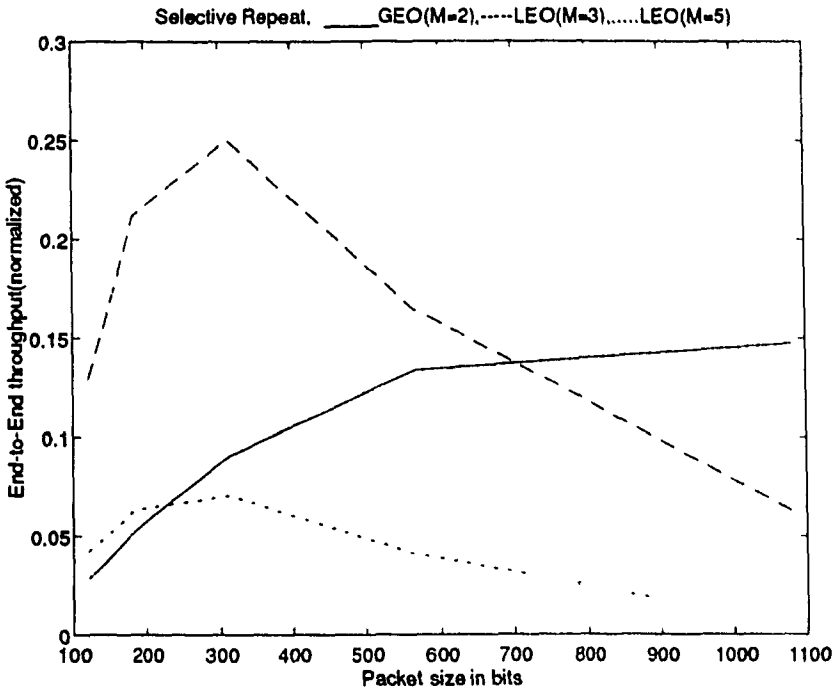

Fig. 2.11 (a) SR scheme, $P_b=10^{-3}$, ρ=0.99;

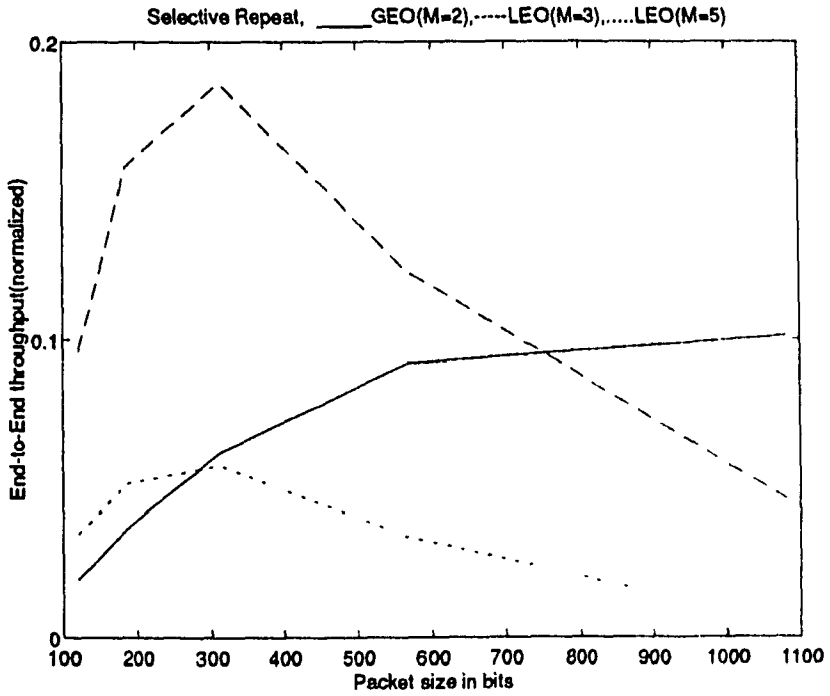

Fig. 2.11 (b) SR scheme, $P_b=10^{-3}$, ρ=0.5;

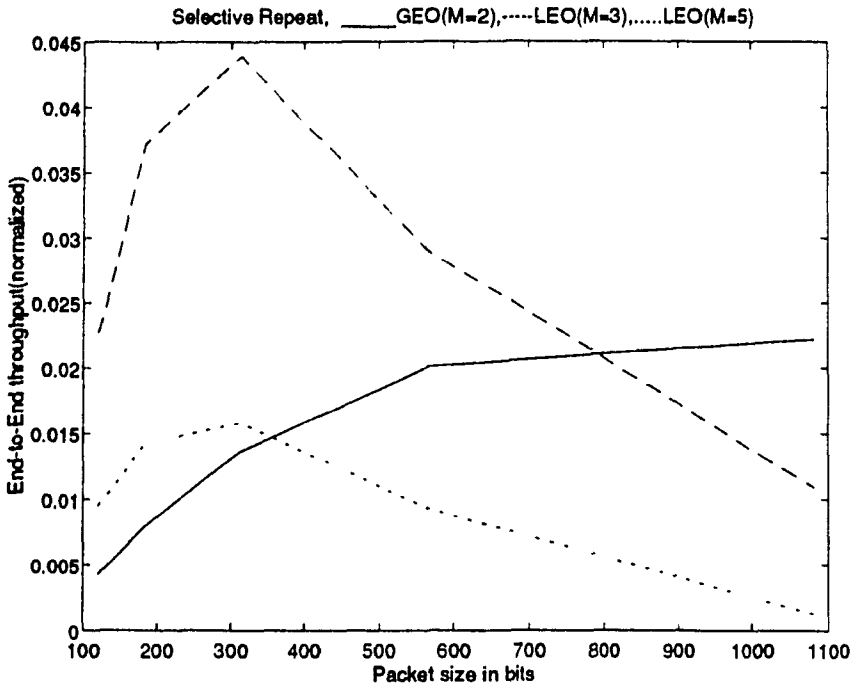

Fig. 2.11 (c) SR scheme, P_b=10^{-3}, ρ=0.1;

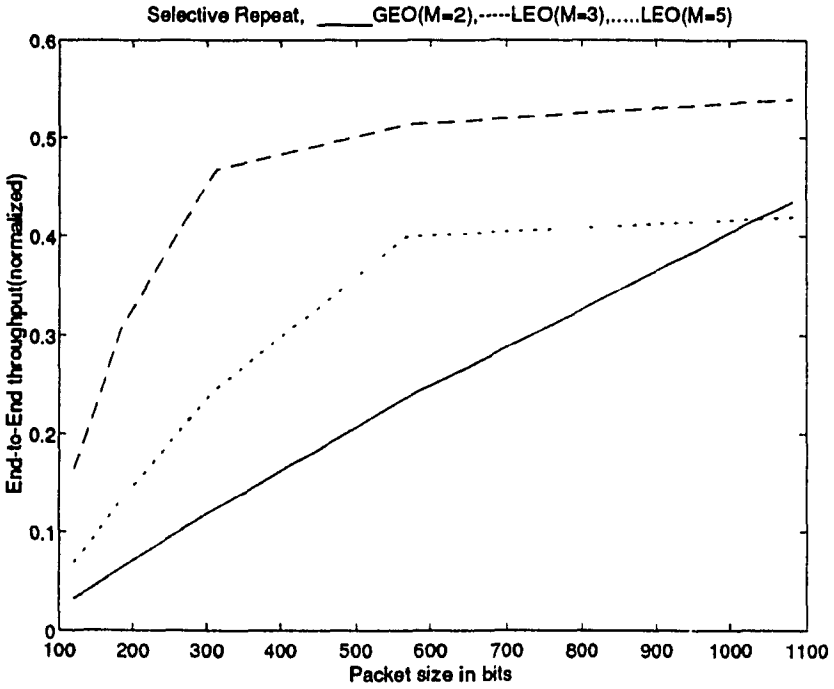

Fig. 2.11 (d) SR scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.11. Throughput for SR scheme with N=4, n_h=56 bits, R=9600 bps.

obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities in Figs 2.11 (a), (b), and (c), respectively. The results in Fig. 2.11 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. The results clearly show that the end-to-end throughput increases as the packet size is increased until $N = 2a+1$. As the packet size is increased furthermore, the probability of packet error increases, and hence the throughput decreases as observed in the results. However, SR scheme provided better throughput when P_b is high, which can be observed in Figs 2.11 (a), (b), and (c). The results in 2.11 (d) do not show significant difference from GBN scheme which were obtained in Fig. 2.9 (d), due to the fact that both schemes provide better end-to-end throughput. Moreover, 2 LEO satellite (M=3) configuration provided better throughput results at lower packet sizes. However, both 2 LEO satellite (M=3) and 4 LEO satellite (M=5) configurations provided better throughput results compared to 1 GEO satellite (M=2) configuration, when P_b is as low as 10^{-6}, which can be observed in Fig. 2.11 (d).

The end-to-end delay characteristics were obtained as a function of packet size (data + header) for SR ARQ scheme in Fig. 2.12. These results were obtained for 9600 bps VC, with a window size of 4 and a packet header size of 7 bytes. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.12 (a), (b), and (c), respectively. The results in Fig. 2.12 (d), were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. It can be noted that the end-to-end delay results in Fig. 2.12 correspond to

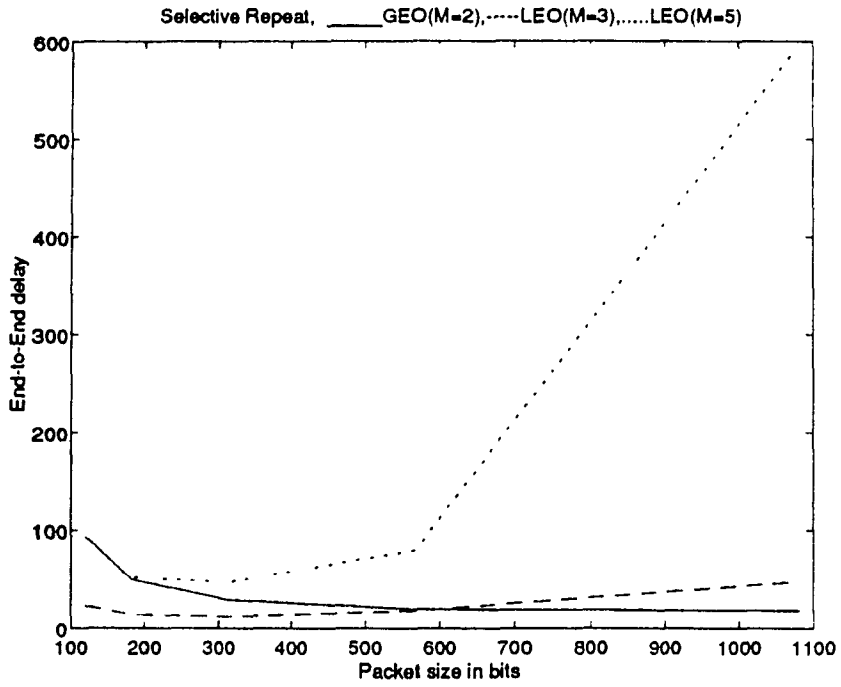

Fig. 2.12 (a) SR scheme, $P_b=10^{-3}$, $\rho=0.99$;

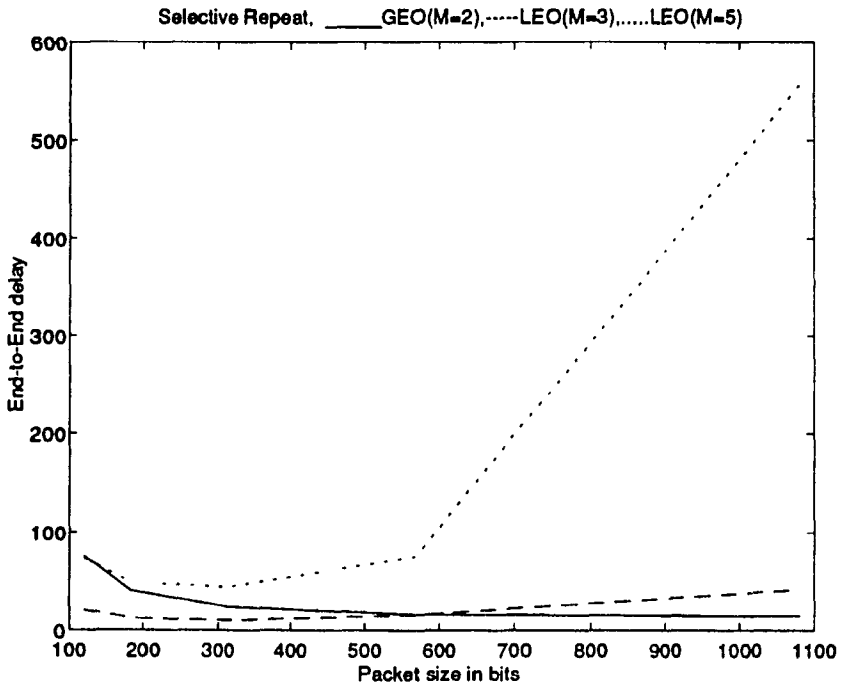

Fig. 2.12 (b) SR scheme, $P_b=10^{-3}$, $\rho=0.5$;

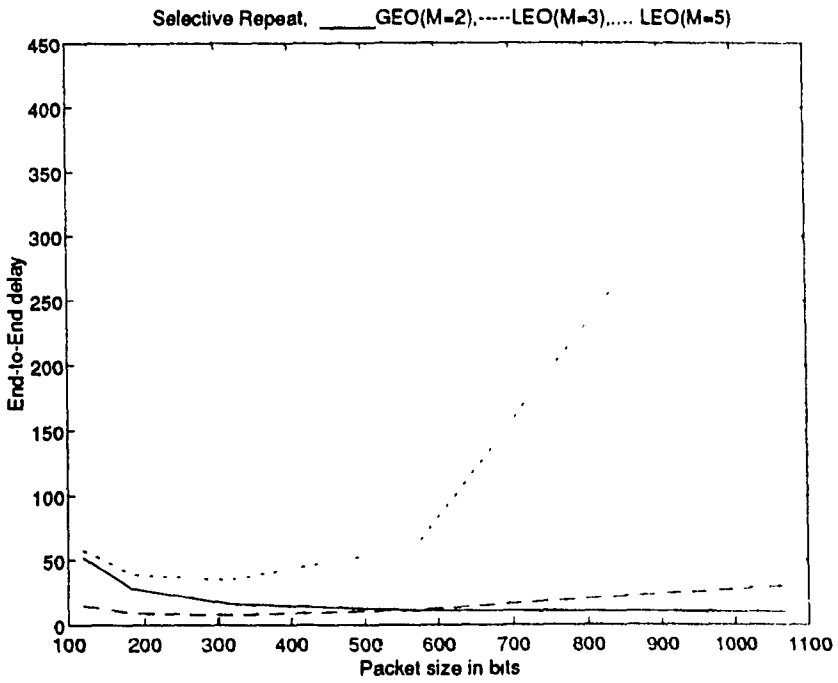

Fig. 2.12 (c) SR scheme, P_b=10^{-3}, ρ=0.1;

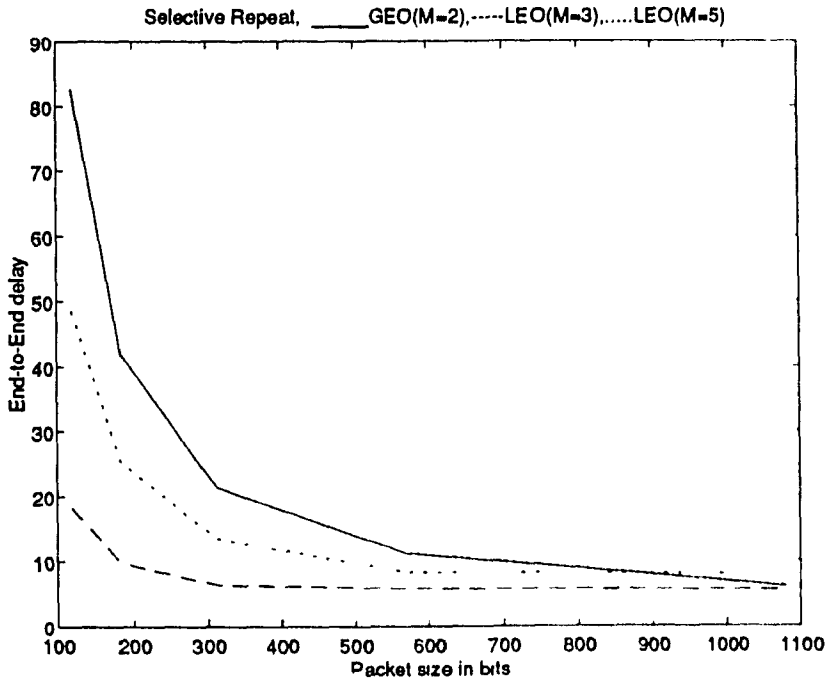

Fig. 2.12 (d) SR scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.12. Delay for SR scheme with N=4, n_h=56 bits, R=9600 bps.

the same set of the protocol parameters that resulted the end-to-end throughput characteristics in Fig. 2.11. When the P_b is as high as 10^{-3}, the results in Figs 2.12 (a), (b), and (c) clearly show that the end-to-end delay increases as the packet size is increased, which is due to several retransmissions due to packet errors. However, these results show that 1 GEO satellite (M=2), and 2 LEO satellite (M=3) configurations provide lesser end-to-end delay compared to 4 LEO satellite (M=5) configuration, due to several hops before reaching the destination. As P_b is decreased to 10^{-6}, the results in Fig. 2.12 (d) exhibited lower end-to-end delay as the packet size is increased, as several packets were transmitted correctly compared to the other Figs 2.12 (a), (b), and (c). Moreover, both 2 LEO satellite (M=3) configuration and 4 LEO satellite (M=5) configuration provided lower end-to-end delay when P_b is 10^{-6}. Hence, the probability of bit error (P_b) plays a major role to control end-to-end delay due to multiple end-to-end retransmissions when P_b is high. The end-to-end delay results of SR scheme in Figs 2.12 (a), (b), and (c) exhibit significant improvement over the respective GBN results in Figs 2.10 (a), (b), and (c), respectively. However, both GBN and SR ARQ schemes provided identical results when P_b is as low as 10^{-6}, due to less number of end-to-end packet retransmissions.

The end-to-end throughput characteristics were obtained as a function of VC transmission speed (in bits) for GBN ARQ scheme in Fig. 2.13. These results were obtained for a packet data size of 64 bytes, packet header size of 7 bytes, and a window size of 4. These results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.13 (a), (b), and (c), respectively.

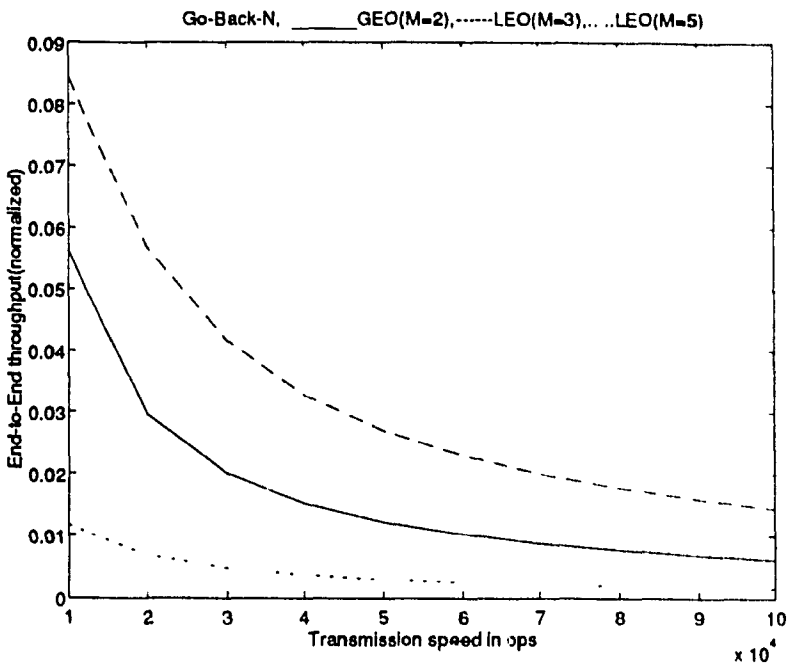

Fig. 2.13 (a) GBN scheme, P_b=10^{-3}, ρ=0.99;

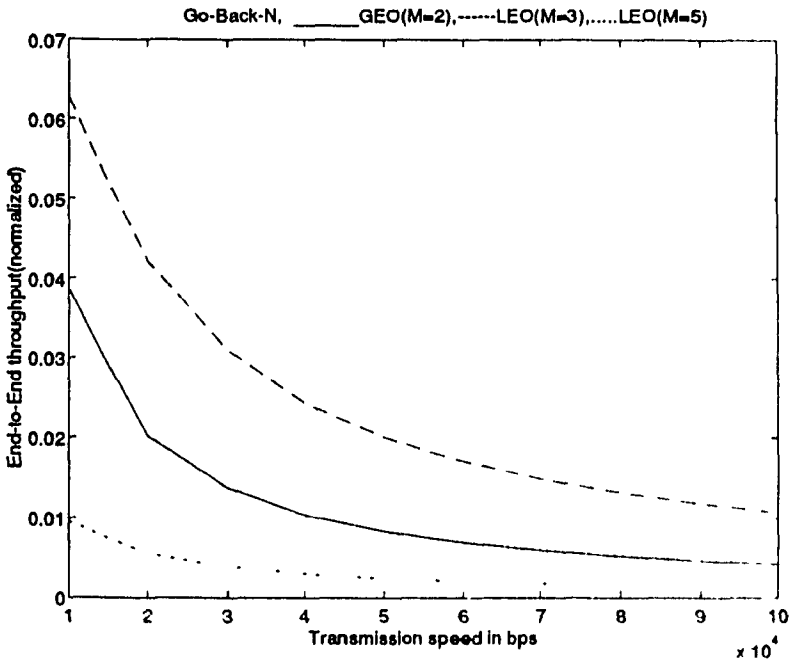

Fig. 2.13 (b) GBN scheme, P_b=10^{-3}, ρ=0.5;

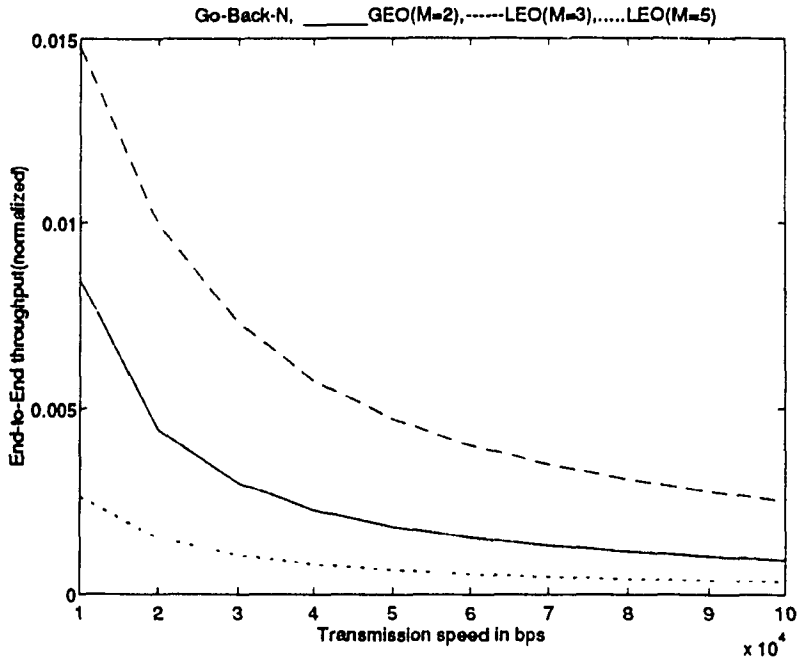

Fig. 2.13 (c) GBN scheme, P_b=10^{-3}, ρ=0.1;

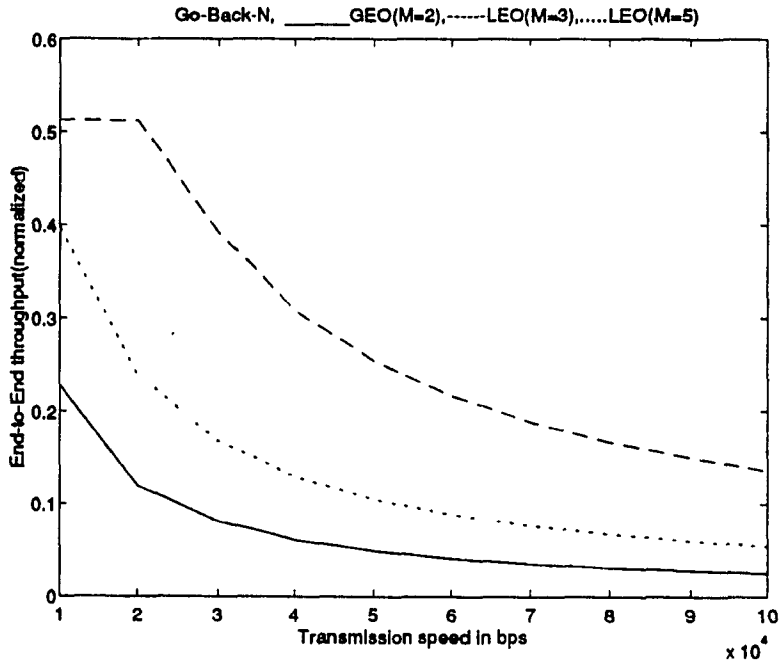

Fig. 2.13 (d) GBN scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.13. GBN Throughput vs Transmission speed with N=4, k=568 bits, n_h=56 bits.

The results in Fig. 2.13 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. For window ARQ schemes, the propagation delay plays a major drawback for higher transmission speeds. As the propagation delay is constant over a end-to-end virtual circuit, any increase in transmission speeds when $N \geq 1+2a$, provides constant utilization of the available bandwidth. But as the transmission speed is increased when $N < 1+2a$, the virtual circuit utilization is reduced, which in turn reduces the end-to-end throughput, as observed in Fig. 2.13. When P_b is reduced to 10^{-6}, better end-to-end throughput can be observed in Fig. 2.13 (d), compared to (a), (b), and (c). The 2 LEO satellite (M=3) configuration provided better throughput results in Fig. 2.13, compared to 1 GEO satellite(M=2), and 4 LEO satellite (M=5) configurations.

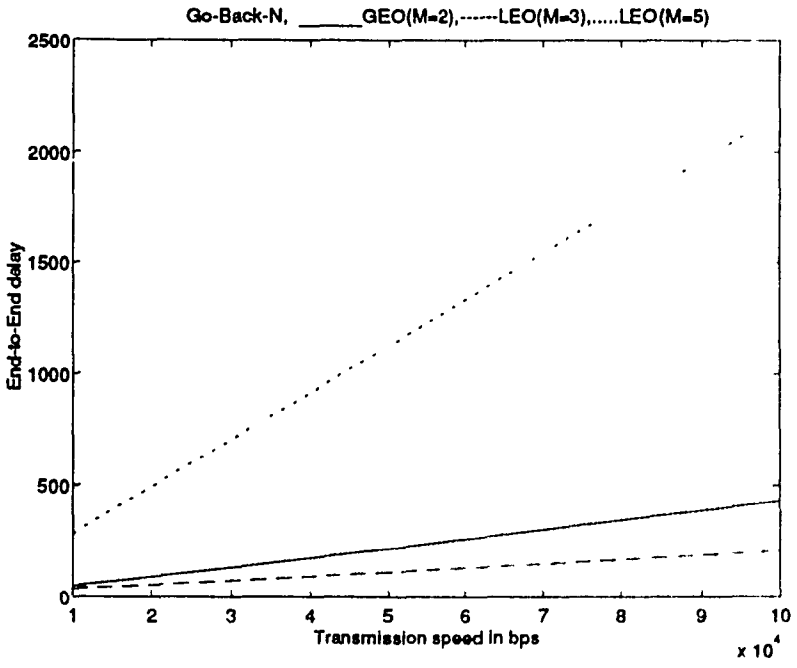

Fig. 2.14 (a) GBN scheme, P_b=10^{-3}, ρ=0.99;

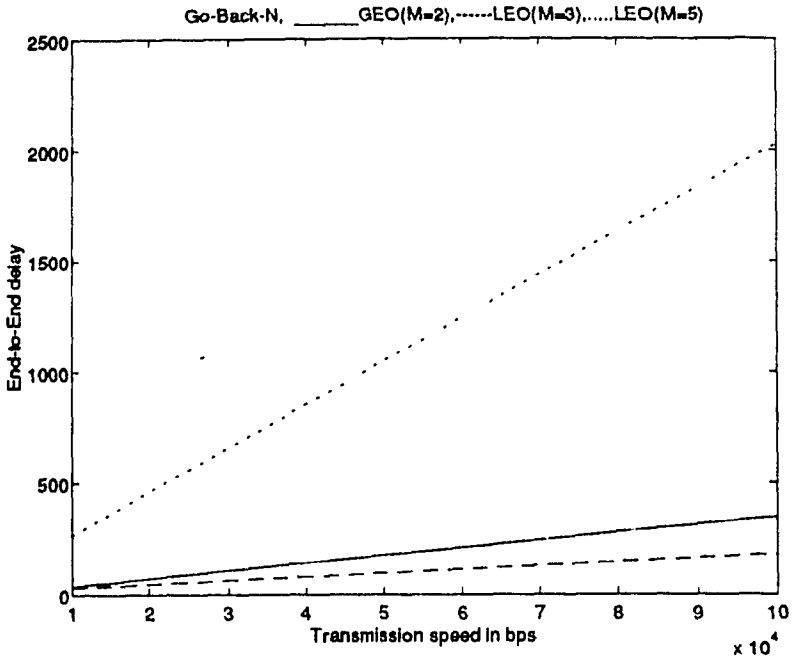

Fig. 2.14 (b) GBN scheme, P_b=10^{-3}, ρ=0.5;

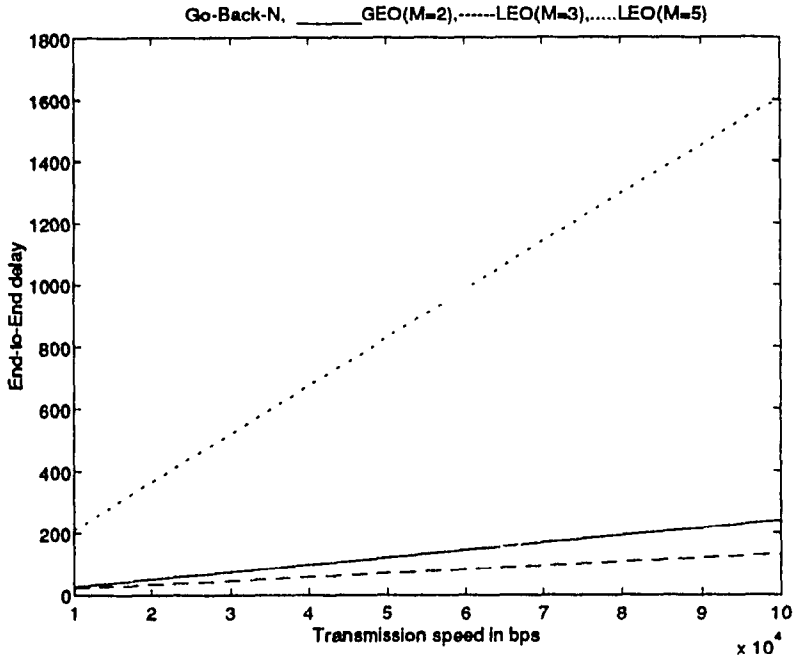

Fig. 2.14 (c) GBN scheme, P_b=10^{-3}, ρ=0.1;

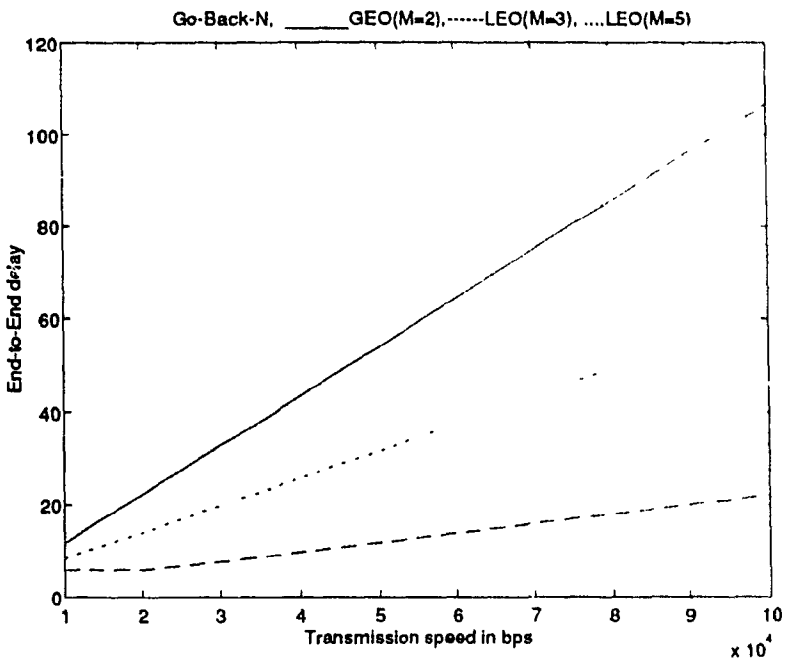

Fig. 2.14 (d) GBN scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.14. GBN Delay vs Transmission speed with N=4, k=568 bits, n_h=56 bits.

The end-to-end delay characteristics were obtained as a function of VC transmission speed (in bits) for GBN ARQ scheme in Fig. 2.14. These results were obtained for a packet data size of 64 bytes, packet header size of 7 bytes, and a window size of 4. The results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.14 (a), (b), and (c), respectively. The results in Fig. 2.14 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. As detailed earlier, the propagation delay plays a major drawback for higher transmission speeds in ARQ window schemes. As $N < 1+2a$ condition is quickly reached in Fig. 2.14, the increase in transmission speeds increased

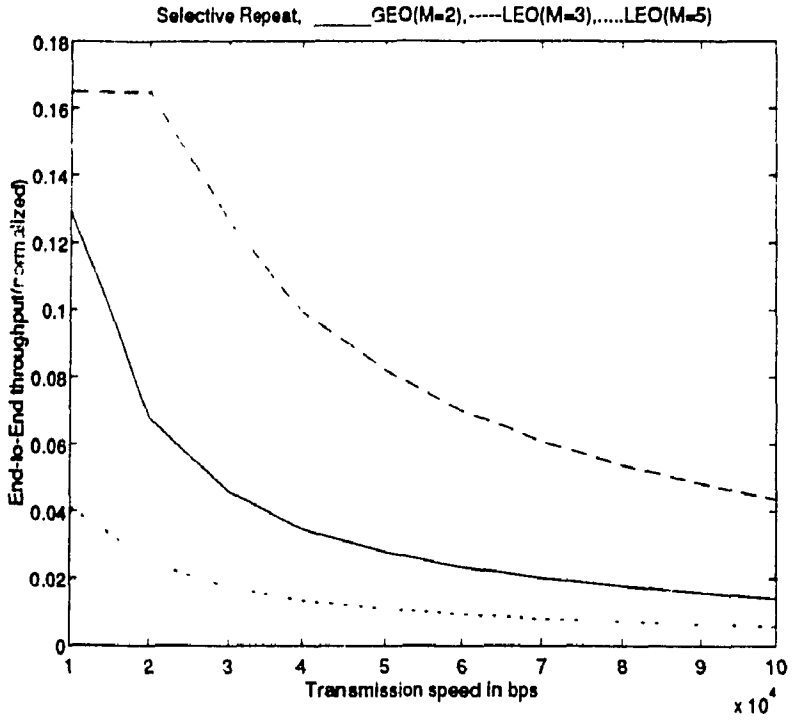

Fig. 2.15 (a) SR scheme, $P_b=10^{-3}$, ρ=0.99;

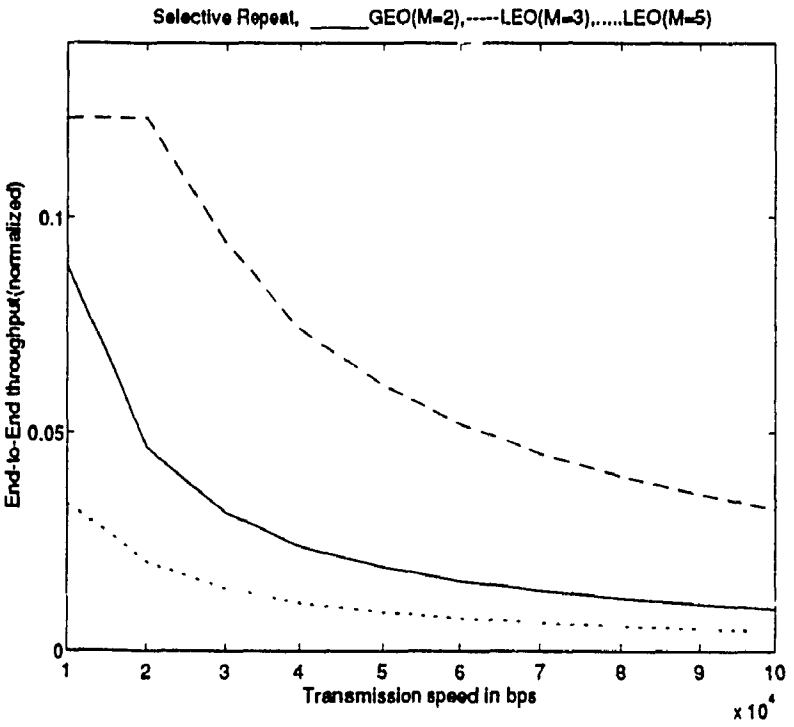

Fig. 2.15 (b) SR scheme, $P_b=10^{-3}$, ρ=0.5;

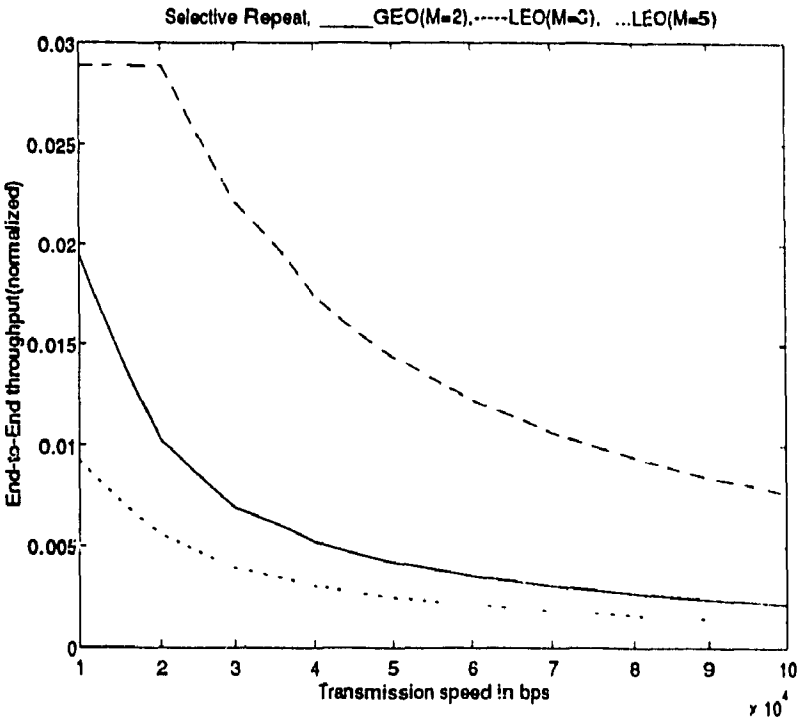

Fig. 2.15 (c) SR scheme, P_b=10^{-3}, ρ=0.1;

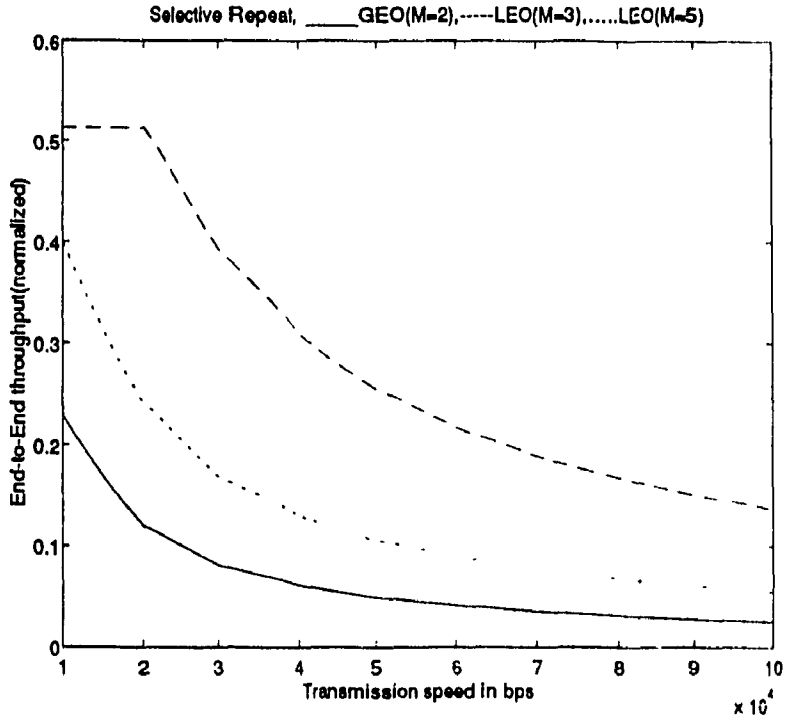

Fig. 2.15 (d) SR scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.15. SR throughput vs Transmission speed with N=4, k=568 bits, n_h=56 bits.

end-to-end delay. The end to end delay can be decreased by choosing a higher window size for better channel utilization. The 2 LEO satellite (M=3) configuration provided better (i.e., lesser) end-to-end delay in Fig. 2.14, compared to 1 GEO satellite (M=2), and 4 LEO satellite (M=5) configurations.

The end-to-end throughput characteristics were obtained as a function of VC transmission speed (in bits) for SR ARQ scheme in Fig. 2.15. These results were obtained for a packet data size of 64 bytes, packet header size of 7 bytes, and a window size of 4. The results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.15 (a), (b), and (c), respectively. The results in Fig. 2.15 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. The end-to-end throughput is decreased when the transmission speed is increased, as observed in Fig. 2.15. It can be clearly observed that better throughput is provided for SR scheme, when compared to GBN scheme. When P_b is reduced to 10^{-6}, better end-to-end throughput can be observed in Fig. 2.15 (d), compared to (a), (b), and (c). The 2 LEO satellite (M=3) configuration provided better throughput results in Fig. 2.15, compared to 1 GEO satellite (M=2), and 4 LEO satellite (M=5) configurations.

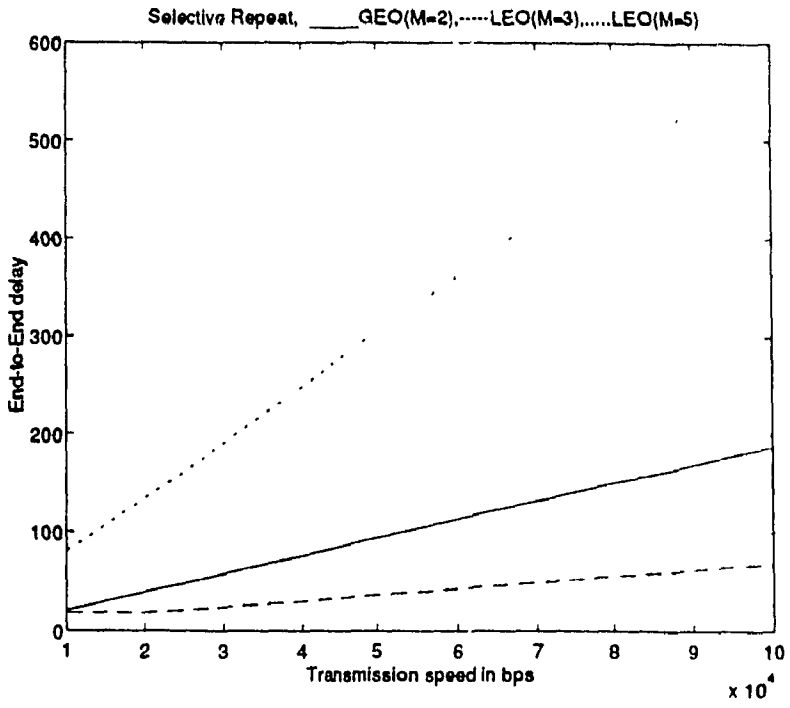

Fig. 2.16 (a) SR scheme, P_b=10^{-3}, ρ=0.99;

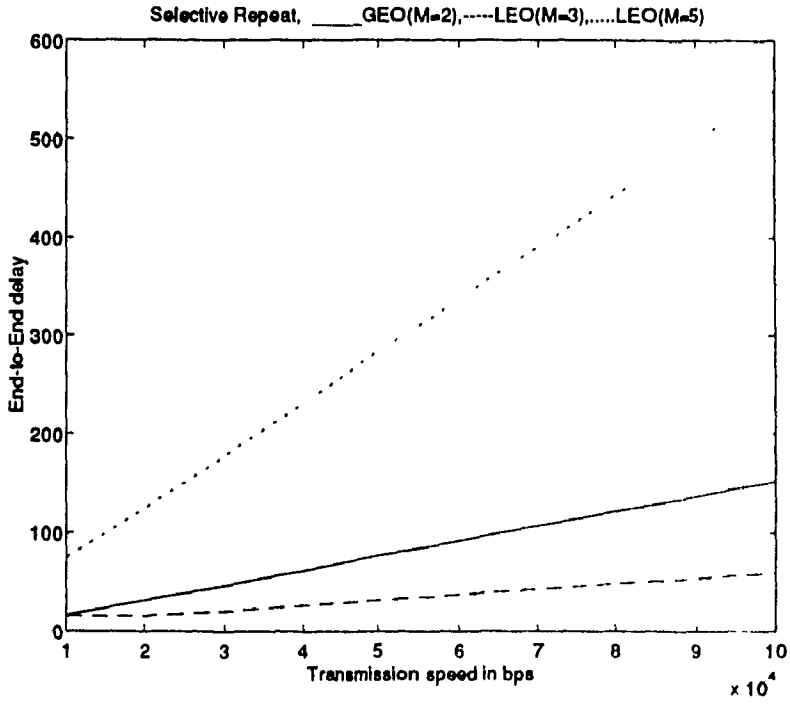

Fig. 2.16 (b) SR scheme, P_b=10^{-3}, ρ=0.5;

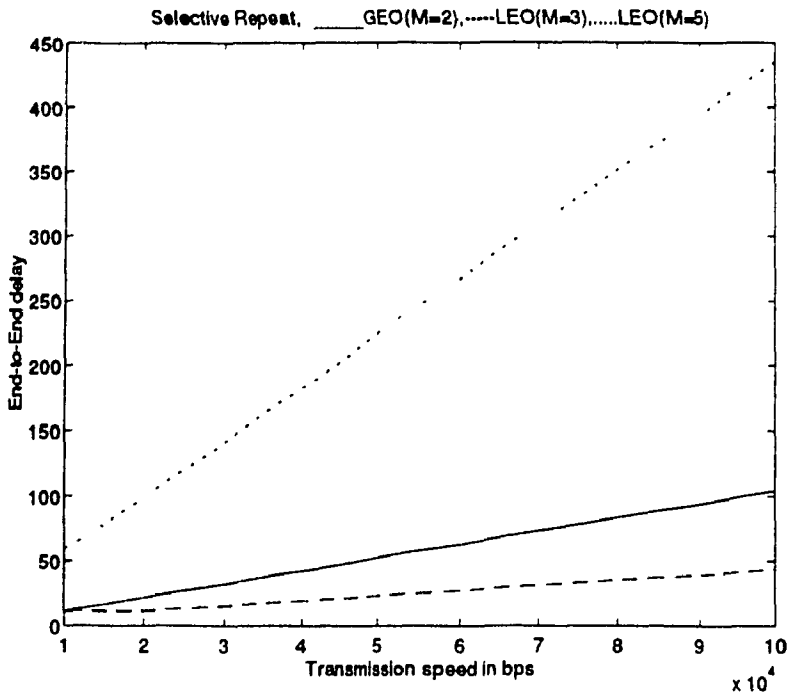

Fig. 2.16 (c) SR scheme, P_b=10^{-3}, ρ=0.1;

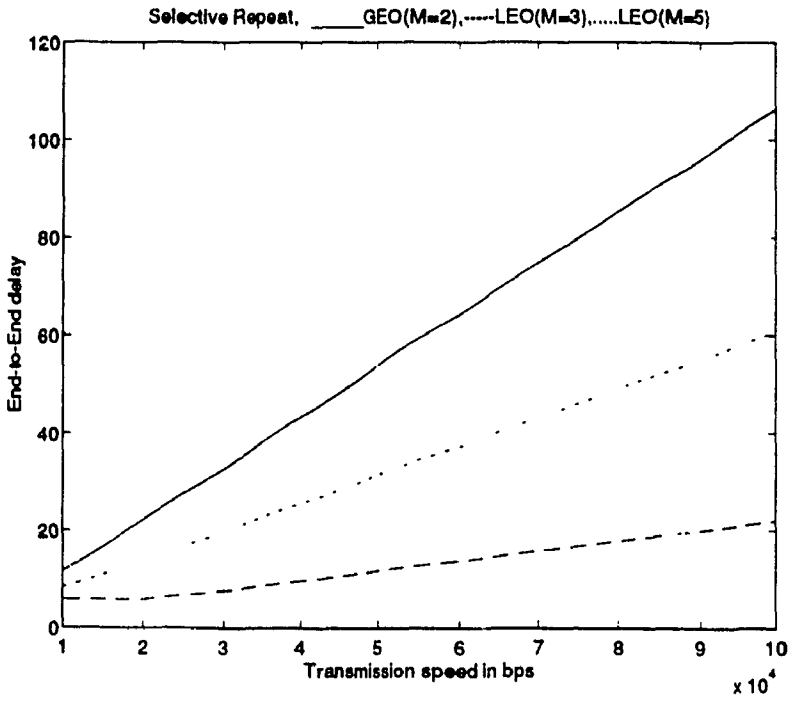

Fig. 2.16 (d) SR scheme, P_b=10^{-6}, ρ=0.99;

Fig. 2.16. SR delay vs Transmission speed with N=4, k=568 bits, n_h=56 bits.

The end-to-end delay characteristics were obtained as a function of VC transmission speed (in bits) for SR ARQ scheme in Fig. 2.16. These results were obtained for a packet data size of 64 bytes, packet header size of 7 bytes, and a window size of 4. The results were obtained at a probability of bit error (P_b) of 10^{-3} and different traffic intensities, i.e., 0.99, 0.5 and 0.1 in Figs 2.16 (a), (b), and (c), respectively. The results in Fig. 2.16 (d) were obtained at a probability of bit error (P_b) of 10^{-6} and a traffic intensity of 0.99 for comparison. As detailed earlier, the propagation delay plays a major drawback for higher transmission speeds in ARQ window schemes. As $N < 1+2a$ condition is quickly reached in Fig. 2.16, the increase in transmission speeds increased end-to-end delay. The end to end delay can be decreased by choosing a higher window size for better channel utilization. The 2 LEO satellite (M=3) configuration provided better (i.e., lesser) end-to-end delay in Fig. 2.16, compared to 1 GEO satellite (M=2), and 4 LEO satellite (M=5) configurations.

2.5 Conclusions

The end-to-end throughput and delay characteristics were analyzed for GBN and SR ARQ transport users using Norton equivalent queuing model, and the ARQ protocol efficiency formulas. The analysis was applied to GEO and LEO satellite links to identify optimum end-to-end ARQ flow control parameters, i.e., packet and window sizes, and bandwidth in bits per second. The results were computed for a set of window size (N) values to identify better ARQ window size range over LEO satellite links for optimum

throughput and end-to-end delay. Once an optimum window size is identified, then the end-to-end throughput and delay performance are computed as a function of packet size. Using an optimum set of ARQ packet and window sizes, the throughput and delay characteristics were obtained as a function of the bandwidth values (in bps) of the end-to-end VC.

It was observed from Figs 2.5 and 2.6 that the GBN ARQ scheme with 2 LEO satellite (M=3) VC provided better throughput and low delay. When P_b is as low as 10^{-6}, both 2 LEO (M=3) and 4 LEO satellite (M=5) VCs performed better than 1 GEO satellite (M=2) VC for fairly larger window size. Using an optimum window size of 4, the end-to-end throughput and delay characteristics were obtained as a function of packet size in Figs 2.9-2.12. When P_b is as low as 10^{-6}, these figures also exhibited better results for 2 LEO satellite (M=3) VC than the other two configurations, i.e., 1 GEO satellite (M=2), and 4 LEO satellite (M=5) VCs for fairly a reasonable range of packet sizes. Moreover, the end-to-end throughput and delay characteristics were obtained as a function of VC bandwidth at a packet data size of 512 bits, packet header size of 56 bits, and a window size of 4 in Figs 2.13-2.16. The throughput and delay performance over 2 LEO satellite (M=3) VC was observed to be better than 1 GEO satellite (M=2) and 4 LEO satellite (M=5) VCs. So, even though there may exist a constellation of LEO satellites, but the throughput and delay results indicate that it is not recommended to hop several LEO satellites to cover wide areas. Hence, the GEO satellites may still provide better throughput and delay results to cover wide areas.

Hence, the results clearly indicated that the 2 LEO satellite (M=3) VC is a better choice for both GBN and SR ARQ schemes. Moreover, the transmission path loss for LEO satellites, is obviously lower than that of GEO satellites for same conditions, that may result even better throughput characteristics for the LEO satellite environment. Hence, the LEO satellite communication may be better for multimedia traffic with low packet sizes. Furthermore, the results conform to the theory that SR ARQ provides better throughput than GBN ARQ when probability of bit error is as high as 10^{-1}.

CHAPTER 3

ATM NETWORKS

3.1 Introduction

The broadband integrated services digital network (BISDN) will provide cost-effective integrated access for diverse customers using a wide variety of applications [MIN-89], [AH2-89], [MIY-91]. It serves as a common carrier for interactive and distributive services, broadband and narrowband rates, bursty and continuous traffic, connection-oriented and connection-less services, digital signal processing, point-to-point and complex multimedia services. Moreover, the future BISDN models will support new characteristics like bandwidth on-demand, burstiness and quality of service performance [SYK-91], [MIN-89], [I121-88]. Asynchronous transfer mode (ATM) is the target transport method for implementing BISDN. As the ATM networks must operate at higher link speeds ranging from several hundreds of megabits or gigabits per second, the design of congestion control becomes an important challenge [ECK-91]. This chapter provides an introduction to ATM, modeling of traffic sources, and congestion strategies for ATM networks.

3.2 Asynchronous transfer mode

The notion of transfer mode is fairly a close combination of multiplexing and switching mechanisms. It shows that only packet-mode provides true independence between application and network, and offers solutions to applications that require variable

bit-rates. The packet transfer mode is not a recent discovery, but *"what is new"* is its usage as a multi-service mode [COU-91]. Traditional mode of information transfer is synchronous transfer mode (STM). In STM, information is divided into frames and transmitted in-sync with the receiver. STM is good for traffic in stream mode at peak rate. However, STM is not well suited for diverse applications that have different QOS requirements. ATM is a packet oriented switching and multiplexing technique based on fixed size packets, which are referred as *cells*. It provides efficient utilization of the bandwidth, to support diverse traffic types with different QOS requirements [SYK-94], [MIN-89]. It confines to the physical layer (layer 1) and basic functions of data-link layer (layer 2) of the ISO OSI reference model [DAY-83], [ZIM-80]. By using small and fixed-size cells, ATM is so efficient that it extends multi-rate circuit switching to allow multiple channels with the data rate of each channel dynamically set on demand [STA-94]. In ATM networks, the transmission medium provides transport service to the virtual path (VP), and the VP provides transport service to virtual circuit (VC). From the protocol point of view, the function supported by the transmission medium is described in the physical layer. The functions supported by the VP and VC are described in the ATM layer [SAI-94]. ATM allows B-ISDN to have logical separation of signaling from user information streams, which allows simultaneous establishment and release of multiple connections belonging to a call [KAW-91].

ATM networks are expected to provide information transport for a rich mixture of services and applications. These services may consist of a broad spectrum of traffic types

and transport performance needs. ATM provides statistical sharing of network resources such as bandwidth, buffers, processing, etc., by multiple users. The statistical multiplexing gain is caused by the aggregation of the cell streams (cell arrival processes) in the resources (cell level or call level) shared by the aggregated stream [SAI-94]. B-ISDN has received increased attention in the past few years due to the increased demand for multimedia communication services, e.g., data, voice and video. Some traffic, such as interactive data and video are highly bursty, while others such as large files, is continuous.

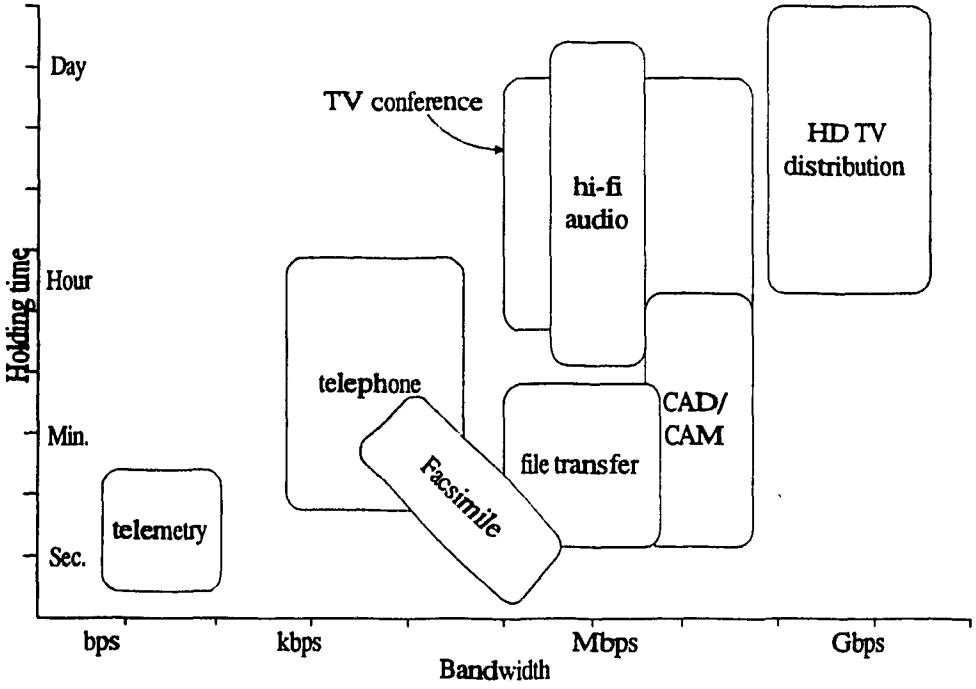

Fig. 3.1. Traffic characteristics of various applications.

Real-time voice for instance, requires rapid transfer through the network, but loss of small amounts of voice information is tolerable. In many data applications, real-time delivery is not of primary importance, but high throughput and strict error control are required. Some services, such as real-time video communications, require error-free transmission as well

as rapid transfer [BAE-91]. The diverse traffic characteristics of various applications are shown in Fig. 3.1. By utilizing virtual circuits and virtual paths, ATM inherits the advantages of both the circuit and packet switching modes [SAI-94].

3.3 ATM Protocol reference model

The ATM protocol reference model consists of a user plane (U-plane), a control plane (C-plane), and a management plane (M-plane) as shown in Fig. 3.2 [I321-91]. The U-plane provides user information transfer, and the C-plane performs connection control functions with signaling (for example, connection setup and release). The M-plane activities include establishment of user VPs and OAM related functions.

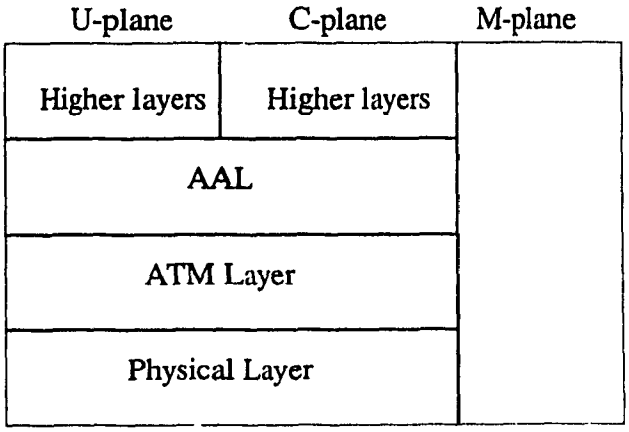

Fig. 3.2. ATM protocol reference model.

The ATM layer carries cells of user and signaling information between a user and a switching node or between switching nodes (the signaling and user information is distinguished from VPI/VCI values). The ATM adaptation layer (AAL) provides service-dependent functions (such as cell-error or cell-loss recovery functions, and cell

segmentation and reassembly) to the layers above the AAL. The ATM layer, which is below the AAL, offers service-independent function of cell-transfer. Four major protocols in the AAL, are provided for the four service classes A, B, C, and D. The AAL types 1, 2, 3, and 4 correspond to class A, B, C, and D, respectively. In addition to these protocol types, AAL protocol type 5, which is a simplified protocol for data communication, may also be used. In general, AAL type 1 is used for constant bit rate (CBR) services such as interactive voice. AAL type 2 is used for variable bit rate (VBR) services such as video [KAD-95]. The connection-oriented data is supported by the type 3 protocol. The connection-less service is supported by the type 4 protocol. The main functions of AAL type 1 are cell segmentation and reassembly for the user information, absorption of cell delay variation, disposal for loss and mis-insertion of cells. The AAL type 2 protocol has functions similar to those of the type 1 protocol, but it segments and reassembles variable-length packets into cells. The AAL type 3 and 4 also have similar functions.

3.4 ATM Traffic models

Source characterization is necessary for the precise definition of the behavior of a traffic source. In general, traffic corresponds to one of the three broad categories, i.e., voice, video, and data. An ATM traffic source can be represented as a high speed cell generator. In ATM networks, there is a general trend to visualize cell generation as a succession of active and silent (idle) periods. Cell generation occurs only during active periods. A group of successive cells that are not interrupted by idle period is called *burst*.

The most prominent paradigm of a source model exhibiting this behavior is the popular *on-off* model [ST3-94], which is discussed in detail later in this section.

SERVICE CLASS	PEAK BIT RATE	BURSTINESS	PEAK DURATION
Telephony	64 Kbps	1	100 s
Text			8 s
Fax			20 s
File transfer			2 s
Document retrieval	64 Kbps	200	250 ms
Data on demand			40 ms
Color fax	2 Mbps	1	3s
File Transfer			1s
Video telephony	10 Mbps	5	1 s
Video retrieval			1 s

Table I - ATM Services characterization.

Typical ATM services defined in RACE projects [GAL-89] are summarized in Table I. In this table, *burstiness* is measured as the peak-to-mean bit rate ratio. A high level of burstiness is assumed in these new broadband services, to take into account the behavior of traffic sources such as video telephony, video retrieval, document retrieval, and data on

demand [GAL-89]. A set of three important parameters that are used for the characterization of ATM sources is as follows:

- Peak bit rate, i.e., the maximum transmission bit rate of a source,
- Mean bit rate, i.e., the mean transmission bit rate of a source averaged over the duration of a call, and
- Mean peak duration, i.e., the mean duration of the time-interval during which the traffic source transmits at the peak rate.

ATM networks offer a specific set of service classes with guaranteed QOS values. The user must request a specific service class, specifying the expected nature of traffic and the required QOS during a connection. Service classes are used by ATM networks to differentiate the types of connections, each with a particular mix of traffic and QOS parameters [ALL-95]. The current QOS classes as per ATM Forum UNI 4.0, are (i) constant bit rate (CBR), (ii) variable bit rate (VBR), (iii) available bit rate (ABR), and (iv) unspecified bit rate (UBR).

In general, the CBR is used for *stream traffic*, whereas the VBR is used for *bursty traffic*. The ATM Forum is currently focusing on ABR service. The primary goal of ABR service is the economic support of applications with vague requirements for throughput and delays, which are best expressed as ranges of acceptable values [BON-95]. Hence, the ABR is expected to support variable rate data transmissions without preserving any timing relationships between source and destination. Unlike the VBR service, the ABR service

does not provide any guaranteed bandwidth to the user. Rather, the network provides a feedback flow control to increase available-bandwidth to the user. Hence, the ABR provides the *best effort service* to its users. The UBR service does not offer any service guarantees. The user is free to send any amount of data up to a specified maximum while the network makes no guarantees on the cell-loss rate, delay, or delay variation that might be experienced. The UBR service is currently the best match for LAN protocols, given that the ABR specification is not yet completed [ALL-95]. The voice, video, data, and superimposed traffic models are detailed in the following sections.

Voice sources:

Voice services have several different QOS requirements. The voice traffic is very sensitive to delay, whereas it can tolerate moderate cell loss. The simplest source model for packetized voice is CBR model, which generates cells constantly at the peak rate during *talk* period. However, a substantial part of voice conversation involves *no talk* or silence. The successive talkspurts and silence periods from a voice source can be characterized by an alternating renewal process. All talkspurt and silence time intervals are independent of each other [ST3-94], [SRI-86] as shown in Fig. 3.3. In ATM networks, CBR model is used for 64 Kbps pulse code modulation (PCM), or 32 Kbps adaptive differential PCM (ADPCM) using AAL type 1 protocol. The other method is to use a SAD, and a digital speech interpolation technique, using VBR model with AAL type 2 protocol. The performance of voice cell multiplexing requires a new model for the

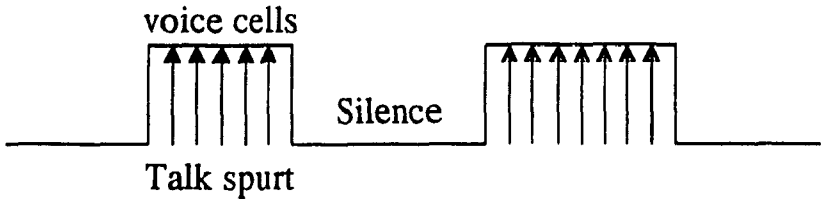

Fig. 3.3. Voice source behavior.

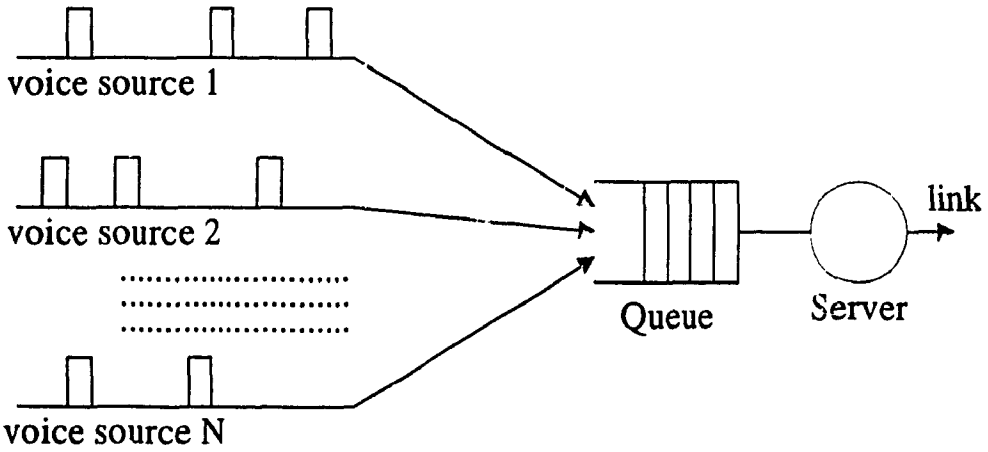

Fig. 3.4. A statistically packetized voice multiplexer.

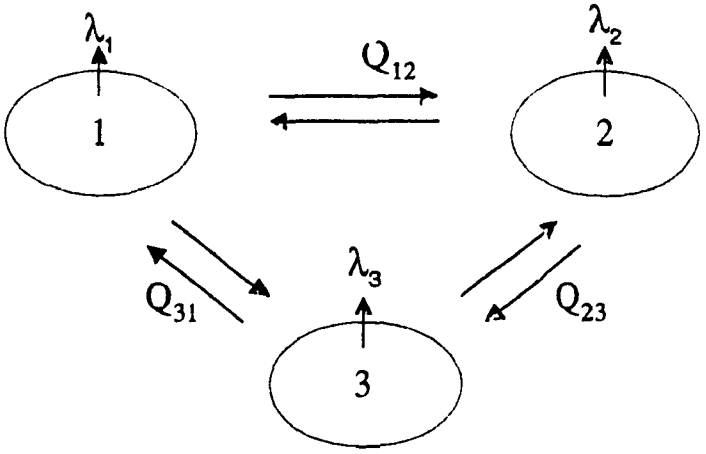

Fig. 3.5. A three state MMPP.

superposed voice arrival processes as shown in Fig. 3.4. The aggregated arrival process is highly correlated. So, a simple M/D/1 model does not provide accurate results [SRI-86]. Moreover, it is not acceptable when the number of voice sources is large [SAI-91], or traffic is light, or number of waiting rooms is small [SRI-86]. A common approach for modeling aggregate arrivals from several voice sources is to use MMPP [AKI-94]. Consider a continuous-time Markov process with state space $\{1,2,....m\}$, and let the infinite generator of this process be Q, an (m,m) matrix with $Q_{ii} < 0$ and $Q_{ij} \geq 0$ for $i \neq j$ such that Q^{-1} exists. During any state i, where $1 \leq i \leq m$, there are Poisson process arrivals of rate λ_i, and the resulting arrival rate matrix is $A=\text{diag}(\lambda_1, \lambda_2, ... ,\lambda_m)$. This is called m-state MMPP as shown in Fig. 3.5 (where m=3). When the state of Markov chain is $\{1,2\}$, the MMPP is called Switched Poisson Process (SPP). In particular, an SPP in which one of the arrival rates λ_1 or λ_2 is 0, then it is called an Interrupted Poisson Process (IPP) [SAI-94]. Let m, c, and z be the mean, coefficient of variation, peak cell arrival, respectively. Let T, α^{-1}, and β^{-1} be cell arrival process, mean silence period, and mean talkspurt period, respectively. The assumption of exponential talkspurt and silence distributions lead to the following:

$$m = (1+\beta/\alpha)T \tag{3.1}$$

$$c = \frac{\sqrt{2\beta T - \beta^2 T^2}}{(\alpha+\beta)T} \tag{3.2}$$

$$z = \left[1 - \left(1 - \beta T + \frac{\alpha\beta T}{m^{-1} + \alpha}\right) e^{-T/m}\right]^{-1} - \frac{m\alpha}{(\alpha + \beta)T} \tag{3.3}$$

Video sources:

The statistics of the traffic generated by a video source are greatly influenced by the nature of the pictures transmitted, the relation among them, the quality of service provided, and the coding technique adopted. Efficient network utilization and constant picture quality can be achieved by VBR coding of video frames [SEN-89], [YEG-93], [MAR-94], [HAB-92]. However, CBR can also be used for video traffic with AAL type 1. A few recent studies have reported on cell arrival processes of cells with a format and length different from the CCITT standard [GRU-91]. These studies are based on the cell-level statistics. An auto-regressive moving average (ARMA) model linked with a zero-memory non-linearity (ZMNL) transform is shown in Fig. 3.6. This combination is used to model the process of the number of cells arriving during 64 time slots, where a single time slot denotes the cell transmission time by a 150 Mbps VP. Because the ARMA can take a negative value, the ZMNL is introduced to keep it from yielding a negative value for the

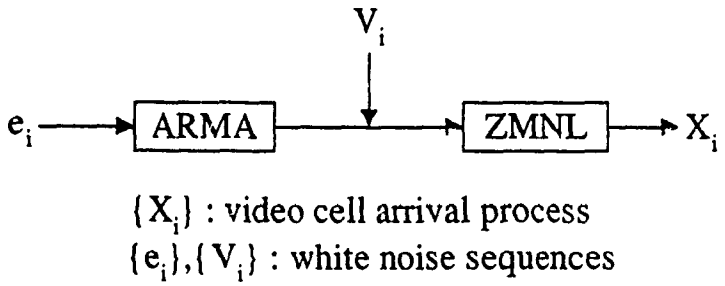

Fig. 3.6. ARMA-ZMNL model.

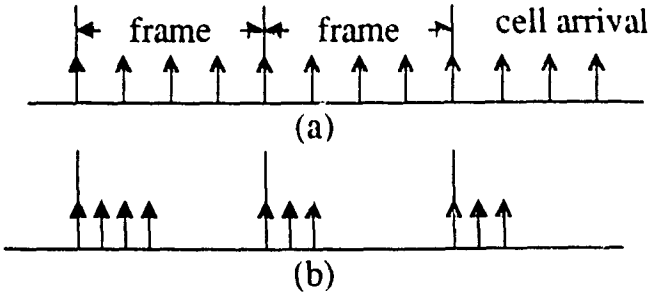

Fig. 3.7. Video cell arrival processes:

(a) *uniform* model; (b) *on-off* model.

number of cells arriving. If the traffic is sufficiently large, the effect of the ZMNL is negligible and the proposed model is therefore identical to the ARMA. The ARMA model is promising for simulation because it is applicable to the number of cells generated during short interval as well as in a frame. The ARMA model is not convenient for queueing analysis. Hence, two other models (Fig. 3.7), i.e., uniform model and on-off model, are generally used for queueing analysis of video sources. In the uniform model, cell arrival times are uniformly distributed with a fixed intensity within a frame. In the on-off model, cell arrival times are either fixed or exponentially distributed with a fixed intensity during *on* period, whereas no cells exist during *off* period. Studies by Maglaris et al., in [MAG-88] and [SEN-89], used the examples of uniform model. In [MAG-88], the number of cells generated within a frame is modeled by the first order auto-regressive model and by the discrete finite level Markov model. The latter is used to evaluate the statistical multiplexing of the video calls. This Markov model assumes that the input level $\lambda(t)$, i.e., the number of cells arriving from N sources per unit of time, takes one of $M+1$ values: 0,

A, ..., MA. It also assumes that the input level transition is given by the following infinitesimal generator Q:

$$Q = \begin{pmatrix} -M\beta & M\beta & & \\ \alpha - [\alpha + (M-1)\beta] & (M-1)\beta & & \\ \cdots\cdots & & & \\ \cdots\cdots & & & \\ & & (M-1)\alpha - [(M-1)\alpha + \beta]\beta & \\ & & M\alpha & -M\alpha \end{pmatrix} \tag{3.4}$$

The parameters A, α, and β are obtained by matching measured input levels to the average, the variance, and the exponential tail of the auto-covariance of the input level of the model. Sen [SEN-89] extended this model that could describe the scene change.

The *on-off* model is presented in [SAI-91], in which one frame is divided into k-phases, and the sojourn time in each phase is exponentially distributed. Hence, the frame length is modeled as the k^{th} order Erlang distribution. During the first p phases, cells arrive in a Poisson process at λ_v (the peak cell rate of the source). During the rest (k-p) phases, no cell arrives. Here p is assumed to be constant to simplify the model. However, p can actually be a random variable. Consequently, the video cell arrival process from a single video source can be modeled by an MMPP. Therefore the infinitesimal generator of the phase process of the MMPP describing an individual process of a video source is

$$Q = \begin{pmatrix} -q & q & 0 & 0 & \dots & 0 \\ 0 & -q & q & 0 & \dots & 0 \\ 0 & & \dots\dots\dots & & & 0 \\ q & 0 & 0 & 0 & \dots & -q \end{pmatrix} \qquad (3.5)$$

and the arrival rate vector describing the Poisson arrival rate in each phase of MMPP is

$$A = diag(\lambda_v, \dots\dots \lambda_v, 0, \dots, 0) \quad . \qquad (3.6)$$

Data sources:

In general, data applications such as image and file transfer generate traffic-flows that move from a burst state, offering a continuous stream of cells at some peak rate, to an idle state [SAA-94]. These applications are usually very sensitive to cell-loss, while their respective delay requirements may not be so important. The generation of data from a single data source is well characterized by a Poisson process for continuous-time case, or by a geometric inter-arrival process for discrete-time case [BAE-91]. However, actual data packet arrival process in ATM networks has not been identified [SAI-94]. Hence the properties of cell arrival processes have been derived from the statistics of data packets. One characteristic of data packet statistics is that the packet length distribution is bimodal [GUS-90]. A data packet can be treated as a burst of cells. The AAL generates cells from a data packet by treating it as an SDU. Since the AAL is fast, the bimodal distribution of packet length suggests that the burst of data cells can be bimodally distributed. The cell arrival process within a burst depends on the implementation of a higher level protocol (such as AAL). Hence it can not be identified easily. However, the data cell arrival process is likely to be a Switched Poisson Process (which includes the Interrupted Poisson

Process), and the geometrically modulated deterministic process (GMDP) [SAI-94]. Recent work on modeling data traffic is to use a two-state doubly stochastic Poisson process with Pareto packet interarrivals [SUB-95], [PAX-95].

Traffic modeling in ATM networks:

The VBR burst-silence source model (Fig. 3.3) is generally used for modeling traffic in ATM networks. It is used not only for packetized voice models, but also for still video, and data messaging applications. It allows relevant parameters, namely maximum cell rate, mean cell rate, and mean peak rate duration, to be varied independently of each other. Cells are generated periodically during a burst. In the discrete time case, the number of cells per burst is assumed to have a geometric distribution with mean *E[X]*. The duration of the silence phases is assumed to be distributed according to a negative exponential distribution with mean *E[S]*, and the inter-cell time duration is given by δ [RAT-91] with

$$\begin{aligned} \beta^{-1} &= E[X]\,\delta \\ \alpha^{-1} &= E[S] \end{aligned} \tag{3.7}$$

where α^{-1} and β^{-1} are the mean silence (*off*) and burst (*on*) periods respectively. The continuous time analog is an exponential distribution using Poisson process. This model for video sources does not take into consideration the inter-frame correlation between successive frames of the video source. This model generates cells at the peak rate of the video codec during each frame. It does not generate any cells during the rest of the frame. When *N* independent traffic sources with the same parameters are multiplexed, aggregate cell arrivals are governed by the number of sources in the burst or *on* state. Assuming a discrete time system, the probability that *n* out of *N* sources are in the *on* state is given by

$$P_n = \binom{N}{n} \left(\frac{\beta}{\alpha+\beta}\right)^n \left(\frac{\alpha}{\alpha+\beta}\right)^{N-n} \tag{3.8}$$

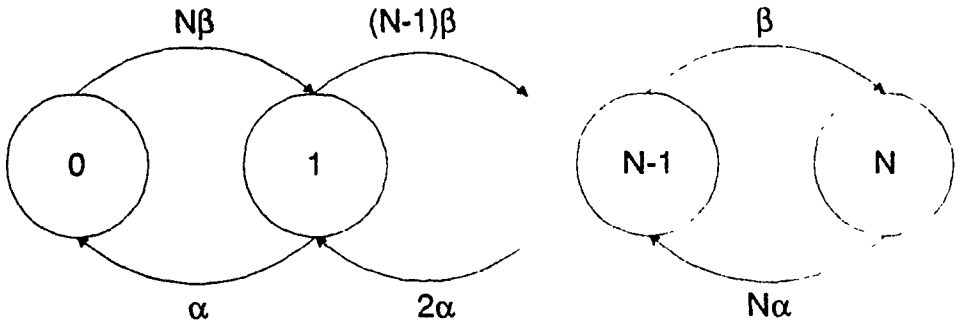

Fig. 3.8. Birth-death model for N sources.

The continuous time analog represents the number of sources in *on* state as the birth-death process (Fig. 3.8) with birth rate $\lambda(n)$ and death rate $\mu(n)$, where

$$\begin{aligned} \lambda(n) &= (N-n)\beta \\ \mu(n) &= n\alpha \qquad \textit{where } 0 \le n \le N \end{aligned} \tag{3.9}$$

For this continuous time case, the probability P_n that n out of N sources are in *on* state, is also given by equation (3.8).

A main characteristic of ATM networks is that the heterogeneous traffic is superimposed. It is therefore important to superpose arrival process models easily when evaluating the performance of ATM networks. To model the superposed traffic, two methods can be considered: statistics-superposition and model-superposition. The statistics superposition evaluates the statistics of the individual processes and then uses them to compute the statistics of the superposed process. Finally, it determines the model

that fits the statistics of the superposed process. The advantages of statistics-superposition are that it does not require the individual process model. Its computation is easy, and the state space does not grow (unlike the state space in model superposition). The model superposition, on the other hand, first determines models for the individual processes and then superposes them. This method is applicable only when the superposed model can be obtained directly from the individual models such as MMPPs, Markovian arrival processes (MAP), or a discrete-time MAP.

The video and voice sources with SAD, generate VBR traffic. They can share a VP and a finite buffer capacity. In other words, superimposed voice and video traffic can be offered in a single server queue with a finite number of waiting rooms. The aggregated voice cell arrival is modeled by a two-state MMPP. The video cell arrival is also modeled by MMPP. The superposed video model is also MMPP. When the voice model and video model are both MMPPs, the superposed voice and video traffic model is an MMPP [SAI-94].

ATM traffic with TDMA service:

As the integrated traffic is digitized, the high bandwidth transmission capacity of the state of art technology can be shared by the integrated traffic (i.e., voice, video, and data) by means of fixed composite time-division multiplexed frames. The channel can be synchronously clocked and divided into frames of fixed duration. Each frame can be further subdivided into slots. A slot duration is designed to hold an ATM cell. The TDM

frame duration is determined by digitization rate and the transmission rate. The TDM frame is further partitioned into several regions based on dynamic sharing of the channel capacity for the classes of ATM traffic. Each partition of the TDM frame is designed to hold cells of a class of ATM traffic. All cells, which can not be transmitted are buffered and serviced on a first-come-first-served basis. The talkspurt of voice, or TPDU of data, or video picture can be treated as VBR generator, which generates bursts of cells. The specific characteristics of a burst of traffic are obtained from the particular class of traffic, which determine arrival and service characteristics of the cells. The steady-state probability p_l that l number of cells arrive during a TDM frame, is found [WID-95], [SRI-83], [TSE-91] to be

$$p_l = \frac{\rho_1^l / l!}{\sum_{i=0}^{nn} \rho_1^i / i!} \qquad where \;\; l = 0.1,2,\ldots,nn \tag{3.10}$$

and where $\rho_1 = \lambda_1 / \mu_1$, and $1/\lambda_1$, $1/\mu_1$ are interarrival-time and service-time of frames respectively, and nn is the total number of slots. The distribution of equation (3.10) accounts for the possible variability of a source bit rate. The simplicity and flexibility of equation (3.10) are further improved in chapter 4 for characterizing the ATM traffic with TDMA service, while analyzing congestion behavior of the ATM multiplexer node in a hybrid ATM/TDMA network.

3.5 Congestion in ATM networks

As ATM networks must operate at higher link speeds ranging from several hundreds of megabits or gigabits per second, congestion is extremely difficult problem to

cope-up with the large propagation delays. The existing congestion control schemes may not be applicable to such high-speed networks. In order to start a new era of multimedia communications, CCITT has recommended two speeds for B-ISDN access on the first issue, one at approximately 155.25 Mbps and the second at approximately 622.08 Mbps. At these speeds, cells, which are short fixed-length blocks, must be switched at a rate greater than one cell per 3 μs or 0.7 μs challenging the cell processing speeds [HO2-91]. Hence, the traditional error-recovery and flow-control between neighboring switching nodes, i.e., link-by-link control, are not introduced in ATM. The ATM layer is considered to provide a physical transmission line, which guarantees that the cell loss probability will be less than certain value between end-to-end nodes. Hence, the existing congestion control schemes shall be redesigned with end-to-end control, or new schemes shall be discovered for ATM networks [MIY-91].

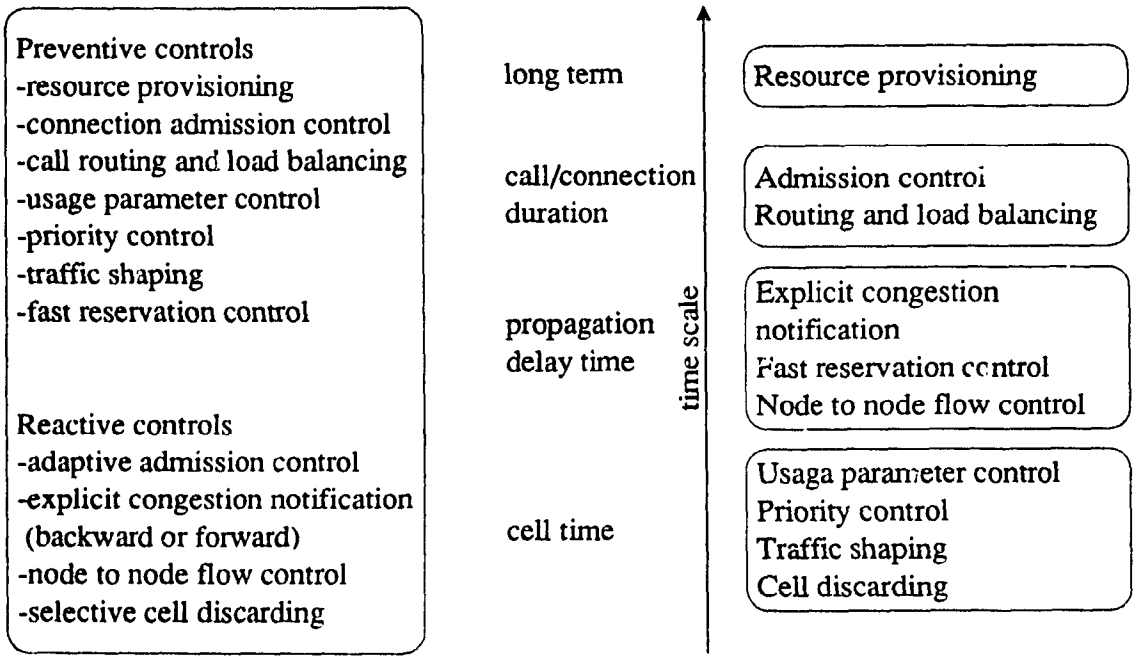

Fig. 3.9. Classification of ATM congestion control schemes.

Congestion control schemes must ensure sufficient bandwidth, satisfactory performance and quality of service to all users of the ATM node. As ATM networks are sensitive to processing speeds, the four major goals of a congestion control architecture are *simplicity* for low processing delays, *robustness* for relative insensitivity of a control architecture to changes in assumptions, *flexibility* for adapting to new situations, and *controllability* for efficient network resource utilization without paying penalty in performance.

The congestion control schemes have been classified into two major categories: preventive and reactive controls as shown in Fig. 3.9. Preventive controls try to prevent the ATM network from reaching an unacceptable level of congestion by controlling the amount of traffic flow at the entry point to the network (i.e., at the access nodes). The resource provisioning ensures sufficient resources to accept all potential virtual connections (VCs), while maintaining a cost-effective network design. The connection admission control (CAC) procedure determines the maximum set of virtual connections that can be admitted based on QOS parameters, while ensuring an acceptable level of performance. Furthermore, cells can be prioritized based on the diverse characteristics of traffic streams. Another preventive scheme called "fast reservation control" allows to negotiate bandwidth at burst level.

In spite of several preventive measures for congestion, network may still be congested due to violation of QOS of the users, which forces loss of cells due to

momentary buffer overflow and sustained overload. The momentary *buffer overflow* or *cell loss* may happen when several users transmit at their peak bandwidth bursts at the same time, exceeding the capacity of the buffers or link, respectively. Reactive congestion control schemes work to prevent momentary buffer overflows turning into sustained congestion events due to the additional traffic caused by end-to-end retransmissions. Explicit congestion notification (ECN) is a reactive control by which the end users are kept informed about the congestion status of the ATM network to avoid generation of new traffic during congestion. In response to the network congestion indicators, the higher layer protocols are expected to start shaping the volume of traffic admitted to the network. This scheme is highly beneficial when the duration of congestion is at least an order of magnitude larger than the propagation delay. Another reactive scheme called "node to node flow control using shutter algorithm" is a rate based node-to-node flow control mechanism that dynamically regulates the flow of traffic between the adjacent nodes of an ATM network.

The reactive control schemes are not suitable for ATM networks due to the fact that propagation delay is longer than switching time. Hence it will be too late to react to the congestion since many cells will already be transmitted before realizing congestion on the adjacent nodes. Another conservative approach to prevent congestion at runtime is *admission control* or *policing*. As ATM network is expected to support large number of bursty sources, network traffic control is essential to provide a desirable level of network performance to the users. Admission control and traffic smoothing are the most promising

control techniques in ATM networks [MUR-90]. The principle of admission control is that any new user or traffic is accepted only when it does not create congestion. Admission control at user level is often called as *call-level admission control*. The admission control decision also depends on the declared traffic descriptor values of the requested connection. The time required to perform admission control is critical in a high-speed environment [ROB-91].

In addition to admission policing, it is necessary to implement some form of cell discard schemes to control congestion in the network. Some of the popular cell discard control schemes are leaky bucket (LB) [JAD-94] and virtual leaky bucket (VLB) [BER-94]. A leaky bucket corresponds to a counter which is incremented each time a cell is generated by the source and is decremented periodically with a suitable leaky rate. A cell, which arrives after the counter has reached a given threshold value, is dropped [BHU-91]. In other words, the cells which violate the allowed rate of arrival are dropped. Hence, a leaky bucket forces a traffic source not to exceed some specified average rate of traffic input and some specified *burst length* (maximum sequence of consecutive cells at the peak input rate) [LEL-89]. An improvement over LB discard scheme is the virtual leaky bucket (VLB) scheme. The VLB bandwidth enforcement scheme overcomes the limitations of LB mechanism by merging policing and congestion control problems [GAL-89]. The VLB scheme "marks" the violating cells for future probable dropping when the overall congestion level reaches a certain threshold over the VC. The motivation behind this scheme is to improve throughput by minimizing cell-discard at the input access node

which is inevitable in the leaky bucket scheme. In this thesis, the LB and VLB schemes are modified to take into account, the probability of violation and ATM multiplexer node congestion, coupled with burst length modulation. The modified LB (mLB) scheme admits traffic from an ATM user as long as the QOS is respected (i.e., cells do not violate QOS). However, the modified VLB (mVLB) scheme admits all cells, including the violating cells, as long as ATM node is not congested. Several other congestion control schemes, i.e., jumping window (JW), triggered jumping window (TJW), exponentially weighted moving average mechanism (EWMA), moving window (MW), etc., are described in [RAT-91].

In spite of several schemes, congestion control is still a challenge in ATM networks due to high transmission speeds and the diverse bandwidth requirements for various multimedia services. It will be quite challenging to use the existing TDMA networks to interface ATM users in the form of hybrid ATM/TDMA networks. The following chapters of this thesis investigate new congestion control schemes for hybrid ATM/TDMA interconnected networks.

CHAPTER 4

LOCAL CONGESTION CONTROL

4.1 Introduction

Asynchronous transfer mode (ATM), due to its ability to handle multimedia services at high speeds, is the target transport method for broadband networks, using virtual connection concept [RAT-91]. As the ATM networks must operate at higher link speeds ranging from several hundreds of megabits or gigabits per second, cells, which are fixed-length blocks, must be switched at a rate faster than a micro second. At these link speeds of ATM networks, the design of congestion control becomes an important challenge [ECK-91]. A promising technique to prevent congestion in ATM networks, is *admission control* [MUR-90]. The principle of admission control is that any new user (call) is accepted only if it does not create congestion. Besides admission control, it is necessary to implement some form of cell discard schemes to control congestion in the network. Some of the popular cell discard schemes are leaky bucket (LB) and virtual leaky bucket (VLB) [GAL-89]. The LB corresponds to a *counter*, which is incremented each time a cell is received and is decremented periodically with a suitable leaky rate. The LB scheme drops all arriving cells when the *counter* is at a pre-defined threshold (N) [BHU-91]. The VLB is an improved version of LB, in which all arriving cells when the *counter* is at *N*, are *marked* for dropping only if the ATM multiplexer node gets congested [GAL-89]. In spite of several schemes mentioned in chapter 3, congestion control is still a

challenge in ATM networks due to high transmission speeds and the diverse bandwidth requirements for various multimedia services.

The objective of this chapter is to propose a new congestion control scheme in interconnected broadband networks such as hybrid asynchronous transfer mode/time division multiple access (ATM/TDMA) networks. The new congestion control scheme shall ensure necessary bandwidth for supporting various classes of service, i.e., video, voice, data, etc.. The ATM multiplexer node is assumed to provide TDMA service (bulk service of cells) to the generic ATM sources, by means of composite TDMA frames. The transport users are assumed to be generic ATM sources, which generate active bursts of traffic (bulk arrival of cells) during the connection. By communicating the ATM congestion criteria to the transport users by means of *service primitives*, the active bursts of traffic can be controlled based on the instantaneous congestion criteria, which is referred as *active burst control*. Moreover, when the transport user and ATM adaptation layers exist in distributed processing systems (probably using different processors), it is very appropriate to use layer interaction to control system resources (especially buffers). The length of an active burst can be controlled by modulating the higher layer protocol data unit (PDU) size in case of video and voice users. As PDU size may not be controlled for data users, the number of active bursts can be controlled by modulating the higher layer window size. By controlling the active bursts judiciously to preserve the quality of picture or speech, the number of cells that are generated during congestion periods, can be effectively controlled, thereby helping to reduce ATM multiplexer node congestion. This is

a significant improvement over window control mechanisms, which are presented in [RAT-91]. Furthermore, the higher layer window control reduces the possible end-to-end re-transmissions, which may otherwise be prompted by the cell loss during congestion, and thereby enhancing the end-to-end throughput for data traffic significantly. Hence, it is clearly a significant improvement over *fluid technique* of [DA1-92], which accepts or rejects an active burst. The analysis of LB scheme is based on discarding all traffic that violates the *quality of service* (QOS) of the service-class. The analysis of VLB scheme is based on admitting all traffic, including that violates the QOS of the service-class, as long as the ATM multiplexer node is not congested. The VLB scheme discards only the traffic that violates the QOS of the service-class during congestion periods. Hence, it leaves the scope for congestion if all ATM users simultaneously transmit stream traffic without violating their assigned bandwidths during the periods of congestion. We improved this limitation by applying *active burst control* on the top of LB and VLB schemes, to derive *modified* LB (mLB) and *modified* VLB (mVLB), respectively. The mLB and mVLB schemes communicate instantaneous congestion criteria to the transport users by means of *service primitives*, while performing the cell discard functions of LB and VLB schemes, respectively. Hence, the mLB and mVLB schemes shall control congestion more effectively than the respective LB and VLB schemes. Both mLB and mVLB schemes enhance the system performance by reducing the ATM multiplexer node congestion and the violation of QOS of the service-class. The admitted cells in all these four schemes, namely LB, VLB, mLB and mVLB, are added to the ATM multiplexer buffer. The ATM multiplexer node buffer is analyzed using a Markov modulated Poisson process (MMPP)

[SAI-94] queuing model with bulk arrival and bulk service of cells. Simple analytical formulas for the ATM multiplexer node congestion criteria, i.e., the mean probabilities of local congestion (P_c), generating a cell which gets admitted for transmission (ξ), discarding a cell (Δ), buffer content (Ψ), i.e., the number of admitted cells, buffer overflow ($P_{overflow}$), i.e., the number of admitted cells during peak buffer occupancy, which may be lost when there is no room in the queue, are also derived. The buffer content criterion is a measure of system delay to transmit an admitted cell. The probability of generating a cell that is admitted for transmission (ξ), is directly proportional to the throughput of the system. The probability of discarding a cell (Δ) is a measure of cell-loss due to violation of QOS. The buffer overflow ($P_{overflow}$) is also a measure of cell-loss due to lack of system buffers for storing the admitted cells. The ATM multiplexer-node congestion performance of mLB and mVLB schemes are compared with LB and VLB schemes, respectively.

4.2 Description of the model

The statistical behavior of ATM traffic can be classified into several levels, i.e., call-level, burst-level, and cell-level, as shown in Fig. 4.1. Since the ATM traffic is highly bursty, a two-phase burst and silence model, as shown in Fig. 4.2, can be applied to an ATM traffic source. The burst and silence periods are represented as active and idle states respectively. The transition probabilities *active-to-idle* and *idle-to-active* are α, and β, respectively. During an active burst of traffic from an ATM source, cells are generated at a constant rate with cell-time duration of δ as shown in Fig. 4.3. Since an ATM network is expected to support a large number of bursty sources with different types of traffic, a

Poisson process may no longer be able to describe the network traffic accurately [MUR-90]. The aggregated cell arrival from multiple active sources can be modeled by a Markov modulated Poisson process (MMPP) [SAI-94]. The statistical characteristics of traffic generated by all calls of a certain class, are assumed to be identical and independent of each other. A bursty source may generate cells at near peak rate for a short-period of time, and a moment later, it may become idle contributing no traffic to the network. Thus, an ATM network traffic can change rapidly, forcing the network to move from one degree of congestion to the other, even when the total number of calls is constant [KAM-89].

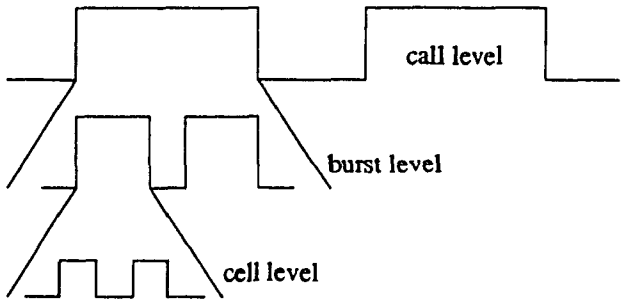

Fig. 4.1. Three levels of ATM traffic.

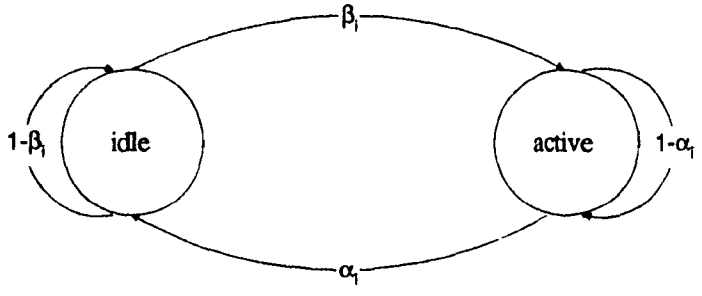

Fig. 4.2. Two-phase burst/silence model for ATM call.

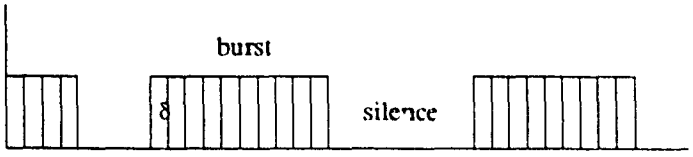

Fig. 4.3. Cell generation of a bursty call.

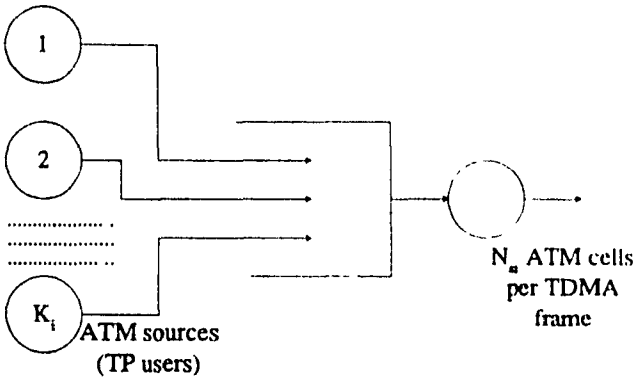

Fig. 4.4. Queueing model for ATM/TDMA.

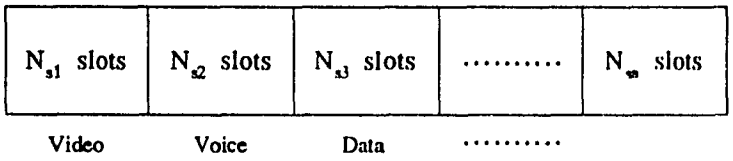

Fig. 4.5. ATM/TDMA composite frame structure.

The ATM multiplexer node is assumed to provide TDMA service to generic ATM sources by means of composite TDMA frames. The transport users are considered as generic ATM sources with diverse bandwidth requirements, as shown in Fig. 4.4. The transport users are classified into several classes of service, i.e., video, voice, data, etc.. A

composite TDMA frame is divided into several fixed subframes, each servicing a class of traffic, as shown in Fig. 4.5. The length of a subframe (N_{si} fixed time-slots) is sufficient to fully support an average number (K_{av}) of active transport users and the quality of service. A time-slot in a subframe is designed to transmit a cell. The virtual path identifier (VPI) and virtual circuit identifier (VCI) of a cell are assumed to uniquely address an ATM user. A transport user is allowed to send up to a maximum of n_{ui} cells, during a TDMA composite frame service-time. A service-class is assigned with a QOS value (q_i). An active transport user may not always generate traffic. As the total bandwidth of a service-class is shared by all active transport users, the ATM multiplexer node can support K_i (where $K_i > K_{av}$) number of *active* transport users of the service-class. As all transport users of a service-class may not be active simultaneously, the ATM multiplexer can support more than K_i number of transport users. It is assumed that an active transport user can generate only one active burst of traffic during a composite TDMA frame service-time. It limits the total number of active bursts generated by all K_i active transport users to k_i (where $0 \le k_i \le K_i$) during a composite frame service-time. However, an active burst of traffic is assumed to generate variable number of cells, depending on the allowed bandwidth of the user. The total number of active bursts (k_i) that are generated during a composite TDMA frame service-time vary rapidly. Hence, the ATM multiplexer buffer content (which is defined as the total number of cells that are waiting for transmission) varies accordingly. A bursty active user may generate traffic, which may momentarily exceed the assigned QOS (q_i), and thereby violating the bandwidth agreement. To handle this situation, the cell-level discard control is used to prevent congestion at runtime. To reduce the complexity of

analysis, a measurement window (N_{fmax}) is assumed at the ATM layer for computing the probability of violation (P_{vi}) in terms of n_{ai}, q_i, N_{si}, etc.. The measurement period (i.e., total time period to transmit N_{fmax} composite TDMA frames) is fairly high when compared to N_{si} time-slots. However it is fairly short when compared to the single user active call period, and thus enabling fairly accurate estimation of P_{vi}. The accuracy of the results can be improved by reducing the N_{fmax}, which increases the computation burden. The probability of *bandwidth violation* (P_{vi}) is used to determine whether to admit cells or to discard them at runtime. In ATM networks, the congestion control schemes shall be simple and robust, for an immediate *reactive* followed by *preventive* controls to keep congestion within the tolerable limits. Hence, the instantaneous congestion criteria of the ATM multiplexer node, is computed using simple analytical formulas for LB and VLB schemes. The instantaneous congestion status is communicated to the internal transport users by means of service primitives. As video and voice traffic can be dropped moderately without significantly affecting the quality of the traffic, the transport users can judiciously drop some part of the traffic preserving the minimum quality of video or voice traffic during congestion periods. Thus, the transport users may modulate the protocol data unit (i.e., PDU) size for video and voice traffic, based on the instantaneous congestion criteria. As it is not recommended to vary PDU sizes within a message or file transfer, the higher layer window size (w) for data users, can be modulated based on the instantaneous congestion criteria. The PDU or window size modulation is referred as *active burst control*. The LB bandwidth enforcement scheme discards all cells that violate the QOS (i.e., based on P_{vi}). The VLB bandwidth enforcement scheme admits all cells, including

those which violate the QOS, as long as the ATM multiplexer node is not congested. The *modified* LB (mLB) and *modified* VLB (mVLB) schemes apply the *active burst control* on the top of LB and VLB schemes, respectively. The admitted cells are added to the ATM multiplexer buffer. The maximum size of the queue for a service-class (i.e., L_iB_i) is assumed to hold cells, sufficient for queueing a fixed number (ll) of composite TDMA frames. The throughput and delay characteristics of the ATM multiplexer node are directly proportional to the maximum size of the queue (i.e., L_iB_i).

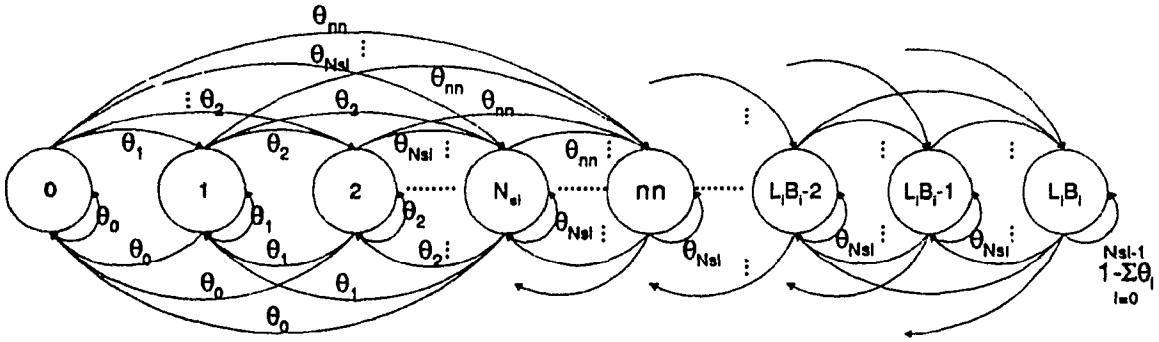

Fig. 4.6 (a) $nn > N_{si}$;

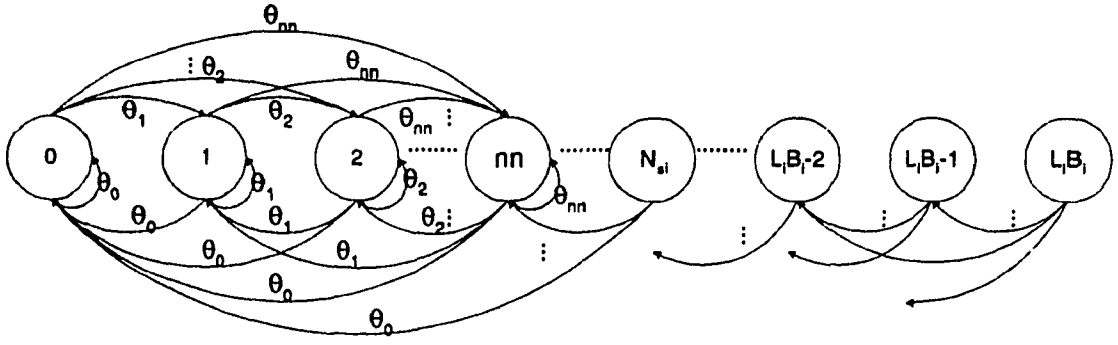

Fig. 4.6 (b) $nn \leq N_{si}$;

Fig. 4.6. State transition diagram for the queue length.

The ATM multiplexer node buffer is analyzed using a Markov modulated Poisson process (MMPP) [AKI-94] queueing model with bulk arrival (up to nn) and bulk service (N_{si}) of cells. As the instantaneous number of active bursts (i.e., k_i) varies continuously, the value of nn varies accordingly. Hence, the ATM multiplexer node buffer may buildup steadily when $nn > N_{si}$, and gets cleared when $nn < N_{si}$ as shown in Fig. 4.6, where θ_i is the probability of admitting i cells from nn. Hence, the ATM multiplexer node buffer moves from one degree of congestion to the other, based on the random arrival of cells during the TDMA service (i.e., N_{si} cells per TDMA frame for service class i). The ATM multiplexer node local congestion is defined as the amount of buffer content that exceeds a congestion threshold value (cng). The congestion threshold value is a percentage of maximum size of the queue (L_iB_i) for a service-class, beyond which the ATM multiplexer node shall take an immediate action to control congestion. As the maximum size of the queue for a service-class (L_iB_i) is fixed at subscription time, the admitted cells can not be queued when the queue is full, and thereby leading to buffer overflow or cell loss. To measure the cell-loss due to buffer overflow, we defined the buffer overflow as the amount of buffer content that exceeds an overflow threshold value (i.e , ofl), which is higher than congestion threshold (i.e., $ofl > cng$). If the value of L_iB_i is chosen to be higher at subscription time, the throughput will increase, and the probability of buffer overflow will decrease, but the transmission delay will increase for the same service-class i. The control schemes to admit cells are detailed in the flowchart of Fig. 4.7. The LB and mLB controls are described in the flowchart of Fig. 4.7 (a). The VLB and mVLB controls are described in the flow chart of Fig. 4.7 (b). The transmission procedure is detailed in the flowchart of

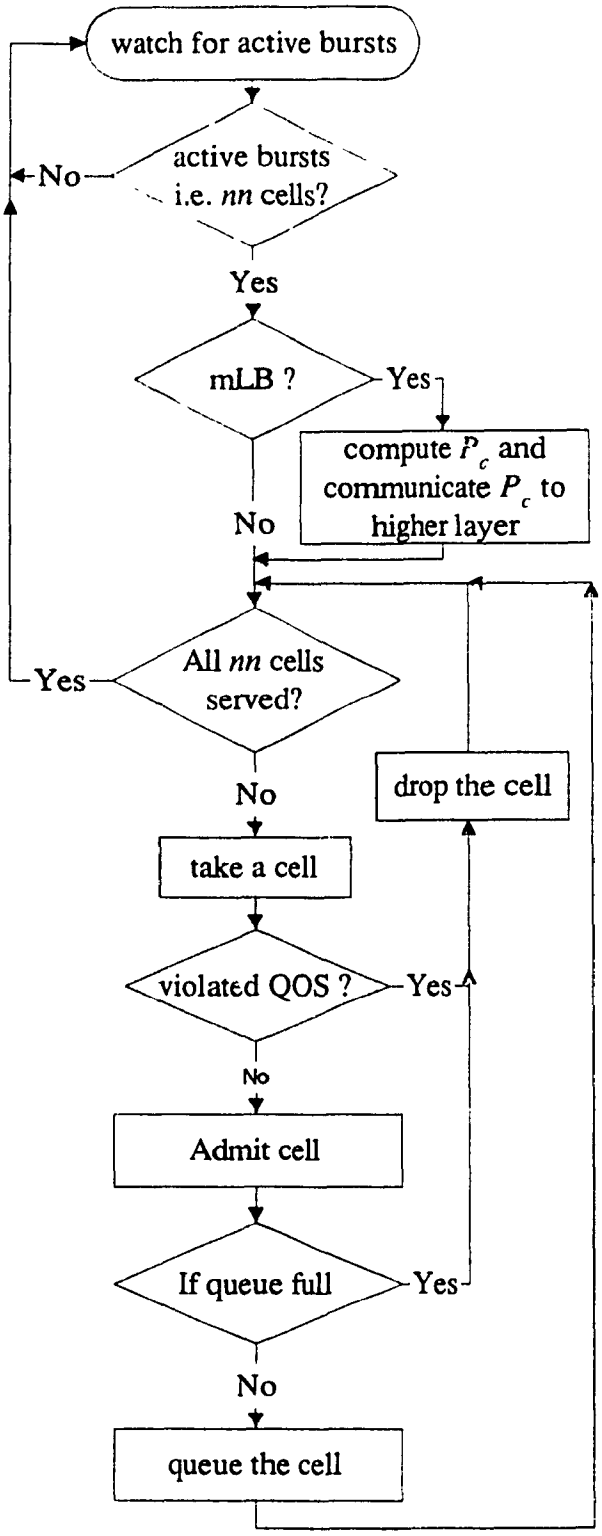

Fig. 4.7 (a) LB and mLB control;

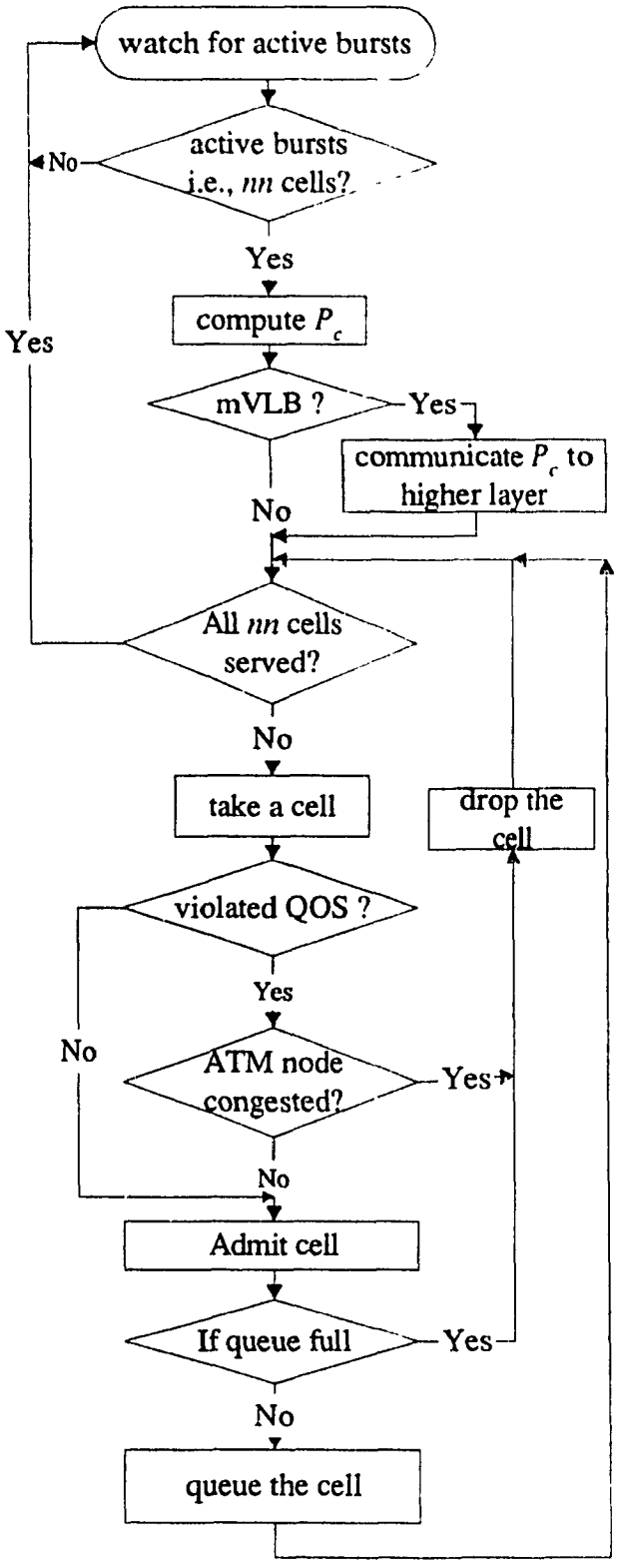

Fig. 4.7 (b) VLB and mVLB control;

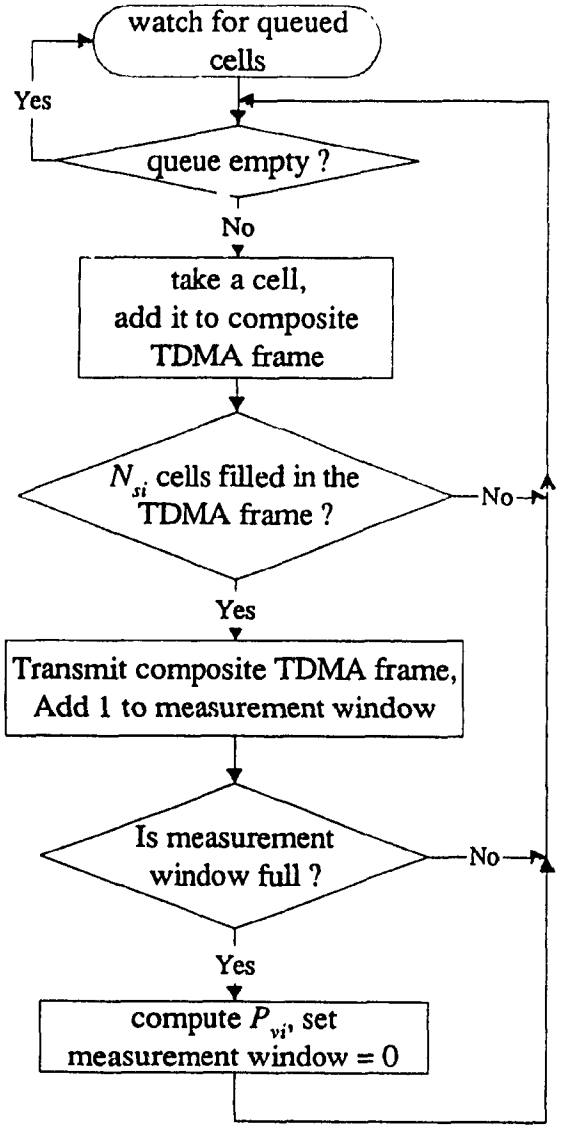

Fig. 4.7 (c) ATM cell transmitter;

Fig. 4.7. Congestion control schemes in ATM multiplexer.

Fig. 4.7 (c) for the admitted-cells. The ATM multiplexer congestion performance criteria, i.e., the mean probabilities of local congestion (P_c), generating a cell (ξ), discarding a cell (Δ), buffer content (Ψ), and buffer overflow ($P_{overflow}$), are obtained as detailed in the flowchart of Fig. 4.8.

4.3 Queueing analysis

In this section, we analyzed the queueing behavior of the ATM multiplexer node buffer with bulk arrival and bulk service of cells. The bulk arrival of cells is by means of active bursts of traffic from the generic ATM sources. The bulk service of cells is by means of composite TDMA frames. The total bandwidth (BW) of the ATM multiplexer node, is distributed among all service-classes based on peak rate assignment, as

$$BW = K_{av} \sum_{i=1}^{S} R_{pi} \tag{4.1}$$

where K_{av} is the average number of active transport users, R_{pi} is the peak service rate for a service-class i, and S is the total number of service-classes. The total number of time-slots for all classes of traffic in a TDMA composite frame are

$$N_s = \sum_{i=1}^{S} N_{si} \, . \tag{4.2}$$

The total number of time-slots in a TDMA subframe for a service-class i, are

$$N_{si} = n_{ui} K_{av} \tag{4.3}$$

where n_{ui} is the maximum number of cells which are allowed from a transport user of class i, during a composite TDMA frame service-time. For stream users, the maximum number of active calls, i.e., K_i, shall be set to K_{av} for full utilization of the channel

capacity. However, the system can support more active bursty users than K_{av} for better utilization of the channel capacity. The total number of active bursts for a service-class (k_i) can range from 0 to K_i (i.e., $0 < k_i \leq K_i$) during a composite TDMA frame service-time. Hence, the instantaneous maximum number of cells, which can be received during one composite frame service-time, are

$$nn = n_{ui}\, k_i \,. \tag{4.4}$$

However, the instantaneous number of cells that are actually generated by k_i active bursts, during a composite TDMA service-time, can be any number between 0 to nn. The maximum size of the queue (in terms of cells) for a service-class i is

$$LiB_i = n_{ui}.\, K_i.\, ll \tag{4.5}$$

where $n_{ui}\, K_i$ is the maximum number of cells which can be received during a composite TDMA frame service-time, and ll is the size of the queue in terms of composite TDMA frames which can be held in the queue of a service-class i. The probability that an active call generates an active burst of traffic, is

$$P_a = \frac{\beta_i}{\beta_i + \alpha_i} \,. \tag{4.6}$$

The probability that an active call does not generate an active burst of traffic, is

$$P_l = 1 - P_a \,. \tag{4.7}$$

There can be a variable number of cells generated by an active burst of traffic. The probability that k_i number of active bursts generate l number of cells of a service-class i, during a composite frame service-time, is assumed from equation (3.10) as

$$\theta_l = \frac{(nn.\xi)^l / l!}{\sum_{j=0}^{nn} [(nn.\xi)^j / j!]} \qquad where\ 0 \le l \le nn \tag{4.8}$$

where ξ (i.e., the probability of generating a cell which is admitted for transmission) will follow shortly. The equation (4.8) is applicable for both constant bit rate (CBR), and variable bit rate (VBR) traffic sources. It accounts for the burst to burst correlation in case of VBR traffic sources. However, the actual value of n_{ui}, and hence the nn of equation (4.4) can be found only from the simulation results of a VBR traffic source, such that the distribution in equation (4.8) can be determined uniquely. The probability that K_i active calls generating k_i active bursts of traffic, is given by

$$P(k_i) = \binom{K_i}{k_i} P_a^{k_i} (1 - P_a)^{K_i - k_i} . \tag{4.9}$$

An active user may momentarily exceed the allocated number of time-slots during an active burst of traffic by generating more cells than allowed, and thus violating the agreed bandwidth. The probability that the instantaneous l cells of service-class i violate the allocated bandwidth is

$$P_{vi} = \sum_{l=\eta_i+1}^{N_{F\max}.N_{si}} \binom{N_{F\max}.N_{si}}{l} \left(\frac{\rho_i}{K_i}\right)^{l} \left(1 - \frac{\rho_i}{K_i}\right)^{N_{F\max}.N_{si}-l}$$

where (4.10)

$$\eta_i = N_{F\max}.N_{si}.q_i / K_i$$

and ρ_i is traffic utilization per time-slot or often called as traffic intensity (which is defined shortly in equation 4.11), q_i is the QOS, K_i is the total number of active calls or transport users, and N_{si} is the total number of cells serviced by a composite TDMA frame, for traffic service-class i. A measurement window size of N_{Fmax} number of composite TDMA frames, is assumed for congestion performance measurement, to generate only η_i cells per measurement-window. The traffic utilization per time-slot for a service-class i can be defined as

$$\rho_i = P_a \frac{n_{ui} K_i}{N_{si}} . \tag{4.11}$$

It is assumed that an instantaneous maximum number of cells nn can be generated by k_i instantaneous active bursts according to equation (4.4). A composite TDMA frame services N_{si} cells to K_i statistically multiplexed active ATM sources of service-class i. As the cells of a service-class are admitted, the queue of a service-class i builds-up when the cell arrival rate is higher than service rate. The state transitions of the ATM multiplexer queue for a service-class follow Fig. 4.6. It can be observed that the queue size may rise (moves in the forward direction) when $nn > N_{si}$, as shown in Fig. 4.6 (a). However, the queue gets cleared (moves in the backward direction) or never builds beyond nn when $nn \leq N_{si}$, as shown in Fig. 4.6 (b). The transition between state diagrams Figs 4.6 (a) and (b) occur as the number of active bursts (k_i) varies with time and thereby varying the value of nn at any instance. The maximum possible number of cells nn that can be generated during a composite TDMA frame service, depends on the total number of active bursts k_i. Hence,

a state transition mainly depends on the total number of generated cells during a composite TDMA frame service. The transition probability from state l to j, when $nn > N_{si}$ can be derived as

$$
\begin{aligned}
b_{lj} = \theta_j \quad & where \quad j < l, l \le N_{si}, (l-j) \le N_{si} \\
= \theta_{N_{si}+j-l} \quad & j < l, l > N_{si}, (l-j) \le N_{si} \\
= \theta_j \quad & j > l, l \le N_{si}, j \le nn \\
= \theta_{N_{si}+j-l} \quad & j > l, l > N_{si}, 0 \le (N_{si}+j-l) \le nn \\
= \theta_j \quad & j = l, l \le N_{si} \\
= \theta_{N_{si}} \quad & j = l \ne L_i B_i, l > N_{si} \\
= 1 - \sum_{i=0}^{N_{si}-1} \theta_i \quad & j = l = L_i B_i, l > N_{si} \\
= 0 \quad & otherwise .
\end{aligned}
\tag{4.12}
$$

In equation (4.12), l represents initial state or the existing queue size, and j represents final state of the queueing system, while servicing one TDMA composite frame. During a composite TDMA service-time, a maximum of nn cells of service-class i can be received. When the ATM multiplexer node services all of the existing queue (i.e., $0 \le l \le N_{si}$) by a TDMA composite frame, it moves to a state j with just the probability receiving j cells, i.e., θ_j, where $0 \le j \le nn$. These conditions are applicable to the cases ($j<l$, $l \le N_{si}$, $l-j \le N_{si}$), ($j>l$, $l \le N_{si}$, $j \le nn$), and ($j=l$, $l \le N_{si}$). When $N_{si} < l < L_i B_i$, since the system services a maximum of N_{si} cells by a TDMA composite frame, leaving ($l-N_{si}$) cells in the queue, which shall be added to the arriving cells, the queue moves to a state j with the probability of receiving $j-(l-N_{si})$, i.e., $\theta_{N_{si}+j-l}$. These conditions are applicable to the cases ($j<l$, $l>N_{si}$, $l-j \le N_{si}$), ($j>l$, $l>N_{si}$, $0 \le N_{si}+j-l \le nn$), and ($j=l \ne L_i B_i$, $l>N_{si}$). Moreover, when queue is full, i.e.,

$l=L_iB_i$, the arrival of any number of cells greater than N_{si} will keep the queueing system in $j=L_iB_i$ state. Similarly, the transition probability from state l to j when $nn \le N_{si}$ can be derived as

$$\begin{aligned} b_{lj} &= \theta_j \qquad where\ j < l, l \le N_{si}, j \le nn \\ &= \theta_{N_{si}+j-l} \qquad j < l, l > N_{si}, 0 \le (N_{si}+j-l) \le nn \\ &= \theta_j \qquad j \ge l, j \le nn \\ &= 0 \qquad otherwise\ . \end{aligned} \tag{4.13}$$

In equation (4.13), the behavior of state transitions are exactly the same as those of equation (4.12) as long as $l \le N_{si}$ and $j \le nn$. However, when $l > N_{si}$, the state transitions will always be in the backward direction, since more number of cells can be serviced by a composite TDMA frame than the total number of received cells. Equations (4.12) and (4.13) depict all possible initial states and all possible number of cells, which can be generated (i.e., 0 to nn). As the queueing system toggles between state transition diagrams Figs 4.6. (a) and (b), the state transitions follow equations (4.12) and (4.13), respectively.

Applying the fundamental law of flow conservation, whereby the probability that the ATM multiplexer queue size for a service-class i assumes state j under equilibrium condition, is equal to the sum of various probabilities of arriving at state j from all possible states times the respective state transition probabilities, we obtain

$$\psi_j\big|_{k_i} = \sum_{l=0}^{L_iB_i} b_{lj}\big|_{k_i} \psi_l\big|_{k_i} \qquad where\ \ 0 \le j \le L_iB_i\ . \tag{4.14}$$

Also, as the sum of all state probabilities is one, giving rise to:

$$\sum_{i=0}^{L_iB_i} \psi(i)\big|_{k_i} = 1\ . \tag{4.15}$$

Furthermore, these equilibrium equations can be written in the following matrix form.

$$[A]_{k_i}.[\psi]_{k_i} = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} \qquad (4.16)$$

where

$A|_{k_i} = [a_{ij}]$,

$a_{ij} = b_{ij}$ *when* $l \neq j$

$= b_{ij} - 1$ *when* $l = j$ *and*

$\Psi|_{k_i} = [\psi_0 \ \psi_1 \ \psi_2 \ \psi_3 \ \cdots\cdots \ \psi_{L,B_i}]^T|_{k_i}$

and $\Psi|_{k_i}$ is the vector of state probabilities conditional to k_i active bursts, of the ATM multiplexer buffer of service-class i.

4.4 Congestion performance analysis

In this section, we propose analytical formulas for the steady state local congestion performance characteristics of a service-class i, under LB, VLB, mLB, and mVLB schemes. An ATM multiplexer node can be declared as congested when its buffer occupancy rises beyond a threshold. A conservative threshold of 50% of L_iB_i is arbitrarily chosen to declare ATM multiplexer node local congestion. However, the threshold shall be tuned depending on the system environment. After solving equation (4.13), the probability of declaring local congestion can be evaluated as

$$P_c|_{k_i} = \sum_{i=cng.\,L_iB_i}^{L_iB_i} \psi(i)|_{k_i} \qquad (4.17)$$

where $\psi(j)$ at a given number of active bursts k_i, is the probability of staying in state j, and *cng* is threshold beyond which *local congestion* is declared. At any given time, k_i number of active bursts from K_i number of active calls may generate traffic. The mean probability of local congestion, with K_i number of active calls, is

$$\overline{P}_c\Big|_{K_i} = \sum_{k_i=0}^{K_i} P_c\Big|_{k_i} P(k_i) \tag{4.18}$$

where $P(k_i)$ has been defined in equation (4.9).

The probability of an active user generating class i traffic, which is admitted for transmission (i.e., introduced to the ATM multiplexer buffer) is given by,

$$\begin{aligned} \xi\Big|_{k_i} &= [(1-P_{vi})+P_{vi}.(1-P_c\Big|_{k_i})]\frac{\rho_i}{K_i} \quad \text{with VLB} \\ &= (1-P_{vi})\frac{\rho_i}{K_i} \qquad\qquad \text{with LB} \\ &= \frac{\rho_i}{K_i} \qquad\qquad \text{without congestion control.} \end{aligned} \tag{4.19}$$

Here we see that, in case of VLB scheme, a cell is introduced to the ATM buffer if it is generated without violating the QOS, i.e., $(1-P_{vi})$, or when it violates the QOS during the periods of no local congestion, i.e., $P_{vi}(1-P_c)$. The LB scheme drops all violating cells irrespective of the congestion status of the ATM multiplexer node. However, congestion develops rapidly when neither of these cell-discard schemes is used. The mean probability of generation of a cell from an active user with the overall burst condition ξ, is given by

$$\overline{\xi}\Big|_{K_i} = \sum_{k_i=0}^{K_i} \xi\Big|_{k_i} P(k_i) \,. \tag{4.20}$$

The probability that an active user generates a cell, which gets discarded, is given by

$$\Delta|_{k_i} = P_{vi} P_c|_{k_i} \frac{\rho_i}{K_i} \quad \text{with VLB}$$

$$= P_{vi} \frac{\rho_i}{K_i} \quad \text{with LB} \qquad (4.21)$$

$$= 0 \quad \text{without congestion control.}$$

Equation (4.21) depicts that VLB scheme discards all violating cells only when the ATM multiplexer node experiences local congestion, i.e., with a probability $P_{vi}P_c|_{ki}$, whereas the LB scheme discards all traffic that violates the QOS with a probability P_{vi}. The mean probability of discarding a cell, averaged over all possible number of bursts k_i, is

$$\overline{\Delta}|_{Ki} = \sum_{k_i=0}^{Ki} \Delta|_{k_i} P(k_i) . \qquad (4.22)$$

Once the steady state distribution $\Psi(j)|_{ki}$ where $0 \leq j < L_i B_i$, is found, by solving the simultaneous equations (4.12) and (4.13), the ATM multiplexer node buffer content, for a given number of active bursts k_i, can be defined as

$$\overline{\psi}|_{k_i} = \sum_{j=0}^{L_i B_i} j \, \psi(j)|_{k_i} \qquad (4.23)$$

where L_iB_i is the total buffer space, which is defined in equation (4.5). Averaging over the distribution of the number of the bursts k_i, i.e., $P(k_i)$, we obtain the ATM multiplexer node mean buffer content of service-class i, for a given number of active calls K_i, as

$$\overline{\psi}|_{Ki} = \sum_{k_i=0}^{Ki} \overline{\psi}|_{k_i} P(k_i) . \qquad (4.24)$$

A conservative approach to define buffer overflow ($P_{overflow}$) is to take higher order states probabilities, instead of only one final buffer state (i.e., at state L_iB_i). The probability of buffer overflow of the ATM multiplexer node, can be written as

$$P_{overflow}\big|_{k_i} = \sum_{j=ofl.L_iB_i}^{L_iB_i} \psi(j)\big|_{k_i} \tag{4.25}$$

where *ofl* is the buffer overflow threshold value, e.g. 90% of L_iB_i . The mean probability of buffer overflow of the ATM multiplexer node, is

$$\overline{P}_{overflow}\big|_{K_i} = \sum_{k_i=0}^{K_i} P_{overflow}\big|_{k_i} \; P(k_i) \; . \tag{4.26}$$

For video and voice traffic, the mLB and mVLB schemes are assumed to have PDU size modulation at higher layers, based on the instantaneous congestion criteria $P_c\big|_{k_i}$ of the ATM multiplexer node. The instantaneous PDU size for mLB and mVLB schemes, is

$$p = p_{max}\left(1 - P_c\big|_{k_i}\right) \tag{4.27}$$

where p is instantaneous PDU size, p_{max} is the subscribed PDU size, and $P_c\big|_{k_i}$ is the instantaneous local congestion.

Similarly, since data traffic favors window transport mechanisms, the mLB and mVLB schemes modulate the transport user's window size, considering the instantaneous congestion criteria $P_c\big|_{k_i}$ of the ATM multiplexer node. The instantaneous transport user window for mLB and mVLB schemes, is

$$w = w_{max}\left(1 - P_c|_{k_i}\right) \tag{4.28}$$

where w_{max} is the subscribed window size.

Equations (4.27) and (4.28), result in modulated active bursts by the same amount of traffic control. Hence, the probability that an active source generates traffic, which was defined in equation (4.6) shall be modified to

$$P_a' = \left(\frac{\beta_i}{\alpha_i + \beta_i}\right)\left(1 - P_c|_{k_i}\right) \tag{4.29}$$

for mLB and mVLB schemes.

The system of equations (4.8) to (4.26), shall be repeated with P_a' instead of P_a to modify the congestion performance criteria of LB and VLB schemes to mLB and mVLB schemes, respectively.

4.5 Computation procedure and results

The congestion performance characteristics, i.e., the probabilities of ATM multiplexer node local congestion (P_c), cell generation (ξ), cell discard (Δ), buffer content (ψ) and buffer overflow ($P_{overflow}$), were obtained for hybrid ATM/TDMA network, using the system of equations (4.1)-(4.29). The computation of congestion performance characteristics is detailed in the flowchart of Fig. 4.8. Once the static system parameters such as K_i, K_{av}, n_{ui}, R_{pi}, S, ll, q_i, N_{fmax}, etc., are known or assumed, the related static parameters such as N_{si}, L_iB_i can be obtained from equations (4.3)-(4.5). The performance

parameters, namely congestion threshold (*cng*), buffer overflow threshold (*ofl*) and convergence of P_c value in terms of percentage deviation from its value in the previous iteration (*%cnv*) are assumed to be 50%, 90% and 5%, respectively. As the averaging procedure needs all possible traffic generation conditions, the factor P_a, i.e., $\beta_i/(\alpha_i+\beta_i)$ as in the equation (4.6) for LB and VLB schemes, or P_a', i.e., $(1-P_c|_{k_i})\beta_i/(\alpha_i+\beta_i)$ as in the equation (4.29) for mLB and mVLB schemes, is varied from 0 to 1, and the same effect is reflected on the traffic utilization (ρ_i) of equation (4.11). Moreover, the active bursts k_i are also varied from 0 to K_i. Once the system parameters are specified, one can assume an initial value of $P_c|_{ki}$ at a certain ρ_i, k_i, and evaluate the values of $P_{vi}|_{ki}$, $\xi|_{ki}$, $\Delta|_{ki}$ from equations (4.10), (4.19) and (4.21), respectively. Subsequently, $\theta_i|_{ki}$, is evaluated from equation (4.8) and the state diagram for the ATM multiplexer node is solved to obtain $\Psi(j)|_{ki}$ where $0 \le j \le LB_i$, and a new value of $P_c|_{ki}$ is then obtained from equation (4.17). With this value of $P_c|_{ki}$, we repeat the same procedure to evaluate new values of $P_{vi}|_{ki}$, $\xi|_{ki}$, $\Delta|_{ki}$ and $\Psi(j)|_{ki}$ where $0 \le j \le LB_i$, and a new value of $P_c|_{ki}$ until it converges to 95% of its value in the previous iteration (i.e., *%cnv* is 5%). By repeating the same procedure for all possible active bursts, i.e., k_i and traffic generation, i.e., P_a for LB and VLB schemes, or P'_a for mLB and mVLB schemes, and averaging all criteria over the distribution of $P(k_i)$, $\overline{P}_c|_{K_i}$, $\overline{\xi}|_{K_i}$, $\overline{\Delta}|_{K_i}$, $\overline{\Psi}|_{K_i}$, and $\overline{P}_{overflow}|_{K_i}$ were obtained and plotted. Moreover, the results were obtained for different values of N_{si} to observe the effect of performance characteristics with the service rate.

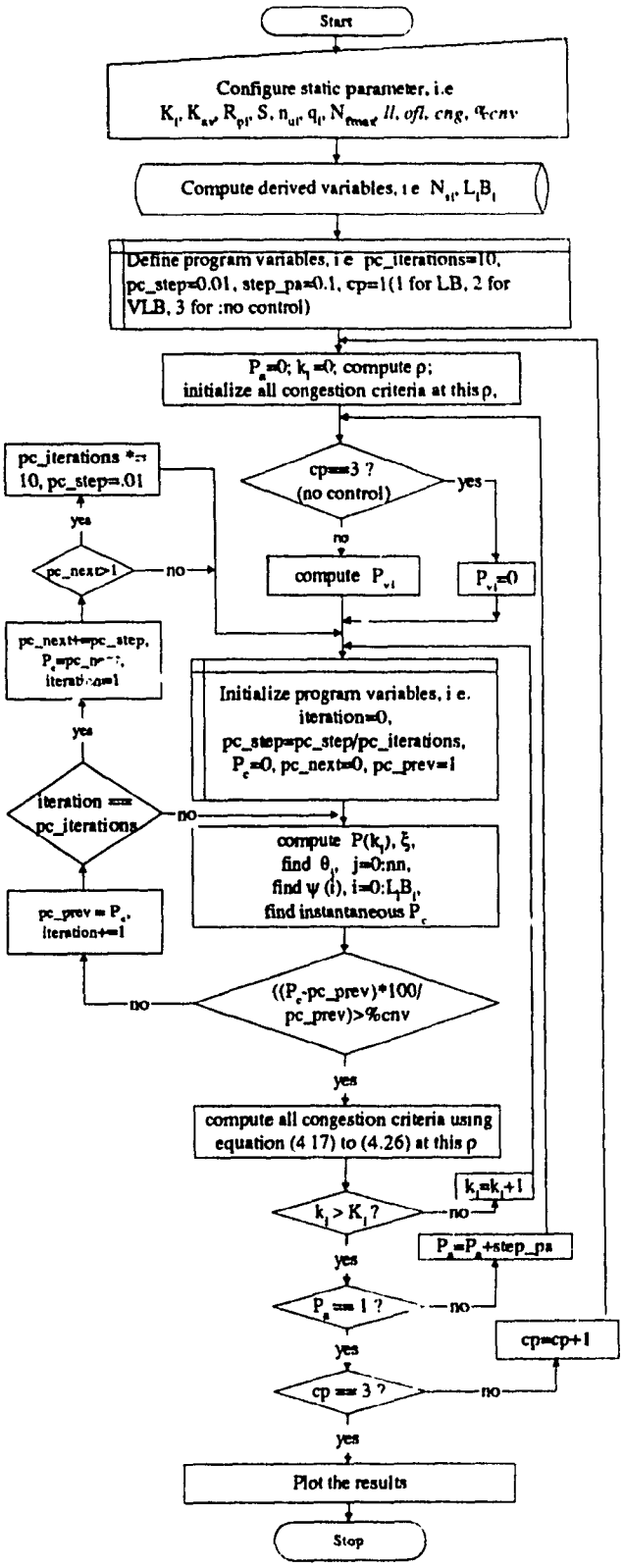

Fig. 4.8. Computation of performance criteria.

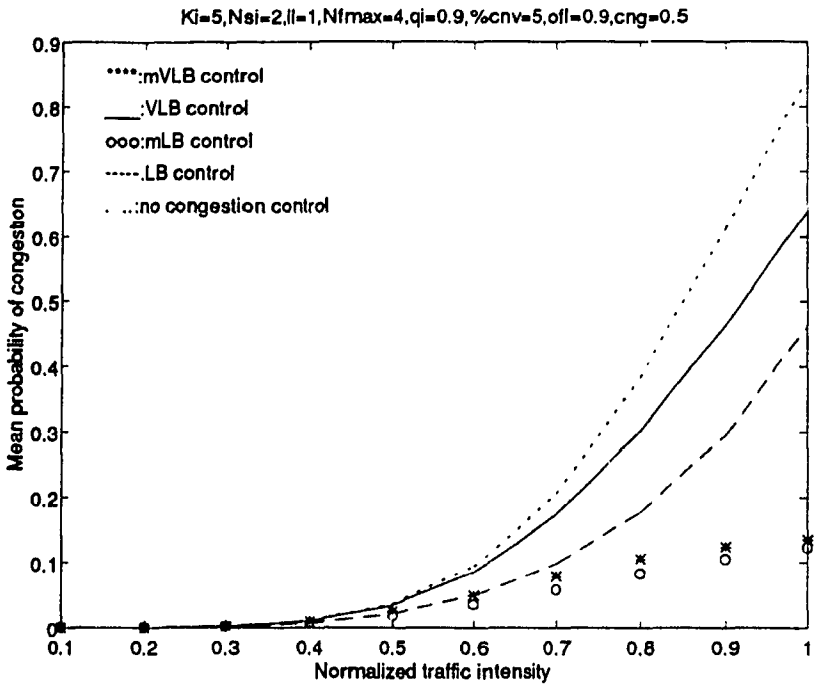

Fig 4.9 (a) $N_{si} = 2$;

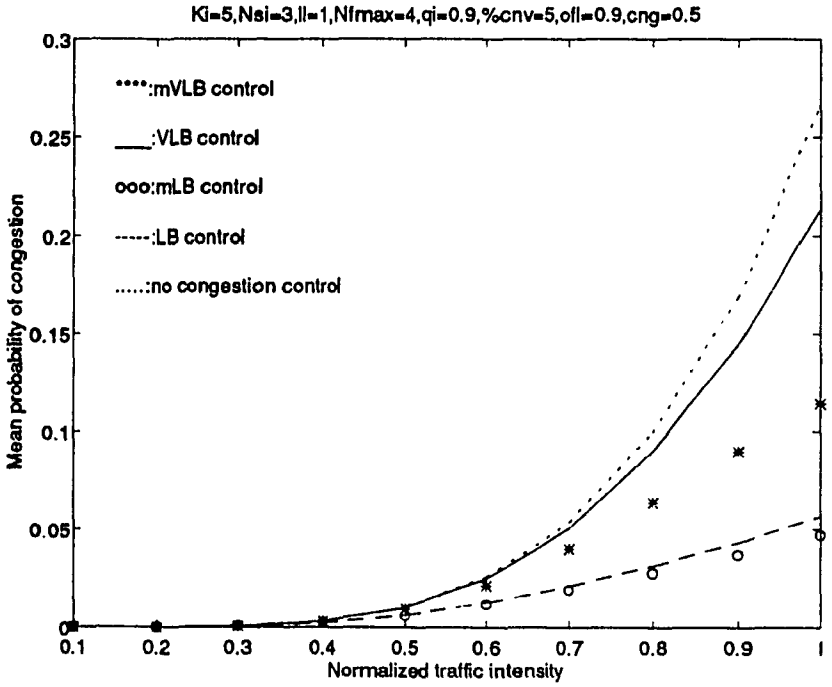

Fig. 4.9 (b) $N_{si} = 3$;

Fig. 4.9. Mean probability of congestion.

The ATM multiplexer node local congestion is obtained as a function of traffic intensity in Fig. 4.9 for N_{si} values of 2 and 3 in (a) and (b) respectively. It is obvious from the results that the ATM multiplexer local congestion increases with the traffic intensity. It can be observed that the congestion develops at a slower rate for LB scheme, as it discards cells faster than VLB to control congestion. By controlling the active bursts at the transport users, the mLB and mVLB schemes further reduce the congestion than the respective LB and VLB schemes. Moreover, the mean probability of congestion reduced when the service rate (N_{si}) is increased, which can be observed from Figs 4.9 (a) and (b). It can also be observed from the results that the ATM multiplexer node congestion develops at a higher rate when the control schemes are not employed.

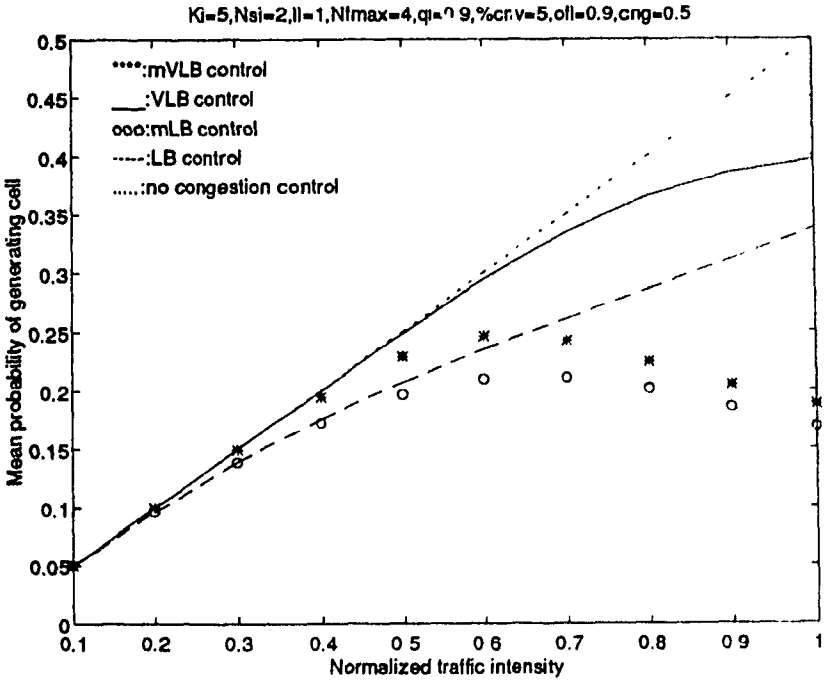

Fig. 4.10 (a) $N_{si} = 2$;

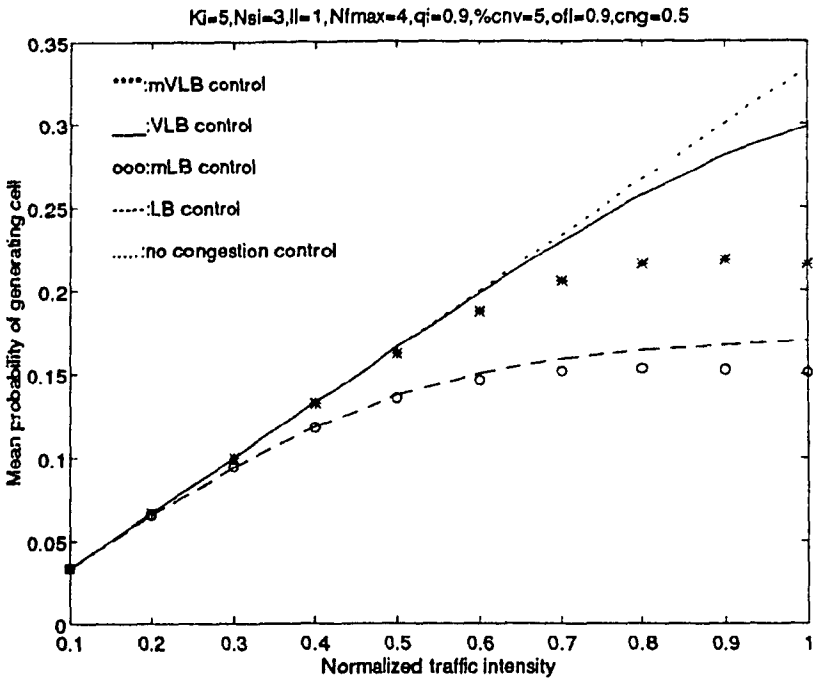

Fig. 4.10 (b) $N_{Si} = 3$;

Fig. 4.10. Mean probability of generating a cell.

The mean probability of generating a cell is obtained as a function of traffic intensity in Fig. 4.10. The mean probability of generating a cell increases linearly when there is no congestion control. However, the LB and VLB schemes drop the traffic based on P_{vi} and P_c as specified in (4.19), thereby reducing the probability of generating a cell during violation or congestion periods. The mLB and mVLB schemes control the active bursts based on the instantaneous congestion, thereby reducing the probability of generating a cell furthermore, by leaving the traffic at the transport users instead of dropping at the ATM layer, as observed in Fig. 4.10. It can also be observed from Figs

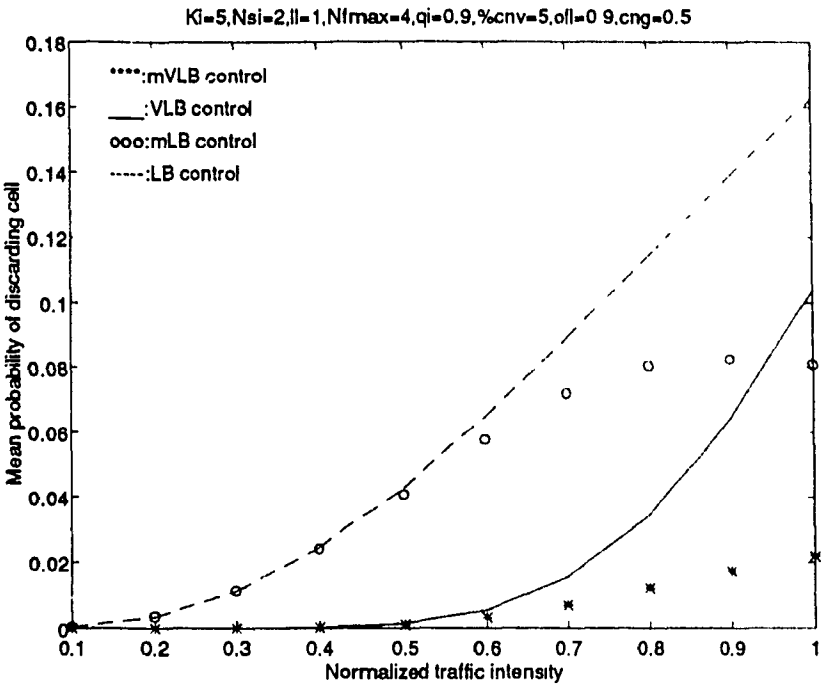

Fig. 4.11 (a) $N_{si} = 2$;

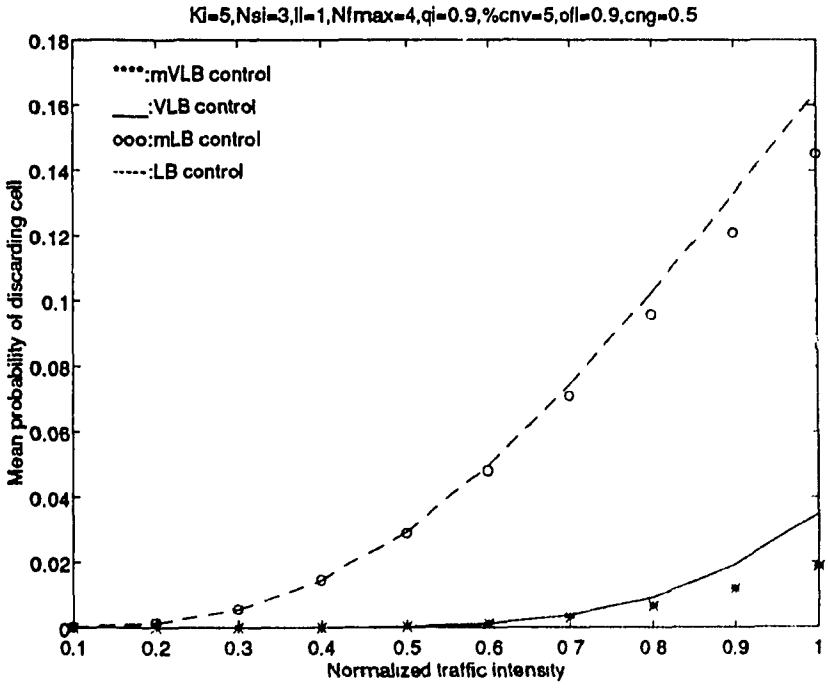

Fig. 4.11 (b) $N_{si} = 3$;

Fig. 4.11. Mean probability of discarding a cell.

4.10 (a) and (b) that higher capacity transport users can be handled when N_{si} is increased, as more number of cells are serviced. The probability of generating a cell rises when a congestion control scheme is not used, which may contribute more "cell discards" and hence end-to-end packet retransmissions.

The mean probability of discarding a cell is obtained as a function of traffic intensity in Fig. 4.11. It can be noted that the cells are not discarded when a congestion control scheme is not used. Hence, we compare only the LB, VLB, mLB, and mVLB congestion control schemes in Fig. 4.11. It is obvious that P_{vi} increases with ρ_i for the same QOS (q_i) of a service-class i. Since the LB cell discard scheme drops all cells which

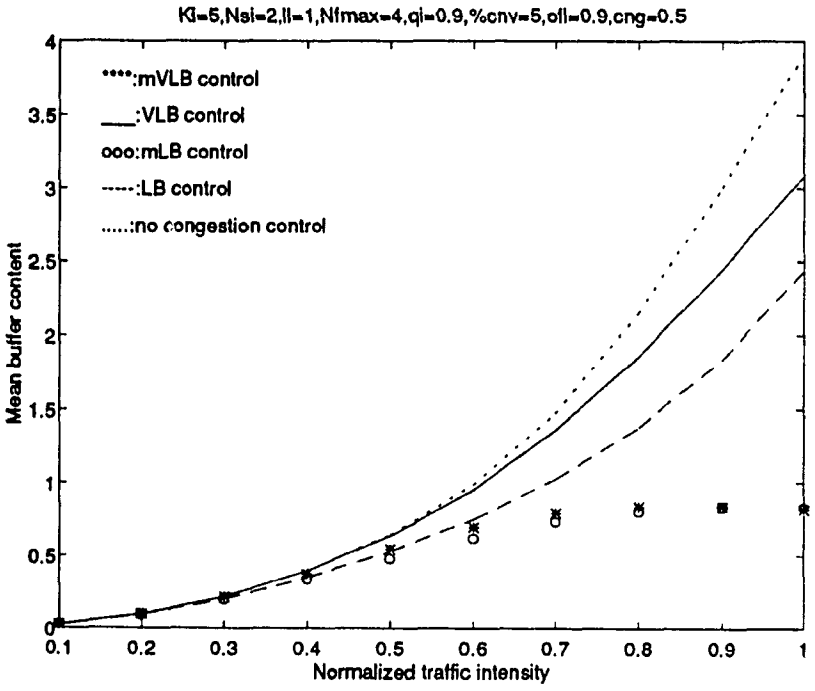

Fig. 4.12 (a) N_{si} =2;

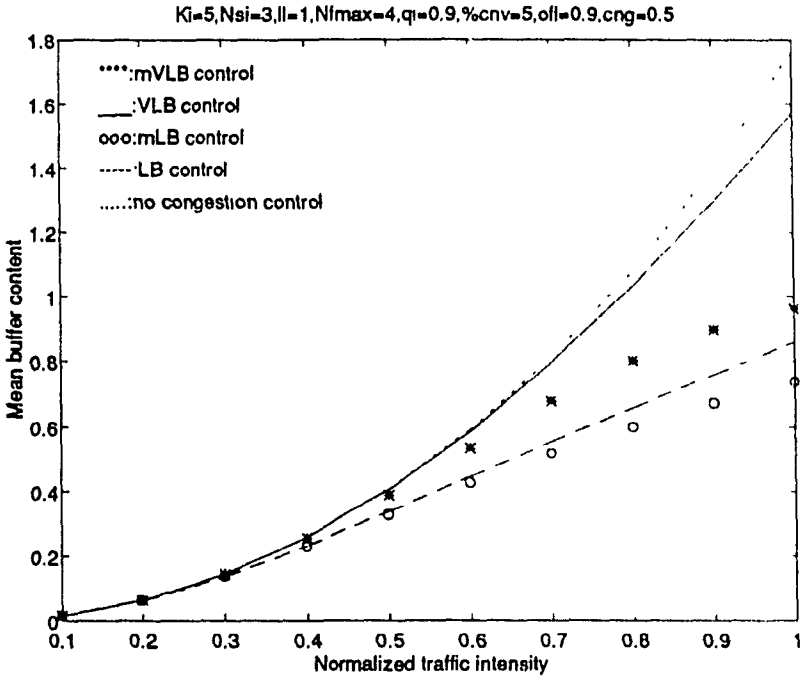

Fig. 4.12 (b) $N_{Si} = 3$;

Fig. 4.12. Mean buffer content.

violate the QOS, the probability of discarding a cell increases with traffic intensity (ρ_i). In case of VLB cell discard scheme, violating cells are not dropped unless local congestion is declared, and hence the probability of discarding a cell is lesser than in the case of LB, for the same traffic intensity, which can be observed in Fig. 4.11. The mLB and mVLB schemes reduce the necessity to drop the traffic at the ATM layer, by controlling the active bursts at the transport users to reduce the instantaneous congestion. It can also be observed in Figs 4.11 (a) and (b) that the probability of discarding a cell decreased as the service rate (N_{si}) increased.

The mean buffer content of the ATM multiplexer node is obtained as a function of traffic intensity in Fig. 4.12. The mean buffer content increases as the traffic intensity increases. As the *active burst control* reduces the traffic during congestion periods, the mLB and mVLB schemes reduce the buffer content, compared to the respective LB and VLB schemes, as in Fig. 4.12. It can also be observed from Figs 4.12 (a) and (b) that mean buffer content decreased as N_{si} increased. The mean buffer content rises more rapidly when the ATM multiplexer node does not use any congestion control scheme.

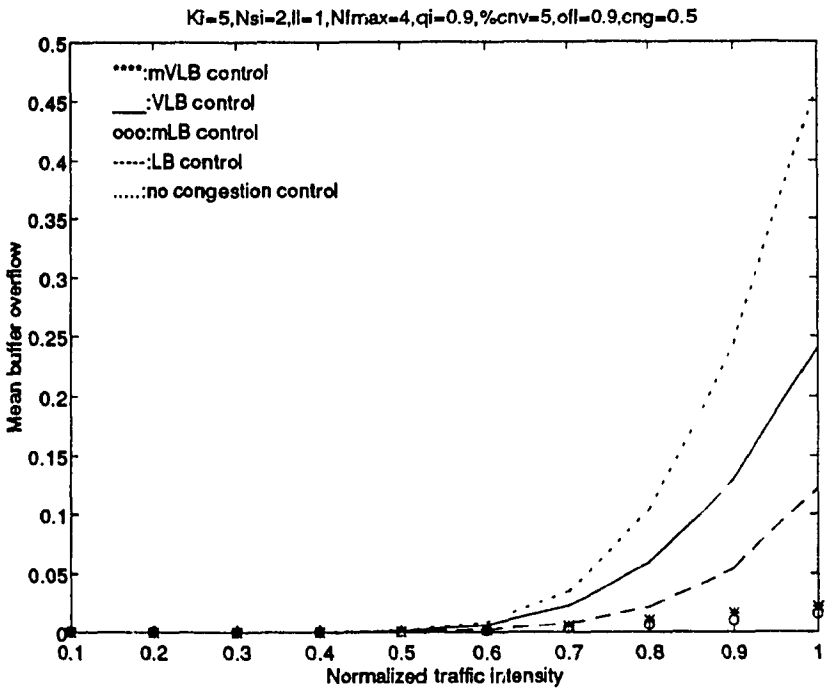

Fig. 4.13 (a) N_{Si} =2;

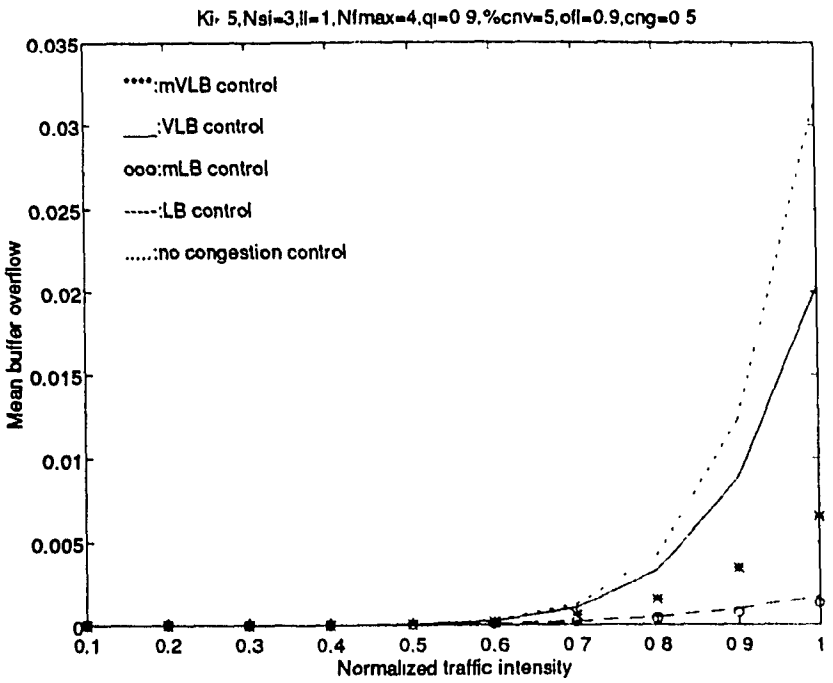

Fig. 4.13 (b) N_{si} =3;

Fig. 4.13. Mean buffer overflow.

The mean buffer overflow of the ATM multiplexer node is obtained as a function of traffic intensity in Fig. 4.13. As the buffer content rises beyond an overflow threshold (*ofl*), the buffer overflow also rises leading to cell-loss. Hence, the results show no buffer overflow (i.e., zero) when the traffic intensities are low. As the LB scheme drops all violating cells, the LB scheme contributes to less buffer overflow, compared to VLB scheme. The mLB and mVLB schemes reduce the buffer overflow for the same traffic intensity, when compared to the respective LB and VLB schemes. It can be observed from the results that the buffer overflow reduces if the service rate (N_{si}) is increased.

4.6 Conclusions

A new local congestion control scheme is developed for an ATM multiplexer node, to provide the required bandwidth for supporting various classes of service in a broadband hybrid ATM/TDMA network. It was observed from Figs 4.9-4.13 that the mLB and mVLB schemes consistently improved congestion performance criteria over the respective LB and VLB schemes. Almost all congestion criteria were interrelated to explain the performance of the congestion control scheme. The VLB discard scheme controlled congestion when ATM multiplexer node buffer-content crossed the congestion threshold. When congestion gradually increased with traffic intensity, the mLB and mVLB schemes reduced the overall local congestion level, by preventing traffic from transport users. The buffer overflow criterion helps to protect the ATM multiplexer node from a possible crash due to prolonged local congestion, by dropping the excessive cells, which are admitted without violating the QOS. Hence, the ATM multiplexer congestion was well controlled by mLB and mVLB schemes, when compared to the respective LB and VLB schemes. It was also observed that the mLB scheme provided a better control of congestion when compared to mVLB scheme, since mVLB scheme discarded the violating cells only when the ATM multiplexer node congestion was intolerable. However, it was also clear from the results that the mVLB scheme provides higher throughput when compared to mLB scheme. Hence, the mVLB scheme can be considered as an optimum choice when compared to mLB scheme. By controlling the active bursts at the transport users, judicious actions (to drop or to hold the traffic) are possible at the transport layer. These actions will improve throughput for the data traffic by reducing end-to-end re-

transmissions that may be needed due to cell-loss. Moreover, the quality of video and voice traffic will be preserved by selectively dropping the excessive traffic. To implement and utilize the results of this chapter in interconnected ATM networks, the ATM multiplexer node shall relay *early warning cells* with congestion criteria periodically to the neighboring nodes. Then, the neighboring ATM multiplexer nodes will be able to control active bursts from their own transport users according to the *early warning cells*, to prioritize the transit traffic over local traffic. In interconnected ATM networks, with large propagation delays and high transmission speeds, such extension of the policy should localize the congestion to fewer ATM multiplexer nodes as opposed to the state-of-art ATM networks that do not have these capabilities.

CHAPTER 5

GLOBAL CONGESTION CONTROL

5.1 Introduction

ATM networks support several diverse types of traffic sources, some of which are highly bursty, at higher link speeds ranging from several hundreds of megabits or gigabits per second. At these link-speeds, the propagation delay is a significant bottleneck to control congestion in broadband interconnected ATM networks. In such networks, congestion develops rapidly, not only due to the highly bursty traffic generated by its own traffic sources but also due to the bursty or sustained peak-rate transit traffic from other nodes. Uncontrolled or under-estimated congestion eventually disconnects several virtual circuits or creates system outage. Hence, it is important to communicate the congestion criteria of an ATM multiplexer node to all neighboring ATM multiplexer nodes periodically, so that faster reactive followed by preventive congestion controls can be applied to the local (internal) traffic sources. The propagation delay becomes a bottleneck, even for congestion-criteria communication in high speed networks, because the network can be flooded before congestion status is propagated. Several new methods were investigated to address the propagation delay problem. These new schemes were implemented by means of congestion notification messages, e.g., global traffic control by harmonious cooperation of call admission control (CAC) and usage parameter control (UPC) [PAR-95]. Several other schemes are fast resource management based congestion control using pilot ATM cells for indicating the payload [DA2-93], [DA3-93], fast

reservation protocols and explicit congestion notification [WER-92], burst-level feedback control using generic flow control (GFC) field of ATM cell header [RUB-94], backward congestion notification using ATM resource management (RM) cells [NEW-94], etc..

In this chapter, we propose a new global congestion control scheme in an ATM multiplexer node. This scheme prioritizes transit traffic from other ATM multiplexer nodes by controlling active bursts from the local transport users, to ensure necessary bandwidth for supporting various classes of service, i.e., video, voice, data, etc., in interconnected hybrid ATM/TDMA networks. This chapter also analyzes the end-to-end throughput and delay characteristics of the transport users with the underlying ATM global congestion control scheme. The ATM multiplexer node is assumed to communicate the congestion criteria by transmitting *early warning cells* to the neighboring ATM multiplexer nodes. The early warning cells are used to compute global congestion criteria by means of a new global congestion control algorithm. The objective of the global congestion control algorithm is to prioritize the transit traffic by controlling the local traffic of the ATM multiplexer node. The global congestion control algorithm uses a set of global congestion *levels or states*, based on the degree of congestion propagation. The congestion propagation is analyzed by means of two groups of adjacent nodes, i.e., how many neighboring nodes are *informed and uninformed* of the congestion criteria, among a cluster of c adjacent ATM multiplexer nodes. In other words, the *informed* nodes constitute the group of nodes, from where the congestion criteria were received by the concerned ATM multiplexer node. It is obvious that the steady-state congestion behavior

of the concerned ATM multiplexer node is identical to all other homogeneous ATM multiplexer nodes of the global network. Hence, the global network congestion can be controlled by accounting the "*informed nodes*" at the concerned ATM multiplexer node for steady-state congestion-criteria computation. A simplified approach for congestion duration is also analyzed, based on the global congestion control algorithm, which is a significant improvement over [KAM-89]. By communicating the ATM congestion criteria to the transport users by means of *service primitives*, the active bursts of traffic can be controlled based on the instantaneous congestion criteria, which is referred as *active burst control*. Moreover, when the transport user and ATM adaptation layers exist in distributed processing systems (probably using different processors), it is very appropriate to use layer interaction to control system resources (especially buffers). As stated in chapter 4, the length of an active burst can be controlled by modulating the transport user protocol data unit (PDU) size, based on the global congestion criteria in case of video and voice users. As PDU size may not be controlled for data users, the number of active bursts can be controlled by modulating the transport user window-size based on the global congestion criteria. By controlling the active bursts judiciously to preserve the quality of picture or speech, the number of cells that are generated during congestion periods, can be effectively controlled, thereby helping to reduce ATM multiplexer node global congestion. Moreover, the video and voice transport users can implement *quality control algorithms* to selectively drop parts of the PDU information during *severe congestion* periods, without sacrificing the quality of the picture and speech information. Such decisions are not possible at ATM layer LB and VLB cell discard schemes. Hence, the PDU size

modulation scheme is a significant improvement over window control mechanisms of [RAT-91]. Moreover, the higher layer window-control scheme reduces the possible end-to-end re-transmissions that may otherwise be prompted by cell-loss during congestion, thereby enhancing the end-to-end throughput for data traffic significantly. Hence, it is clearly a major improvement over the *fluid technique* of [DA3-93], which accepts or rejects an active burst. Simple analytical formulas for the ATM multiplexer node global congestion criteria, i.e., the mean probabilities of local congestion (P_c), congestion duration (CD), generating a cell which gets admitted for transmission (ξ), discarding a cell (Δ), buffer content (Ψ), and buffer overflow ($P_{overflow}$) are also derived. The ATM multiplexer congestion performance of mLB and mVLB are compared with LB and VLB schemes, respectively. Moreover, the throughput and delay characteristics were analyzed for the virtual circuit transport user with and without the underlying global congestion control scheme.

5.2 Description of the model

The ATM network is assumed with homogeneous ATM multiplexer nodes. An ATM multiplexer node is assumed to interface a cluster of c neighboring ATM multiplexer nodes to receive transit-traffic as shown in Fig. 5.1. In this section, *early warning cells* are proposed to propagate congestion criteria to the cluster of all c neighboring ATM multiplexer nodes as shown in Fig. 5.1. A global congestion control algorithm is designed for both local-traffic and transit-traffic by means of *early warning cells* from all neighboring c ATM multiplexer nodes, to take reactive followed by preventive measures

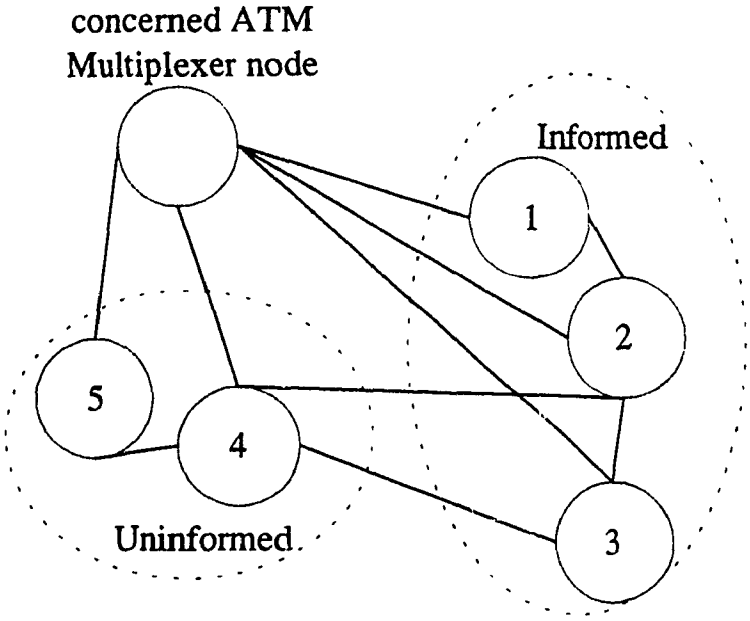

Fig. 5.1. A cluster of c (c=5) neighboring nodes in ATM Network.

to control congestion. The *early warning cells* are transmitted periodically, and at the beginning of each active burst of traffic. As it is impossible to inform congestion criteria to all c neighboring ATM multiplexer nodes instantaneously, there will be v nodes informed and c-v nodes uninformed of the congestion criteria at any given instance. For simplicity, only three global congestion states (i.e., l=0, 1 *and* 2) are assumed to reflect *local congestion* (l=0), *light global congestion* (l=1), and *severe global congestion* (l=2), as shown in Fig. 5.2. The ATM multiplexer node moves from one global congestion state to another state, based on congestion information from the adjacent nodes and active bursts from the local transport users. The global congestion state-transition probabilities are $H_{i,j}$, where i is the initial state and j is the final state. The $H_{i,j}$ (i=0, 1, *or* 2; j =0, 1, *or* 2) are found from the amount of buffer content, and the number of *informed nodes* v (as detailed in section 5.3). The global congestion state l=0, *i.e., local conges ion*, refers to low traffic

conditions or node initialization when the ATM multiplexer node is unaware of congestion status of other ATM multiplexer nodes. During local congestion state l=0, fairly a large portion of the ATM multiplexer node buffer and line capacity are assigned to the local ATM sources, whereas a smaller portion of these resources is assigned to the transit traffic. Two threshold levels $TH_{l,1}$ and $TH_{l,2}$ are defined with hysteresis to fallback to a lower congestion state, or to move to a higher congestion state. Hence, the values of $TH_{l,1}$ shall be assumed in a descending order while l increases, to allocate less amount of buffer for local transport users. However, the values of $TH_{l,2}$ may also be assumed in a descending order while l increases, for faster transition to a higher global congestion state. As l increases, the ATM multiplexer node will slowdown the local transport users by means of active burst control. The portion of buffer capacity assigned to local traffic when l=0, is $TH_{0,1}$ (say 0.6 L_iB_i, i.e., 60% of the buffer is allocated to the local traffic, whereas the remaining 40% is allocated to the transit traffic). When all ATM multiplexer nodes are in l=0, if the local traffic of an ATM multiplexer node exceeds $TH_{0,1}$ but is still less than $TH_{0,2}$ (say 0.8 L_iB_i), then it switches to a new congestion state l=1, i.e., *light global congestion*. When l=0, the ATM multiplexer node may be switched to a higher congestion state l=2, if the buffer-content of (*i*) two ATM multiplexer nodes or more, each exceeding $TH_{0,1}$ but less than $TH_{0,2}$, or (*ii*) one ATM multiplexer-node exceeds $TH_{0,2}$. When l=1, the ATM multiplexer node may be switched to l=2, when at least one ATM multiplexer node's local-traffic exceeds $TH_{1,2}$ (say 0.6 L_iB_i). While in global congestion state l=2, the ATM multiplexer node remains in the same state as long as the local traffic exceeds $TH_{2,2}$ (say 0.4 L_iB_i). When the ATM multiplexer node is in a higher global

congestion state (l=1 or 2), it may fallback to a lower congestion state when the local traffic falls below $TH_{l,l}$. The ATM multiplexer node follows the same queueing analysis of buffer state diagram of Fig. 4.6 in each global state l. The buffer state diagram of Fig. 5.2 shall be solved to find the probabilities of global congestion states. The mean congestion criteria in each global congestion state shall be averaged over all global congestion states to find the effective global congestion criteria.

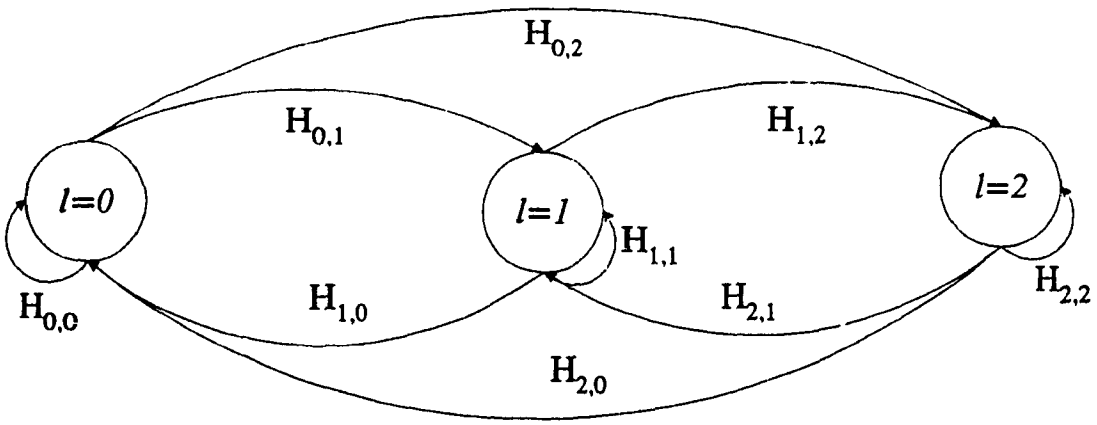

Fig. 5.2. Global congestion control state diagram.

5.3 Analysis of global congestion control

The queueing behavior of ATM multiplexer node is analyzed in chapter 4. In this section, the ATM multiplexer node buffer is further analyzed based on global congestion algorithm. The ATM multiplexer node global state-variable l starts with a value of zero, i.e., *local congestion* state, during the node initialization. Later, the ATM multiplexer node moves to appropriate global state depending on the instantaneous global congestion status. Once the steady-state buffer distribution is found in a global congestion state l

(i.e., $\psi(j)\big|_{k_i,\rho_i,l}$ where $0 \le j < L_i B_i$) by solving the simultaneous equations (4.14)-(4.18), the buffer content is averaged over active bursts k_i (i.e., $0 \le k_i \le K_i$), as

$$\psi'_j\big|_{K_i,\rho_i,l} = \sum_{k_i=0}^{K_i} \psi_j\Big|_{k_i,\rho_i,l} P(k_i) \quad where\ j = 0\ to\ L_i B_i\ . \tag{5.1}$$

As mentioned earlier, the state $l=0$ can be arrived when all informed nodes stay in *local congestion* state, or simply the concerned ATM multiplexer node is unaware of the congestion status of other nodes. In state l=1, the ATM multiplexer node reduces its buffer limit for local traffic, by lowering the aforementioned thresholds to $TH_{l,1}$ and $TH_{l,2}$. Thus, the ATM multiplexer node allows more transit traffic to pass-by, to relieve the network from congestion (by back-pressure action) during l=1. Similarly, these buffer thresholds are adjusted to $TH_{2,1}$ and $TH_{2,2}$ in state $l=2$. The mean probabilities of holding buffer content within the ranges 0-$TH_{l,1}$, $TH_{l,1}$-$TH_{l,2}$, and $TH_{l,2}$- L_iB_i, i.e., $h_{l,0}$, $h_{l,1}$, and $h_{l,2}$ respectively, are defined as

$$\begin{aligned} h_{l,0} &= \sum_{j=0}^{TH_{l,1}-1} \psi'_j\Big|_{K_i,\rho_i,l} \\ h_{l,1} &= \sum_{j=TH_{l,1}}^{TH_{l,2}} \psi'_j\Big|_{K_i,\rho_i,l} \quad and \\ h_{l,2} &= \sum_{j=TH_{l,2}}^{L_iB_i} \psi'_j\Big|_{K_i,\rho_i,l} \end{aligned} \tag{5.2}$$

where $\psi'_j\big|_{K_i,\rho_i,l}$ is defined in equation (5.1).

The ATM multiplexer node moves from a global congestion state l to a new state 0, or 1, or 2, with the state transition probability of $H_{l,0}$, or $H_{l,1}$, or $H_{l,2}$ respectively, as shown in

Fig. 5.2. Assuming v nodes are informed of the global congestion status, the probability that none of the v ATM multiplexer nodes are experiencing any global congestion, i.e., accounting for only *local congestion*, is

$$H_{l,0} = (h_{l,0})^{v} \ . \tag{5.3}$$

The v ATM multiplexer nodes, which are informed of the global congestion status, will prevent congestion by adjusting the buffer thresholds accordingly. As v increases towards c, more ATM multiplexer nodes will be aware of the global congestion status, and hence $H_{l,0}$ will further decrease. The probability that exactly one of the ATM multiplexer node is experiencing *light global congestion* ($l=1$) while all other nodes are experiencing only local congestion, is

$$H_{l,1} = v\, h_{l,1}\ \ h_{l,0}^{v-1} \ . \tag{5.4}$$

The probability that two or more ATM multiplexer nodes are experiencing *light global congestion ($l=1$), or* one or more ATM multiplexer nodes are experiencing *severe global congestion ($l=2$),* is

$$H_{l,2} = 1 - H_{l,0} - H_{l,1} \quad . \tag{5.5}$$

Applying the fundamental law of flow conservation, wherein the probability that an ATM multiplexer assumes global congestion state j under equilibrium condition, is equal to the sum of various probabilities of arriving the state j from all possible states times the respective state transition probabilities. Hence, we obtain

$$\chi_j = \sum_{l=0}^{2} H_{lj}\chi_l \qquad where \;\; 0 \le j \le 2 \; . \tag{5.6}$$

Also, as the sum of all state probabilities is one, giving rise to

$$\sum_{i=0}^{2} \chi_i \;\; = \;\; 1 \; . \tag{5.7}$$

We obtain χ_0, χ_1, χ_2 by solving the above equations as

$$\chi_0 = \frac{H_{2,1}(H_{1,0} - H_{2,0}) - H_{2,0}(H_{1,1} - 1 - H_{2,1})}{(H_{0,0} - 1 - H_{2,0})(H_{1,1} - 1 - H_{2,1}) - (H_{0,1} - H_{2,1})(H_{1,0} - H_{2,0})}$$

$$\chi_1 = \frac{H_{2,1}(H_{0,0} - 1 - H_{2,0}) - H_{2,0}(H_{0,1} - H_{2,1})}{(H_{1,0} - H_{2,0})(H_{0,1} - H_{2,1}) - (H_{1,1} - 1 - H_{2,1})(H_{0,0} - 1 - H_{2,0})} \qquad and \tag{5.8}$$

$$\chi_2 = 1 - \chi_0 - \chi_1 \quad .$$

Furthermore, simple analytical formulas are proposed using equation (5.8), for the steady-state global congestion performance characteristics of a service-class *i* with LB, VLB, mLB, and mVLB schemes. An ATM node shall be declared *congested* when its buffer occupancy increases beyond a threshold. The threshold to declare congestion to the local transport users, varies with the global congestion state *l*. The ATM multiplexer node buffer is analyzed by the set of equations (4.1)-(4.16), and (4.28)-(4.29) in a global congestion state *l*. As the ATM multiplexer node moves to higher global congestion states (i.e., *l*=1, or *l*=2), the congestion threshold value for local transport users, shall be decreased to allocate larger buffer space for the transit traffic. Hence a set of congestion threshold values, i.e., *Th(l)* (say 0.6, 0.4 and 0.2 for *l=0, 1 and 2*, respectively), are used to detennine the probability of congestion during global congestion states. The probability

to declare congestion in a given state l with active bursts k_i (where $0 \le k_i \le K_i$), i.e., $P_c|_{k_i,\rho_i,l}$ can be evaluated as

$$P_c|_{k_i,\rho_i,l} = \sum_{j=Th(l).L_iB_i}^{L_iB_i} \psi(j)|_{k_i,\rho_i,l} \tag{5.9}$$

where $\psi(j)$ at a given number of active bursts k_i, is the probability of staying at a buffer state j. Only k_i active bursts are generated from K_i active calls, at any time. Hence, the mean probability of congestion for K_i number of active calls at a global congestion state l is

$$\overline{P}_{cR}|_{K_i,\rho_i,l} = \sum_{k_i=0}^{K_i} P_c|_{k_i,\rho_i,l}\ P(k_i) \tag{5.10}$$

where $P(k_i)$ has been defined in equation (4.9). Hence, the mean probability of congestion, averaged over all global congestion states, is

$$\overline{P}_{cR}|_{K_i,\rho_i} = \sum_{l=0}^{2} \overline{P}_{cR}|_{K_i,\rho_i,l}\ \chi_l\ . \tag{5.11}$$

The ATM multiplexer node buffer content for a given number of active bursts k_i, averaged over all possible buffer states of Fig. 4.6, can be defined as

$$\psi^{-}|_{k_i,\rho_i,l} = \sum_{j=0}^{L_iB_i} j\ \psi(j)|_{k_i,\rho_i,l} \tag{5.12}$$

where L_iB_i is the total buffer space, which is defined in equation (4.5). Averaging over the distribution of active bursts k_i, i.e., $P(k_i)$, we obtain the ATM multiplexer node mean buffer content of traffic class i, for a given number of active calls K_i, in a global congestion state l, as

$$\overline{\Psi}\big|_{K_i,\rho_{i,l}} = \sum_{k_i=0}^{K_i} \Psi^{-}\big|_{k_i,\rho_{i,l}} \quad P(k_i) \,. \tag{5.13}$$

Then, the mean buffer content, averaged over all global congestion states, is

$$\overline{\Psi}\big|_{K_i,\rho_i} = \sum_{l=0}^{2} \overline{\Psi}\big|_{K_i,\rho_{i,l}} \chi_l \,. \tag{5.14}$$

The probability that there exist j consecutive independent frames in congestion, is

$$\begin{aligned} B_j &= \overline{P}_{cg}^{\,j}(1-\overline{P}_{cg}) \qquad when \quad v=0 \\ &= P_g^j(1-P_g) \qquad when \quad v\neq 0 \end{aligned} \tag{5.15}$$

where $P_g = \chi_1 + \chi_2$ using equation (5.8). The congestion duration (CD) is defined as the probability that j consecutive frames are lost due to congestion. Then the mean CD can be derived as

$$\begin{aligned} CD = T_F \sum_{j=0}^{\infty} jB_j &= \frac{T_F\,\overline{P}_{cg}}{1-\overline{P}_{cg}} \qquad when \quad v=0 \\ &= \frac{T_F\,P_g}{1-P_g} \qquad when \quad v\neq 0 \end{aligned} \tag{5.16}$$

$$where \quad frame\ size \quad T_F = \sum_{i=1}^{S} N_{si} \quad slots \,.$$

Equation (5.16) is a great simplification over [KAM-89]. It can be inferred that when $CD > \tau$ will lead c-v to zero, i.e., all nodes are informed of the congestion status. But when $CD < \tau$, then there exist two groups, i.e., congestion status informed (v) and uninformed (c-v) over the VC.

The probability that an active user generating class i traffic in a global congestion state l, which is accepted for transmission (i.e., introduced to the ATM multiplexer buffer) is given by

$$\xi\big|_{k_i,\rho_i,l} = [(1-P_{vi}\big|_{\rho_i}) + P_{vi}\big|_{\rho_i}\ (1-P_c\big|_{k_i,\rho_i,l})]\frac{\rho_i}{K_i} \quad \text{with VLB}$$

$$= (1-P_{vi}\big|_{\rho_i})\frac{\rho_i}{K_i} \quad \text{with LB} \qquad (5.17)$$

$$= \frac{\rho_i}{K_i} \quad \text{without congestion control .}$$

In case of VLB scheme, a cell is introduced to the ATM buffer if it is generated without violating the QOS, i.e., $(1\text{-}P_{vi})$, or when it violates the QOS during the periods of no congestion, i.e., $P_{vi}(1\text{-}P_c)$. The LB scheme drops all violating cells irrespective of the congestion status of the ATM multiplexer node. However, congestion develops rapidly when neither of these cell-discard schemes is used. The mean probability of generating a cell in a global congestion state l, from an active user with the overall burst condition, is given by

$$\bar{\xi}\big|_{K_i,\rho_i,l} = \sum_{k_i=0}^{K_i} \xi\big|_{k_i,\rho_i,l}\ P(k_i)\ . \qquad (5.18)$$

Then, the mean probability of generating a cell, averaged over all global congestion states, is

$$\bar{\xi}\big|_{K_i,\rho_i} = \sum_{l=0}^{2} \bar{\xi}\big|_{K_i,\rho_i,l}\ \chi_l\ . \qquad (5.19)$$

The probability that an active user generates a cell in a global congestion state l, which gets discarded, is given by

$$\Delta\big|_{k_i,\rho_i,l} = P_{vi}\big|_{\rho_i} \cdot P_c\big|_{k_i,\rho_i,l} \frac{\rho_i}{K_i} \quad \text{with VLB}$$

$$= P_{vi}\big|_{\rho_i} \cdot \frac{\rho_i}{K_i} \quad \text{with LB} \tag{5.20}$$

$$= 0 \quad \text{without congestion control .}$$

Equation (5.20) depicts that VLB scheme discards all violating cells only when the ATM node experiences local congestion with a probability $P_{vi}P_c\big|_{ki}$, whereas the LB scheme discards all traffic which violates the QOS with a probability P_{vi}. The mean probability of discarding a cell in a global congestion state l, averaged over all possible number of active bursts k_i, is

$$\overline{\Delta}\big|_{K_i,\rho_i,l} = \sum_{k_i=0}^{K_i} \Delta\big|_{k_i,\rho_i,l} P(k_i) \ . \tag{5.21}$$

Then, the mean probability of discarding a cell, averaged over all global congestion states, is

$$\overline{\Delta}\big|_{Ki,\rho i} = \sum_{l=0}^{2} \overline{\Delta}\big|_{Ki,\rho i,l} \chi_l \ . \tag{5.22}$$

Another important congestion performance criterion is the ATM multiplexer node buffer-overflow. The probability of ATM buffer overflow, which is defined as the buffer content beyond a predefined threshold, is

$$P_{overflow}\big|_{k_i,\rho_i,l} = \sum_{j=ofl.LB_i}^{LB_i} \psi(j)\big|_{k_i,\rho_i,l} \tag{5.23}$$

where ofl is overflow factor or overflow threshold (say $0.9L_iB_i$), to declare buffer-overflow when the buffer content reaches beyond this threshold. The mean probability of ATM buffer overflow in a global congestion state l, is

$$\overline{P}_{overflow}\Big|_{K_i,\rho_{i,l}} = \sum_{k_i=0}^{K_i} P_{overflow}\Big|_{k_i,\rho_{i,l}} \; P(k) \; . \tag{5.24}$$

Then, the mean probability of ATM buffer overflow, averaged over all global congestion states, is

$$\overline{P}_{overflow}\Big|_{K_i,\rho_i} = \sum_{l=0}^{2} \overline{P}_{overflow}\Big|_{K_i,\rho_{i,l}} \chi_l \; . \tag{5.25}$$

The above analysis is applied to video/voice, and data traffic by using P_a as in equations (4.6) and (4.29), respectively.

5.4 Global congestion performance results

The congestion performance characteristics, i.e., probabilities of ATM multiplexer node congestion (P_c), cell generation (ξ), cell discard (Δ), buffer content (ψ), congestion duration (CD), and buffer overflow ($P_{overflow}$) were obtained using the system of equations (5.1)-(5.25). The computation of congestion performance criteria is detailed in flowchart of Fig. 4.8. Once the static system parameters such as K_i, K_{av}, n_{ui}, R_{pi}, S, ll, q_i, N_{fmax}, etc., are known or configured, the related static parameters such as N_{si}, L_iB_i can be obtained from equations (4.3)-(4.5). The congestion measurement thresholds, i.e., $Th(l)$ to determine P_c using equation (5.9), are assumed to be 0.6, 0.4, and 0.2 for global congestion states l=0, 1, and 2, respectively. The first level thresholds $TH_{l,1}$ are assumed to be 0.6, 0.4 and 0.2 for l=0, 1, and 2, respectively. The second level thresholds $TH_{l,2}$ are assumed to be 0.8, 0.6 and 0.4 for l=0, 1, and 2, respectively. The performance parameters, namely buffer overflow threshold (ofl) and convergence of P_c value in terms of percentage deviation from its value in the previous iteration ($\%cnv$), are assumed to be

90% and 5%, respectively. As the averaging procedure needs all possible traffic generation conditions, the factor $\beta_i/(\alpha_i+\beta_i)$, i.e., P_a in equation (4.6) is varied from 0 to 1, and the same effect is reflected on the traffic intensity (ρ_i) of equation (4.11). As the probability of violation depends on traffic intensity, the set of $P_{vi}|_{Ki}$ are obtained for measurement traffic intensities (ρ_i). Moreover, the active bursts k_i are also varied from 0 to K_i to reflect all possible occurrences for each value of ρ_i. The system starts with initial state $l=0$, and the buffer state diagram (Fig. 4.6) of the ATM multiplexer node is solved to obtain $\Psi(j)|_{ki}$ where $0 \le j \le L_iB_i$, and then a new values of $P_c|_{ki}$ is arrived. With this value of $P_c|_{l,ki}$, we repeat the same procedure to evaluate new values of $\xi|_{l,ki}$, $\theta|_{l,ki}$ and $\Psi'(0)$-$\Psi(L_iB_i)|_{l,ki}$ and a new value of $P_c|_{l,ki}$ until it converges with its previous value at an accuracy of $\%cnv$ (say 5%) within the same l. By repeating the same procedure for all possible active bursts, i.e., $0 \le k_i \le K_i$, the average buffer distribution as specified in equation (5.1), is found in the relevant l, and the global state transition characteristics are obtained from equations (5.2)-(5.5). The global congestion state diagram (Fig. 5.2) is then solved from equation (5.6) to obtain global state probabilities, i.e., χ_0, χ_1, and χ_2, from equation (5.8). By repeating the same procedure for all possible traffic generation, i.e., $0 \le P_a \le 1$. The mean congestion performance characteristics, i.e., $\overline{P}_c|_{Ki}$, $\overline{\xi}|_{Ki}$, $\overline{\Delta}|_{Ki}$, $\overline{\Psi}|_{Ki}$, $\overline{P}_{overflow}|_{Ki}$ and CD, are obtained by averaging over the distribution of $P(k_i)$ (where $0 \le k_i \le K_i$), and by averaging over all global congestion states ($l=0$, 1, and 2).

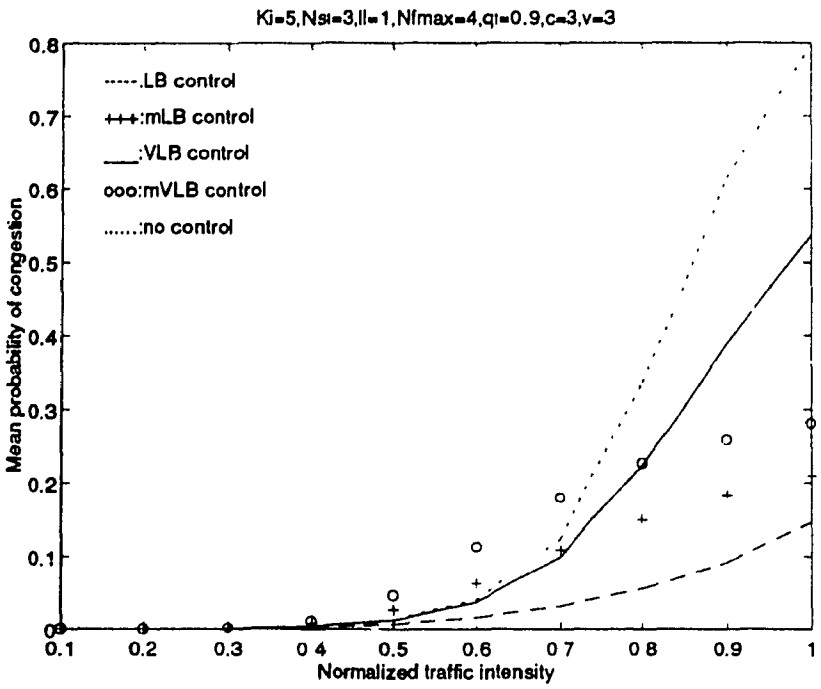

Fig. 5.3 (a) v = 3, N_{si}=3;

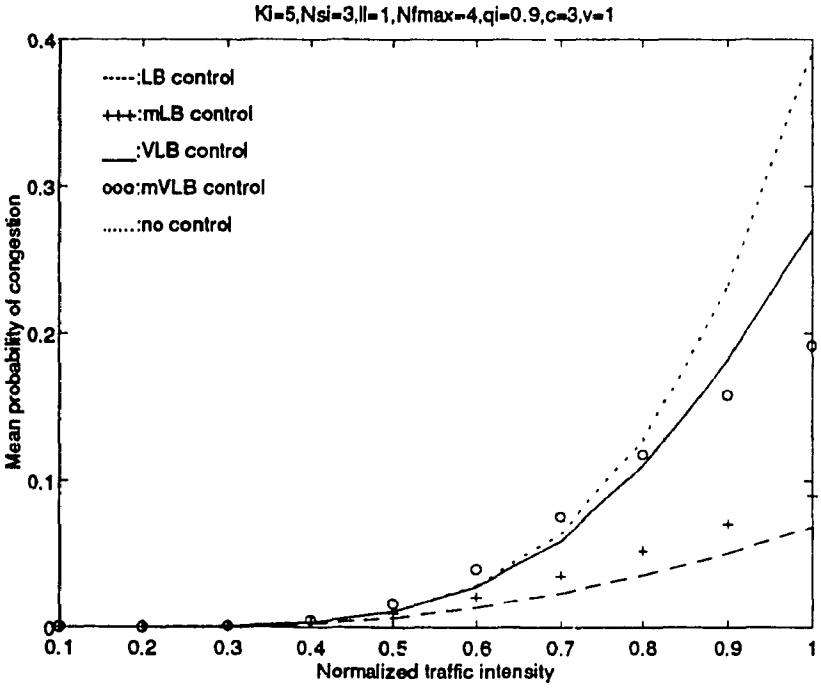

Fig. 5.3 (b) v = 1, N_{si}=3;

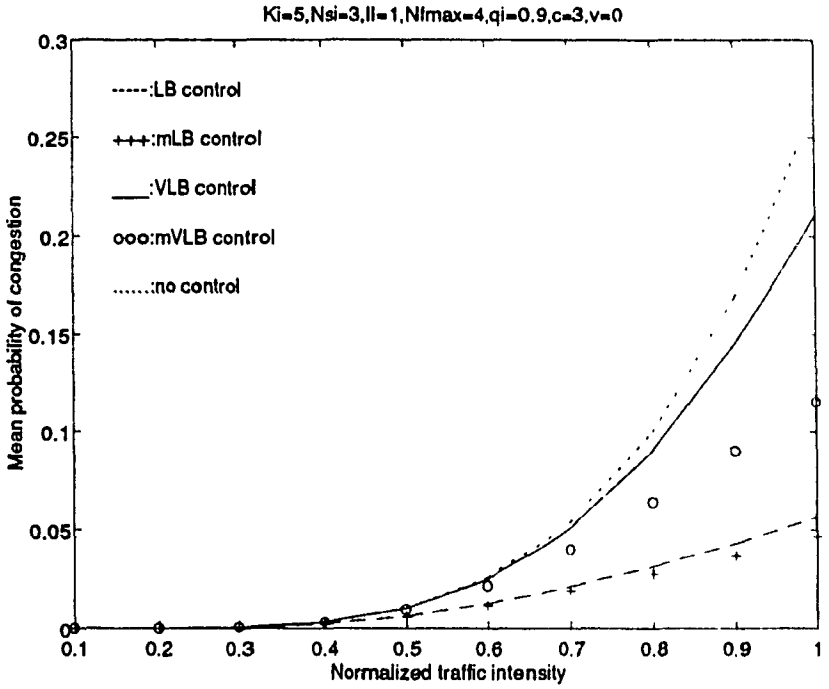

Fig. 5.3 (c) v = 0, N_{si}=3;

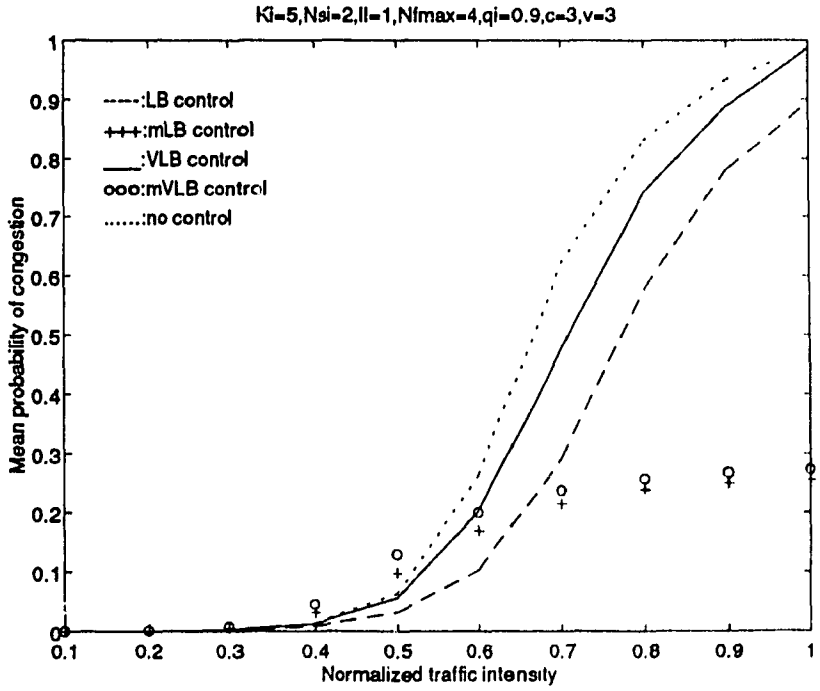

Fig. 5.3 (d) v = 3, N_{si}=2;

Fig. 5.3. Mean probability of global congestion.

The ATM multiplexer global congestion is obtained as a function of traffic intensity in Fig. 5.3. The results in Figs 5.3 (a), (b), and (c) were obtained for an N_{si} value of 3, with different congestion communication status, i.e., v=0, 1, and 3. The results in Fig. 5.3 were obtained for an N_{si} value of 2 for comparison. It is obvious from the results that the ATM multiplexer global congestion increases with the traffic intensity. It can be observed that the congestion develops at a slower rate for LB scheme, as it discards cells faster than VLB, to control congestion. By controlling the active bursts at the transport users, the mLB and mVLB schemes further reduce the congestion than the respective LB and VLB schemes. Moreover, the mean probability of congestion reduces, when the service rate, i.e., N_{si} is increased, which can be observed from Figs 5.3 (a) and (d) for the same conditions of global congestion communication status (v=3, c=3, i.e., all nodes are informed of congestion status). The results in Figs 5.3 (a), (b), and (c) were obtained when the ATM multiplexer node is fully informed, partially informed, and unaware of congestion status, respectively. The congestion status in Fig. 5.3 (a) is observed to be higher than (b) and (c), because of the fact that the ATM multiplexer node reacts to the developing congestion of the neighboring nodes, in order to reduce the traffic from the local transport users, if the overall (global) congestion reaches beyond a threshold. The results in Fig. 5.3 (c) are identical to the local congestion results of Fig. 4.9 (b) of chapter 4, because the ATM multiplexer node assumes that the neighboring ATM multiplexer nodes are not experiencing any congestion. It can also be observed from the results that the ATM multiplexer node congestion develops at a higher rate when the control schemes are not employed.

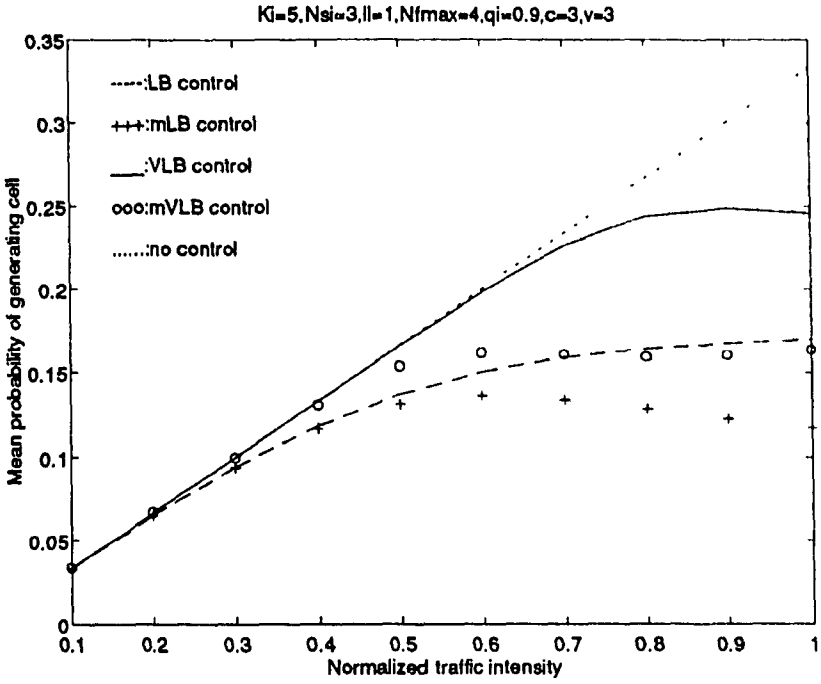

Fig. 5.4 (a) v=3, N_{si}=3;

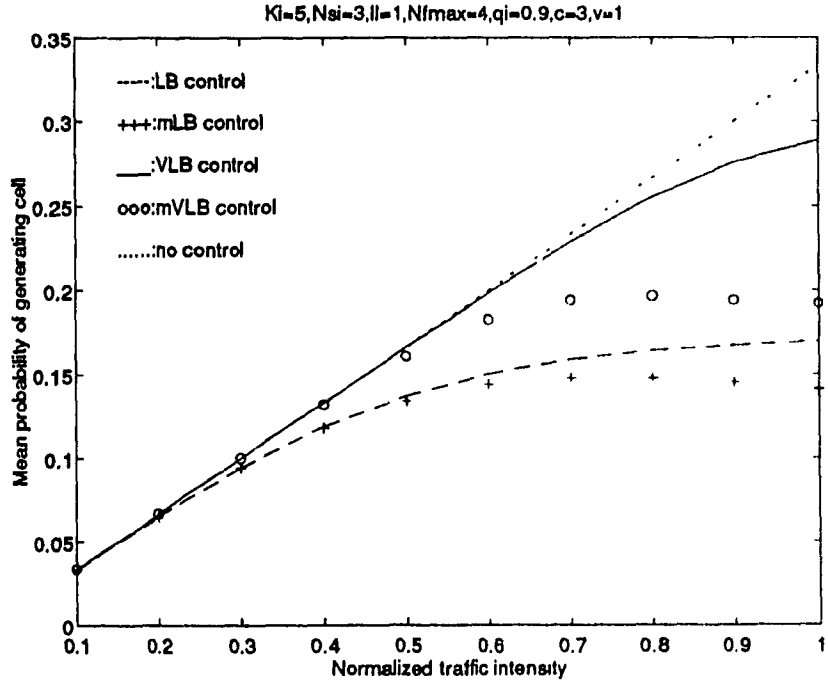

Fig. 5.4 (b) v = 1, N_{si}=3;

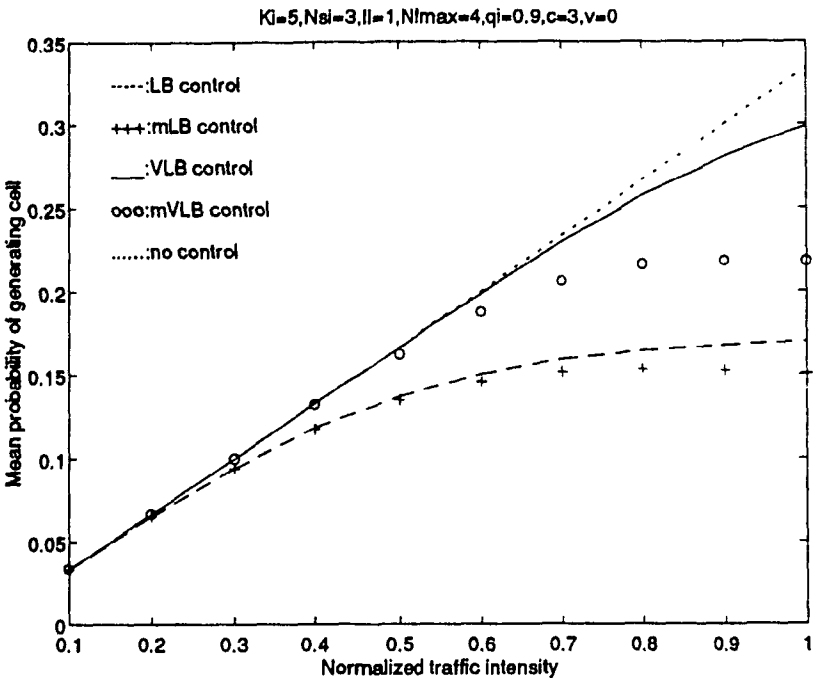

Fig. 5.4 (c) v = 0, N_{si}=3;

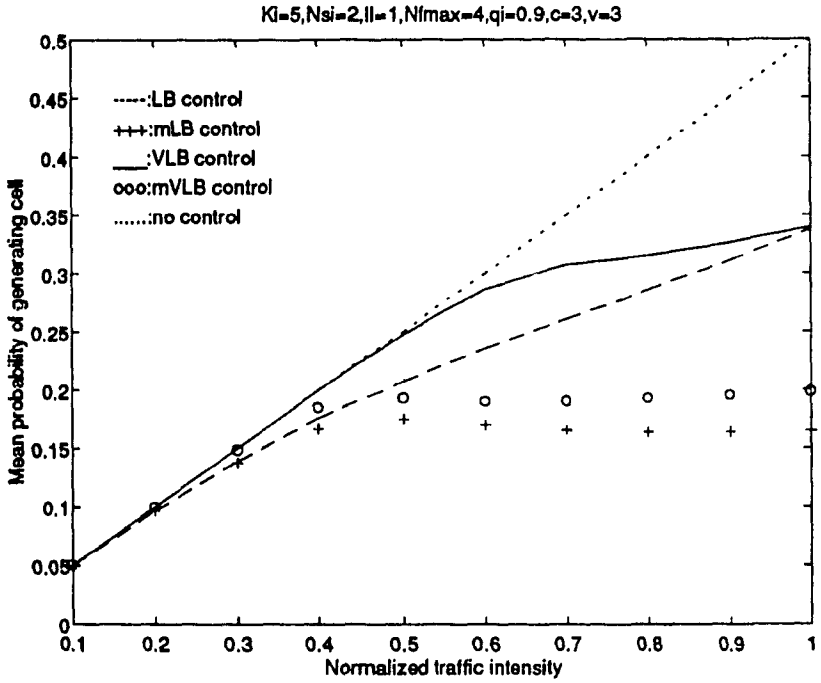

Fig. 5.4 (d) v = 3, N_{si}=2;

Fig. 5.4. Mean probability of generating a cell.

The mean probability of generating a cell is obtained as a function of traffic intensity in Fig. 5.4. The results in Figs 5.4 (a), (b), and (c) were obtained for an N_{st} value of 3, with different congestion communication status, i.e., v=0, 1, and 3. The results in Fig. 5.4 were obtained for an N_{st} value of 2 for comparison. The mean probability of generating a cell increases linearly when there is no congestion control. However, the LB and VLB schemes drop the traffic based on P_{vt} and P_t, thereby reducing the probability of generating a cell during violation and congestion periods. The mLB and mVLB schemes control the active bursts based on the instantaneous global congestion level, thereby reducing the probability of generating a cell furthermore, by leaving the traffic at the transport users instead of dropping at the ATM layer, as observed in Fig. 5.4. Moreover, higher capacity transport users can be handled when N_{st} is increased, as more number of cells are serviced, which can be observed from Figs 5.4 (a) and (d) for the same conditions of global congestion communication (v=3, c=3, i.e., all nodes are informed of congestion status). The results in Figs 5.4 (a), (b), and (c) were obtained when the ATM multiplexer node is fully informed, partially informed, and unaware of congestion status, respectively. It can be observed that cell-generation is unaffected, when a congestion control scheme is not used. However, it can be observed that the probability of generating a cell reduces as the ATM multiplexer node is aware of congestion status of the neighboring ATM multiplexer nodes, by accounting the transit traffic into the computation of global congestion. The results in Fig. 5.4 (c) are identical to the cell generation results of Fig. 4.10 (b) of chapter 4, because the ATM multiplexer node assumes that the neighboring ATM multiplexer nodes are not experiencing any congestion.

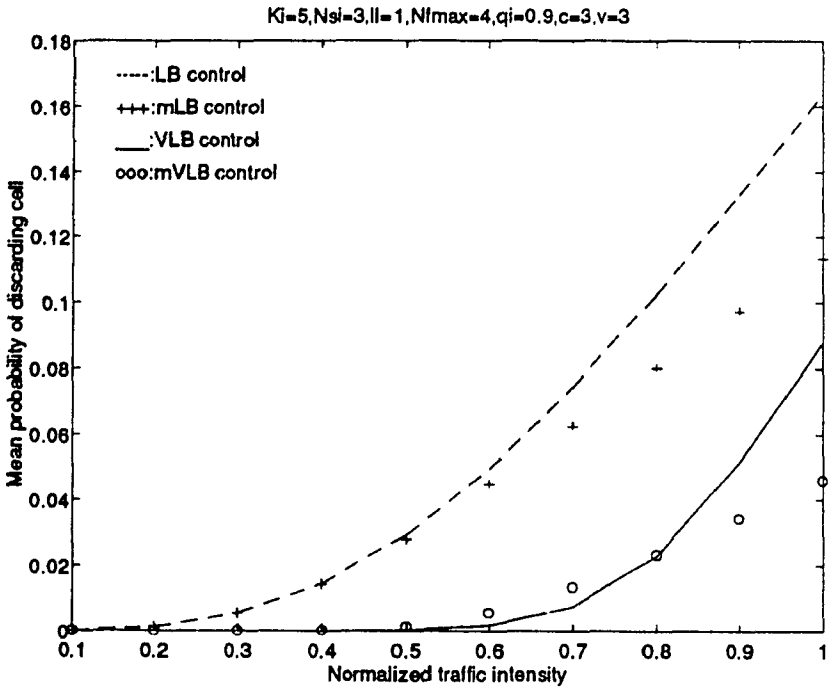

Fig. 5.5 (a) v = 3, N_{si}=3;

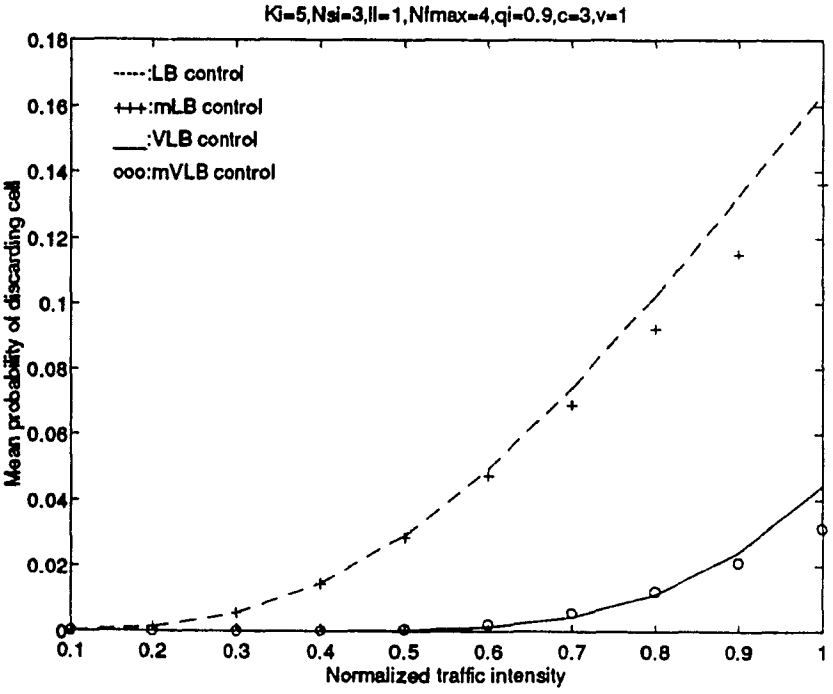

Fig. 5.5 (b) v = 1, N_{si}=3;

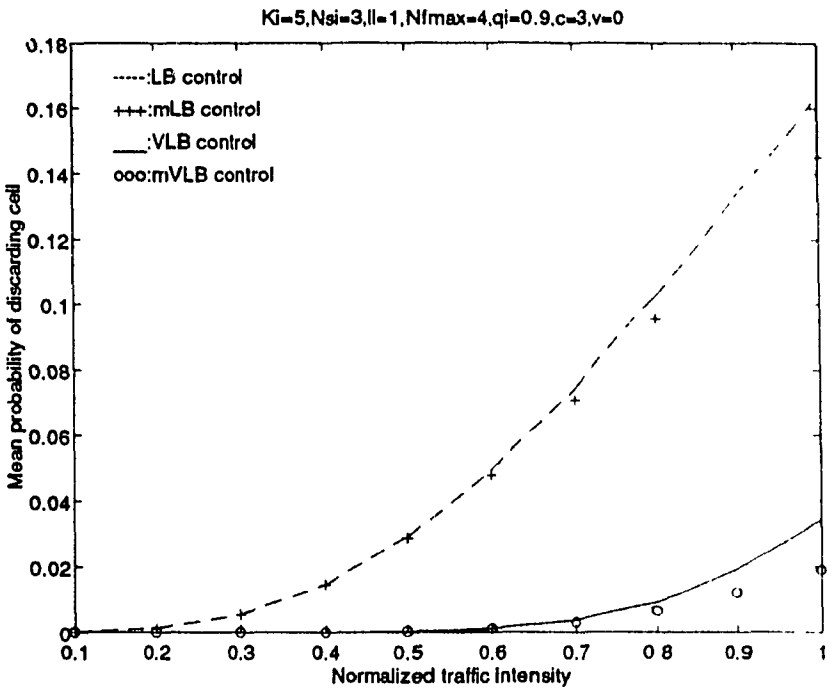

Fig. 5.5 (c) v = 0, N_{si}=3;

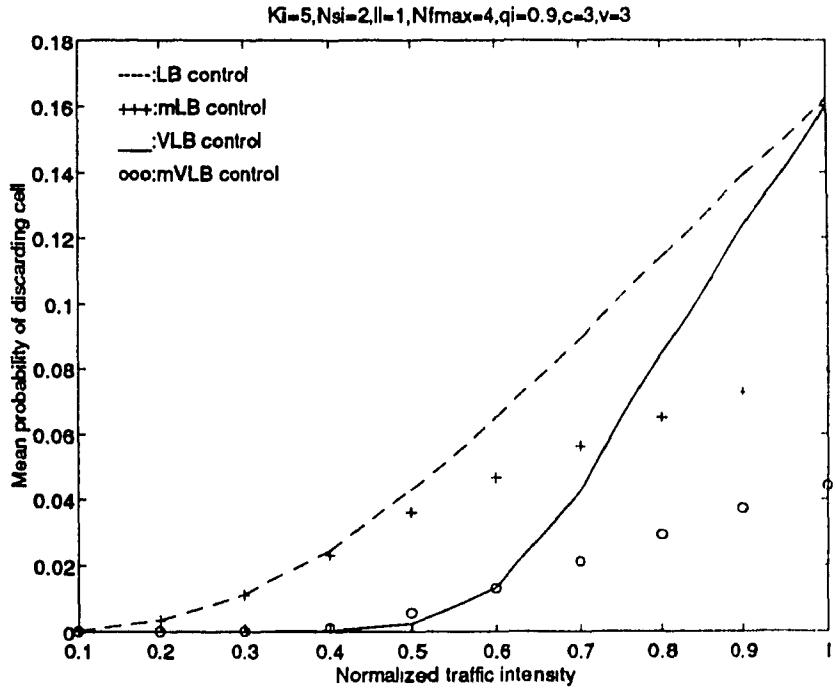

Fig. 5.5 (d) v = 3, N_{si}=2;

Fig. 5.5. Mean probability of discarding a cell.

The mean probability of discarding a cell is obtained as a function of traffic intensity in Fig. 5.5. The results in Figs 5.5 (a), (b), and (c) were obtained for an N_{si} value of 3, with different congestion communication status, i.e., v=0, 1, and 3. The results in Fig. 5.5 were obtained for an N_{si} value of 2 for comparison. It can be noted that the cells are not discarded when a congestion control scheme is not used. Hence, we compare only the LB, VLB, mLB, and mVLB congestion control schemes in Fig. 5.5. Since the LB cell discard scheme drops all cells which violate the QOS, and since the P_{vi} increases with ρ_i for the same QOS, the probability of discarding a cell increases with traffic intensity (ρ_i). In case of VLB cell discard scheme, violating cells are not dropped unless local congestion is declared, and hence the probability of discarding a cell is lesser than in the case of LB, for the same traffic intensity, which can be observed in Fig. 5.5. The mLB and mVLB schemes reduce the necessity to drop the traffic at the ATM layer, by controlling the active bursts at the transport users to reduce the instantaneous congestion. It can also be observed in Figs 5.5 (a) and (d) that the probability of discarding a cell decreases as the service rate (N_{si}) increases for the same conditions of global congestion communication (v=3, c=3, i.e., all nodes are informed of congestion status). The results in Figs 5.5 (a), (b), and (c) were obtained when the ATM multiplexer node is fully informed, partially informed, and unaware of congestion status, respectively. It can be observed that more cells are discarded when the ATM multiplexer node is aware of congestion status of the neighboring ATM multiplexer nodes. The results in Fig. 5.5 (c) are identical to the cell

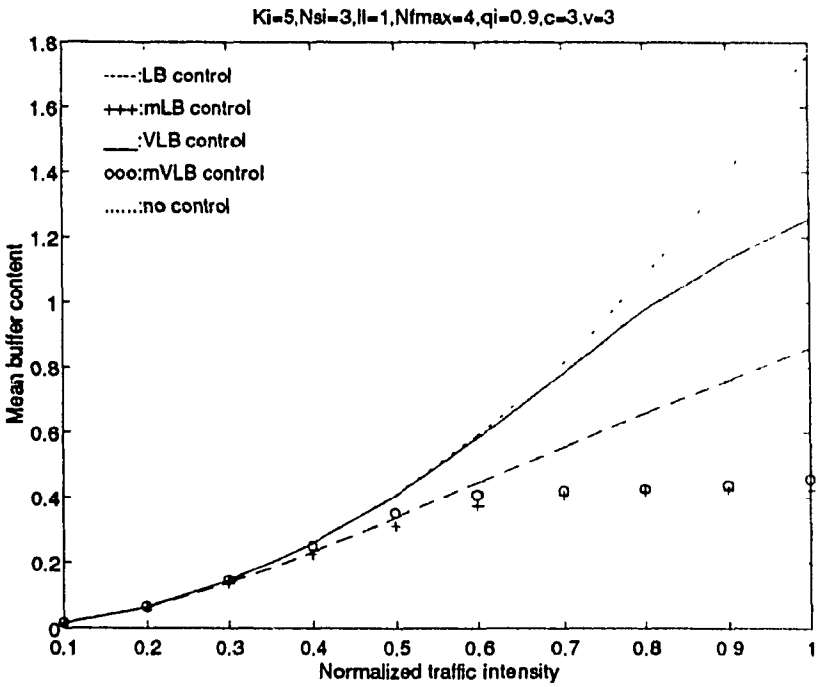

Fig. 5.6 (a) v = 3, N_{si}=3;

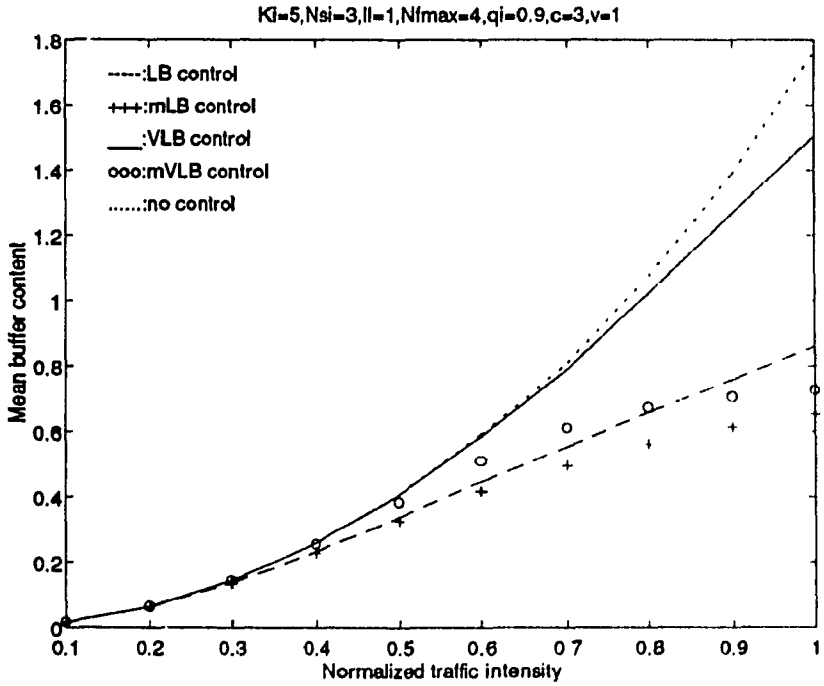

Fig. 5.6 (b) v = 1, N_{si}=3;

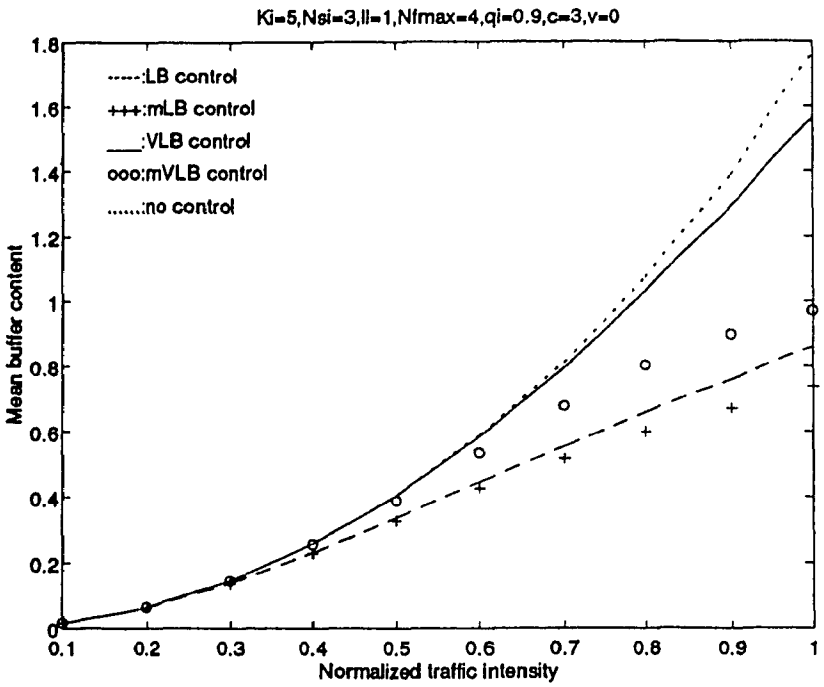

Fig. 5.6 (c) v = 0, N_{si}=3;

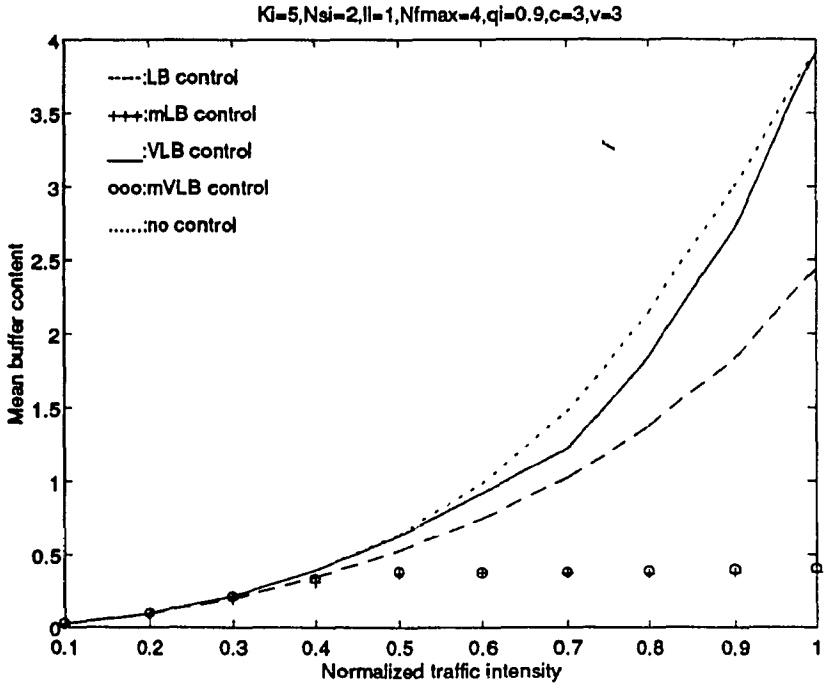

Fig. 5.6 (d) v = 3, N_{si}=2;

Fig. 5.6. Mean buffer content.

discard results of Fig. 4.11 (b) of chapter 4, because the ATM multiplexer node assumes that the neighboring ATM multiplexer nodes are not experiencing any congestion.

The mean buffer content of the ATM multiplexer node is obtained as a function of traffic intensity in Fig. 5.6. The results in Figs 5.6 (a), (b), and (c) were obtained for an N_{si} value of 3, with different congestion communication status, i.e., v=0, 1, and 3. The results in Fig. 5.6 were obtained for an N_{si} value of 2 for comparison. The mean buffer content increases as the traffic intensity increases. As the active burst control reduces the traffic from the local transport users during congestion periods, the mLB and mVLB schemes reduce the buffer content, compared to the respective LB and VLB schemes, as observed in Fig. 5.6. It can also be observed from Figs 5.6 (a) and (d) that mean buffer content decreases as N_{Si} increases, for the same conditions of global congestion communication (v=3, c=3, i.e., all nodes are informed of congestion status). The results were obtained when the ATM multiplexer node is fully informed, partially informed, and unaware of congestion status, in Figs 5.6 (a), (b), and (c), respectively. It can be observed that buffer content rises less rapidly for LB and VLB cell discard schemes when the ATM multiplexer node is aware of congestion status of the neighboring ATM multiplexer nodes, to prevent global congestion. The results in Fig. 5.6 (c) are identical to the buffer content results of Fig. 4.12 (b) of chapter 4, because the ATM multiplexer node assumes that the neighboring ATM multiplexer nodes are not experiencing any congestion. The mean buffer content is observed to be same in (a), (b), and (c), when there is no congestion control in the ATM multiplexer node.

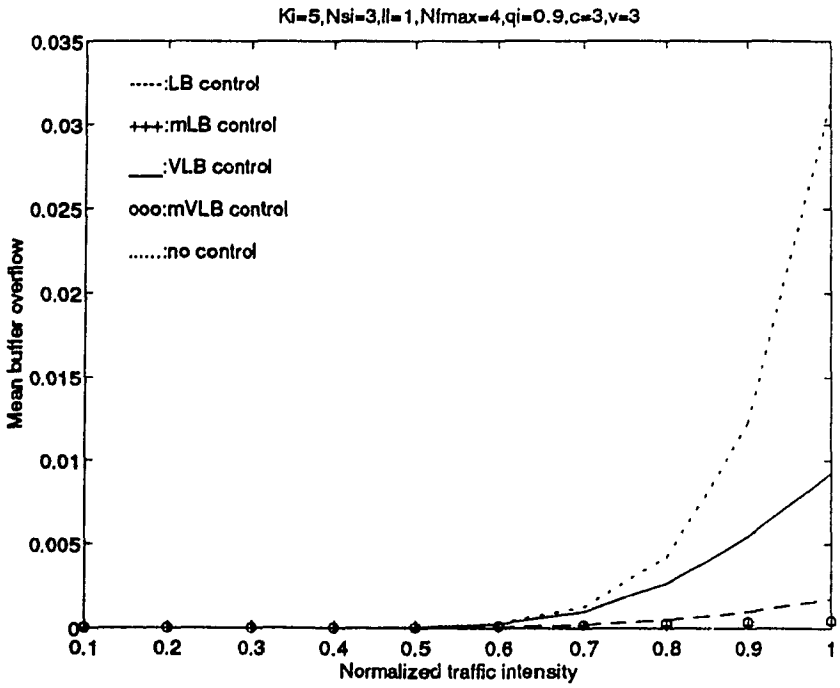

Fig. 5.7 (a) v = 3, N_{si}=3;

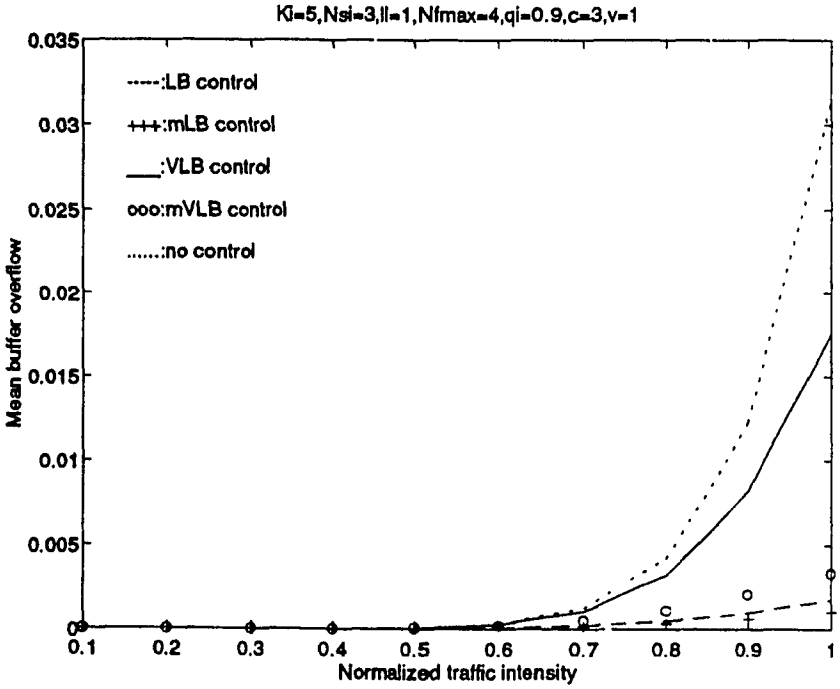

Fig. 5.7 (b) v = 1, N_{si}=3;

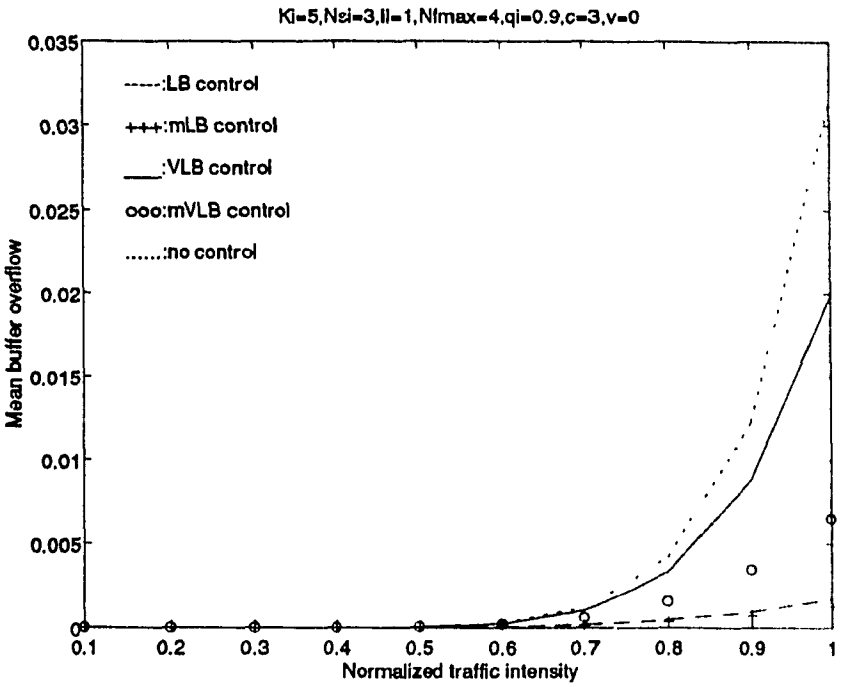

Fig. 5.7 (c) v = 0, N_{si}=3;

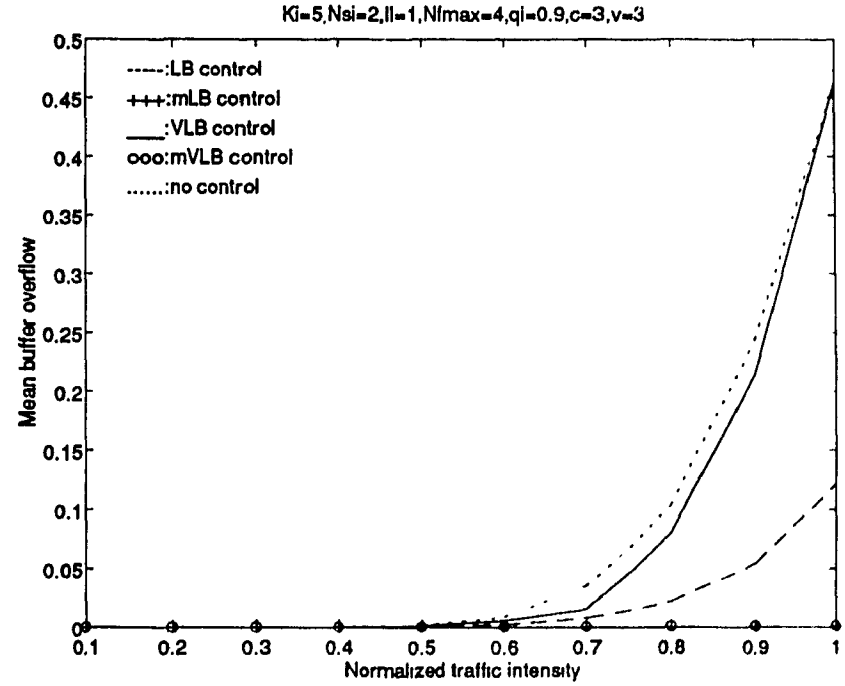

Fig. 5.7 (d) v = 3, N_{si}=2;

Fig. 5.7. Mean buffer overflow.

The mean buffer overflow of the ATM multiplexer node is obtained as a function of traffic intensity in Fig. 5.7. The results were obtained at an N_{si} value of 3, with different congestion communication status, i.e., v=0, 1, and 3, in Figs 5.7 (a), (b), and (c), respectively. The results in Fig. 5.7 were obtained for an N_{si} value of 2 for comparison. As the buffer content rises beyond an overflow threshold (*ofl*), the buffer overflow also rises leading to cell-loss. Hence, the results show no buffer overflow (i.e., zero) when the traffic intensities are low. As the LB scheme drops all violating cells, it contributes to less buffer overflow, compared to VLB scheme. The mLB and mVLB schemes reduce the buffer overflow for the same traffic intensity, when compared to the respective LB and VLB schemes. It can be observed from Figs 5.7 (a) and (d) that the buffer overflow reduces if the service rate (N_{si}) is increased, for the same conditions of global congestion communication (v=3, c=3, i.e., all nodes are informed of congestion status). The results in were obtained when the ATM multiplexer node is fully informed, partially informed, and unaware of congestion status in Figs 5.7 (a), (b), and (c), respectively. It can be observed that buffer overflow rises less rapidly for LB and VLB cell discard schemes when the ATM multiplexer node is aware of congestion status of the neighboring ATM multiplexer nodes, to prevent global congestion. The results in Fig. 5.7 (c) are identical to the buffer overflow results of Fig. 4.13 (b) of chapter 4, because the ATM multiplexer node assumes that the neighboring ATM multiplexer nodes are not experiencing any congestion. The mean buffer overflow is observed to be same in (a), (b), and (c), when there is no congestion control in the ATM multiplexer node.

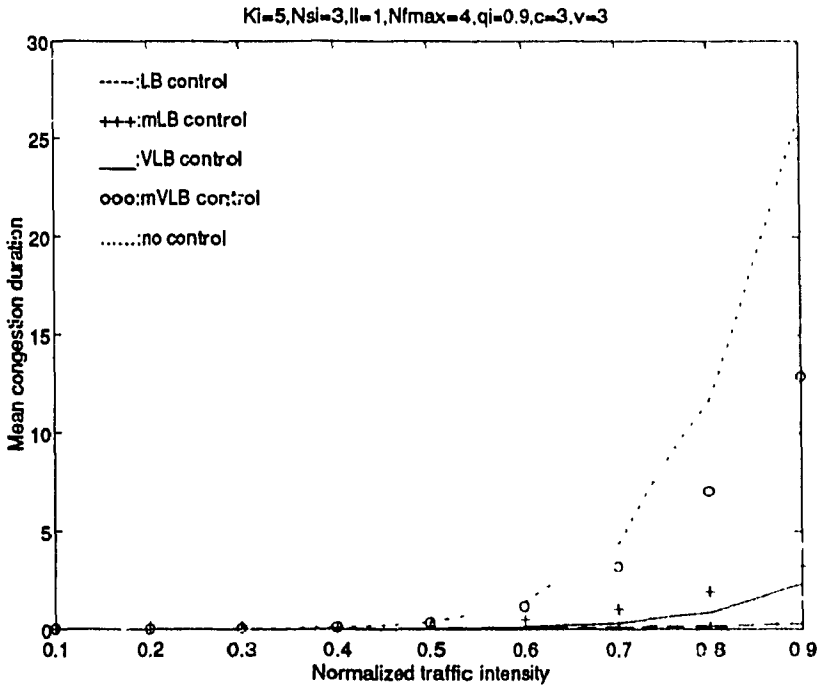

Fig. 5.8 (a) v = 3, N_{si}=3;

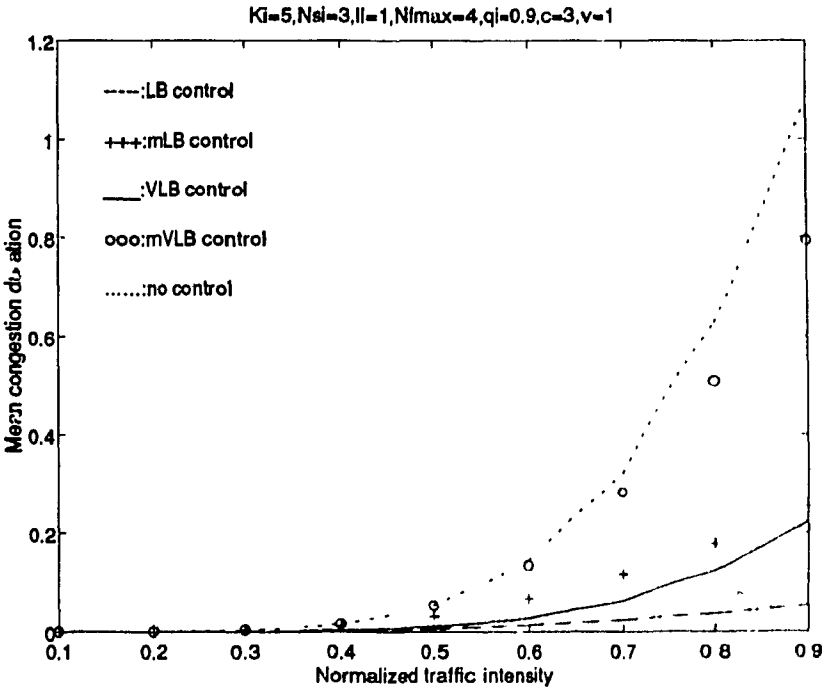

Fig. 5.8 (b) v = 1, N_{si}=3;

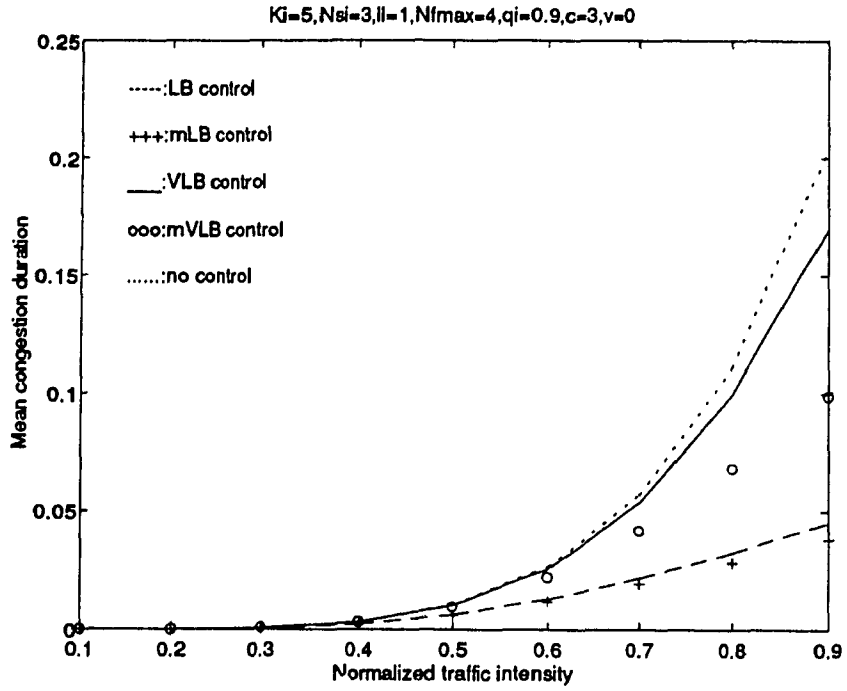

Fig. 5.8 (c) v = 0, N_{si}=3;

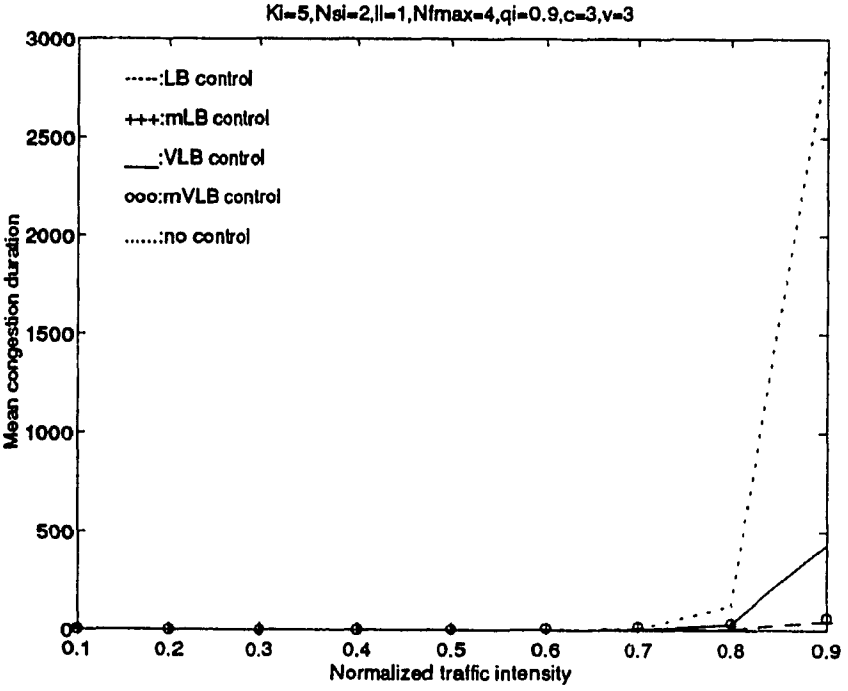

Fig. 5.8 (d) v = 3, N_{si}=2;

Fig. 5.8. Mean congestion duration.

The mean congestion duration is obtained as a function of traffic intensity in Fig. 5.8. The results were obtained at an N_{si} value of 3, with different congestion communication status, i.e., v=0, 1, and 3, in Figs 5.8 (a), (b), and (c) respectively. The results in Fig. 5.8 were obtained for an N_{si} value of 2 for comparison. As the buffer content rises beyond a threshold, the probability of congestion increases as observed in Fig. 5.3. Moreover, the increase in traffic intensity from the local transport users and the congestion communication from other ATM multiplexer nodes, contribute to the probability of congestion, which in turn increases the congestion duration. As the LB scheme drops all violating cells, it contributes to less congestion, compared to VLB scheme. The mLB and mVLB schemes reduce the ATM multiplexer global congestion for the same traffic intensity, when compared to the respective LB and VLB schemes. Hence, the congestion duration is obtained to be less for mLB, and mVLB schemes compared to LB, and VLB schemes in Fig. 5.8 (c), when v=0, i.e., ATM multiplexer node is unaware of the congestion status of other ATM multiplexer nodes. It can be observed from Figs 5.8 (a), (b) and (c) that the transit traffic is prioritized over local traffic (when $v \neq 0$), which increased *CD* for mLB, and mVLB schemes compared to LB and VLB schemes. It can be observed from Figs 5.8 (a) and (d) that the congestion duration reduces if the service rate (N_{si}) is increased, for the same conditions of global congestion communication (v=3, c=3, i.e., all nodes are informed of congestion status).

As observed from Figs 5.3-5.7, the mLB and mVLB schemes reduced $\overline{P}_r$, $\overline{\xi}$, $\overline{\Delta}$, $\overline{\Psi}$, and $\overline{P}_{overflow}$ compared to the respective LB and VLB schemes, for the same traffic

intensity, by controlling active bursts from the local transport users. As *CD* observed in Fig. 5.8, the congestion duration for the local traffic depends mainly on the transit traffic. The *CD* is higher when the ATM multiplexer node is aware of congestion status of other ATM multiplexer nodes. The global congestion control scheme reacts to control congestion based on congestion status communication. Hence, the congestion is observed to be more when the ATM multiplexer node is well informed of the congestion status of other ATM nodes, i.e., $v = c$, so that preventive measures were applied to control the global congestion. When the ATM multiplexer node is unaware of congestion status of other ATM multiplexer nodes, i.e., $v = 0$, only intrinsic (local) traffic contributes to congestion measurement. The mean probability of generating a cell ($\overline{\xi}$) increases linearly with traffic intensity when there is no congestion control as obtained in Fig. 5.4. However, the LB and VLB schemes reduce the amount of traffic when congestion rises beyond the threshold. As LB scheme acts faster to drop all violating cells without depending on the congestion status of the ATM multiplexer node, the $\overline{\xi}$ is obtained to be lower than VLB scheme, for the same traffic intensity. Furthermore, when N_{si} is increased then the $\overline{\xi}$ is increased, since more traffic is serviced for the same QOS. The mean probability of discarding a cell ($\overline{\Delta}$) increases with traffic intensity as observed in Fig. 5.5, since increased traffic intensity rapidly increases the P_{vi} and P_c. The mean Λ_i decreases when the service rate (N_{si}) is increased. The mean buffer content ($\overline{\psi}$) indicates the delay in the ATM multiplexer node, which increases with traffic intensity. As VLB scheme does not discard all violating cells until congestion crosses the threshold, the $\overline{\psi}$ increases, compared to LB scheme at the same traffic intensity. Incessant increase in buffer content may lead to node-

crash, which hampers service to users for longer duration. Hence, it is better to declare buffer overflow beyond an ultimate overflow threshold (*ofl*) on the buffer content as obtained in Fig. 5.6. The ATM multiplexer buffer overflow is observed only for higher traffic intensities as obtained in Fig. 5.7. The probabilities of cell discard and buffer overflow can be used to determine the probability of *cell loss*.

5.5 Transport user performance with underlying ATM congestion control

The traffic control based on ATM cells is one of the important functions to provide high-speed and broadband multimedia communication efficiently on ATM networks. This function is required for guaranteed QOS of various traffic types [YOK-95]. Since ATM does not provide media access control, it has been a concern that the throughput will be low if an ATM network experiences congestion [ROM-95]. Hence, as stated in previous sections, congestion control has been an important challenge in ATM networks. The ATM network can be used for any type of transport users [ROM-95] exercising window based or rate based transport control for the information transfer. Chapter 4 analyzes a local congestion control scheme for hybrid ATM/TDMA networks. The previous sections of this chapter described a global congestion control algorithm to prioritize the transit traffic by regulating the traffic from the local transport users. This scheme eventually achieves global congestion control with the aid of *early warning cells* for congestion communication. In this section, we analyze the throughput and delay characteristics of end-to-end ARQ transport users over a congestion controlled ATM/TDMA network as shown in Fig. 5.9. The underlying ATM layer is assumed to use LB, VLB, mLB, and mVLB *cell discard* schemes. The results are compared with throughput and delay characteristics of end-to-end ARQ transport users without any underlying ATM congestion control, i.e., with only TPDU discard at transport layer, which is referred as *packet discard*. Similar work has been simulated in [ROM-95] for transport users using TCP/IP with and without an ATM backbone network. The underlying global congestion control scheme improves the throughput and delay characteristics for any transport

mechanism, i.e., TCP/IP, or UDP, or DECnet, or VMTP, or AppleTalk, or a GBN ARQ or SR ARQ scheme. The involved transport protocol mechanism provides an efficient usage of the available bit rate (ABR) or QOS.

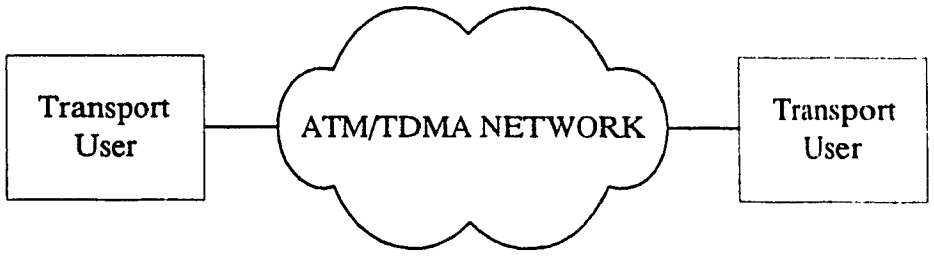

Fig. 5.9. Transport users over ATM/TDMA network

The ARQ transport users are assumed over LEO and GEO satellite links for analyzing throughput and delay characteristics of ARQ transport users in chapter 2. In this section, the transport user packets are segmented into ATM cells before transmitting them over the broadband ATM network. The ATM network transmits cells of several service classes by means of composite TDMA frames as described in chapter 4. The ATM global congestion scheme is applied to control congestion as described in the sections 5.2, 5.3 and 5.4. The resulting ATM multiplexer node buffer is analyzed using MMPP queueing model with bulk arrival and bulk service of cells, as analyzed in chapter 4. The LB, VLB, mLB, and mVLB cell discard schemes are used to discard cells to control congestion at run-time. The transport users are assumed to use GBN or SR end-to-end ARQ scheme. The end-to-end throughput and delay characteristics are analyzed with and without ATM cell discard schemes, which are refened as *cell discard* and *packet discard* respectively.

Due to the limited buffer space of a service-class i $(L_i B_i)$, and the employed cell discard scheme, the ATM cells may be lost due to buffer overflow ($P_{overflow}$), violation of QOS (P_{vi}), and *preventive* congestion (P_c) control. The mean probability of cell loss is defined as

$$\overline{P}_{cell-loss}|_{K_i,\rho_i} = \overline{\Delta}|_{K_i,\rho_i} + \overline{P}_{overflow}|_{K_i,\rho_i} \tag{5.26}$$

where $\overline{\Delta}|_{K_i,\rho_i}$ is the probability of cell discard, and $\overline{P}_{overflow}|_{K_i,\rho_i}$ is the probability of buffer overflow, which were analyzed in equations (5.22) and (5.25), respectively.

Equation (5.26) approximates the cell loss as a sum of higher buffer state probabilities, and the probability of cell-discard due to the congestion control scheme. This is a conservative approach, since the other approaches depend only on the final buffer state. Hence, equation (5.26) provides *best effort preventive* congestion control by taking *early* actions.

Assuming that an average number of n cells per packet, the probability of correct packet transmission is

$$P_{correct-packet} = (1 - P_{cell-loss})^n. \tag{5.27}$$

where n is the number of cells per packet.

The service rate equations (2.18) and (2.21) assume $P_{correct-packet}$ to be 1, i.e., a packet is always correctly transmitted. However, the underlying ATM global congestion control

scheme will introduce *cell loss* as defined in equation (5.26). Hence, the service rate equations (2.18) and (2.21) will be modified to

$$\mu'|_{GBN} = \mu|_{GBN} \; P_{correct-packet}, \qquad and$$
$$\mu'|_{SR} = \mu|_{SR} \; P_{correct-packet} \tag{5.28}$$

where $P_{correct\text{-}packet}$ is defined in equation (5.27), $\mu|_{GBN}$ is defined in equation (2.18), and $\mu|_{SR}$ is defined in equation (2.21).

Hence, the transport user throughput and delay results are obtained from equations (2.9) and (2.10), by replacing μ by μ' in equations (2.9) and (2.10), respectively. The end-to-end throughput and delay results with the underlying ATM global congestion control, are obtained in Figs 5.10-5.17.

The end-to-end throughput characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for GBN ARQ scheme in Fig. 5.10. These results were obtained for 9600 bps VC, with a packet size of 256 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The ATM multiplexer node allocates more buffer for the transit traffic when v=c (i.e., all other nodes are using global control). Hence, the results in Fig. 5.10 are obtained by assuming fully informed scenario (v=c), for local traffic. It can be observed from results that the end-to-end throughput increases with the window size (N),

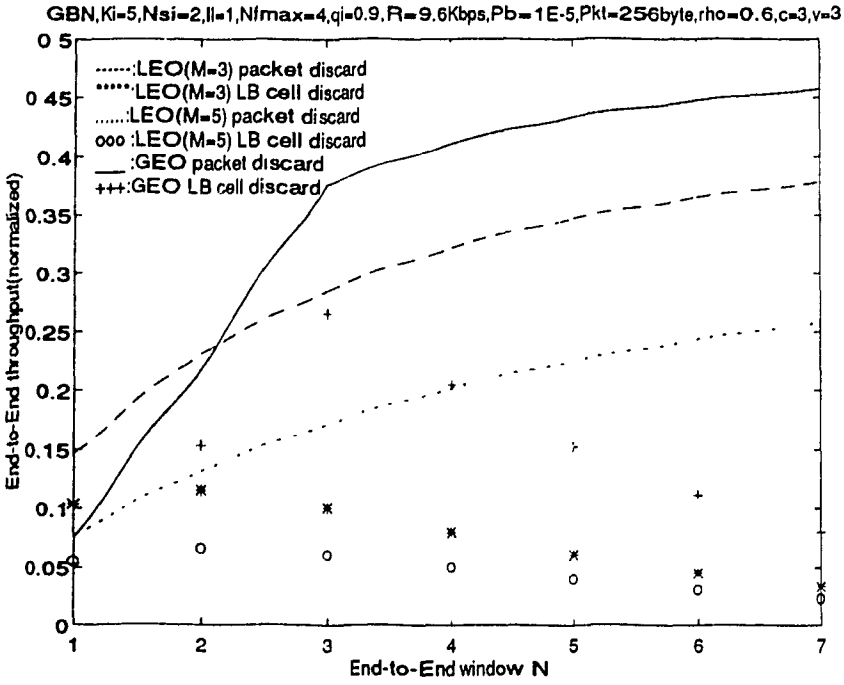

Fig. 5.10 (a) LB cell discard;

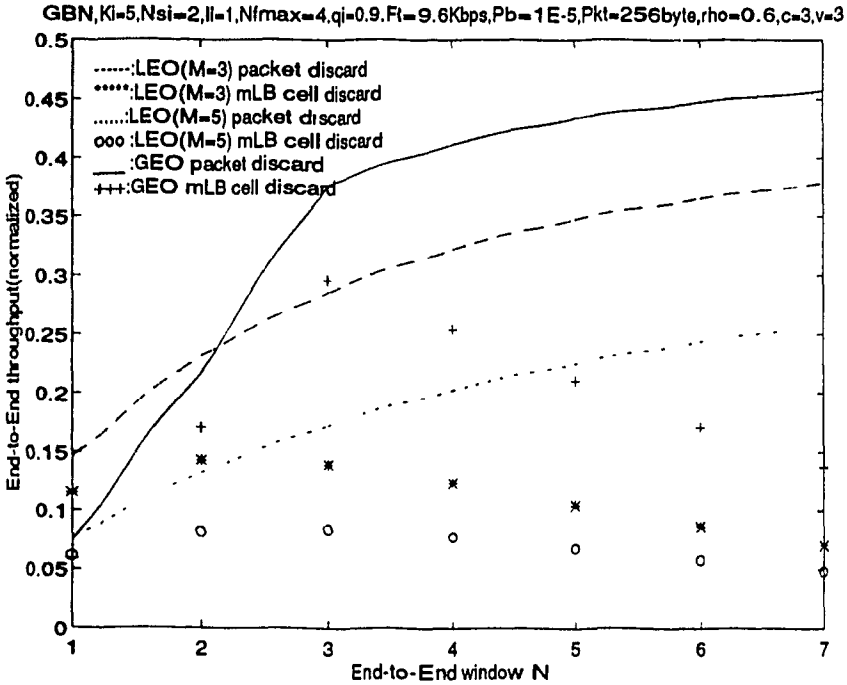

Fig. 5.10 (b) mLB cell discard;

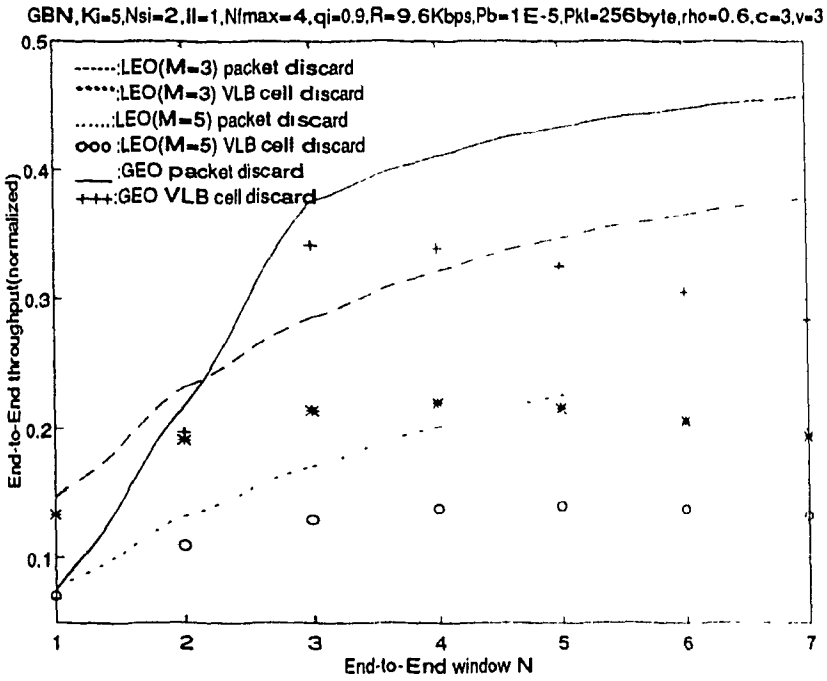

Fig. 5.10 (c) VLB cell discard;

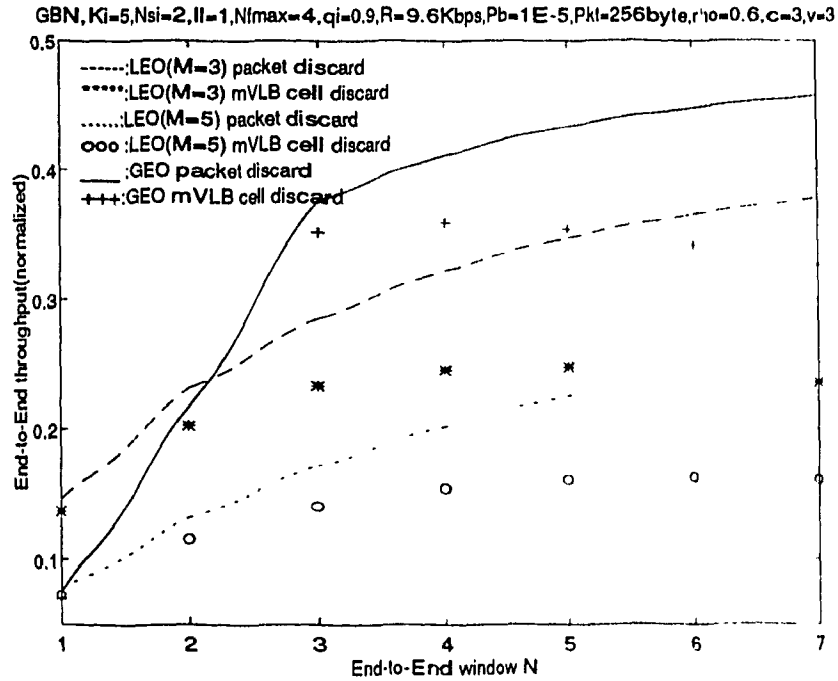

Fig. 5.10 (d) mVLB cell discard;

Fig. 5.10. Go-Back-N end-to-end throughput, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps.

when ATM congestion control scheme is not used, or the underlying congestion control is not considered for analyzing the end-to-end throughput. The effective end-to-end throughput decreases when the underlying congestion control schemes (i.e., LB, or VLB, or mLB, or mVLB) act to reduce local traffic based on the instantaneous congestion status. It can be observed from the results that the mVLB congestion control scheme in Fig. 5.10 (d) provides optimum end-to-end throughput, when compared to other congestion control schemes of Figs 5.10 (a), (b), and (c). Moreover, the 2 LEO satellite (M=3) configuration is observed to be better at very low window sizes, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configuration.

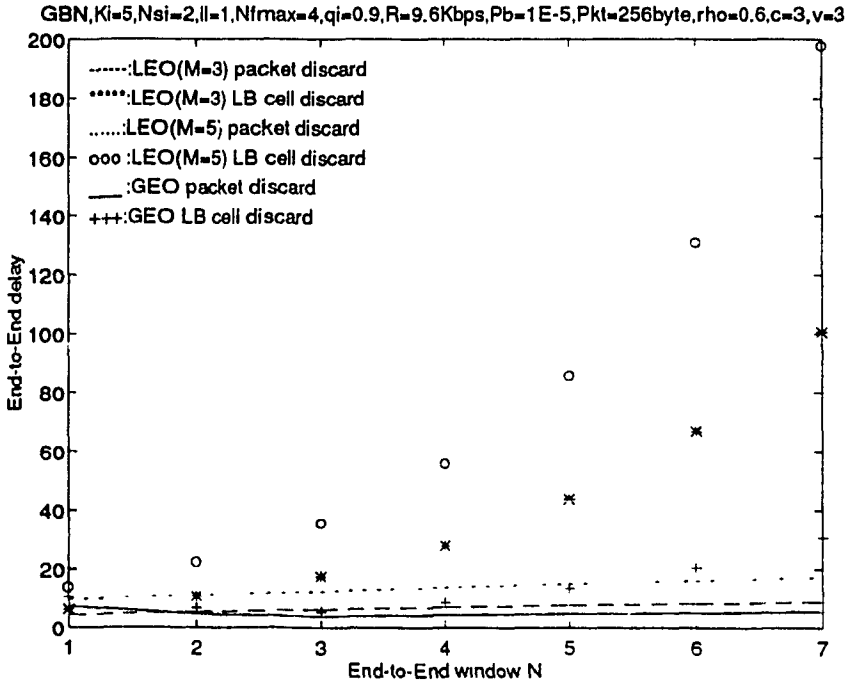

Fig. 5.11 (a) LB cell discard;

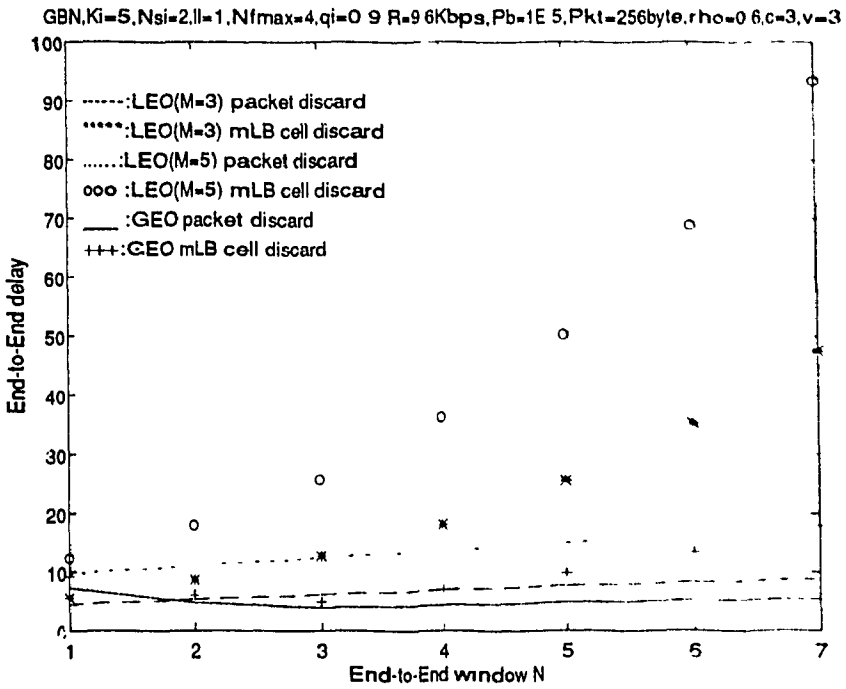

Fig. 5.11 (b) mLB cell discard;

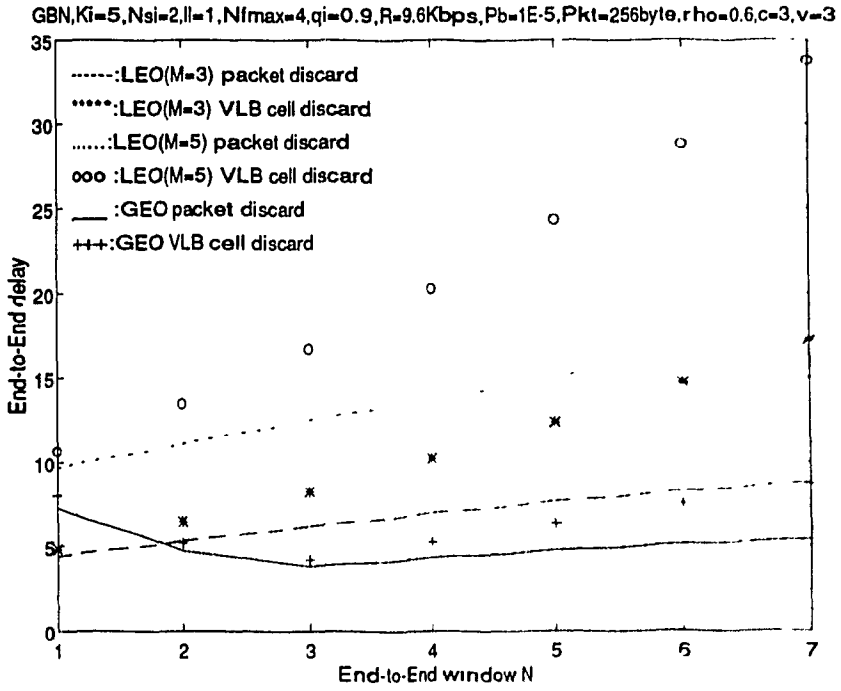

Fig. 5.11 (c) VLB cell discard;

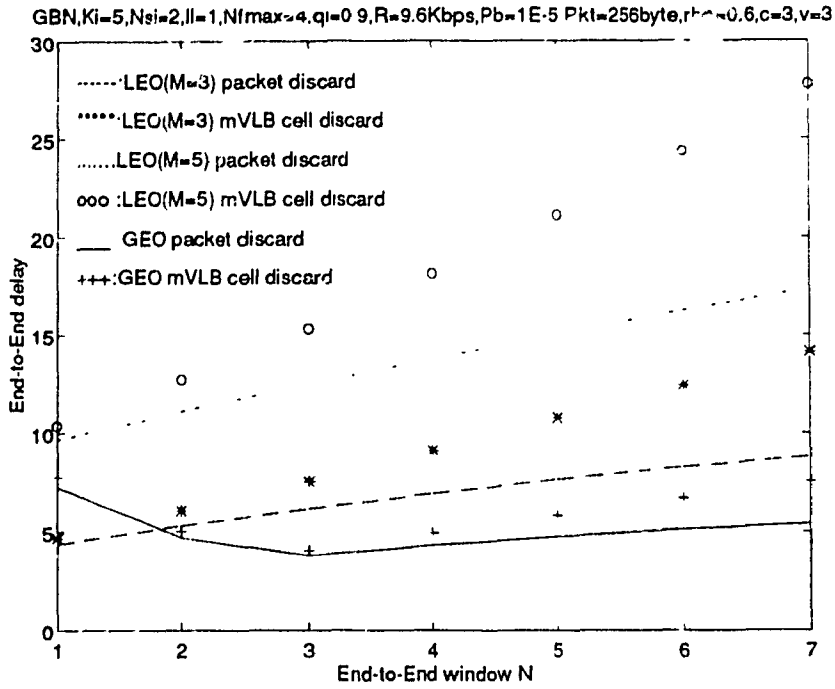

Fig. 5.11 (d) mVLB cell discard;

Fig. 5.11. Go-Back-N end-to-end delay, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps.

The end-to-end delay characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for GBN ARQ scheme in Fig. 5.11. These results were obtained for 9600 bps VC, with a packet size of 256 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The results in Fig. 5.11 are obtained by assuming fully informed scenario (v=c), for local traffic. These results correspond to the same set of protocol parameters that resulted the end-to-end throughput characteristics of Fig. 5.10. It can be observed from results that the end-to-end delay increases with the window size (N), when ATM congestion control scheme is not used, or the underlying congestion control is not considered for analyzing the end-to-end delay. The effective end-to-end delay increases as

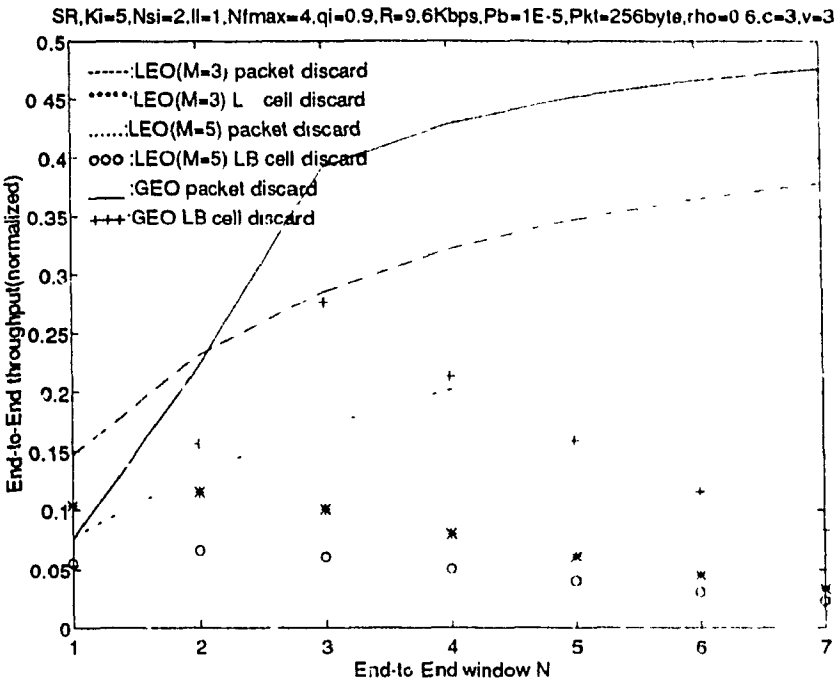

Fig. 5.12 (a) LB cell discard;

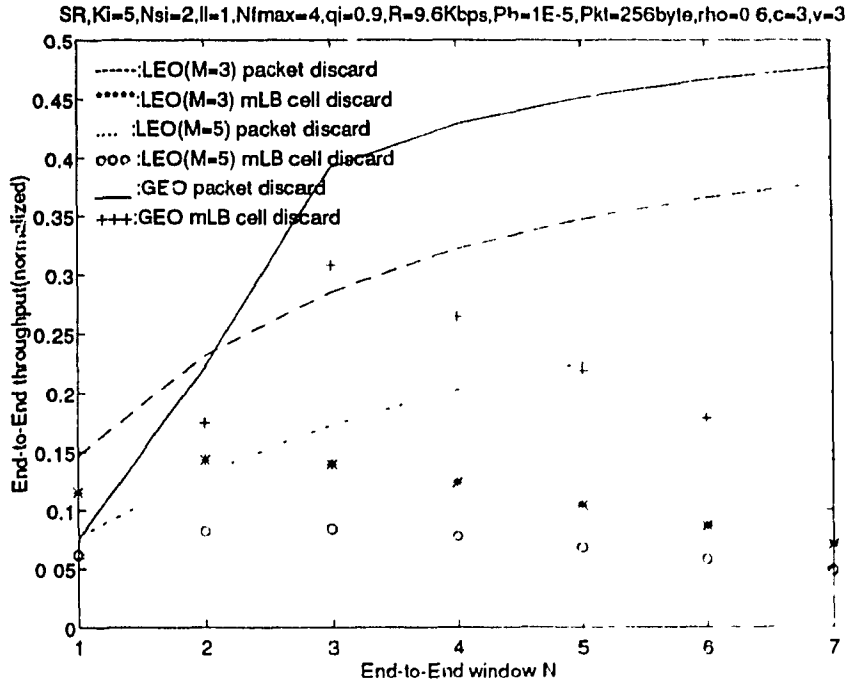

Fig. 5.12 (b) mLB cell discard;

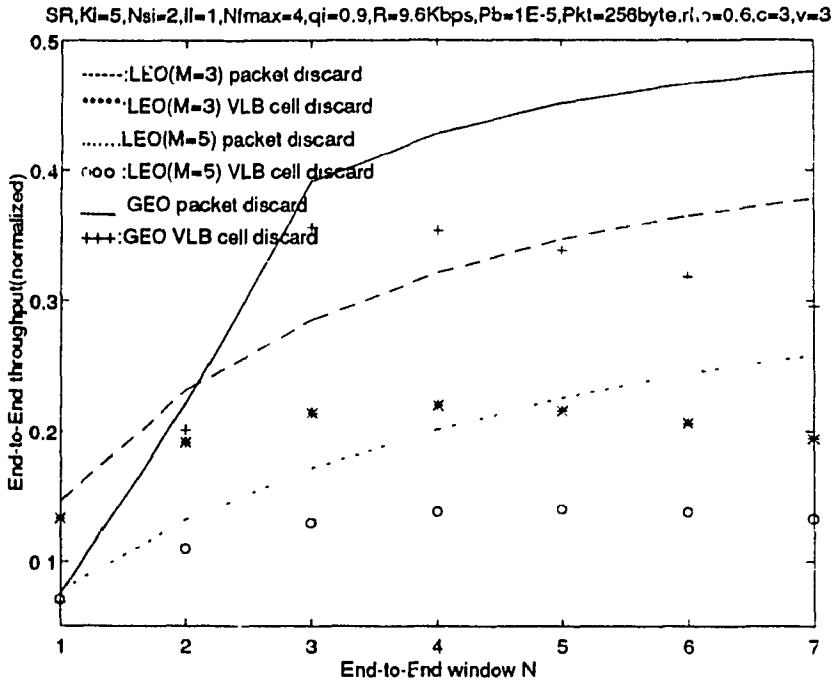

Fig. 5.12 (c) VLB cell discard;

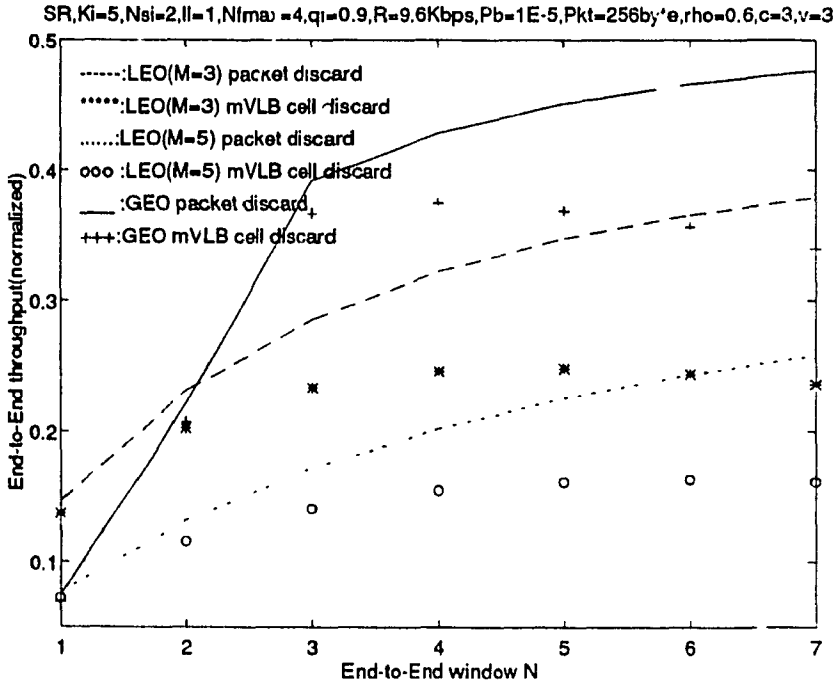

Fig. 5.12 (d) mVLB cell discard;

Fig. 5.12. SR end-to-end throughput, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps.

the underlying congestion control schemes act to reduce local traffic based on the instantaneous congestion status. It can be observed from the results that the mVLB congestion control scheme in Fig. 5.11 (d) provides optimum end-to-end delay when compared to other congestion control schemes of Figs 5.11 (a), (b), and (c). Moreover, the 2 LEO satellite (M=3) configuration is observed to be better at very low window sizes, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configuration.

The end-to-end throughput characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for SR ARQ scheme in Fig. 5.12. These results were obtained for 9600 bps VC, with a packet size of 256 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The results in Fig. 5.12 are obtained by assuming fully informed scenario (v=c), for local traffic. These results of SR ARQ scheme correspond to the same set of protocol parameters that resulted the end-to-end throughput for GBN ARQ scheme in Fig. 5.10. It can be observed from results that the SR ARQ scheme with an underlying congestion control scheme, provides better end-to-end throughput performance compared to GBN ARQ scheme with the same underlying congestion control scheme. It can be observed from the results that the mVLB congestion control scheme in Fig. 5.12 (d) provides optimum end-to-end throughput performance, compared to other congestion control schemes of Figs 5.12 (a), (b), and (c). Moreover, the 2 LEO satellite (M=3) configuration is observed to be better at very low window sizes, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configuration.

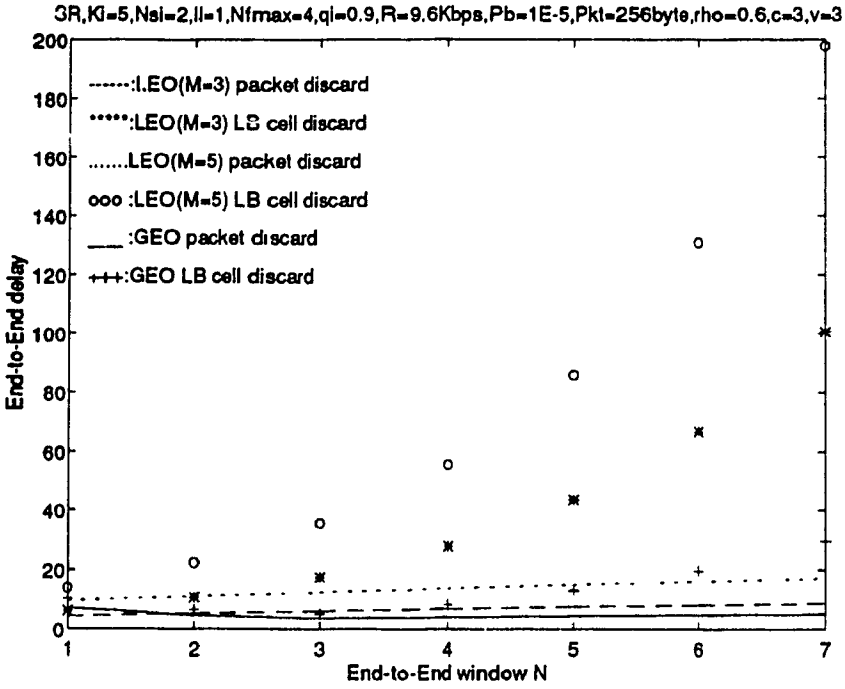

Fig. 5.13 (a) LB cell discard;

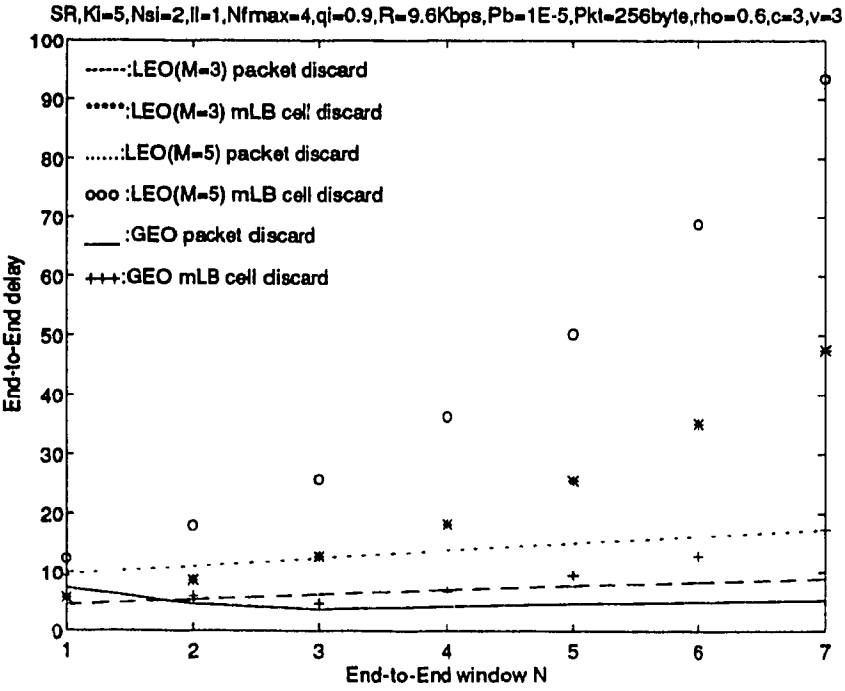

Fig. 5.13 (b) mLB cell discard;

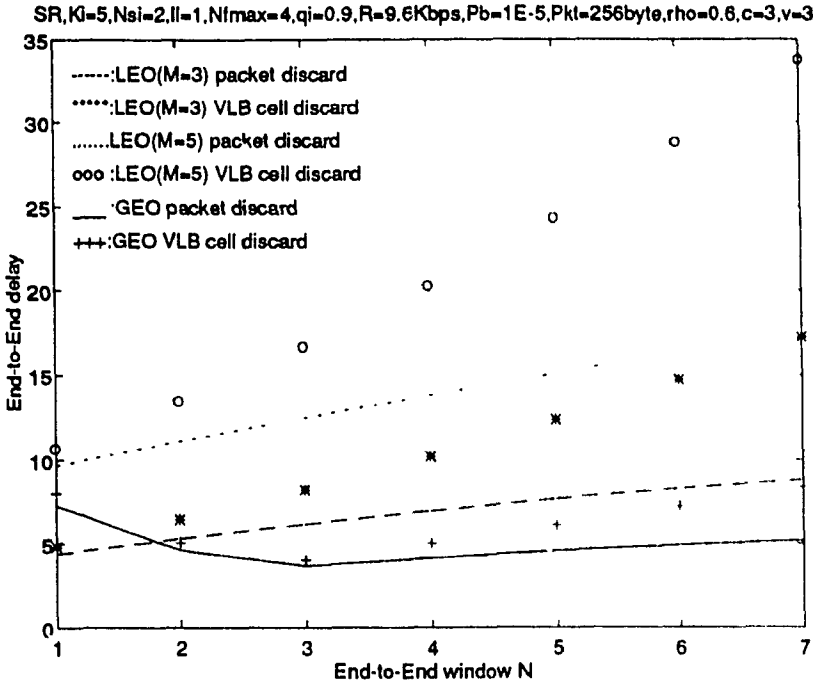

Fig. 5.13 (c) VLB cell discard;

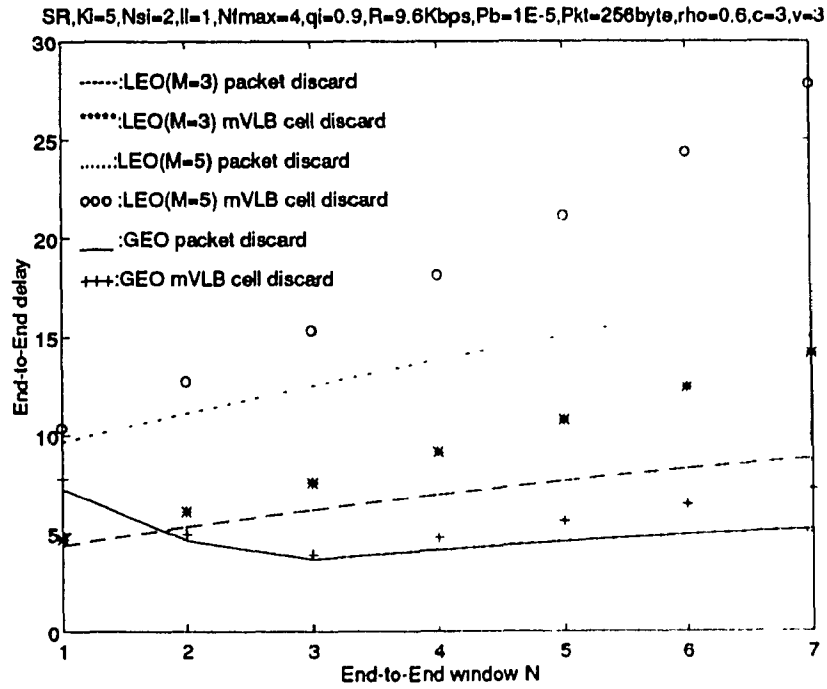

Fig. 5.13 (d) mVLB cell discard;

Fig. 5.13. SR end-to-end delay, packet=256 bytes, n_h=7 bytes, R=9.6 Kbps.

The end-to-end delay characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for SR ARQ scheme in Fig. 5 13. These results were obtained for 9600 bps VC, with a packet size of 256 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The results in Fig. 5.13 are obtained by assuming fully informed scenario (v=c), for local traffic. It can be observed from the results that the SR ARQ scheme with an underlying mVLB congestion control provides low end-to-end delay compared to LB, mLB, and VLB congestion control schemes. Moreover, the 2 LEO satellite (M=3) configuration is observed to be better at very low window sizes, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configuration.

As detailed above, the transport user throughput and delay characteristics were obtained in Figs 5.10-5.13, using a typical set of protocol parameters. i.e., packet size of 256 bytes at 9600 bps, which can be a portion of the link bandwidth. These parameter values are generally used for data applications [WU-87]. It was consistently observed that mLB and mVLB schemes provide better performance when compared to LB and VLB schemes, respectively. However, both VLB and mVLB cell discard schemes were observed to be better when compared to LB and mLB cell discard schemes, respectively. The performance of GBN and SR ARQ schemes were almost identical, when the probability of bit error (P_b) is low (10^{-5}) for VC transport user at 9600 bps, low window (i.e., 1-7), and a low packet size of 256 bytes. The LEO links were observed to be better

at lower window sizes, whereas GEO link was observed to be better for higher window sizes. The results were obtained by using the same P_b value for both LEO and GEO satellite links. In fact, the LEO satellite links provide lower P_b compared to GEO satellite links. By using the correct P_b values for the LEO and GEO satellite links, further improvement in the performance can be observed for LEO. As GBN ARQ scheme is window sensitive, SR ARQ will result better performance at higher window sizes (which can be observed shortly in the following results).

The end-to-end throughput characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for GBN ARQ scheme in Fig. 5.14. These results were obtained for 64 Mbps VC, with a packet size of 4096 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and

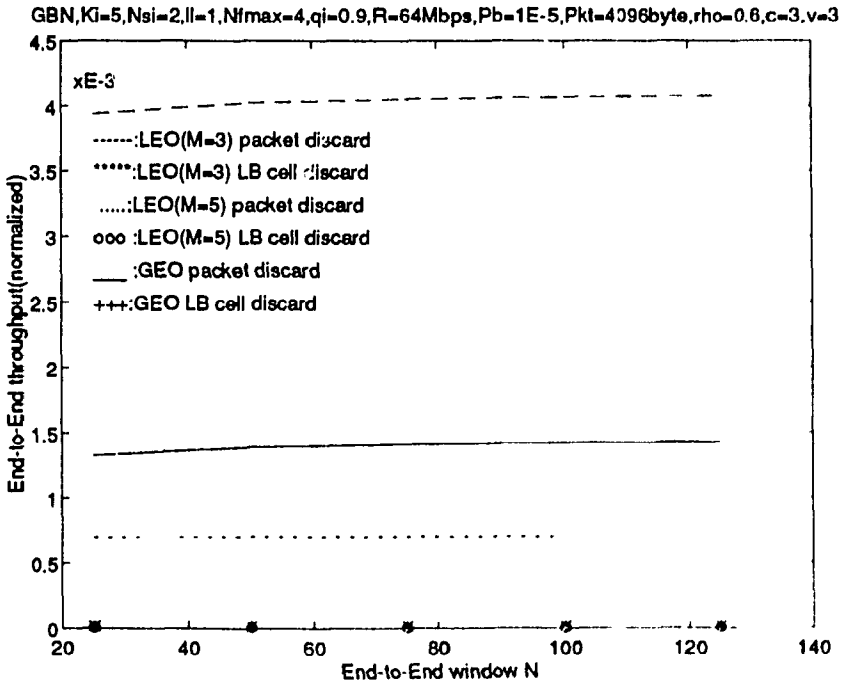

Fig. 5.14 (a) LB cell discard;

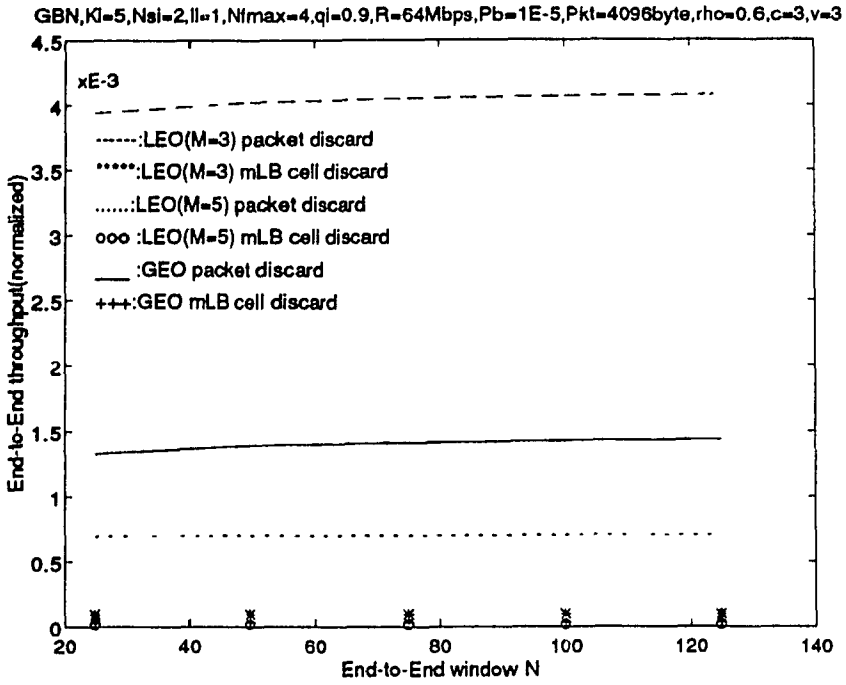

Fig. 5.14 (b) mLB cell discard;

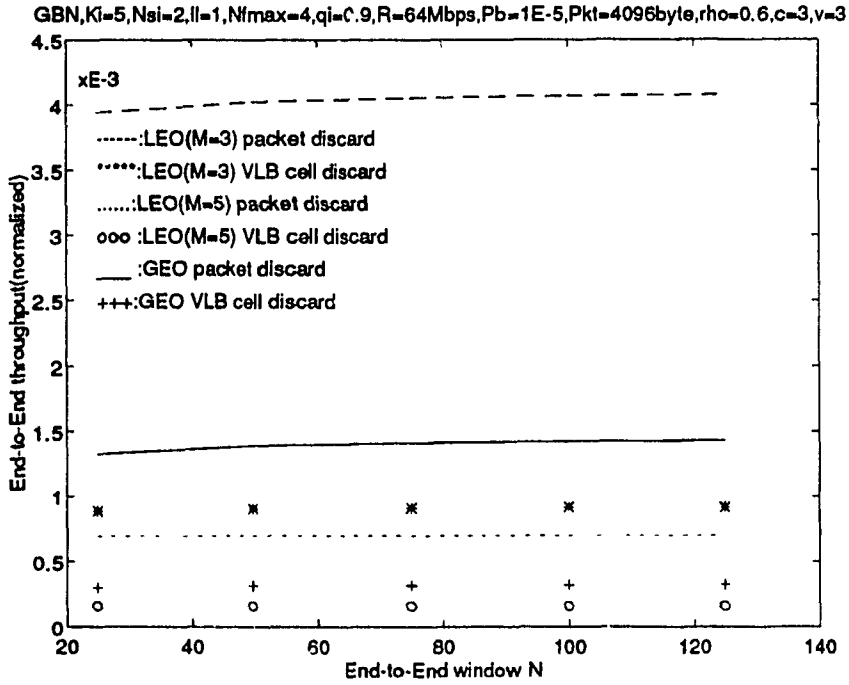

Fig. 5.14 (c) VLB cell discard;

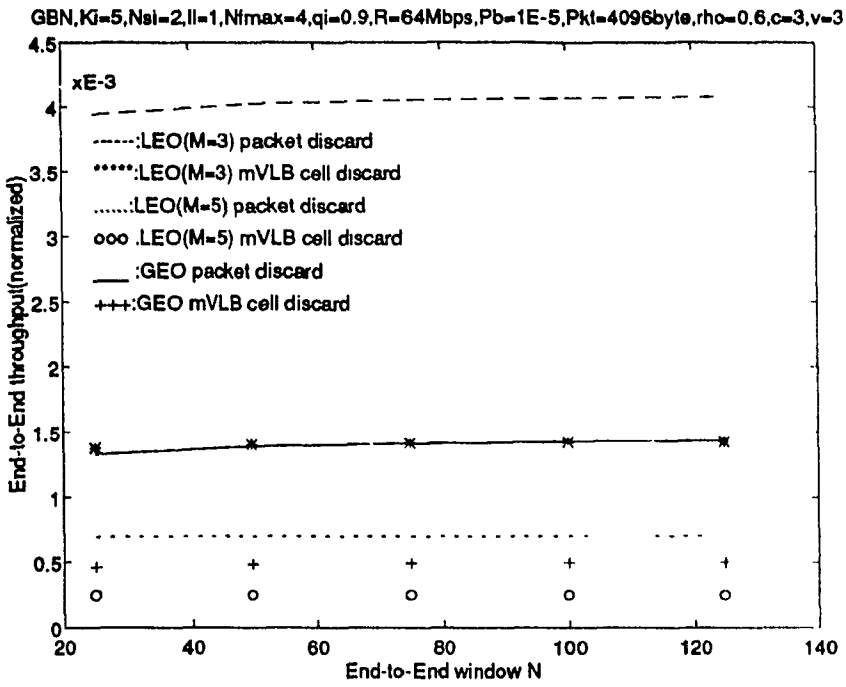

Fig. 5.14 (d) mVLB cell discard;

Fig. 5.14. Go-Back-N end-to-end throughput, packet=4096 bytes, n_h=7 bytes,R=64 Mbps.

a traffic intensity of 0.6. The results in Fig. 5.14 are obtained by assuming fully informed scenario for local traffic. It can be observed from results that the GBN ARQ scheme is inferior to SR ARQ at high bandwidths. Moreover, the end-to-end throughput does not significantly change with the window size (N), when ATM congestion control scheme is not used, or the underlying congestion control is not considered for analyzing the end-to-end throughput. The effective end-to-end throughput is observed to be significantly low, since the underlying congestion control schemes act to reduce local traffic based on the instantaneous congestion status. It can be observed from the results that the mVLB congestion control scheme in Fig. 5.14 (d) provides optimum end-to-end throughput, when compared to other congestion control schemes of Figs 5.14 (a), (b), and (c). It can

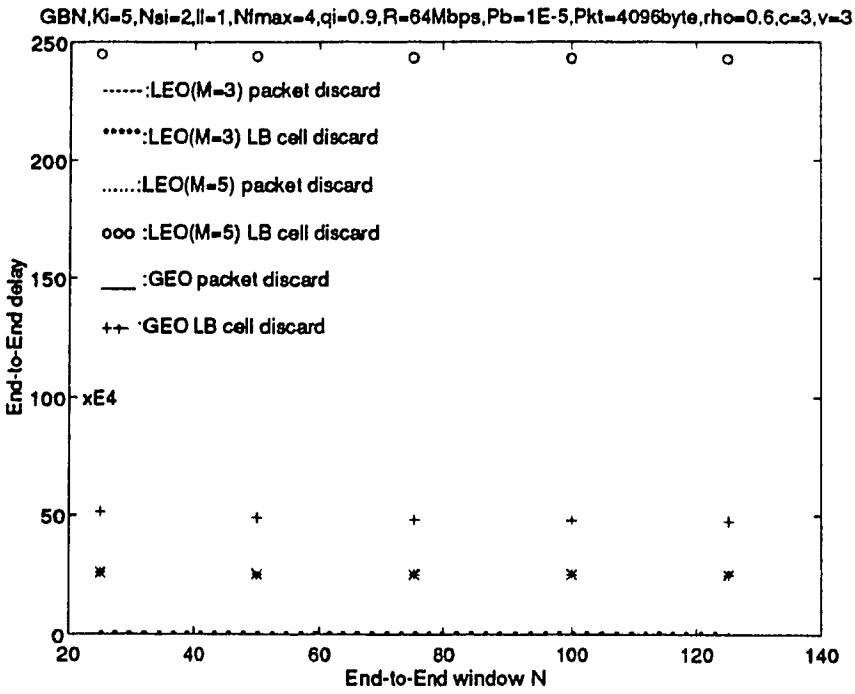

Fig. 5.15 (a) LB cell discard;

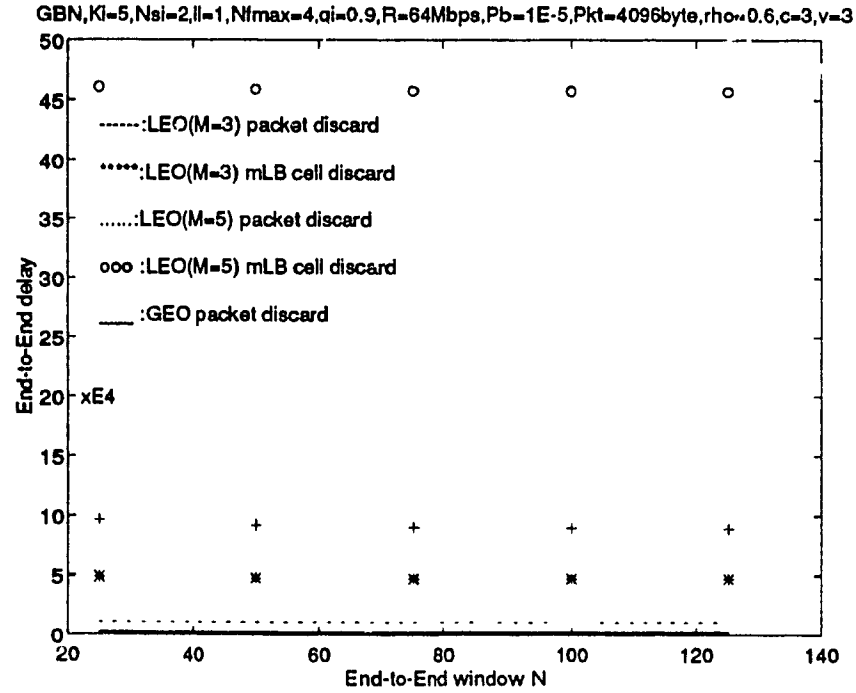

Fig. 5.15 (b) mLB cell discard;

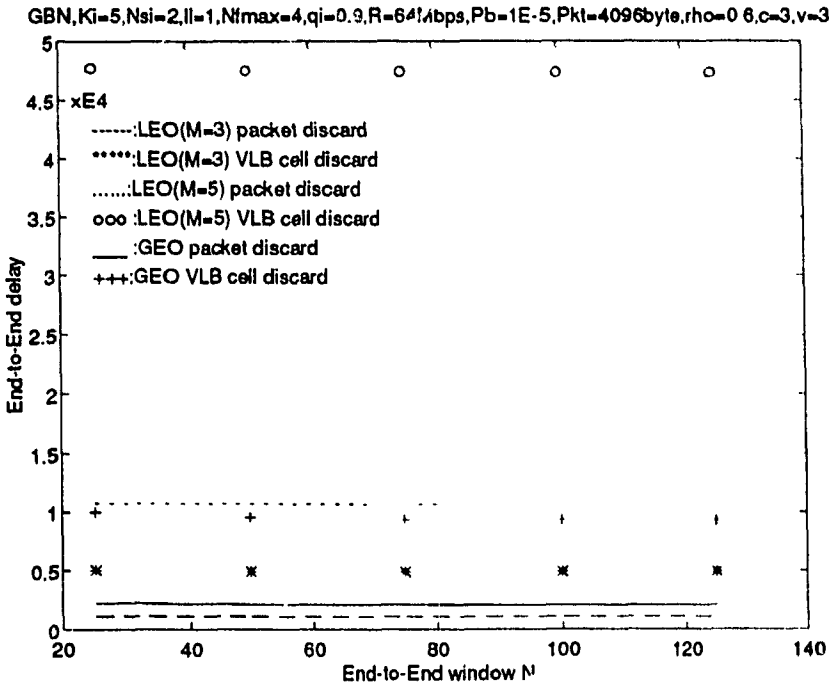

Fig. 5.15 (c) VLB cell discard;

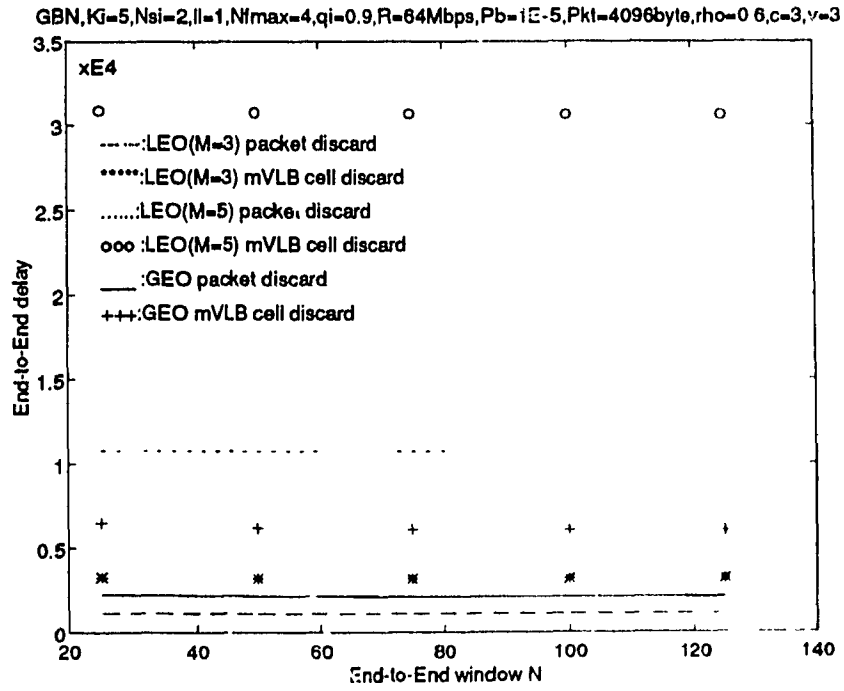

Fig. 5.15 (d) mVLB cell discard;

Fig. 5.15. Go-Back-N end-to-end delay, packet=4096 bytes, n_h=7 bytes, R=64 Mbps.

be clearly observed from Fig. 5.14 that the 2 LEO satellite (M=3) configuration provides better throughput results for the high bandwidths, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configurations.

The end-to-end delay characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for GBN ARQ scheme in Fig. 5.15. These results were obtained for 64 Mbps VC, with a packet size of 4096 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The results in Fig. 5.15 are obtained by assuming fully informed scenario (v=c), for local traffic. These results correspond to the same set of protocol parameters of Fig. 5.14. It can be observed from results that the GBN ARQ scheme is inferior to SR ARQ at high bandwidths. Moreover, the results show that the end-to-end delay does not significantly change with the window size (N), when ATM congestion control scheme is not used, or the underlying congestion control is not considered for analysis. The effective end-to-end delay was observed to be significantly high since the underlying congestion control schemes act to reduce local traffic based on the instantaneous congestion status. It can be observed from the results that the mVLE congestion control scheme in Fig. 5.15 (d) provides optimum end-to-end delay, compared to other congestion control schemes of Figs 5.15 (a), (b), and (c). It can be clearly observed from Fig. 5.15 that the 2 LEO satellite (M=3) configuration provides better end-to-end delay results for the high bandwidths, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configurations.

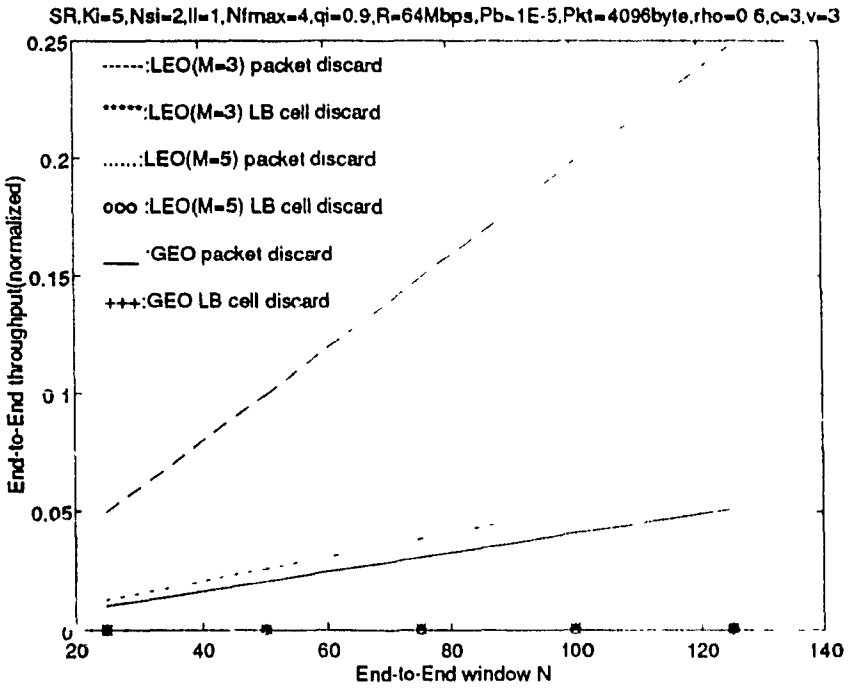

Fig. 5.16 (a) LB cell discard;

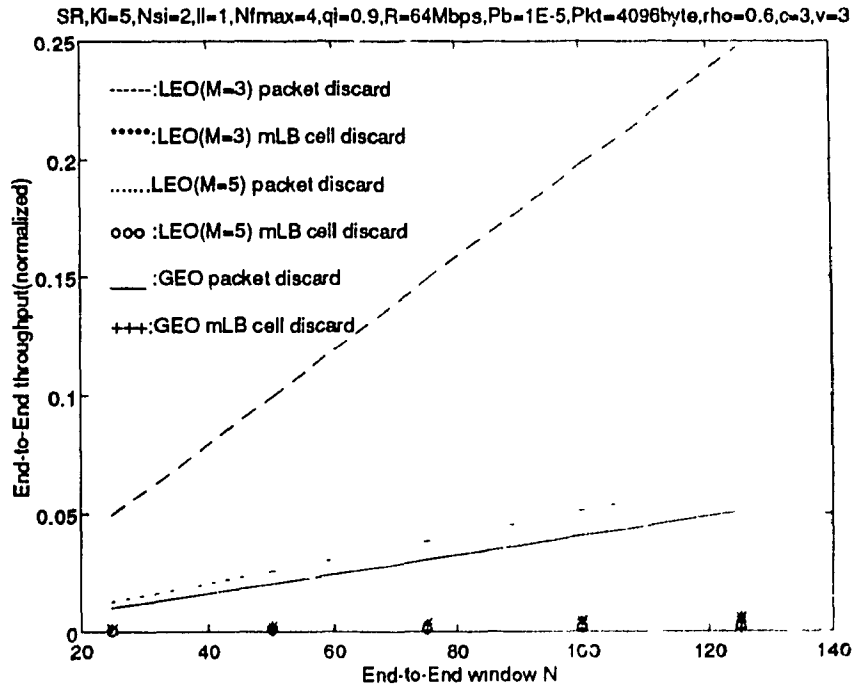

Fig. 5.16 (b) mLB cell discard;

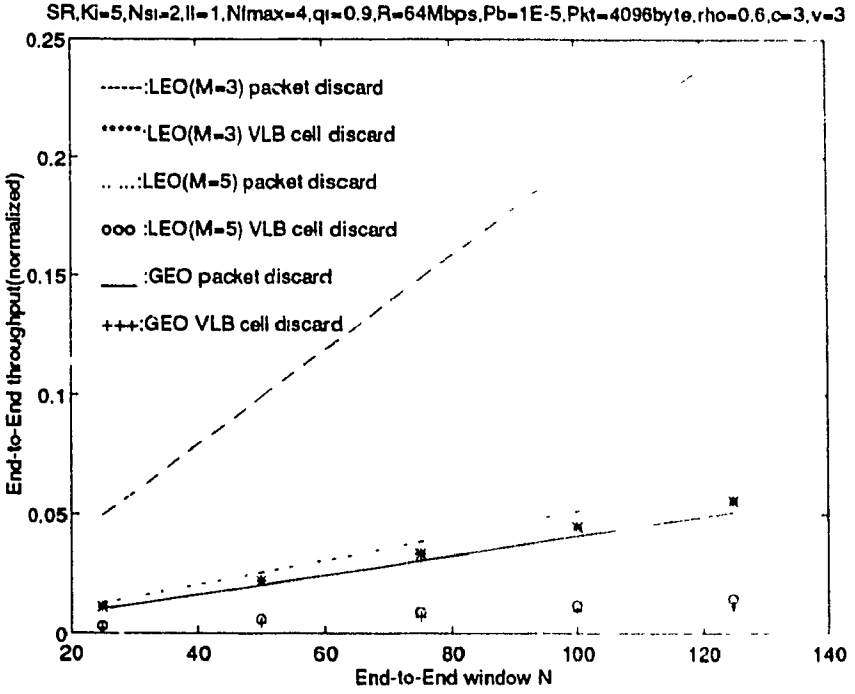

Fig. 5.16 (c) VLB cell discard;

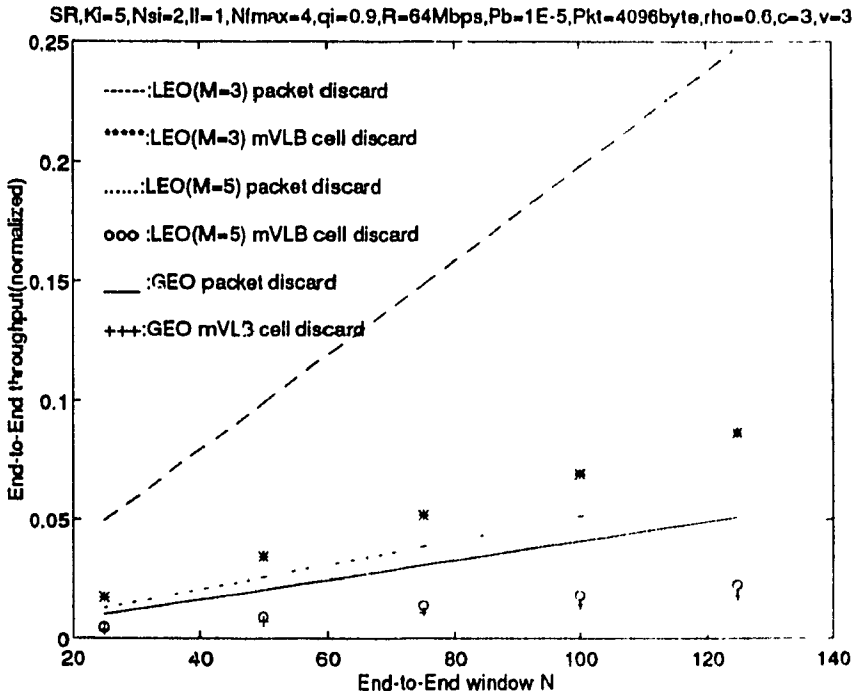

Fig. 5.16 (d) mVLB cell discard;

Fig. 5.16. SR end-to-end throughput, packet=4096 bytes, n_h=7 bytes, R=64 Mbps.

The end-to-end throughput characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for SR ARQ scheme in Fig. 5.16. These results were obtained for 64 Mbps VC, with a packet size of 4096 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The results in Fig. 5.16 are obtained by assuming fully informed scenario (v=c) for local traffic. These results are obtained for the same protocol parameters of GBN ARQ scheme of Fig. 5.14, for comparison. It can be clearly observed that the end-to-end throughput significantly increases with the window size (N). The effective end-to-end throughput was observed to be significantly low when the underlying congestion control schemes act to reduce the congestion based on the instantaneous congestion status. It can be observed from the results that the mVLB congestion control scheme in Fig. 5.16 (d) provides optimum end-to-end throughput, when compared to

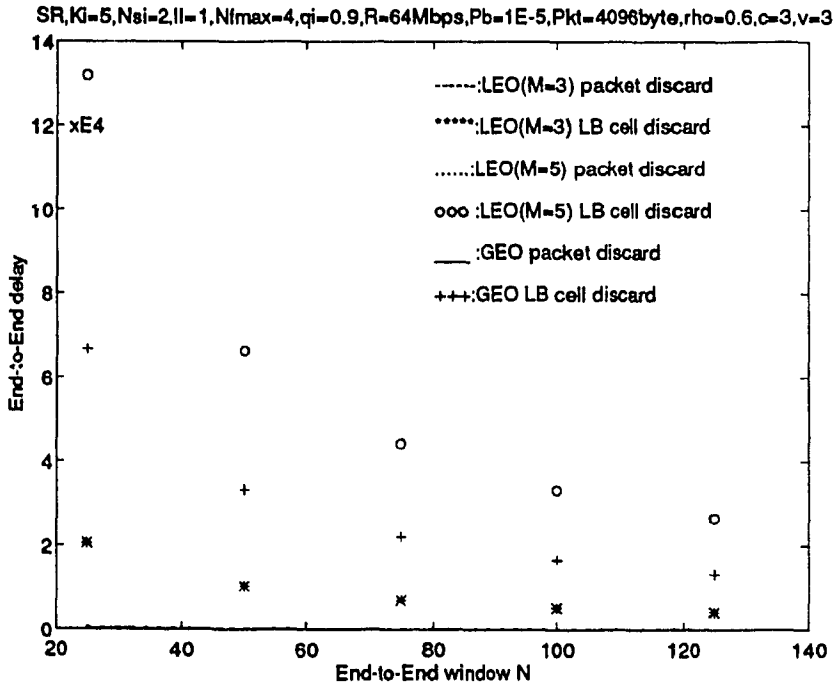

Fig. 5.17 (a) LB cell discard;

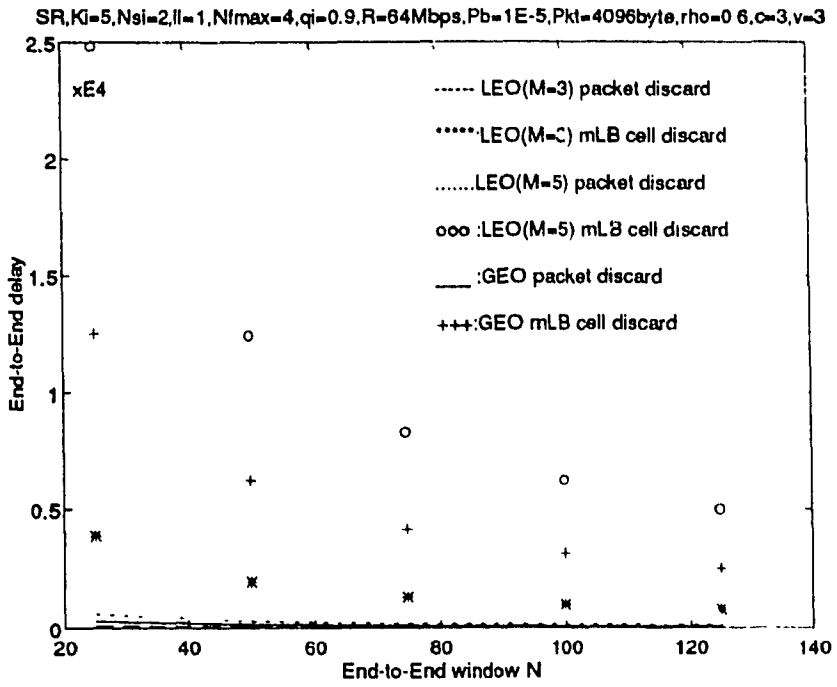

Fig. 5.17 (b) mLB cell discard;

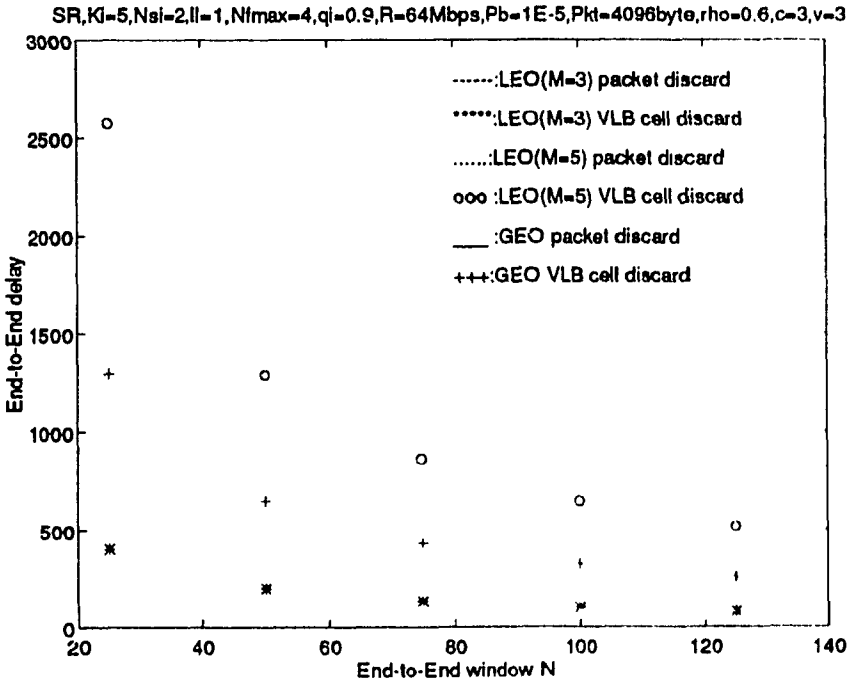

Fig. 5.17 (c) VLB cell discard;

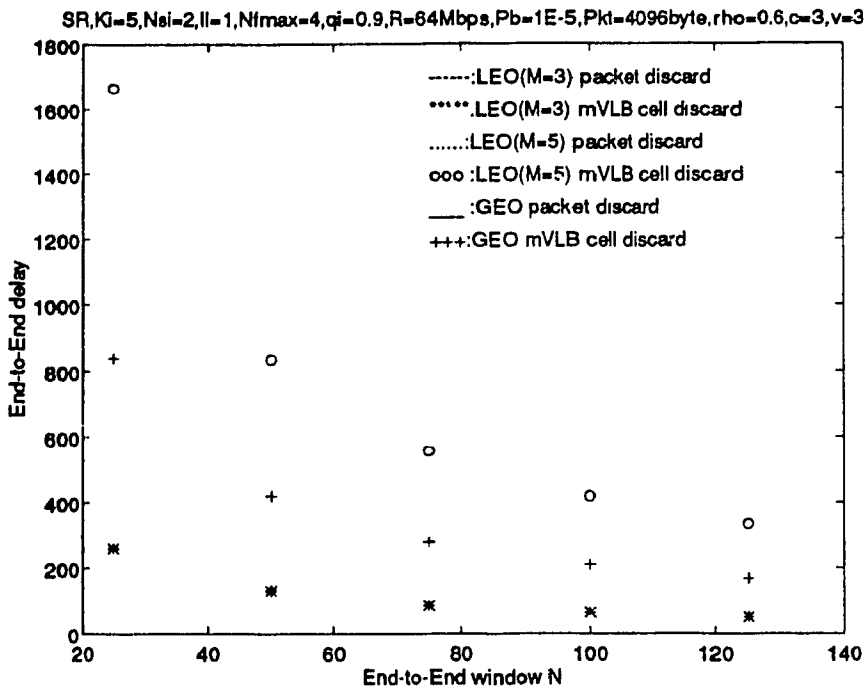

Fig. 5.17 (d) mVLB cell discard;

Fig. 5.17. SR end-to-end delay, packet=4096 byte, n_h=7 bytes, R=64 Mbps.

other congestion control schemes of Figs 5.16 (a), (b), and (c). It can be clearly observed from Fig. 5.16 that the 2 LEO satellite (M=3) configuration provides better results for the high bandwidths, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configurations.

The end-to-end delay characteristics of a transport user with the underlying ATM congestion control scheme were obtained as a function of window size (N) for SR ARQ scheme in Fig. 5.17. These results were obtained for 64 Mbps VC, with a packet size of 4096 bytes, packet header size of 7 bytes, probability of bit error (P_b) of 10^{-5}, and a traffic intensity of 0.6. The results in Fig. 5.16 are obtained by assuming fully informed scenario

(v=c), for local traffic. It can be clearly observed that the end-to-end delay significantly decreases when the window size (N) is increased. The effective end-to-end delay was observed to be significantly high when the underlying ATM layer introduced additional delay by means of queueing. It can be observed from the results that the mVLB congestion control scheme in Fig. 5.17 (d) provides optimum end-to-end delay, when compared to other congestion control schemes of Figs 5.17 (a), (b), and (c). It can be clearly observed from Fig. 5.17 that the 2 LEO satellite (M=3) configuration provides better results for the high bandwidths, compared to 1 GEO satellite (M=2) or 4 LEO satellite (M=5) configurations.

The Figs 5.14-5.17 were obtained, using a typical set of protocol parameters, i.e., packet size of 4096 bytes at 64 Mbps which can be a portion of the link bandwidth, assigned to the VC. These parameter values are generally used for video or voice applications [WU-87], [BAE-91]. The VLB and mVLB cell discard schemes were observed to be better than LB and mLB cell discard schemes, for optimum end-to-end throughput and delay. At high bandwidths (64 Mbps), the SR ARQ scheme was observed to be significantly better than GBN ARQ. It was also observed from in Figs. 5.14 and 5.16 that the LEO (M=3) configuration was better than GEO (M=2) configuration.

5.6 Conclusions

The ATM multiplexer global congestion was analyzed by utilizing *early warning cells* from other ATM multiplexer nodes, to ensure necessary bandwidth for supporting

various service classes, i.e., video, voice, and data, in a hybrid ATM/TDMA network. The transport user throughput and delay characteristics were also analyzed with the underlying ATM global congestion control. The congestion performance analysis was simple for faster runtime decisions, since the applied controls, i.e., LB, VLB, mLB, and mVLB, were simple, even though the performance computation was little cumbersome. The congestion performance computation was justified over the measurement period, which was demonstrated to be quite accurate. As observed in the results, the proposed global congestion control algorithm adapted to the varying intensities of transit traffic, based on congestion status communication. The three-state global congestion control algorithm was also observed to be very simple, and fast for state transitions to control the global congestion. Similar analysis can be applied to any other congestion state diagram with moderately higher number of states, to improve the accuracy of results. The global congestion performance results clearly indicated that congestion was well controlled using LB, VLB, mLB, and mVLB congestion control schemes. It was also observed that LB scheme provided a better control of congestion, when compared to VLB scheme, because the VLB scheme discards the violating cells only when congestion is intolerable. This fact was well observed throughout the congestion performance results. The mLB and mVLB schemes improved congestion performance over the respective LB and VLB schemes, by controlling active bursts from the local transport users. It was clear from the results that VLB and mVLB schemes provided higher throughput when compared to LB and mLB schemes, respectively. Even though LB and mLB schemes provided better congestion performance, these schemes degraded the throughput and delay performance of the

transport user. When the probability of bit error is decreased by assuming better communication media (such as optical fibers) or LEO satellites at lower altitudes, the VLB and mVLB schemes were observed to be optimum choice for better throughput and delay performance. By controlling active bursts at the transport users, the mVLB scheme with SR ARQ scheme was observed to be a better choice for optimum congestion, end-to-end throughput and delay performance. Hence, the virtual-circuit transport user with an underlying mVLB ATM cell discard scheme was observed to be the best choice for end-to-end throughput and delay performance.

CHAPTER 6

CONCLUSIONS

6.1 Summary

The ARQ transport user end-to-end throughput and delay characteristics were analyzed using protocol efficiency formulas and Norton equivalent queueing model, to derive more realistic results of the system. New congestion control strategies were also derived for an ATM multiplexer node, to provide the required bandwidth for supporting various classes of service in a broadband hybrid ATM/TDMA network. The congestion control schemes were designed to improve the performance of traditional leaky bucket and virtual leaky bucket congestion control schemes. The ATM multiplexer node buffer was analyzed using Markov modulated Poisson process (MMPP) with bulk arrival and bulk service of cells. Simple analytical formulas were derived to analyze the congestion criteria of the ATM multiplexer node. A three-state global congestion algorithm was also designed to regulate traffic from the local transport users to prioritize transit traffic from the adjacent ATM multiplexer nodes, by utilizing *early warning cells*. Finally, the ARQ (GBN, and SR) transport user throughput and delay were analyzed with an underlying ATM *global* congestion control scheme, to derive more practical results of the system.

6.2 Conclusions

We have analyzed throughput and delay performance of the transport user in a hybrid ATM/TDMA network. The analysis was very general, and can be applied to any

other high performance transport mechanism. First, the transport user performance was analyzed in chapter 2, without taking into account the overhead introduced by the underlying protocol layers. The results were obtained for GBN, and SR ARQ schemes over LEO and GEO satellite links. The results clearly indicated that LEO with M=3 (i.e., 2 LEO satellites) performed better than GEO with M=2 (i.e., 1 GEO satellite), at low window sizes and low packet sizes. The SR scheme, due to its obvious advantages, performed better than GBN ARQ scheme, which was clearly observed in the results. The end-to-end performance results were obtained by assuming the same P_b value for both LEO and GEO, and still observed better throughput and delay results for VC with M=3 (i.e., 2 LEO satellites). The results were observed to be very sensitive to P_b value. Hence, by using the accurate P_b values of LEO and GEO (i.e., a lesser P_b value for LEO than GEO), significant improvement may be observed with LEO satellites for short and medium range communications.

To extend the advancements of ATM networks to the state-of-art TDMA networks, we analyzed new congestion control schemes (i.e., local and global congestion control schemes) for an ATM multiplexer node in a hybrid ATM/TDMA network. The traditional LB and VLB schemes were improved by *active burst control* to derive modified versions, i.e., mLB, and mVLB. The new global congestion control scheme has provided the *best effort preventive congestion control* to the ATM multiplexer node. The global congestion control scheme not only controlled the traffic from the local transport users, but also prioritized the transit traffic from other ATM multiplexer nodes. The

results of local congestion control scheme in chapter 4, clearly indicated that the ATM multiplexer congestion was well controlled by mLB and mVLB schemes, when compared to the respective LB and VLB schemes. It was also observed from the results that the mLB scheme provided better control of congestion when compared to mVLB scheme, since the mVLB scheme discards the violating cells only when there is congestion. However, it was also clear from the results that the mVLB scheme provides higher throughput when compared to mLB scheme. Hence, the mVLB scheme was considered as an optimum choice compared to mLB scheme. These new congestion control schemes improved the end-to-end throughput and delay performance by reducing the end-to-end re-transmissions for data traffic. Furthermore, the transport user will have a chance to preserve the quality of video and voice information by selectively dropping the traffic during congestion periods.

The three-state global congestion control algorithm was also observed to be very simple for state transitions, which anticipate transit traffic to effectively control the congestion. Similar analysis can also be applied to any other congestion state diagrams with moderately higher number of states, to improve the accuracy of the results. The results of chapter 5 also indicated that the ATM multiplexer congestion was well controlled using LB, VLB, mLB, and mVLB congestion control schemes. It was observed from the results that the mVLB scheme provided optimum end-to-end throughput. Hence, the mVLB scheme is a better compromise for both end-to-end throughput and delay.

Finally, the transport user throughput and delay performance were analyzed with the underlying ATM global congestion control scheme. The results were obtained for an ATM multiplexer node, which is fully aware of the congestion criteria from the other ATM nodes (i.e., $v=c$). The ATM multiplexer node anticipates more transit traffic when $v=c$. In fact, it is hard to achieve complete awareness of the other ATM multiplexer nodes at the same time. The results reflected lower performance, to take into account the transit traffic from other ATM multiplexer nodes, and thereby controlling congestion of the global network. The SR ARQ protocol scheme has provided better performance compared to GBN ARQ protocol scheme. Hence, the high-speed networks can only use SR ARQ protocol scheme, when compared to GBN ARQ protocol.

Hence, the virtual circuit transport user with an underlying mVLB ATM congestion control and an end-to-end SR ARQ protocol provides optimum end-to-end throughput, delay, and global network congestion performance.

6.3 Further research work

The research work in this thesis completely analyzed the throughput and delay performance of the transport user with the underlying ATM global congestion control scheme. Future research can be conducted by introducing error detection, error correction methods [LIM-95], [OHT-91], [DRA-91], [MIL-89], and by improving the assumptions of this research work, some of which are listed below.

* The steady-state congestion criteria were computed within a measurement window (N_{fmax}). However, transient analysis may improve the accuracy of congestion criteria.

* The total buffer space of a service class i, was assumed to be a constant, i.e., L_iB_i. However, it can be varied during system runtime, so that additional buffer space can be added whenever needed based on a tolerable latency. This change can judiciously reduce buffer overflow, and hence can improve system performance.

* The end-to-end window can be varied in a two-dimensional fashion [TA7-94], so that additional credits can be borrowed from other virtual circuits of the system.

* The service classes can be prioritized to provide more throughput and less delay for important traffic.

* An adaptive threshold value can be used to declare congestion, based on the instantaneous congestion status of all service classes of the ATM multiplexer node.

* The total number of connections (i.e., active calls) was assumed to be constant K_i It can be a variable to take advantage of buffer-sharing among several classes of traffic.

* The global congestion states can be increased from 3 to another acceptable number, so that better global congestion control can be achieved.

* A new hybrid mLB/mVLB scheme can be derived as a compromise for throughput and delay performance of the system.

* New congestion communication algorithms can be designed to achieve high-level of awareness of the network.

* Call admission by traffic descriptors to ensure the required QOS, call blocking, and tight-coupling of transport-window to leaky-bucket [REI-95] can be further analyzed with the global congestion control of this thesis.

* The transport user performance analysis can also be derived with the underlying frame relay with variable frame sizes.

* The selective repeat protocol can be modified for further improvements by employing check point control [AHM-89], or by constantly communicating the transport level state information [NET-90] between the end-to-end communicating entities.

* The throughput and delay analysis of this thesis can be applied to other high

performance transport protocols such as XTP [PRG-92], TCP, UDP, VMTP [CHE-86], etc..

* The application of the throughput and delay analysis to LEO/GEO satellite environment can be improved by considering several other important aspects, such as coverage, payload, etc..

* The analysis can be repeated by reflecting the difference in the probability of bit error (P_b) value over LEO and GEO satellite links.

* Reliability of the steady-state results can be compared after analyzing transient results (mentioned above as future work).

* The service primitives may be propagated to the other higher layers to slow-down the user whenever it is necessary to control congestion.

Nonetheless, the end-to-end protocol parameters may still be tuned for further improvements in throughput and delay performance of the transport user.

REFERENCES

[ABB-94] M. Abbas, and Z. A. Ahmad, "Performance analysis of ATM statistical multiplexer with batch arrivals". *IEE Proc. Commun.*, Vol. 141, No. 4, pp 190-195, June 1994.

[AHM-89] Hamid Ahmadi, Parviz Kermani and Phuoc Tran-gia, "Throughput analysis of a class of selective repeat protocols in high speed environments", *IEEE Proc. of GLOBECOM'89* , pp. 930-938, 1989.

[AH2-89] Hamid Ahmadi, and W. E. Denzel, "A survey of modern high-performance switching techniques", *IEEE J. Select. Areas Commun.*, Vol. 7, No. 7, pp. 1091-1103, September 1989.

[AKI-94] Haruo Akimaru, Takashi Okuda, and Kazunori Nagai, "A simplified performance evaluation for bursty multiclass traffic in ATM systems", *IEEE Trans. Comm.*, Vol. 42, No. 5, pp. 2078-2083, May 1994.

[ALL-95] Anthony Alles, "ATM Internetworking", *Engineering InterOP*, Las Vegas, March 1995.

[BAE-91] Jaime Jungok Bae, and Tatsuya Suda, "Survey of traffic control schemes and protocols in ATM networks", *Proc. of the IEEE*, Vol. 79, No. 2., February, 1991.

[BAH-94] Saewoong Bahk, and M. Ei. Zark, "Congestion control based dynamic routing in ATM networks", *J. Computer Commun.*, Vol. 17, No. 12, pp. 825-835, December 1994.

[BER-94] Roberto Beraldi, Antonio Iera, Salvatore Marano, and Roberto Sestito, "A study on the management of on/off loss sensitive sources traffic in ATM networks", *Proc. of ICCS'94*, pp. 530-534, November 1994.

[BHU-90] Brij Bhushan, "Frame relay, fast packet, and packet switching - convergence or coexistence ?", *Telecommun.*, pp. 51-54, Dec. 1990.

[BHU-91] Milena Butto, Elisa Cavallero, and Alberto Tonietti, "Effectiveness of the leaky bucket policing mechanism in ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 3, pp. 335-342, Apr. 1991.

[BON-95] Flavio Bonomi, and Kerry W. Fendick, "The rate-based flow control framework for the available bit rate ATM service", *IEEE Network Mag.*, pp. 25-39, March 1995.

[CHE-86] D. Cheriton, "VMTP:A protocol for the next generation of communication systems", *ACM SIGCOMM Symp.*, pp. 406-415, August 1986.

[CIN-94] Maurizio Cinotti, Enzo Dalle Mese, Stefano Giordano, and Franco Russo, "Long-range dependence in Ethernet traffic offered to interconnected DQDB MANs", *Proc. ICCS'94*, pp. 479-484, November 1994.

[COU-91] J-P. Coudreuse, "General principles of ATM", *l'Ecole des Recherches*, pp. 363-373, 1991.

[DA1-92] *DARPA technical report NXS.TR.003.01*, "ATM network routing techniques, ATM flow control techniques", July 1992.

[DA2-93] *DARPA technical report NXS.TR.006.01*, "ATM network routing techniques, ATM flow control techniques", April 1993.

[DA3-93] *DARPA technical report NXS.TR.007.01*, "ATM network routing techniques, ATM flow control techniques", July 1993.

[DAY-83] J. D. Day and H. Zimmermann, "The OSI reference model", *Proc. of IEEE*, Vol. 71, No.12, December 1983.

[DEL-95] Enrico Del Re, Romano Fantacci, and Giovanni Giambene, "Efficient dynamic channel allocation technique with handover queuing for mobile satellite networks", *IEEE J. Select. Areas Commun.*, Vol. 13, No. 2, pp. 397-405, February 1995.

[DIT-91] Lars Dittmann, Soren B. Jacobson, and Klaus Moth, "Flow enforcement algorithms for ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 3, pp. 343-350, April 1991.

[DRA-91] Subrahmanyam Dravida, and Ravi Damodaram, "Error detection and correction options for data services in B-ISDN", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 9, pp. 1484-1494, December 1991.

[ECK-91] Adrian E. Eckberg, Bharat T. Doshi, and Richard Zoccolillo, "Controlling congestion in B-ISDN / ATM : Issues and Strategies", *IEEE Commun. Mag.*, pp. 64-69, September 1991.

[ELH-94] A. K. Elhakeem, R. Girolamo, L. B. Bdira, and M. Talla, "Delay and throughput characteristics of TH, CDMA, TDMA, and Hybrid networks for multipath faded data transmission channels", *IEEE J. Select Areas Commun.*, Vol. 12, No. 4, pp. 622-637, May 1994.

[FAN-90] Romano Fantacci, "Performance evaluation of efficient continuous ARQ protocols", *IEEE Trans. Commun.*, Vol. 38, No. 6, June 1990.

[FOR-94] Samuel W. Fordyce, and Willium W. Wu, "The myth and reality of mobile satellite communications", *Proc. ICCS'94*, pp. 393-399, November 1994.

[GAL-89] G. Gallassi, G. Rigolio, and L. Fratta, "ATM : bandwidth assignment and bandwidth enforcement policies", *Proc. IEEE GLOBECOM'89*, pp. 1788-1793, 1989.

[GOU-92] M. G. Gouda, "The two-dimensional window protocol", 12th ISPSTV international symposium, Florida, U.S.A., June 1992.

[GRU-91] R. Grunenfelder, J. P. Cosmos, S. Manthorpe, and Odinma-Okafor, "Characterization of video codecs as autoregressive moving average processes and related queueing system performance", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 3, pp. 284-293, 1991.

[GUS-90] R. Gusella, "A measurement study of diskless workstation traffic on an Ethernet", *IEEE Trans. Commun.*, Vol. 38, No. 9, pp. 1557-1568, 1990.

[HAB-92] Ibrahim W. Habib, and Tarek N. Saadawi, "Multimedia traffic characteristics in broadband networks", *IEEE Communication Mag.*, pp. 48-54, July 1992.

[HAS-92] Yukio Hashimoto, Ryutaro Suzuki, Ryu Miura, Kimio Kondo, Tomoki Obuchi et al, "A study on a simple communication network by microsats", *American Institute of Aeronautics and Astronautics*, AIAA-92-1928-CP, pp. 896-902, 1992.

[HON-91] Duke Hong, Tatsuya Suda, and Jaime Jungok, "Survey of techniques for prevention and control of congestion in an ATM network", *Proc. of ICC'91*, pp. 204-210, 1991.

[HO2-91] Duke Hong, and Tatsuya Suda, "Congestion control and prevention in ATM networks", *IEEE Network Magazine*, pp. 10-16, July 1991.

[HUA-94] T-Y. Huang., and J. - L. C. Wu, "Performance analysis of ATM switches using priority schemes", *IEE Proc. Commun.*, Vol. 141, No. 4, pp. 248-254, August, 1994.

[I121-88] CCITT recomm. I.121, "Broadband aspects of ISDN", Melbourne, 1988.

[I321-91] CCITT recomm. I.321, "BISDN protocol reference model and its applications", Geneva, 1988.

[ILA-95] Andy Ilana, "Satellites : going wireless among the stars", *Canadian Telecom Mag.*, pp. 22-25, May 1995.

[INU-92] Thomas Inukai, Faris Faris, and Dong-Jye Shyy, "On-board processing satellite network architectures for broadband ISDN", *American Institute of Aeronautics and Astronautics, AIAA-92-2003-CP*, pp. 1471-1484, 1992.

[JAD-94] T. M. Jadoon, and D. A. Harle, "Performance comparison of peak bit rate and bursty traffic specification and allocation policier functions", *Proc. of ICCS'94*, pp. 525-529, November 1994.

[JAI-86] Raj Jain, and Shawn A. Routhier, "Packet trains-measurements and a new model for computer network traffic", *IEEE J. Select. Areas Commun.*, Vol. SAC-4, No. 6, pp. 986-995, April 1991.

[JA2-86] Raj Jain, "A timeout based congestion control scheme for window flow-controlled networks", *IEEE J. Select. Areas Commun.*, Vol. SAC-4, No. 7, pp. 1162-1167, October 1986.

[KAD-95] Michel Kadoch, "ATM signaling : a tutorial", *Proc. of IEEE Canadian conference on Elec. & Computer Eng., Montreal*, pp. 421-424, September 1995.

[KAM-89] Takashi Kamitake, and Tatsuya Suda, "Evaluation of an admission control scheme for an ATM network considering fluctuations in cell loss rate", *Proc. GLOBECOM'89*, pp. 1774-1780, 1989.

[KAN-92] Jaidev Kaniyil, Jun Takei, Shigeru Shimamoto, Yoshikuni Onozato et al., "Preliminary aspects of a global message network employing low earth-orbiting satellites", *American Institute of Aeronautics and Astronauts, AIAA-92-1926-CP*, pp. 876-884, 1992.

[KAW-91] Masatoshi Kawarasaki, Bijan Jabbari, "B-ISDN architecture and protocol", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 9, pp. 1405-1414, December 1991.

[KAW-95] Kenji Kawahara, Yuji Oie, Masayuki Murata, and Hideo Miyahara, "Performance analysis of reactive congestion control for ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 13, No. 4, pp. 651-661, May 1995.

[KUN-95] H. T. Kung, and Robert Morris, "Credit based flow control for ATM networks", *IEEE Network*, pp. 40-48, March/April 1995.

[KWA-92] R. K. Kwan, and K. M. Price, D. M. Chitre, L. W. White, and T. R. Henderson, "Satellite delivery of B-ISDN services", *American Institute of Aeronautics and Astronautics, AIAA-92-1831-CP*, pp. 228-239, 1992.

[LEL-89] Will E. Leland, "Window-based congestion management in broadband ATM networks: The performance of three access-control policies", *IEEE GLOBECOM'89*, pp. 1794-1800, 1989.

[LEI-92] T. C. Leisgang, and M. D. Benz, "High power lightstats for low earth orbit applications", *American Institute of Aeronautics and Astronautics, AIAA-92-1952-CP*, pp. 1065-1073, 1992.

[LIM-95] Hyo Taek Lim, and Joo Seok Song, "Cell loss recovery method in B-ISDN/ATM networks", *Electronics letters*, Vol. 31, No. 11, pp. 849-851, 25th May, 1995.

[LIN-91] Arthur Y.- M. Lin, and John Silvester, "Priority queueing strategies and buffer allocation protocols for traffic control at an ATM integrated

broadband switching system", *IEEE Select. Areas Commun.*, Vol. 9, No. 9, December 1991.

[LIU-95] Zhen Liu, and Don Towsley, "Burst reduction properties of rate-control throttles : Dwonstream queue bahavior", *IEEE/ACM Trans. Networking*, Vol. 3, No. 1, pp. 82-90, February 1995.

[MAG-88] Basil Maglaris, Dimitris Anastassiou, Prodip Sen, Gunnar Karlson, and John Robbins, "Performance analysis of statistical multiplexing for packet video sources", *IEEE Trans. Commun.*, Vol. 36, No. 7, pp. 834-844, 1988.

[MAH-87] Christoph E. Mahle, Geoffrey Hyde, and Thomas Inukai, "Satellite scenarios and technology for the 1990's", *IEEE J. Select. Areas Commun.*, Vol. SAC-5, No. 4, pp. 556-570, May 1987.

[MAR-94] Nasser M. Marafih, Ya-Qin Zhang, and Raymond L. Pickholtz, "Modeling and queueing analysis of variable-bit-rate coded video sources in ATM networks", *IEEE Trans. circuits and systems for video technology*, Vol. 4, No. 2, April 1994.

[MIL-81] M. J. Miller, and S. Lin, "The analysis of some selective repeat ARQ schemes with finite receiver buffer", *IEEE Trans. Commun.*, Vol COM-29, pp. 1307-1315, September 1981

[MIL-89] Michael J. Miller, "Error control techniques for integrated services packet networks", *IEEE J. Select. Areas Commun.*, Vol. 7, No. 5, pp. 690-697, June 1989.

[MIN-89] Steven E. Minzer, "Broadband ISDN and asynchronous transfer mode (ATM)", *IEEE Commun. Mag.*, pp. 17-24, September 1989.

[MIY-91] Hideo Miyahara, "ATM : A most promising switching technique for B-ISDN", *ICICE Trans. Commun.* , Vol. E-74, No. 4, April 1991.

[MUR-90] Masayuki Murata, Yuji Oie, Tatsuya Suda, and Hideo Miyahara, "Analysis of a discrete-time single-server queue with bursty inputs for traffic control in ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 8, No. 3, pp. 447-453, April 1990.

[MUR-91] Tutomu Murase, Hiroshi Suzuki, Shohei Sato, and Takao Takeuchi, "A call admission control scheme for ATM networks using a simple quality estimate", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 9, December 1991.

[NET-90] Arun N. Netravali, W. D. Roome, and K. Sabnani, "Design and implementation of high-speed transport protocol", *IEEE Trans. Commun.*, Vol. 38, No. 11, November 1990.

[NEW-94] P. Newman, "Traffic management for ATM local area networks", *IEEE Commun. Mag.*, Vol. 32, No. 8, pp. 44-50, Aug. 1994.

[NOH-93] Mitsuo Nohara, Yoshinori Arimoto, Wataru Chujo, and Masayuki Fujise, "A link study of a low-earth orbit satellite communications system using optical intersatellite links", *IEICE Trans. Commun.*, Vol. E76-B, No. 5, May 1993.

[OGI-95] Nagao Ogino, and Yasushi Wakahara, "Application of neural network in ATM call admission control based on cell transfer state monitoring with dynamic threshold", *IEICE Trans. Commun.*, Vol. E78-B, No. 4, pp. 465-475, April 1995.

[OHN-88] Hirokazu Ohnishi, Tadanobu Okada, and Kiyohiro Noguchi, "Flow control schemes and delay/loss tradeoff in ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 6, No. 9, pp. 1609-1616, December 1988.

[OHT-91] Hiroshi Ohta, and Tokuhiro Kitami, "A cell loss recovery method using FEC in ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 9, pp. 1471-1483, December 1991.

[PAR-95] Hong-Shik Park, Dong-Yong Kwak, Woo-Seop Rhee, Man-Yeong Jeon, and Jae-kyoon Kim, "Global traffic control in ATM networks", *IEICE Trans. Commun.*, Vol. E-78-B, No. 4, pp. 476-484, April 1995.

[PAX-95] Vern Paxson, and Sally Floyd, "Wide area traffic: The failure of Poisson modeling", *IEEE/ACM Trans. Networking*, pp. 226-244, June 1995.

[PRI-89] F. Delli Priscoli, M. Listanti, A. Roveri, and A. Vernucci, "Access and switching techniques in an ATM user-oriented satellite system", *Proc. of IEEE ICC'89*, pp. 632-640, 1989.

[PRO-92] "XTP Protocol definition", Protocols Engineers Inc., rev. 3.6, January 1992.

[PRY-88] Martin De Prycker, "Definition for the Belgian ATM broadband experiment", *IEEE J. Select. Areas Commun.*, Vol. 6, No. 9, pp. 1538-1544, December 1988.

[RAH-91] Moe Rahnema, "Frame relaying and fast packet switching concepts and issues", *IEEE Network Magazine*, pp. 18-23, July 1991.

[RAM-95] K.K. Ramakrishnan, and Peter Newman, "Integration of rate and credit schemes for ATM flow control", *IEEE Network*, pp. 49-56, March/April 1995.

[RAS-91] Carsten Rasmussen, Jens H. Sorensen, Kenn S. Kvols, and Soren B. Jacobson, "Source-independent call acceptance procedures in ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 3, pp. 351-358, April 1991.

[RAT-91] Erwin P. Rathgeb, "Modeling and performance comparison of policing mechanisms for ATM networks", *IEEE J. Select. Areas Commun.*, Vol. 9, No. 3, pp. 325-334, April 1991.

[REI-95] Amy R. Reibman, and Arthur W. Berger, "Traffic descriptors for video teleconferencing over ATM networks", *IEEE/ACM Trans. Networking*, Vol. 3, No. 3, pp. 329-339, June 1995.

[ROB-91] James W. Roberts, "Variable-bit-rate traffic control in B-ISDN", *IEEE Communications Magazine*, pp. 50-55, September 1991.

[ROB-95] James W. Roberts, "Rate envelope multiplexing and rate sharing in B-ISDN", *IEICE Trans. Commun.*, Vol. E78-B, No. 4, pp. 431-438, April 1995.

[ROM-95] Allyn Romanow, and Sally Floyd, "Dynamics of TCP traffic over ATM Networks", *IEEE J. Select. Areas Commun.*, Vol. 13, No. 4, pp. 633-641, May 1995.

[ROU-94] J. K. Roussos, E. G. Economou, and P. Georgiadis, "Congestion avoidance in high speed interconnected voice data traffic LANs", *J. Computer Commun.*, Vol. 17, No. 4, pp. 242-250, April 1994.

[RUB-94] Izhak Rubin, K. David Lin, "Input rate flow control for high speed communications networks using burst level feedback control", *European Trans. Telecommun.*, Vol. 5, No. 1, pp. 107-123, Jan-Feb. 1994.

[SAA-94] T. Saadawi, M. Ammar, and A. K. Elhakeem, *"Fundamentals of Telecommun. Networks"*, Wiley Back, 1994.

[SAI-91] H. Saito, M. Kawarasaki, and H. Yamada, "An analysis of statistical multiplexing in an ATM transport network", *IEEE J. Select Areas Commun.*, Vol. 9, No. 3, pp. 359-367, 1991.

[SA2-91] Hiroshi Saito, Konosuke Kawashima, and Ken-ichi Sato, "Traffic control technologies in ATM networks", *IEICE Trans.*, Vol. E-74, No. 4, pp. 761-771, April 1991.

[SAI-94] Hiroshi Saito, *"Teletraffic Technologies in ATM Networks"*, Artech House, 1994.

[SCH-92] M. Schwartz, *"Telecommun. Networks : Protocols, Modelling and Analysis"* , Addison-Wesley publishing company, 1992.

[SEN-89] Prodip Sen, Basil Maglaris, Nasser-Eddine Rikli, and Dimitris Anastassiou, "Models for packet switching of variable-bit-rate video sources", *IEEE J. Select. Areas Commun.* , Vol. 7, No. 5, pp. 865-869, June 1989.

[SHI-95] Kohei Shiomoto, and Shin-ichiro Chaki, "Adaptive connection admission control using real-time traffic measurements in ATM networks", *IEICE Trans. Commun.*, Vol. E78-B, No. 4, pp. 458-463, April 1995.

[SRI-83] Kotikalapudi Sriram, Pramod K. Varshney, and J. George Shanthikumar, "Discrete-time analysis of integrated voice/data multiplexers with and without speech activity detectors", *IEEE J. Select. Areas Commun.*, Vol. SAC-1, No. 6, pp. 1124-1130, December 1983.

[SRI-86] Kotikalapudi Sriram, and Ward Whitt, "Characterizing superposition arrival processes in packet multiplexers for voice and data", *IEEE J. Select. Areas Commun.*, Vol. SAC-4, No. 6, pp. 833-845, September 1986.

[STA-94] William Stallings, "*Data and computer communications*", Fourth edition, Macmillan publishing company, 1994.

[ST2-94] Georgios M. Stamatelos, and Jeremiah F. Hayes, "Admission-control techniques with application to broadband networks", *J. Computer Commun.*, Vol. 17, No. 9, pp. 663-673, September 1994.

[ST3-94] G. D. Stamoulis, M. E. Anagnostou, and A. D. Georgantas, "Traffic source models for ATM networks : a survey", *J. Computer Commun.*, Vol. 17, No. 6, pp. 428-438, June 1994.

[SUB-95] Selvakumaran N. Subrahmanian, and Tho Le-Ngoc, "Traffic modeling in a Multi-media environment", *IEEE Canadian conf. on elec. & computer eng.*, pp. 838-841, September 1995.

[SUN-94] W. Sun, M. N. Sweeting, and M. S. Hodgart, "In-orbit channel measurement and equalization for low earth orbit communications using digital signal processing techniques on the POSAT-1 microsatellite", *Proc. IEEE ICCS*, pp. 415-419, November 1994.

[SYK-91] E. D. Sykas, K. M. Vlakos, and M. J. Hillyard, "Overview of ATM networks functions and procedures", *J. Computer Commun.*, Vol. 14, No. 10, December 1991.

[SYK-94] E. D. Sykas, K. M. Vlakos, K. P. Tsoukatos, and E. N. Protonotarios, "Performance evaluation of analytical models for effective bandwidth allocation in ATM networks", *European Trans. Telecommun.*, Vol. 5, No. 3, pp. 391-396, May-June 1994.

[TAL-94] M. Talla, A. K. Elhakeem, and M. Kadoch, "Throughput and delay characteristics of ARQ transport protocols over LEO and GEO satellite networks", *Proc. IEEE ICCS*, pp. 410-414, January 1994.

[TA2-95] M. Talla, A. K. Elhakeem, and M. Kadoch, "A Novel congestion control in hybrid ATM/TDMA networks", *accepted* for publication by *Int. J. Computers and Elect. Eng.*, 1995.

[TA3-96] M. Talla, A. K. Elhakeem, and M. Kadoch, "A Global congestion control for hybrid ATM/TDMA networks", *submitted to IEEE ICC'96.*

[TA4-95] M. Talla, A. K. Elhakeem, and M. Kadoch, "A new congestion control scheme for hybrid ATM/TDMA networks", *IEEE GLOBECOM*, 1995.

[TA5-95] M. Talla, A. K. Elhakeem, and M. Kadoch, "Quality of service based congestion control scheme in interconnected ATM/TDMA networks", *submitted* to *European Trans. Commun., 1995.*

[TA6-95] M. Talla, A. K. Elhakeem, and M. Kadoch, "Optimal protocol parameters of ARQ schemes over frame relay LEO and GEO satellite networks", *Proc. IEEE conference on Computers in Elect. Eng.*, Montreal, pp. 137-140, September 1995.

[TA7-94] M. Talla, R. Ahooja, and A. K. Elhakeem, "The two dimensional window protocol for high speed networks", *Proc. IEE ICT'94*, pp. 148-150, January 1994.

[TA8-95] M. Talla, A. K. Elhakeem, and M. Kadoch, "Throughput and delay performance of transport user in congestion controlled hybrid ATM/TDMA networks", *to be submitted* to a journal.

[TSE-91] Kent H. Tseng, and Man-Tung Hsiao, "Admission control of voice / data integration in an ATM network", *IEEE Proc. of ICC'91*, pp. 188-192, 1991.

[WAR-95] Christopher Ward et. al., "A data link control protocol for LEO satellite networks providing a reliable datagram service", *IEEE/ACM Trans. Networking*, Vol. 3, No. 1, February 1995.

[WEL-82] E. J. Weldon, "An improved selective repeat ARQ strategy", *IEEE Trans. Commun.*, Vol. COM-30, pp. 480-486, March 1982.

[WER-92] Marek Wernik, Osama Aboul-Magd, and Henry Gilbert, "Traffic management for B-ISDN services", *IEEE Network Mag.*, pp. 10-19, September 1992.

[WER-95] Markus Werner, Axel Jahn, Erich Lutz, and Axel Bottcher, "Analysis of system parameters for LEO/ICO-satellite communication networks", *IEEE J. Select. Areas Commun.*, Vol. 13, No. 2, pp. 371-381, December 1995.

[WHI-92] B. E. White, R. G. Bland, C-H. Chen, and R. A. Pomponi, "Low earth orbit satellite concepts for air traffic control applications", *American Institute of Aeronautics and Astronautics*, AIAA-92-1927-CP, pp. 1065-1073, 1992.

[WID-95] I. Widjaja, "Performance analysis of burst admission-control protocols", *IEE Proc. Commun.*, Vol. 142, No. 1, pp. 7-14, February 1995.

[WU-87] William W. Wu, David Haccoun, Robert Peile, and Yasuo Hirata, "Coding for satellite communication", *IEEE J. Select Areas Commun.*, Vol. SAC-5, No. 4, May 1987.

[YAM-95] Naoaki Yamanaka, Toyofumi Takenaka, Youichi Sato, and Ken-ichi Sato, "Jitter tolerant usage parameter control method for ATM-based B-ISDN", *IEICE Trans. Commun.*, Vol. E78-B, No. 4, pp. 485-493, April 1995.

[YEG-93] Ferit Yegenoglu, Bijan Jabbari, and Ya-Qin Zhang, "Motion-classified autoregressive modeling of variable bit rate video", *IEEE Trans. on Circuits Systems for video technology*, Vol. 3, No. 1, pp. 42-53, February 1993[WAN-95] Chia-Jiu Wang, "Delivery time analysis of a low earth orbit satellite network for seamless PCS", *IEEE J. Select. Areas Commun.*, Vol. 13, No. 2, pp. 389-396, February 1995.

[YOK-95] Tetsuya Yokotani, Tatsuki Ichihashi, Chikara Matsuda, and Michihiro Ishizaka, "Selectable traffic control scheme for burst data transmission using TCP/IP on ATM networks", *IEICE Trans. Commun.*, Vol. E78-B, No. 4, pp. 531-538, April 1995.

[ZHO-90] Xiao You Zhou, and Ahmed E. Kamal, "Automatic repeat-request protocols and their queuing analysis", *Computer Commun. Mag.*, Vol. 13, No. 5, pp. 298-311, June 1990.

[ZIM-80] H. Zimmermann, "OSI reference model: The ISO architecture for open systems interconnections", *IEEE Trans. Commun.*, Vol. COM-28, No. 4, pp. 425-432, April 1980.