Multipath Routing and Quality of Service Support for Mobile Ad hoc Networks

Mohammad Haseeb Zafar

A thesis submitted for the degree of Doctor of Philosophy to the
Centre for Intelligent Dynamic Communications
Department of Electronic and Electrical Engineering
University of Strathclyde, Glasgow, UK

July 2009

Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has lead to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50. Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Mohammad Haseeb Zafar 10 July 2009

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible. I would like to begin with thanking my supervisor Dr. David Harle for giving me an excellent opportunity to carry out a very innovating and challenging research on mobile ad hoc networks. Dr. Harle provided me endless source of ideas and encouragement. His enthusiasm in research, systematic organisation in work and optimistic attitude towards life positively influenced my study and life. His assistance and instructions during my time at University of Strathclyde has been invaluable.

I wish to express my warm and sincere thanks to Professor Ivan Andonovic for being my second supervisor and actively giving me enthusiastic encouragement, guidance and sincere support throughout the entire PhD. I am indebted to all my colleagues at University of Strathclyde, especially Dr. Christos Tachtatzis for his feedback, advice and friendship.

I would like to extend my sincere gratitude to my employer and sponsor N-W.F.P. University of Engineering & Technology, Peshawar, Pakistan for giving me the opportunity to pursue my PhD in University's joint Research programme with University of Strathclyde.

My deepest and sincere thanks are always to my parents, parent-in-laws, brothers, sisters, brother-in-laws and sister-in-laws for their constant support and prayers. I am so lucky and so proud to have such a wonderful family.

Lastly, and most importantly, I owe my loving thanks to my wife Nazish and daughter Mahum without their understanding, endless patience and encouragement it would have been impossible for me to finish this work.

Abstract

Mobile Ad hoc Networks (MANETs) are a key part of the ongoing evolution of wireless communications. MANETs are a collection of wireless mobile nodes that dynamically form a temporary wireless network without an infrastructure. The design of an efficient and reliable routing scheme and Quality of Service (QoS) support for MANETs is a major challenge. Unlike traditional routing schemes that seek only single path, multipath routing allows the establishment of multiple paths for routing between a source-destination pair. Multipath routing exploits the resource redundancy and diversity in the underlying network to provide benefits such as fault tolerance, load balancing, capacity aggregation and the improvement in QoS metrics such as delay. In the first part of the thesis, a multipath routing scheme, referred to as Shortest Multipath Source (SMS) routing based on Dynamic Source Routing (DSR) protocol is proposed. The mechanism has two novel aspects compared to other on-demand multipath routing schemes: it achieves shorter multiple partial-disjoint paths and allows more rapid recovery from route breaks. This scheme addresses the problem of wireless broadcast storms by simple hop count mechanism. The performance differentials are investigated using Network Simulator version 2 (NS-2). Results show the superiority of SMS under certain scenarios in terms of goodput of up to 85% and end-to-end delay of up to 99% when compared to the competing schemes. Although SMS is designed to find multiple shorter routes, these routes have no information about the network traffic or application requirements. The second aspect of the thesis addresses QoS support. Two novel capacity-constrained routing schemes based on SMS, which allow nodes to depend on their estimation of the residual capacity to make correct admission control decisions, are presented. The performance evaluation demonstrates the merits of the proposed schemes with a 20% increase in goodput while end-to-end delay is reduced by 47% and the necessity of QoS-aware multipath routing schemes in Mobile Ad hoc Networks becomes more apparent.

Publications

The following lists the author's publications in reverse chronological order.

Journal Publication

[J1] H. Zafar, D. Harle, I. Andonovic, Y. Khawaja, "Performance Evaluation of Shortest Multipath Source Routing Scheme", *IET Communications in Special Issue on Wireless Ad hoc Networks*, 3, (5), May 2009, pp. 700–713. (Featured Paper)

Conference Publications

[C4] H. Zafar, D. Harle, I. Andonovic, M. Ashraf, "SMS: Shortest Multipath Source Routing for Mobile Ad hoc Networks", *Proc. 2007 IEEE International Conference on Signal Processing and Communications (ICSPC)*, Dubai, UAE, November 2007, pp. 97–100.

[C3] H. Zafar, D. Harle, I. Andonovic, M. Ashraf, "Partial-Disjoint Multipath Routing for Wireless Ad hoc Networks", *Proc. 32nd IEEE International Conference on Local Computer Networks (LCN)*, Dublin, Ireland, October 2007, pp. 258–259.

[C2] H. Zafar, D. Harle, I. Andonovic, M. Ashraf, "On-demand Partial-Disjoint Multipath Routing with Low Delay for Wireless Ad hoc Networks", *Proc. 8th Annual Postgraduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting (PGNet)*, Liverpool, UK, June 2007, pp. 233–236.

[C1] H. Zafar, D. Harle, I. Andonovic, M. Ashraf, "Performance Evaluation of On-demand Multipath Routing Protocols for Mobile Ad hoc Networks", *Proc. Seventh IASTED International Conference of Wireless & Optical Communications (WOC)*, Montreal, Canada, May 2007, pp. 325–329.

Contents

List of Figures

List of Tables

Abbreviations

ACK	Acknowledgement
AIMD	Additive Increase Multiplicative Decrease
AODV	Ad hoc On-demand Distance Vector
AOMDV	Ad hoc On-demand Multipath Distance Vector
API	Application Programming Interface
AQOR	Ad hoc QoS On-demand Routing
AQR	Adaptive QoS Routing
ARP	Address Resolution Protocol
CBR	Constant Bit Rate
CBRP	Cluster Based Routing Protocol
CDMA	Code Division Multiple Access
CEDAR	Core Extraction Distributed Ad hoc Routing
CRC	Cyclic Redundancy Checksum
CS	Carrier Sensing
CSMA	Carrier Sense Multiple Access
CSMA/CA	CSMA with Collision Avoidance
CSMA/CD	CSMA with Collision Detection
CTS	Clear To Send
CW	Contention Window
DBPSK	Differential Binary Phase Shift Keying
DCF	Distributed Coordination Function
DDD	Data Display Debugger
Diffserv	Differentiated Services
DIFS	DCF Inter Frame Spacing
DMQR	Disjoint Multipath QoS Routing
DoS	Denial-of-Service
DQPSK	Differential Quadrature Phase Shift Keying
DSDV	Destination Sequenced Distance Vector

DSR	Dynamic Source Routing
DSSS	Direct Sequence Spread Spectrum
DYMO	Dynamic MANET On-demand Routing
ECN	Explicit Congestion Notification
ETSI	European Telecommunication Standard Institute
FCS	Frame Check Sequence
FDMA	Frequency Division Multiple Access
FHSS	Frequency Hopping Spread Spectrum
FIFO	First-In First-Out
FQMM	Flexible QoS Model for MANETs
FTP	File Transfer Protocol
GDB	GNU Debugger
GFSK	Gaussian Frequency Shift Keying
GloMo	Global Mobile Information Systems
GloMoSim	Global Mobile Information Systems Simulation Library
GSR	Global State Routing
HRDSSS	High Rate Direct Sequence Spread Spectrum
HTTP	Hypertext Transfer Protocol
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IFQ	InterFace Queue
IFS	Inter Frame Spacing
INSIGNIA	In-band Signalling for QoS in Ad hoc mobile networks
IntServ	Integrated Services
IP	Internet Protocol
IR	Infrared
ISM	Industrial Scientific Medical
LAN	Local Area Network
LBHBF	Largest Bandwidth-Hop-Bandwidth First
LL	Link Layer
LLC	Logical Link Control
MAC	Medium Access Control

MAN	Metropolitan Area Network
MANET	Mobile Ad hoc NETwork
M-DSR	Multipath Dynamic Source Routing
MPR	Multi Point Relay
MSDU	MAC Service Data Unit
MSR	Multipath Source Routing
NAM	Network Animator
NAV	Network Allocation Vector
NDMR	Node-Disjoint Multipath Routing
NetIF	Network InterFace
NS-2	Network Simulator 2
NTDR	Near-term Digital Radio
OLMQR	On-demand Link State Multipath QoS Routing
OLSR	Optimised Link State Routing
OMNeT++	Objective Modular Network Testbed in C++
OPNET	Optimized Network Engineering Tool
OQR	On-demand QoS Routing
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First
OTcl	Object Tool command language
PAN	Personal Area Network
PCF	Point Coordination Function
PDA	Personal Digital Assistant
PDU	Protocol Data Unit
PHY	Physical
PPM	Pulse Position Modulation
PRNet	Packet Radio Networks
Q-DSR	QoS-enabled Dynamic Source Routing
QoS	Quality of Service
QRBE	QoS-aware Routing based on Bandwidth Estimation
QRP	QoS Routing Protocol
QRREP	QoS Route Reply

QRREQ	QoS Route Request
Q-SMS	QoS-aware Shortest Multipath Source routing
RAN	Regional Area Network
RERR	Route Error
RFC	Request For Comments
RIP	Routing Information Protocol
RREP	Route Reply
RREQ	Route Request
RTP	Real-time Transport Protocol
RTS	Request To Send
SDMA	Space Division Multiple Access
SIFS	Short Inter Frame Spacing
SMORT	Scalable Multipath On-demand RouTing
SMR	Split Multipath Routing
SMS	Shortest Multipath Source routing
SMTP	Simple Mail Transfer Protocol
SURAN	Survivable Adaptive Radio Networks
SWAN	Stateless Wireless Ad hoc Networks
TBR	Ticket-Based QoS Routing
TCP	Transmission Control Protocol
TDMA	Time Division Multiple Access
TORA	Temporally Ordered Routing Algorithm
TTL	Time To Live
VBR	Variable Bit Rate
WAN	Wide Area Network
WIFS	Waiting for Interframe Spacing
WiMax	Worldwide Inter-operability for Microwave Access
WLAN	Wireless Local Area Network
WRP	Wireless Routing Protocol
ZRP	Zone Routing Protocol

1. Introduction

Current technology is characterised by an increasing need for mobile computing and communication devices (e.g. cell phones, laptops, handheld digital devices and personal digital assistants). The traditional wired networks cannot accommodate the mobility demand, which has led to the introduction of a considerable number of wireless networking technologies such as cellular networks [1], Wireless Local Area Networks (WLANs) [2], Bluetooth networks [3], Ultra-Wideband (UWB) networks [4], Mobile Ad hoc Networks (MANETs) [5] and WiMax [6]. Among these, cellular networks, Bluetooth networks and WLANs are the most widely used. Cellular networks and WLANs are based on single-hop communications, i.e. nodes[1] participating in the network must be in direct communication with a base station or an access point respectively. This fact implies a restriction on the networks, in that they depend on a centralised infrastructure. Using Bluetooth technology, hosts can connect to each other in an ad hoc fashion, but this technology is only targeted at low power short-range wire replacement. Therefore, a further development in wireless networking is to allow participating nodes to autonomously configure themselves and relay traffic for other nodes, thus enabling wireless multi-hop communications. Such networks are known as Mobile Ad hoc Networks (MANETs) and have obtained significant attention in recent years.

1.1 Mobile Ad hoc Networks

Mobile Ad hoc Networks (MANETs) are a key part of the ongoing evolution of wireless communications. In contrast to the traditional infrastructure based cellular systems, ad hoc networks comprise of mobile/semi-mobile nodes that do not rely on infrastructure

---

[1] The terms node, host, peer are used interchangeably.

and are free to move, appear and disappear randomly. Each node has the capability to communicate directly with other nodes, acting not only as a mobile wireless host but also as router, forwarding data packets for other nodes. In other words, ad hoc networks are self-creating, self-organising and self-administrating multi-hop wireless networks, with a dynamically changing topology. This paradigm is depicted in Figure 1.1. There are eight mobile nodes taking part in an ad hoc network. The circles around nodes show the current transmission radius of a given node. It is obvious that if two nodes lie in the transmission radii of each other they can communicate.

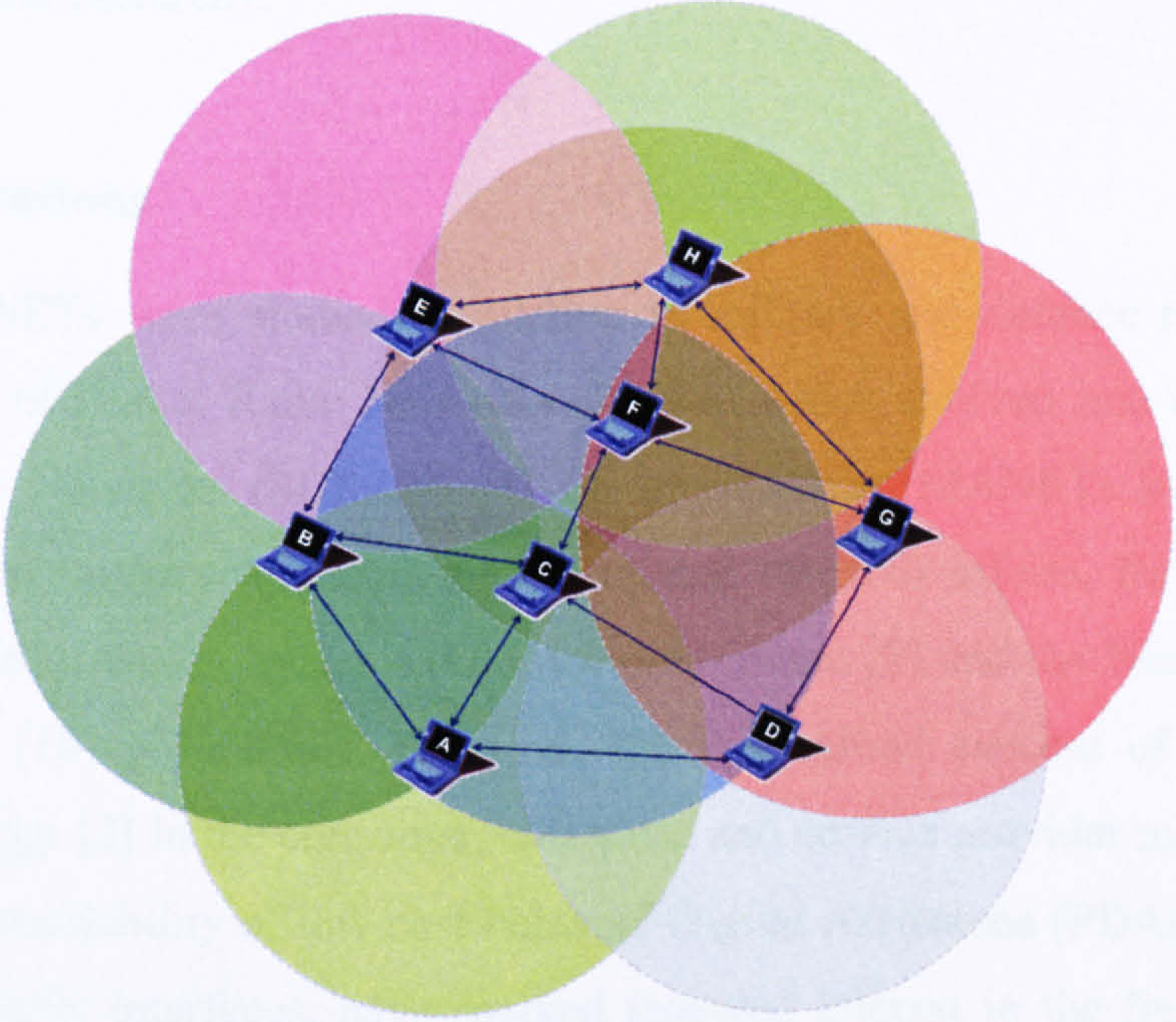


Figure 1.1 Illustration of an ad hoc network

Since ad hoc networking has been deemed as one of the most vibrant and actively evolving fields today, the IETF MANET working group [5] was chartered in 1997, to discuss and develop solutions in this area. According to S. Corson et al [7], “A MANET consists of mobile platforms (e.g. a router with multiple hosts and wireless communication devices) – herein simply referred to as ’nodes’ – which are free to move about randomly”. Nodes in ad hoc networks are often mobile, but can also consist of stationary nodes, such as access points to the Internet. In the later operational mode, it is typically envisioned to operate as a 'stub' network connecting to a fixed inter-network. Stub networks carry traffic originating at and/or destined for internal nodes, but do not permit exogenous traffic to 'transmit' through the stub network.

With the above definition, MANETs are most likely to be deployed, in the real world, as stub networks connecting to fixed infrastructure networks. In such cases, the network as a whole consists of two distinct categories of nodes: infrastructure nodes and client nodes. The former's principal function is to aggregate and transport the traffic for the client nodes in the network (e.g. access points). The latter represent the application users of the network. In other words, instead of the flat hierarchy of pure ad hoc networks there is a two level hierarchy.

1.1.1 Applications

Originally MANETs were studied in relation to military and defence research, often under the name of Packet Radio Networks (PRNet) which evolved into the Survivable Adaptive Radio Networks (SURAN) [8]. In this context, MANETs have played an important role in military applications and related research efforts, for example, the Global Mobile Information Systems (GloMo) programme [9] and the Near-term Digital Radio (NTDR) [10] programme. However, the widespread success of IEEE 802.11 WLAN technology [2] in the consumer, enterprise and service provider markets, as well as the common availability of low cost Personal Digital Assistance (PDAs), laptops and palmtops with radio interfaces, have sparked renewed interest in the field. Internet or intranet connectivity, therefore, are significant factors to be taken into account in the utilisation of ad hoc network technology. Therefore, the IETF MANET working group's main task is to develop a framework of IP-based routing protocols for ad hoc networks.

There are a growing number of real time applications using wireless ad hoc and sensor networks, and they are being taken seriously by the industries. Some of the potential applications of ad hoc networks that might provide the basis for commercially successful products are:

- *Conferencing:* Perhaps the prototypical application requiring the establishment of an ad hoc network is mobile conferencing enabling notebook or palmtop for spreading or sharing information among participants in a conference.
- *Home Networking:* It might be possible to deploy ad hoc technology to enable direct communication between devices at home. This would make possible the exchange of information such as voice, video-alarms and configuration updates.
- *Internet Hot Spots:* Ad hoc networks can be linked to a fixed infrastructure via access points to provide extended wireless Internet access.
- *Personal Area Networks:* Short-range ad hoc networks can be formed to simplify intercommunication between various mobile devices (such as a cellular phone and a laptop) by forming a personal area network (PAN).
- *Emergency Services:* Ad hoc networks can help to overcome network impairment during disaster emergencies. Mobile units will probably carry networking equipment in support of routine operations for the times when the Internet is available and the infrastructure has not been impaired.
- *Vehicular Networks:* Vehicles on a highway can form an ad hoc network in order to propagate information such as traffic and road conditions. This information can be generated by an individual vehicle and subsequently broadcast to other vehicles. Alternatively, the information can be transmitted to and received from fixed network access points placed near the road.
- *Sensor Dust:* Recent advances in sensor, computing and networking technology have enabled the mass production of intelligent, wireless communicating sensors. Networks of these sensors can be used in many different ways:
    - *Monitoring Space:* Environmental and habitat monitoring, precision agriculture, indoor climate control, surveillance and intelligent alarms.
    - *Monitoring Objects:* Structural monitoring, condition-based equipment maintenance, medical diagnostics, etc.
    - *Monitoring Interactions:* Between objects and between objects and their environment, e.g. wildlife habitats, disaster management, emergency response, healthcare and manufacturing process flow.

1.1.2 Open Challenges

MANETs pose numerous challenges and generate new research problems compared with the fixed wireless networks. These are due to the following reasons:

- *Mobility:* Each node in MANETs tends to have a mobility pattern with changeable speeds. This phenomenon adds another aspect to the problems that is of routing and supporting Quality of Service (QoS).
- *Variable Topology:* The network topology can change rapidly and unpredictably. This is because, as previously stated, nodes are free to move arbitrarily. Moreover, radio propagation conditions can change rapidly.
- *Inexact State Information:* The link state information required for effective (QoS) routing is subject to change mainly due to user mobility and changeable channel conditions.
- *Capacity Constraints:* Wireless links have significantly lower capacity than wired links and hence congestion is more problematic.
- *Variable Link Capacity:* The capacity of wireless links can vary over time due to effects such as multiple access, multipath fading, noise and signal interference.
- *Energy Constrained Nodes:* Nodes participating in the network rely on batteries for power. If the energy in the batteries is depleted, there is an adverse effect on the network's performance.
- *Limited Security:* Mobile wireless networks are generally more vulnerable than wired networks to security threats, such as eavesdropping, spoofing and denial-of-service (DoS) attacks.
- *Scalability:* Because MANETs do not typically allow the same kind of aggregation techniques that are available to standard Internet routing protocols, they are vulnerable to scalability problems. This issue in MANETs can be generally defined as whether the network is able to provide an adequate level of service to packets even in the presence of a larger number of mobile nodes in the network.

Keeping in view the aforementioned challenges, there are still quite a number of open issues. These include medium access scheme, transport layer protocol, energy management, mobility management, security and, of principal interest here, efficient routing and Quality of Service (QoS) issues.

1.2 Research Motivation

The provisioning of real-time applications such as voice and video over ad hoc networks have received a lot of attentions among researchers mainly due to the increasing demand on this service among users [11][12][13][14]. This is particularly challenging due to capacity requirements and stringent delay constraints. In general, wireless nodes have limited resources like capacity and battery power. In multi-hop wireless mobile networks, one of the key issues is how to route packets efficiently. Some of the important factors that need to be considered in designing a routing scheme for MANETs are: minimum delivery latency, higher probability of packet delivery, energy efficiency and adaptability. Therefore, the design of an efficient and reliable routing scheme and QoS support for such applications is a major challenge.

1.2.1 Multipath Routing

Routing is one of the most fundamental aspects of any network. Routing in ad hoc wireless networks play an important role for data forwarding, where each mobile node can act as a relay in addition to being a source or destination node. Because nodes are usually multiple hops away from each other, routing schemes are usually needed for a source to find a route to the destination before it can send any data to the destination.

Reactive or on-demand routing protocols have been widely studied because they consume less capacity than their pro-active or table-driven counterparts [15]. The reason is that table-driven protocols waste the limited system resources to discover routes that are not needed. On the other hand, on-demand routing protocols have been proposed as an effective solution to this problem. Their main advantage is that a route discovery is performed only when there is a request for communication between two network nodes. However, on-demand routing protocols does not exploit the fact that the route discovery has already been performed and does not discover multiple paths. This results in a higher frequency of route discoveries, which in turn, increases delay and overheads.

Multipath routing has the potential to alleviate these problems by establishing multiple paths between source and destination within a single route discovery process. Most of the multipath routing schemes presented in the literature build multiple disjoint routing paths between source and destination, but encounter a broadcast storm of routing packets in the process of discovering multiple disjoint routing paths. This considerably increases delay and routing overhead in the network. In this thesis, a novel and practical routing scheme called Shortest Multipath Source (SMS) routing scheme is proposed. The principle objective of SMS is to address the problem of wireless broadcast storm and build multiple partial-disjoint paths from source to destination in order to avoid the overhead of additional route discoveries and to recover quickly in case of route breaks.

1.2.2 Quality of Service (QoS) Support

The second aspect studied in this thesis is the QoS-aware routing to support QoS in MANETs. QoS is defined as a set of service requirements that needs to be met by the network while transporting a packet stream from a source to its destination [16]. The network requirements are administered by the service requirements of the end user applications. The network is likely to guarantee a set of measurable pre-specified QoS parameters to the users in terms of available capacity, probability of packet loss, delay variance (jitter, unpredictable delay), end-to-end delay (accumulation of jitter along the path) and power consumption or battery charge. As different applications have different requirements, their level of QoS and the associated QoS parameters also differ from application to application.

For real-time applications, the capacity and delay are the key parameters. However, delay is associated with network load and degree of congestion. When capacity is sufficient, delay is relatively small but, when congestion occurs, delay increases significantly [17]. Therefore, in this thesis only capacity constraint is studied; solving the capacity problem inherently helps in solving the delay problem.

Capacity estimation is a key component of any admission control scheme required to support QoS provision in MANETs. A range of routing schemes has been previously proposed to estimate residual capacity that is derived from window-based measurements of channel estimation. In this thesis, a new capacity-constrained multipath routing scheme is proposed. The proposed scheme modifies and extends the route discovery phase and route maintenance phase of SMS scheme to provide QoS assurance. The novel part of this QoS-aware SMS (Q-SMS) routing scheme is a simple and improved mechanism to estimate residual capacity in IEEE 802.11-based ad hoc networks. The scheme proposes the use of a 'forgiveness factor' to weight these previous measurements and is shown through simulation-based evaluation to provide accurate utilisations estimation and improved available capacity based admission control.

1.3 Thesis Contributions

In this thesis, multipath routing and QoS support for MANETS are addressed. The major contributions are as follows:

- A Shortest Multipath Source (SMS) routing scheme is proposed which exploits the route propagation procedure to discover multiple partial-disjoint paths in order to overcome the limitations of current unipath and multipath routing schemes. The advantages of the proposed routing schemes are: a) its simplicity, b) it does not increase the route discovery overhead and c) it recovers quickly in case of route breaks. Additionally, the scheme addresses the problem of wireless broadcast storms through the use of a hop count mechanism.

- A QoS-aware Shortest Multipath Source (Q-SMS) routing scheme based on residual capacity estimation is designed. The scheme modifies and extends the route discovery procedure and route maintenance procedure of SMS scheme to provide QoS assurance.
- A simple mechanism (Q-SMS-F) with a lower reliance upon window size to estimate residual capacity in IEEE 802.11-based ad hoc networks is proposed. The mechanism proposes the use of a 'forgiveness factor' to weight previous window-based measurement techniques in order to provide appropriate utilisation measurements to make improved admission control decisions.

1.4 Thesis Organisation

The outline of the thesis indicating the relationships between different chapters is schematically depicted on Figure 1.2.

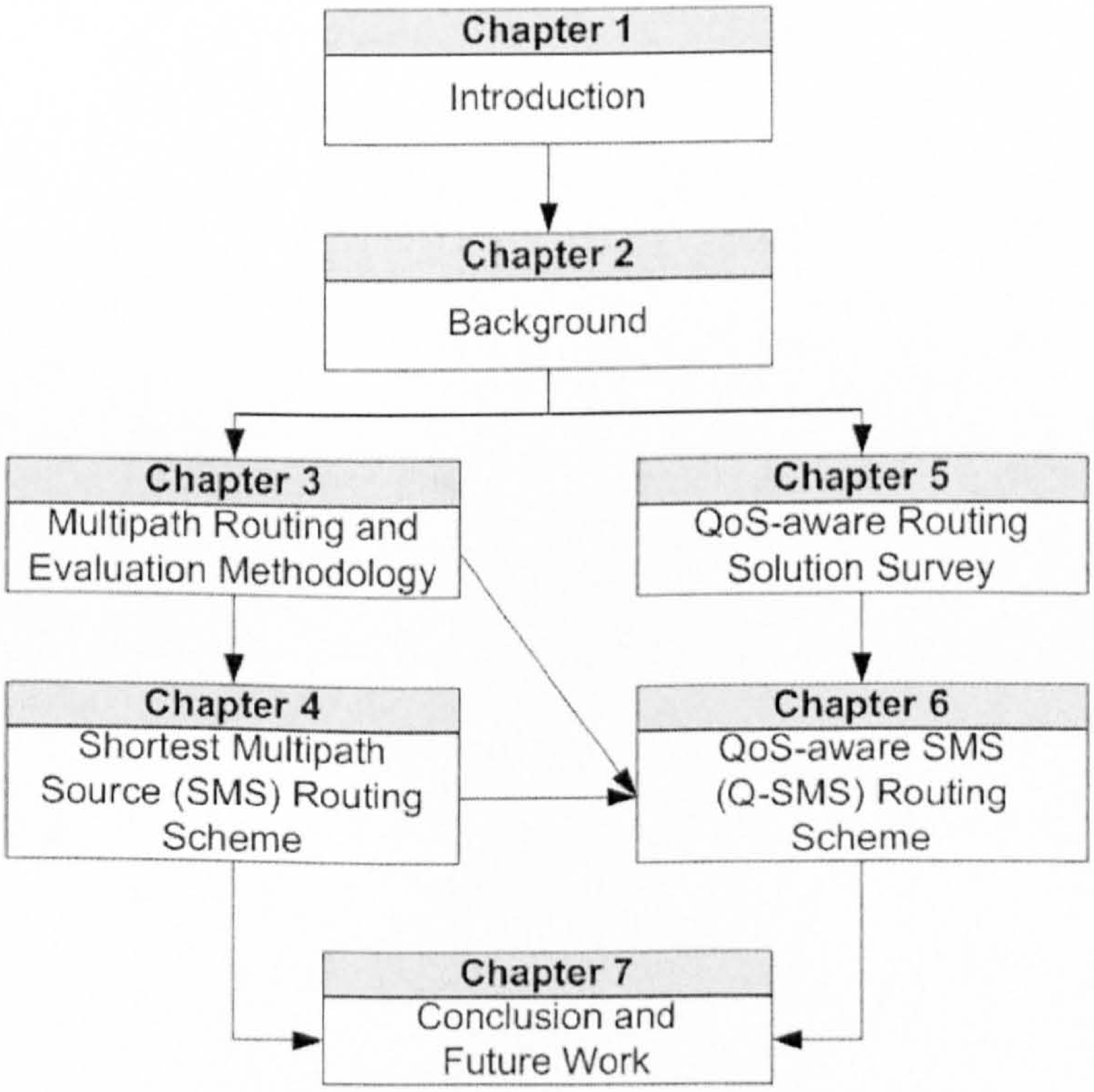


Figure 1.2 Schematic overview of thesis outline

The remaining chapters of this thesis are organised as follows: Chapter 2 provides background information and describes related research efforts in mobile ad hoc networks with particular emphasis on ad hoc routing protocols. Chapter 3 presents the concept of multipath routing and describes a number of multipath routing schemes proposed for MANETs. Additionally, the chapter presents the evaluation methodology and models adopted in this thesis. A novel and practical multipath routing scheme called Shortest Multipath Source (SMS) routing scheme is proposed in Chapter 4. The important aspects of the scheme are explained and a comprehensive performance evaluation of the proposed scheme is performed using NS-2 based model. Chapter 5 presents a survey on existing QoS-aware routing schemes and highlights key routing design issues to support QoS in MANETs. Chapter 6 proposes a novel QoS-aware routing scheme based on capacity estimation to ensure QoS support in mobile ad hoc networks. The performance evaluation of the proposed scheme is performed to demonstrate the merits of the proposed algorithm. Finally, conclusions of the thesis and possible future work are presented in Chapter 7.

2. Background

2.1 Introduction

This chapter provides background and describes related research efforts in Mobile Ad hoc Networks (MANETs). Section 2.2 gives basic definitions used in ad hoc networks. This is then followed by a review of radio propagation models in section 2.3. Section 2.4 presents ad hoc network architecture from the perspective of mobile wireless node. Section 2.5 explains several important concepts of IEEE 802.11 Wireless Local Area Network (WLAN). Section 2.6 highlights common routing protocols used in fixed networks and explains why such protocols cannot be used in ad hoc networks. Section 2.7 presents taxonomy of ad hoc routing protocols and provides a review of commonly used routing protocols which has the status of Internet Engineering Task Force (IETF) experimental Request for Comments (RFCs). Section 2.8 concludes this chapter with a summary that highlights the unique aspects of ad hoc routing strategy and provides a discussion on possible further work related to routing.

2.2 Basic Definitions

Mobile Ad hoc Network (MANET) is typically represented as a dynamic graph $G = \{V, E(t)\}$, where V is the set of vertices or nodes and $E(t)$ is the set of edges at time t [18]. Let $n = |V|$ be the number of mobile nodes participating in wireless communication. Node $i \in V$ can hear node $j \in V$ if node i is within radio range of j. Let $H(i)$ to be a set of nodes which node i can hear and $H(j)$ to be a set of nodes which node j can hear. It is obvious that nodes i and j can hear each other if and only if $i \in H(j)$ and $j \in H(i)$.

The radio range of a node is the geographic distance over which packets sent by the node can be received. The distance used is the Euclidean distance or Euclidean metric. Thus, if the range of a node *A* is *r*, then a packet sent by *A* can be received only by the nodes that are within or on the circle of radius *r* centred at the point occupied by *A*. Different nodes may have different ranges. Therefore, in the light of the above definition, it is not true that if $i \in H(j)$ then $j \in H(i)$ or vice-versa, though it is a frequent assumption for many routing and Medium Access Control (MAC) protocols.

In an ad hoc network, a 'hop' refers to the movement of a packet directly from one node *A* to another node *B* which is within the range of node *A* as shown in Figure 2.1.

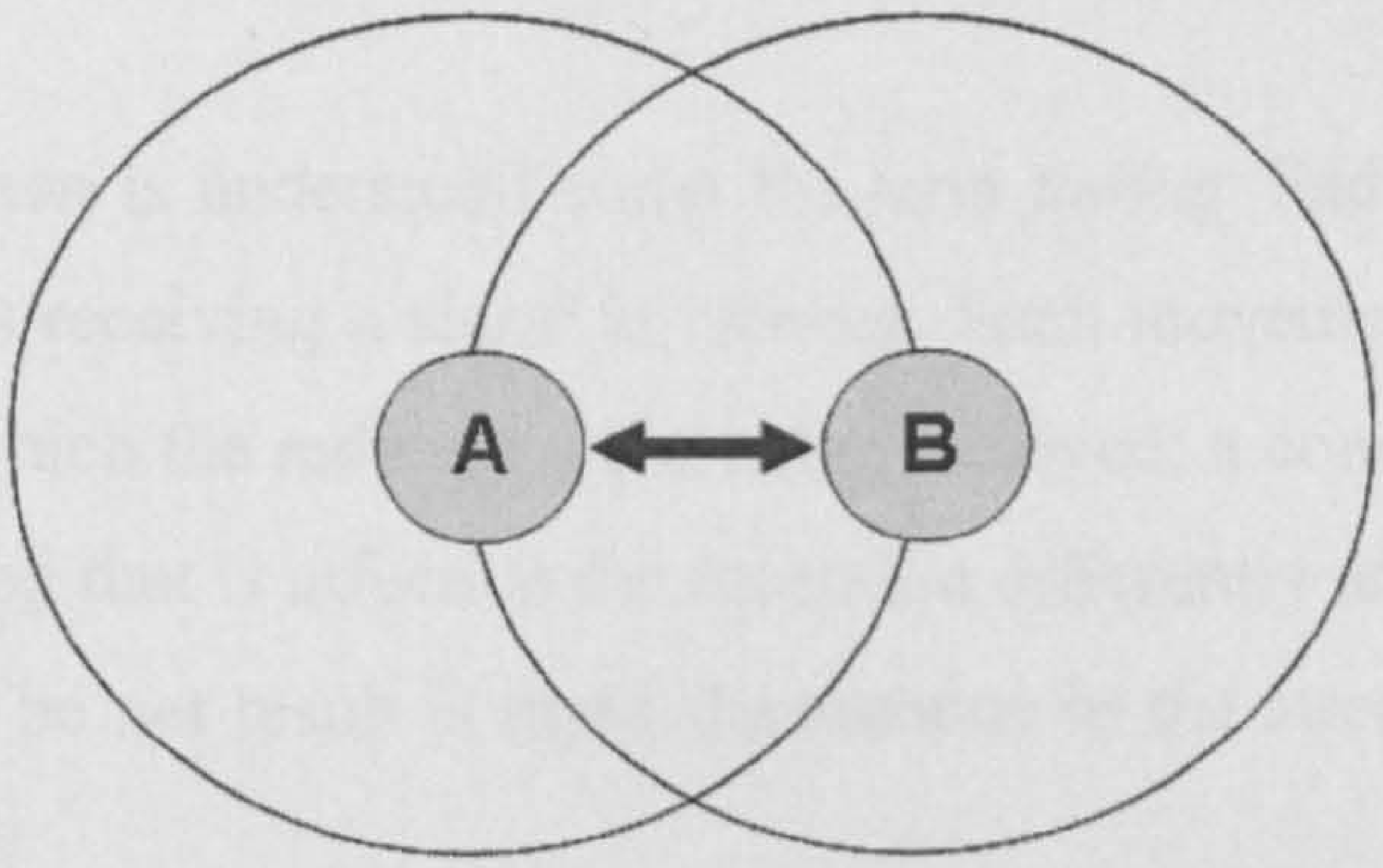


Figure 2.1 Simple ad hoc network with two participating nodes

2.3 Mobile Radio Propagation

Radio range of a mobile node in real conditions is subject to many restrictions. Most wireless systems are expected to operate in areas with a profusion of obstructions in form of walls, buildings, trees or mountains. Other important limiting factors are the over-ground elevation of the transmitter and receiver, speed of motion, atmospheric conditions, etc.

According to [1], the mechanisms behind radio propagation can generally be attributed to reflection, diffraction and scattering:

- *Reflection* occurs when direct radio propagation is not possible as a result of an obstructing object having very large dimensions when compared to the wavelength of the propagating wave. Reflections occur from the surface of the earth and from walls and buildings.
- *Diffraction* occurs when direct radio propagation is obstructed by a surface that has sharp irregularities or edges. The direction of radio waves is changed due to these edges and often results in the bending of waves around the obstacle.
- *Scattering* occurs when the medium through which a radio wave is being propagated consists of many objects that are of comparable size or smaller than the wavelength, and at the same time, the number of these objects is high. Examples of such objects are sign posts, foliage or even people.

A different phenomenon is understood under the term *fading*. Fading can occur when a wireless device that is receiving a signal is moving. Such movement results in a change of conditions under which the radio signal is being received; a consequence of reflection, diffraction or scattering that is affecting the reception differently at each new position of the wireless device. The net result is rapid fluctuations in the strength and phase of the received signal.

Models that attempt to describe the above mentioned propagation qualities can be divided into large-scale and small-scale propagation models. Large-scale models describe situations when an arbitrary transmitter-receiver separation distance is many orders of magnitude higher than the wavelength, usually measured in meters or kilometres. Small-scale models describe situations where propagation over distances proportional to wavelength is considered.

A number of common propagation models used in ad hoc networks are described below. Details on other propagation models can be found in [1].

2.3.1 Free Space Propagation Model

The free space propagation model attempts to describe radio wave propagation when the transmitter and receiver have a clear, unobstructed line-of-sight path between them. This model predicts received power decays as a function of the transmitter-receiver separation distance raised to some power, i.e. the decay obeys a power law function. The free space power received $P_r(d)$ with respect to distance is then given by the following equation:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{2.1}$$

where P_t is the transmitted power, G_t is the transmitter antenna gain, G_r is the receiver antenna gain, λ is the wavelength, d is distance between the transmitter and the receiver and L is a system loss factor not related to propagation ($L \geq 1{:}0$). System losses are usually due to line attenuation, filter losses and antenna losses. Gain of antenna is related to its effective aperture A_e by:

$$G = \frac{4\pi A_e}{\lambda^2} \tag{2.2}$$

The effective aperture is then related to the physical size of the antenna and λ is related to the carrier frequency f by:

$$\lambda = \frac{c}{f} \tag{2.3}$$

where c is the speed of light. The above equations show that signal strength decays proportionally to the inverse of square of the distance between the transmitter and the receiver.

2.3.2 Two-ray Propagation Model

In a mobile radio channel, a single direct path between the transmitter and receiver is seldom the only physical means for propagation, and hence the free space propagation model of equation (2.1) is in most cases inaccurate when used alone. The two-ray ground reflection model is a useful propagation model that takes into account both the direct path and a ground reflected propagation path between transmitter and receiver. These aspects are depicted in Figure 2.2. The signal strength at the receiver for the two-ray ground model can be expressed as:

$$P_r(d,h_t,h_r) = \frac{P_t G_t G_r h_t^2 h_r^2}{d^4} \tag{2.4}$$

where h_t, h_r is elevation of the transmitter and receiver, respectively. Thus, the received signal strength for large distances ($d >> \sqrt{h_t h_r}$) decays with the fourth power of the distance d. This is a much faster rate of decay than with the free space propagation model. At large values of d, the received power and path loss become independent of frequency. The two-ray propagation model is the mostly used model in MANET research community.

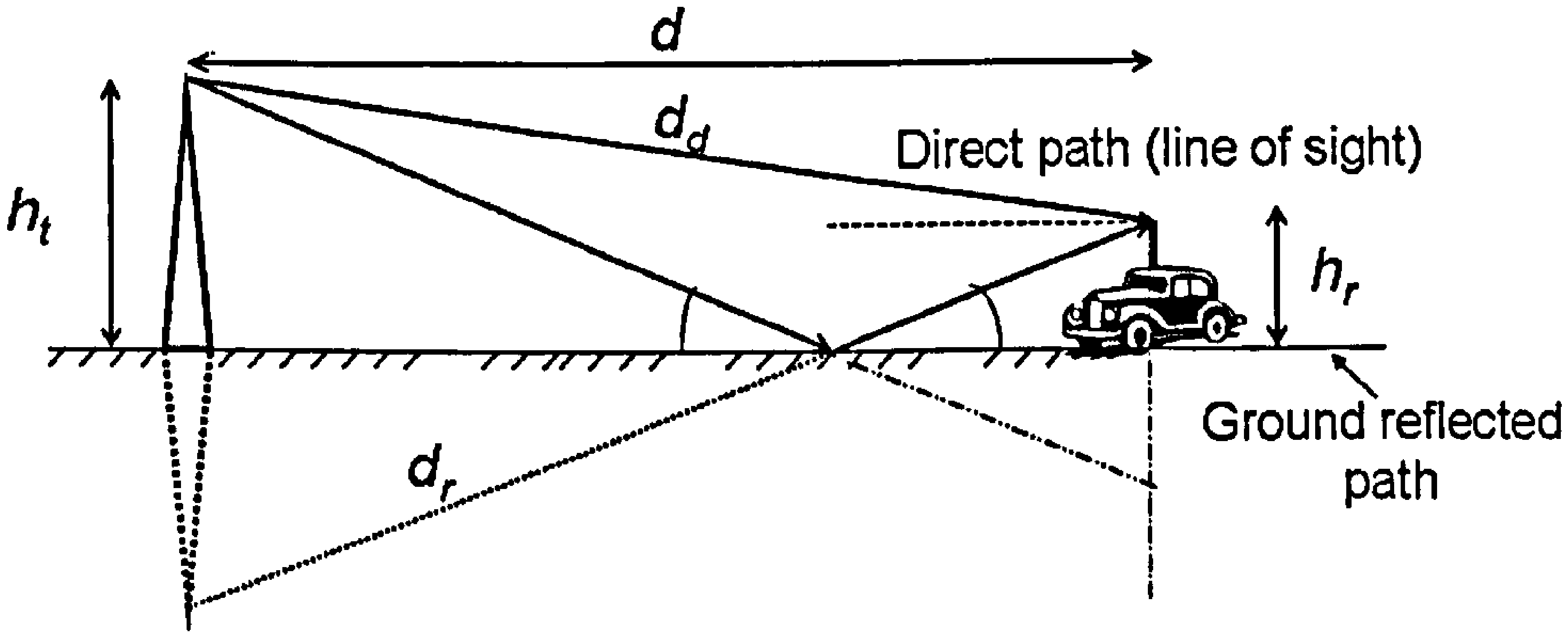


Figure 2.2 Two-ray radio propagation [1]

2.4 Network Architecture

The layers, along with their respective protocols, constitute the network architecture. The Open Systems Interconnection (OSI) model [19] is depicted in Figure 2.3. A packet is sent from the source node to the destination node. The information travels from the application layer of the source node downwards through the intermediate layers, travel up to the application layer of the destination. The intermediate nodes act as routers and so are concerned only with the functions associated with the lower three layers.

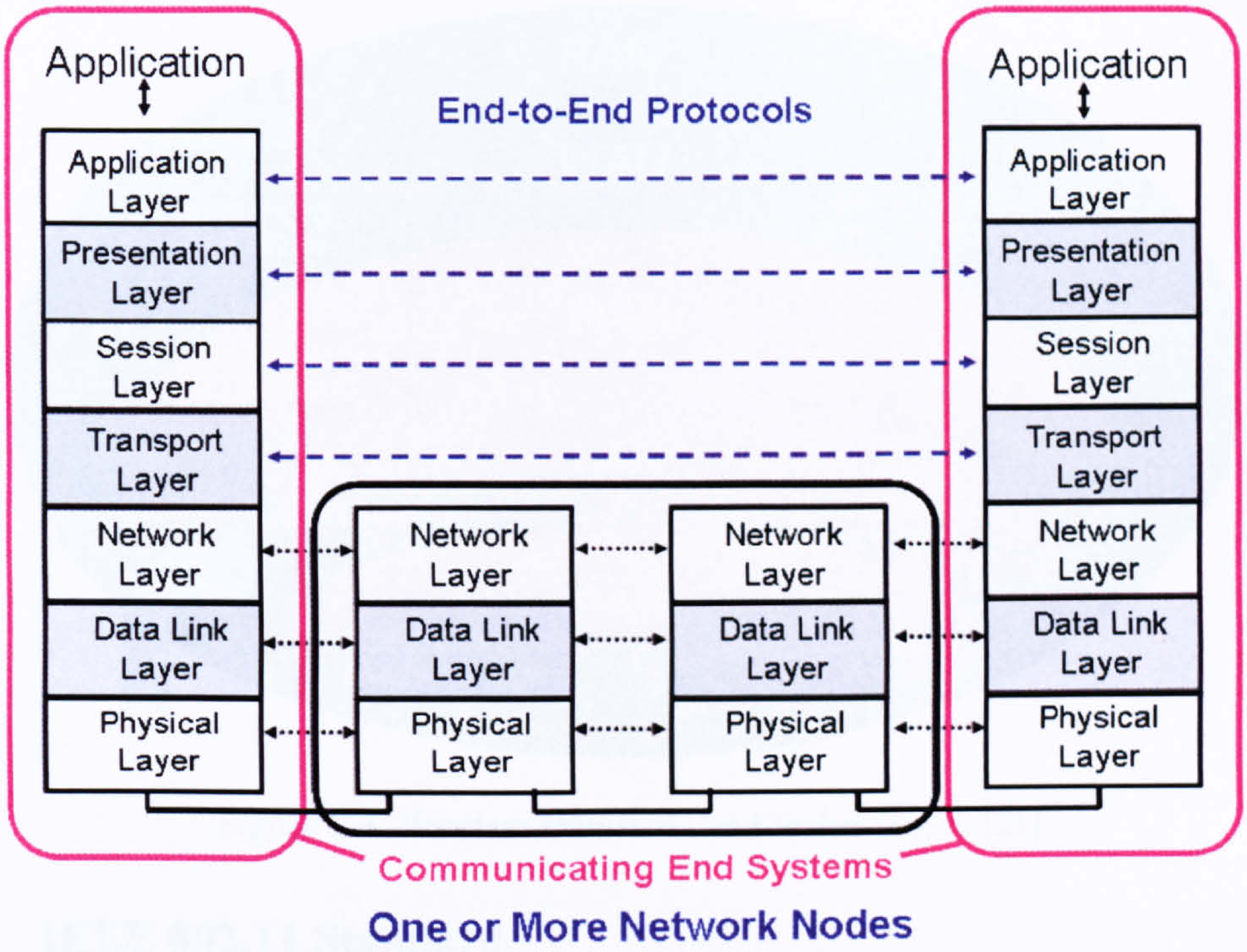


Figure 2.3 Communication between nodes w.r.t. OSI reference model [19]

The network architecture does not pertain to any specific hardware or software. Using the network architecture, an implementer would be able to design their hardware and software for the network. For example, the IEEE 802.11 [2] WLAN standard describes the physical (PHY) and the data link or MAC layers of the OSI model. Whereas technologies like Bluetooth [3] and Zigbee [20] specify the whole protocol stack. It is not in the scope of this thesis to analyse various wireless standards in depth. However, for a purpose of completeness a brief overview of IEEE 802 and European Telecommunication Standard Institute (ETSI) standards and technologies are depicted in Figure 2.4 [21]. This also serves to show how IEEE 802.11 (used in the majority of ad hoc networking simulations and test-bed implementations) fits with the rest of the enabling technologies. Details of IEEE 802 and ETSI standards and technologies can be found in [22] and [23] respectively. The next sections describe the IEEE 802.11 standard (PHY and MAC layers) and the routing protocols (network layer) for ad hoc networks.

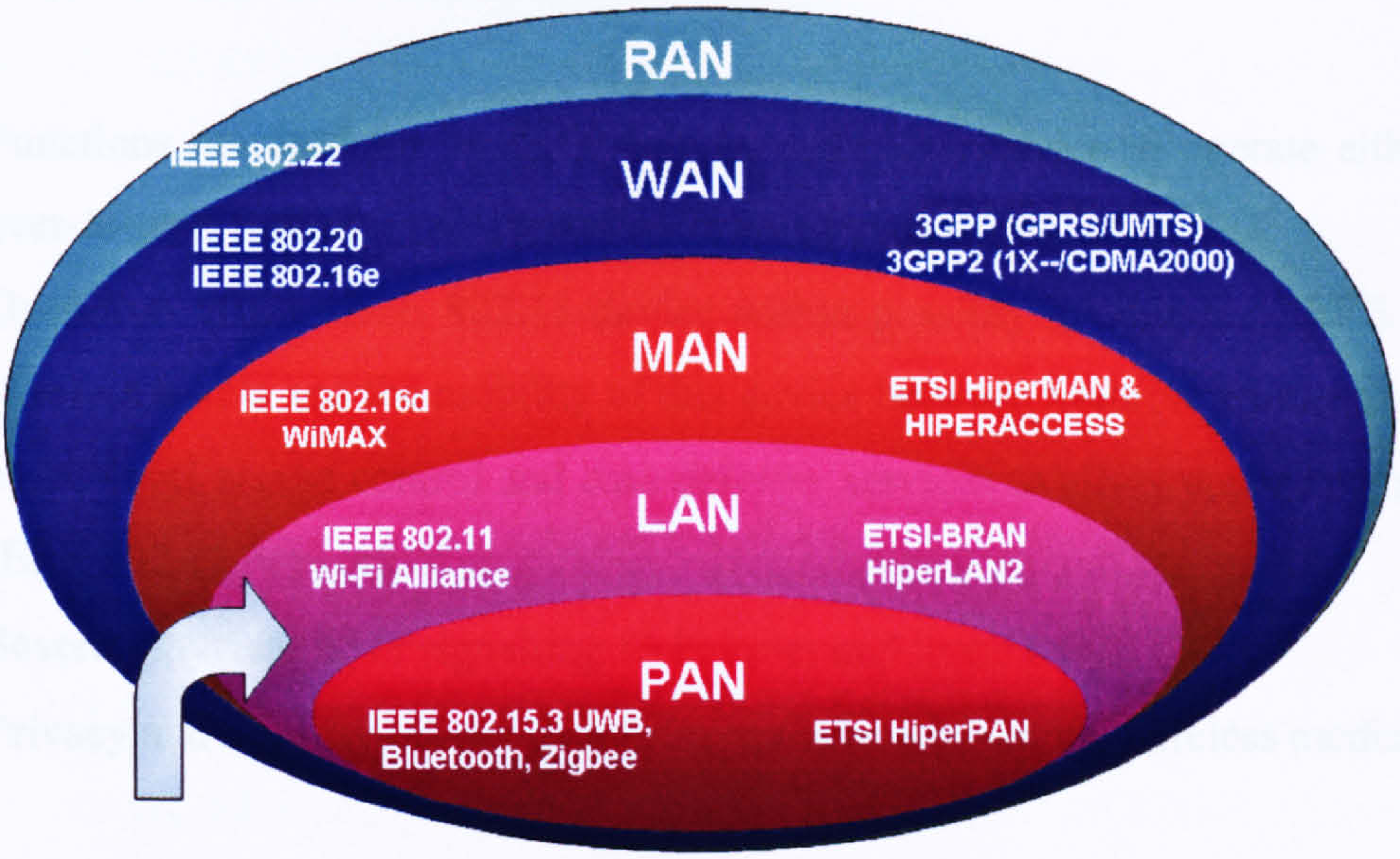


Figure 2.4 Wireless standards and technologies [21]

2.5 IEEE 802.11 Standard

In 1997 the IEEE adopted IEEE 802.11 standard, the first WLAN standard. This standard has been accepted widely and rapidly for many different environments. The main characteristics of the IEEE 802.11 are its simplicity, scalability and robustness against failures due to its distributed nature. This standard defines the PHY and MAC layers for a LAN with wireless connectivity. It addresses local area networking where the connected devices communicate over the air to other devices that are within close proximity to each other. IEEE 802.11 does not specify any special nodes that support routing, forwarding of data or exchange of topology information. Figure 2.5 illustrates the IEEE 802.11 standard mapped to the OSI reference model.

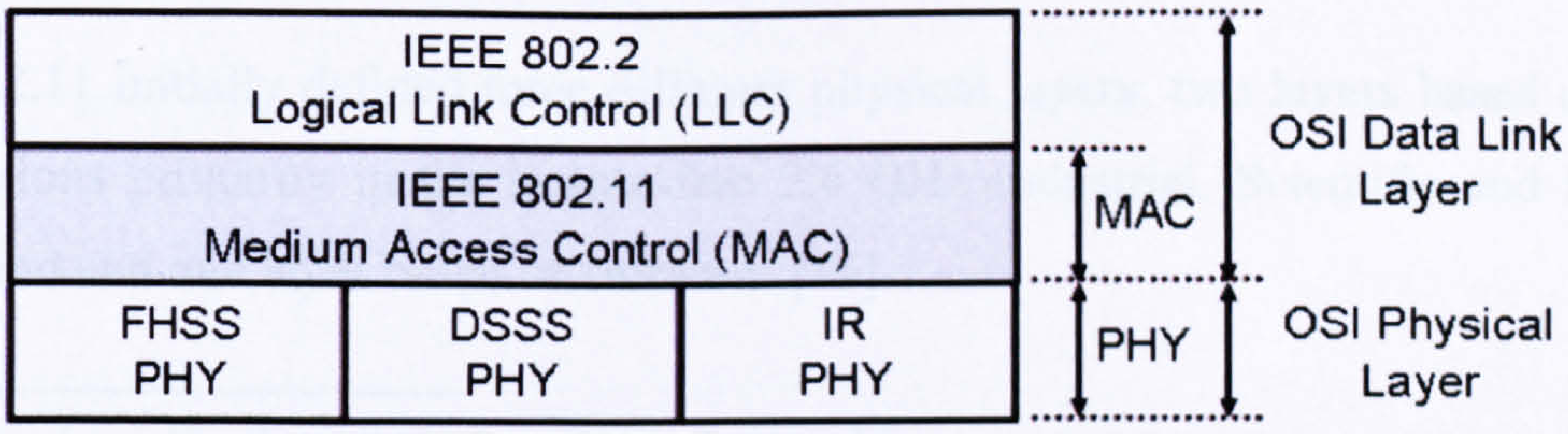


Figure 2.5 IEEE 802.11 standards mapped to the OSI reference model

The standard is similar in most respects to the IEEE 802.3 Ethernet standard [24]. Specifically, the IEEE 802.11 standard addresses:

- Functions required for an IEEE 802.11 compliant device to operate either in a peer-to-peer fashion or integrated with an existing wired LAN.
- Operation of the IEEE 802.11 device within possibly overlapping IEEE 802.11 wireless LANs and the mobility of this device between multiple wireless LANs.
- MAC level access control and data delivery services to allow upper layers of the IEEE 802.11 network.
- Several physical layer signalling techniques and interfaces.
- Privacy and security of user data being transferred over the wireless media.

2.5.1 IEEE 802.11 Physical Layer (PHY)

The IEEE 802.11 PHY is the interface between the MAC and the wireless media (air in ad hoc networks) where frames are transmitted and received. The PHY provides three functions.

- It provides an interface to exchange frames with the upper MAC layer for transmission and reception of data.
- It uses signal carrier and spread spectrum modulation to transmit data frames over the transmission media.
- It provides a carrier sense indication back to the MAC to verify activity on the media.

IEEE 802.11 initially defined three different physical layers; two layers based on radio transmissions primarily in the license-free 2.4 GHz Industrial, Scientific and Medical (ISM) band and one layer based on infrared[2] [15].

---

[2] The original specifications did include IR but that has fallen into abeyance. IR is out of the thesis scope and hence is not discussed further.

- *FHSS:* Frequency Hopping Spread Spectrum (FHSS) operating in the 2.4 GHz ISM band, at data rates of 1 Mbps using 2-level Gaussian Frequency Shift Keying (GFSK) modulation scheme and 2 Mbps using 4-level GFSK.
- *DSSS:* Direct Sequence Spread Spectrum (DSSS) operating in the 2.4 GHz ISM band, at data rates of 1 Mbps using Differential Binary Phase Shift Keying (DBPSK) modulation scheme and 2 Mbps using Differential Quadrature Phase Shift Keying (DQPSK).

There are several specifications in the IEEE 802.11 family: IEEE 802.11a uses OFDM transmission technology, operates at the 5 GHz band, and offers a nominal rate of up to 54 Mbps. IEEE 802.11b defines 11 Mbps and 5.5 Mbps data rates (in addition to the 1 and 2 Mbps rates) utilising an extension to DSSS called High Rate DSSS (HR/DSSS). It uses an unlicensed portion of the radio spectrum at 2.4 GHz. IEEE 802.11g is similar to 802.11a and uses OFDM. It operates at the 2.4 GHz band, and is backward compatible to IEEE 802.11b. IEEE 802.11n is a proposed amendment to the IEEE 802.11 standard to significantly improve network throughput over previous standards with a significant increase in the maximum raw data rate from 54 Mbps to a maximum of 600 Mbps. The current state of the art supports a PHY rate of 300 Mbps, with the use of 2 spatial streams at a channel width of 40 MHz. Table 2.1 provides the frequencies, data rates and ranges for the IEEE 802.11 family of standards [25].

Table 2.1 IEEE 802.11 family of standards

Standard	**Operating frequency**	**Throughput (Typical)**	**Data rate (Maximum)**	**Range (Indoor)**
IEEE 802.11a	5 GHz	23 Mbps	54 Mbps	~35 m
IEEE 802.11b	2.4 GHz	4.5 Mbps	11 Mbps	~38 m
IEEE 802.11g	2.4 GHz	19 Mbps	54 Mbps	~38 m
IEEE 802.11n	5 GHz and/or 2.4 GHz	74 Mbps	300 Mbps (2 streams)	~70 m

2.5.2 Classification of MAC Schemes

The MAC layer provides the network layer with a connection between two nodes. One of the main functions when the communicating entities work in a shared medium is to coordinate their respective use of the medium for effective communication. This is the task of MAC schemes. Based on the technique adopted, MAC schemes are classified into two categories: contention-free and contention-based [15].

- *Contention-free scheme:* The category uses reservation-based channel access mechanisms such as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA), Code Division Multiple Access (CDMA) and Space Division Multiple Access (SDMA). However, contention-free MAC requires effective time synchronisation between all nodes in the network. Applying synchronisation in an ad hoc network is expensive and synchronisation is very likely to fail when the nodes are mobile.
- *Contention-based scheme:* This scheme is based on random channel access policy, such as Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA). A node does not make any resource reservation a priori. Whenever it receives a packet to be transmitted, it contends with its neighbour modes for access to the shared channel. When several neighbouring nodes transmit simultaneously, a collision will occur. When a node detects a collision, it tries to access the channel again after a random backoff time. The MAC techniques of interest in ad hoc networking are contention-based in order to avoid the cost of time synchronisation in MAC layer.

2.5.3 IEEE 802.11 MAC Layer (MAC)

The following two network architecture modes have been defined for the IEEE802.11 standard:

- *PCF:* The Point Coordination Function (PCF) mode uses the centralised approach in which a network access point controls all traffic in the network, including local traffic between wireless nodes in the network.
- *DCF:* The Distributed Coordination Function (DCF) mode supports direct communication between wireless nodes. Each node gets an equal share of the channel through contention. It is clear that, for ad hoc networks, DCF mode is the one used.

The mandatory access mechanism of IEEE 802.11[3] is based on CSMA/CA [15][26][27], which is a random access scheme with carrier sense and collision avoidance through random backoff. The basic CSMA/CA mechanism is shown in Figure 2.6. If the medium is idle for the duration of DIFS[4] (DCF Inter Frame Spacing), a node can access the medium for transmission. Thus, channel access delay at very light loads is equal to the DIFS. If the medium is busy, the node uses a backoff mechanism, which leads to fair time distribution of the transmissions. If a node senses the medium as busy, it defers until the ongoing transmission finishes. At this time, the node initializes a backoff timer by selecting a random interval (backoff interval). The backoff timer decrements when the channel is sensed as idle but crucially is passed when an ongoing transmission is heard. The time resumes when the channel is detected to be idle for duration greater than DIFS. When the backoff reaches zero, the node commences transmitting.

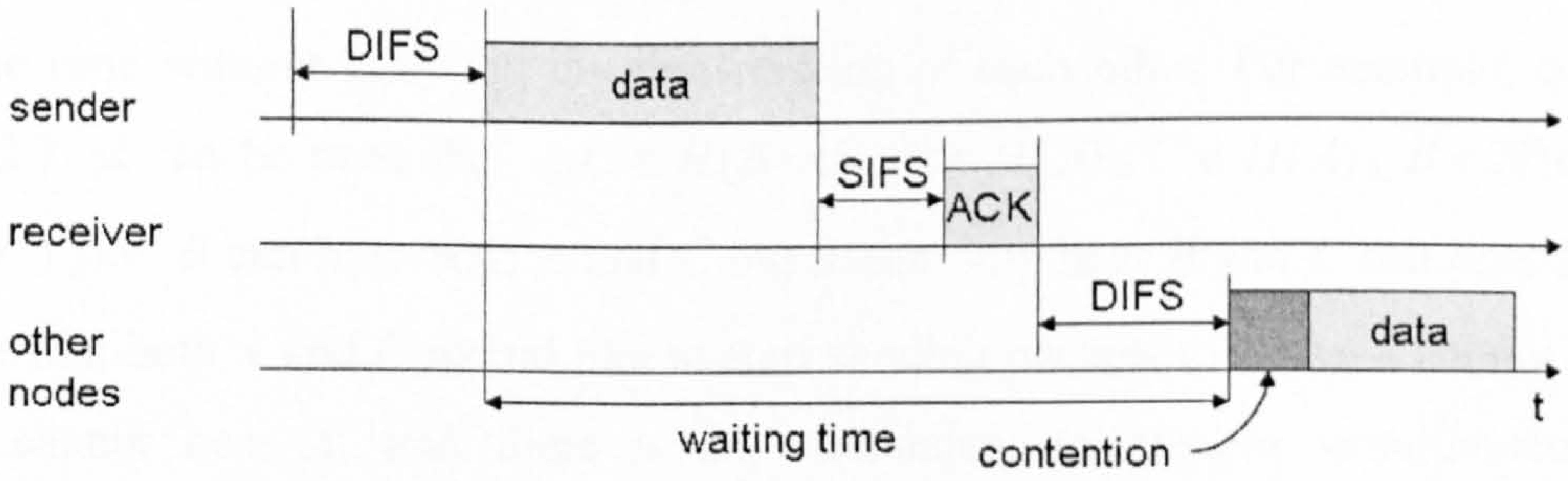


Figure 2.6 Basic access scheme [27]

---

[3] IEEE 802.11 DCF and IEEE 802.11 are used interchangeably for simplicity.
[4] DIFS denotes the longest waiting time and has the lowest priority for medium access.

A further improvement is to set priorities depending on the time spent by a node waiting for the medium. IEEE 802.11 uses a binary-exponential backoff for this purpose. The initial backoff window (also called contention window) is established at (0, *CWmin*). The interval is important, since choosing a too large interval could result in greater overload while choosing a too short interval could result in more collisions. In IEEE 802.11 the contention window is set dynamically, depending on collision occurrence.

Acknowledgements (ACK) for data packets must be sent in order to ensure their correct delivery. For unicast packets, the receiver accesses the medium after waiting for a SIFS[5] (Short Inter Frame Spacing) and then sends an ACK. Other nodes wait for a duration equivalent to DIFS plus their backoff time. Cyclic Redundancy Checksums (CRC) is used to ensure frame/ACK integrity. If no ACK is received by the sender, then a retransmission takes place. The number of retransmissions is limited and failure is reported to the higher layer after the retransmission count exceeds this limit.

In MANETs, when a node transmits a packet, only neighbouring nodes that are within its transmission range can receive it. However, this characteristic leads to the hidden terminal and exposed terminal problems [26]. The hidden terminal problem refers to the collision of packets at a receiving node due to the simultaneous transmissions of those nodes that are not within the direct transmission range of the sender, but are within the transmission range of the receiver. Collision occurs when both nodes transmit packets at the same time without knowing the transmission of each other. For example, consider Figure 2.7. It can be seen that $A, C \in H(B)$ but $B \in H(A)$, $C \notin H(A)$, $B \in H(C)$ and $A \notin H(C)$, i.e. B can hear both A and C but A can only hear B and C can only hear B. Suppose that both A and C would like to start sending packets to B. As A cannot hear C and C cannot hear A, and there is no mechanism to prevent simultaneous data transmission, the wireless medium at B may be corrupted by transmission from nodes A and C. This will result in lost data packets and wasted bandwidth. This behaviour is

---

[5] The shortest waiting time for medium access (so the highest priority) is defined for short control messages.

characteristic to the ALOHA protocol [28] that was designed to work on very sparse networks in which the hidden terminal problem is partially eliminated.

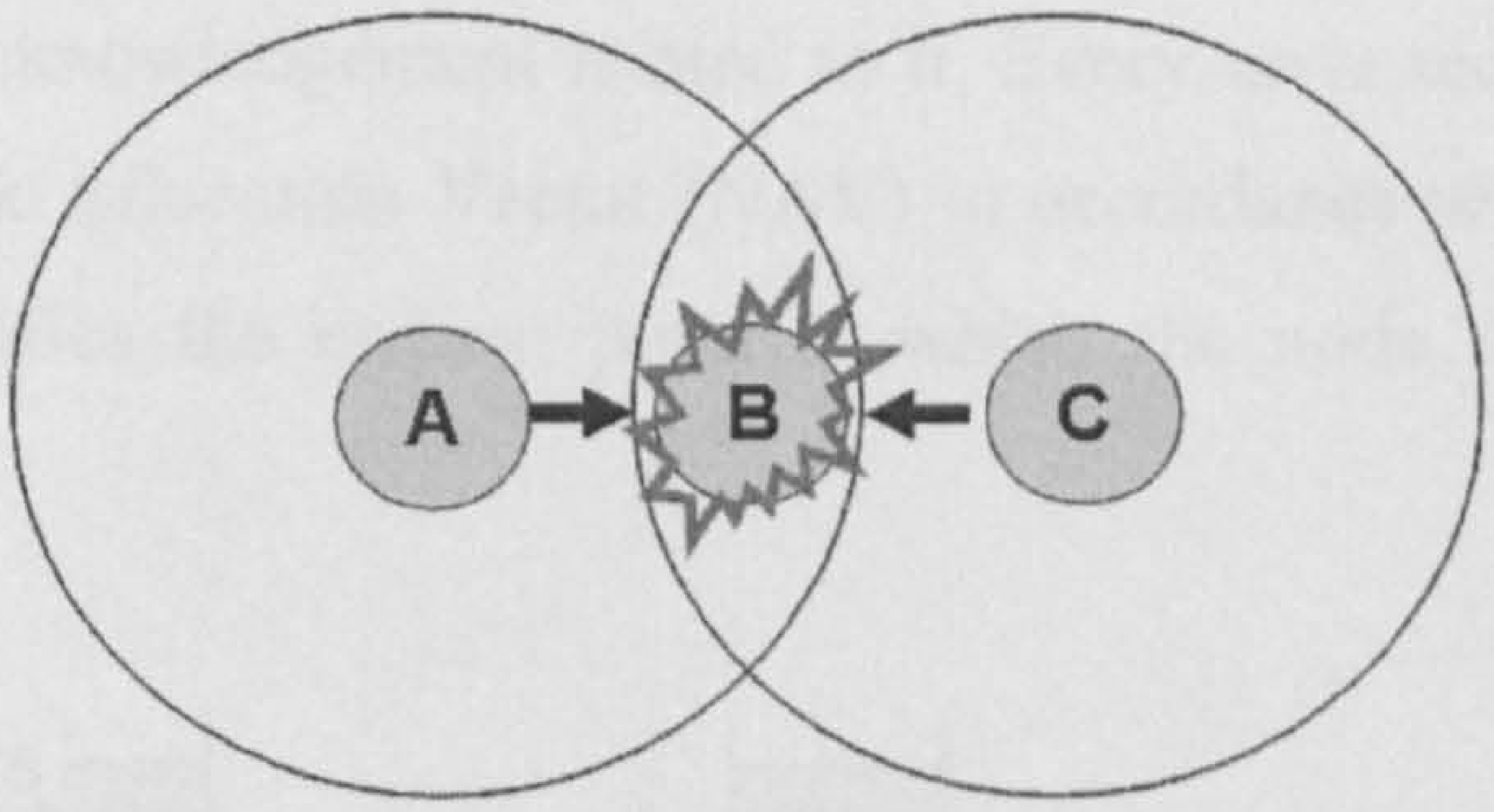


Figure 2.7 Hidden terminal problem

The exposed terminal problem refers to the inability of a node, which is blocked due to transmission by a nearby transmitting node, to transmit to another node. Consider the example in Figure 2.8 which shows four nodes *A*, *B*, *C*, *D*. The situation can be expressed as $A \in H(B)$, $B \in H(A)$ and $B, D \in H(C)$; *B* is transmitting to *A* and *C* would like to start a transmission to *D*. However, *C* senses the carrier of *B* and therefore postpones its own transmission when it could have transmitted, i.e. *C* could have started its own transmission but due to $B \in H(C)$ defers.

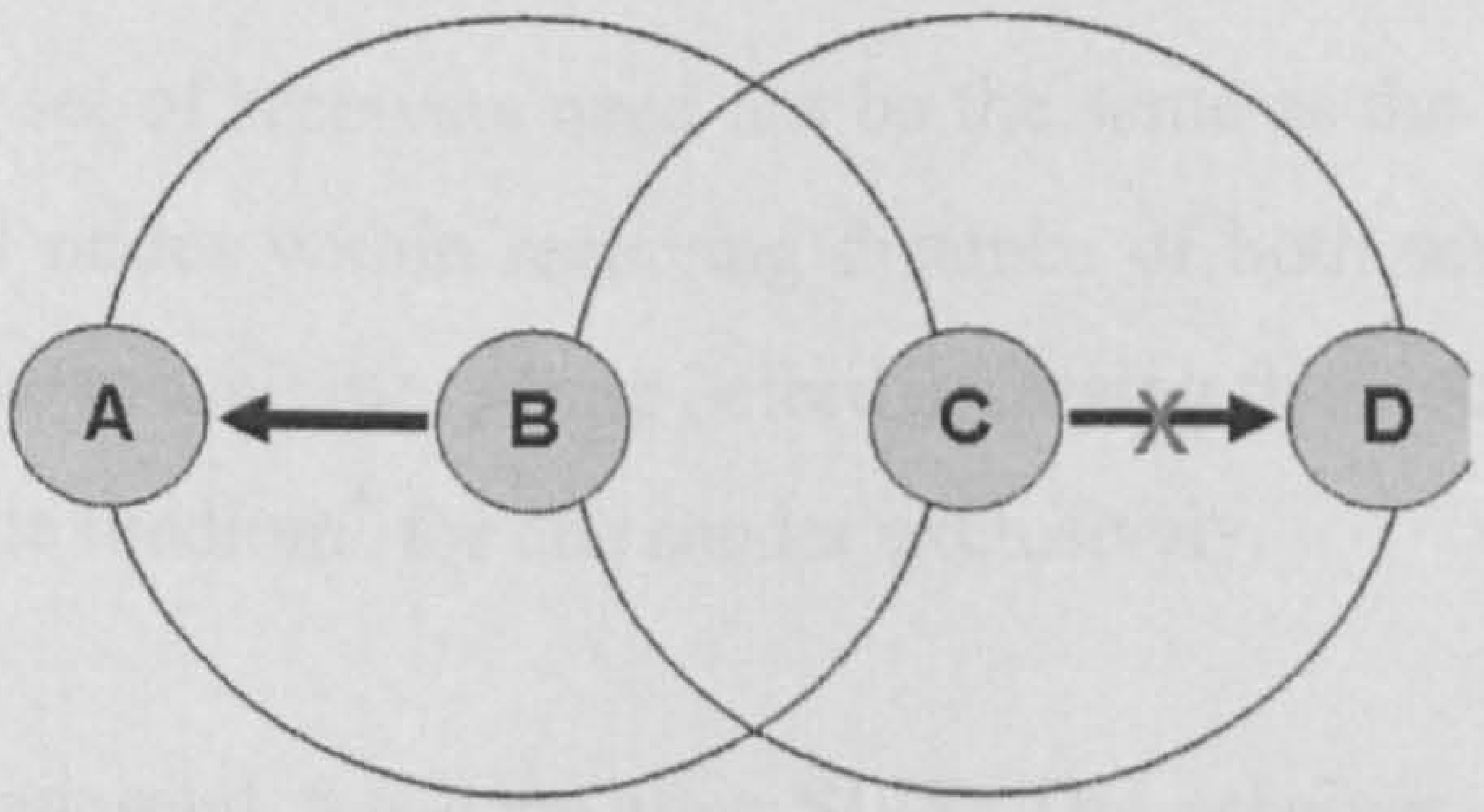


Figure 2.8 Exposed terminal problem

To elevate the hidden terminal and exposed terminal problems, an optional Request to Send (RTS) – Clear to Send (CTS) scheme is used in addition to the previous basic scheme, as shown in Figure 2.9. After waiting for DIFS (plus a random backoff time if the medium was busy), the sender can issue an RTS control packet. The RTS packet is not given any higher priority compared to other data packets. The RTS packet includes the receiver of the data transmission to come and the duration of the whole data transmission. This duration specifies the time interval necessary to transmit the whole data frame and the acknowledgement related to it. Every node receiving this RTS now has to set its Network Allocation Vector (NAV) in accordance with the duration field. The NAV then specifies the earliest point at which the node can try to access the medium again.

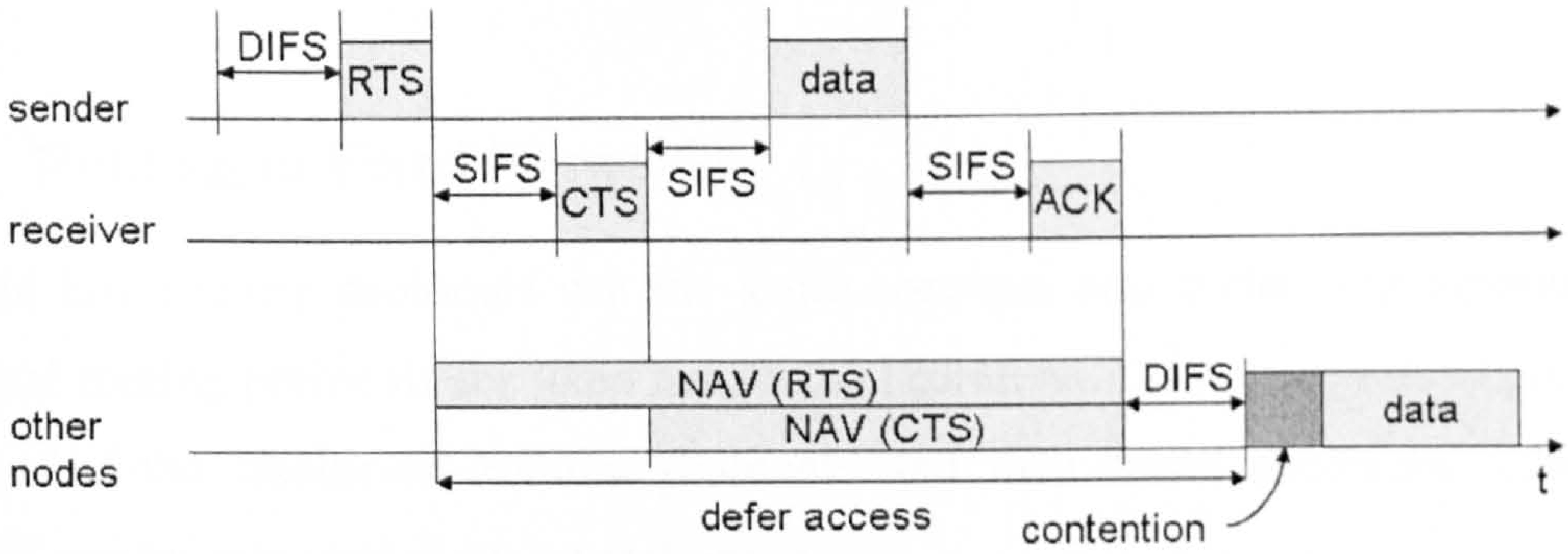


Figure 2.9 RTS-CTS access mechanism [27]

If the receiver of the data transmission receives the RTS, it answers with a CTS message after waiting for SIFS. This CTS packet contains the duration field again and all nodes receiving this packet from the receiver of the intended data transmission must adjust their NAV. The latter set of receivers need not be the same as the first set receiving the RTS packet. Now, all nodes within receiving distance of both sender and receiver are informed that they have to wait more time before accessing the medium. Essentially, this mechanism reserves the medium[6] for one sender exclusively.

Finally, the sender can send the data after SIFS. The receiver waits for SIFS after receiving the data packet and then acknowledges whether the transfer was correct. The transmission has now been completed; NAV in each node marks the medium as free and the cycle can start again.

---

[6] It is some times called a virtual reservation scheme.

It can be observed that the above mechanism is similar to reserving the medium prior to a particular data transfer sequence in order to avoid collisions during this transfer. But transmission of RTS-CTS can result in non-negligible overhead. Therefore, the RTS-CTS mechanism is used judiciously. An RTS threshold is used to determine whether to start the RTS-CTS mechanism or not. Typically, if the frame size is more than the RTS threshold, the RTS-CTS mechanism is activated and a four-way handshake (i.e. RTS-CTS-DATA-ACK) follows. If the frame size is below the RTS threshold, the nodes resort to a two-way handshake (DATA-ACK).

2.6 Routing in Fixed Networks

Many ad hoc routing protocols use the same concepts and underlying algorithms as traditional routing protocols for fixed networks. Therefore, it is appropriate to present an overview of the traditional routing protocols and their basic operation. Traditional protocols can be categorized by several criteria:

- *Static vs. Dynamic:* In static protocols, the routing tables are constructed a priori (possibly by a network administrator). In contrast to dynamic protocols, these tables do not change in response to small topological changes (caused by broken links and node failures) or changes in traffic patterns.
- *Centralised vs. Distributed:* In centralised routing a single node collates all the information about the network topology, computes the relevant routes and then passes them to the other nodes of the network. In distributed routing, adjacent nodes exchange information to update routing tables.
- *Single-path vs. Multipath:* Single-path routing is performed by obtaining the one best possible path for a packet to travel from source to destination. Multipath routing, on the other hand, acknowledges that there is more than one possible route between source and destination. It is therefore more reliable than single-path routing but may not always use the best possible route.

From the above, it is clear that, with respect to the characteristics of ad hoc networks, the most suitable protocols are dynamic and distributed. The two most prevalent examples of dynamic, distributed routing are distance vector and link state routing.

2.6.1 Distance Vector

In distance vector routing each node maintains a routing table containing the next hop and length of the shortest path to every other node in the network. Nodes determine the 'distance' between themselves and adjacent nodes using periodic message exchanges. Furthermore, each node's routing table is periodically broadcast to all adjacent nodes. Upon receiving a neighbouring node's routing table, a node can use this information to update the shortest path in its routing table.

Compared to link state, distance vector is more computation efficient, easier to implement and requires much less storage space. However, it can create long-lived routing loops, due to the fact that nodes choose their next hops to a completely distributed manner, based on information that can be out of date. Distance vector is used in the Routing Information Protocol (RIP) [29] that is used for Internet routing.

2.6.2 Link State

In link state routing, each node maintains a table of links to adjacent nodes, with some measure of the state, or cost, of each link. This table is periodically flooded throughout the network, such that each node in the network has full knowledge of the network topology. With this knowledge, the nodes are able to construct and update routing tables containing the next hops of the best paths to all other nodes in the network using an appropriate shortest path algorithm. Typically the shortest path algorithm is that developed by Dijkstra [19].

Link state routing overcomes some of the shortcomings of distance vector, such as long-lived routing loops that can cause packets to circulate in the network indefinitely. It forms the basis of the Open Shortest Path First (OSPF) protocol [30], which is used for Internet routing.

2.7 Routing in Ad hoc Networks

Design of appropriate routing protocols for ad hoc networks is one of the major challenges. To support ad hoc mobile communications, an ad hoc routing protocol will need to perform the following functions:

- Determine and detect changes in network topology.
- Maintain network topology and connectivity.
- Schedule packet transmissions and channel assignment.
- Routing.

2.7.1 Characteristics of an Ideal Ad hoc Routing Protocol

The existing routing protocols, designed for conventional wired packet switched networks, are very mature, having been in use for a considerable amount of time. These protocols, which make use of distance vector or link state algorithms, could be used in ad hoc networks. Nevertheless, the particular characteristics of the wireless medium make the following properties desirable for ad hoc routing protocols:

- *Distributed implementation:* MANETs are autonomous and self-organising systems, which do not rely on centralised authorities. Therefore, routing protocols must be based on distributed routing techniques.
- *Adaptability to changing topology:* Routing protocols should be able to react to changes in topology and provide new and stable routes promptly. For example, in distance vector routing, the speed of convergence to a new route is slow, leading to inaccurate route information due to the presence of out of date routes.

- *Efficient capacity utilisation:* Since the capacity of wireless networks is limited, reduction of control overhead is an important factor. That is, if a routing protocol generates excessive control traffic, the capacity available for data traffic will be greatly reduced. Moreover, the larger the network, the more capacity is consumed during the propagation of routing information. The above can have a serious impact on communication performance. For example, table driven ad hoc routing protocols transmit routing information (distance vector or link state tables) periodically, generating significant control overhead.
- *Energy conservation:* The lifetime of mobile nodes has a strong impact on the performance of MANETs, since nodes need to relay their messages towards their destinations through other nodes. Therefore, a decrease in the number of available nodes could lead to a degradation of network performance or even partitioning of the original network into smaller networks. Consequently, one of the most important considerations of ad hoc routing protocols should be power consumption. For example, routing protocols should be able to accommodate sleep periods without causing any adverse consequences.
- *Freedom from loops:* Since routing algorithms can cause a small fraction of packets circulating around the network for an arbitrary period of time (temporary loops). Looping of packets can result in considerable capacity and power consumption. These cause many problems and are highly undesirable, as discussed previously.
- *Unidirectional link support:* A number of factors, such as the presence of different radio capabilities and signal interference, can cause wireless links to be unidirectional. Hence, the ability to utilise unidirectional links is important.

2.7.2 Taxonomy of Ad hoc Routing Protocols

The recognition of the requirements mentioned in the previous section has led to a large body of work on the subject of ad hoc routing. A large number of competing protocols have emerged. Some of them have become Internet Engineering Task Force (IETF)

Request for Comments (RFCs); others have draft status in the IETF, while many have been dropped by the research community.

Consequently there are many ways in which the protocols can be categorised according to various characteristics they have. The most prevalent taxonomies found in literature [15][31][32] are depicted in Figure 2.10. This taxonomy is based on to divide protocols according to following criteria, reflecting fundamental design and implementation choices:

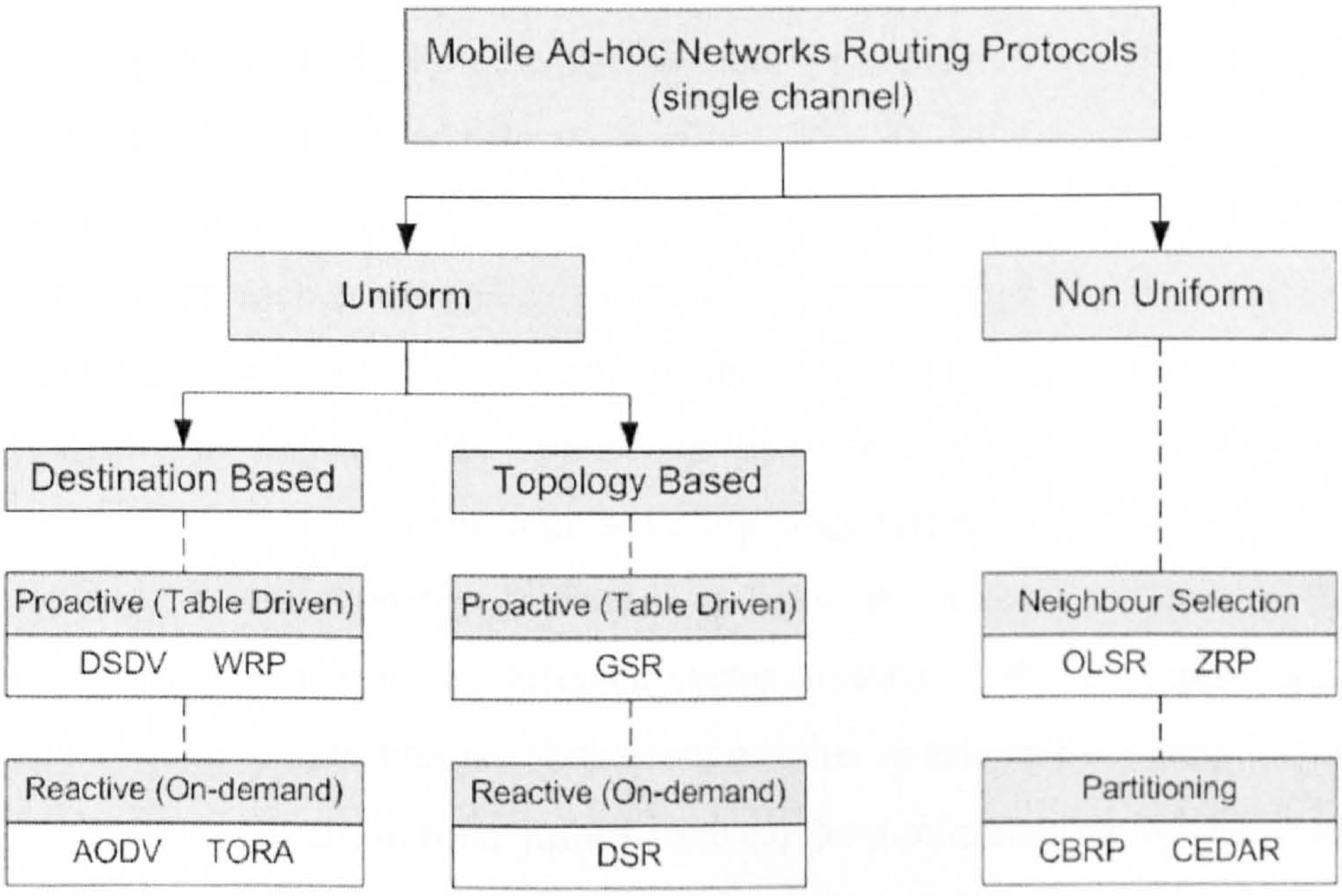


Figure 2.10 Taxonomy of ad hoc routing protocols

- *Communication Model:* Protocols can be divided according to communications model to protocols that are designed for *single-channel* or *multi-channel* communications. A large class of protocols assumes that nodes communicate over a single logical wireless channel. The IEEE 802.11 MAC method is the most widely used example for such a shared channel link layer. *Multi-channel* protocols are low level routing protocols which combine channel assignment and routing functionality. *Multi-channel* protocols utilize CDMA, FDMA or TDMA to form specific channels. Although communication can be much more efficient using such a method, it is difficult to be used in an ad hoc network since, usually, a distinguished controlling station is required to assign the channels.

- *Structure:* Routing protocols can be categorised according to node uniformity. Some protocols treat all the nodes uniformly, other make distinctions between different nodes. In *uniform* protocols, there is no hierarchy in network; all nodes send and respond to routing control messages at the same manner. *Non-uniform* protocols attempts to limit routing complexity by reducing the number of nodes participating in the route computation. *Non-uniform* protocols fall into two categories: protocols in which each node focuses routing activity on a subset of its neighbours (*neighbour selection*) and protocols in which the network is topologically partitioned (*partitioning*).
- *State Information:* Protocols may be described in terms of the state information obtained at each node and/or exchanged among nodes. Nodes participating in *topology-based* protocols maintain large scale topology information. This principle is the same as applied in link state protocols. *Destination-based* protocols do not maintain large scale topology information. They only maintain topology information that is needed to know the nearest neighbours. The best known such protocols are distance vector protocols, which maintain a distance and a vector to a destination (hop count or other metric and next hop).
- *Scheduling:* Obtain in route information can be continuous or regular or it can be triggered on-demand. Thus, protocols can be classified to proactive and on-demand protocols. *Proactive* or *table-driven* protocols constantly maintain knowledge of the status of the whole network, regardless of communication requests. In *on-demand* or *reactive* protocols, the routes are only calculated on an on-demand basis which means no unnecessary routing information is maintained. The route calculation process is divided into route discovery and route maintenance phases. The route discovery process is initiated when a source needs a route to a destination. The route maintenance process deletes failed routes and re-initiates route discovery. Table 2.2 summarises and compares the proactive and reactive characteristics.

Table 2.2 Proactive and reactive comparison

Parameters	Proactive (table-driven)	Reactive (on-demand)
Route availability	Always available irrespective of need	Computed when needed
Routing philosophy	Flat	Flat, except for CBRP
Periodic updates	Always required	Not required
Handling mobility	Updates occur at regular intervals	Use localized route discovery
Control traffic generated	Usually higher than on-demand	Increases with mobility of active routes
Storage requirements	Higher than on-demand	Depends on the number of routes maintained or needed
Delay	Small as routes are predetermined	High as routes are computed when needed
Scalability	Usually up to 100 nodes	Usually higher than table-driven

Table 2.3 summarises the salient feature of ad hoc routing protocols mentioned in Figure 2.10 that cover a range of design choices. Discussing in detail all of the above is not appropriate for thesis and thus only a selection will be considered. The prerequisites used for selection are:

- The protocols must be mature, with valid and well documented performance evaluation.
- They must be under active development.
- For obvious reasons, it is preferable to have protocols adopted by the IETF MANET working group.
- The protocols must be as representative of different categories as possible.
- Protocols too similar to each other or which have been shown to perform poorly with respect to the others are excluded; their inclusion would not add to existing knowledge.

Table 2.3 Salient feature of ad hoc routing protocols

Protocol	Salient Feature
DSDV (Destination Sequenced Distance Vector) [33]	DSDV uses destination sequence numbers to solve routing loop problem.
WRP (Wireless Routing Protocol) [34]	WRP utilise information about distance and second-to-last hop (predecessor) along the path to each destination.
AODV (Ad hoc On-demand Distance Vector) [35][36]	AODV minimizes the number of required broadcasts by creating routes on a demand basis.
TORA (Temporally Ordered Routing Algorithm) [37]	TORA is a highly adaptive loop-free distributed routing algorithm based on the concept of link reversal.
GSR (Global State Routing) [38]	GSR reduces the cost of disseminating link-state information by relying on periodic exchange of sequenced data rather than flooding.
DSR (Dynamic Source Routing) [35][39][40][41]	DSR is based on the concept of source routing and mobile nodes are required to maintain route caches that contain the source routes of which the mobile is aware.
OLSR (Optimised Link State Routing) [42][43]	OLSR requires the selection of MPR (multipoint relay) nodes, which has a special role.
ZRP (Zone Routing Protocol) [44]	ZRP protocol defines a routing zone; there is also some sort of clustering.
CBRP (Cluster Based Routing Protocol) [45]	CBRP forms clusters and thus requires cluster-heads, which are distinguished nodes.
CEDAR (Core-Extraction Distributed Ad hoc Routing) [46]	CEDAR forms a core network, which requires a special role for the nodes, which are part of the core.

The next sections of this chapter introduce three commonly used routing protocols: Dynamic Source Routing (DSR), Ad hoc On-demand Distance Vector (AODV) and Optimised Link State Routing (OLSR), which has the status of IETF Experimental RFCs. Finally, IETF is developing Dynamic MANET On-demand Routing (DYMO) [47] routing protocol which is successor to the popular AODV and DSR protocols and shares many of its benefits.

2.7.3 Dynamic Source Routing (DSR)

Dynamic Source Routing [35][39][40][41] is a reactive protocol. It has a flat structure, with all nodes treated equally by the routing algorithm. When sending a data packet with

DSR, the originating node includes the route in the packet's header, a mechanism known as source routing. Routes between nodes are determined dynamically, on an as-needed basis. Therefore, there is no need for periodic route broadcasts. The basic mechanisms behind DSR are known as route discovery and route maintenance.

Route discovery mode is entered when a source node wants to send a data packet to the destination for which it has no route in the cache. When this happens, a *route-request* (RREQ) packet is flooded throughout the network. Each node, upon receiving a RREQ packet, rebroadcasts the packet to its neighbours if it has not forwarded already or if the node is not the destination node, provided the packet's time to live (TTL) counter has not been exceeded. Each RREQ carries a unique ID generated by the source node and the path it has traversed. A node, upon receiving a RREQ packet, checks the ID on the packet before forwarding. The packet is forwarded only if it is not a duplicate RREQ. The ID in the packet is used to prevent loop formations and to avoid multiple transmissions of the same RREQ by an intermediate node that receives it through multiple paths. Thus, all nodes who accept the destination forward a RREQ packet during the route discovery phase. A destination node, after receiving the first RREQ packet, replies to the source node through the reverse path the RREQ packet has traversed. In Figure 2.11, source node S initiates a RREQ packet to obtain a path for destination D. This protocol uses a route cache that stores all possible information extracted that is contained about the source route in a data packet. Nodes can also learn about the neighbouring routes traversed by data packets if operated in the promiscuous mode[7]. The route cache is also used during the route discovery phase. If an intermediate node receiving a RREQ has a route to the destination node in its route cache, it then replies the source node by sending a *route-reply* (RREP) with the entire route information from the source node to the destination node. RREQ packets propagate through the network until reaching the destination node. On receiving a RREQ, the

---

[7] The mode of operation in which a node can receive the packets that are neither broadcast nor addressed to itself.

destination node sends a RREP packet back to the source node as shown in Figure 2.12. Alternatively, it can allow the original route discovery procedure if no route is available.

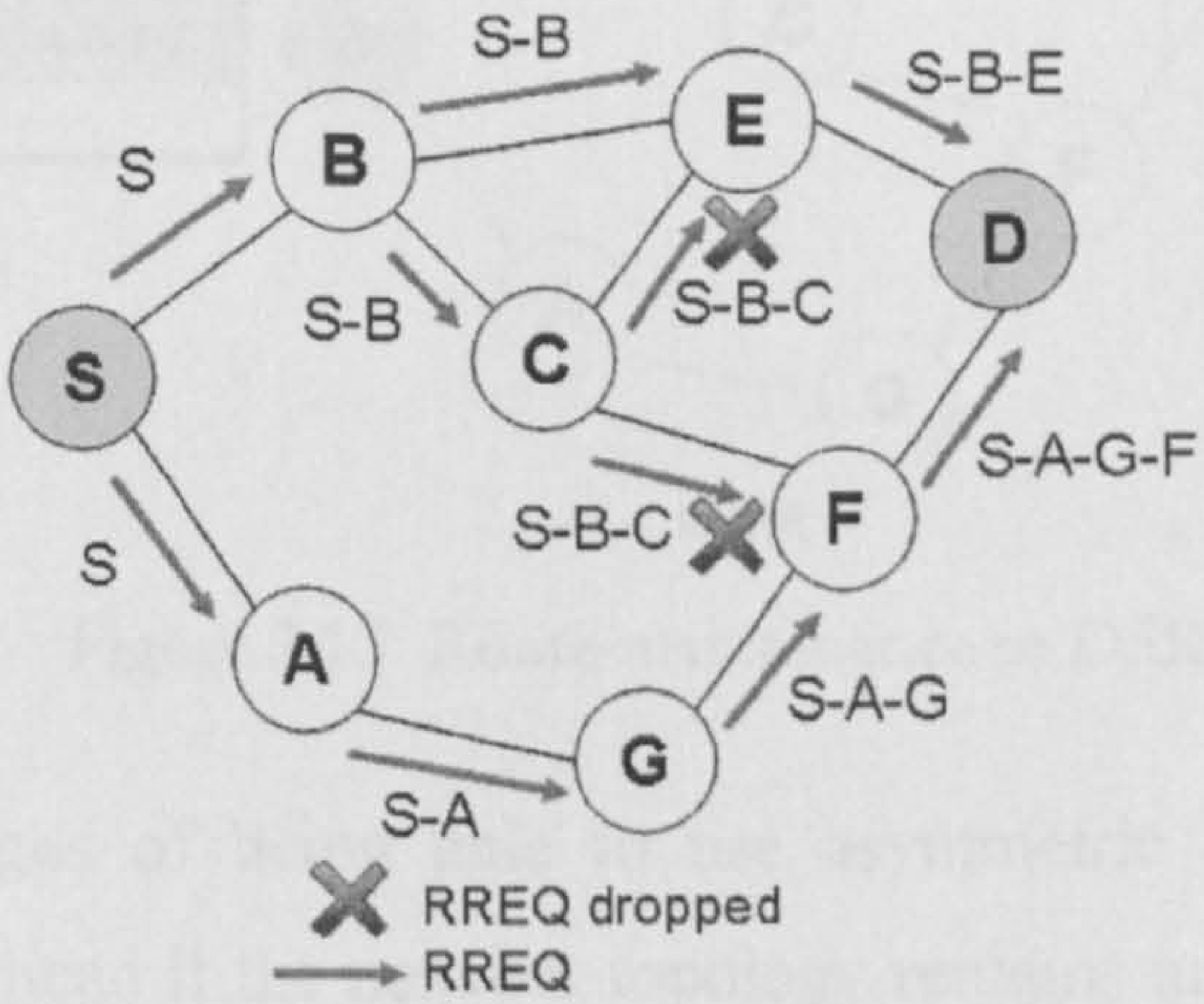


Figure 2.11 Route request in DSR

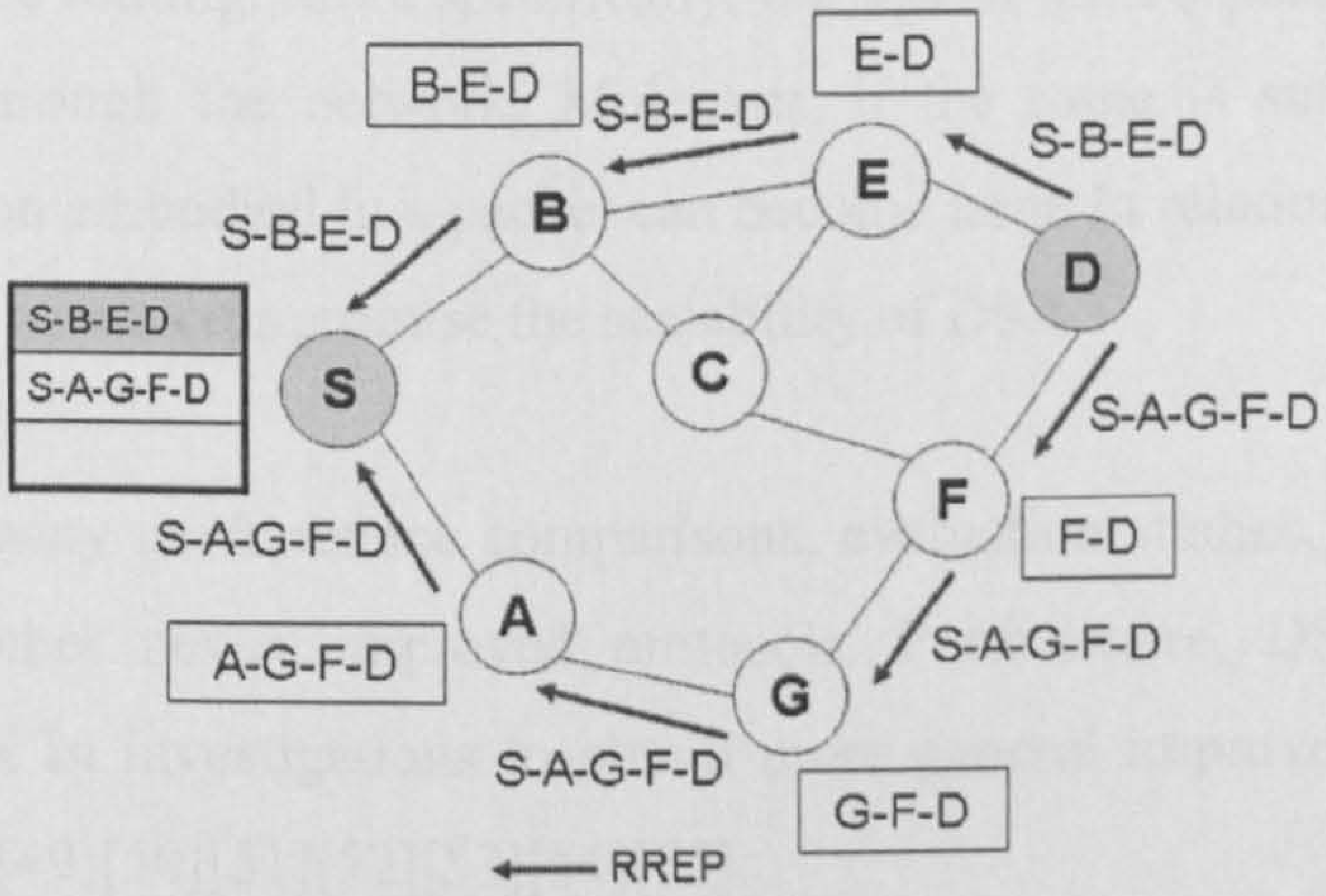


Figure 2.12 Route reply in DSR

The route maintenance procedure is called when a broken link is encountered while forwarding data packets on a previously known route, for example, the link between nodes E and D in Figure 2.13. According to this procedure, the last node (E) on the route to successfully receive the data packet sends a *route-error* (RERR) packet back to the data packet's originator. On receiving the RERR packet, the originator removes routes including the broken link from its cache. To transmit a packet to the required destination, an alternative route from the originator's cache will be used or the route discovery procedure will be initiated again.

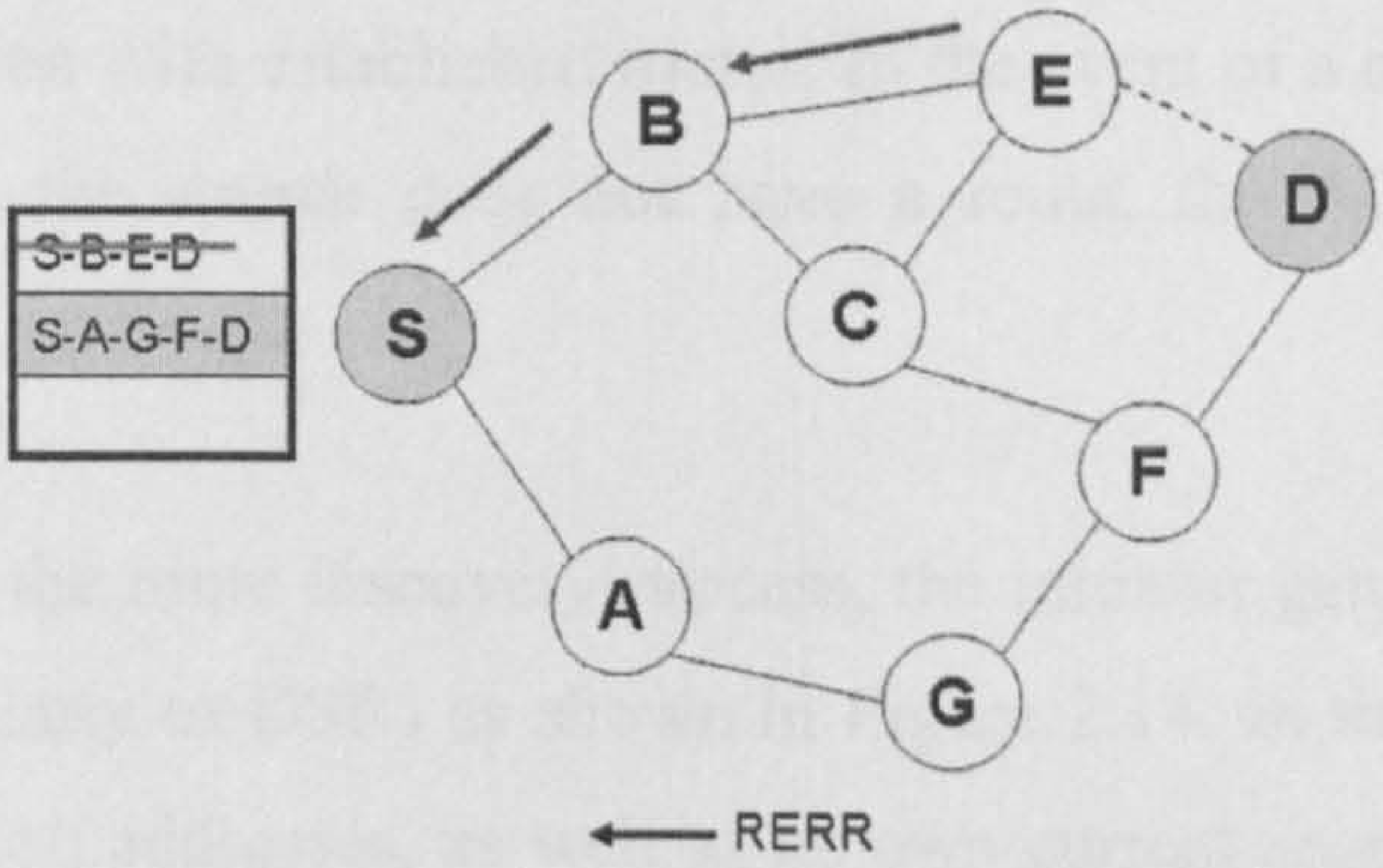


Figure 2.13 Route maintenance in DSR

DSR has the advantages of being able to use asymmetric links and generating no additional routing overhead if the network topology remains unchanged. Moreover, it is easily implemented and supports multipath routing. The main disadvantage of DSR lies in the use of source routing. More specifically, the size of RREQ packets can increase as they propagate through the network. Moreover, if the route is sufficiently long, the routing information embodied in a packet can become large in relation to data carried by the packet. These issues compromise the scalability of DSR.

DSR is used in many performance comparisons, evaluation studies, and is used as the benchmark for other newer improved protocols. Furthermore, DSR is used as the reference protocol in investigations to obtain more general improvements in MANET performance [48][49][50][51][52][53][54][55].

2.7.4 Ad hoc On-demand Distance Vector (AODV)

Ad hoc On-Demand Distance-Vector (AODV) routing [35][36] can be considered as a reactive derivative of the earlier, proactive DSDV protocol [33]. It employs destination sequence numbers to identify the most recent path. The major difference between AODV and DSR stems from the fact that DSR uses source routing in which a data packet carries the complete path to be traversed. However, in AODV, the source node and the intermediate nodes store the next-hop information in a routing table corresponding to each flow for data packet transmission. This routing table is used during normal operation with established routes. In the event of a destination node being requested, for which the source does not have a route, the AODV route discovery process is initiated.

More analytically, in the route discovery process, the initiator generates a *route-request* (RREQ) packet (similarly to DSR) as shown in Figure 2.14. In this packet, it stores its own and the destination addresses, as well as its own current sequence number and the destination's last known sequence number. Per node sequence numbers are a characteristic feature of AODV, utilised to ensure loop freedom. Each sequence number is incremented by the node whenever there is a change in its local connectivity information. Next, the source node floods the RREQ packet into the network. On reception of a RREQ, a node creates a reverse route entry for the source node in its routing table and checks for a valid route to the destination. In the situation where the node finds a valid route to the destination, with a sequence number greater than or equal to the one contained in the RREQ, does the node reply. The node also sends a reply if it is the destination itself. If neither of these conditions is met, the node rebroadcasts the RREQ.

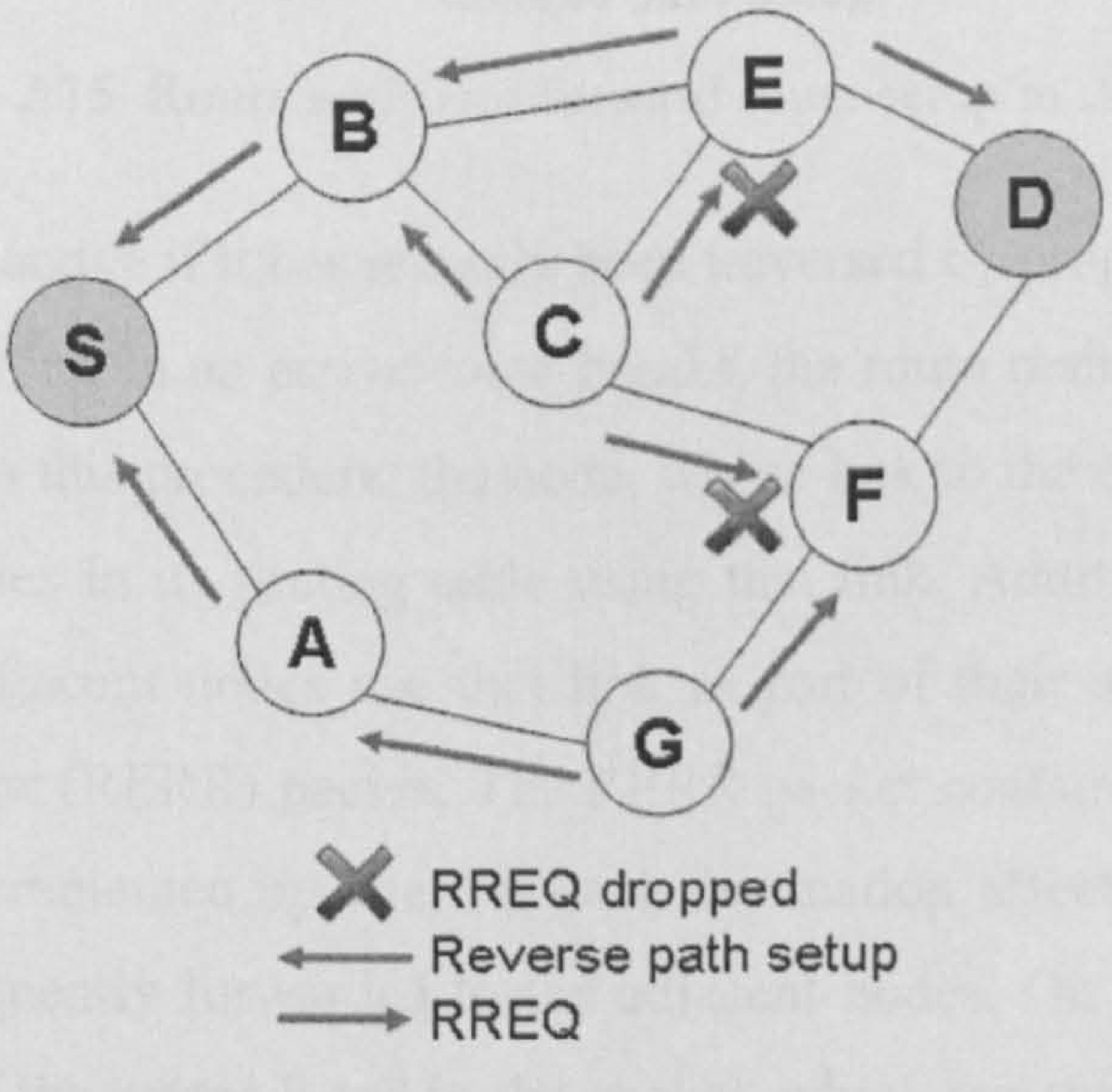


Figure 2.14 Route setup and reverse route setup in AODV

Conversely, if one of these conditions is met, the node generates a *route-reply* (RREP) packet. The RREP, which contains the current sequence number of, and distance to, the destination, is than sent back to the source through the reverse path. On receiving the RREP, an intermediate node creates a forward route entry for the destination node in its route table, and subsequently forwards the RREP to the source node. Only when the source node receives the RREP, can it begin using the route to transmit data packets to the destination. RREP and forward route setup is depicted in Figure 2.15.

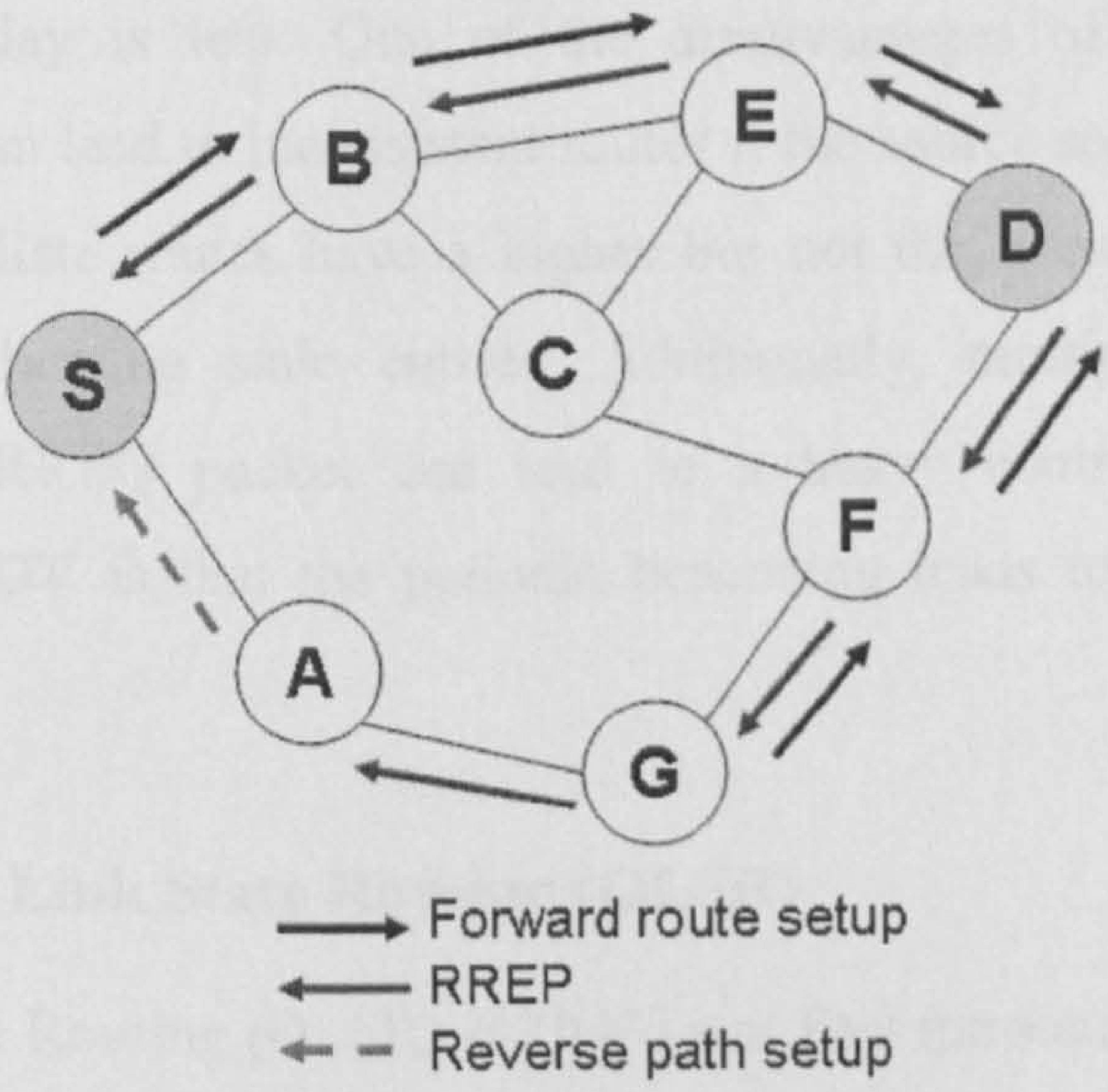


Figure 2.15 Route reply and forward route setup in AODV

A route is considered active if it has recently been traversed by propagating data packets. When a link participating in an active route breaks, the route maintenance procedure is initiated. According to this procedure, the node, whose link to the next hop is the broken one, disable any entries in its routing table using that link. Additionally, it determines whether any of its adjacent nodes use that link as part of their active routes. If so, it generates a *route-error* (RERR) packet. The RERR packet contains the address and the sequence number, incremented by one, for each destination affected by the link break. This RERR is subsequently forwarded to the adjacent nodes. On receiving a RERR, a node disables each of the routes listed in the packet, which use the source of the RERR as a next hop. If one or more routes are deleted, the node than generates and broadcasts it own RERR message. If a source node receiving a RERR determines that it still needs any of the disabled routes, it re-initiates route discovery for that route.

A large number of ad hoc related papers cite AODV as a reference. However, some papers did an independent comparison between some ad hoc routing protocols including AODV [50][51][52]. The main advantage of this protocol is that routes are established on demand and sequence numbers are used to find the latest route to the destination. The connection setup delay is less. One of the disadvantages of this protocol is that intermediate nodes can lead to inconsistent routes if the source sequence number is very old and the intermediate routes have a higher but not the latest destination sequence number, thereby indicating stale entries. Additionally, multiple RREP packets in response to single RREQ packet can lead to a heavy control overhead. Another disadvantage of AODV is that the periodic beaconing leads to unnecessary capacity consumption.

2.7.5 Optimised Link State Routing (OLSR)

Optimised Link State Routing (OLSR) [42][43] was first introduced as an IETF draft to the MANET working group in 1998. It is a proactive point-to-point routing protocol based on traditional link-state algorithms such as OSPF. In this scheme, each node maintains topology information about the network by periodically exchanges link-state messages. The originality of OLSR is that it minimises the size of each control message as well as the number of rebroadcast nodes during each route update by employing a multipoint relaying strategy. More specifically, during each topology update, each node in the network selects a set of neighbouring nodes to forward topology information. Nodes included in this set are called the multipoint relays (MPRs) of that node. Any node that is not included in the set can read and process each packet but does not forward it.

To select the MPRs, each node periodically broadcasts a list of one-hop of neighbours using hello messages. A node receiving a hello message can determine, from the contents of the message, the subset of the two-hop neighbours, which are accessible through the originator of the message. From its knowledge of all two-hop neighbours and the one-hop neighbours through which they are accessible, the node calculates the set of MPRs as the minimum set of one-hop neighbours required to reach all two-hop neighbours (see Figure 2.16). Each node forwards the set of MPRs to its neighbours in the hello message. This information is then used to update each node's record of which nodes are using it as an MPR, i.e. its MPR selectors. It is obvious from the above that the OLSR requires bi-directional links.

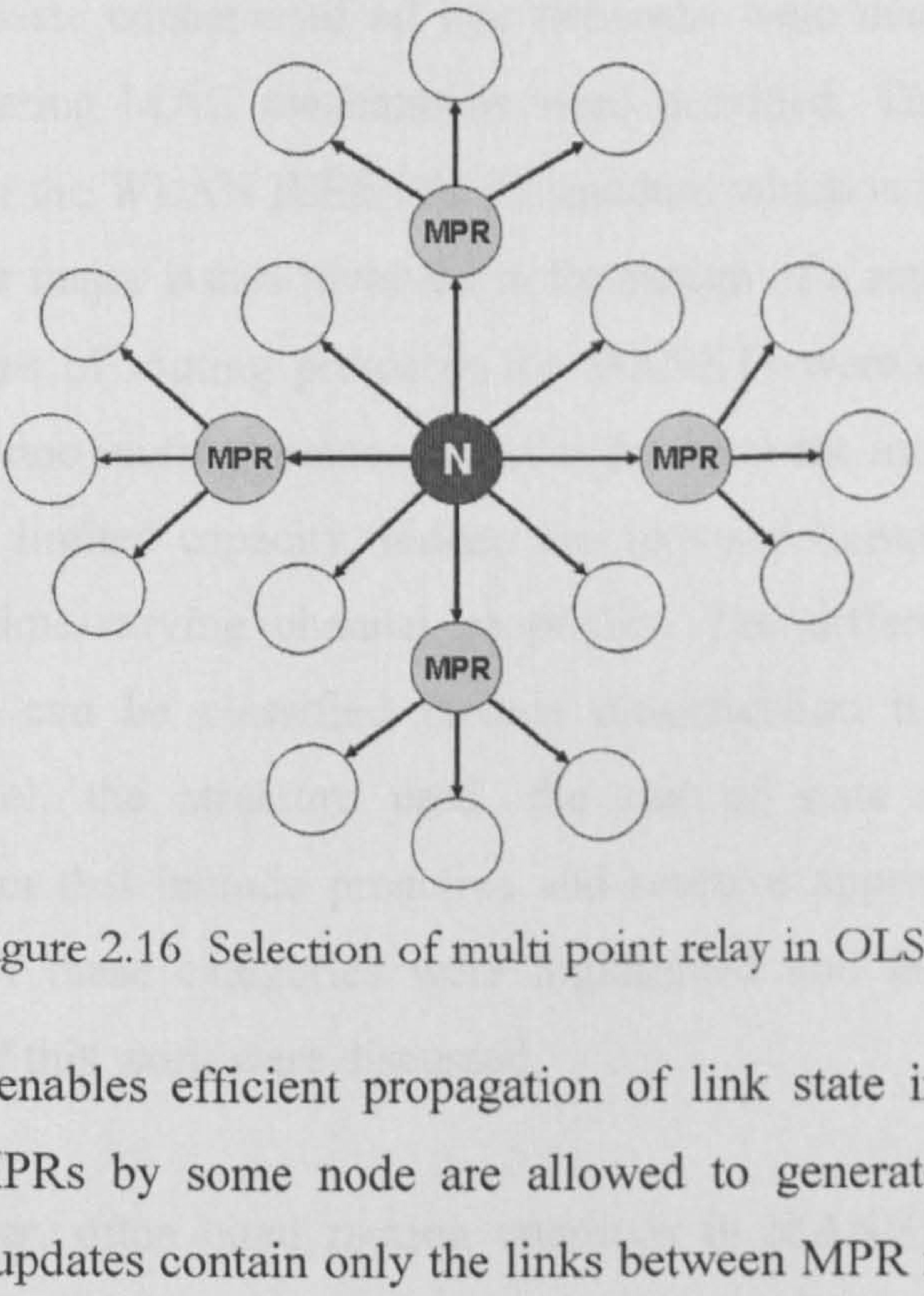


Figure 2.16 Selection of multi point relay in OLSR

The MPR selection enables efficient propagation of link state information; only the nodes selected as MPRs by some node are allowed to generate link state updates. Moreover, link state updates contain only the links between MPR nodes and their MPR selectors, in order to keep the update size small. Thus, only partial topology information is made available at each node. However, this information is sufficient for each to locally compute shortest hop path to every other node because at least one such path consists of only MPR nodes.

The main advantage of OSLR against other proactive protocols is the reduced number of broadcasts, due to the use of MPRs, which leads to reduced control overhead. One of its main disadvantages is the overlapping MPR sets and MPR selectors which lead to duplicate messages. There were few a performance comparisons: P. Jacquet et al [56] undertook an analytical comparison of OLSR with DSR in a random graph model, and L. Christensen et al [57] undertook a detailed comparison of OLSR with AODV where OLSR appeared to offer in many cases improved performance.

2.8 Summary

In this chapter, the basic concepts of ad hoc networks were discussed. The different classifications of existing MAC mechanisms were provided. This chapter also gave detailed description of the WLAN IEEE 802.11 standard which is widely used in ad hoc networks. Further, the major issues involved in the design of a routing protocol and the different classifications of routing protocols for MANETs were described. The major challenges that an ad hoc routing protocol must address are the mobility of nodes, rapid changes in topology, limited capacity, hidden and exposed terminal problems, limited battery power and time-varying channel properties. The different approaches upon which the protocols can be classified include classification based on the type of communication model, the structure used, the use of state information and the scheduling mechanism that include proactive and reactive approaches. The protocols belonging to each of these categories were highlighted and key protocols that are possible candidates of this work were discussed.

A particularly popular, often cited routing approach in MANETs is reactive or on-demand routing. Most proposed schemes of this type use a single route for each session. The routes that a source has acquired can deteriorate due to mobility and packets can be lost. On-demand multipath routing schemes can alleviate these problems by establishing multiple paths between a source and a destination within a single route discovery process. Chapter 3 discusses multipath routing and performance evaluation methodology, which form the basis through out this study.

3. Multipath Routing and Evaluation Methodology

3.1 Introduction

The provisioning of real-time services such as voice and video over ad hoc networks is problematic since wireless links are unreliable and are of limited bandwidth. Compared to the requirements of data-only applications, these real-time requirements demand high packet delivery rates, low delays, negligible jitter and lower routing overheads. The design of an efficient and reliable routing protocol for such applications is a major challenge.

A particularly popular, often cited routing approach in MANETs is reactive or on-demand routing (discussed in section 2.7). Reactive schemes build routes only when a source node seeks to transmit data packets to a destination node. The source floods the network with packets that search the destination and discover the route. Most proposed schemes of this type (for example, Dynamic Source Routing (DSR) [41] and Ad hoc On-demand Distance Vector (AODV) [36]) use a single route for each session. However, in case of a failure of an active link between source and destination, the routing protocol must invoke a route maintenance process followed by route discovery process and, in so doing, additional delay is incurred and overheads increase [50].

On-demand multipath routing schemes [58][59][60][61][62][63][64][65][66][67][68] alleviate problems with single-path on-demand approaches by establishing multiple paths between a source and a destination within a single route discovery process. Such schemes are typically proposed in order to increase the reliability of data transmission (i.e., fault tolerance) or to provide load balancing. Although these schemes build multiple routes on-demand, most only consider node-disjoint or link-disjoint paths. In this thesis, a multipath routing scheme is proposed, which discovers multiple partial-disjoint paths to the destination.

The first part of this chapter give a full description to the implementation issues associated with the development and design of a multipath routing and discusses related published work (Section 3.2). The second part of the chapter presents the evaluation methodology adopted in this thesis and is organised is follows. Section 3.3 highlights common evaluation techniques. Section 3.4 presents an overview of NS-2 packet level simulation used throughout this work. Section 3.5 presents simulation model of DSR and its implementation in NS-2. Section 3.6 discusses methods used during the analysis of the generated results. Section 3.7 presents validation of the routing protocol simulation models. Finally, the chapter is concluded in section 3.8.

3.2 Multipath Routing

Multipath routing is a concept that has been applied to several applications. However, the fundamental difficulties presented in MANETs caused by node mobility and communication over a wireless medium make new study of multipath routing schemes within this context interesting. This section briefly describes benefits and components of multipath routing, related published work and open issues.

3.2.1 Benefits of Multipath Routing

- *Fault tolerance:* Multipath routing schemes can provide fault tolerance in the presence of route failures by routing redundant information to the destination through alternative paths. This increase in route resiliency is largely dependent on factors such as the degree of the network and disjointness of the available paths.

- *Load balancing:* Load balancing is of special importance in MANETs because of the limited bandwidth between the nodes. Load balancing can be achieved by distributing the traffic along multiple routes [64][65]. This can alleviate congestion and bottlenecks and improve the overall quality of service (QoS).
- *Bandwidth aggregation:* The effective bandwidth can be aggregated by distributing traffic to the same destination into multiple paths. This approach is particularly useful when a node has multiple low bandwidth links but requires a bandwidth greater than an individual link can provide. End-to-end delay may also be reduced directly as a result of the availability of increase aggregate bandwidth.
- *Reduced delay:* Single path on-demand routing protocols use a single route for each session. However, in case of a failure of an active link between source and destination, the routing protocol must invoke a route maintenance process followed by route discovery process and, in so doing, additional delay is incurred. Multipath routing schemes alleviate this problem by establishing multiple paths between a source and a destination within a single route discovery process.

3.2.2 Components of Multipath Routing

The following components are identified as fundamental to multipath routing algorithms [66][67][68][69]. Although all of them need not be present in a routing scheme, they all ought to be considered in any multipath design. These concepts can be used as building blocks for designing a new multipath scheme or as features for comparison among existing multipath routing schemes.

3.2.2.1 Route Discovery

Route discovery is the process of determining the available routes between a source and destination node. This procedure generally includes provisions to avoid route looping and heuristics to generate disjoint multipath sets. There are three main types of path disjointness, namely *link-disjoint*, *node-disjoint* and *partially-disjoint*. A set of link-disjoint routes have no common links, but may share some common intermediate nodes (Figure 3.1). Link-disjoint routes are not as resilient to geographically localised and correlated failures. Node-disjoint routes have no common nodes except the source and the destination (Figure 3.2). Node-disjoint routes are least abundant and are hardest to find. It has been shown that in moderately dense networks, only a small number of node-disjoint routes may exist between any two arbitrary nodes, especially as the distance between the nodes increases [63]. A more relaxed form of the multipath routing scheme is to construct partially-disjoint (also known as non-disjoint) routes (Figure 3.3). The formation of partially disjoint routes is mostly for reliability purposes. The primary route is used for data transmission, while other partially-disjoint routes are backup routes for failure recovery. Because there are no restrictions that require the routes to be node-disjoint or link-disjoint, a greater number of partially-disjoint routes exist in a given network than node-disjoint or link-disjoint routes.

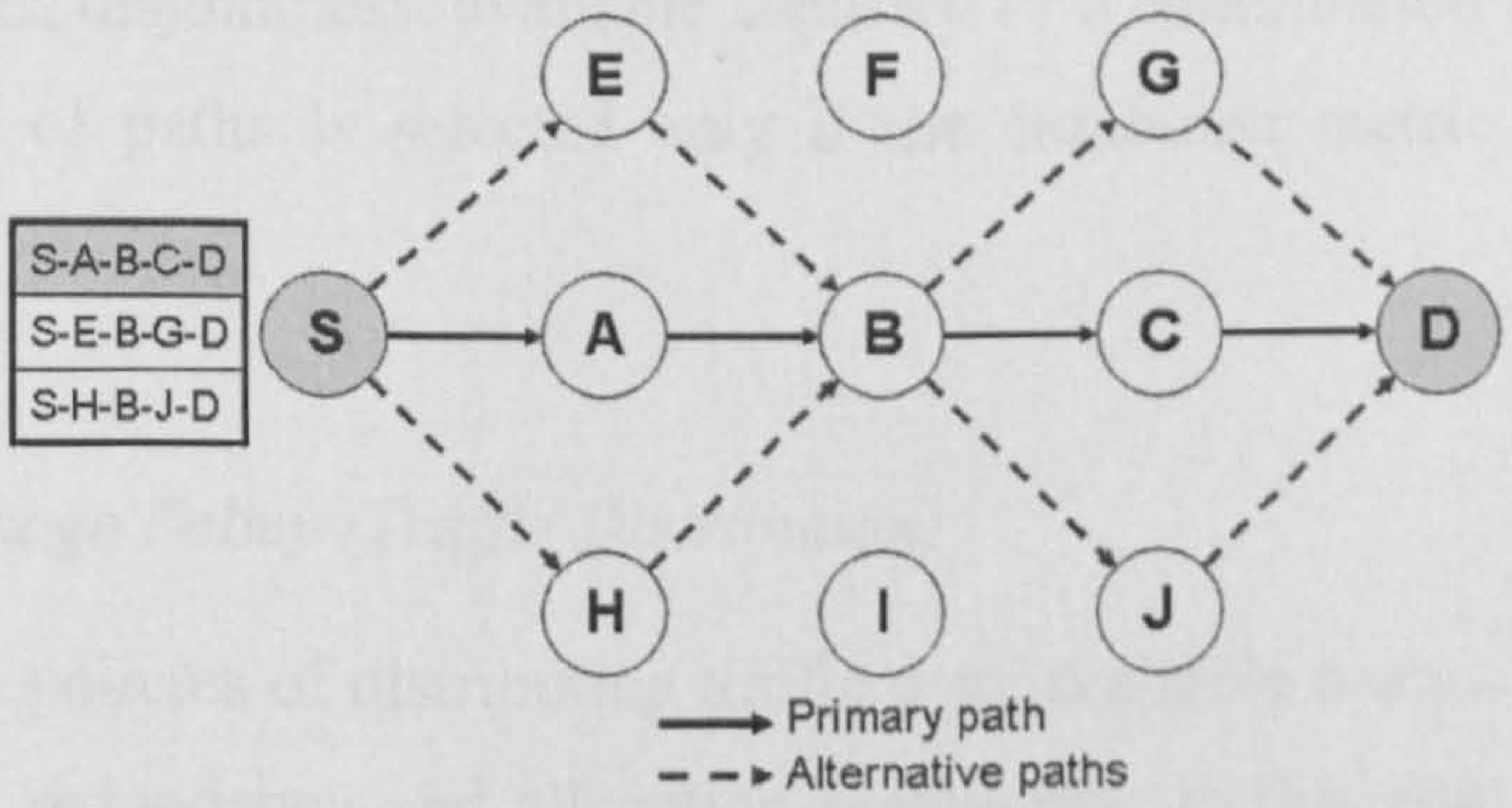


Figure 3.1 Link-disjoint multiple paths

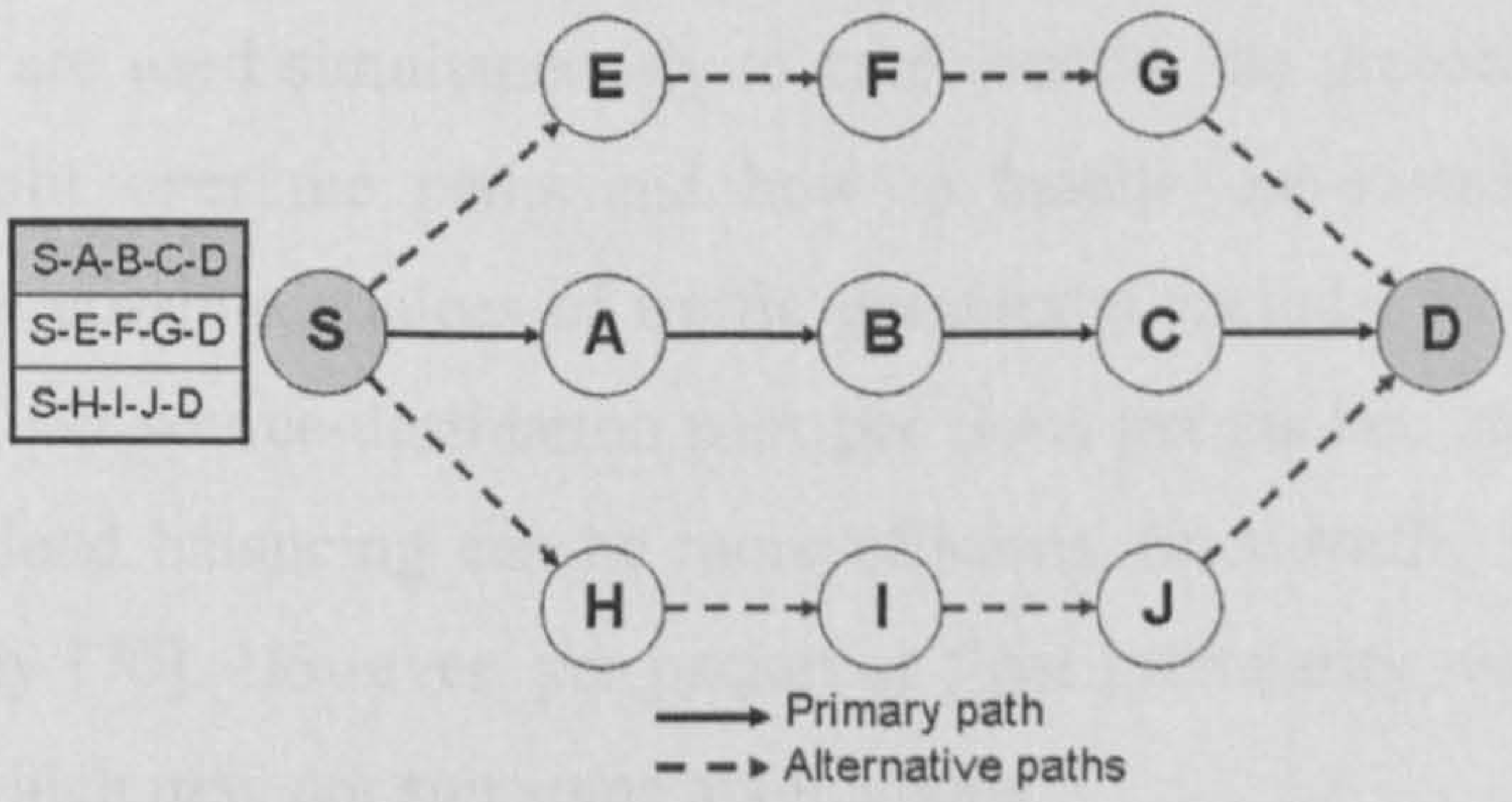


Figure 3.2 Node-disjoint multiple paths

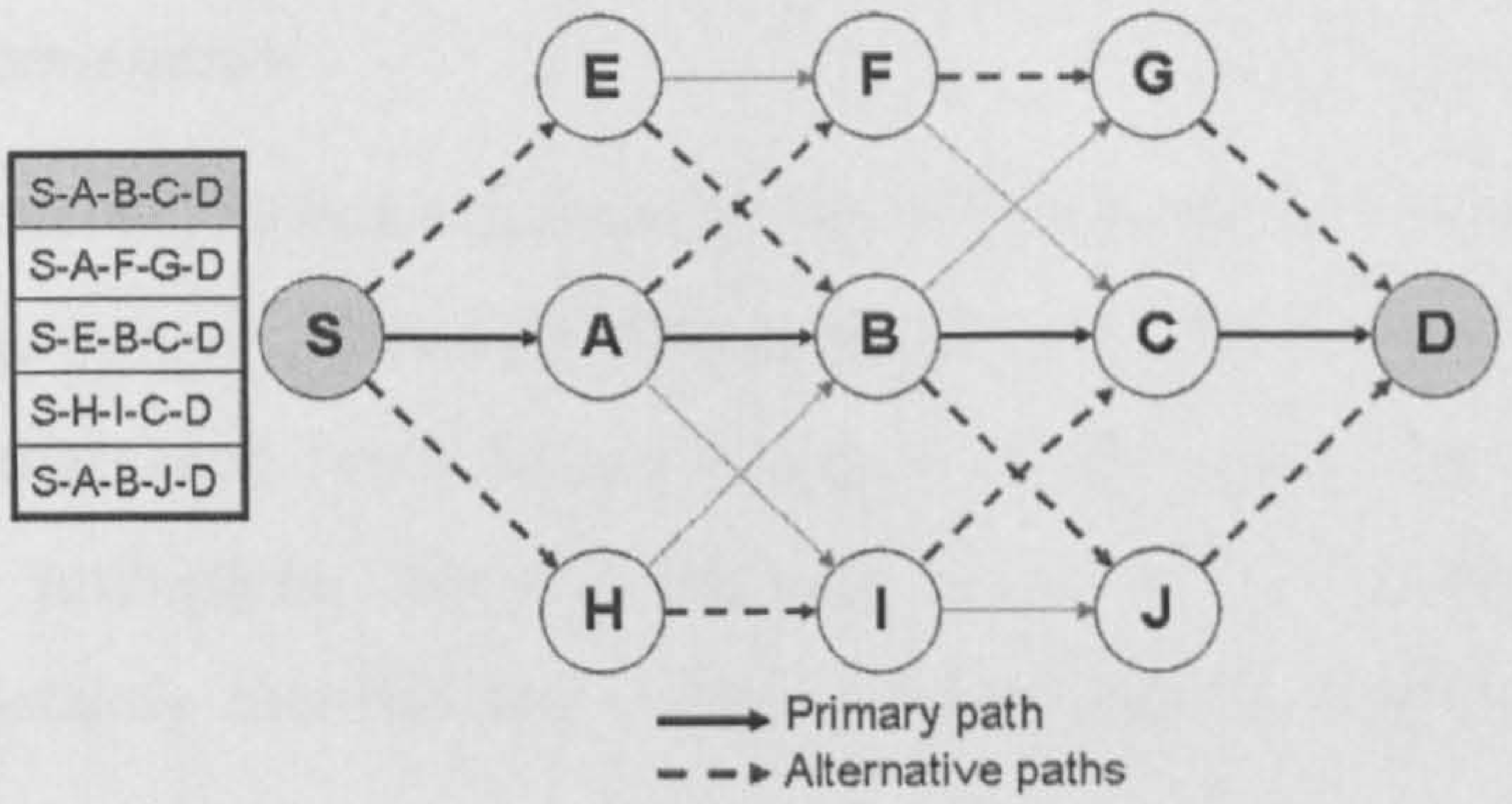


Figure 3.3 Partial-disjoint multiple paths

3.2.2.2 *Filtering Provision*

Filtering provision is the option of choosing a subset of alternative paths according to certain quality of the paths. Some filtering provisions are path length or hop-count, shortest multipaths, disjointness, available capacity, or a combination of metrics. In QoS routing, a subset of paths is selected only if the combined metric satisfies the QoS requirement.

3.2.2.3 *Route Usage Policy (Traffic Distribution)*

There are various policies of distributing traffic over available routes. Examples of such a policy are path redundancy and allocation granularity. In the case of redundant path usage, the source node can choose to use a single path and keep the rest as backups, or it can utilise multiple paths in a round-robin fashion, with only one path sending at a time. If multiple paths are used simultaneously to carry traffic, the protocol needs to decide how traffic is split over the paths and how to handle out-of-order packets at the destination. Some possible choices of traffic granularity include, in order of increased control overhead, per source-destination pair, per flow, per packet, per segment. With a fine granularity, load balancing can be more efficient, since traffic fluctuation can be adapted to quickly [70]. However, per packet or finer granularity require reordering at the destination, which may not suit some applications.

3.2.2.4 Route Maintenance

Routes may fail due to link/node failures or, in ad hoc networks, node mobility. Route maintenance is the process of re-establishing routes after the initial route discovery. It can be initiated after each route failure or when all the routes have failed to enable efficient use of multipaths. Some multipath schemes use dynamic maintenance algorithms to constantly monitor and maintain the quality or combined QoS metric of available paths.

3.2.3 Related Work

Numerous multipath routing schemes have been proposed for ad hoc networks [58][59][60][61][62][63][64][65][66][67][68]. Most of these schemes compute link-disjoint and node-disjoint paths. This section presents a selection based on DSR and AODV.

3.2.3.1 Schemes based on DSR

Multipath-DSR (M-DSR) [58] presents two slightly varied versions of multipath extensions of the popular DSR. Instead of replying when the first RREQ is received as with DSR, the destination node sends an additional RREP for a RREQ which carries a link-disjoint route compared with the primary path. In the first of the proposed extensions, only the source node has the responsibility to switch to an alternative route in response to intermediate link failures. Such a situation will cause a temporary loss of route until the source receives a RERR message and switches to a new route. The second proposed extension alleviates this problem by allowing intermediate nodes to have one alternative route and to switch route as soon as the primary path fails. Upon link failure, intermediate nodes do not send a RERR message to the source node but adjust the packet's header with the alternative route. In such a way, traffic currently in transit is not lost. It was shown analytically that multipath outperforms single-path routing in terms of route re-discovery and longer alternative routes are less beneficial. However, in many cases, M-DSR cannot compute link-disjoint paths because intermediate nodes drop every duplicate RREQ that may invoke another link-disjoint path.

Multipath Source Routing (MSR) [59] extends DSR's route discovery and route maintenance phases to compute multiple node-disjoint paths which are used simultaneously to distribute traffic. It proposes a mechanism to distribute load over multiple paths, based on the round-trip time measurement such that paths with lower delays receive a greater proportion of traffic. The SRPing tool[8] is used to estimate the end-to-end delay. A multipath table is used to maintain information of routes to a destination. This table contains for each route to the destination: the index of the path in the route cache, the destination ID, the delay and the calculated cost weight based on the delay of all paths to the same destination. Such an approach has the disadvantage, pertinent to real-time services, that traffic flows over the different paths may be unbalanced and requires re-sequencing and subsequently incurs additional delay at the receiver.

The Split Multipath Routing (SMR) [60] scheme is an on-demand multipath source routing that provides a way of determining maximally disjoint paths. Unlike DSR intermediate nodes are not allowed to respond to RREQs from their cache. In order to maximise the number of RREQs reaching the destination, it uses a modified RREQ packet flooding scheme in the route query process. Intermediate nodes do not necessarily discard any duplicate RREQ packets received. The destination node, upon reception of first RREQ, responds immediately to the source node and then waits for a pre-allocated time to receive additional RREQs to determine a second maximally disjoint route. Once the source node receives the first RREP, it starts sending traffic to the destination to avoid additional delay to the currently buffered packets. When the second RREP is received, the source node splits traffic using per-packet allocation scheme [70][64] in a round-robin fashion. This approach has the disadvantage of transmitting a large number of RREQ packets.

---

[8] SRPing is a measurement tool to get the round trip time between two arbitrary nodes.

3.2.3.2 Schemes based on AODV

Ad hoc On-demand Multipath Distance Vector (AOMDV) [61] routing is an extension to AODV. It computes multiple loop-free, link-disjoint or node-disjoint paths during route discovery. AOMDV uses the notion of advertised hop-count and last hop that ensures loop freedom and path disjointness respectively. Unlike AODV, in AOMDV the intermediate node does not simply drop the duplicate RREQ packet but examines it to see if it contains a node-disjoint path to the destination. If so, then the node checks to see if the reverse path to the source is available. If this is also true, then the path is added in the table. In the case of a link-disjoint path, the node applies a slightly lenient policy and replies to a certain number of RREQs, which come from disjoint neighbours. AOMDV can only discover equal length multiple paths, which may not be possible when the network is moderately to sparsely connected.

Node-Disjoint Multipath Routing (NDMR) [62] modifies and extends AODV to enable the path accumulation feature of DSR to build multiple node-disjoint paths. It exhibits a low routing overhead during the route discovery process by recording the shortest routing hops of loop free paths. The source node selects one of the alternative paths in case of a link-failure. This approach, however, has similar disadvantages as schemes based on AODV and DSR and suffers from scalability issues.

Scalable Multipath On-demand Routing (SMORT) [63] scheme reduces the routing overhead incurred in recovering from route breaks through the use of secondary paths. SMORT (pre-) computes fail-safe multiple paths, providing all intermediate nodes on the primary path with multiple routes (if they exist) to the destination. A path between source and destination is said to be fail-safe with respect to the primary path, if it bypasses at least one intermediate node on the primary path. The number of RERR packets transmitted during route break recovery is reduced which, in turn, reduces end-to-end delay and overheads whilst providing improved scalability. This approach has the disadvantage of generating/recording multiple copies of RREQ packets at each node, thus increasing memory overheads.

3.2.4 Comparison and Open Issues

Table 3.1 summarises and compares the main characteristics of multipath routing schemes presented.

Table 3.1 Summary of multipath routing schemes

	M-DSR [58]	MSR [59]	SMR [60]	AOMDV [61]	NDMR [62]	SMORT [63]
Base protocol	DSR	DSR	DSR	AODV	AODV	AODV
Alternative / backup paths	Link-disjoint	Node-disjoint	Maximally link- or node-disjoint	Link- or node-disjoint	Node-disjoint	Fail-safe
Routing choice made at	Source and intermediate nodes	Source	Source	Intermediate nodes	Source	Intermediate nodes
Complete routes known at source	Yes	Yes	Yes	No	Yes	No
Delay known	No	Yes	No	No	No	No
Paths used simultaneously	Possibly (not specified)	Yes	Yes	Possibly (not specified)	Possibly (not specified)	No
TTL limitation	Yes	Yes	Yes	Yes	Yes	Yes
QoS support	No	Possibly	No	No	No	No
Multicast support	No	No	No	No	No	No
Power management	No	No	No	No	No	No
Security support	No	No	No	No	No	No
Motivation / application	Reduce frequency of route discovery floods	QoS applications with soft end-to-end reliability	Splitting traffic provides better load distribution	Discovers disjoint paths without using source routing	Reduce routing overheads using path accumulation feature of DSR	Discovering fail-safe paths provides better scalability

Although on-demand multipath routing schemes can build multiple routing paths between source and destination, most of them will encounter a broadcast storm[9] of routing packets in the process of discovering multiple disjoint routing paths. This increases considerably routing overhead in the network [59][60][61][62][63]. Also, disjoint multipath routing schemes discussed previously compute maximally disjoint multiple paths, whose availability is lower due to the disjointness restriction imposed in the path selection. Hence, disjoint multiple paths cannot provide efficient fault-tolerance towards route breaks[10]. Also, these schemes involve significant delay and overhead in selecting disjoint multiple paths.

In the next chapter, a novel and practical Shortest Multipath Source (SMS) routing scheme is proposed to solve these issues.

3.3 Performance Evaluation Methodology

Performance is often a key issue in the design, development and configuration of systems. It is not always enough to know that systems work properly; they must also work effectively and fairly. Performance analysis studies are conducted to evaluate existing or designed systems, to compare alternative configurations or to find an optimal configuration of a system. The primary step in performance evaluation is to select the right evaluation technique. In general, there are three techniques used when investigating the protocols for MANETs [71]: measurements, analytical modelling and simulation.

- *Measurements* involve physically monitoring the actual system. However, such approach incurs high cost or disruptive changes to the system. For example, to measure the performance of the routing schemes for a 50 node network moving over a 700x500 meters flat space considered in this thesis, hardware such as notebooks could be deployed to measure the protocol performance, but the cost per notebook may make such approach exorbitant by expensive. Also there will

---

[9] The broadcast storm problem is specifically discussed later in Section 4.2.2.
[10] The fault-tolerance analysis is specifically discussed later in Section 4.2.1.

be difficulty associated when measuring performance of mobile scenarios in the said field dimensions.

- *Analytical modelling* approach has, essentially, two components; first come up with a way to describe a system mathematically with the help of applied mathematical tools such as queuing and probability theories, and then second apply numerical methods to gain insight from the developed mathematical model. When the system is quite simple and relatively small, analytical modelling would be preferable (over simulation). Here the model needs to be mathematically tractable. The numerical solutions to this model in effect require lightweight computational efforts. If properly employed, analytical modelling can be highly cost-effective and therefore can potentially provide an abstract view of the components interacting with one another in the system. However, if many simplifying assumptions on the system are made during the modelling process, analytical models may not give an accurate representation of the real system.
- *Simulation* uses a computer to simulate real networks in terms of structures and operations. Compared to analytical modelling, simulation usually requires lower degrees of abstraction in the model (i.e., fewer simplifying assumptions) since more details of the specifications of the system can be included in the simulation model. When the system is rather large and complex, a simple mathematical formulation may not be feasible. In this case, the simulation approach is usually preferred to the analytical approach. Further, even if a system is available for measurement, a simulation model may be preferred over measurements because it allows the alternatives to be compared under a wider range of workloads and environments.

Consequently, a simulation model is used throughout this work, as the network structure being represented is complex.

3.3.1 Types of Simulations

Among the variety of simulation methods described in the literature [71][72][73], two important methods are Trace-Driven Simulation and Discrete-Event Simulation.

3.3.1.1 Trace-Driven Simulation

A trace is a time-ordered record of events on a real system. A simulation using a trace as its input is a trace-driven simulation. Trace-driven simulations are common in computer systems since they generally have built-in tracing programs which monitor the activities of the system and the sequences of processes pertinent to the planned simulation. The major advantage of trace-driven simulation is that it avoids much of the statistical work. This removes a potentially difficult and cumbersome activity. A second advantage of such approach is that it is relatively easy to verify the model. The main problem with trace-driven approach is that the scope of application is very small. This approach is really only applicable to the performance modelling of computer systems and even then they can only be used when the aim is to make moderate changes to a currently running system, which is not the case for the work carried out in this thesis.

3.3.1.2 Discrete Event Simulation

In discrete event simulation, the process of a system is represented as a chronological sequence of events. Each event occurs at an instant in time and marks a change of state in the system. A discrete event simulation model has an internal source of random numbers. The random numbers derive the components of the simulation model; they are used to determine the occurrence time of between system events, branching probabilities and so on. For example, random number generators are used to generate network topologies, traffic profiles and mobility patterns. The essential feature is that the model is self-contained; it requires no external inputs to operate. All discrete-event simulations have a common structure. The most basic components in a discrete-event simulation are [64][71][73]: simulation clock, events, event handler, event list and time-advancing routine.

One of the difficulties associated with discrete event simulation derive from its inherent stochastic nature. For example, in a discrete event simulation model inter-arrival times are normally generated from independent, identically distributed, random variables. However, if significant interactive effects between the underlying processes are present, then detailed statistical analysis must be performed in order to discover their characteristics [72]. This information must then be used to appropriately adjust the random variables. Also, verification and validation of discrete event simulation models can be very difficult. Issues such as selecting appropriate levels of detail, determining the modelling assumptions and fixing model parameters will all complicate the task [74][75].

Despite the problems of performing an accurate simulation, the discrete event simulation method of system analysis is very attractive for the following reasons [72]:

- Cost-effectiveness: The ability to perform for relatively low costs (compared to other methods).
- Accessibility: The ability to be performed by non-mathematicians.
- Power: The ability to apply the difficult problems.
- Scope: The ability to apply to diverse problems.

The benefits of discrete event simulation outweigh its limitations and, for complex systems, provides the only feasible means for evaluating protocol performance and operation and hence is considered for carrying out this work. As a notational convenience from now on the terms 'model' and 'simulation' should be understood to refer to discrete event model and to discrete event simulation respectively.

3.3.2 Simulation Approach

Simulation packages are widely used to develop a simulation model. Such packages are particularly valuable in having the potential to reduce time to develop the model as a result of enhanced development, debugging and execution environment. Many packages also have excellent visualisation feature. There are several well-known packages that can be used for the simulation of MANETs such as Network Simulator version 2 (NS-2) [76][77][78], GloMoSim [79], QualNet [80], OPNET [81] and OMNeT++ [82]. Full comparisons between such packages are presented by G.A. DiCaro [83], M. Halvardsson et al [84] and L. Begg et al [85]. This work adopts the NS-2 as simulation platform for several reasons:

- It is widely utilised in the MANETs domain.
- It is an open source.
- It has no license cost.
- It provides both good manuals and tutorials.
- The source code can be compiled on different platforms, e.g. Unix and Windows.
- Many wireless extensions have been contributed from the UCB Daedelus, the CMU Monarch projects and Sun Microsystems [83].
- New modifications can be easily included.

A new version of Network Simulator version 3 (NS-3) [86] was released in June 2008. NS-3 is intended as an eventual replacement for the popular NS-2 simulator [86]. NS-3 is not an extension of NS-2; it is a new simulator. The two simulators are both written in C++ but NS-3 is a new simulator that does not support the NS-2 APIs. Given the work started in 2005, NS-2 was the only option available but perhaps future work could use NS-3 as a simulation platform. A short overview of NS-2 is given in the next section.

3.4 NS-2 Overview

Network Simulator 2 (NS-2) developed at UC Berkeley, is an object-oriented, discrete event driven network simulator targeted at networking research. Simulation of wired as well as wireless network functions and protocols (e.g. routing schemes, TCP, UDP) can be done using NS2. In general, NS2 provides users with a way of specifying such network protocols and simulating their corresponding behaviours. NS-2 has been significantly improved by the open source community and its current release includes wireless network support provided by the CMU Monarch extensions [76].

The simulator is written in C++ and MIT's object extension to Tool command language (OTcl) [87]. C++ is a powerful programming language, which enables fast execution of applications. However, some modifications may be requested in order to perform several simulations, that is, retaining the main structure of the simulation but modifying some parameters, with the purpose of comparing different results. This implies additional time; recompiling C++ code every time a modification is requested. OTcl is an interpreted language, and offers the key advantage that such modifications do not need additional time recompiling but on the other hand, the execution time for an interpreted language is slower than compiled languages. NS-2 makes this unification feasible through tclcl, i.e. OTcl linkage. The main objects of a simulation such as nodes and protocols are implemented using C++ and the configuration of the parameters such as number or position of nodes, time of the simulation, etc. are implemented in OTcl. Hence the combination of the two languages offers a compromise between performance and ease of use. Implementation and simulation under NS-2 consists of 4 steps:

- Implement the protocol by adding a combination of C++ and OTcl code to NS-2's source base.
- Describe the simulation in an OTcl script.
- Start the NS-2 simulation engine.
- Analyse the generated trace files that can be used to do data processing (goodput, end-to-end delay) and to visualise the simulation with a program called Network Animator (NAM).

Implementing a new protocol requires adding C++ code to represent the protocol's functionality, as well as updating key NS-2 OTcl configuration files in order for NS2 to recognise the new protocol and its default parameters. The C++ code also describes which parameters and methods are to be made available for OTcl scripting. The NS-2 simulation process adopted for this work is summarised in Figure 3.4.

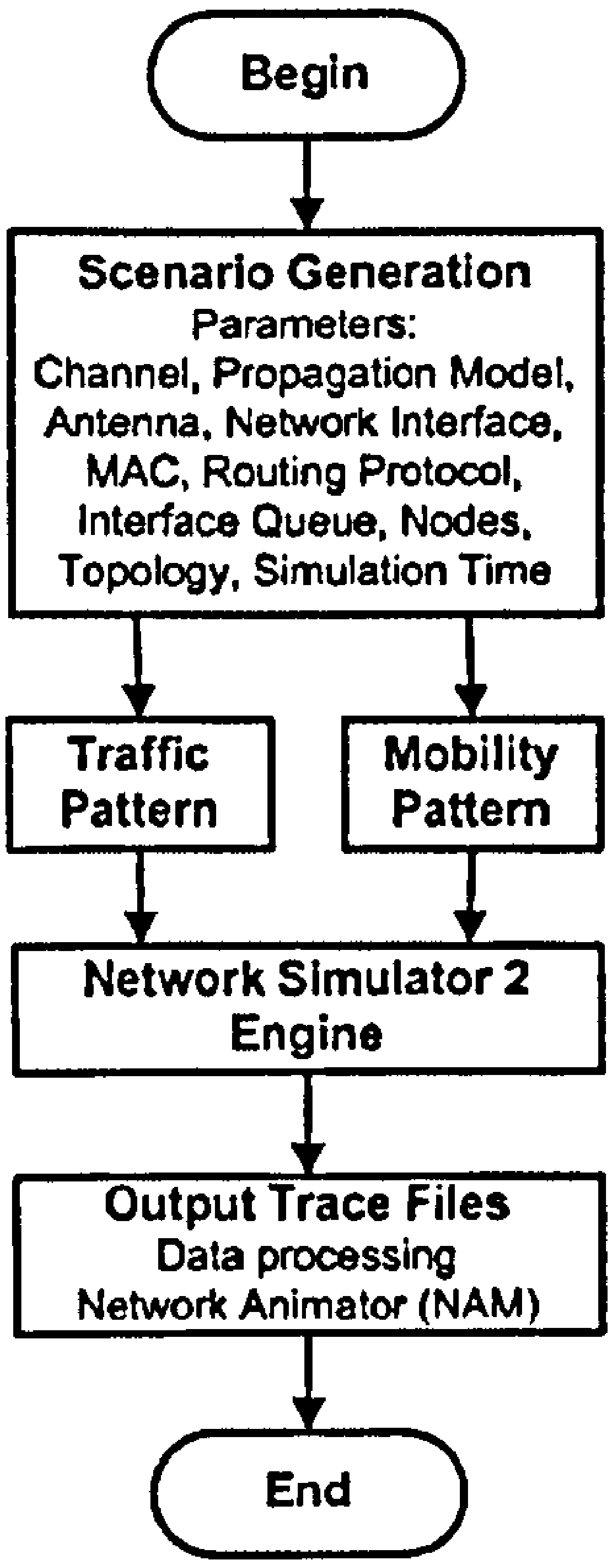


Figure 3.4 NS-2 simulation proccss

3.5 Simulation Implementation

The main aspects studied in this thesis are multipath routing and quality of service (QoS) support in MANETs. The techniques developed in the thesis are based on DSR routing protocol presented in section 2.7.3. The reasons for choosing DSR among the other routing protocols are as follows:

- In a range of scenarios, source routing is shown to outperform table-driven approaches [50].
- DSR is widely considered as a benchmark source-based scheme [50][58][59][60] and standard, validated implementations of this protocol are available with the NS-2.
- When the flow-state is disabled, the end-to-end path is appended to the packet's header. This feature permits more flexible monitoring and more flexible tracing of the path, when the packet is propagated from its source to destination [64].

- Intermediate nodes need not maintain up-to-date routing information in order to route the packets that they forward; the packets themselves already contain all the routing decisions.
- It is convenient to build multiple paths using source routing, since the destination node knows the complete path of all available routes.

This section provides a detailed descriptions of the network model, mobile node model, DSR routing process model state machine and mobility process model implemented in NS-2.

3.5.1 Network Model

Figure 3.5 shows a screenshot of NAM depicting the network model. The scenario contains 50 mobile nodes; a value common in the literature [50][60] and used here for comparison. The nodes can move around the specified area and communicate over wireless links with a transmission range of 250 m.

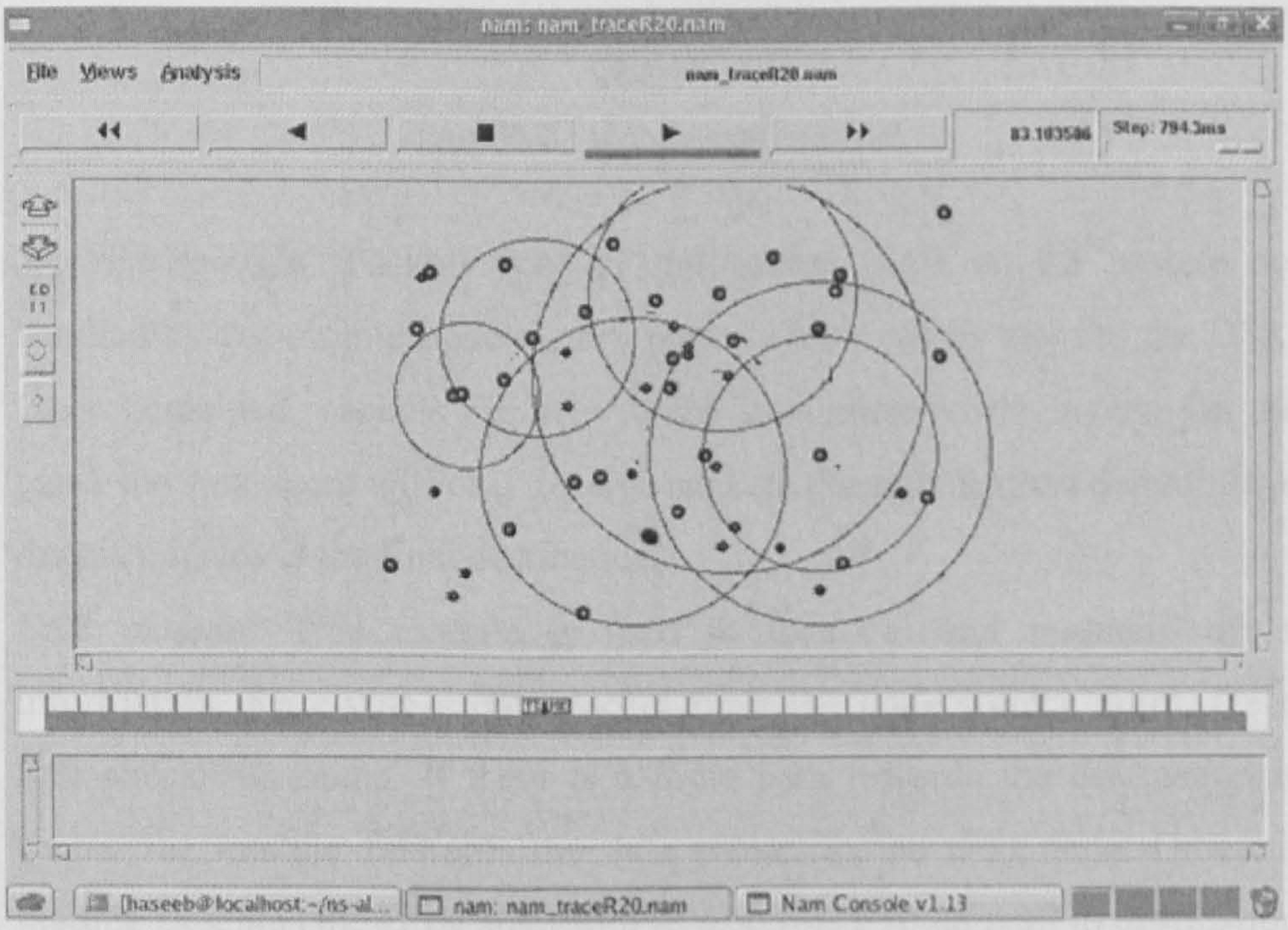


Figure 3.5 Network model

3.5.2 Mobile Node Model

Figure 3.6 shows the DSR mobile node model [76], which simulates the protocol stack. Below are the details of modules that make up the DSR mobile node model:

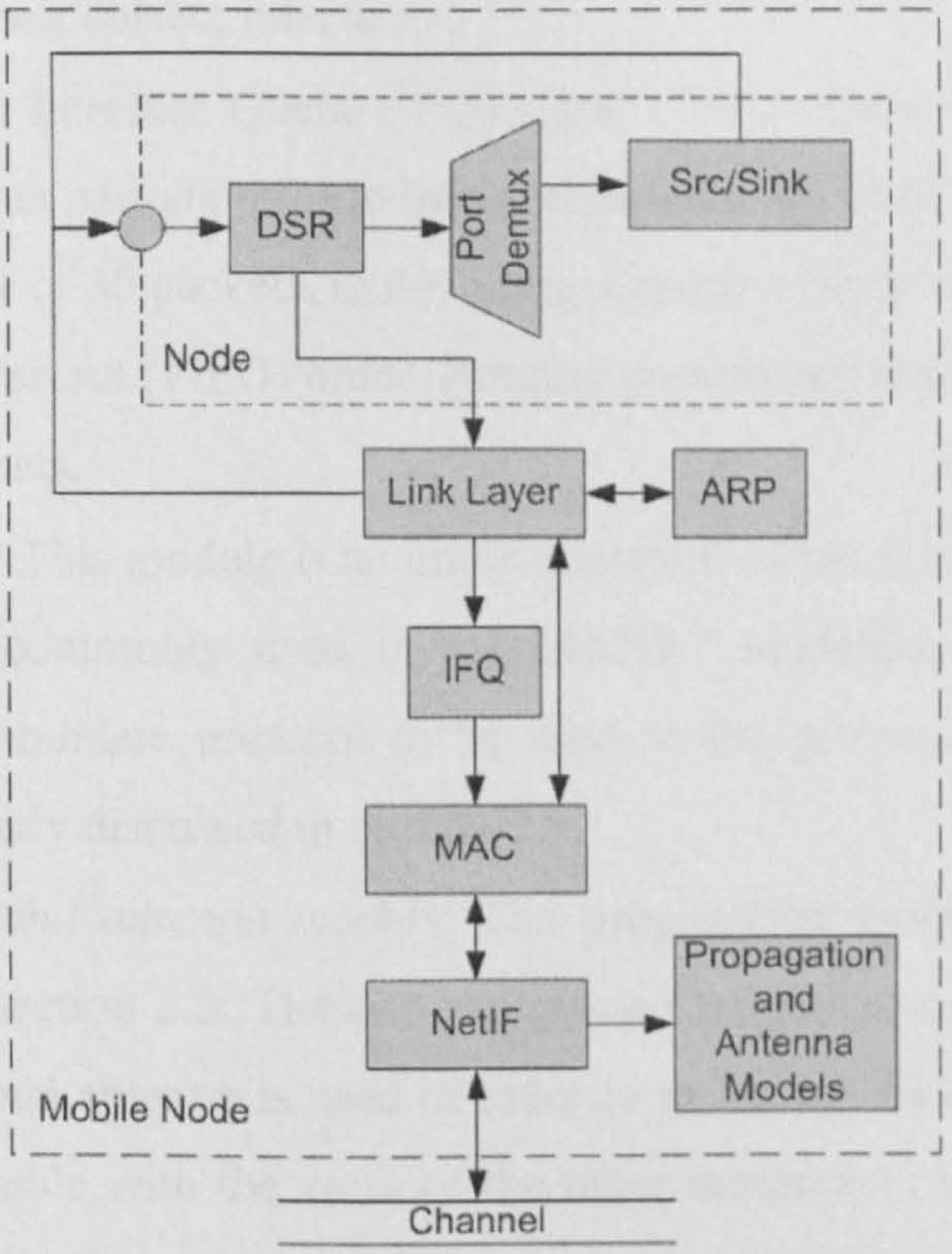


Figure 3.6 Mobile node model

- *Src/Sink module:* Packets sent by the source (Src) on the mobile node are handled by the mobile node's entry point, which passes them to the DSR agent. Once generated, packets are sent to the immediate lower layers. On the other hand, the sink agent will only receive packets through the port demuliplexer (port demux), if this is the final destination.
- *DSR module:* This module is used to discover and maintain the routing information. Receiving a data packet from the node's entry point, the module first checks its cache. If there is a route path towards the destination node in cache, the module forwards the data packet to the next node. Otherwise, the module executes the DSR algorithm to discover a route path to destination node. The DSR process model is discussed in the next sub-section.

- *LL class and ARP module:* The link-layer (LL) class is responsible for simulating the data link protocols. The LL class uses an Address Resolution Protocol (ARP) to determine the hardware addresses of the neighbouring nodes and map IP addresses to their correct interfaces.
- *IFQ class:* An Interface Queue (IFQ) class 'CMUPriQueue' is used to queue all routing and data packets prior to being transmitted by MAC layer. The IFQ has a maximum size of 50 packets, maintaining a queue with two priorities each served in a first-in first-out (FIFO) order. Routing packets are assigned a higher priority than data packets.
- *MAC module:* This module is an implementation of the IEEE 802.11 standard [2] the protocol commonly used in the MANET modelling community, and an appropriate candidate protocol to be used in the proposed applications. IEEE 802.11 is already discussed in section 2.5.
- *Propagation and antenna models:* The propagation models used in NS-2 are described in section 2.3. The two-ray ground reflection model and a unity gain omni-directional antenna is used in order to make the results obtained from this study comparable with the work of the other researchers who have adopted the same approach. The antennas of the transmitter and the receiver are placed at a height of 1.5 meters above the ground. The antenna parameters for type *Antenna/OmniAntenna* are summarised in Table 3.2.

Table 3.2 Antenna parameters

Parameter	**Tcl class variable**	**Value**
Position X (meters)	X_	0
Position Y (meters)	Y_	0
Position Z (meters)	Z_	1.5
Transmitting Antenna Gain	Gt_	1.0
Receiving Antenna Gain	Gr_	1.0

- *NetIF module:* The wireless network interface (NetIF) uses the characteristics of the 914MHz Lucent WaveLAN Direct Sequence Spread Spectrum (DSSS) radio [88]. WaveLAN is modelled as a shared-media radio with minimum range of 250m and nominal bit rate of 2Mbps. The configuration parameters that make the wireless interface *Phy/WirelessPhy* model the Lucent WaveLAN radio interface is given in Table 3.3.

Table 3.3 Wireless interface parameters

Parameter	**Tcl class variable**	**Value**
Raw bit rate (bps)	Rb_	2*1e6
Power of transmission (W)	Pt_	0.2818
Frequency (Hz)	freq_	914e+6
System loss factor	L_	1.0
Carrier sense threshold (W): min power required to detect another node's transmission	CSThresh_	1.559e-11
Receive threshold (W): min power required to receive a packet	RXThresh_	3.652e-10
Capture threshold (dBm)[11]: signal ratio required to maintain receiver capture of incoming packet in face of collision	CPThresh_	10.0

3.5.3 DSR Process Model State Machine

The DSR process (Figure 3.7) implements the DSR algorithm. The role of each state in this finite state machine is described below [89]:

- *Pre-init State:* This state pre-initializes the DSR process model by managing the DSR address of the current node and by checking that this address is valid within the network.
- *Init State:* This state initializes every variable, statistic, cache and user parameter that is used by the DSR process model. Once the initialisation setup is accomplished, the process transits to the idle state.

---

[11] The reader may observe inconsistencies in the units used. These are the previously reported values and for consistency with other work the same values are listed here.

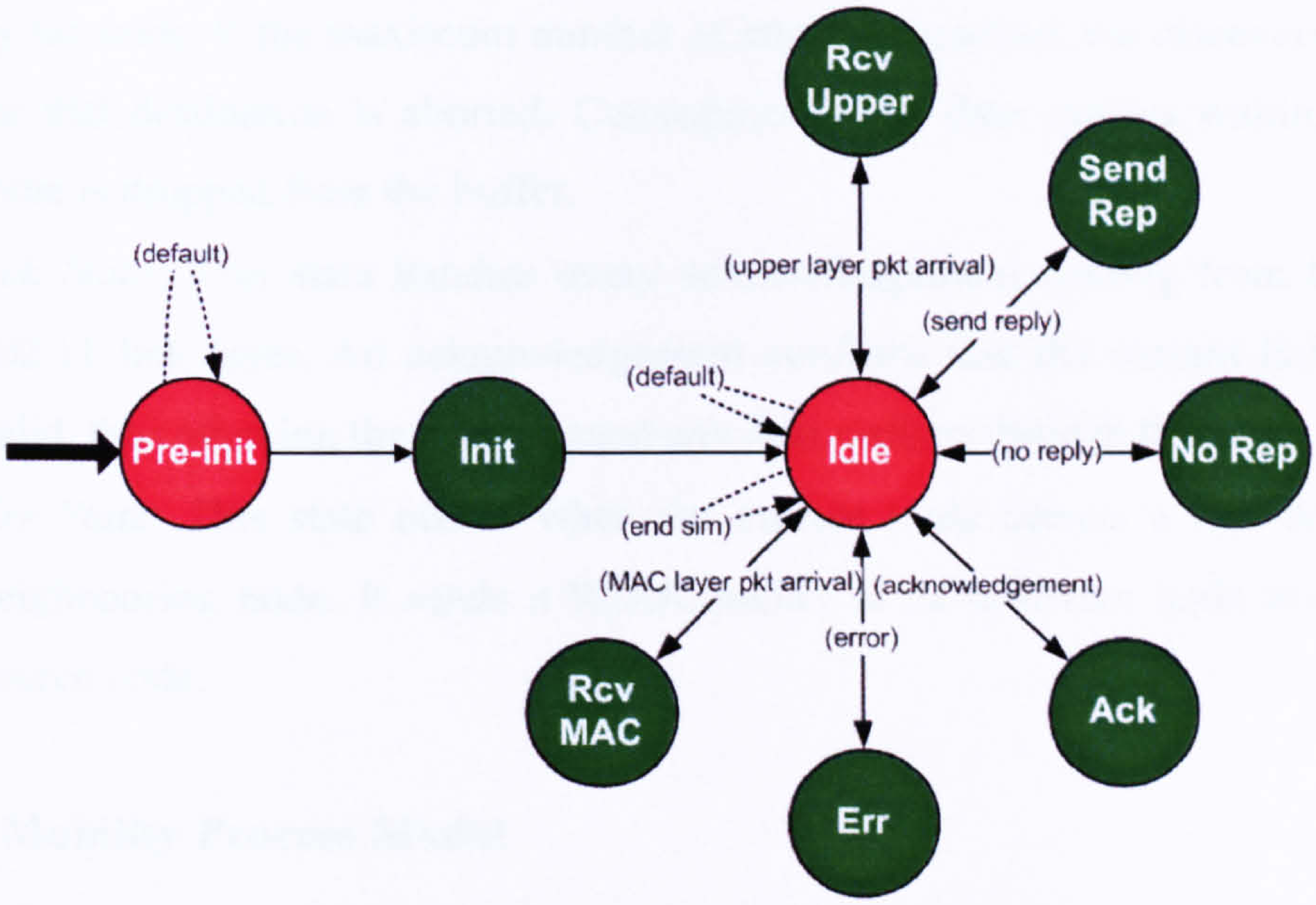


Figure 3.7 DSR process model state machine

- *Idle State:* This is the default state where the process waits for an event.
- *Rcv Upper State:* This state handles every packet generated by the upper layer that the DSR protocol must carry to a given destination. The current state first extracts the destination address from the received packet, then checks the cache. If a route path exists in its cache, the current state forwards it to the next hop node towards destination. Otherwise, the current state saves the packets to a waiting queue and initiates a route discovery process by sending an RREQ packet.
- *Rcv MAC State:* This state handles every packet received from the IEEE 802.11 link layer. The packet is then processed according to its type: RREQ, RREP, data or RERR.
- *Send Rep:* This state occurs when the current (destination) node sends a scheduled RREP packet to the source node using the reverse path identified within the RREQ packet.
- *No Rep:* This state is called when the RREP wait timer expires. That means that the source node still did not receive a RREP to its RREQ. Thus, the previous step of the route discovery has failed and that a new request packet must be generated

by the node. If the maximum number of retries is reached, the discovery process for that destination is aborted. Consequently, any data packets waiting for the route is dropped from the buffer.

- *Ack State:* This state handles every acknowledgement coming from the IEEE 802.11 link layer. An acknowledgement confirms that the current link used is valid, thus allowing the node to send any data packets through this link.
- *Err State:* This state occurs when the current node detects a link break to a neighbouring node. It sends a RERR packet to its upstream node towards the source node.

3.5.4 Mobility Process Model

The mobility process model (Figure 3.8) implements a random waypoint mobility model [50]. The random waypoint mobility model is widely used model when evaluating MANETs [35][39][48][52][58][59][60][61][62][64] and hence is considered in this thesis. Each state of the model is described below.

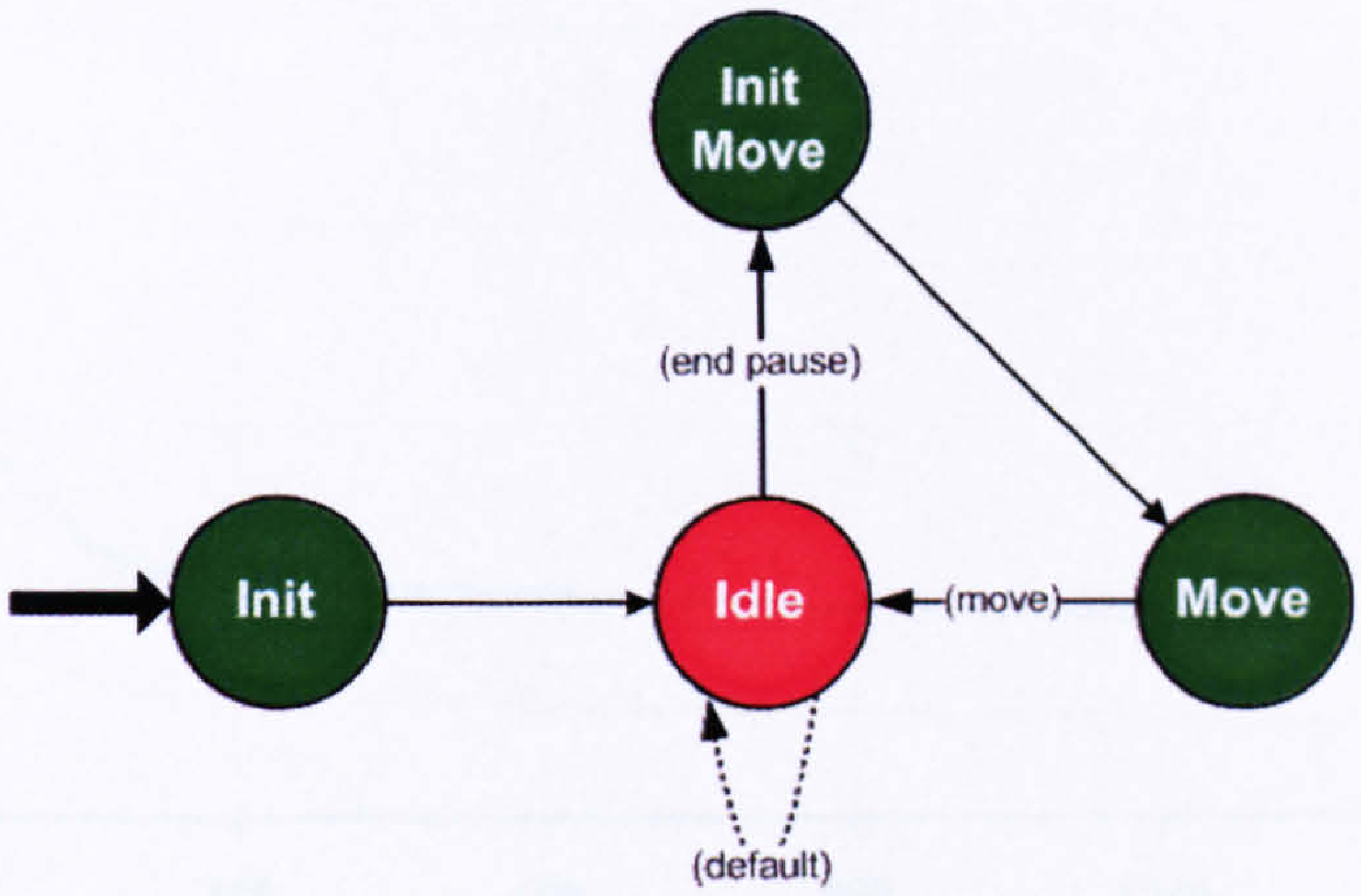


Figure 3.8 Mobility process model

- *Init State:* In the initialisation state, each node picks a random location inside the simulation area.
- *Idle State:* This is the default state where the node remains stationary for 'pause time' seconds.

- *Init Move State:* In this state, the node selects a random destination in the simulation area and a random speed distributed uniformly between 0 and maximum speed.
- *Move State:* In this state, the node travels toward the chosen destination at the selected speed. Upon arrival, the node returns to the idle state and pauses for a specified time period before restarting the process.

In the random waypoint mobility model, the average node speed is often believed to be half of the maximum speed because node speeds are chosen from a uniform distribution [0, maximum speed]. An example generated by the random waypoint mobility model with maximum speed of 10 and 20 m/s and zero pause time is plotted in Figure 3.9. This is the average over 30 different scenarios and the average speed is calculated every 5 seconds. The simulation duration is 1500 seconds.

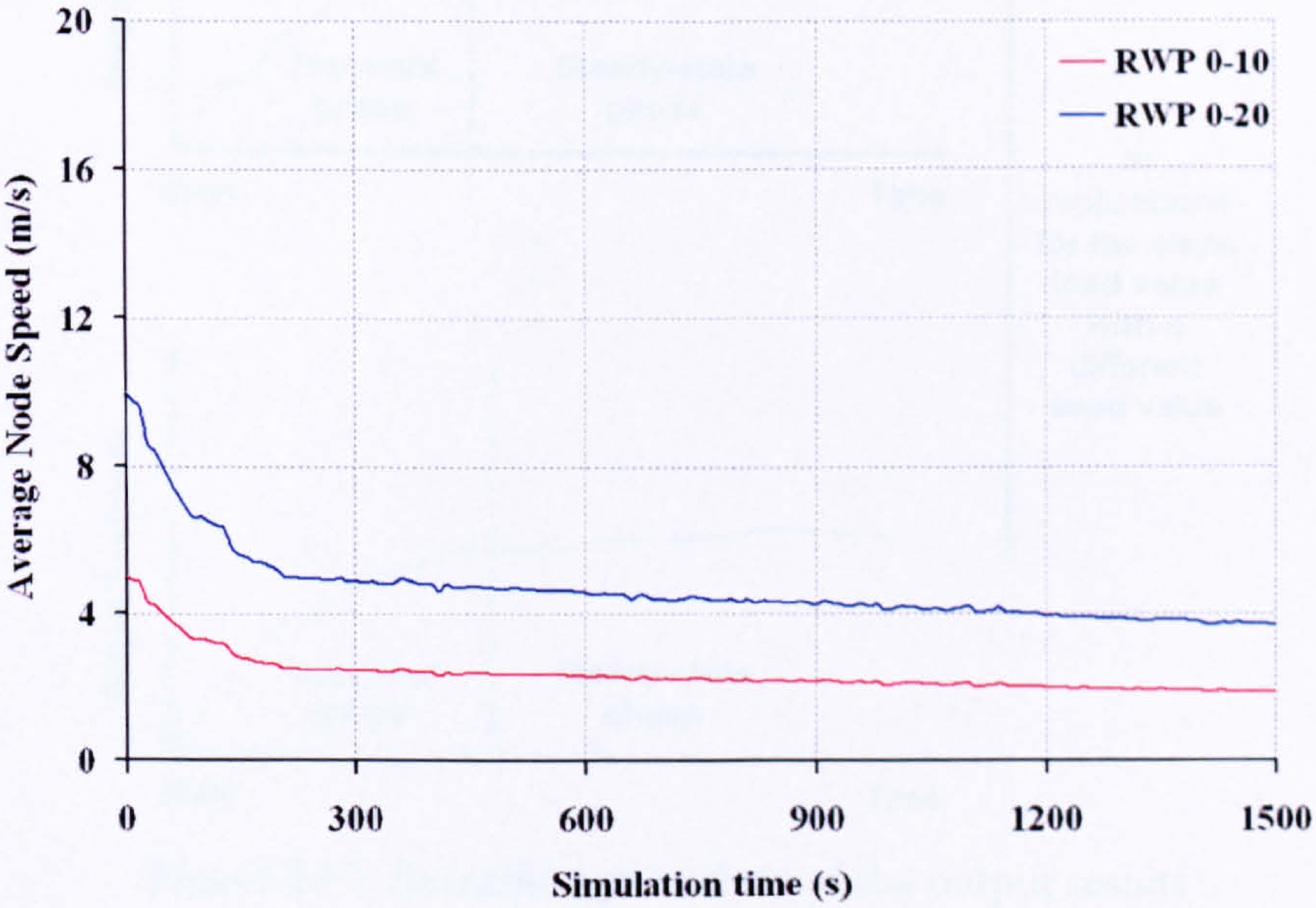


Figure 3.9 Average node speed versus time

It can be observed in Figure 3.9 that the average node speed is consistently decreasing. This is because some nodes select very low speeds as they proceed towards their destination. These nodes become 'trapped' because of the low speed and it can take long times to reach their destination [64][90]. As the simulation progresses, an increasing number of nodes will be 'trapped' to slower trips resulting in the speed decay. Running the simulation for longer does not really improve the position as the average speed continues to reduce albeit at a reduced rate. The lack of stability[12] can represent a significant drawback of the standard random waypoint mobility model. Therefore, care must be taken to ensure that the standard random waypoint mobility model has not decay significantly.

3.6 Simulation Results Analysis

This section deals with the procedure for obtaining output results in terms of accuracy and repeatability. Figure 3.10 illustrates the recording procedure of the output results.

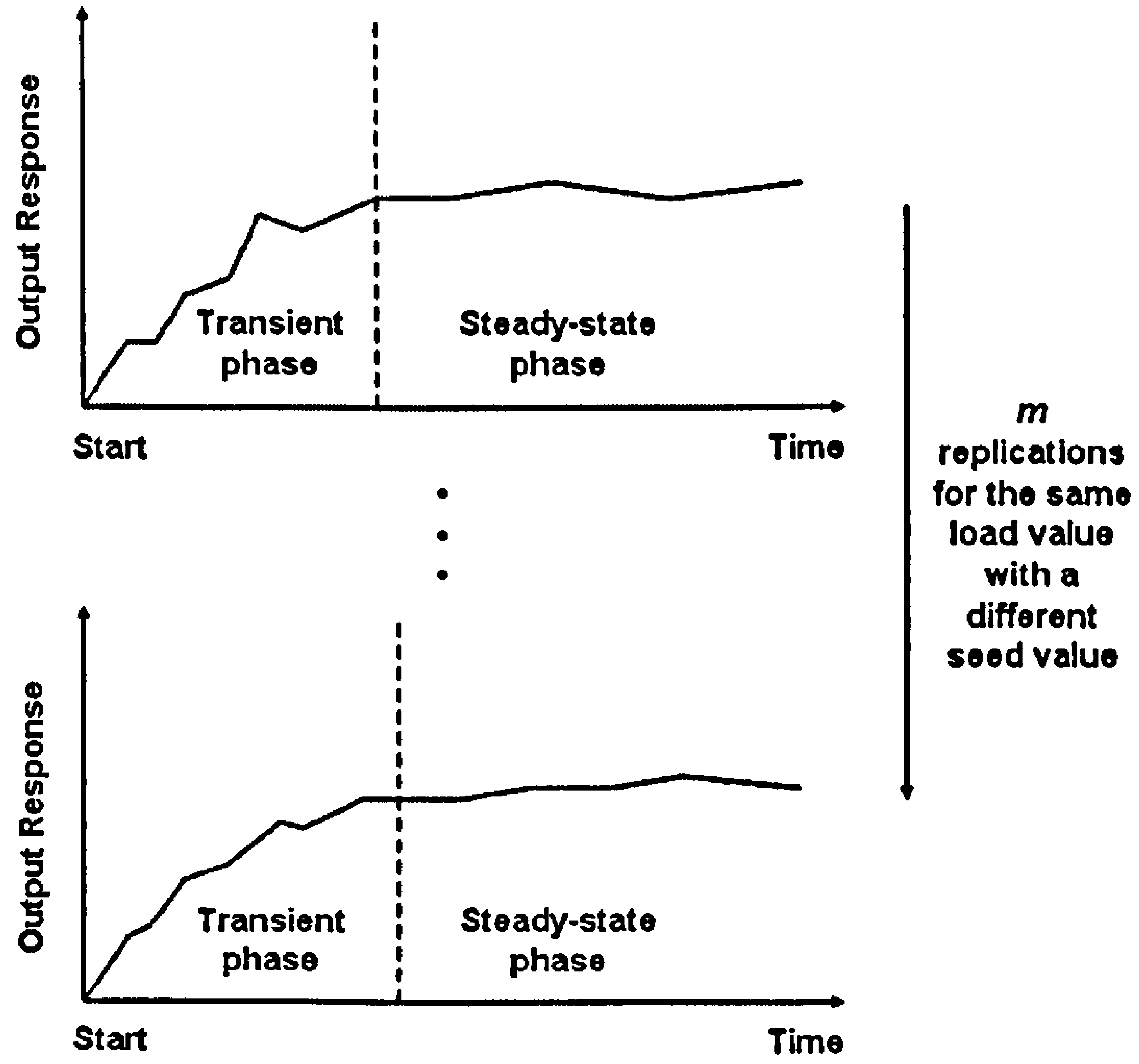


Figure 3.10 Recording procedure of the output results

The simulation involves two phases: firstly, the transient (warm-up) phase whereby the average fluctuation experiences significant variation, and secondly, the steady-state

---

[12] The stability is discussed in Section 3.6.1.

phase in which the average fluctuation tends to remain within a narrow range. The data collection process must begin at the steady-state phase and it is essential that the simulation is repeated several times for the same load value with a different random number generator seed value.

There are several issues which must be considered in order to achieve a higher level of accuracy including steady-state phase detection, simulation termination and number of replications. Although there are generally few deterministic methods to calculate these conditions, it is possible to calculate estimated values using some statistical techniques as described in the following sub-sections.

3.6.1 Steady-State Detection

In most simulations, only the steady-state performance, that is, the performance after the system has reached a stable state, is of interest. It is important from an efficiency perspective that the onset of steady-state is detected as soon as possible. Two common techniques used for detecting the steady-state phase are Batch Means and Moving Average of Independent Replications [64][71][72].

3.6.1.1 Batch Means

The method of batch means requires running very long simulations and then dividing it up into several parts of equal duration. Each part is called a batch or sub-sample. The mean of observations in each batch is called the batch mean. The method requires analysing the variance of these batch means as a function of the batch size.

As shown in Figure 3.11, a long run of N observations can be divided into m batches of size n each, where $m = N/n$. Let x_{ij} denote the jth observation in the ith batch, $\bar{x}_i$ is the batch mean and $\bar{\bar{x}}$ is the overall mean. The steps for steady-state detection using batch means are illustrated in Table 3.5 [71].

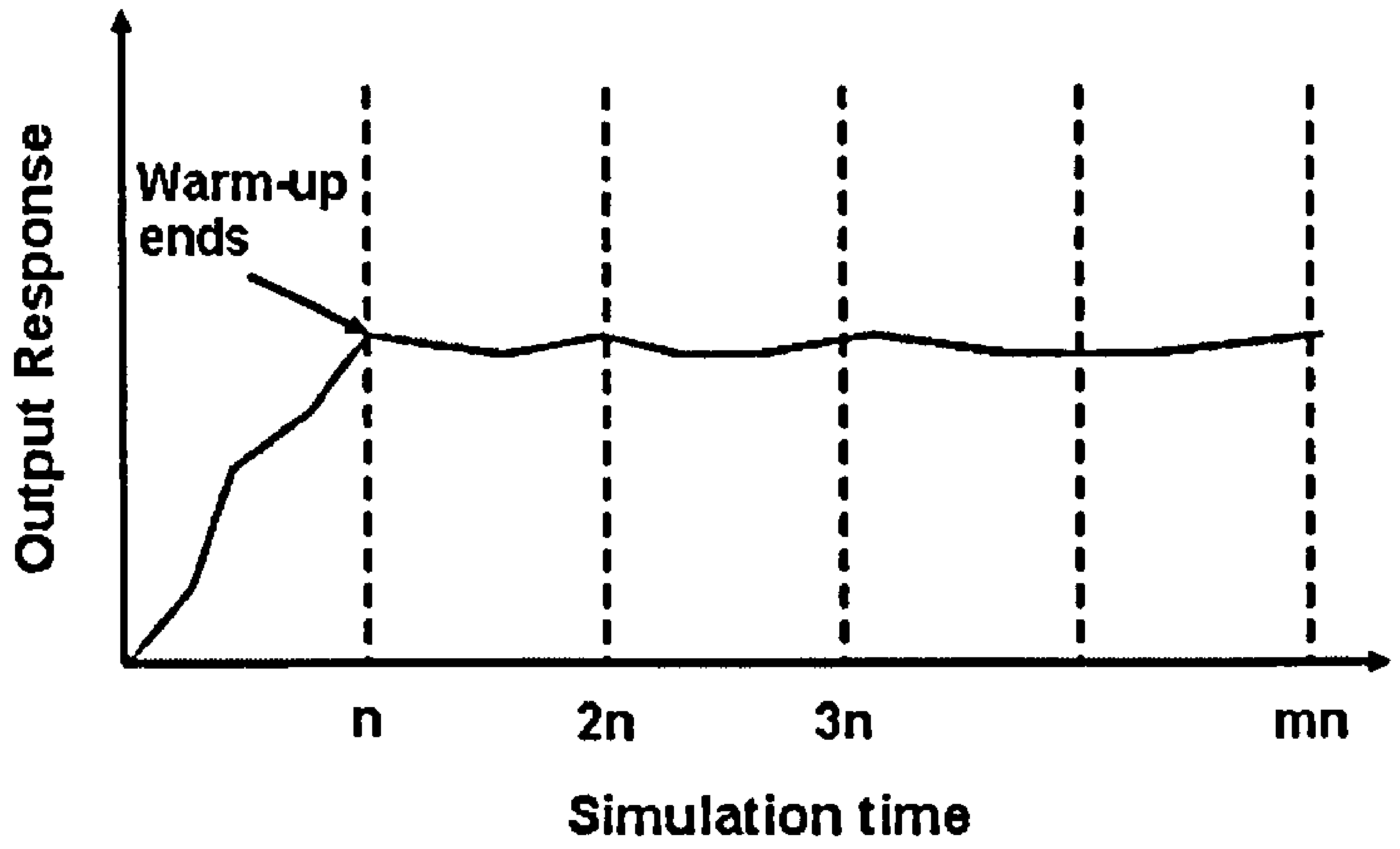


Figure 3.11 Batch means technique

Table 3.4 Steady-state detection using batch means [71]

Step	Description	Equation	
1.	For each batch, compute a batch mean.	$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$	for $i = 1,2,\ldots,m$
2.	Compute the overall mean.	$\bar{\bar{x}} = \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i$	
3.	Compute the variance of the batch means.	$Var(\bar{x}) = \frac{1}{m-1}\sum_{i=1}^{m} (\bar{x}_i - \bar{\bar{x}})^2$	
4.	Repeat steps 1 to 3 for increasing batch size *n*.		
5.	Plot the variance as a function of batch size *n*.		
6.	The transient interval is the value of *n* for which the variance starts decreasing[13].		

An advantage of batch mean technique is that only one transient (warm-up) interval needs to be removed during the process of recording observations. On the other hand, the limitation of this approach is that the batches, as illustrated in Figure 3.11, may not

---

[13] Suppose the length of the transient period is *T*. If the batch size *n* is much less than *T*, initial batches bring the overall mean toward the initial batch means and the variance is small. As the batch size is increased, the variance increases. At *n* larger than *T*, only the first batch mean is different; other batch means are approximately equal. This results in the decrease of the variance [71].

really be statistically independent. This would be especially true if the batch sizes are not large enough and will lead to significant correlation between successive batches. Since the confidence estimation measures discussed in section 3.6.2 are based on the independence of the individual simulation runs, this lack of independence between successive batches, if sufficiently serious, may have a serious impact on the accuracy of the confidence estimation procedures.

3.6.1.2 Moving Average of Independent Replications

In moving average of independent replications technique, the average is computed over a moving window interval. Multiple independent replications are performed which are subsequently averaged. Each of the independent replications is performed with common load values but different seed values. This results in a smoother trajectory mean. Given m replications of size n each, let x_{ij} denote the jth observation in the ith replication, $\bar{x}_j$ is the trajectory mean across m replications and $\bar{\bar{x}}_j$ is the mean of observations within the window of length k. The steps for steady-state detection using moving average of independent replications are illustrated in Table 3.5 [71].

Table 3.5 Steady-state detection using moving average of independent replications [71]

Step	Description	Equation	
1.	Compute the trajectory mean by averaging across replications.	$\bar{x}_j = \frac{1}{m}\sum_{i=1}^{m} x_{ij}$	for $j = 1,2,...,n$
2.	Set $k = 1$ and plot a trajectory of the moving average of successive $2k+1$ values.	$\bar{\bar{x}}_j = \frac{1}{2k+1}\sum_{l=-k}^{k} \bar{x}_{j+l}$	for $j = k+1, k+2,...,n-k$
3.	Repeat step 2, with $k = 2,3,...$ until the plot is sufficiently smooth.		
4.	Find the knee of the plot (as shown in Figure 3.12), the value j at the knee gives the length of the transient phase.		

The key advantage of independent replications technique is that it ensures that the samples are independent. On the other hand, the limitation of this approach is that running through the warm-up phase for each replication extends the length of time to perform the replications. Furthermore, there is a possibility that the length of the warm-up period is under estimated, resulting in bias for each instantiation. It is highly desirable in the independent replications technique that the warm-up period is properly detected and removed.

As a case study, the average node speed is examined in order to determine the warm-up time of the mobility scenario. Section 3.5.4 showed the average node speed for the random waypoint mobility model. When the moving average technique is applied to the average node speed, the graph of Figure 3.12 is obtained. The knee of the plot occurs at approximately 200 seconds. In order to ensure that the warm-up period is completely removed, the first 300 seconds of the simulation time are ignored. This is also one of the solutions suggested previously [64][90] and hence used in this work.

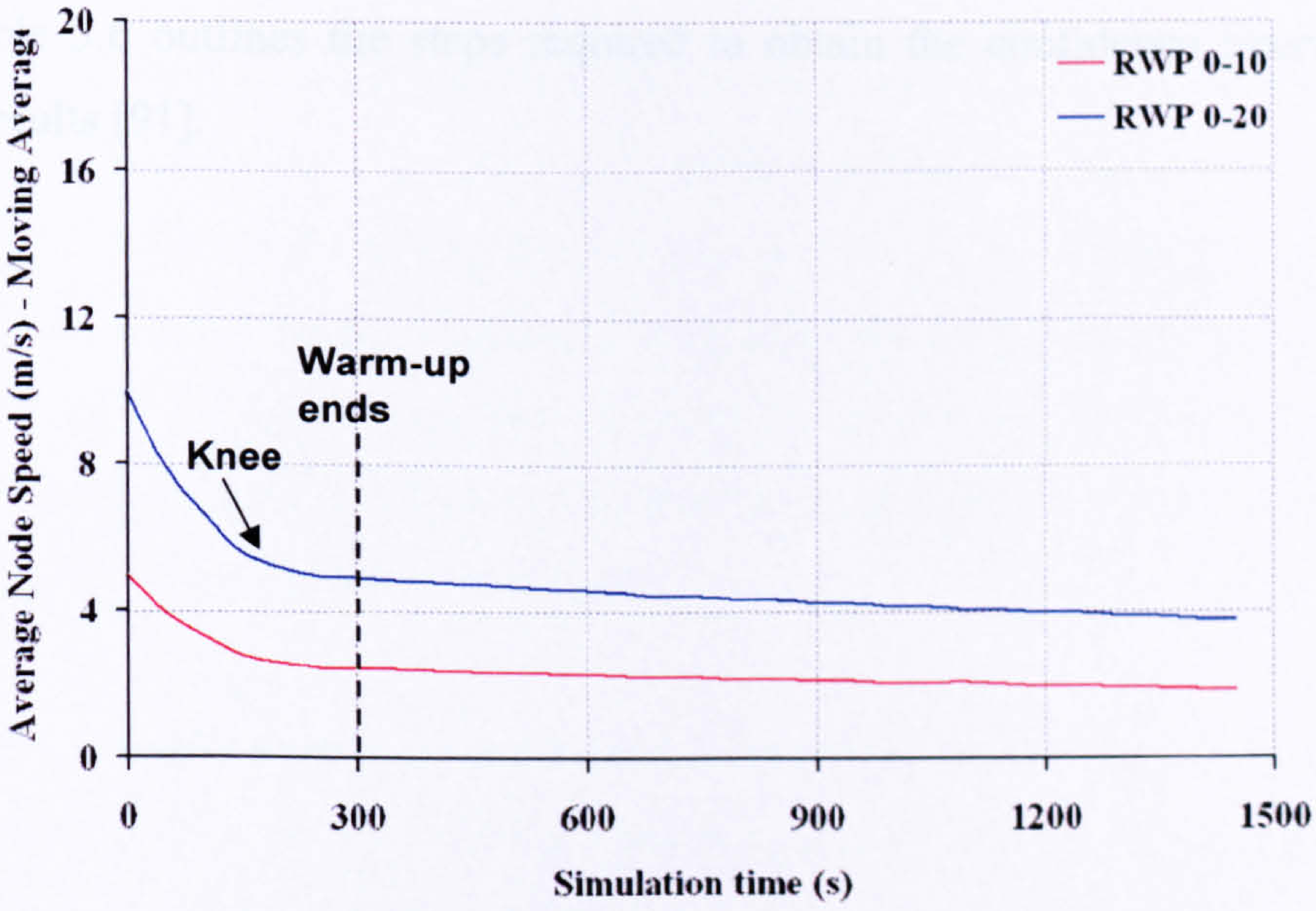


Figure 3.12 Moving average applied to the average node speed

3.6.2 Simulation Time and Confidence Intervals

The simulation time should be properly chosen to guarantee a sufficiently accurate estimate of the final results. If the simulation is too short, the results may be highly variable. On the other hand, if the simulation is too long, computing resources may be unnecessarily wasted. Better to cut towards the latter in that results are free of bias and are stable albeit having taken longer to obtain. The simulation should be repeated with a different random number generator seed value. Twenty replications (m=20) are considered in this work of size $n+n_t$ each, where n_t is the length of transient phase. The first n_t observations of each replication are discarded. An overall mean is calculated across the replications. The variance of the replicate means is then calculated and this provides the confidence interval for the mean response. Finally the variance of the replicate means is utilised to find the upper and lower confidence intervals denoted with vertical bars across the mean values. This is achieved by using ·the Student's t-distribution table since the number of replications is less than thirty [71][91]. If the number of replications was greater than thirty, a normal distribution could have been used. A confidence interval of 95% (confidence level α of 0.05) is used throughout this thesis. Table 3.6 outlines the steps required to obtain the confidence intervals for the sampled results [91].

Table 3.6 Independent replications [91]

Step	Description	Equation
1.	Compute the mean for each replication.	$\bar{x}_i = \frac{1}{n}\sum_{j=n_i+1}^{n_i+n} xij$ for $i = 1,2,\ldots,m$
2.	Compute the overall mean for all replication.	$\bar{\bar{x}} = \frac{1}{m}\sum_{i=1}^{m}\bar{x}_i$
3.	Calculate the variance of the replicate means.	$Var(\bar{x}) = \frac{1}{m-1}\sum_{i=1}^{m}(\bar{x}_i - \bar{\bar{x}})^2$
4.	An approximate 100(1-*a*)% confidence interval for the mean response is:	$C = \bar{\bar{x}} \mp t\sqrt{\frac{Var(\bar{x})}{m}}$ (where $t \equiv t_{\frac{\alpha}{2},m-1}$ is chosen such that $P\{T_{m-1} \leq t\} = \frac{\alpha}{2}$ and T_{m-1} is a Student-*t* random variable with *m*-1 degrees of freedom)

3.7 Model Validation

For simulation results to be credible the simulation models in use must undergo verification and validation. O. Balci [92] defines verification as substantiating that a model is built from a problem formulation accurately, where validation is substantiating that the model behaves with satisfactory accuracy within its domain. J.S. Carson [93] and R.G. Sargent [94] define the two terms to be similar and both note that sufficient accuracy is recited when a model can be used instead of a real system for purposes of experimentation and analysis.

This section presents results of preliminary simulation-based experiments in order to ensure that the implementation of the Dynamic Source Routing (DSR) [50] and Split Multipath Routing (SMR) [60] schemes inside the simulator (NS-2) are faithful to the scheme's specifications. Each scheme implementation is validated by comparing the simulation results with that of a known baseline. These schemes will be further used for performance comparison with the proposed multipath routing scheme presented in the next chapter.

3.7.1 DSR Model Validation

DSR model validation is carried out with the same simulation parameters as stated by J. Broch et al [50]. The simulation environment consists of 50 wireless nodes forming an ad hoc network, moving according to the random waypoint mobility model within a 1500x300 meters rectangular flat space, which can be setup using a scenario generator script *setdest* [78]. All simulations are run for 15 minutes (900 seconds) of real time. Data traffic is generated using constant bit rate (CBR) using a connection pattern generator script *cbrgen.tcl* [78]. 20 CBR nodes are chosen randomly from the full set of nodes generating 64 byte data packets at a rate of 4 packets per second.

Figure 3.13 and Figure 3.14 shows the goodput (the overall percentage of the UDP data packets originated by nodes that are successfully delivered by DSR) and routing overhead (the number of routing overhead packets generated by DSR to achieve this level of data packet delivery) as a function of pause time for maximum node movement speed of 20 m/s. Each graph represents the average of 10 random mobility and traffic scenarios for the given pause time, and the error bars represent the corresponding confidence interval of 95%, calculated as described in section 3.6. At a pause time of 0 (on the left of each graph), all nodes in the network are in constant motion; as the pause time increases from left to right, the average node movement rate in the network decreases. At a pause time of 900 (on the right of each graph), all nodes are stationary because each simulation is run for 900 simulated seconds of operation of the ad hoc network. The results are at the same level when compared to the reported results as both simulations are conducted using same simulation package, i.e. NS-2. DSR delivers almost all data packets, regardless of pause time, with the goodput rising to 100% at pause time 900 (a stationary network). Similarly, the routing overhead is very small at pause time 900, rising only slowly as pause time decreases (as the average node mobility rate in the network increases). The slight difference in routing overhead can be due to the dynamic nature of the network and underlying (and possibly undocumented) parameter settings and not the schemes being compared.

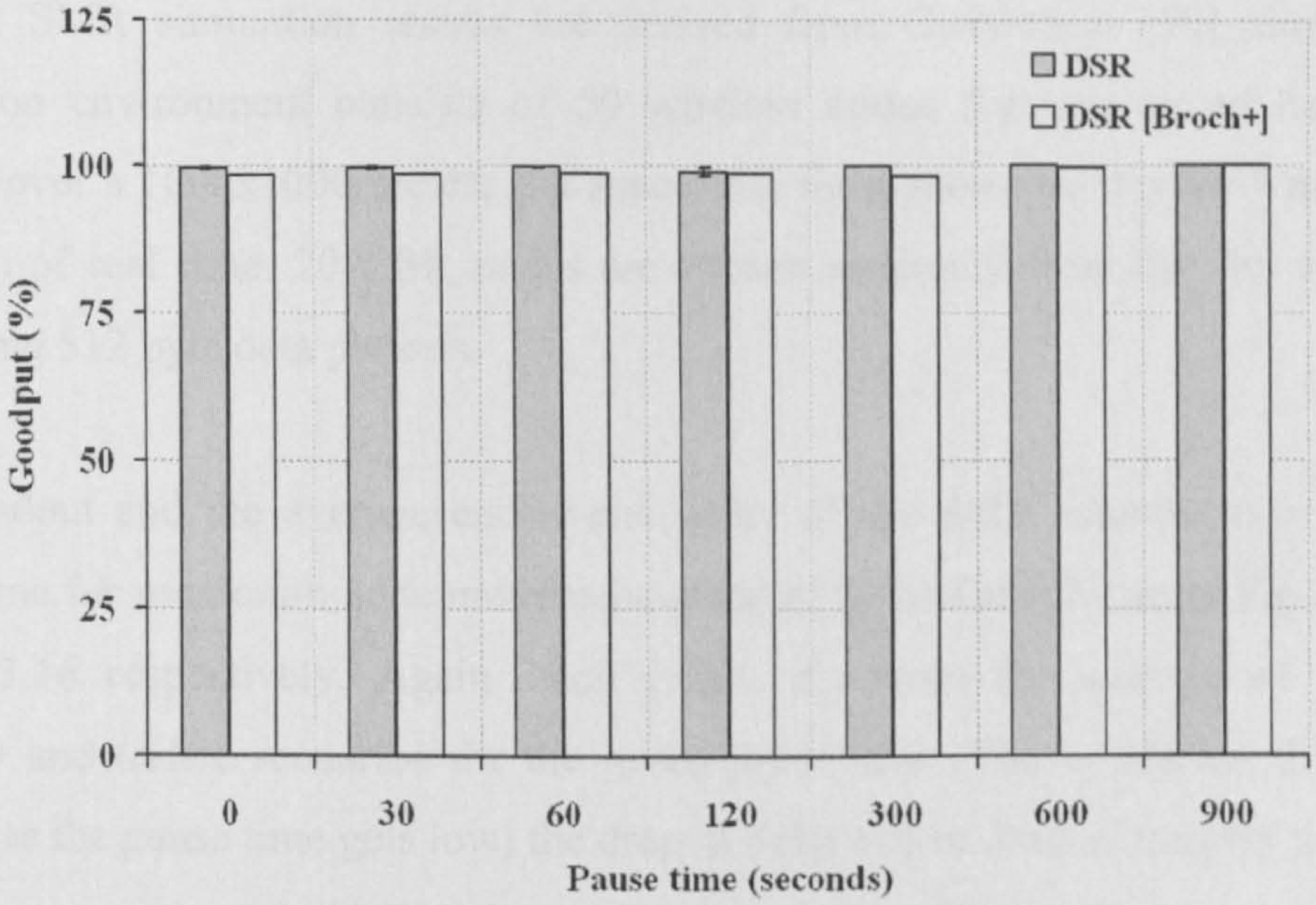


Figure 3.13 Goodput as a function of pause time

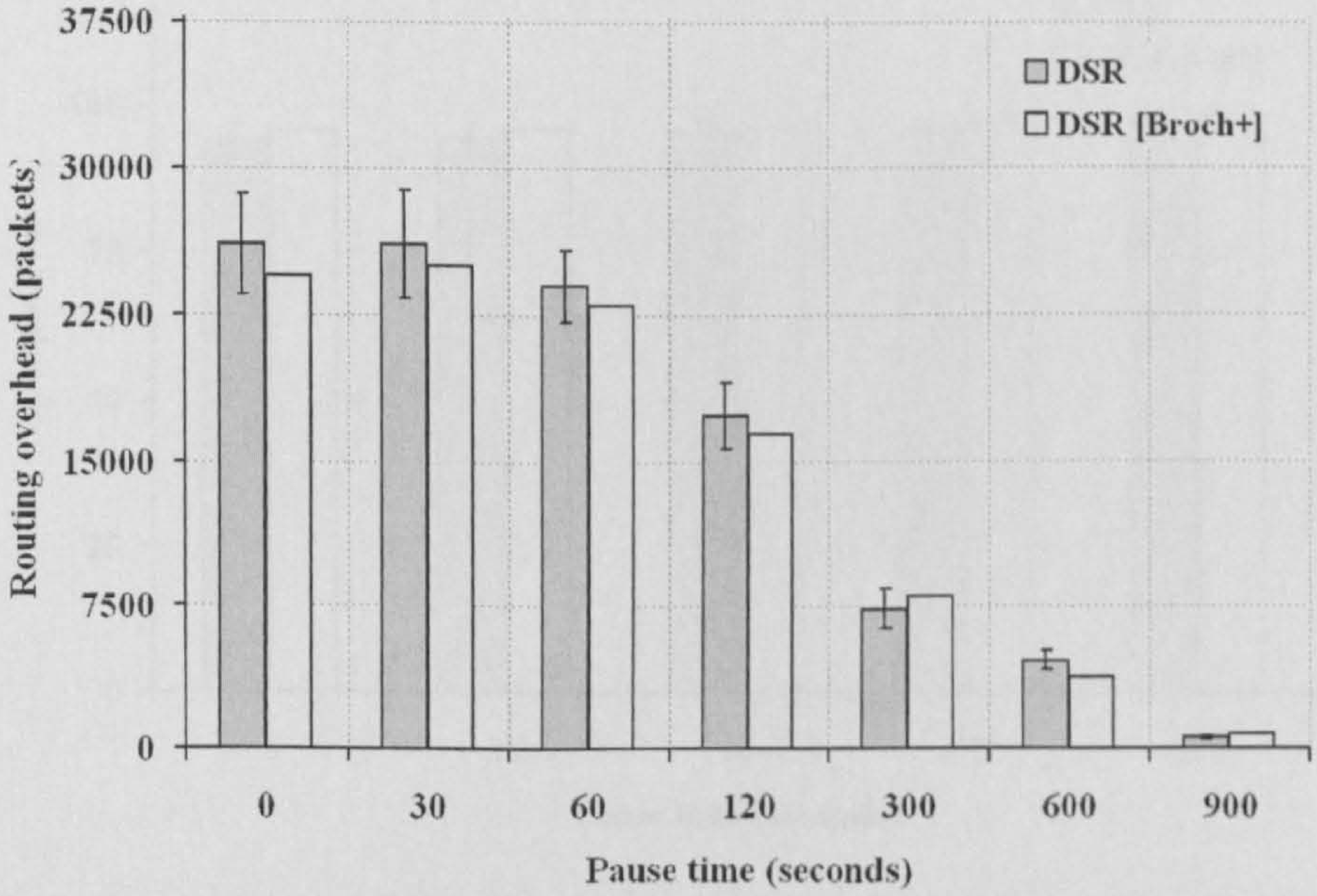


Figure 3.14 Routing overhead as a function of pause time

3.7.2 SMR Model Validation

SMR model validation is carried out with the same simulation parameters as stated in [60]. The only difference is that SMR simulations are conducted using NS-2 whereas the reported SMR simulation results are derived from GloMoSim [79] simulator. The simulation environment consists of 50 wireless nodes forming an ad hoc network, moving over a 1000x1000 meters flat space. All simulations are run for 5 minutes (300 seconds) of real time. 20 CBR nodes are chosen randomly from the full set of nodes generating 512 byte data packets.

The goodput and the average end-to-end delay of the SMR scheme as a function of pause time for maximum node movement speed of 20 m/s are shown in Figure 3.15 and Figure 3.16 respectively. Again, each graph represents the average of 10 random mobility and traffic scenarios for the given pause time. The trends are the same and indeed (as the pause time gets low) the drop in delay (up to 3ms) is roughly the same.

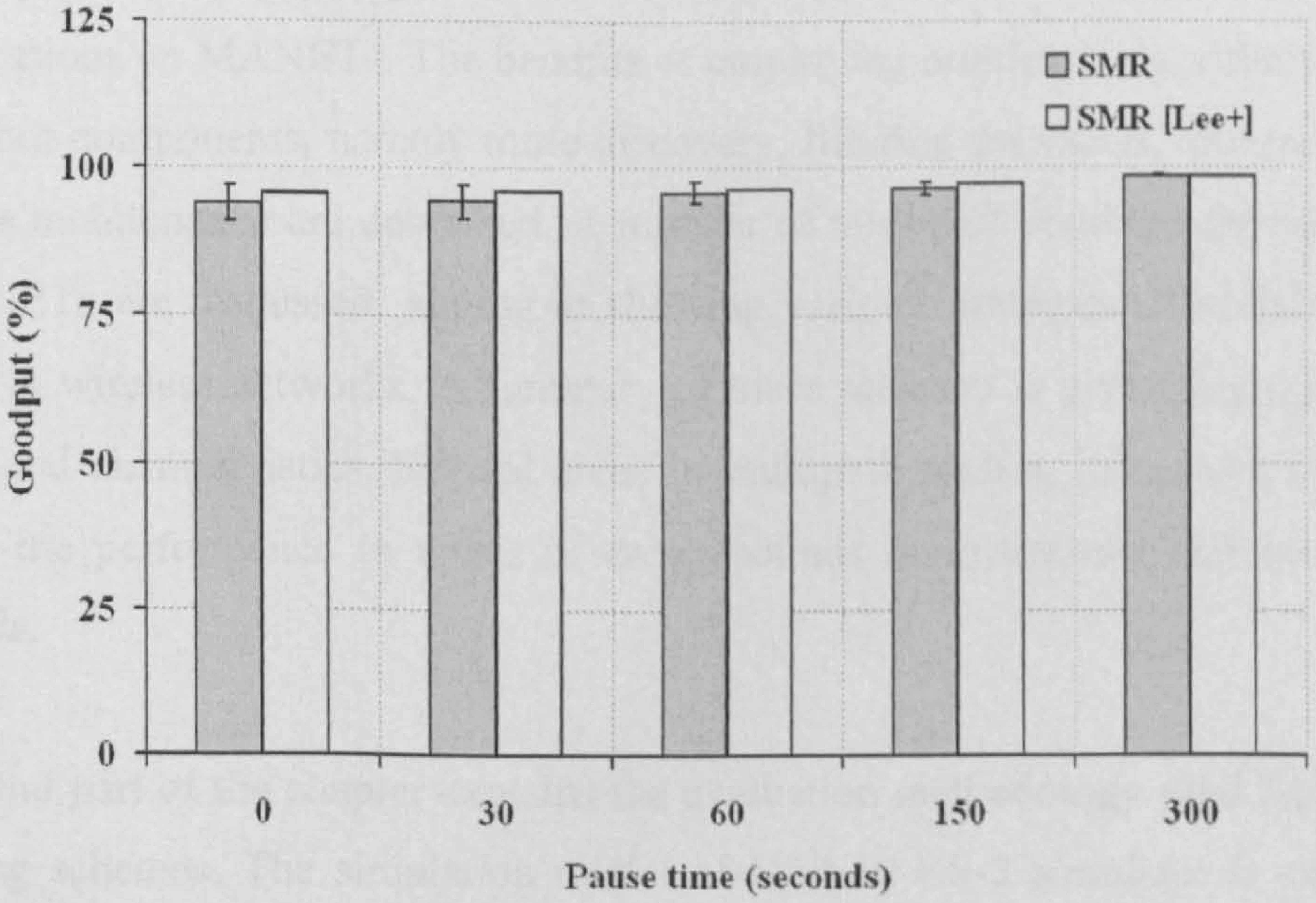


Figure 3.15 Goodput as a function of pause time

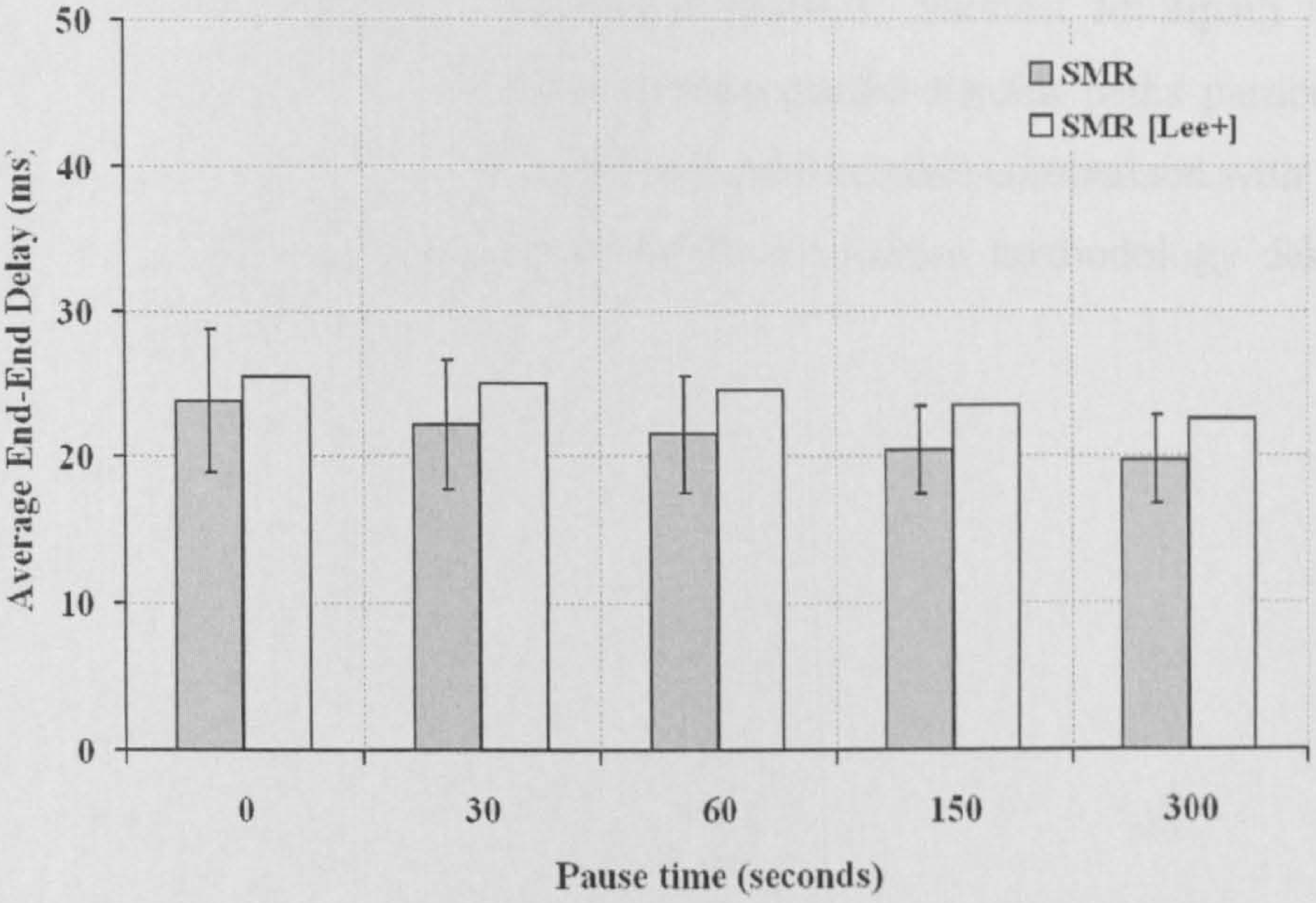


Figure 3.16 Average end-to-end delay as a function of pause time

3.8 Summary

The first part of the chapter presents the concept of multipath routing with emphasis on its applications on MANETs. The benefits of employing multipath algorithms in routing, and its four components, namely route discovery, filtering provision, route usage policy and route maintenance are described. A number of multipath routing schemes proposed for MANETs are discussed, aiming at showing various strategies of utilising multiple routings in wireless networks. A summary of these schemes is given, highlighting their features and characteristics. Several areas in multipath routing have been identified to improve the performance in terms of delay bounds or guarantees and lower routing overheads.

The second part of the chapter explains the evaluation methodology used for evaluation of routing schemes. The simulation model of DSR in NS-2 simulator is described. A number of preliminary simulation-based experiments for DSR and SMR are presented to raise confidence in the model behaviour and assumptions.

The next chapter proposes a novel and practical Shortest Multipath Source (SMS) routing scheme that builds multiple shortest partial-disjoint paths particularly suitable for real-time service delivery. An extensive performance comparison with the competing source-based routing schemes based on the evaluation methodology described in this chapter is performed.

4. Shortest Multipath Source (SMS) Routing Scheme

4.1 Introduction

In this chapter, a novel and practical routing scheme called Shortest Multipath Source (SMS) routing scheme, an extension to the DSR, is proposed. The principle objective of SMS is to build multiple partial-disjoint paths from source to destination in order to avoid the overhead of additional route discoveries and to recover quickly in case of route breaks. The performance differentials are investigated using NS-2 based model under conditions of varying mobility, offered load and network size. Section 4.2 describes Shortest Multipath Source (SMS) routing scheme in detail. Section 4.3 presents SMS routing process model. Section 4.4 presents verification and validation of the SMS scheme. In Section 4.5, the simulation framework is described. Section 4.6 discusses the simulation results and compares performances between unipath and multipath routing schemes. Section 4.7 concludes this chapter with a summary that highlights the unique aspects of SMS routing scheme and provides a discussion of possible further work related to SMS routing scheme.

4.2 Shortest Multipath Source (SMS) Routing Scheme

Motivated by limitations in on-demand multipath routing approaches (discussed in section 3.2), a novel and practical on-demand multipath routing scheme referred to as the Shortest Multipath Source (SMS) routing scheme is now proposed. The mechanism modifies and extends DSR, to build multiple partial-disjoint paths from source to destination in order to avoid the overhead of additional route discoveries and to recover quickly in case of route breaks.

SMS computes shortest multiple partial-disjoint paths that will bypass at least one intermediate node on the primary path. In other words, the alternate path need not necessarily be completely another than the primary path. Figure 4.1 illustrates partial-disjoint paths; the paths S-A-F-G-D, S-E-B-C-D, S-H-I-C-D and S-A-B-J-D are partial-disjoint compared to the primary path S-A-B-C-D. Partial-disjoint paths are different from both link-disjoint and node-disjoint multiple paths in the sense that partial-disjoint paths can have both nodes and links in common. This, less restrictive, constraint allows the computation of more partial-disjoint than node-disjoint or link-disjoint multiple paths and provides better fault-tolerance as a result of faster and efficient recovery from route breaks. The availability is represented as the average distance (in number of hops) between multipath nodes on the primary path. Multipath nodes are the nodes that have at least two paths to destination. The lower the inter-distance between nodes, higher is the availability of multiple paths, as more number of nodes on the primary path possesses alternative paths. Availability of link-disjoint or node-disjoint multiple paths drop more rapidly with increase in network size, when compared to partial-disjoint multiple paths. Also node-disjoint or link-disjoint paths tend to be longer than partial-disjoint multiple paths. This is because fully node-disjoint or link-disjoint multiple paths generally span a greater number of nodes or links in order to maintain disjointness.

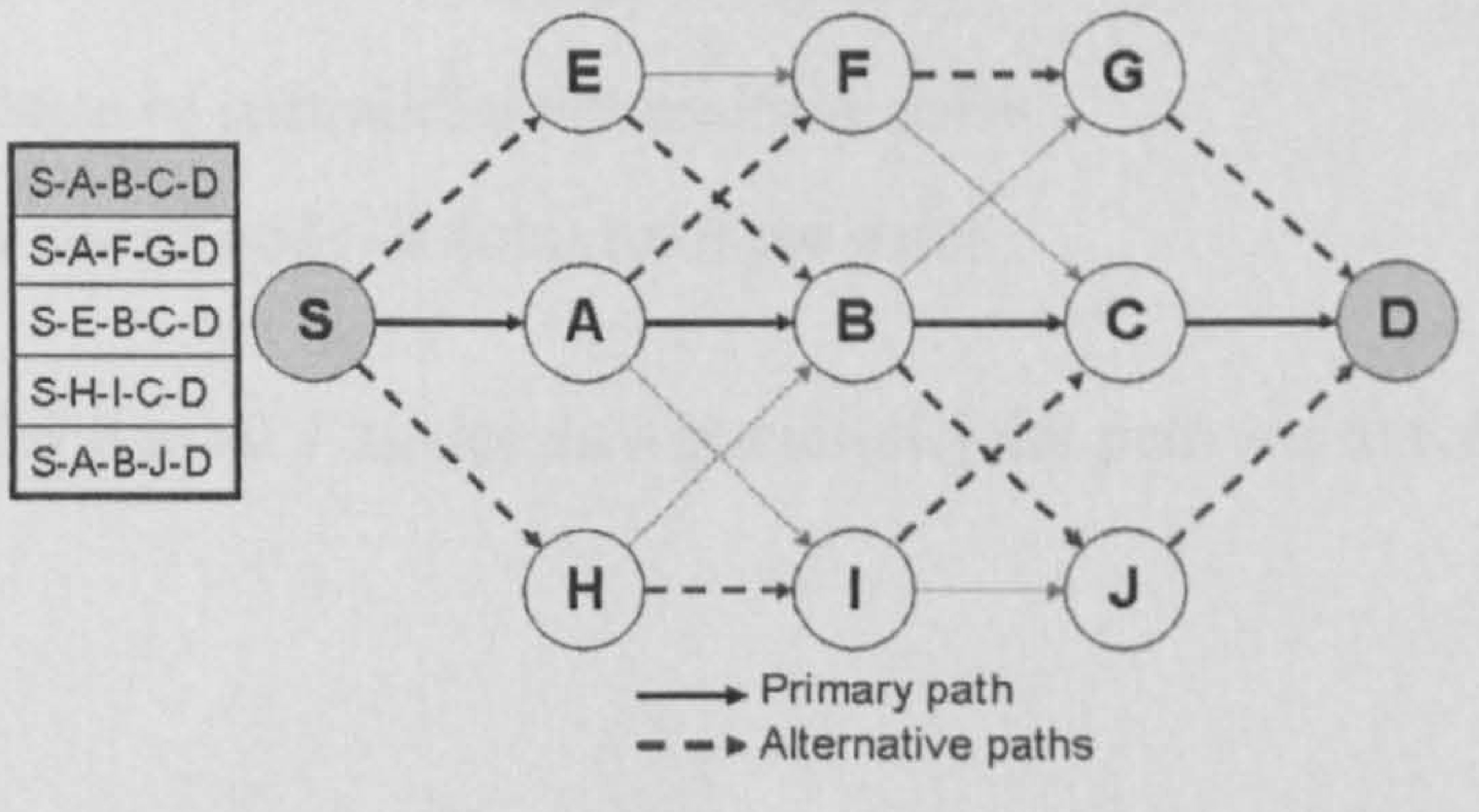


Figure 4.1 Partial-disjoint multiple paths

This section now describes the fault-tolerance analysis, route discovery phase, route reply phase, route maintenance phase and partial-disjoint path selection of the proposed algorithm.

4.2.1 Fault-tolerance Analysis

The fault-tolerance of a set of multiple paths between source and destination for a given node failure probability can be defined as the probability that at least one of the paths in the set are intact [63]. The following probabilistic analysis, which is an enhancement of the probabilistic analysis for a similar problem derived by L.R. Reddy et al [63], shows that partial-disjoint multiple paths have a greater fault-tolerance to route breaks due to their higher availability when compared to that of node-disjoint multiple paths.

The derivation applies the following definitions:

n number of nodes in a network
m number of partial-disjoint paths
m'' number of node-disjoint paths
l average length of partial-disjoint path (in number of nodes)
l'' average length of node-disjoint path (in number of nodes)
p probability that a node on a path is functioning
P_i probability that a partial-disjoint path i is intact
P_i'' probability that a node-disjoint path i is intact
F fault-tolerance of partial-disjoint multiple paths
F'' fault-tolerance of node-disjoint multiple paths

P_i is the probability that all l nodes on a partial-disjoint path are functioning,

$$P_i = p^l$$

Similarly,

$$P_i'' = p^{l''}$$

According to the definition of fault-tolerance,

$$F = 1 - (1 - p^{l})^{m}$$

Similarly,

$$F'' = 1 - (1 - p^{l''})^{m''}$$

As discussed in the previous section $l'' > l$ therefore,

$$(1 - p^{l''})^{m''} > (1 - p^{l})^{m''} \tag{4.1}$$

If $F - F'' > 0$ is proved, then the fault-tolerance of partial-disjoint multiple paths is greater than that of node-disjoint paths i.e. $F > F''$. Considering,

$$F - F'' = \{1 - (1 - p^{l})^{m}\} - \{1 - (1 - p^{l''})^{m''}\} \text{ or}$$

$$F - F'' = (1 - p^{l''})^{m''} - (1 - p^{l})^{m}$$

From (4.1),

$$F - F'' > (1 - p^{l})^{m''} - (1 - p^{l})^{m} \text{ or}$$

$$F - F'' > (1 - p^{l})^{m''} \left\{1 - \frac{(1 - p^{l})^{m}}{(1 - p^{l})^{m''}}\right\} \text{ or}$$

$$F - F'' > (1 - p^{l})^{m''} \{1 - (1 - p^{l})^{m - m''}\} \tag{4.2}$$

As discussed earlier $m > m''$ therefore,

$$\{1 - (1 - p^{l})^{m - m''}\} > 0 \tag{4.3}$$

For $m'' > 0$,

$$(1 - p^{l})^{m''} > 0 \tag{4.4}$$

Substituting (4.3) and (4.4) in (4.2),

$$F - F'' > 0 \text{ or}$$

$$F > F''$$

Hence, partial-disjoint multiple paths provide better fault-tolerance than node-disjoint multiple paths.

4.2.2 Route Discovery Phase

When a source seeks to communicate with a destination, it searches its cache to find any known routes to the destination. If no routes are available, the node initiates route discovery by flooding route-request (RREQ) packets, containing an ID which, along with source address, uniquely identifies the current discovery. When the RREQ packets are forwarded by the network nodes, each node appends its own address to the RREQ packet. Figure 4.2 shows the flowchart associated with initiating the route discovery process.

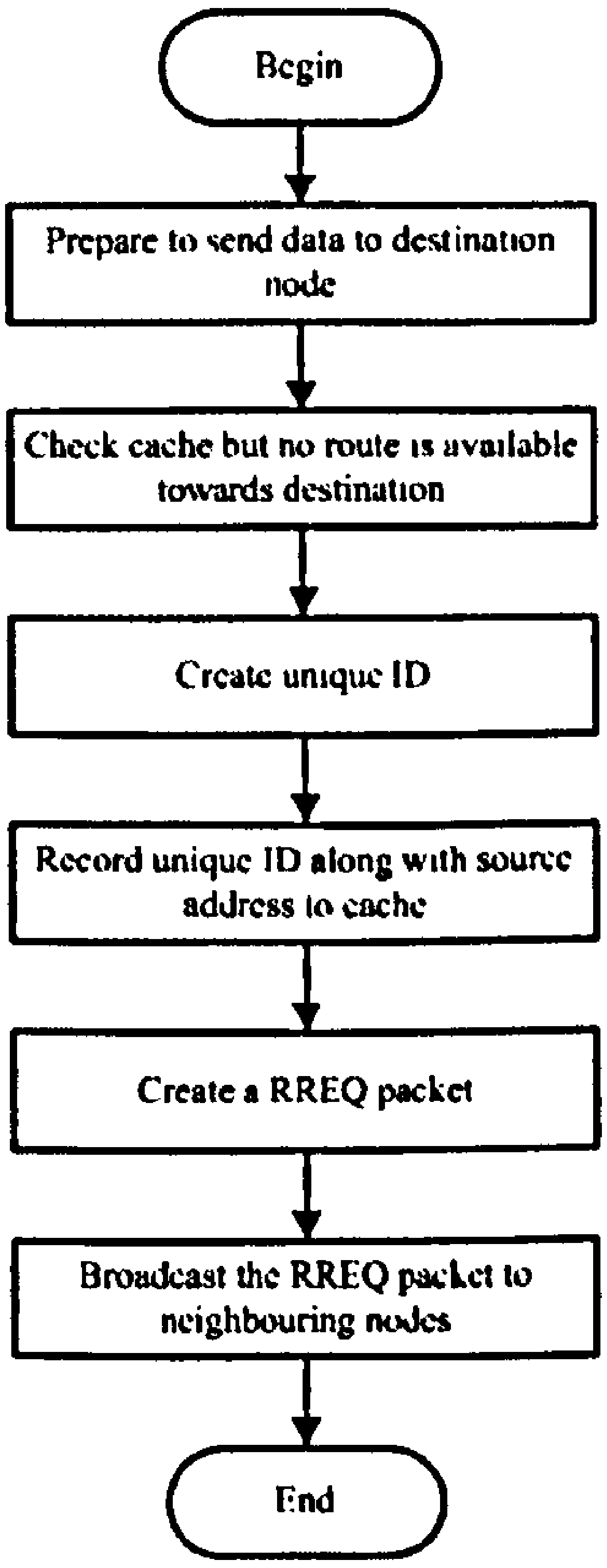


Figure 4.2 Flowchart of initiating a route discovery process

In multipath routing schemes, duplicate RREQ packets are broadcast to identify multiple paths from source to destination. However, if all duplicate RREQ packets are propagated, a broadcast storm can result and the performance of the network can degrade significantly. In order to overcome this problem [95], the SMS scheme introduces a novel approach to reduce the broadcast overhead. When a node receives a RREQ packet for the first time, it checks the route path from the packet and calculates the number of hops from the source to itself and records in its cache, the number as the reverse shortest hop count. If the RREQ duplicate is received again, the node computes the number of hops from the source to itself and compares this number to the reverse shortest hop count recorded in its cache. If the number of hops is less than or equal to the reverse shortest hop count, the node appends its own address to the route path list of the RREQ packet and broadcasts the RREQ packet to its neighbouring nodes. Otherwise, the node drops the RREQ packet.

The hop count is a simple mechanism which introduces minimum overhead. It may have limitations at high loads but represents an ideal first order approach to preventing broadcast storms. There are other storm prevention mechanisms such as location awareness [96]; dominant pruning [97] and connected dominating set [98]. But these do add complexity and can be considered as second order option.

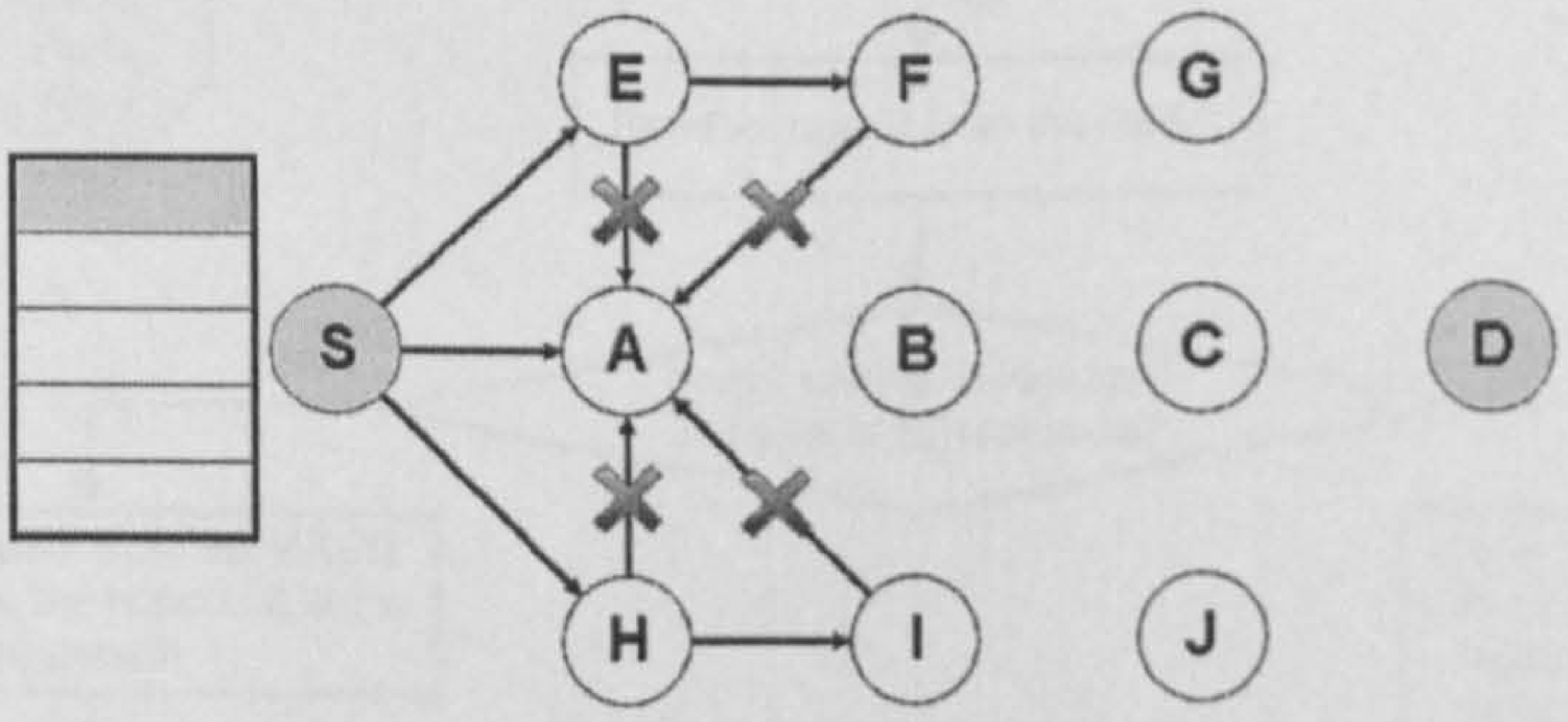


Figure 4.3 Reducing broadcast overhead

Figure 4.3 illustrates the computation to reduce broadcast overheads. When node A, in the first instance, receives the RREQ packet from path S-A, it records 1 as the reverse shortest hop count in its cache. When node A receives duplicate RREQ packets from the other four nodes, it calculates the number of hops and compares the current hop count to the reverse shortest hop count in its cache. As the numbers of hops of the four route paths are all greater than 1, the four duplicate RREQ packets are discarded. It is evident from the example that recording the ‘reverse shortest hop count’ approach ensures that duplicate RREQ packets are discarded during multiple partial-disjoint path discovery process. The flowchart describing the reduction in the broadcast routing overhead is given in Figure 4.4.

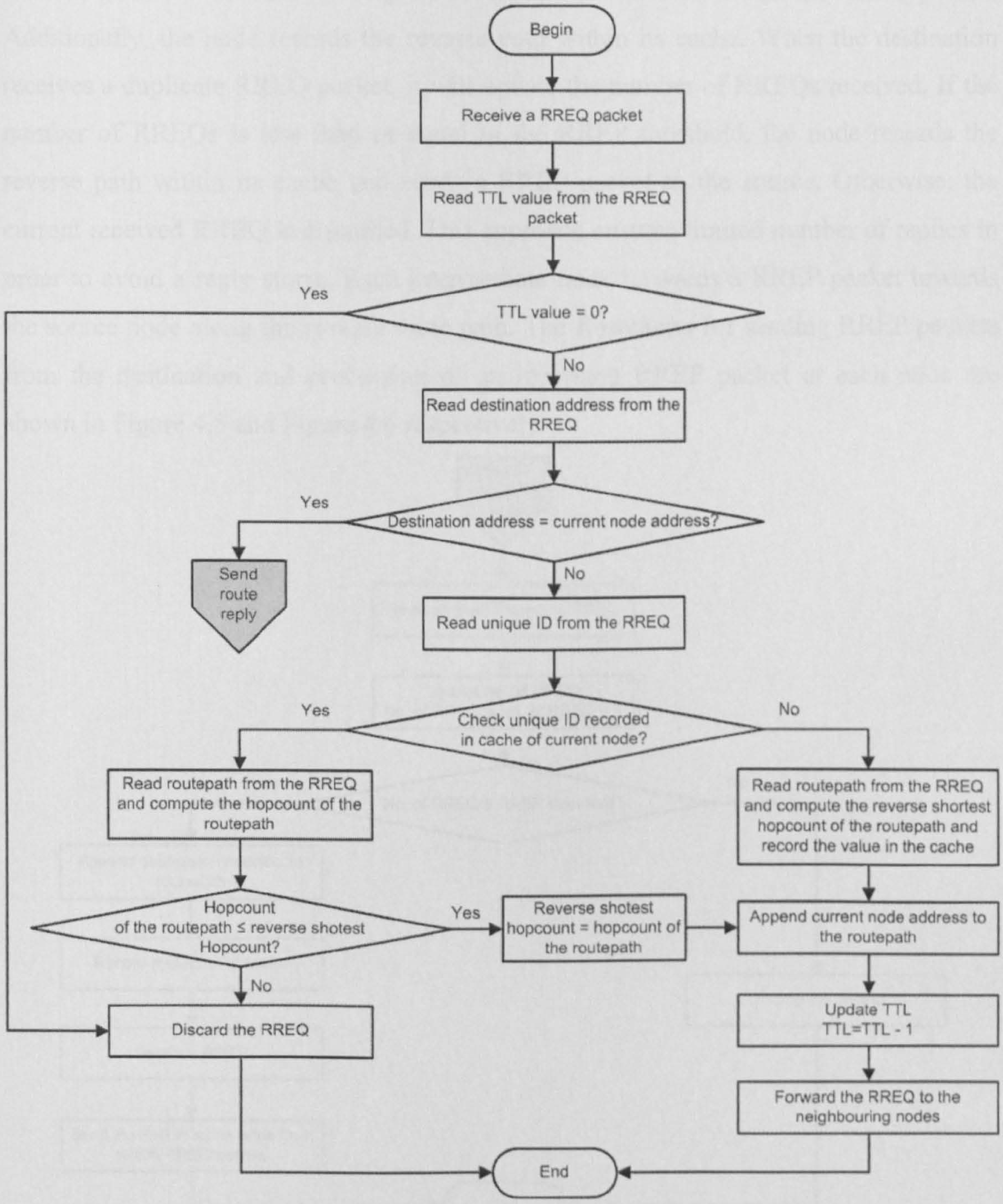


Figure 4.4 Flowchart of reducing broadcast routing overhead

4.2.3 Route Reply Phase

When the first RREQ packet arrives at its destination, the node sends a route-reply (RREP) packet to the source using the reverse path identified within the RREQ packet. Additionally, the node records the reverse path within its cache. When the destination receives a duplicate RREQ packet, it will update the number of RREQs received. If the number of RREQs is less than or equal to the RREP threshold, the node records the reverse path within its cache and sends a RREP packet to the source. Otherwise, the current received RREQ is discarded. This approach ensures limited number of replies in order to avoid a reply storm. Each intermediate node forwards a RREP packet towards the source node along the reverse route path. The flowcharts for sending RREP packets from the destination and processing of an incoming RREP packet at each node are shown in Figure 4.5 and Figure 4.6 respectively.

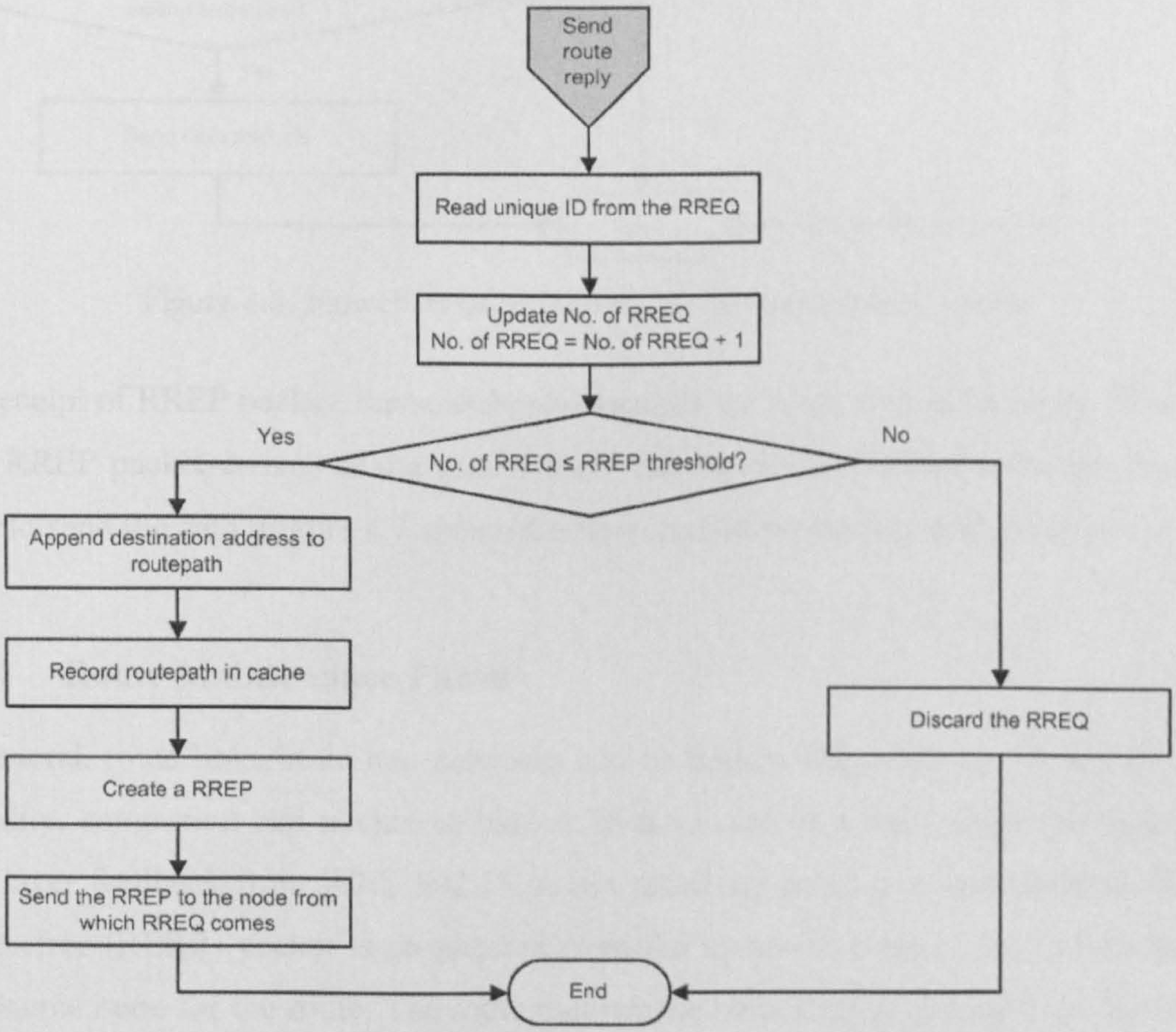


Figure 4.5 Flowchart of sending RREP packets from the destination

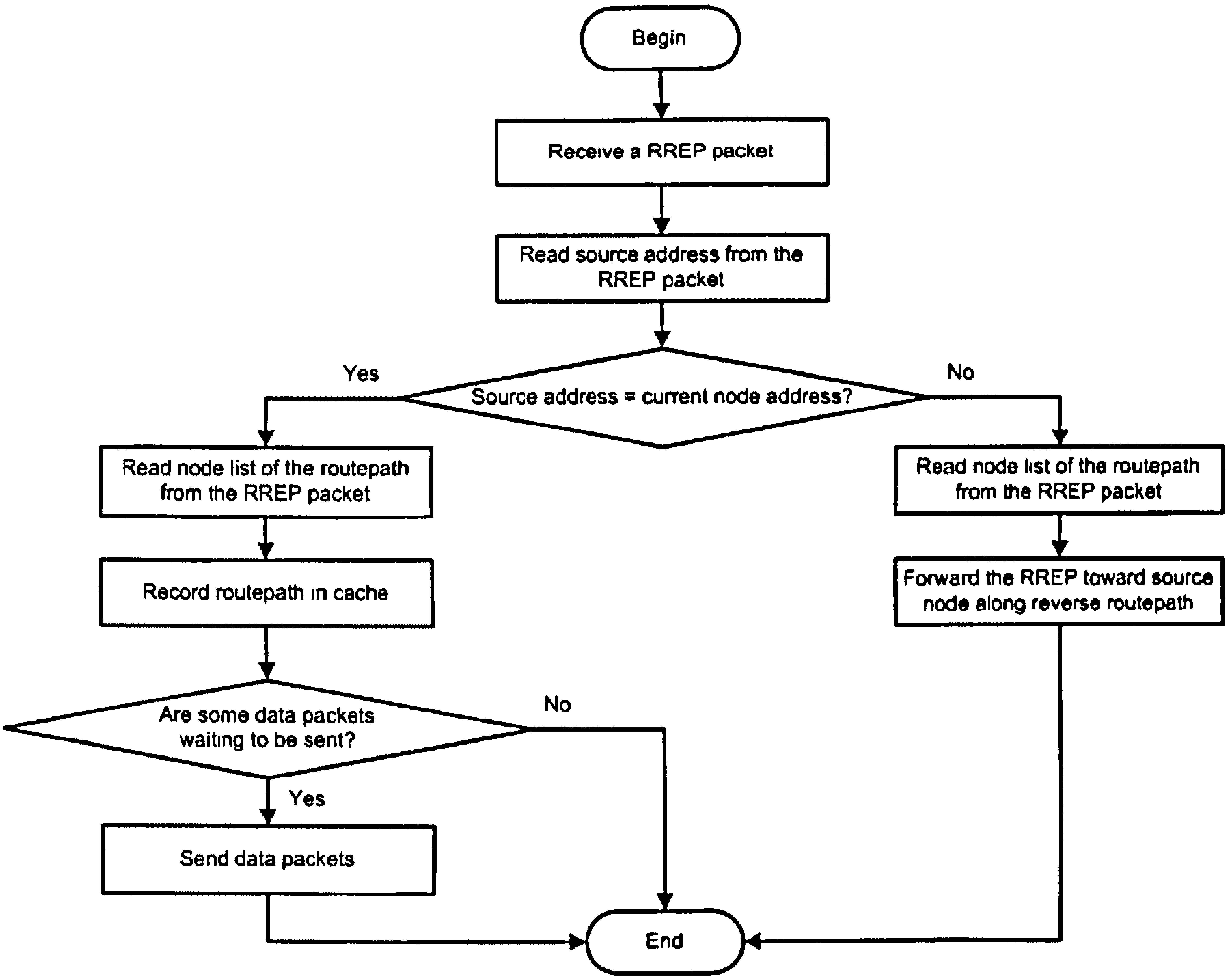


Figure 4.6 Flowchart of processing an incoming RREP packet

On receipt of RREP packet, the source node records the route path in its cache. Once the first RREP packet arrives at the source node, the newly established route can then be used to send the data. Figure 4.7 shows the flowchart of processing data packets.

4.2.4 Route Maintenance Phase

In general, route links in ad hoc networks can be broken frequently as a result of node mobility, congestion and packet collisions. In the event of a link failure (by receiving link layer feedback from IEEE 802.11 or not receiving positive acknowledgements), a route-error (RERR) packet is propagated from the upstream node of the link failure to the source node for the route. The route maintenance mechanism does not locally repair a broken link because stale route cache information at the intermediate node could result in inconsistencies during the route reconstruction phase. After receiving the RERR, the source invalidates every entry in its cache that uses the broken link (regardless of the destination) and then selects a new valid alternate routing path, randomly, from the candidate paths over which to forward any data packets. When there is only a single or no routing path available in the cache of the source node, the source instigates a new route discovery process to discover multiple partial-disjoint paths.

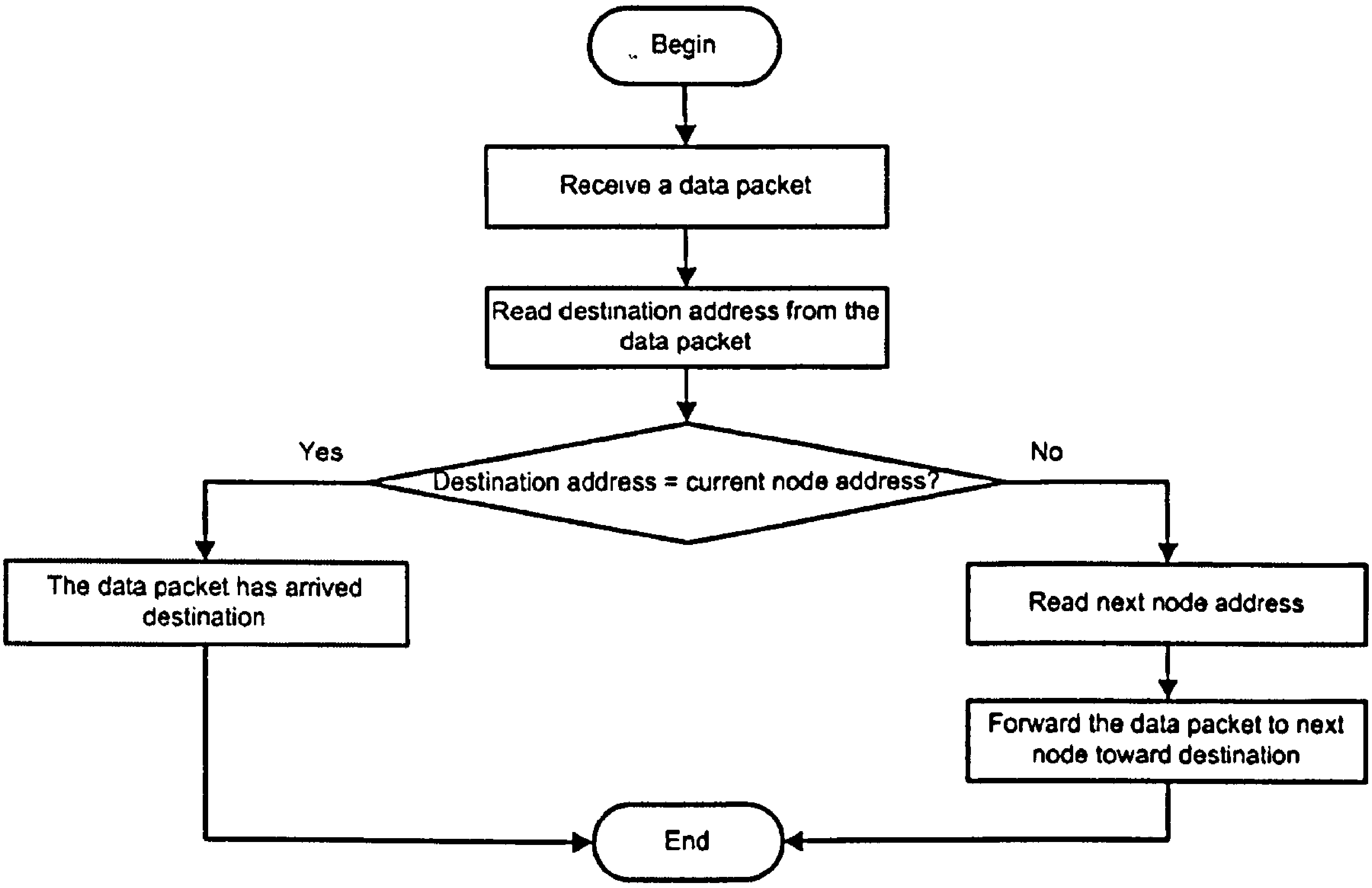


Figure 4.7 Flowchart of processing data packets

4.2.5 Partial-disjoint Path Selection

In the algorithm for selecting partial-disjoint paths, the source is responsible for selecting and recording multiple partial-disjoint route paths. An example that illustrates the selection of partial-disjoint paths is shown in Figure 4.8. Consider the case of traffic flowing between nodes S and D using link S-A-B-C-D as primary path. In the case of link failure between A-B, the source node will discard route paths S-A-B-C-D and S-A-B-J-D as it uses the broken link A-B. The path S-A-F-G-D is selected randomly from the candidate paths S-A-F-G-D, S-E-B-C-D and S-H-I-C-D. The flowchart for selecting partial-disjoint paths is given in Figure 4.9.

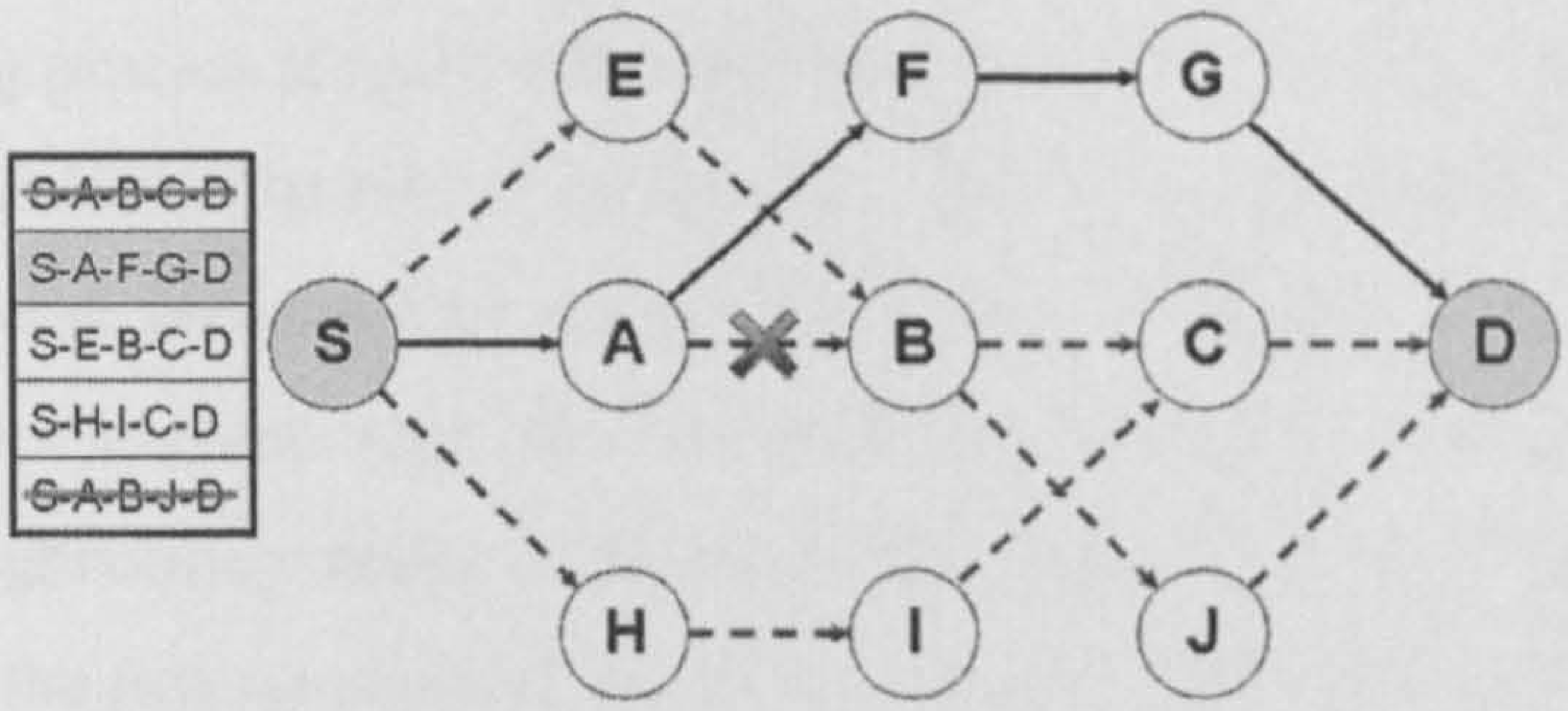


Figure 4.8 Selection of alternative path in case of link failure

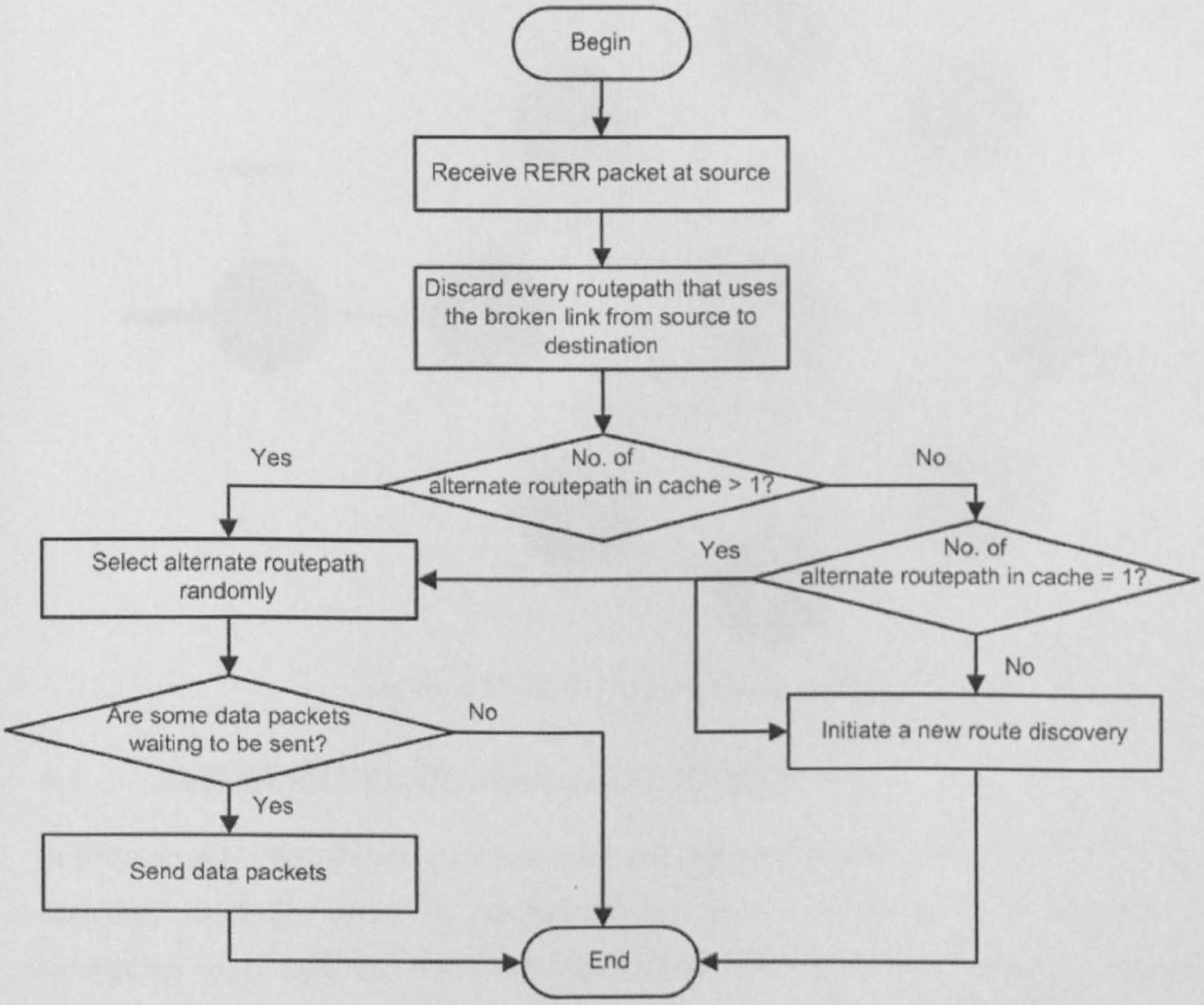


Figure 4.9 Flowchart of selecting partial-disjoint paths

4.3 SMS Routing Process Model State Machine

The SMS routing process (Figure 4.10) implements the SMS routing algorithm proposed in the previous section. The role of each state of this finite state machine is similar to the DSR process model presented in section 3.5.3 with an additional *Rebroad State*. This state occurs when current node receives a RREQ packet. It rebroadcasts the RREQ packet to its neighbouring nodes if the number of hops from the source to itself is less than or equal to the reverse shortest hop count recorded in its cache. Otherwise, the node drops the RREQ packet.

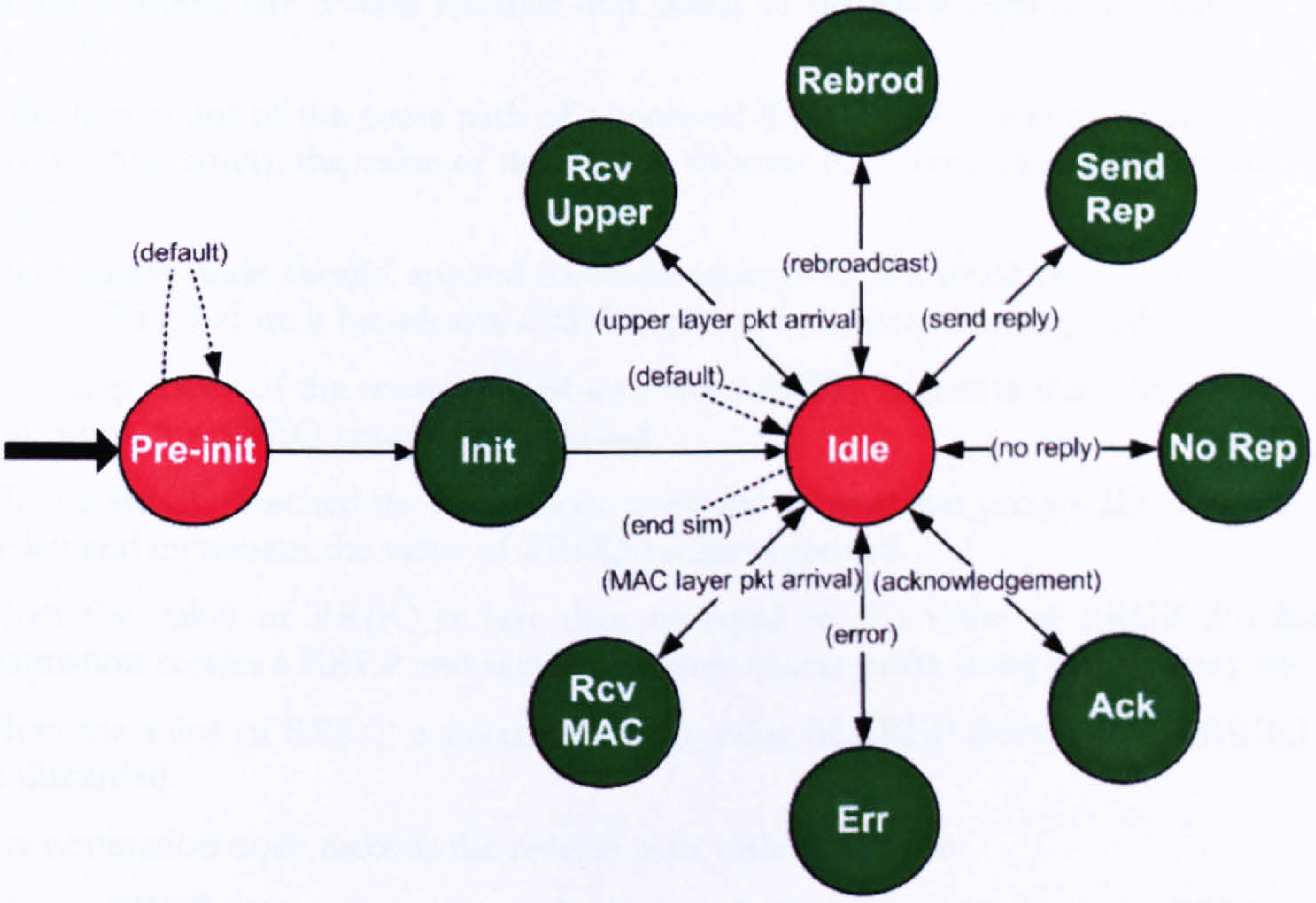


Figure 4.10 SMS routing process model

4.4 SMS Model Verification and Validation

In order to show that the simulation model and implementation of SMS are performing correctly, validation must be carried out to verify and validate the simulation. Debugging tools such as GNU Debugger (GDB) [99] and Data Display Debugger (DDD) [100] are used to trace and monitor the process of the simulation. The overall simulation results and intermediate results can be achieved in traces, breakpoints and files. A set of verification rules are defined as listed in Table 4.1 and the results are

'True' for all cases. The SMS model is evaluated by extensive performance comparison with the like-with-like source-based routing schemes presented in the next section.

Table 4.1 SMS model verification rules

#	Verification Rules
1.	If a node prepares to send data packets to a destination and no route is available towards the destination, it creates a RREQ packet containing its unique ID along with source address and broadcasts the RREQ packet to its neighbouring nodes.
2.	The value of Hop Limit / Time-to-Live (TTL) from any received RREQ packet should be greater than 1. Otherwise, the received RREQ packet should be discarded.
3.	If the intermediate node has no record of the unique ID received in the RREQ packet, it should compute the reverse shortest hop count of the route path and record the value in the cache.
4.	If the hop count of the route path of a received RREQ is less than or equal to the reverse shortest hop count, the value of the reverse shortest hop count should be updated into the cache.
5.	Intermediate node should append its node address to the route path and decrement the value of TTL before it broadcasts a RREQ packet to its neighbouring nodes.
6.	If the hop count of the route path of a received RREQ is greater than the reverse shortest hop count, the RREQ should be discarded.
7.	When a RREQ reached its destination, it should append the unique ID from the RREQ packet and increment the value of RREQ packets received.
8.	When the value of RREQ is less than or equal to the value of RREP threshold, the destination creates a RREP and sends it towards source node along the reverse path.
9.	When the value of RREQ is greater than the value of RREP threshold, the RREQ should be discarded.
10.	The destination node records the reverse path within its cache.
11.	When a RREP arrives in source node, it records the route path from the RREP packet in cache, before it forwards data packets waiting to be sent.
12.	When an intermediate node receives a data packet, it forwards the data packet to next node toward destination.
13.	When an intermediate node receives a RERR packet, it marks its route to the destination invalid and then propagates the RERR to its precursor node along the reverse route path.
14.	After receiving a RERR, source invalidates every route path to destination that uses the broken link and chooses a valid alternate routing path as active routing path from cache to forward any data packets waiting to be sent.
15.	When there is only a single or no routing paths available in the cache of a source node to send data packets, the source invalidates all route paths and initiates a route discovery process to get a new set of multiple partial-disjoint paths.

4.5 Simulation Framework

An NS-2 model forms the basis of the evaluation of the proposed SMS scheme. As SMS is a partial-disjoint multipath source routing scheme, comparing its performance with maximally node-disjoint split multipath routing (SMR) scheme [60] highlights the real strength of SMS. SMR is one of the more generally accepted reactive multipath source routing schemes [101][102][103]. SMS is also compared with a compatible unipath routing scheme, which, in this work, is DSR [41]. The latter ensures that results produced in this work are consistent and provide a meaningful benchmark on which to evaluate the proposed SMS scheme. Furthermore, it is important to determine whether the multipath scheme provides a better solution than its unipath counterpart.

4.5.1 Mobility and Traffic Model

The random waypoint mobility model [50] was used to simulate node movement in the network through varying speed, which can be setup using a scenario generator script *setdest* [78]. According to this model, a node randomly selects a location within the terrain area and moves towards it with a speed uniformly chosen between a pre-defined minimum and maximum values. The node stays in that position for a period (called pause time) and then selects to move to another random location. The less the pause time is, the more active the mobile nodes are in moving. The pause time is therefore kept constant at 30 seconds for all the simulations; a common value used in MANET research community [50][60][62].

Random constant bit rate (CBR) traffic and transmission control protocol (TCP) sessions are established between mobile nodes using a connection pattern generator script *cbrgen.tcl*.[78]. The traffic patterns used are widely acceptable benchmarks; CBR traffic commonly encompasses real-time traffic. CBR traffic more effectively stresses a network as there are no control mechanisms to consider when flows are delayed or packets lost. TCP can be unsuitable for most real-time applications because the protocol needs extra time to verify packets and request retransmissions. Traffic sources are CBR nodes chosen randomly from the full set of nodes generating 512 byte data packets at a rate of 4 packets per second.

Simulations are run for 900 seconds of real time[14]. Each data point represents an average from twenty runs each using different seeds, with a corresponding confidence interval of 95% [91].

4.5.2 Performance Metrics

Four key performance metrics [7][39][48][50][62][65] are used in varying scenarios to evaluate the three different schemes:

- *Goodput:* The ratio of the data packets delivered to destinations to those generated by the CBR sources. Goodput (packet delivery fraction) is a key metric since it shows the loss rate which, in turn, affects the maximum throughput of the network. This metric characterises both the completeness and correctness of a routing scheme.
- *Average End-to-End Delay:* Average End-to-End Delay includes all possible delays caused by buffering during route discovery, queuing at the interface queue, re-transmission delays at the MAC, propagation and transfer times.
- *Normalised Routing Load:* The number of routing packets transmitted per data packet delivered at the destination. Each hop-wise transmission of a routing packet is counted as one transmission. Normalised Routing Load is an important metric, as it measures the efficiency of a routing scheme; the degree to which it will function in congested or low-bandwidth environments.
- *Packet Loss:* The failure of one or more transmitted packets to arrive at their destination. The packet loss metric has characteristics similar to those of goodput.

---

[14] The total simulation time is initialized to 1200 seconds, and only the last 900 seconds involve actual data communication between nodes.

4.6 Simulation Results and Discussion

The main objective of the investigation is to measure the ability of the routing mechanisms to react to network topology changes while successfully delivering data packets to their destinations. To assess this ability, routing mechanism performance is evaluated and compared under a range of conditions.

Performance metrics must be measured against some parameter that describes the characteristics behaviour of an ad hoc network and can be varied in a controlled manner. The parameters chosen are:

- Mobility (speed of the nodes),
- Offered Load (number of sources) and
- Network Size (number of nodes, the size of the area that the nodes are moving within).

In each simulation, there are a range of key parameters that are configured; these are:

- *Speed:* The speed of a node is chosen randomly from the uniform range [0, maximum speed].
- *Number of nodes:* 50 nodes are used in all simulations except the network size simulation where the set of nodes was varied as cited in Table 4.4.
- *Terrain size:* A medium-sized MANET of 700x500 meters[15] is used for all simulations except the network size simulation where terrain was varied as cited in Table 4.4.
- *Simulation time:* The duration over which the simulations are run; a simulation time of 900 seconds was used for all simulations.
- *Pause time:* Pause time is the time for which every mobile node is still. A pause time of 30 seconds was adopted in all simulations.

---

[15] A rectangular space was chosen for the simulation, in order to explicitly make the path lengths between communicating nodes longer than would occur in a square space for greater variations.

4.6.1 Varying Mobility

Mobility is one of the important characteristics of an ad hoc network and can be defined as the average change in distance between all nodes over the simulation time. This effectively creates a dynamic topology; links will go up and down. Mobility is varied to investigate the impact it has on the different metrics being measured. This can be characterised by two parameters: pause time and speed. The speed parameter is used to control the scenario. By increasing the speed, the mobility will also increase. For randomized simulations, the speed is varied between the interval of 0 to 20 m/s. A speed of 0 m/s corresponds to a static network, whereas speed of 20 m/s corresponds to vehicle speed which represents high mobility. The simulation parameters used for the mobility simulations are shown in Table 4.2.

Table 4.2 Simulation Parameters - Mobility

Parameter	**Value**
Transmitter range	250m
Nominal channel bandwidth	2Mbps
Simulation time	900sec
Number of nodes	50
Pause time	30sec
Terrain size	700x500m^2
Traffic type	CBR
Packet rate	4 packets/sec
Packet size	512 bytes
Number of sources	25
Maximum speed	0, 2, 4, 8, 12, 16, 20 m/s

4.6.1.1 Goodput

The goodput of the three schemes is shown in Figure 4.11 which depicts the variation of the goodput as a function of node speeds. As the speed of the nodes increases, the probability of link failure increases and hence the number of packet drops also increases.

In stationary scenarios, SMS and SMR both exhibit a slight improvement over DSR. In fact, when the nodes are stationary, all schemes should show the same goodput performance, as there are no mobility induced route breaks. The reason for this behaviour is attributed to the congestion experienced by nodes when all 25 sessions are active. Congestion increases packet collisions and nodes attempt to retransmit the collided packets. Finally, nodes drop the packets after maximum number of retransmission attempts and the source node initiates a fresh route discovery to find a new path to destination. Although, the same situation occurs in SMS and SMR, the source node uses an alternate path to forward packets, avoiding extra route discovery attempts and error packet transmissions. As mobility increases, the schemes behave as expected. SMS has much higher goodput than both SMR and DSR due to the availability of shorter multiple routing paths with partial-disjointness. Although SMR also uses multiple paths to redirect unfinished sessions, relatively few intermediate nodes possess secondary paths compared to the SMS case. This indicates the more robust nature of the SMS as nodes become more mobile.

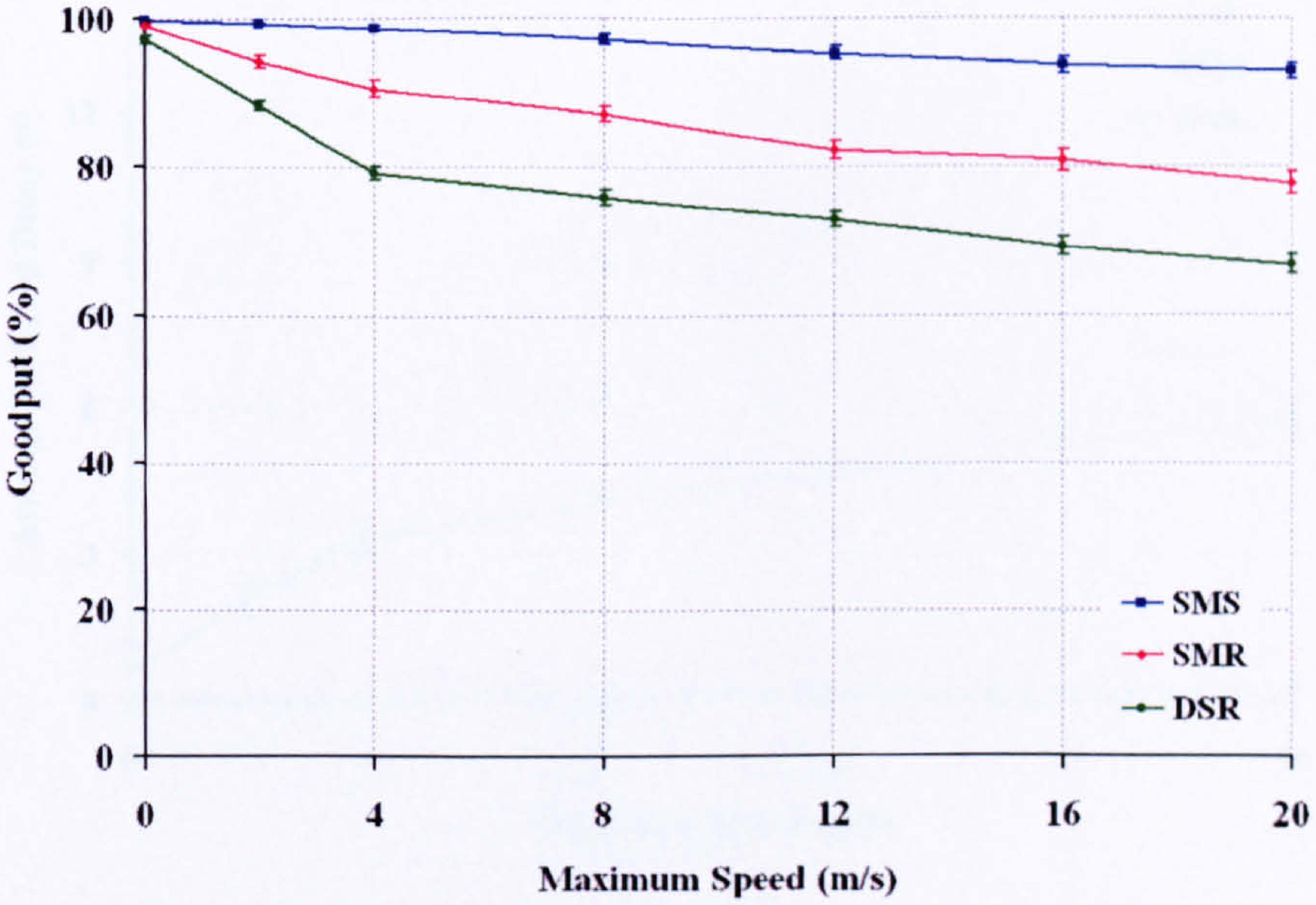


Figure 4.11 Goodput versus maximum speed

4.6.1.2 Average End-to-End Delay

The average end-to-end delay includes all possible delays from the instant the packet is generated to the instant it is received by the destination node. Generally, there are three factors affecting end-to-end delay of a packet: route discovery time, caused by packets waiting in the queue before a route path is found; buffer waiting time, caused by packets waiting in the queue before they can be transmitted; the length of routing path. The more hops a data packet has to traverse, the longer it will take to reach its destination node. As shown in Figure 4.12 the average end-to-end delay of SMS is lower than SMR and DSR. The availability of shorter alternate routing paths in the former eliminates route discovery latency that contributes to the delay when an active route fails. In addition, when a congestion state occurs in a path, the source node, to manage congestion can re-distribute incoming data packets to alternative routing paths, thereby reducing the queue waiting time of data packets. The increase in DSR's delay is the result of the resurrection of data packets by source nodes when the MAC-managed transmissions fail.

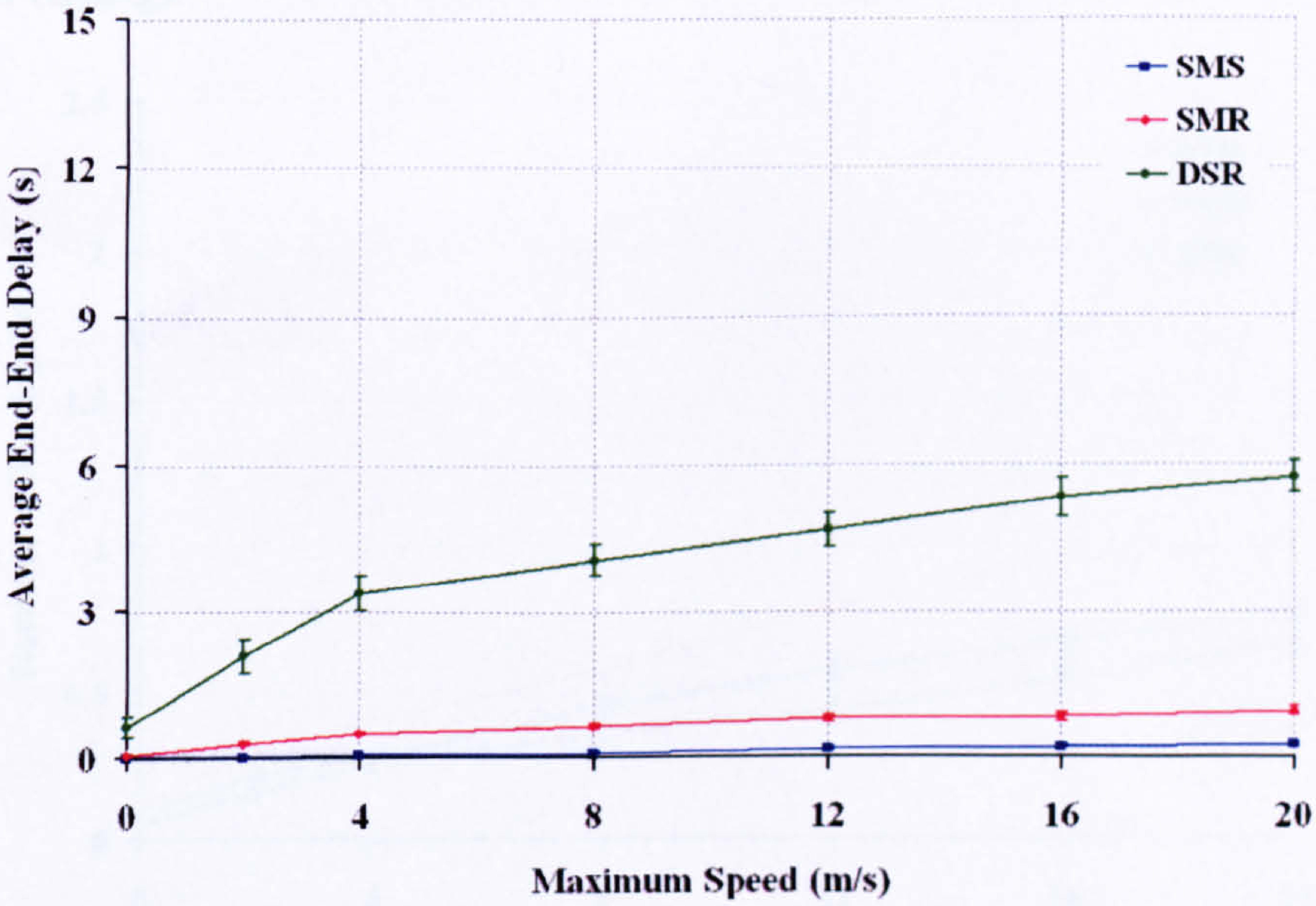


Figure 4.12 Average end-to-end delay versus maximum speed

4.6.1.3 Normalised Routing Load

Normalised routing load is an important metric since yields a measure of the efficiency of protocols; particular important in a low bandwidth and congested wireless environments. Protocols that transmit a large number of routing packets can also increase the probability of packet collisions and the waiting time of data packets in transmission buffer queues. Figure 4.13 presents the normalised routing load characteristics of the 50 node network showing that the normalised routing load in SMS is lower than SMR. SMS finds the shortest alternative routes in a single route discovery process, thus reducing delay and routing overheads incurred when recovering from route breaks. SMR has higher routing overheads as a result of excessive broadcasts of duplicate RREQ packets in order to determine maximally disjoint paths. DSR, by virtue of aggressive caching, is more likely to find a route in the cache and, hence, resorts to route discovery less frequently; however DSR generates more replies and errors (gratuitous or otherwise). Thus, all routing load savings for DSR are a consequence of a saving in RREQs.

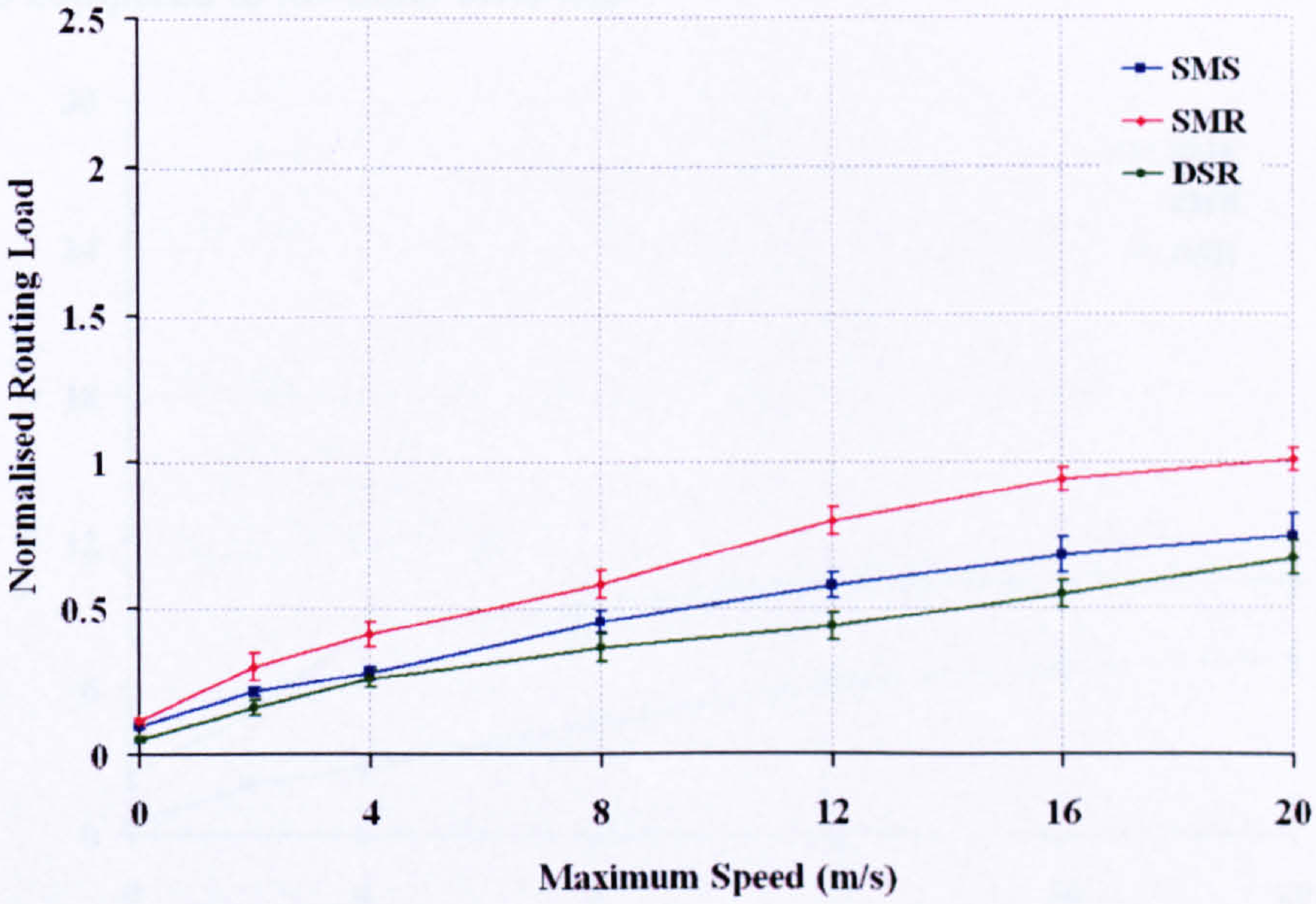


Figure 4.13 Normalised routing load versus maximum speed

4.6.1.4 Packet Loss

Packet loss represents another important measure that quantifies performance and also qualifies the goodput characteristics presented earlier. Packet loss mechanisms are much more complicated in mobile ad hoc networks because wireless links are subject to transmission errors and the network topology changes dynamically. A packet may be lost due to transmission errors, the non-existence of a route to the destination, broken links, congestions, etc. The cumulative effects of these factors are tightly associated with the network context, e.g. host mobility, number of connections, traffic load, etc. Building an approximate model to analytically evaluate packet loss is difficult and, thus, the preferred evaluation mechanism is via simulation. Packet loss of the three schemes as function of mobility is shown in Figure 4.14. Packet loss takes into account packet drops at the MAC and the network layer. Packet losses follow increases in mobility because the protocol is sending RREQ packets on a broken route that it still considers being valid and, thus packets in node buffers are dropped as a result of congestion and timeouts. Packet loss in SMS is lower, simply as a consequence of better link fault-tolerance compared to the other schemes.

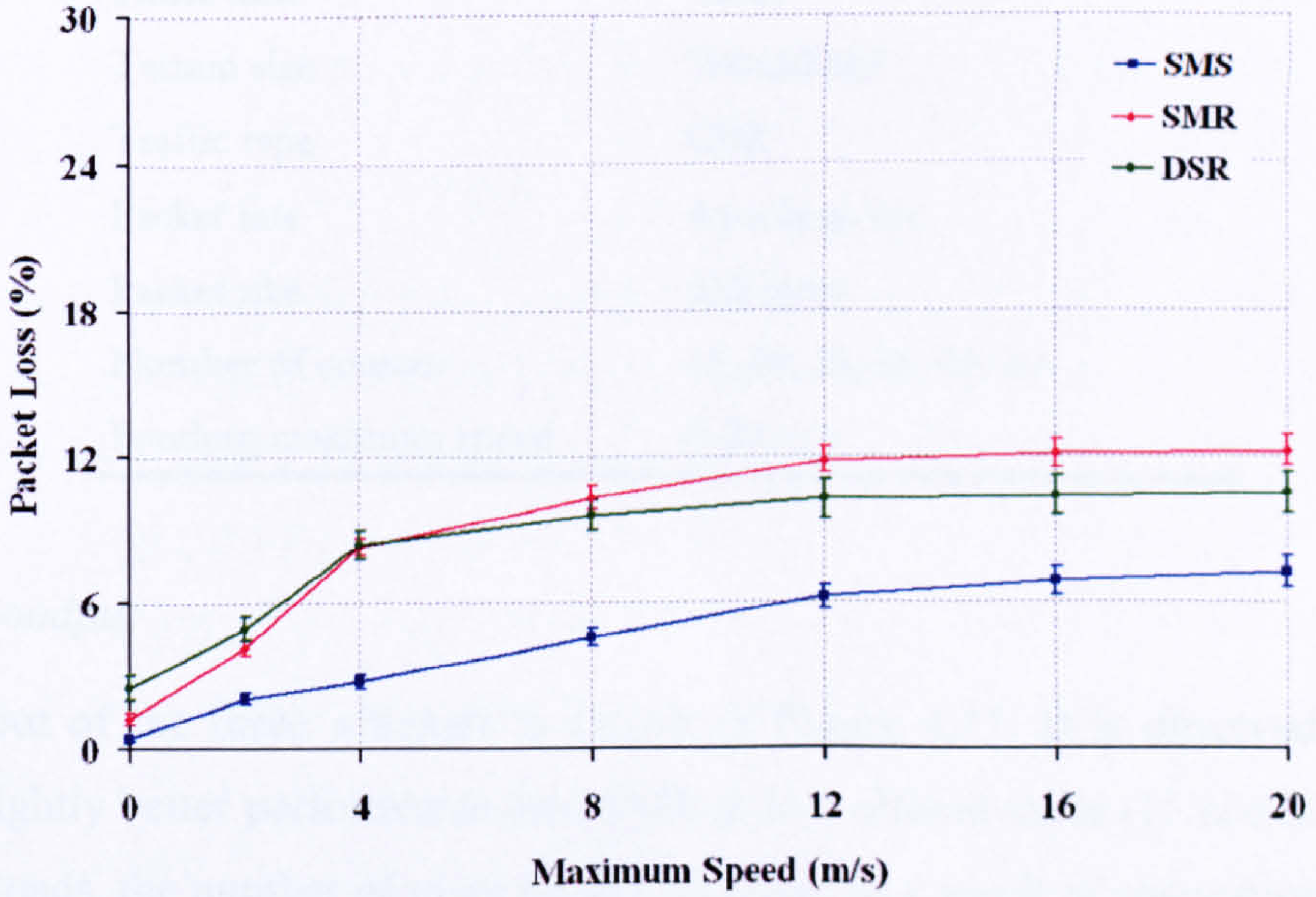


Figure 4.14 Packet loss versus maximum speed

4.6.2 Varying Offered Load

In this section, the offered load is varied to see how the mechanisms behave when, for instance, the load is high. This can be characterised by three parameters: packet size, number of connections and send packet rate. In this set of results, the network load is varied by changing the number of active connections (sources) between 15 and 40 with mobility defined by a random speed of 0-20 m/s. The number of active sessions in the network affects protocol performance because it determines the level of overhead in the network; the greater the number of sessions, the greater the number of route discovery attempts and hence increased routing packet transmissions. The parameters used in the load based simulations are listed in Table 4.3.

Table 4.3 Simulation Parameters – Offered Load

Parameter	Value
Transmitter range	250m
Nominal channel bandwidth	2Mbps
Simulation time	900sec
Number of nodes	50
Pause time	30sec
Terrain size	700x500m^2
Traffic type	CBR
Packet rate	4 packets/sec
Packet size	512 bytes
Number of sources	15, 20, 25, 30, 35, 40
Random maximum speed	0-20 m/s

4.6.2.1 Goodput

The goodput of the three schemes is shown in Figure 4.15. It is observed that SMS exhibits slightly better performance than SMR at low offered loads (15 and 20 sessions). At higher loads, the number of route breaks increases as a result of congestion caused by the increased numbers of active sessions. Route breaks occur as there are packet drops caused by collisions that result from channel congestion, even though the route between source and destination remain intact. DSR's overhead then increases as it initiates a fresh route discovery for every route break. SMS outperforms SMR and DSR by using shorter secondary paths to repair route breaks. SMR produces a large number of control packets which in turn significantly overloads the network.

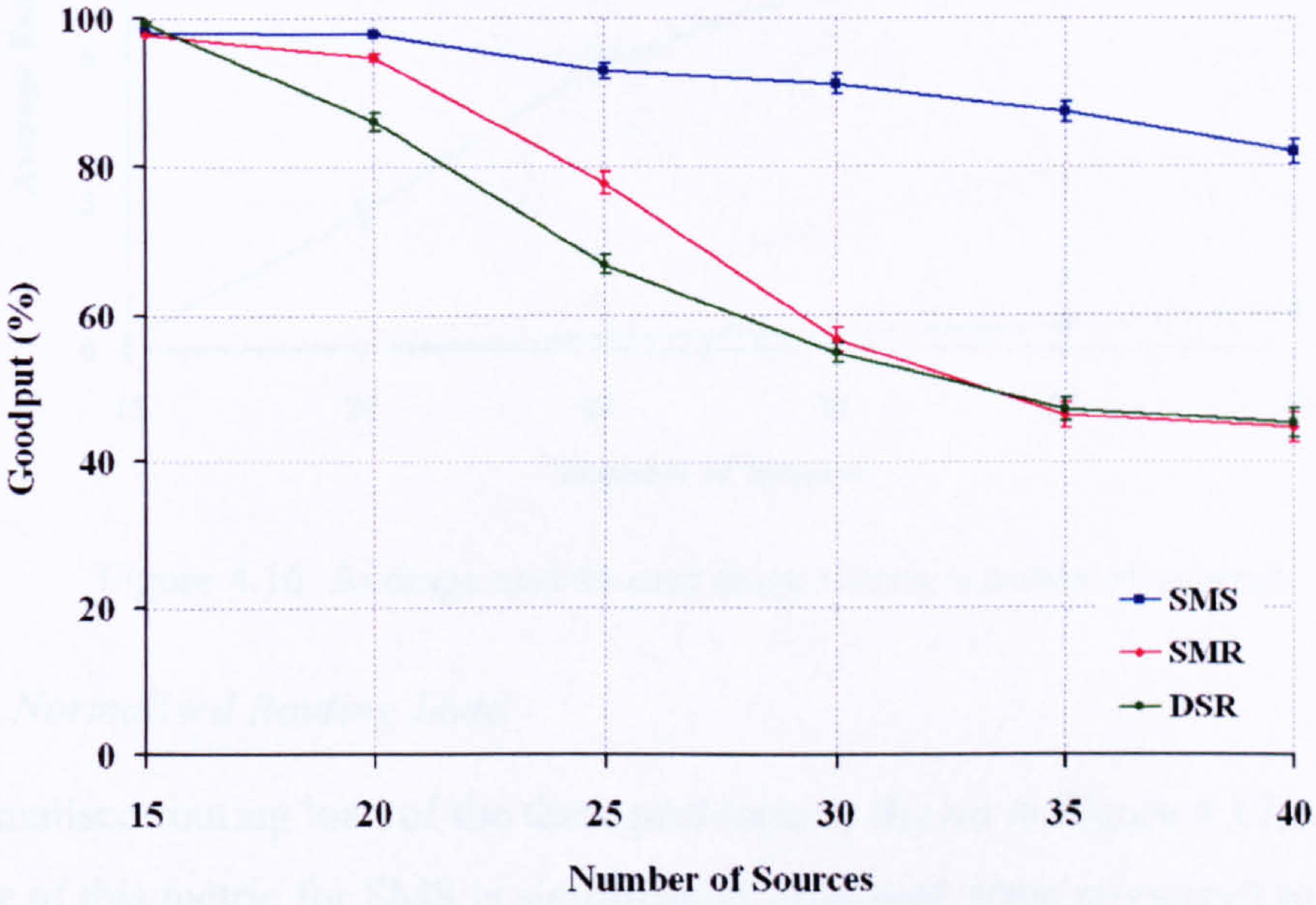


Figure 4.15 Goodput versus number of sources

4.6.2.2 *Average End-to-End Delay*

The average end-to-end delay as a function of the number of sources is shown in Figure 4.16 depicting that SMS has a lower average delay over all number of sources. Although SMS offers the shortest alternative routing paths, its delay also increases gradually with an increase in the number of sources which may in turn result in the bottleneck of the node. DSR exhibits high delays compared to SMS and SMR since an increase of the numbers of sources leads to higher network load traffic and consequently more routes will become invalid, further necessitating new requests. While these new requests propagate through the network in search for a new route, buffers become full and packets are then dropped.

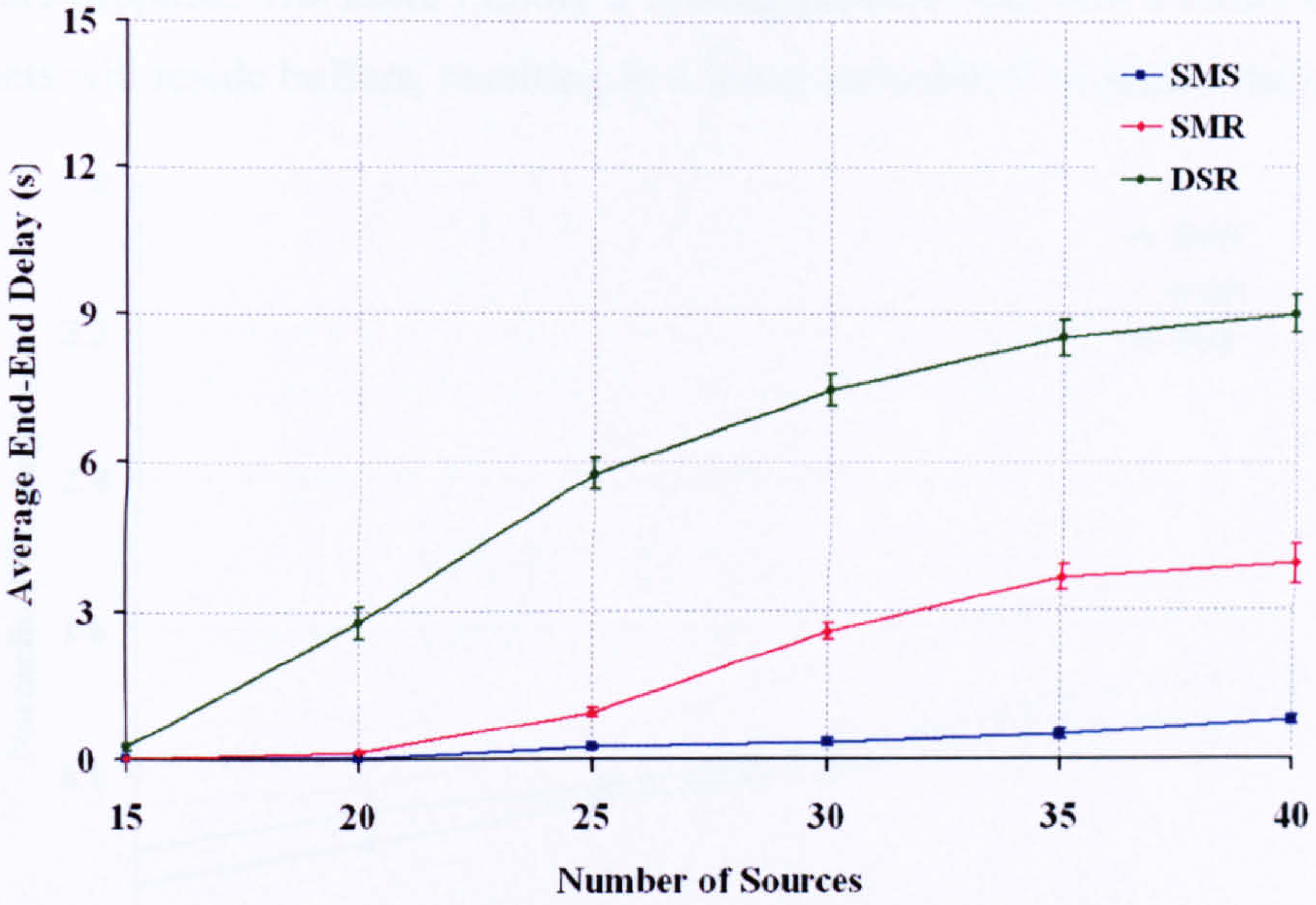


Figure 4.16 Average end-to-end delay versus number of sources

4.6.2.3 Normalised Routing Load

The normalised routing load of the three protocols is shown in Figure 4.17, showing that the value of this metric for SMS is significantly improved when compared to SMR. With an increase of the number of sources, the probability of packet collisions and packet congestion increases, leading to an increase in normalised routing load. DSR has the lowest value; multipath protocols generate more control packets while building multiple routes. DSR builds single routes for each session and minimizes the flooding overhead by allowing intermediate nodes to reply from their cache.

4.6.2.4 Packet Loss

Packet loss for the three schemes is shown in Figure 4.18; SMS has the lowest packet loss, a consequence of lower route discovery latency. However, for all schemes, increasing communication requests has an impact on packet loss, i.e. losses increase in a manner similar to the case when traffic loads are increased. Congestion occurs and thus packets are dropped. The more rapidly a routing protocol can find a route, the less time the packets will reside buffers, resulting in a lower probability of packet drop.

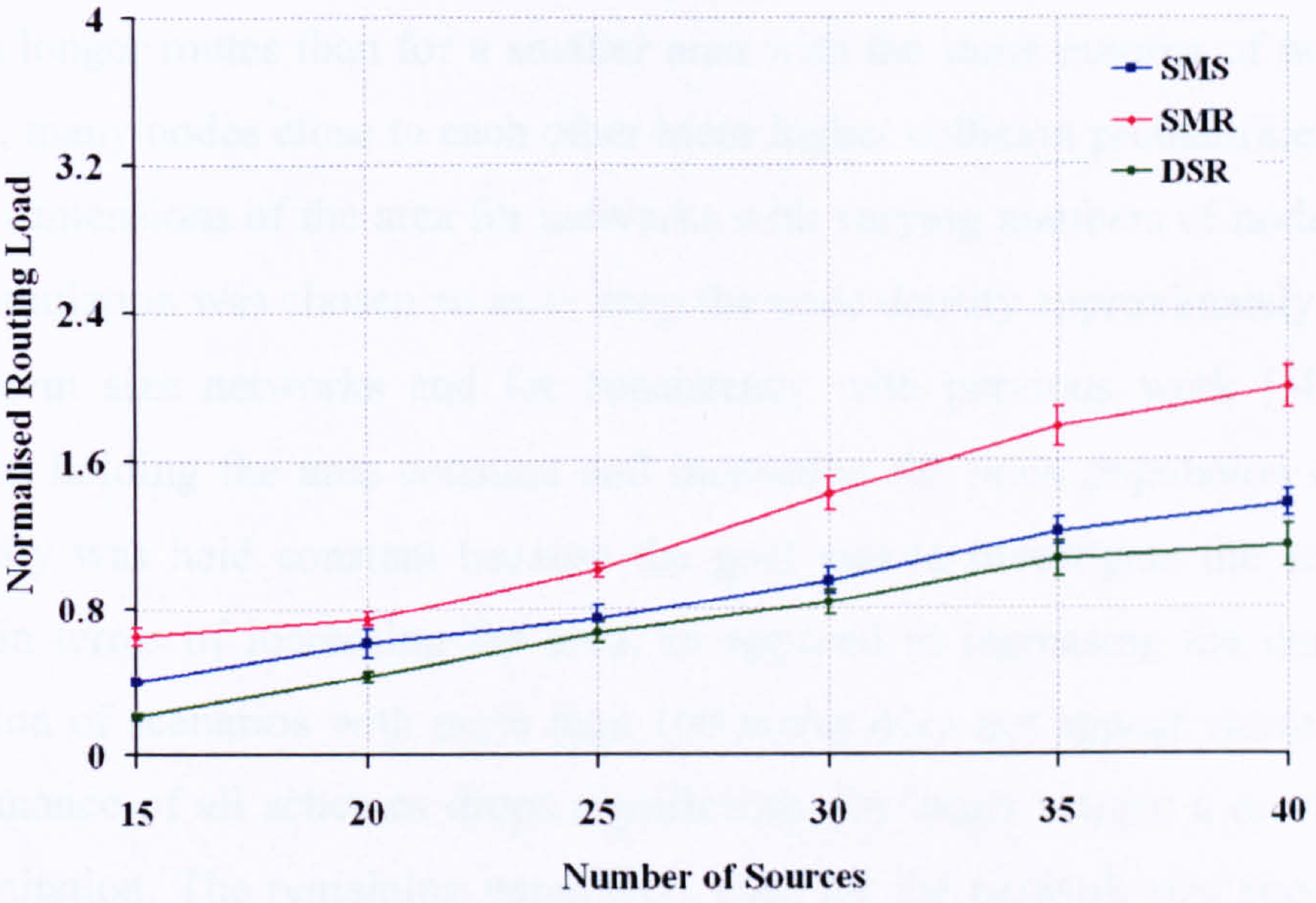


Figure 4.17 Normalised routing load versus number of sources

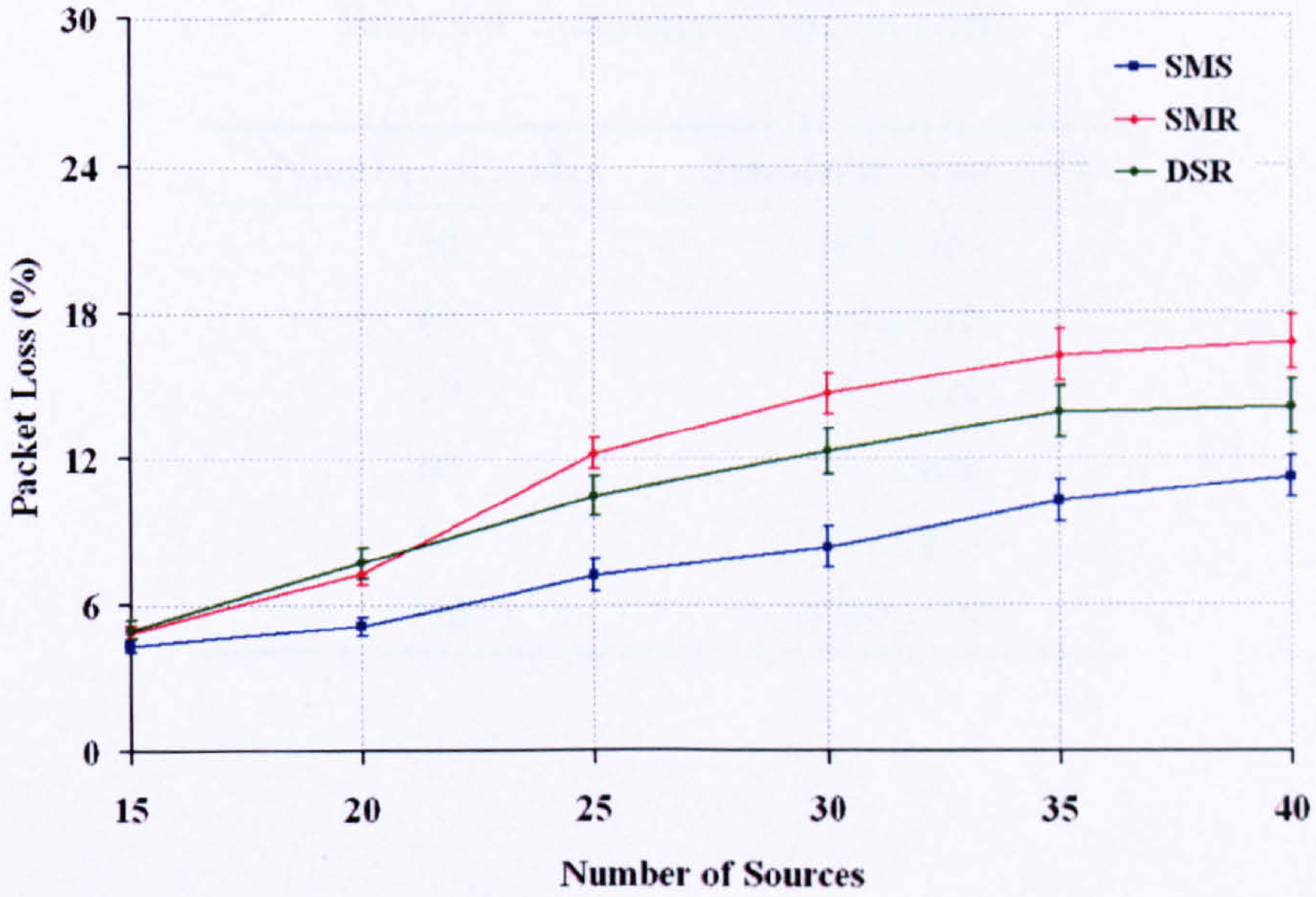


Figure 4.18 Packet loss versus number of sources

4.6.3 Varying Network Size

Goodput, average end-to-end delay, normalised routing load and packet loss are computed by increasing the number of nodes in the network from 50 to 100 nodes. Changing network size shows scalability performance. A large area with many nodes may mean longer routes than for a smaller area with the same number of nodes. At the same time, many nodes close to each other incur higher collision probabilities. Table 4.4 shows the dimensions of the area for networks with varying numbers of nodes. The area for each simulation was chosen so as to keep the node density approximately constant in such different size networks and for consistency with previous work [50][58][104]. Rather than holding the area constant and increasing the node population density, the node density was held constant because the goal was to investigate the scalability of networks in terms of increasing the area, as opposed to increasing the density [104]. Investigation of scenarios with more than 100 nodes does not appear worthwhile since the performance of all schemes drops significantly for larger setups; a result of source routing limitation. The remaining parameters used for the network size simulations are same as used in load based simulations (Table 4.3) except the number of active sources which is kept constant at 20 CBR nodes, chosen at random from the full set of nodes.

Table 4.4 Summary of terrain areas

Number of nodes	Simulation area (m^2)
50	700 x 500
60	780 x 540
70	850 x 580
80	910 x 620
90	960 x 660
100	1000 x 700

4.6.3.1 Goodput

The goodput of the three schemes is shown in Figure 4.19, showing that the ability of the schemes to deliver packets to their destination degrades as the network size increases. The path length is greater in larger networks because, whilst the area and the number of nodes increase, the node density remains constant. Routes are more prone to disconnections in mobile networks when path lengths are longer. Because any link failure along the path results in an inability of the source to reach the destination, longer routes have a greater probability of route disconnect than shorter hop routes. SMS, due to its shorter paths outperforms DSR and SMR. Although, SMR also uses multiple paths at intermediate nodes to redirect unfinished sessions, the numbers of intermediate nodes that possess secondary paths is small when compared to SMS. SMR goodput becomes worse than DSR at higher network sizes because the longer secondary paths in SMR fail more quickly due to their higher probability of route breakage.

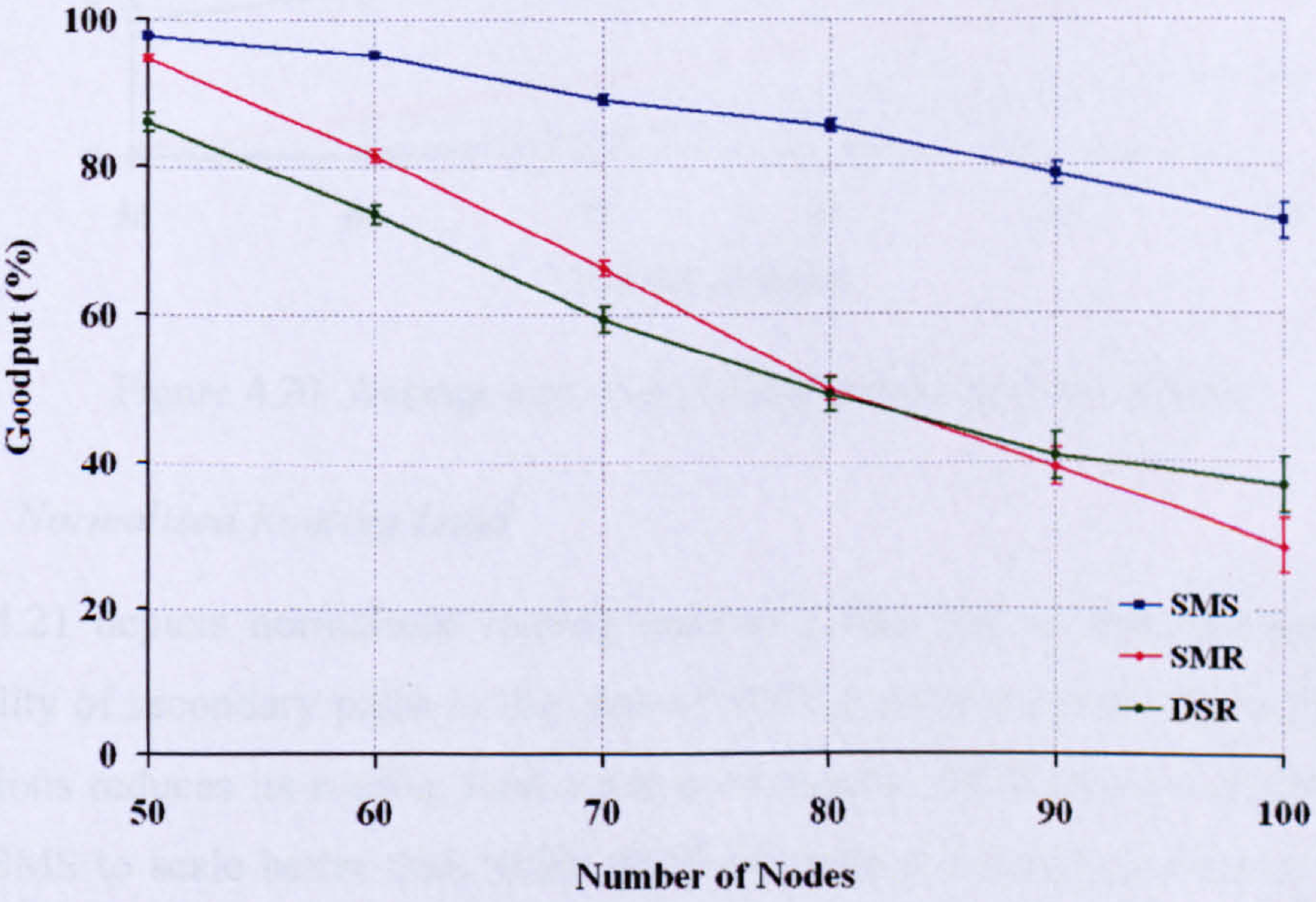


Figure 4.19 Goodput versus number of nodes

4.6.3.2 Average End-to-End Delay

As can be seen in Figure 4.20, the average end-to-end delay of SMS is lower than both DSR and SMR delay. This metric reflects the delay incurred in resuming the sessions after route breaks have occurred. The delay is higher for DSR when compared to SMS and SMR because DSR has to re-initiate the route discovery process in order to resume the session. Additionally, in large networks, the broken secondary paths of SMR delay the initiation of fresh route discovery. SMS has the lowest delay value at all network sizes, as it finishes the session with the lowest number of route computations compared to the other two schemes.

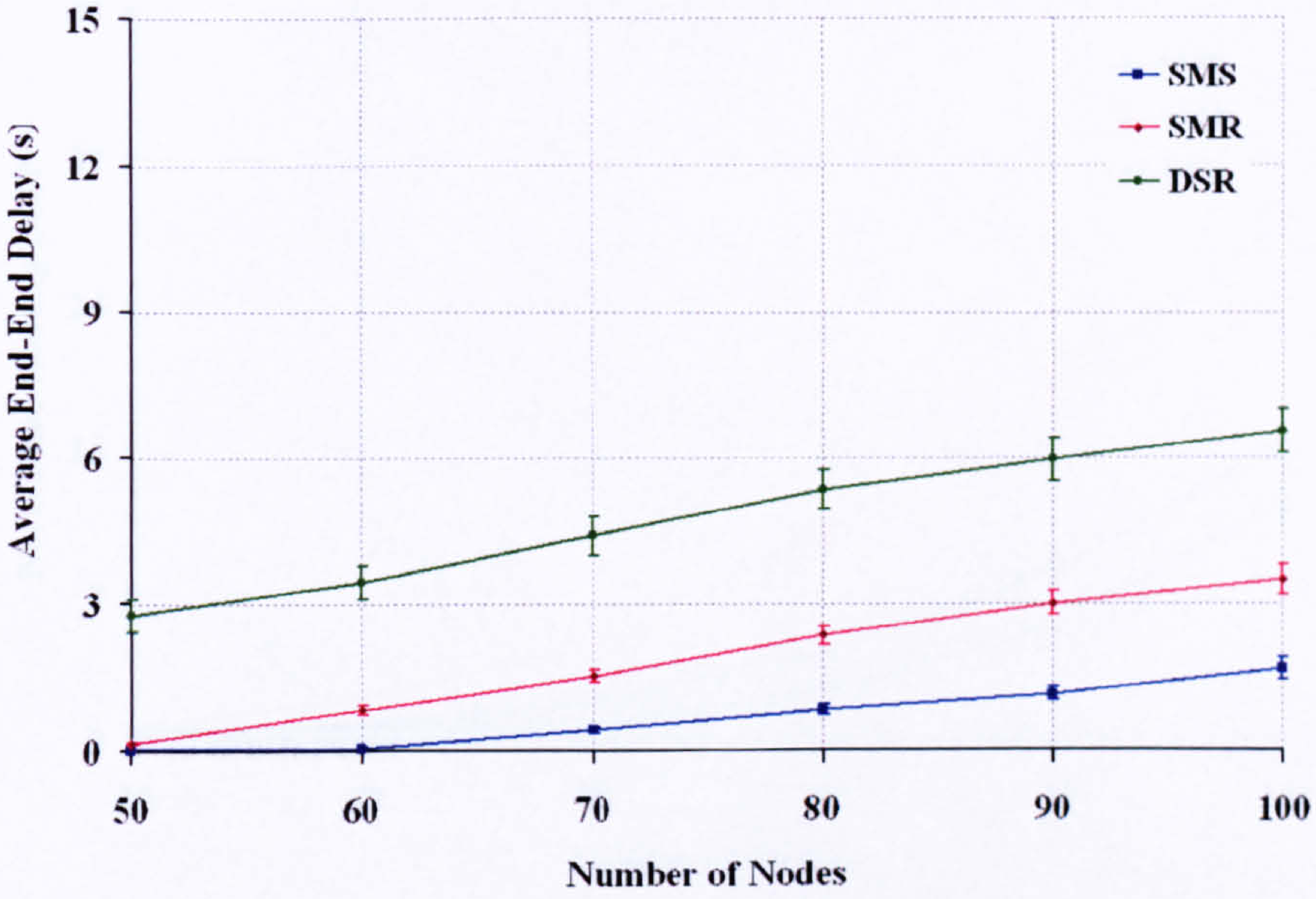


Figure 4.20 Average end-to-end delay versus number of nodes

4.6.3.3 Normalised Routing Load

Figure 4.21 depicts normalised routing load as a function of network size. The high availability of secondary paths in the case of SMS to deliver every single data packet to destinations reduces its routing load when compared to SMR. Such a reduction in load allows SMS to scale better than SMR. SMR has higher normalised routing load due to the higher routing overheads and lower availability of disjoint paths. As mentioned in Section 4.6.2.3, the routing load savings for DSR derive from a reduced number of RREQs. In the highly mobile scenarios with large numbers of nodes, a significant amount of information is stored in node caches, but it can be very difficult to profit from such knowledge. Although nodes appear to have a significant amount of information, frequent changes in topology, as would be caused by nodes moving around would render the stored information redundant. The apparent availability of alternative routes would be compromised by the fact that the stored information is out of date and the indicated routes may no longer be available. The major obstacle to scalability within source routing protocols is high levels of control overhead that result from large topologies, high mobility and frequent route maintenance.

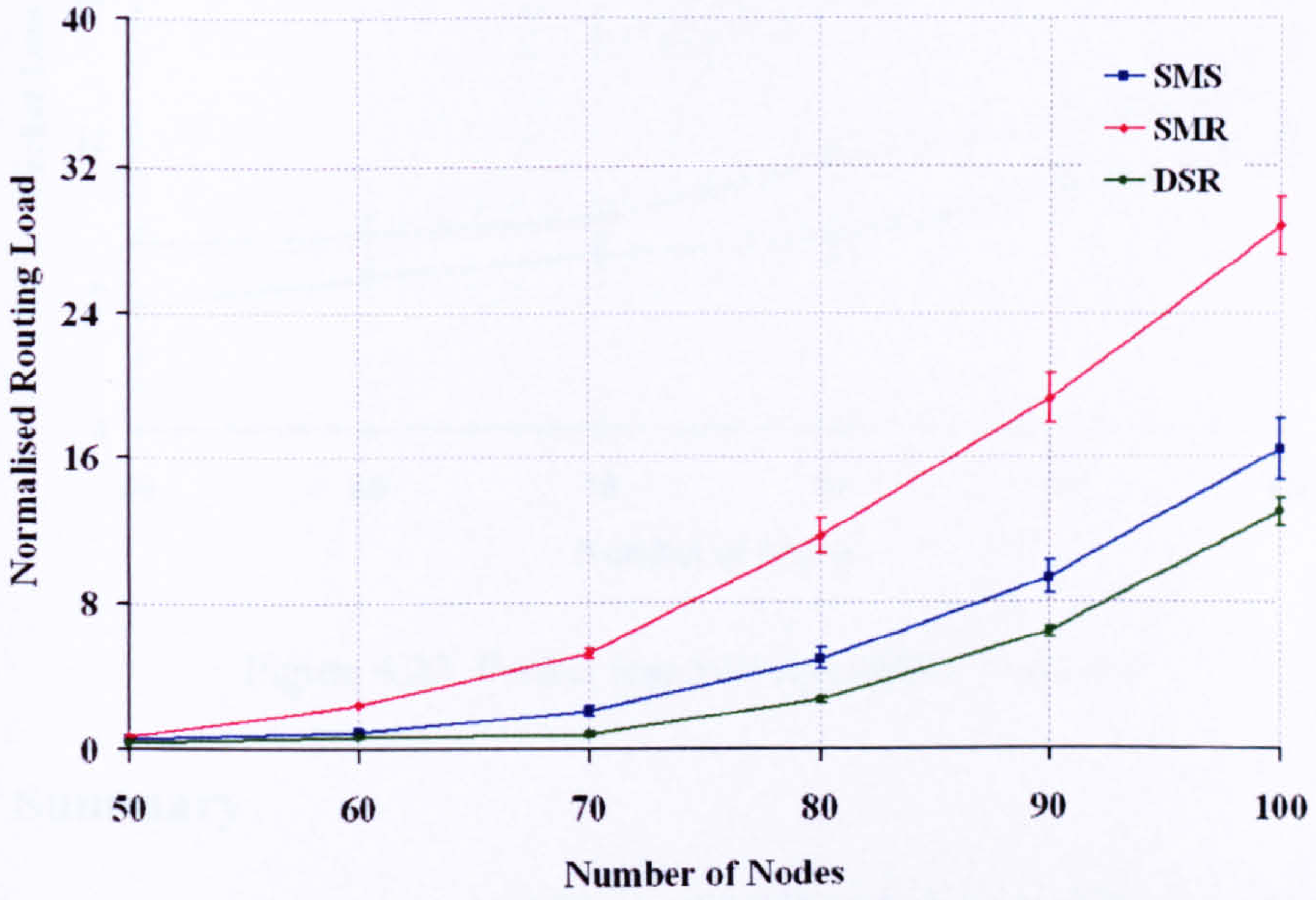


Figure 4.21 Normalised routing load versus number of nodes

4.6.3.4 Packet Loss

Figure 4.22 shows the relationship between the network size and the packet loss of the schemes indicating the degree of reliability of each protocol. Packet loss of the schemes increases with network size due to wireless link transmission errors, mobility and congestion. Packet loss due to transmission errors is affected by the physical condition of the channel, the terrain where networks are deployed, etc. A packet may be dropped at the source if a route to the destination is not available or if the buffer that stores pending packets is full. It may also be dropped at an intermediate host if the link to the next hop has broken. As discussed in Section 4.6.1.4 and 4.6.2.4, SMS performs well for all network sizes as a consequence of shorter multiple partial-disjoint paths and rapid recovery from route breaks.

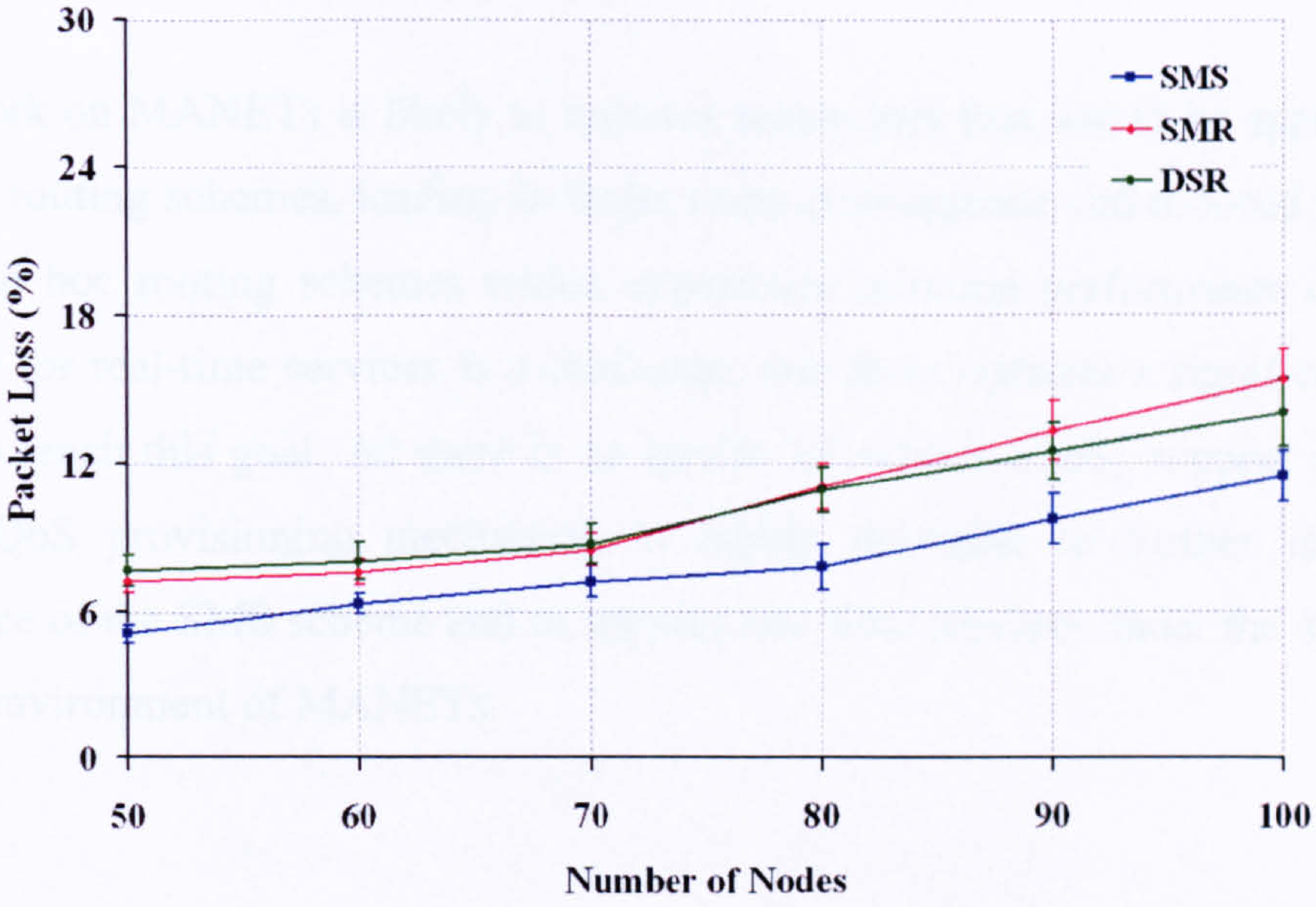


Figure 4.22 Packet loss versus number of nodes

4.7 Summary

In this chapter, a novel and practical Shortest Multipath Source (SMS) routing scheme is proposed for mobile ad hoc networks. The main aim of such a mechanism is to build multiple partial-disjoint paths from source to destination in order to reduce delays, overhead of additional route discoveries and recover quickly in case of route breaks. This scheme addresses the problem of wireless broadcast storm by simple hop count mechanism. The SMS model is implemented within the NS-2 simulator environment and SMS performance is compared with SMR, which uses node-disjoint secondary paths for correcting route breaks. The results are also compared with DSR because it is important to determine whether the multipath schemes provide a better solution than its unipath counterpart. SMS outperforms SMR due to usage of partial-disjoint multiple paths, provides more availability than node-disjoint multiple paths. Performance evaluation of the schemes shows that SMS performs better than both DSR and SMR when considering a variety of metrics. With an increase in mobility, offered load and network size, the performance of SMS as expected also begins to degrade but shows superior performance when compared to competing and directly compatible schemes. It is evident from simulation results that SMS is a candidate that offers improved routing performance, making the scheme particularly suitable for real-time services delivery.

Further work on MANETs is likely to uncover techniques that would be appropriate for stabilizing routing schemes, leading to faster route convergence and reduced route flaps. Creating ad hoc routing schemes which experience minimal performance degradation when used for real-time services is a challenge, and there remains a significant amount of work to reach this goal. As there is no quality of service (QoS) support in SMS, an effective QoS provisioning mechanism is highly desirable to further improve the performance of the SMS scheme and to support real-time services under the dynamically changing environment of MANETs.

5. QoS-aware Routing Solution Survey

5.1 Introduction

The demand for supporting real-time applications such as voice telephony, video-on-demand and video conferencing over MANETs has triggered a spurt of research on how to satisfy the quality of service (QoS) requirements of these applications in terms of bandwidth or more strictly speaking capacity, delay, jitter, packet loss and reliability. One of the key issues in providing QoS guarantees is how to determine paths that satisfy QoS constraints; solving this problem is known as QoS-aware routing or constrained-based routing. Although difficult, it is relevant and challenging to design and develop QoS-aware routing techniques for MANETs.

This chapter provides a survey of the state of the art in this area. Several important research issues and open questions need to be addressed to facilitate QoS support in MANETs. Resource estimation (commonly capacity), route discovery, route maintenance and feasible route selection are some of the issues that are currently being examined and require further exploration. Effective and efficient solutions to these issues will facilitate the design and development of QoS support in MANETs.

After a brief introduction, this chapter first reviews the state of art in the field of QoS-aware routing in section 5.2. This is followed by section 5.3 which provides research motivation by identifying key design considerations for a QoS-aware routing algorithm. Concluding remarks are made in section 5.4.

5.2 MANET QoS-aware Routing Approaches Survey

As mentioned before, most MANET routing schemes support only best-effort traffic. QoS-aware routing schemes play a vital role in QoS mechanisms in general, since it is their task to find a route with sufficient available resources to satisfy the QoS constraints. There are many ways in which the QoS-aware routing approaches can be categorised according to their various characteristics. QoS approaches can be classified [15] based on the:

- interaction between the routing scheme and the QoS provisioning mechanism (coupled or decoupled),
- interaction with the MAC layer, either dependent or independent and
- routing information update mechanism (proactive, reactive or hybrid).

The majority of the QoS-aware routing solutions proposed in the literature until now have focused on providing QoS based on two metrics: capacity (throughput) and delay [105][106][107][108]. Of these, the most common is capacity. This is possibly because most real-time applications require some degree of guaranteed throughput in addition to their other constraints. Figure 5.1 illustrates the classification of selected QoS approaches based on interaction between the routing scheme and the QoS provisioning mechanism by considering three classes of QoS solutions:

- *QoS Solutions Decoupled from the Routing Scheme – QoS Frameworks:* a complete system attempts to provide required services to each user or application.
- *QoS Solutions Coupled with the Routing Scheme – Schemes based on Contention-free MAC:* rely on resource availability (commonly capacity and delay) and resource reservation and therefore require a contention-free MAC protocol such as TDMA and CDMA.
- *QoS Solutions Coupled with the Routing Scheme – Schemes based on Contended MAC:* rely solely on the available resources or achievable performance to be

statistically estimated (commonly residual capacity) and therefore require a contention-based MAC protocol such as IEEE 802.11.

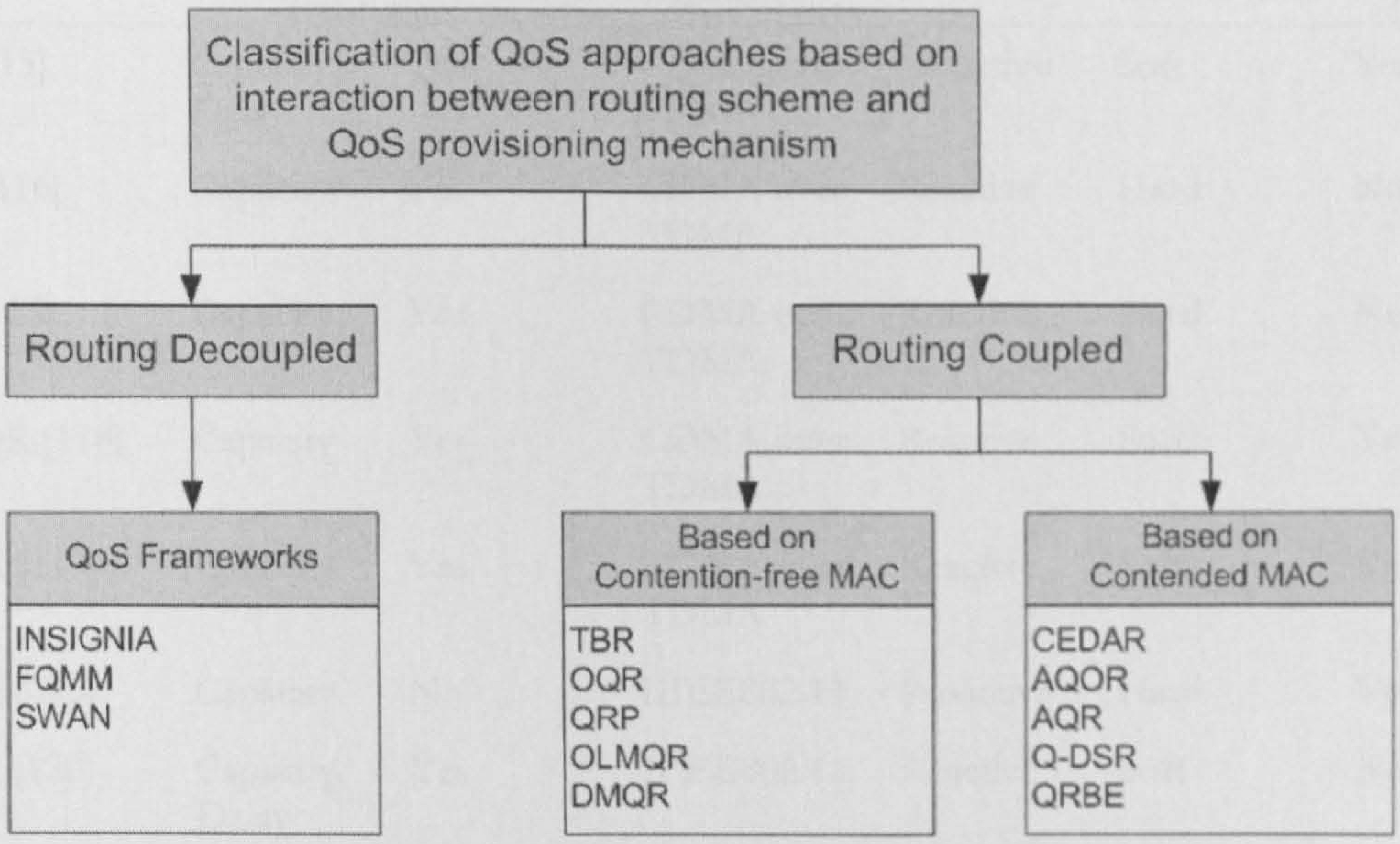


Figure 5.1 Classification of MANET QoS approaches

Table 5.1 and Table 5.2 summarise the salient features of the selected QoS solutions, whose operation, advantages and limitations are briefly described in the next section.

Table 5.1 Summary of QoS Solutions Decoupled from the Routing Scheme

QoS Scheme	QoS metric	Capacity estimation	MAC dependency	Route discovery	Resource reservation	Multiple paths
INSIGNIA [109]	Capacity	Yes	No	Proactive/ Reactive	Hard	No
FQMM [111]	Capacity	Yes	No	Proactive/ Reactive	Hard	No
SWAN [114]	Delay	No	No	Proactive/ Reactive	Soft	No

Table 5.2 Summary of QoS Solutions Coupled with the Routing Scheme

QoS Scheme	QoS metric	Capacity estimation	MAC dependency	Route discovery	Resource reservation	Multiple paths
TBR [115]	Capacity, Delay	Yes	CDMA over TDMA	Proactive	Soft	Yes
OQR [116]	Capacity	Yes	CDMA over TDMA	Reactive	Hard	No
QRP [117]	Capacity	Yes	CDMA over TDMA	Reactive	Hard	No
OLMQR [118]	Capacity	Yes	CDMA over TDMA	Reactive	Soft	Yes
DMQR [119]	Capacity	Yes	CDMA over TDMA	Reactive	Soft	Yes
CEDAR [46]	Capacity	No	IEEE802.11	Proactive	Hard	Yes
AQOR [120]	Capacity, Delay	Yes	IEEE802.11	Reactive	Soft	No
AQR [121]	Capacity, Delay	Yes	IEEE802.11	Reactive	Soft	No
Q-DSR [122]	Capacity	Yes	IEEE802.11	Reactive	Soft	No
QRBE [123]	Capacity	Yes	IEEE802.11	Reactive	Soft	No

5.2.1 QoS Frameworks

5.2.1.1 In-band Signalling for QoS in Ad hoc mobile networks (INSIGNIA)

In-band Signalling for QoS in Ad hoc mobile networks (INSIGNIA) framework [109] is the first scheme designed for MANETs, where resources are reserved in an end-to-end manner using a resource reservation protocol (RSVP)-like signalling mechanism [110]. This QoS framework is specifically designed to support adaptive services in ad hoc networks. It allows packets of real-time applications to specify their maximum and minimum capacity needs, and supports resource allocation, restoration control and session adaptation between communicating mobile hosts. Based on the residual capacity[16], QoS mechanisms attempt to provide assurances in support of adaptive services. To support an adaptive service, the INSIGNIA framework establishes and maintains reservations for continuous media flows and micro-flows. To support these communication services, the INSIGNIA QoS framework is composed of in-band signalling, admission control, packet forwarding, routing protocol, packet scheduling and medium access control modules as illustrated in Figure 5.2 [109].

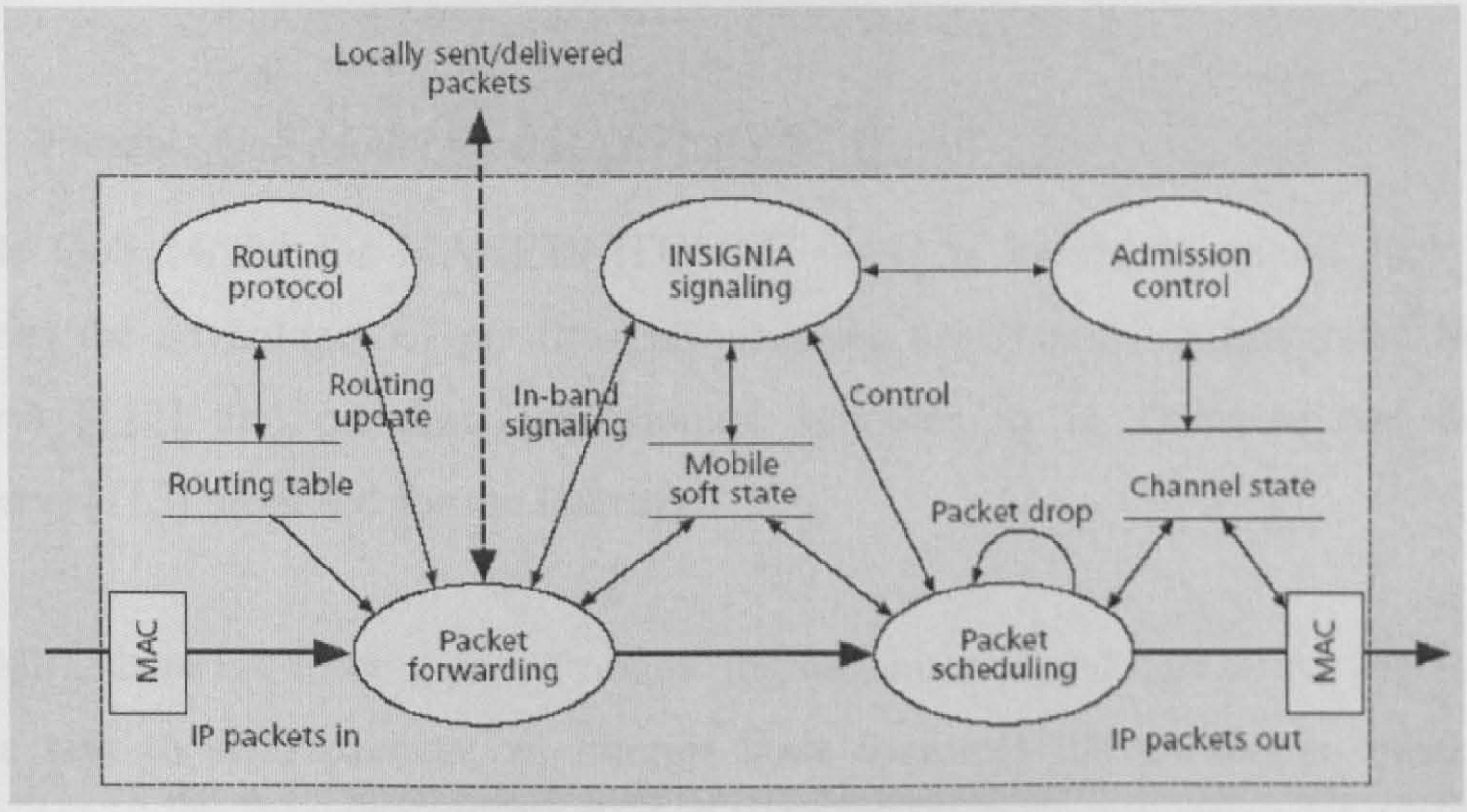


Figure 5.2 INSIGNIA QoS framework [109]

A major element of this QoS framework is the INSIGNIA signalling system that supports fast reservation, restoration and adaptation algorithms that are specifically designed to deliver adaptive service. The admission control module is responsible for allocating capacity to flows based on the maximum and minimum capacity requested. Once resources have been allocated they are periodically refreshed by a mobile soft-state mechanism via the reception of data packets. The packet-forwarding module classifies incoming packets and forwards them to the relevant module (viz. MANET routing, in-band signalling, wireless packet scheduling modules).

---

[16] The residual capacity is often termed available bandwidth in the literature.

INSIGNIA is not a routing protocol but an effective signalling protocol since it is an in-band signalling protocol and the allocation of resources is 'soft-state'. Although INSIGNIA can work in static MANETs, it involves a high signalling cost in more dynamic MANETs. As a result, for highly dynamic ad hoc networks with time varying topology and link capacities, the cost of connection establishment and maintenance would be significant. Hence, INSIGNIA is not particularly suitable for real-time applications that have stringent QoS requirements.

5.2.1.2 Flexible QoS Model for MANETs (FQMM)

Flexible QoS Model for MANETs (FQMM) [111] is hybrid in nature such that it combines the advantages of per-flow provisioning schemes as in Integrated Services (IntServ) [112] and per-class provisioning schemes as in Differentiated Services (DiffServ) [113] proposed for the Internet.

In FQMM, there are three types of nodes: ingress, interior and egress. An ingress node has the task to send packets; an interior node forwards the packets to other nodes whereas the egress node is a destination node. The role of the node changes dynamically based on its position and network traffic. An ingress node is responsible for such processes as classification, metering and marking of its own traffic. The interior nodes perform traffic shaping according to those marks. FQMM model tries to improve the per-class granularity of DiffServ to per-flow granularity for certain classes of traffic. Accordingly, high-priority traffic is given per-flow provisioning, while other lower priority traffic is given per-class provisioning. However, per-flow granularity is preserved for a small portion of traffic. The FQMM model suggests a relative and adaptive differentiation traffic profile in order to maintain consistent differentiation among sessions. Since it is deemed that an absolute traffic profile is not possible due to the inadequate capacity availability, FQMM proposes a traffic profile being defined as the relative percentage of the effective link capacity in order to keep the differentiation among sessions predictable and consistent. Hence, a token bucket metering algorithm is used to allow packets to be marked as in-profile and out-of-profile. In case of network congestion, out-of-profile packets are discarded with a higher probability than in-profile packets.

The FQMM further argues that best-effort routing is not sufficient and, hence, additional constraints need to be imposed on routing. However, it needs to be ensured that the constraints to be imposed on the routing protocols should be consistent with the provisioning policy. For example, per-class provisioning requires that all routers along the determined path make sure that traffic of a particular class injected into a given route should not be greater than the total percentage of capacity assigned in the traffic profile.

A drawback of this approach is that the source nodes have to take great care in regulating their traffic, since the rate of in-profile traffic must be processable in all network regions, including bottleneck areas where traffic from different sources accumulates. FQMM also lacks clear explanation for various aspects, for instance, how the differentiation among classes or flows is made predictable and consistent? How the source nodes should determine parameters for the token bucket metering?

5.2.1.3 Stateless Wireless Ad hoc Networks (SWAN)

Stateless Wireless Ad hoc Networks (SWAN) model [114] uses distributed control algorithms to support soft real-time services and service differentiation in MANETs. Figure 5.3 presents the SWAN architecture which consists of a packet classifier, an admission controller and a rate controller. The classifier distinguishes traffic into real time UDP traffic and best-effort TCP traffic, forcing the shaper to process best-effort traffic but not real-time traffic. It uses admission control for real-time traffic, rate control of TCP traffic and Explicit Congestion Notification (ECN) to ensure that all real-time packets meet QoS bounds. In addition, SWAN uses an 'Additive Increase Multiplicative Decrease' (AIMD) rate control mechanism to improve the performance of real-time

UDP traffic. Unlike TCP that uses packet loss as feedback to avoid network congestion, SWAN uses MAC delay as a feedback to local rate controllers.

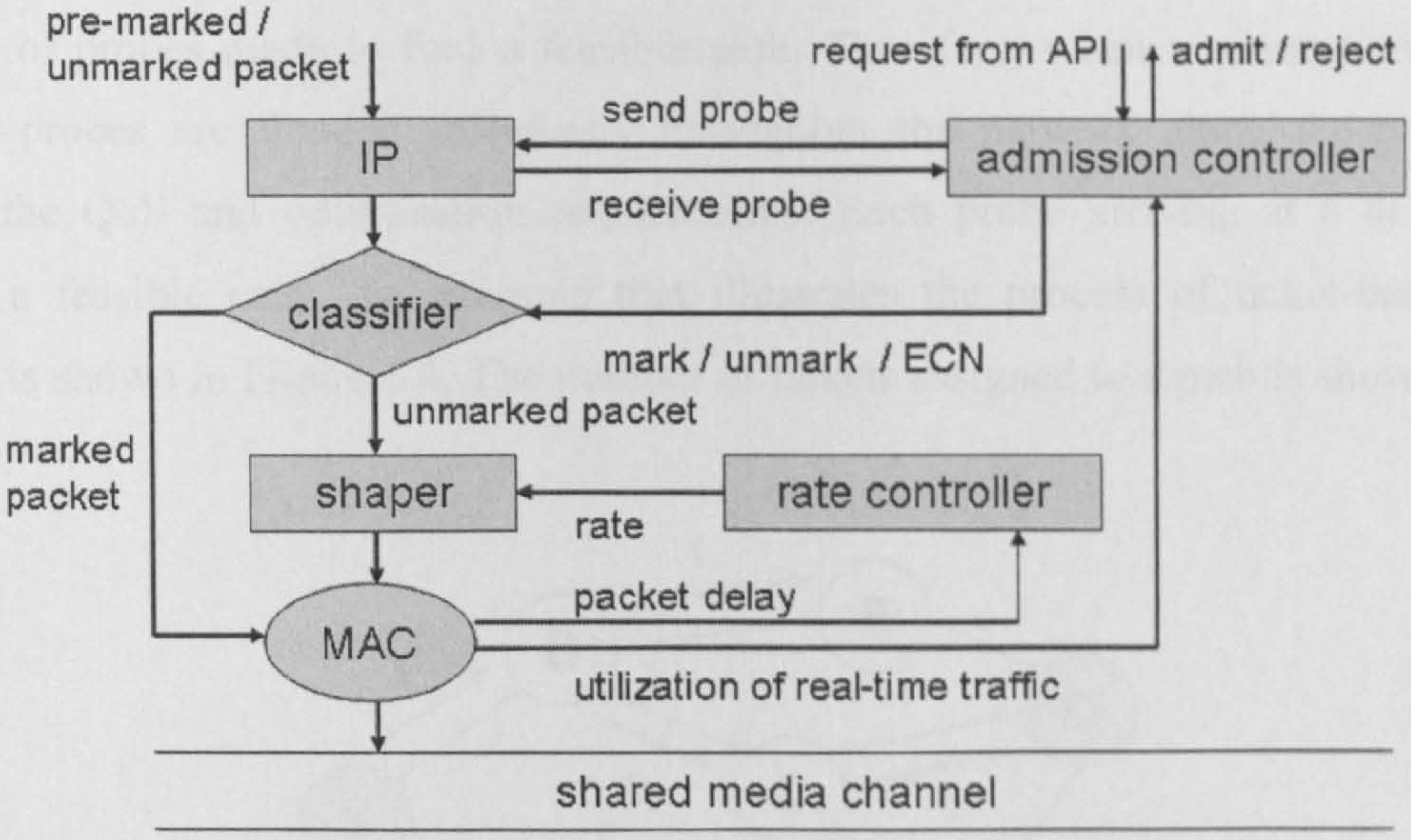


Figure 5.3 Key Components of the SWAN Architecture [114]

The drawback of the SWAN approach is that it can only provide weak service guarantees. ECN based traffic control is ineffective in dynamic MANETs. Although SWAN uses stateless approach, intermediate nodes may be required to remember whether the flows that traverse them are new or old in order to regulate traffic [114]. Furthermore, source-based admission control using probing-packets is again impractical and ineffective in a dynamic environment of MANETs. Also, capacity calculations may lead to a false estimation of the residual capacity, as SWAN does not take best-effort traffic into consideration.

5.2.2 QoS-aware Routing Schemes based on Contention-free MAC

5.2.2.1 Ticket-Based QoS Routing (TBR)

Ticket-Based QoS Routing (TBR) [115] is a distributed multipath QoS routing scheme for MANETs. Each node maintains end-to-end state information as capacity, delay and cost for every possible destination through the use of tickets. Two types of tickets are defined to be used during route discovery known as: yellow and green tickets. Yellow tickets are used for finding a feasible path with certain capacity/delay constraints. Green tickets are used for determining low cost paths. The number of tickets indicates the number of probes made to find a feasible path. Therefore, when a connection request arrives, probes are flooded selectively throughout the network along the paths that satisfy the QoS and optimisation requirements. Each probe arriving at a destination detects a feasible path. An example that illustrates the process of ticket-based QoS routing is shown in Figure 5.4. The number of tickets assigned to a path is shown by the number.

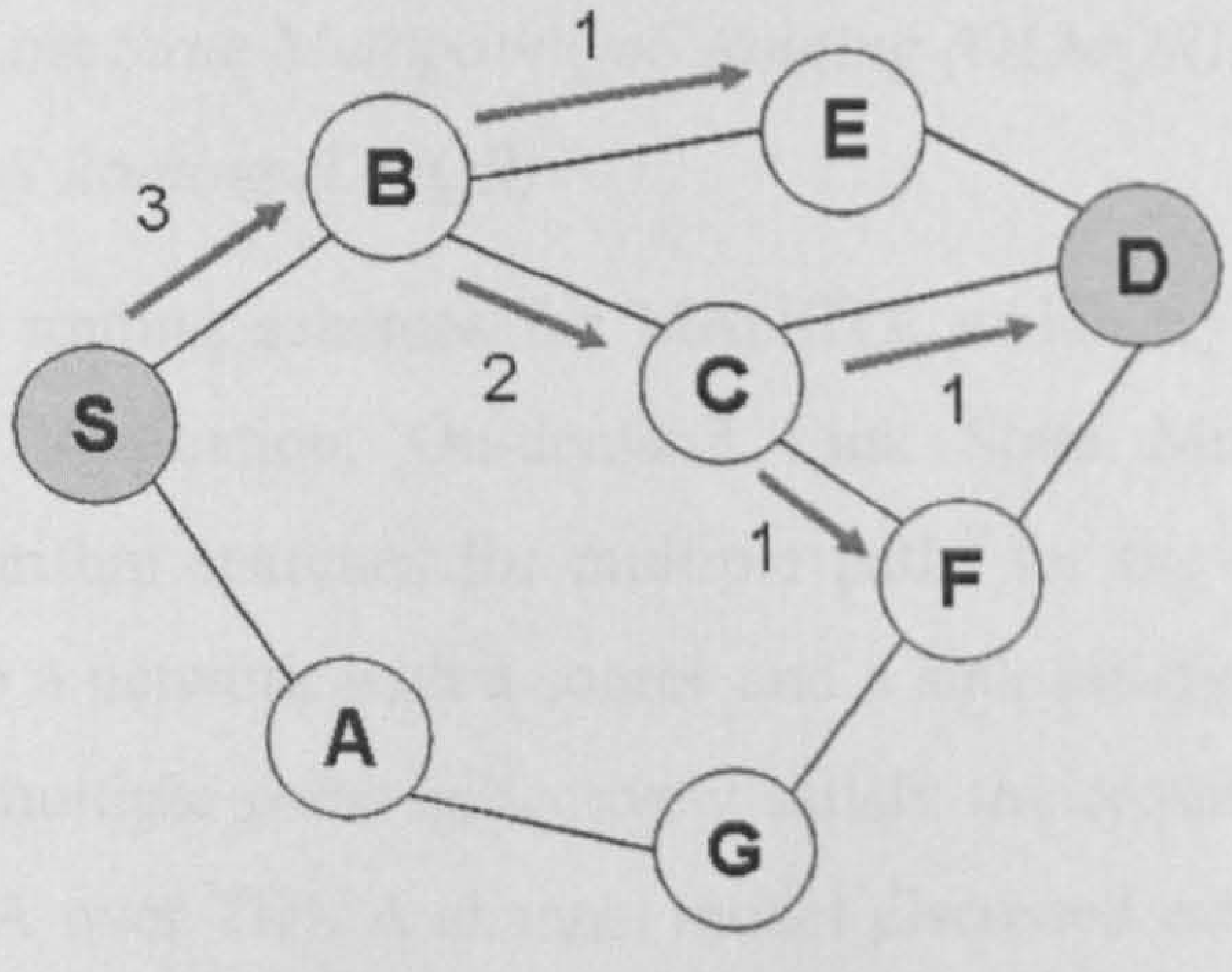


Figure 5.4 Ticket-based QoS routing

The path establishment process, restoration process in case of link-failures and the need to maintain state information with a use of distance-vector protocol lead to a high signalling cost and a subsequent effect upon routing performance. In addition, this work considers only the type of ad hoc networks whose topologies are not changing drastically and unpredictably and hence the proposed mechanism is more applicable to semi-stationary ad hoc networks. Another unrealistic feature of this scheme is its assumption that the underlying MAC is contention-free.

5.2.2.2 On-demand QoS Routing Schemes (OQR and QRP)

A feasible path selection based upon route discovery mechanism similar to that applied in AODV coupled with virtual circuit establishment using slotted channels is the On-demand QoS Routing (OQR) scheme [116]. It uses a combination of time-division multiple access (TDMA) and code-division multiple access (CDMA) as the underlying MAC and, hence, requires accurate network-wide slot-synchronisation and conflict-free code assignment – both tasks are extremely difficult in a highly volatile MANETs. OQR does consider residual capacity in the route selection process. A similar scheme (QRP) is also presented by C. Zhu et al [117]. Since both schemes heavily depend on such processes as flooding, slot synchronisation and code-assignment, their applicability in large-scale highly volatile MANETs is questionable.

5.2.2.3 On-demand Link State Multipath QoS Routing (OLMQR) and Disjoint Multipath QoS Routing (DMQR)

Unlike other existing routing schemes for MANETs, which try to find a single path between source and destination, On-demand Link State Multipath QoS Routing (OLMQR) [118] algorithm searches for multiple paths for the QoS route, where the multiple paths refer to a network with a source and a sink satisfying particular capacity requirements. These multiple paths collectively satisfy the required QoS. This scheme also adopts the CDMA over TDMA channel model discussed earlier to find routes that satisfies the QoS in terms of capacity specified by the source. The destination takes as many route requests it can and perform calculations to find the best multipath that satisfies the capacity back to the source. A similar algorithm, Disjoint Multipath QoS Routing (DMQR) [119] searches multiple disjoint paths, which can offer most aggregate resources to meet the QoS requirement of a call. DMQR uses Largest Bandwidth-Hop-Bandwidth First (LBHBF) method for choosing feasible paths.

The multipath QoS routing algorithms are suitable for ad hoc networks with limited capacity where single paths satisfying the QoS requirements are unlikely to exist. However, they inherit disadvantages similar to CDMA over TDMA supported QoS routing.

5.2.3 QoS-aware Routing Schemes based on Contended MAC

5.2.3.1 Core Extraction Distributed Ad hoc Routing (CEDAR)

Core Extraction Distributed Ad hoc Routing (CEDAR) [46] considers QoS routing to some extent, but in the main concentrates on backbone issues. The basic objective of the CEDAR mechanism is to construct a dynamically organisable virtual backbone infrastructure called core or spine for performing route computations and topology management in MANETs. These core nodes (also called as dominator nodes) may then function as route servers for their dominated nodes as shown in Figure 5.5. The fundamental mechanisms behind CEDAR are known as core extraction, link state propagation and QoS route computation.

- *Core extraction:* A group of nodes is dynamically elected to form the core of the network which maintains local topology information and perform route computations.
- *Link state propagation:* The capacity availability information of stable high capacity links is propagated to core nodes, while information regarding low capacity and unstable links remains local.
- *QoS route computation:* A core path is established first from the dominator (neighbouring core node) of source to the dominator of destination. Using up-to-date local topology, the dominator of source finds a path satisfying the requested QoS from source to furthest possible core node. This furthest core node then becomes the source of next iteration. The above process repeats until destination is reached or the computation fails to find a suitable path.

Although this work present the idea of maintaining a virtual backbone dedicated for control plane tasks in MANETs, the sheer amount of updates (in the form of flooding) they need in order to maintain a global topology and state information makes these approaches undesirable, particularly in the light of nodes' random mobility patterns. It is, however, not clear how adaptive the virtual backbone mechanism is to node-mobility and whether the virtual backbone formation approach takes mobility into consideration at all. In CEDAR, nodes that have higher connectivity with other nodes are likely to be elected as dominator nodes and, since this criterion varies mainly due to node mobility, this may trigger regular and frequent core extraction and, thus, the relevant control traffic will be significant.

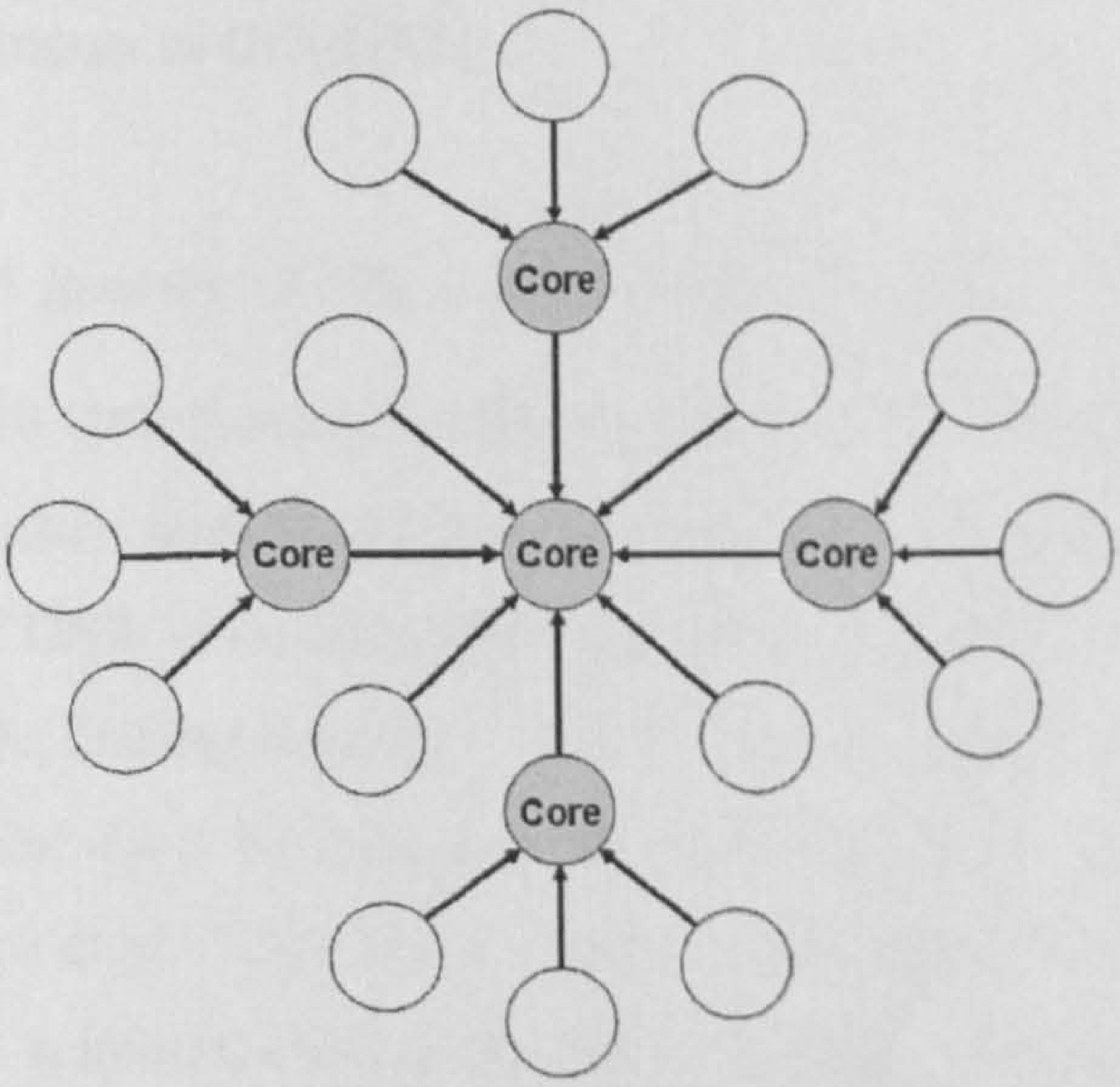


Figure 5.5 Core network found by CEDAR

5.2.3.2 Ad hoc QoS On-demand Routing (AQOR)

Ad hoc QoS On-demand Routing (AQOR) [120] is a resource reservation and signalling algorithm. AQOR is an on-demand QoS-aware routing scheme that provides mechanisms for residual capacity and end-to-end delay measurement, resource reservation and adaptive route recovery. AQOR propose computation algorithms to estimate residual capacity and end-to-end delay in an unsynchronised wireless environment. The residual capacity estimation is based on the aggregate traffic of the neighbourhood and is performed on the MAC Layer. AQOR uses periodic Hello messages to keep an updated view of the neighbourhood. It reserves capacity on each node along a path that is being used by the source. The reservation has been done in the route discovery phase but does not actually take place until the first packet has been forwarded at a node. The adaptive route recovery procedure includes detection of broken links and route recovery at the destination which occurs when the destination node detects a QoS violation or a time-out of the destination's resource reservation. This procedure causes the destination to involve a reverse route exploration. In AQOR, the estimated total traffic at each node, which is the sum of a node's neighbours' traffic, can be larger than the real overall traffic. This overestimation leads to a 'stiff' capacity admission control threshold. Another limitation of this scheme is resource reservation which has similar problems to INSIGNIA.

5.2.3.3 Adaptive QoS Routing (AQR) and QoS-DSR (Q-DSR)

The proposed Adaptive QoS Routing (AQR) mechanism [121] is based on predicting the local performance of each node in ad hoc networks. The route discovery mechanism is very similar to that of DSR, although additional fields are added to the route request and reply packets to ensure finding feasible routes that satisfy capacity and end-to-end delay requirements. A similar QoS Routing scheme based on DSR (Q-DSR) mechanism is presented by L. Hanzo et al [122] which considers the throughput requirements of data sessions. Since these schemes utilises the basic routing functions of DSR, paths may break as a result of mobility and, hence, there would be a need to repair and reconstruct paths, leading to increased packet loss and latency.

5.2.3.4 QoS-aware Routing based on Bandwidth Estimation (QRBE)

QoS-aware Routing based on Bandwidth Estimation (QRBE) [123] scheme is built on AODV. The algorithm either admits a flow with the requested capacity (admission control scheme) or provides feedback about the residual capacity to the application (feedback scheme). Both the admission-control scheme and the feedback scheme require knowledge of the end-to-end capacity available along the route between source-destination pair. This proposed scheme inherits some of the drawbacks associated with AODV discussed in section 2.7.4: the use of flooding in the route discovery; the construction of end-to-end paths being ineffective in highly volatile MANETs unless augmented with accurate mobility predictions and route repair (reconstruction process) that incurs latency and leads to packet drops.

5.3 Open Issues

The previous sections highlighted a selection of schemes proposed to provide QoS support at the routing level. Several important research issues and open questions need to be addressed to facilitate QoS support in MANETs. Capacity estimation, route discovery, route maintenance and feasible path selection are some of the issues that require further exploration.

5.3.1 Capacity Estimation

There is still potential in the area of capacity-constrained QoS routing. To offer a capacity-guaranteed route, the key idea is to obtain information specifically about the residual capacity or achievable throughput from lower layers. This throughput information aids in facilitating admission control.

Some QoS-aware routing schemes, such as TBR and CEDAR, assume that the residual capacity is known. Most TDMA supported QoS routing schemes monitor and schedule free time slots for estimating residual capacity at the nodes. Other routing schemes, such as AQOR, AQR, Q-DSR and QRBE, exploit the carrier-sense capability of IEEE 802.11 to estimate the residual capacity. QRBE adopts two schemes for capacity estimation: 'Listen' capacity estimation and 'Hello' capacity estimation. The 'Listen' scheme requests each node to listen to the channel and estimate the residual capacity based on channel utilisation by using a moving window. However, the choice of window size is a major impediment to such window-based schemes. In the 'Hello' scheme, the 'Hello' packet includes the capacity consumptions of a node and its one-hop neighbours'. This scheme has significant extra overhead in 'Hello' packets.

Therefore, in MANETs, the two fundamental issues in capacity estimation are a) how exactly to estimate the residual capacity and b) how frequently to perform such estimations. Also, the trade-off between any advantage achieved from using capacity estimation and the cost in terms of packet overhead and computing resources used for capacity estimation is another vital issue.

5.3.2 Route Discovery

There are two main approaches to routing in MANETs: proactive routing and reactive routing. TBR and CEDAR are proactive approaches, while OQR, QRP, OLMQR, DMQR, AQOR, AQR, Q-DSR and QRBE are reactive approaches. Generally, reactive routing schemes perform better than proactive counterparts in term of overhead. However, route discovery using flooding in reactive routing schemes requires a large amount of delay. Therefore, incorporating limited flooding for route discovery, such as in the scheme used in Shortest Multipath Source (SMS) routing can improve the routing performance.

5.3.3 Route Maintenance

Node mobility in MANETs causes frequent topology changes in the network, making it difficult to acquire the QoS requirements. Incorporating a quick route maintenance scheme into QoS-aware routing is another challenging problem. Predicting route breaks allows better utilisation of capacity, since packets are not sent via a route that will be disconnected soon. Another useful method is to provide multiple paths, which offer alternative paths when the primary route is broken.

The typical approach to route maintenance in OQR, QRP, AQOR, AQR, Q-DSR and QRBE, which involves waiting for the host to discover a new route, significantly affects the routing performance. TBR, OLMQR and DMQR maintain secondary paths to use when the primary path fails. However, there is a trade-off between path redundancy and overhead and to balance path redundancy with overhead is an open issue.

5.3.4 Path Selection

QoS-aware routing has more rigorous requirements on route stability, since frequent route failures will adversely affect the end-to-end QoS. Thus, in some sense the path with the largest residual capacity (as used in QRBE) shall not be the only consideration but path reliability should also be considered when selecting a suitable route for a QoS-aware routing scheme.

5.4 Summary

This chapter presented the operation, strengths and limitations of the major contributions to the QoS-aware routing solutions published in the period 1997-2007. The schemes are selected in such a way as to highlight key approaches to QoS routing in MANETs. Several open issues have been pointed out such as residual capacity estimation, route discovery, route maintenance and path selection that need to be addressed in QoS-aware routing scheme. Moreover, to increase the user's perceived QoS, multipath routing techniques can be used to improve the previously proposed solutions. The next chapter presents an effective QoS-aware Shortest Multipath Source (Q-SMS) routing scheme based on residual capacity estimation which is an extension to the SMS routing scheme presented in chapter 4.

6. QoS-aware SMS (Q-SMS) Routing Scheme

6.1 Introduction

Routing schemes have prompted a great deal of interest from the beginning of MANET research until the present time. Early work focused on finding feasible routes without considering information about the network status. In addition, without knowing the bottleneck capacity or throughput, the source may send more data than the bottleneck node on the route can accommodate. The overloaded node ultimately drops data which wastes energy and unnecessarily consumes capacity. Also, time is expended in transmitting such data. Therefore, data that eventually reaches its destination would have had to wait longer in packet queues resulting in a significantly increased delay. Although this may be acceptable for data only applications, many real-time applications require QoS support from the network. Possible QoS support can be achieved by finding a route to satisfy the application requirements.

QoS-aware routing takes into consideration multiple QoS requirements, link dynamics, as well as the implication of the selected routes on network utilisation, rendering QoS routing a particularly challenging problem. However, the unique features of MANETs, namely dynamically varying network topology, imprecise state information, lack of central coordination, error-prone shared radio channel, hidden terminal problem and time-varying capacity exacerbate the already complex routing problem [15]. More importantly, node mobility causes frequent failure and reactivation of links, effecting a reaction from the network's routing to the changes in topology, thus increasing network control traffic and saturating the already congested links. Hence, all these aspects necessitate a cost-effective solution for any QoS-aware routing scheme.

Capacity estimation is a key component of any admission control scheme required to support QoS provision in MANETs. A number of schemes have been previously proposed to estimate residual capacity that is derived from window-based measurements of channel estimation. In this chapter, a new capacity-constrained QoS-aware SMS (Q-SMS) routing scheme is proposed. The novel part of this QoS-aware routing scheme is a simple additional mechanism to estimate residual capacity in IEEE 802.11-based ad hoc networks. The scheme further proposes the use of a 'forgiveness factor' to weight these previous measurements and, it is shown through simulation-based evaluation to provide improved utilisations estimation and better available capacity based admission control. The remaining chapter is organised as follows. Section 6.2 presents Q-SMS routing scheme for MANETs. Section 6.3 discusses the simulation results and compares performances between SMS and Q-SMS routing schemes. In Section 6.4, the forgiven capacity estimation scheme (Q-SMS-F) is described. Performance evaluation and comparison of Q-SMS-F and Q-SMS are presented in Section 6.5. Finally, Section 6.6 summarises the chapter.

6.2 QoS-aware SMS (Q-SMS) Routing Scheme

A novel and practical QoS routing scheme referred to as the QoS-aware Shortest Multipath Source (Q-SMS) routing scheme is proposed. The proposed scheme modifies and extends the route discovery and maintenance of SMS to provide QoS assurance. The QoS extension allows nodes to use their estimation of the residual capacity to make better admission control decisions. The Q-SMS routing scheme achieves high goodput and low delays and overheads in the presence of mobility and traffic load and enables natural integration with the local residual capacity estimation.

This section now describes the residual capacity estimation, route discovery with admission control, QoS route reply phase, QoS route maintenance phase and path selection of the proposed scheme.

6.2.1 Residual Capacity Estimation

Residual capacity estimation is a fundamental component in the provision of QoS in MANETs. The residual capacity of a link relates to the unused capacity of the link during a particular predetermined time period. So, even though the capacity of a link depends on the underlying transmission technology and propagation medium, the residual capacity of a link additionally depends on the traffic load at that link and is typically a time-varying metric.

However, accurate capacity estimation can be difficult; because each host has only imprecise knowledge of the network status; thus an effective estimation scheme is highly desirable. Many previously proposed schemes [122][123][124][125] adopt 'Listen'-based estimation techniques derived from IEEE 802.11 MAC/PHY specification [2]. The 'Listen' scheme requires each node to listen to the channel and estimate the local residual capacity based on the measurement of the local channel utilisation. Given the local channel utilisation ($u(t)$) and the maximum achievable channel capacity ($C_{\max}$), the local residual capacity (C_{res}) is estimated using the following equation:

$$C_{res} = (1 - u(t)).C_{\max} \tag{6.1}$$

where $0 \leq u(t) \leq 1$ is a measure of the channel utilisation. A simple and direct technique for determining channel utilisation is to measure channel busy time at nodes within its carrier sensing range. In previous studies [122][123][124][125], the model shown in Figure 6.1 was employed when defining the transmission range and the carrier sensing range. The following definitions for the transmission range and the carrier sensing range can be given [126]:

- The transmission range of a node is the range within which a transmitted frame can be successfully received. The transmission range is mainly determined by the transmission power and the radio propagation properties.

- The carrier sensing range of a transmitting node is the range within which the other nodes detect a transmission. It mainly depends on the sensitivity of the receiver and the radio propagation properties. Nodes inside the carrier sensing range of a node are termed its carrier sensing neighbours. The carrier sensing range is typically larger than the transmission range.

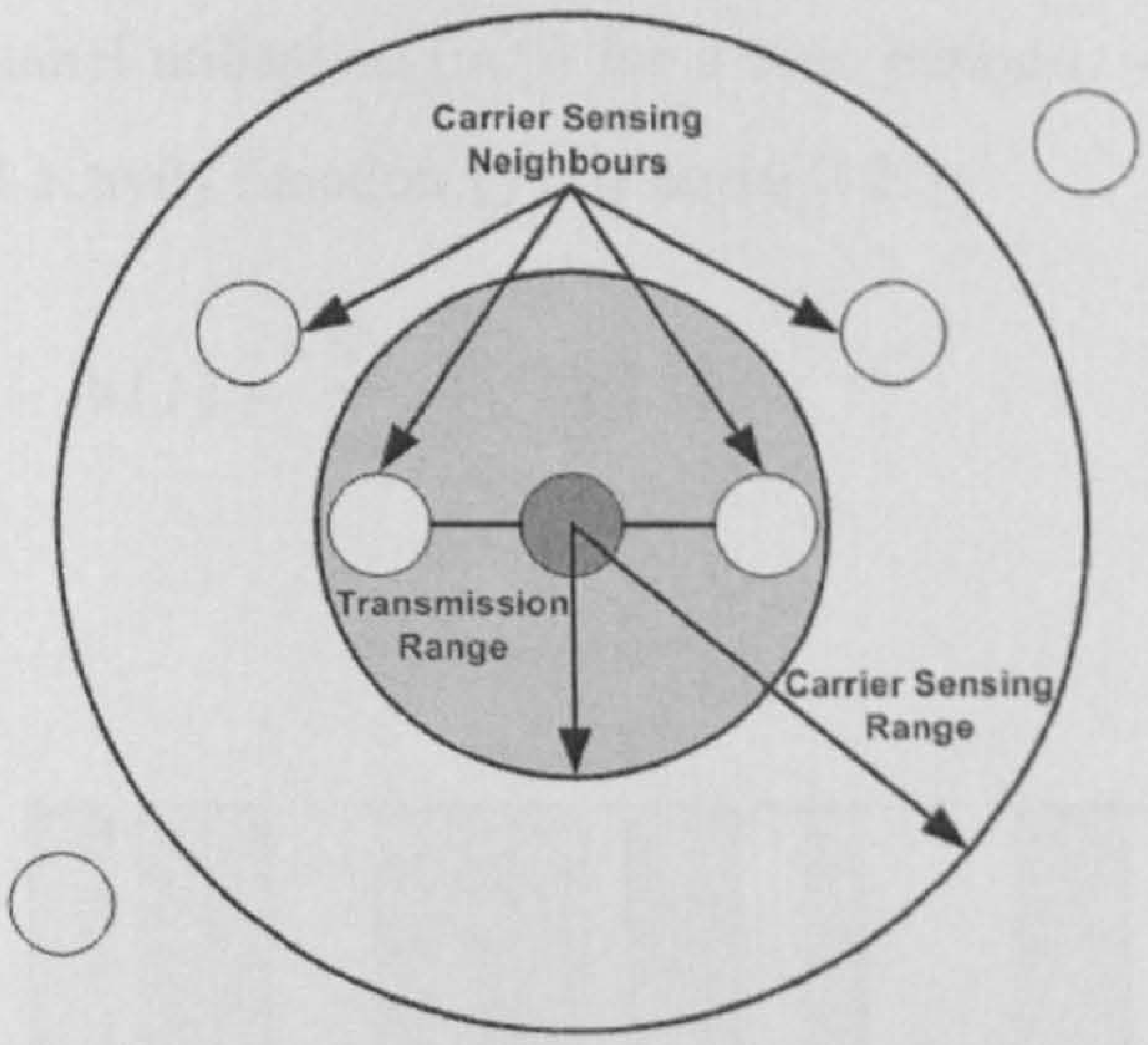


Figure 6.1 Wireless communication ranges

The channel can be in any of the following states:

- *noCSnoNAV:* No carrier is sensed and the Network Allocation Vector (NAV) is clear; channel is idle.
- *noCSNAV:* No carrier is sensed, but the network allocation vector is set; channel is busy.
- *CSnoNAV:* A carrier is sensed and the network allocation vector is clear; channel is busy.
- *CSNAV:* A carrier is sensed and the network allocation vector is set; channel is busy.
- *WIFS:* Waiting for Interframe Spacing (IFS) timer to expire; channel is busy. Essentially, IFS is a self enforced NAV.

All states except the *noCSnoNAV*, prohibit a node to send something in the channel and hence, in these states, the channel is considered as busy. A typical activity graph is shown in Figure 6.2. At any specific instant in time, a link is either transmitting a packet or it is idle, so the channel activity of a link can only be either 0 or 1. Thus, some meaningful measurement of the channel activity requires node to keep track of the busy channel periods over a time window (w) which is the time interval of interest. Consequently, the channel utilisation ($u(t)$) for a time period $(t-w,t)$ is given by the area under the channel activity function ($f(t)$) curve [127]:

$$u(t) = \frac{1}{w} \int_{t-w}^{t} f(t).dt \tag{6.2}$$

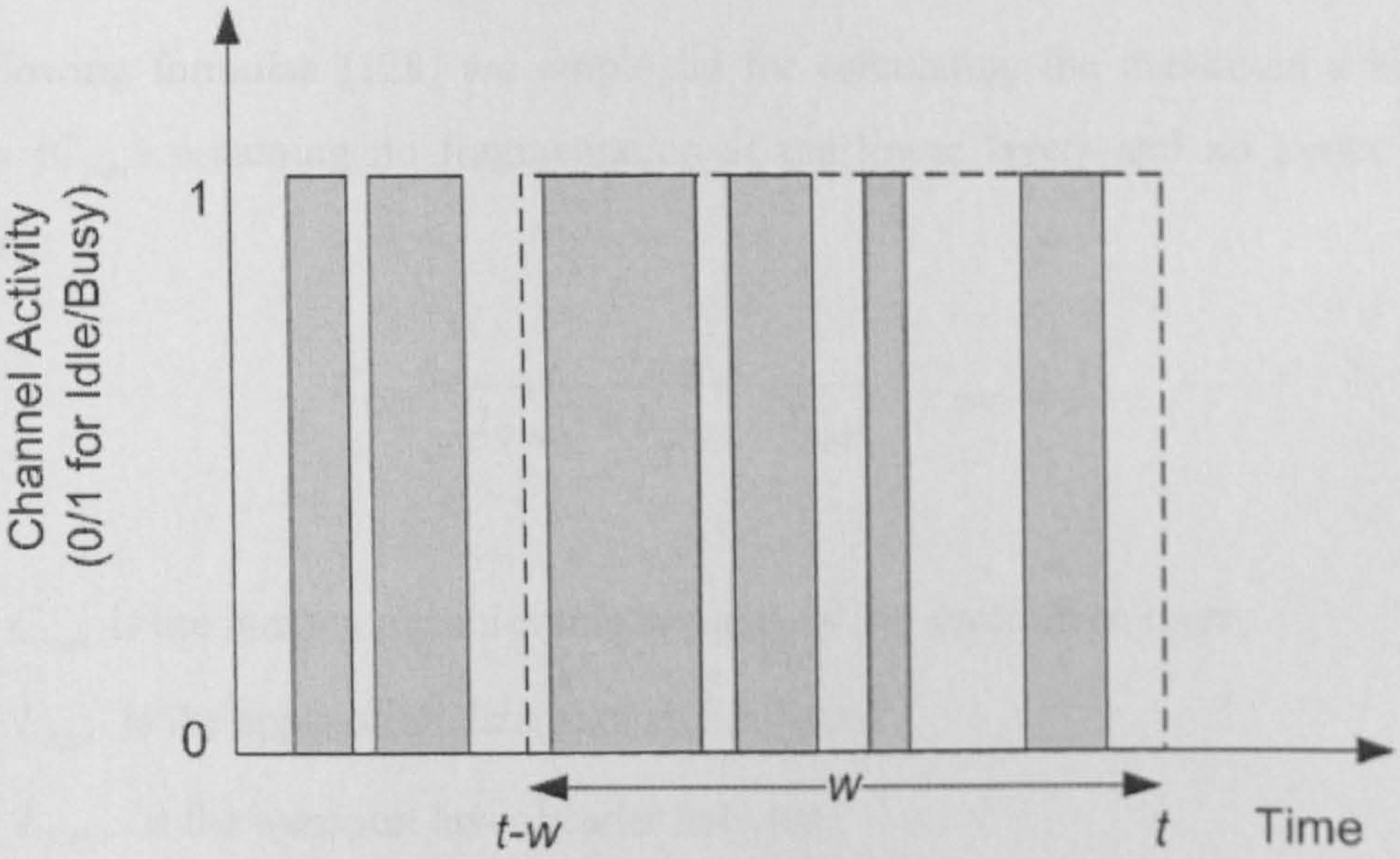


Figure 6.2 Example of the channel activity in an IEEE 802.11 network

6.2.1.1 *Maximum Capacity Calculation*

The maximum capacity that can be used for data transmission is significantly lower than the nominal bit rate due to the overheads associated with the transmission of each packet as shown in Figure 6.3.

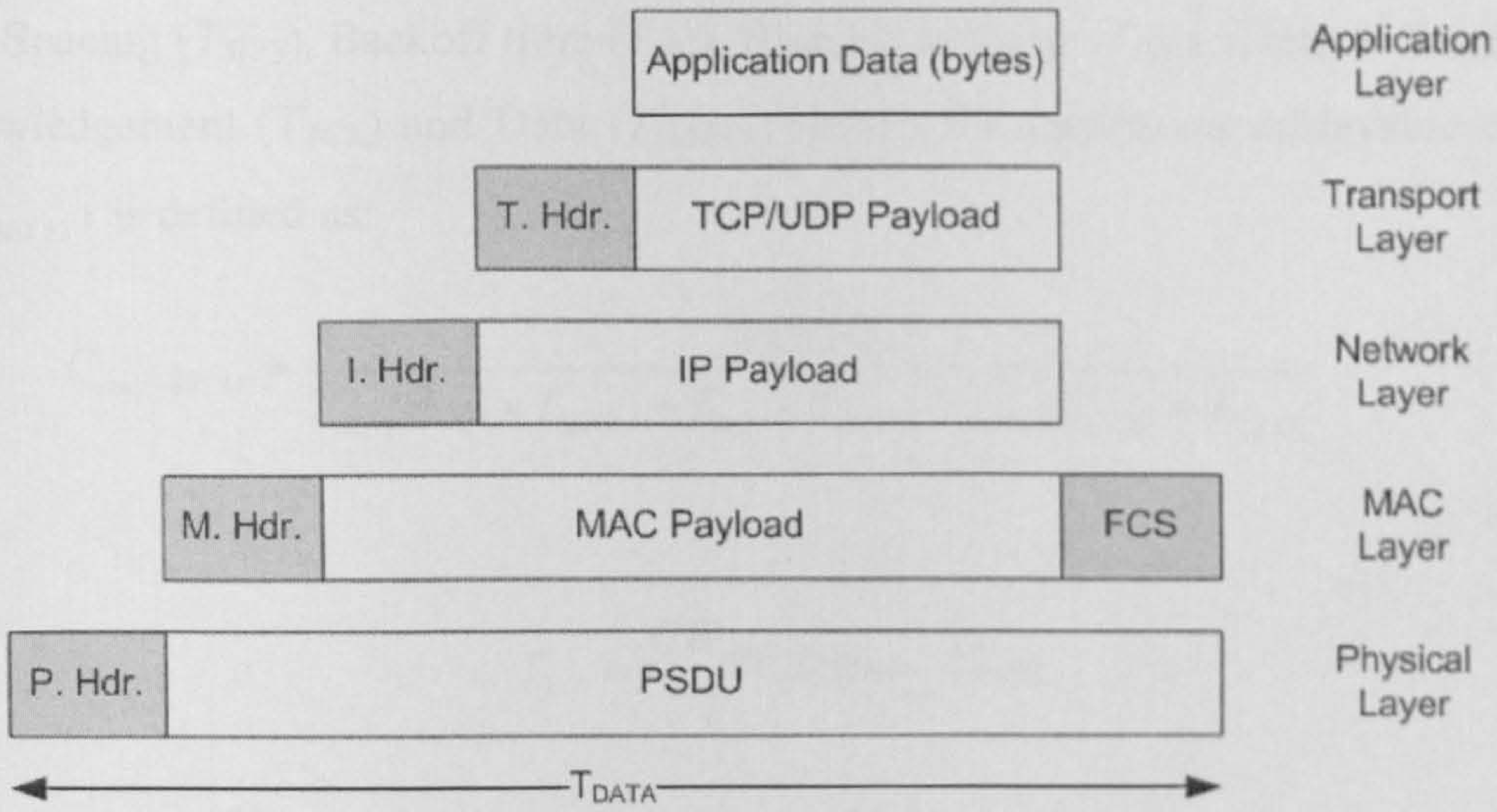


Figure 6.3 Encapsulation overheads

The following formulae [128] are employed for calculating the maximum achievable capacity ($C_{\max}$) assuming no fragmentation at the lower layers and no losses due to collisions:

$$C_{\max} = \frac{L_{Appl.}}{L_{T.Hdr.} + L_{I.Hdr.} + L_{Appl.}} \times C_{\max-802.11} \tag{6.3}$$

where,

$C_{\max}$ is the maximum achievable capacity of the application layer,

$L_{Appl.}$ is the application datagram size in bytes,

$L_{T.Hdr.}$ is the transport layer header in bytes,

$L_{I.Hdr.}$ is the IP layer header in bytes and

$C_{\max-802.11}$ is the maximum achievable capacity of the IEEE 802.11 MAC layer.

The maximum achievable capacity of the IEEE 802.11 MAC layer using RTS/CTS mechanism ($C_{\max-802.11}$) can be obtained by dividing the MAC Service Data Unit (*MSDU*) length by the total time taken to transmit the *MSDU*. The total delay is composed of time required for Distributed Inter Frame Spacing (T_{DIFS}), Short Inter

Frame Spacing (T_{SIFS}), Backoff time (T_{BO}), Request to Send (T_{RTS}), Clear to Send (T_{CTS}), Acknowledgement (T_{ACK}) and Data (T_{DATA}). Hence, the maximum achievable capacity ($C_{\max-802.11}$) is defined as:

$$C_{\max-802.11} = \frac{MSDU}{T_{DIFS} + (3 \times T_{SIFS}) + T_{BO} + T_{RTS} + T_{CTS} + T_{ACK} + T_{DATA}} \tag{6.4}$$

where,

$$T_{BO} = \frac{CW_{\min}}{2} \times Slot\_Time \tag{6.5}$$

and

$$T_{DATA} = \frac{P.Hdr.}{data\_rate_{PHY}} + \frac{M.Hdr. + MSDU}{data\_rate_{MAC}} \tag{6.6}$$

The data rate is not always the same even within the same physical layer protocol data unit (*P PDU*). The *P.Hdr.* is transmitted at 1 Mbps. The data rate of a *MAC-PDU* is determined by its type. Control frames such as RTS, CTS and ACK can be transmitted at 1 or 2 Mbps while data frames can be transmitted at 1, 2, 5.5 and 11 Mbps.

The contention window size (*CW*) does not increase exponentially since there are no collisions. Thus, *CW* is always equal to the minimum contention window size ($CW_{\min}$) and T_{BO} is selected randomly following a uniform distribution from (0, $CW_{\min}$) giving the expected value of $CW_{\min}/2$. Table 6.1 gives the IEEE 802.11 parameters values for Direct Sequence Spread Spectrum (DSSS-2) and High Rate Direct Sequence Spread Spectrum (HR-11) [2][126][128]. The maximum achievable capacity ($C_{\max}$) is computed by applying equations (6.3), (6.4), (6.5) and (6.6), and assuming 512 byte sized packets, the $C_{\max}$ of 2 Mbps system is just 1.47 Mbps ($C_{\max-802.11} = 1.56\text{Mbps}$) and the $C_{\max}$ of 11 Mbps system is just 4.24 Mbps ($C_{\max-802.11} = 4.52\text{Mbps}$). This maximum achievable capacity can be used, along with the measured utilisation and equation (6.1), to estimate the residual capacity for admission control.

Table 6.1 IEEE 802.11 parameter values

Parameter	Value
T_{DIFS}	50μs
T_{SIFS}	10μs
T_{RTS}	352μs
T_{CTS}	304μs
T_{ACK}	304μs
$CW_{\min}$	31
Slot_Time	20μs
P.Hdr.	24 bytes
M.Hdr.	34 bytes
MSDU	1500 bytes
Control bit rate	1 Mbps
Nominal data rate	2 Mbps and 11 Mbps

6.2.2 Route Discovery with Admission Control

Q-SMS is an on-demand QoS-aware routing scheme that utilises a cross-layer design. Therefore, the routing scheme depends on the application requirements. Q-SMS finds a route to the destination by flooding the network with a QoS route request (QRREQ). Q-SMS extends RREQ packet format of SMS with the capacity constraint. The capacity constraint consists of the required capacity (C_{req}) and minimum available capacity (C_{min}) representing maximum capacity of the application and minimum available capacity (bottleneck capacity) of an outgoing link useful for path selection. So QRREQ has the format:

$$QRREQ = RREQ_{SMS} \cup \{C_{req}, C_{\min}\}$$

The source node records C_{req} and compares with the local residual capacity (C_{res}) of the outgoing link (by receiving link layer feedback from IEEE 802.11). If C_{res} is higher than C_{req}, the source node records the value of C_{res} in the C_{min} field which is initially infinity (or a very large number) and broadcast the QRREQ packet to its neighbour nodes.

On receiving QRREQ packet, an intermediate node calculates its residual capacity C_{res}. If C_{res} is greater than C_{req}, the node forwards this QRREQ. Otherwise this QRREQ is discarded. The node also updates the C_{min} field if C_{res} is less than preceding C_{min}. The changes in SMS route discovery procedure to incorporate admission control based on capacity estimation are shown in Figure 6.4.

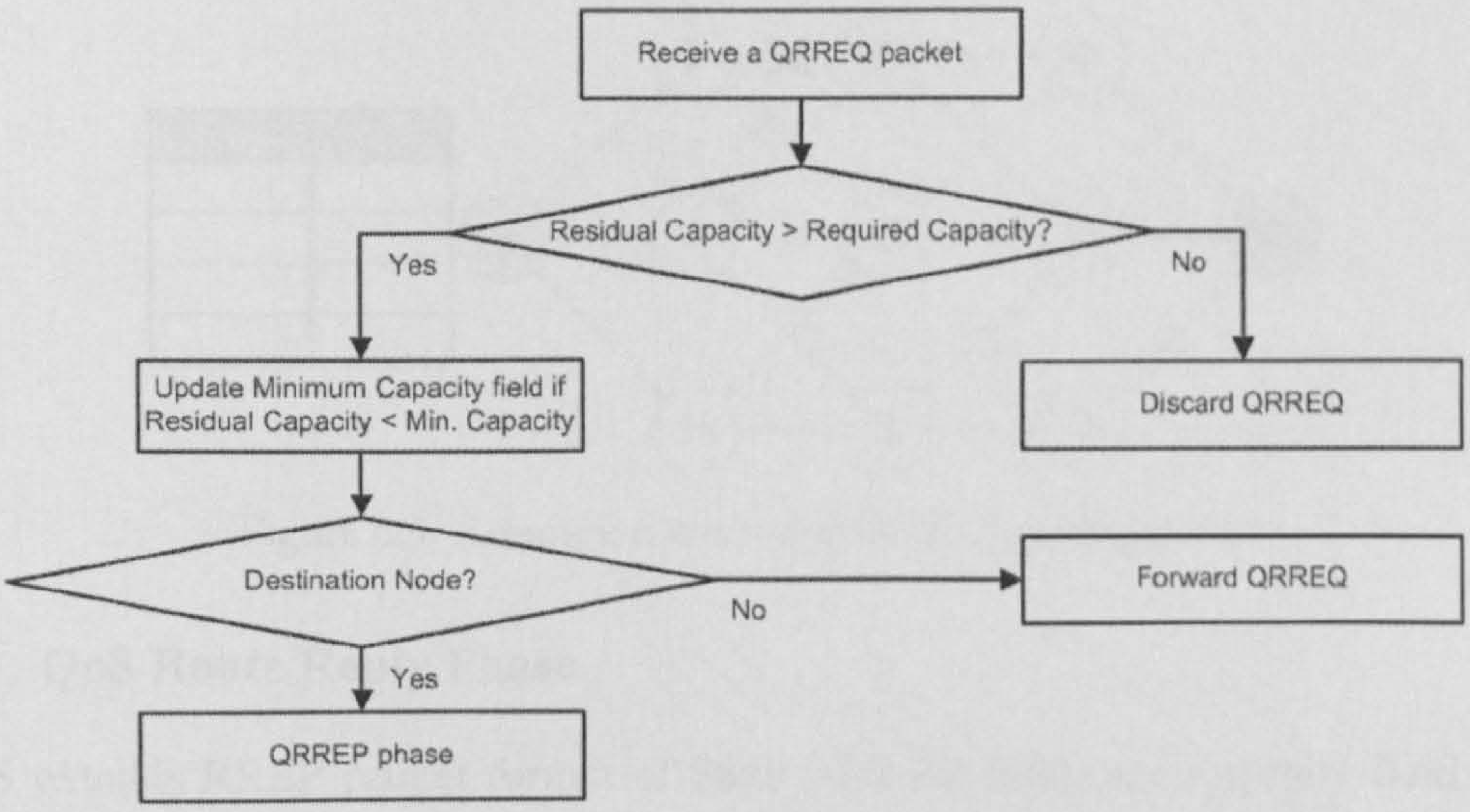


Figure 6.4 Q-SMS route discovery with capacity constraint

In DSR, during a route discovery phase, any intermediate node having a valid route to destination, can reply to the route request. However, state information such as C_{min} to the destination will be required and maintained at the intermediate node of the network to reply to a QREEQ. It is thus preferable to use the reply from the destination to estimate the current network conditions. Therefore, in order to provide QoS guarantees, QRREP always come from the destination in Q-SMS.

Figure 6.5 shows an example of Q-SMS route discovery with admission control, where S is the source node and D is the destination node. S initiates the route discovery phase by sending QRREQ packet to the neighbours with reverse shortest hop count and capacity constraints. In this example, node A has sufficient capacity on its QoS portion of allocated capacity, so it forwards QRREQ to neighbouring nodes. However, node E does not have enough capacity in the QoS portion, so it does not forward QRREQ to D through B and F even though it can reach D. In this fashion if QRREQ reaches the destination by satisfying the reverse shortest hop count and the capacity constraint, then the destination replies with a QRREP packet back to source node using the reverse path identified within the QRREQ packet.

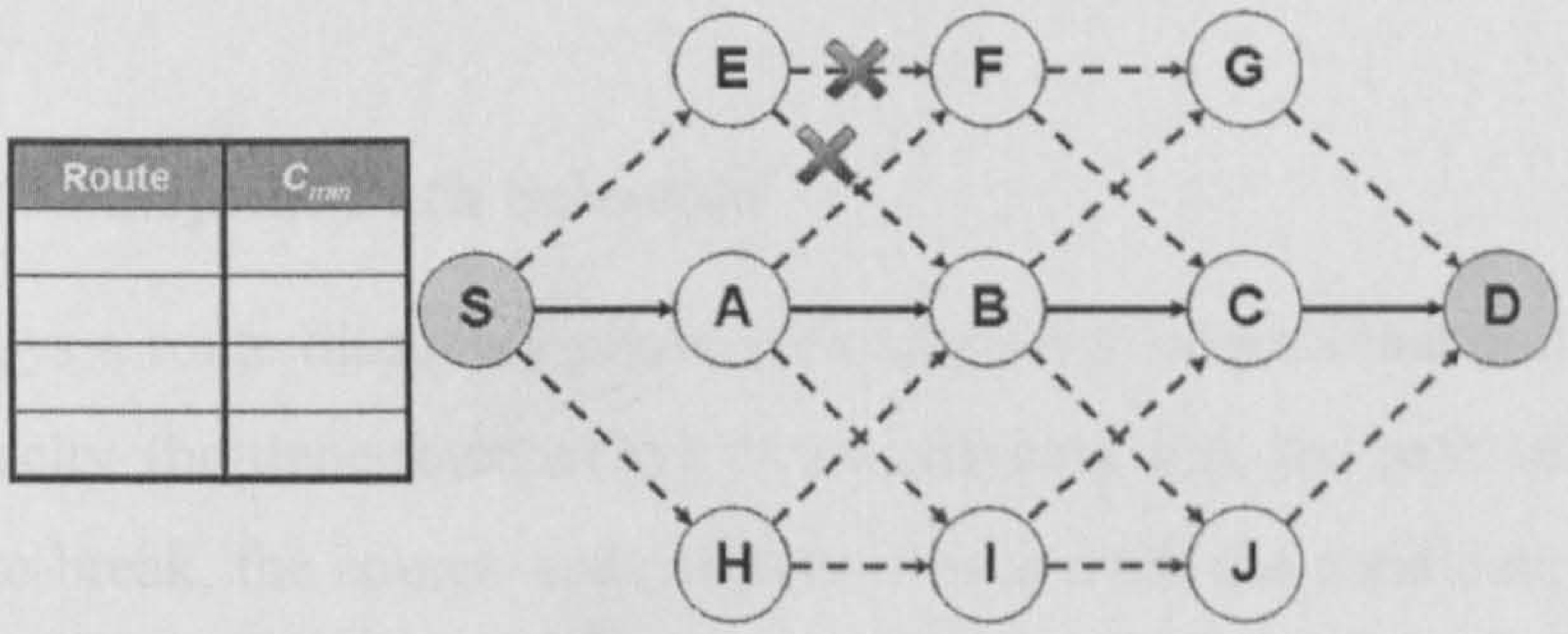


Figure 6.5 Admission control in RREQ propagation

6.2.3 QoS Route Reply Phase

Q-SMS extends RREP packet format of SMS with the minimum capacity field (C_{min}). The QRREP has the format:

$$QRREP = RREP_{SMS} \cup C_{\min}$$

The destination sends a QRREP to the first QRREQ that is received, afterward only limited QRREPs are sent in order to avoid route reply storms. Q-SMS does not update the QoS during route reply since the QoS does not change significantly during this time. Updating QoS both ways would consume battery power of the mobile node, add processing overheads and make the source wait longer for a QRREP.

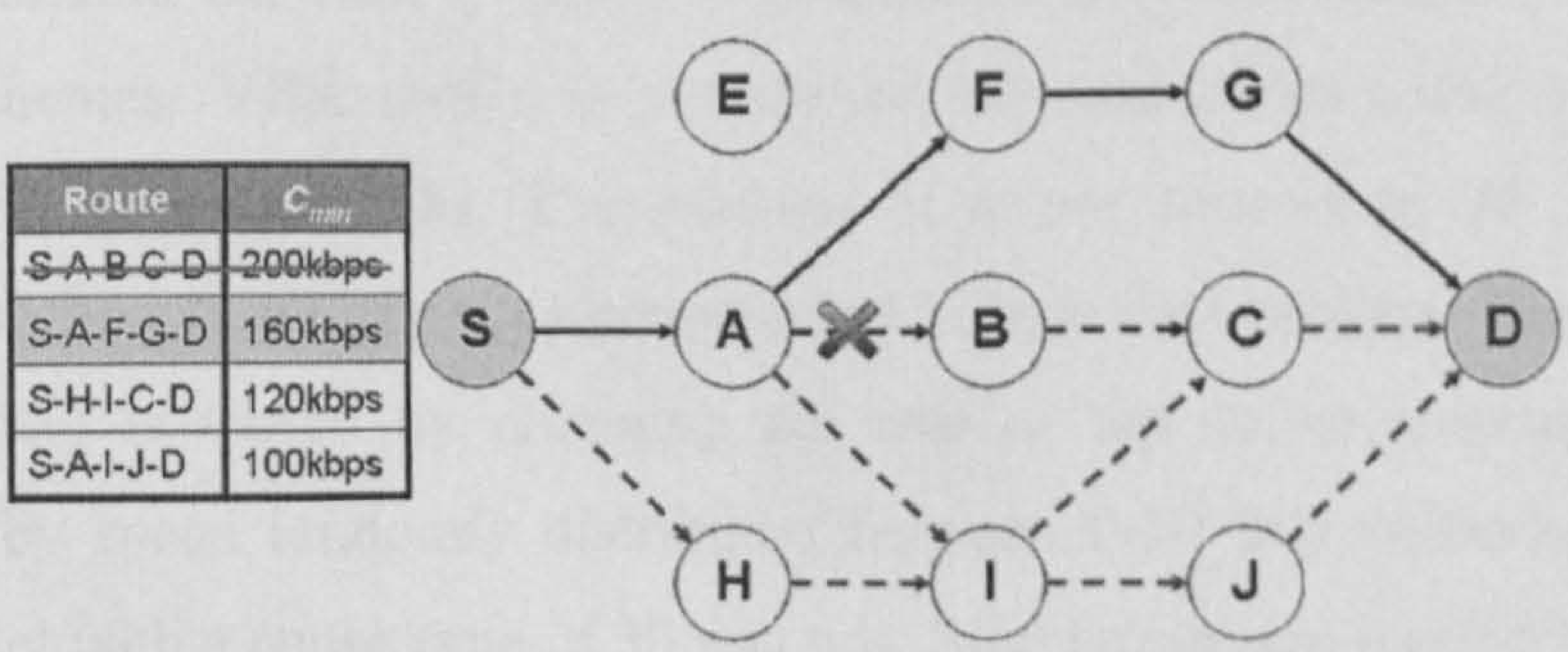
6.2.4 QoS Route Maintenance Phase

Q-SMS adopts the route maintenance approach used in SMS scheme but with a slight modification. When a node encounters a fatal transmission problem at its MAC layer, it generates a RERR packet back to the source node identifying the broken link. When the source node receives a RERR packet, it removes the path containing broken link from the route cache. All routes that contain the broken link in error are truncated at that point. The source node then selects a new valid alternate routing path, with the maximum bottleneck capacity (C_{min}), from the candidate paths over which to forward any data packets. When there is only a single or no routing path available in the cache of the source node, the source instigates a new route discovery process to discover multiple partial-disjoint paths satisfying QoS constraint.

6.2.5 Partial-disjoint Path Selection

Q-SMS employs a route discovery phase with admission control that collects minimum available capacity (bottleneck capacity) of an outgoing link for path selection. In the case of a route break, the source node selects a route from the candidate paths that has maximum bottleneck capacity (C_{min}).

An example that illustrates the selection of partial-disjoint paths is shown in Figure 6.6. Routes are sorted according to maximum bottleneck capacity (C_{min}). Consider the case of traffic flowing between nodes S and D using link S-A-B-C-D as primary path. In the case of link failure between A-B, node A will send a RERR packet back to the source node which will discard route path S-A-B-C-D. The path S-A-F-G-D is selected from the candidate paths S-H-I-C-D and S-A-I-J-D as it has the maximum bottleneck capacity (C_{min}).



Route	C_{min}
~~S-A-B-C-D~~	~~200kbps~~
S-A-F-G-D	160kbps
S-H-I-C-D	120kbps
S-A-I-J-D	100kbps

Figure 6.6 Partial-disjoint path selection

6.3 Simulation Results and Discussion

In this section, simulation based experiments are conducted to evaluate the improvements offered by Q-SMS and contrasted with SMS which has no QoS support. The previously proposed QoS-aware routing schemes [122][123][124][125] obtained results assuming IEEE 802.11 original standard operating at 2 Mbps. Currently, however, the IEEE 802.11b is the de facto reference technology for MANETs. There have been relatively few previous studies into the ability of the IEEE 802.11b standard operating at 11 Mbps to support QoS in MANETs. Therefore, in this study, IEEE 802.11b standard operating at 11 Mbps is considered. The key metrics used in measuring the schemes' performance are: goodput, average end-to-end delay and normalised routing load. The average hop count or mean path length is also discussed here to validate and understand the obtained results.

A dedicated 'Utilisation Monitor' module (*UMon*) is integrated into the IEEE 802.11Ext[17] implementation [129] within network simulator *NS-2.33* [78]. *UMon* effectively monitors and records the state of the interface and is linked to the channel activity function to estimate the channel utilisation. The simulation environment consists of 50 wireless nodes forming an ad hoc wireless network, moving over a 700 x 500 m^2 flat space. The propagation model is a two ray ground with 250 m transmission range, 550 m carrier sensing range and a nominal bit rate of 11 Mbps. Both Constant Bit Rate (CBR) and Variable Bit Rate (VBR) are considered as traffic sources to analyse the contending schemes. VBR traffic is established between nodes using an exponential ON/OFF traffic generator [78]. The number of active sources is 20 nodes, chosen randomly from the full set of nodes generating 512 byte data packets. The network load (traffic intensity) is varied by changing the rate of the active sources. Mobility is characterised by speed randomly distributed between 0-10 m/s following the random waypoint model with a pause time of 30 seconds. Simulations are run for 900 seconds of

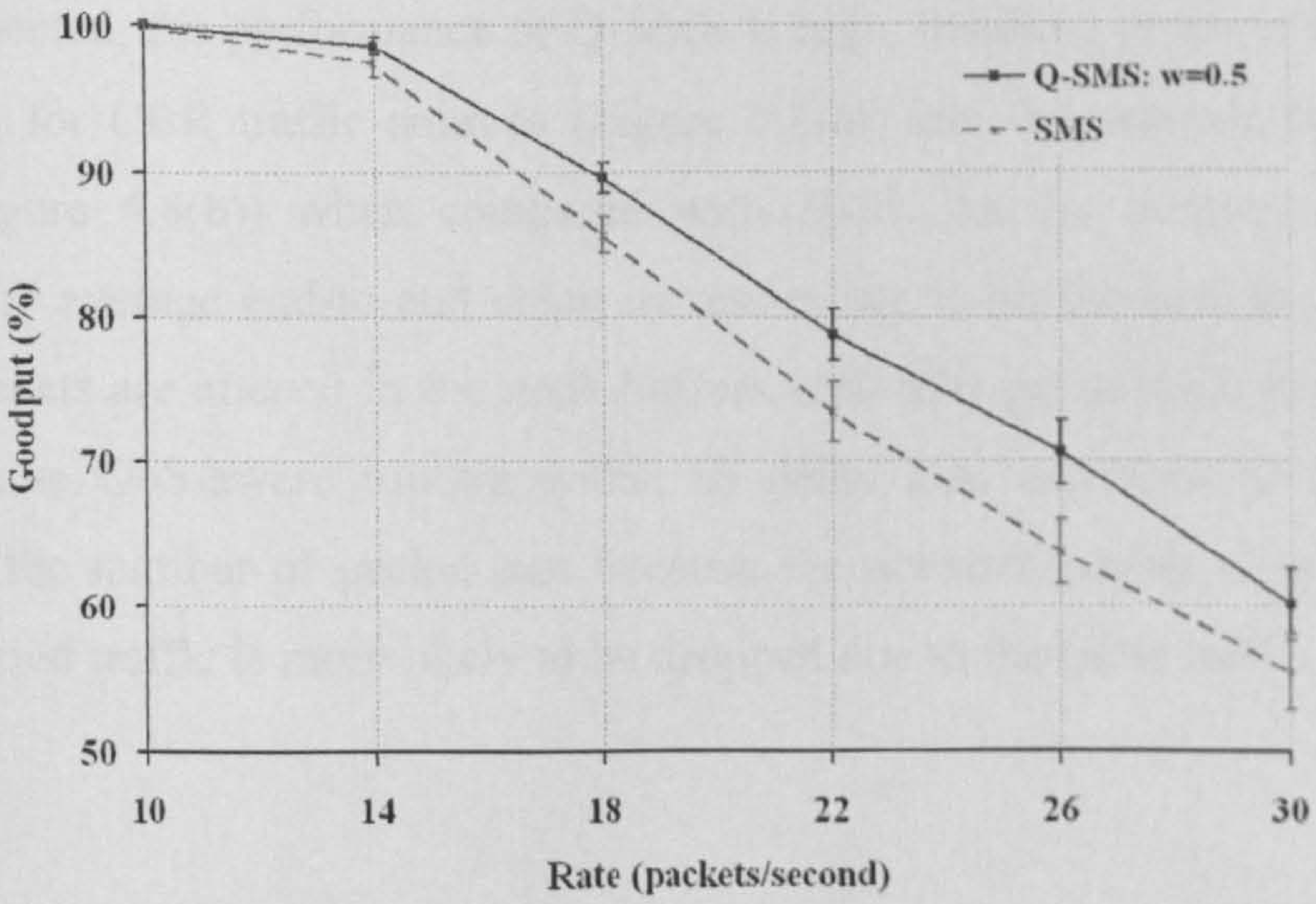
---

[17] The IEEE 802.11Ext MAC models transmission and reception coordination, backoff management and channel state monitoring in a structured and modular manner. The reader is referred to [129] for details about this modelling.

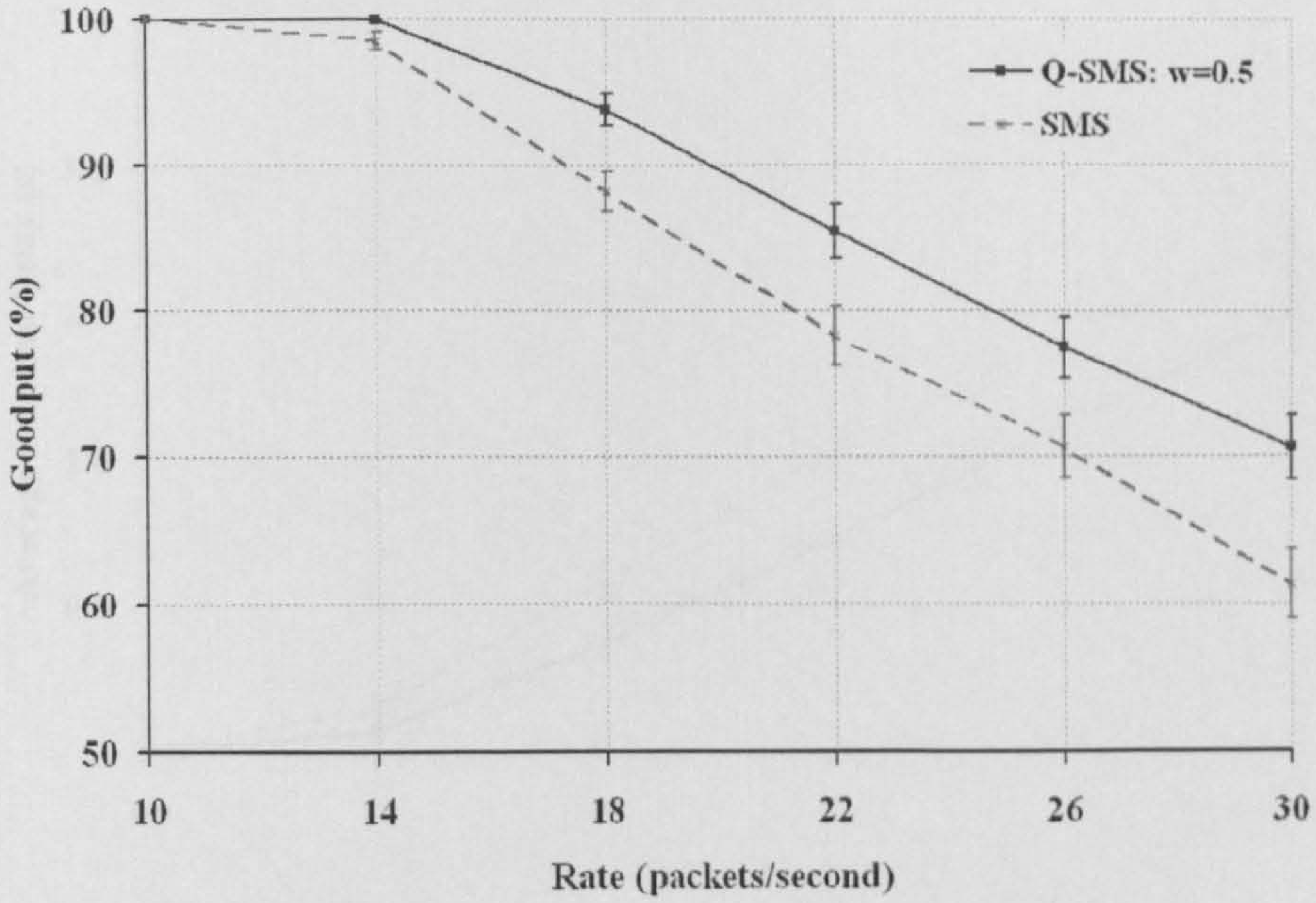
real time. Each data point represents an average of twenty runs using different seeds with the corresponding confidence interval of 95%. As the residual capacity can change over time, it is important to measure it rapidly. A moving window of *w*=0.5 seconds is utilised to estimate the residual capacity.

6.3.1 Goodput

Figure 6.7 shows the goodput versus the load applied to the network for Q-SMS and compared with SMS. Q-SMS with CBR traffic sources (Figure 6.7(a)) performs better (up to 11%) and Q-SMS with VBR traffic sources (Figure 6.7(b)) performs better (up to 15%) when compared with SMS. Initially at low loads (10 and 14 packets per second), the goodput of Q-SMS and SMS is almost identical, with Q-SMS performing slightly better. As the rate of the sources increases to intermediate loads, SMS suffers from congestion in spite of using shorter alternative paths to repair route breaks and, as a result, the successfully delivered traffic drops, resulting in reduced goodput. On the other hand, Q-SMS suffers lower congestion by using capacity constrained alternative paths to repair route breaks. As the rate increases further to high loads, any performance improvement reduces, because the network becomes saturated and any further increase in the injected traffic is more likely to be dropped.



(a) Goodput – CBR sources

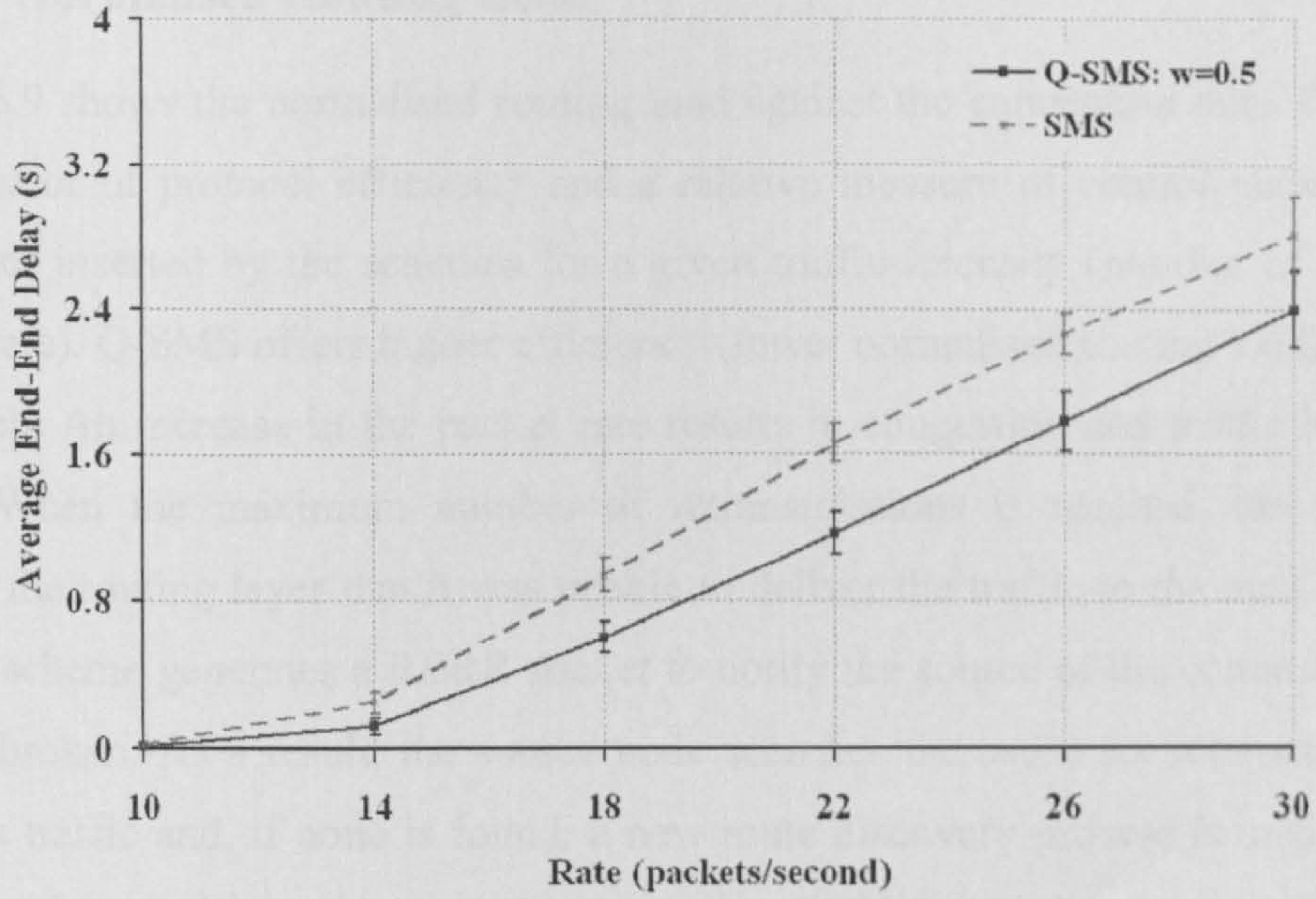


(b) Goodput – VBR sources

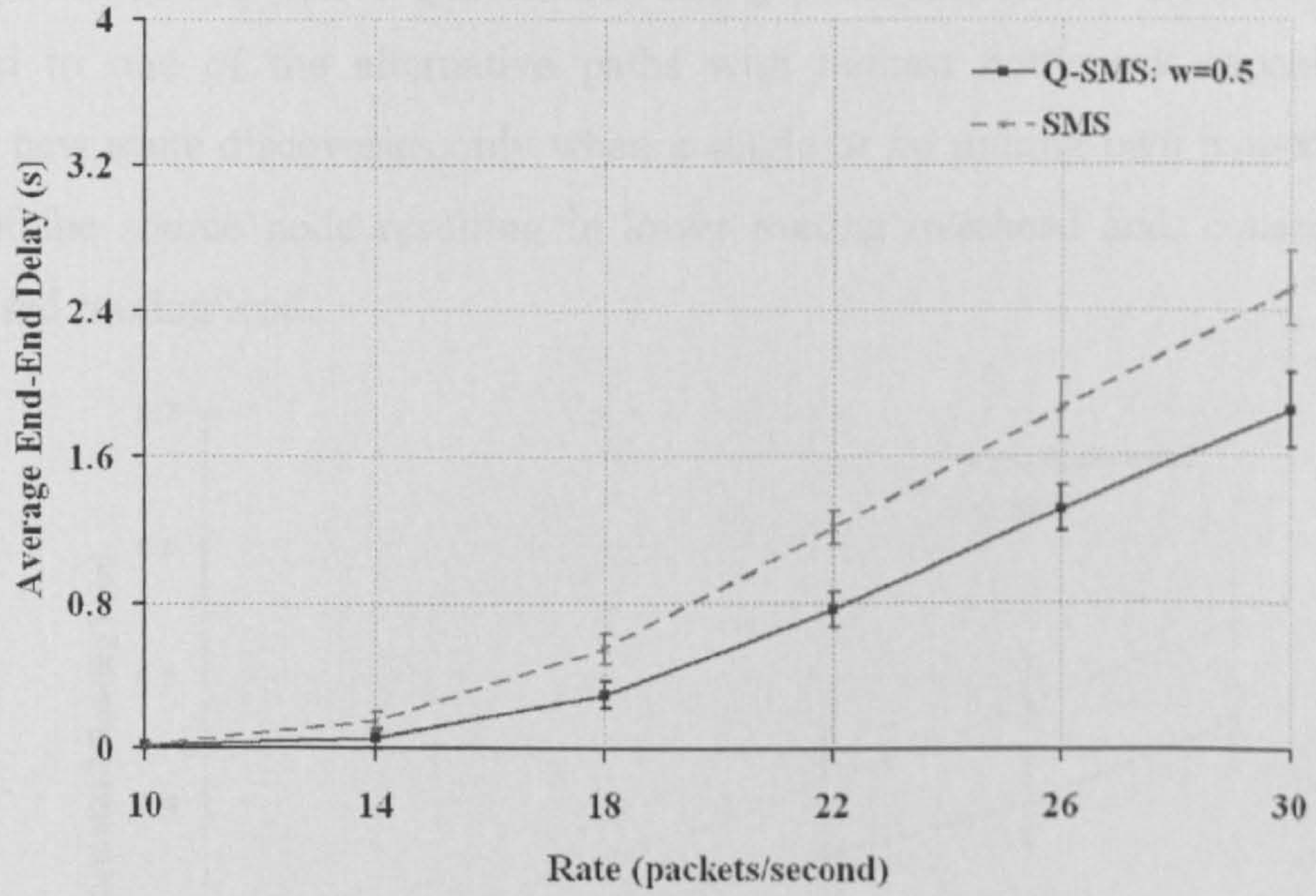
Figure 6.7 Goodput versus packet rate

6.3.2 Average End-to-End Delay

Figure 6.8 shows how the average end-to-end delay varies against the number of packet rate. As expected, the performance of Q-SMS is high, resulting in lower delays of up to 0.5 seconds for CBR traffic sources (Figure 6.8(a)) and 0.7 seconds for VBR traffic sources (Figure 6.8(b)) when compared with SMS. As the number of packet rate increases, the average end-to-end delay increases due to an increase in congestion and the data packets are queued in the node buffers until they get served. Although Q-SMS has alternative QoS-aware routing paths, its delay also increases gradually with an increase in the number of packet rate because the network begins to saturate and any further injected traffic is more likely to be dropped due to the finite buffer size.



(a) Average end-to-end delay – CBR sources

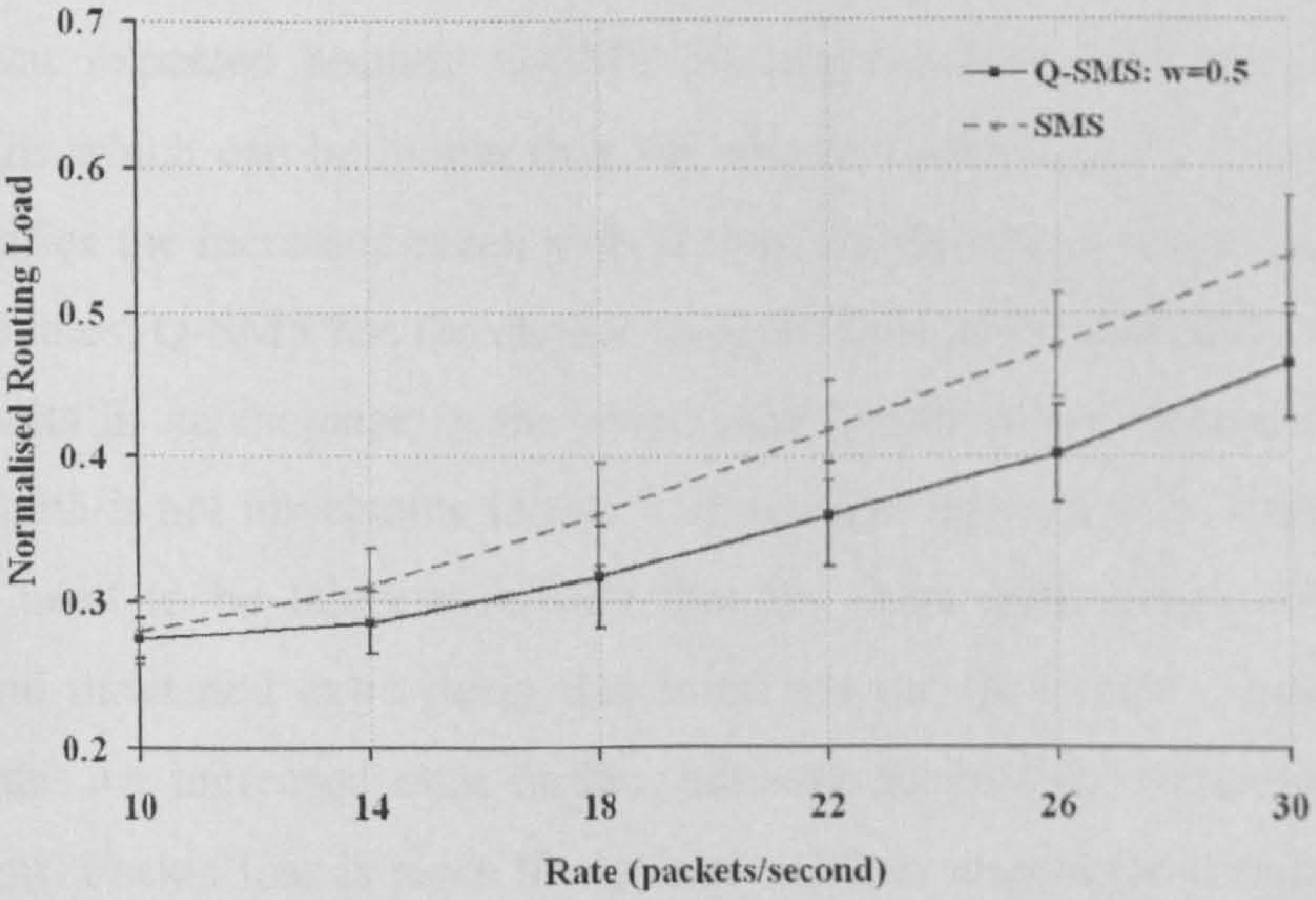


(b) Average end-to-end delay – VBR sources

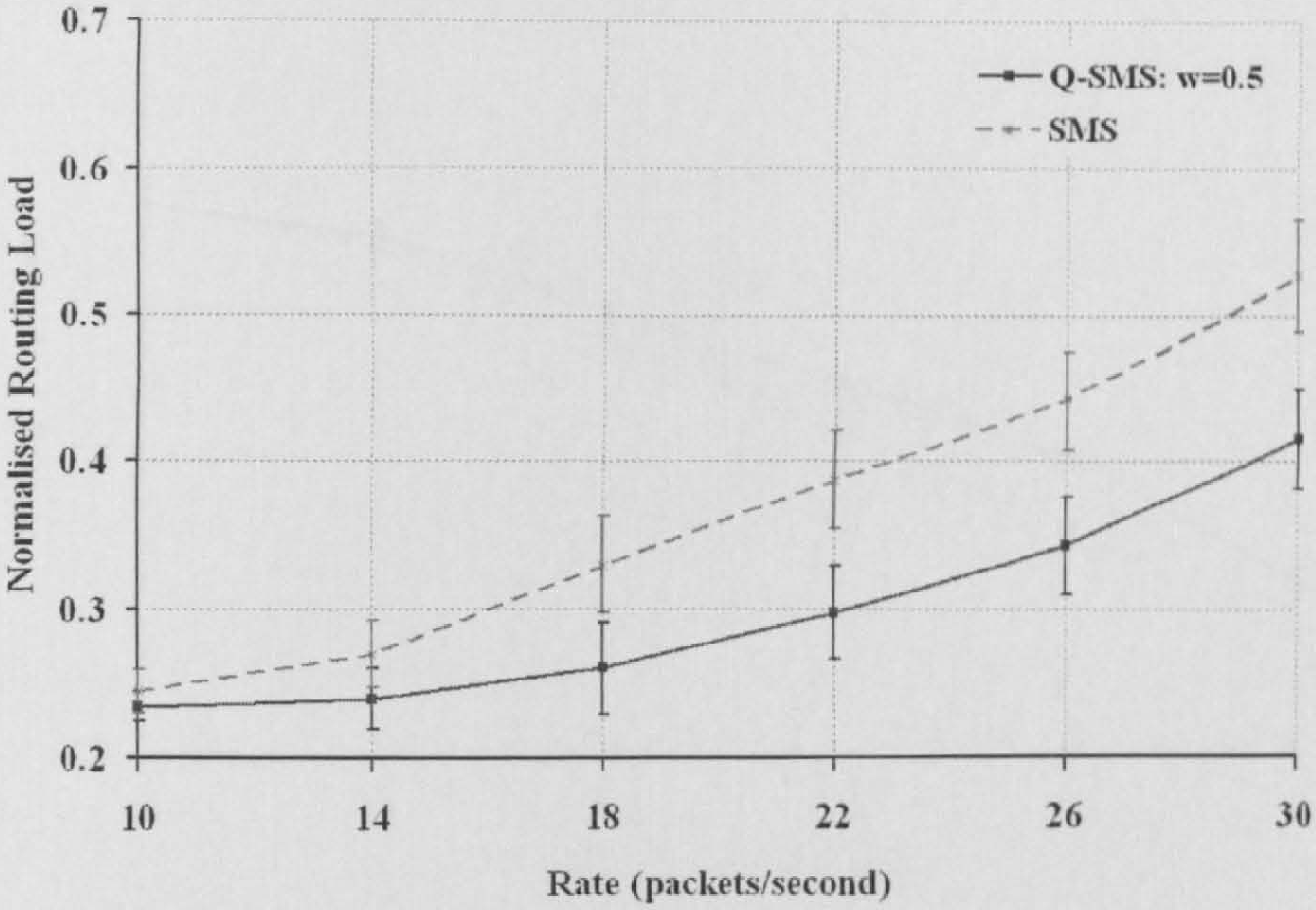
Figure 6.8 Average end-to-end delay versus packet rate

6.3.3 Normalised Routing Load

Figure 6.9 shows the normalised routing load against the connection rate. The metric is an indicator of protocol efficiency and a relative measure of control packets (routing overhead) inserted by the schemes for a given traffic intensity (number of sources and packet rate). Q-SMS offers higher efficiency (lower normalised routing load) throughout the graph. An increase in the packet rate results in congestion and traffic loss starts to occur. When the maximum number of retransmissions is reached, the MAC layer notifies the routing layer that it was unable to deliver the traffic to the next hop and the routing scheme generates a RERR packet to notify the source of the connection that the path is broken. As a result, the source node searches the cache for alternative paths to route its traffic and, if none is found, a new route discovery process is instigated. SMS has alterative routing paths cached but may not fulfil the application requirements and triggers new route discoveries which increase the normalised routing load. On the other hand, Q-SMS has alterative QoS-aware routing paths cached and the affected traffic is switched to one of the alternative paths with highest bottleneck capacity. Q-SMS triggers new route discoveries only when a single or no routing path is available in the cache of the source node resulting in lower routing overhead and, consequently, the normalised routing load.



(a) Normalised routing load – CBR sources

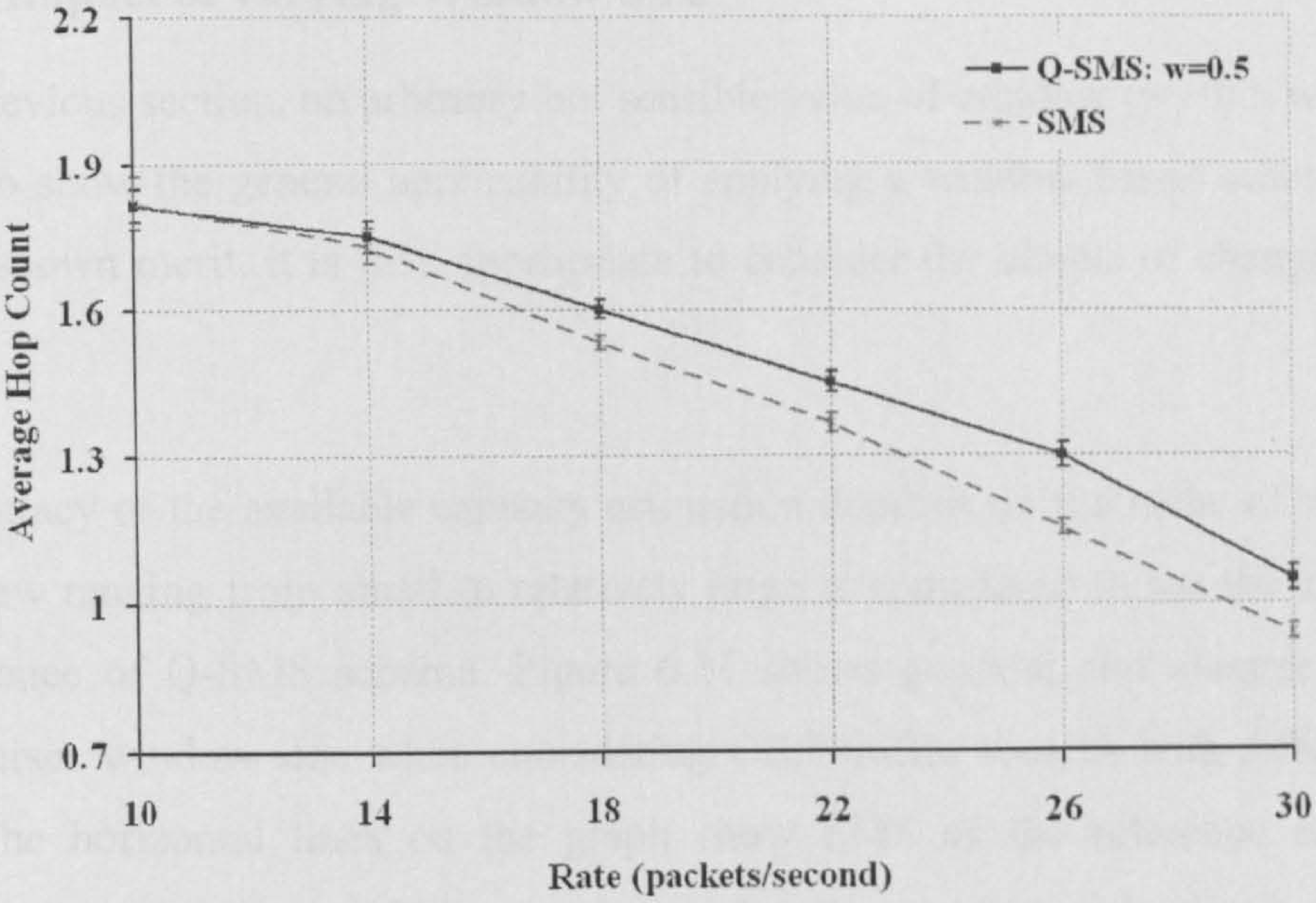


(b) Normalised routing load – VBR sources

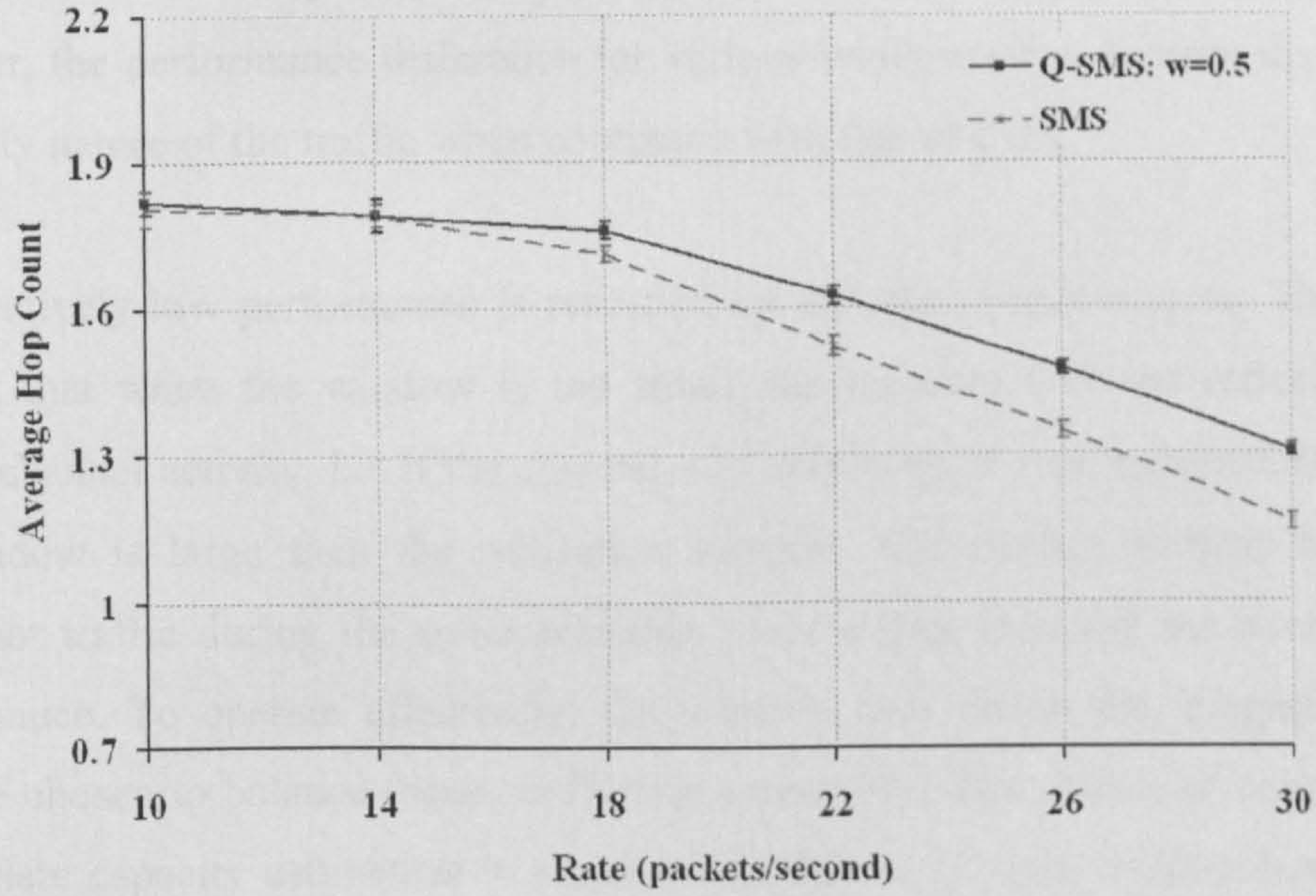
Figure 6.9 Normalised routing load versus packet rate

6.3.4 Average Hop Count

The average hop count is the average distance between any pair of nodes or the mean path length in the network. Figure 6.10 shows the average hop count against the packet rate. The average hop count of Q-SMS is higher than that of SMS for all rates. This behaviour was expected because Q-SMS discovers and exploits capacity-constrained multiple paths which can be longer than the shorter multiple paths discovered by SMS and this justifies the increased mean path length. As the rate of the sources increases to intermediate rates, Q-SMS has the chance to exploit the additional discovered QoS paths and this results in an increase in the mean path length when compared to SMS. The increased length is not inwelcome in that it allows connectivity to be retained. However, care would need to be taken to ensure that the extra path length did not result in excessive and unwanted extra delay that impacted the QoS requirements of the flow. When the rates are increased even further, network congestion is increased, packet loss starts to occur. Packet loss is more likely start to occur in packets propagated over long alternative paths, leading to decreased mean path length.



(a) Average hop count – CBR sources



(b) Average hop count – VBR sources

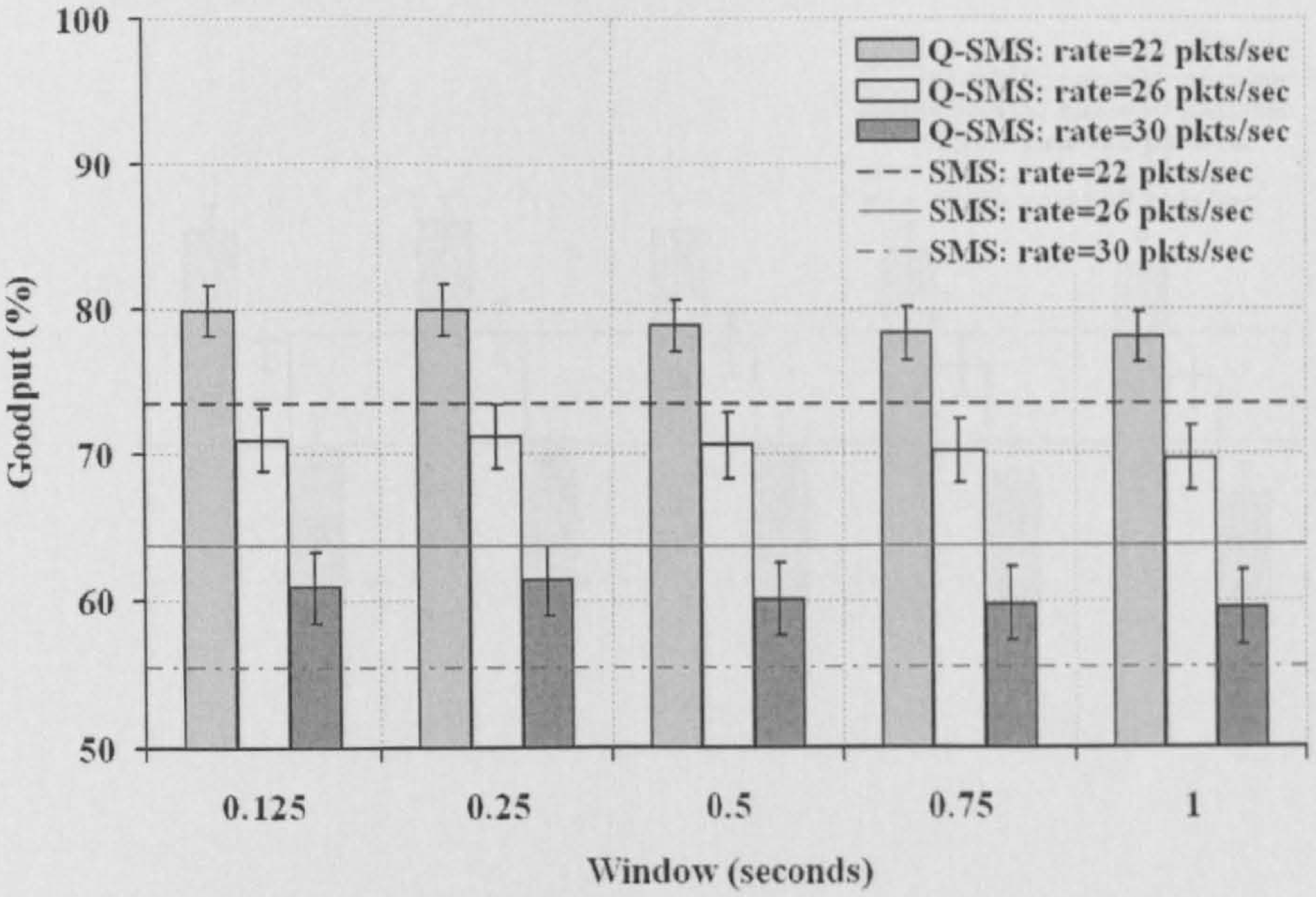
Figure 6.10 Average hop count versus packet rate

6.3.5 Impact of varying Window Size

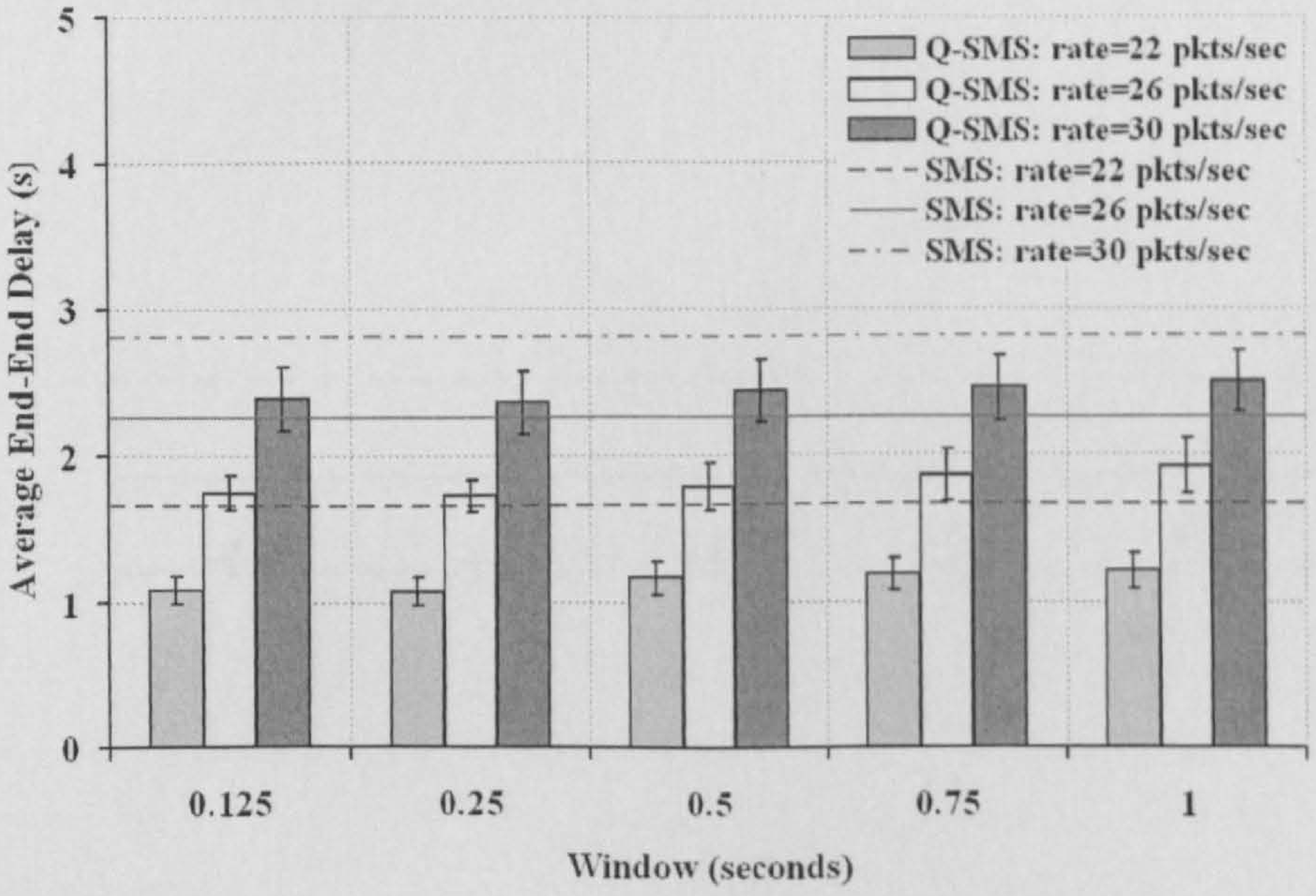
In the previous section, an arbitrary but sensible value of window (w) 0.5 was adopted; simply to show the general applicability of applying a window based scheme to SMS. Having shown merit, it is then appropriate to consider the effects of changing window size.

The accuracy of the available capacity estimation depends on the value of window (w). A window ranging from small to relatively large is considered to see the affect on the performance of Q-SMS scheme. Figure 6.11 shows goodput and average end-to-end delay verses window size when considering CBR traffic sources with different packet rates. The horizontal lines on the graph show SMS as the reference scheme with corresponding packet rates. Better goodput and average end-to-end delay is observed for window size at 0.25. Additionally, similar observations can be made for the goodput and average end-to-end delay when considering VBR traffic sources as shown in Figure 6.12. However, the performance difference for various window sizes is more distinct due to the bursty nature of the traffic when compared with that of CBR.

Comparatively low performance is recorded for all other window sizes. This is due to the fact that when the window is too small, the measure will not reflect accurately overall channel activity, i.e. if the channel was only busy or free momentarily. While, if the window is large then the utilisation measure will contain historic but possibly redundant traffic during the route selection process thus reducing the overall network performance. To operate effectively, the window over which the integration is done must be chosen to balance these conflicting constraints. The choice of window size for appropriate capacity estimation is a major impediment to such window-based schemes and it is this aspect that is addressed in the next section. A new scheme for effective capacity estimation is presented where an additional function is introduced to mitigate the limitation of the previous approaches.

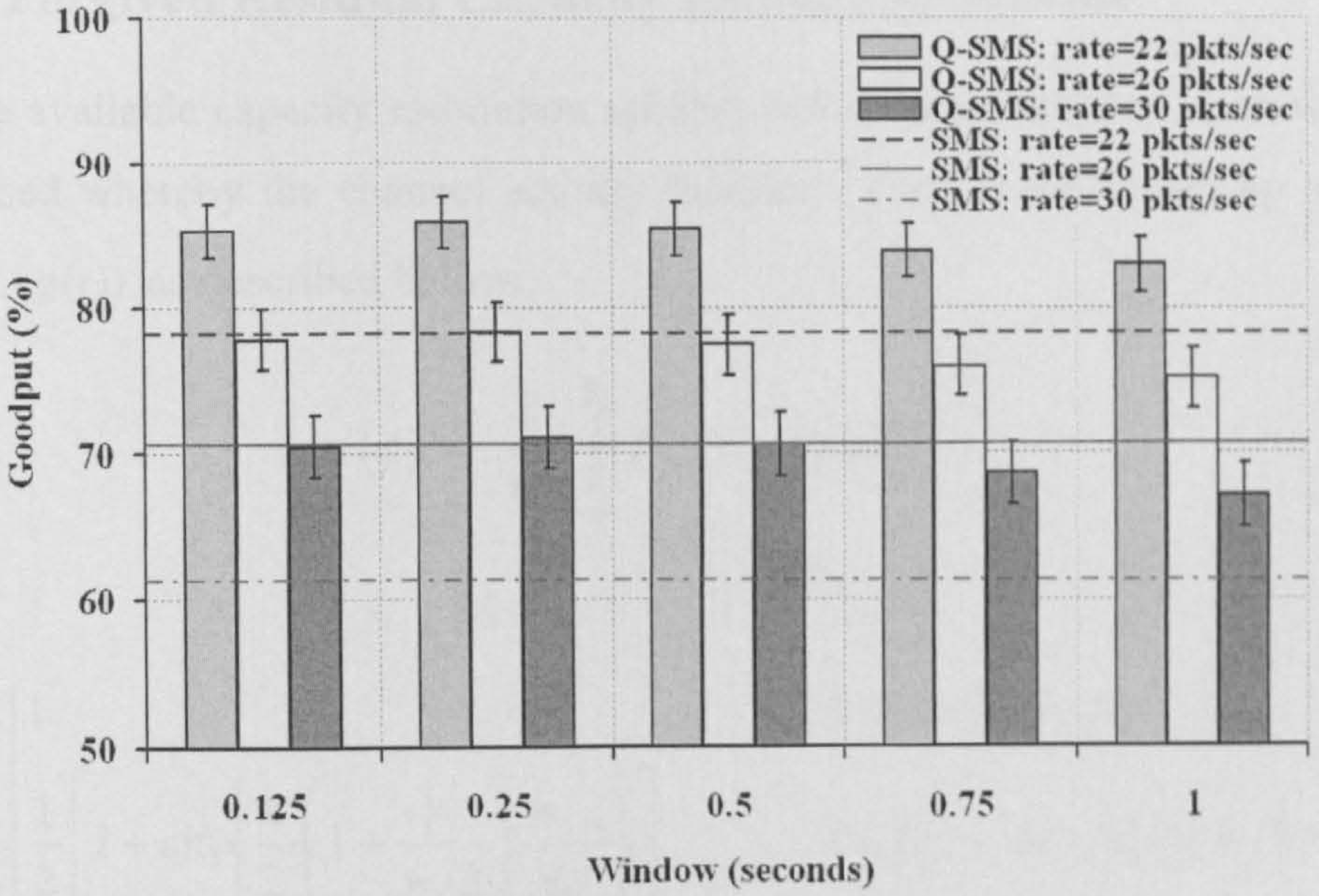


(a) Goodput versus window size

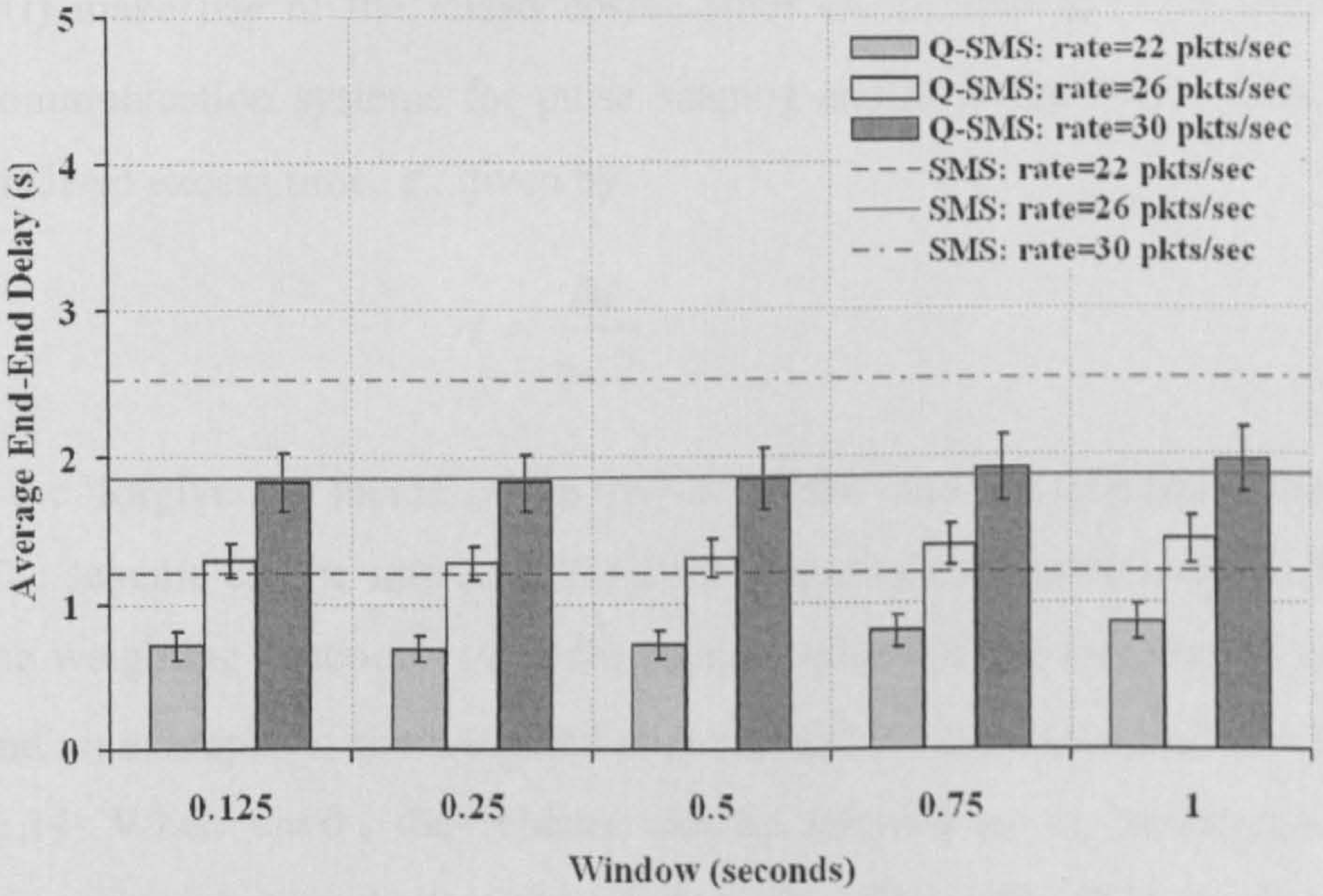


(b) Average end-to-end delay versus window size

Figure 6.11 Results for Q-SMS with different *w* – CBR sources



(a) Goodput versus window size



(b) Average end-to-end delay versus window size

Figure 6.12 Results for Q-SMS with different *w* – VBR sources

6.4 Forgiven Residual Capacity Estimation Scheme

A simple available capacity estimation scheme with a lower reliance upon window size is proposed whereby the channel activity function ($f(t)$) is multiplied by a weighting function ($g(t)$) as described below:

$$u(t) = \frac{1}{w}\int_{-w}^{0} f(t).g(t)\,dt \tag{6.7}$$

and

$$g(t) = \begin{cases} 1 & |t| \le (w/2) - \Delta t \\ \dfrac{1}{2}\left[1 + \sin\left\{\dfrac{\pi}{2}\left(1 - \dfrac{|t|}{w/2}\right)\dfrac{w/2}{\Delta t}\right\}\right] & (w/2) - \Delta t < |t| < (w/2) + \Delta t \\ 0 & |t| \ge (w/2) + \Delta t \end{cases} \tag{6.8}$$

where $g(t)$ make use of the raised cosine filter characteristics [130] widely used in digital communication systems for pulse shaping and Δt is the excess (absolute) time. The normalised excess time, τ, given by:

$$\tau = \frac{\Delta t}{w/2} \tag{6.9}$$

is called the 'forgiveness factor' or the 'roll-off factor' and can take any value between 0 and 1. The scheme can be said as a 'forgiven' capacity estimation scheme. Figure 6.13 shows the weighting function ($g(t)$) for several values of the forgiveness factor (τ) at w=0.5 and an example of the weighted channel activity for w=0.5 and τ=1 is shown in Figure 6.14. When $\tau = 0$, the scheme can be referred as an 'unforgiven' capacity estimation scheme but for half of the window size. The value of τ can be determined through testing to obtain a good estimation of utilisation such that it can give more emphasis to the recent information rather than the past information over a time window (w).

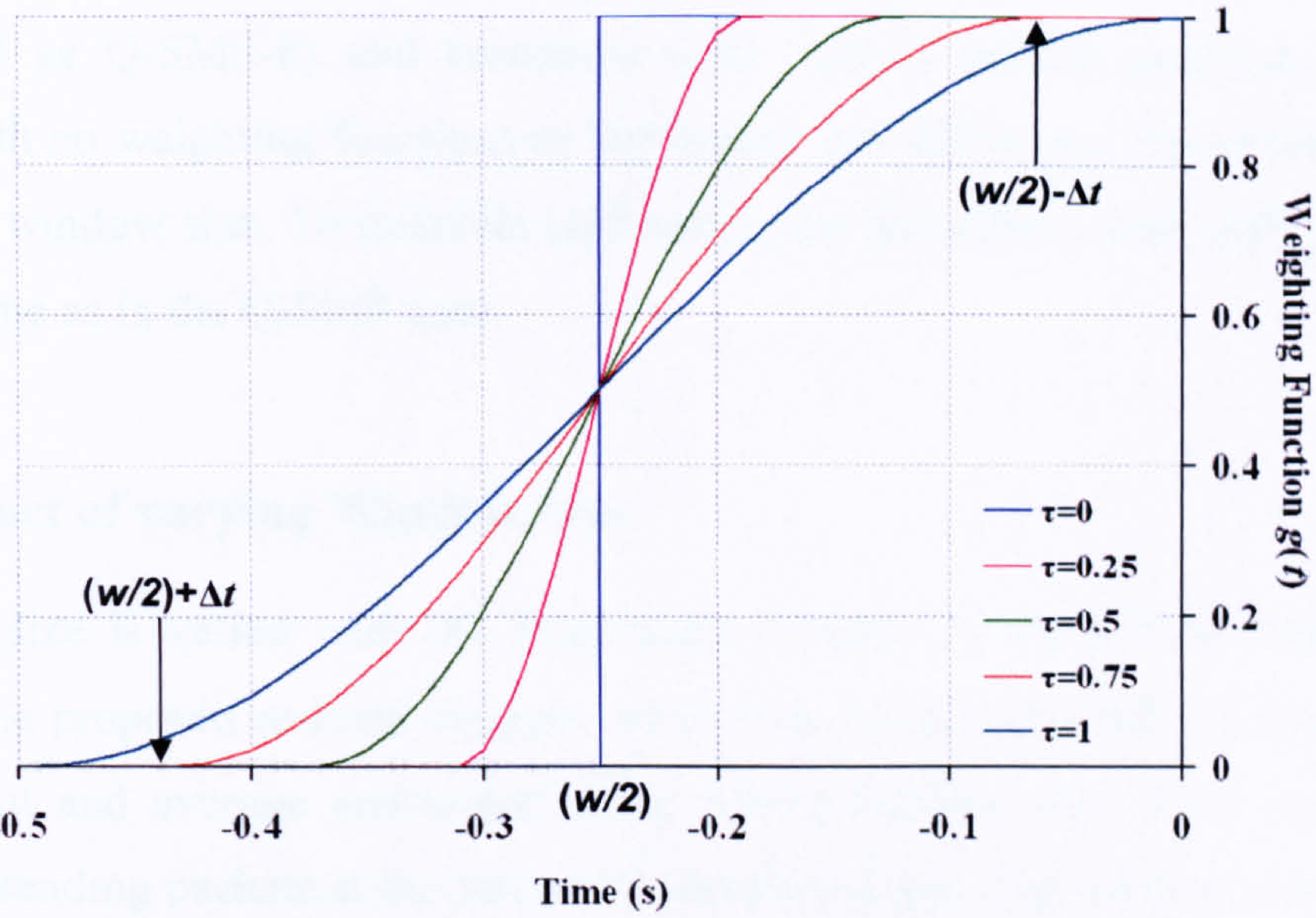


Figure 6.13 Illustration of forgiveness factor (*w*=0.5)

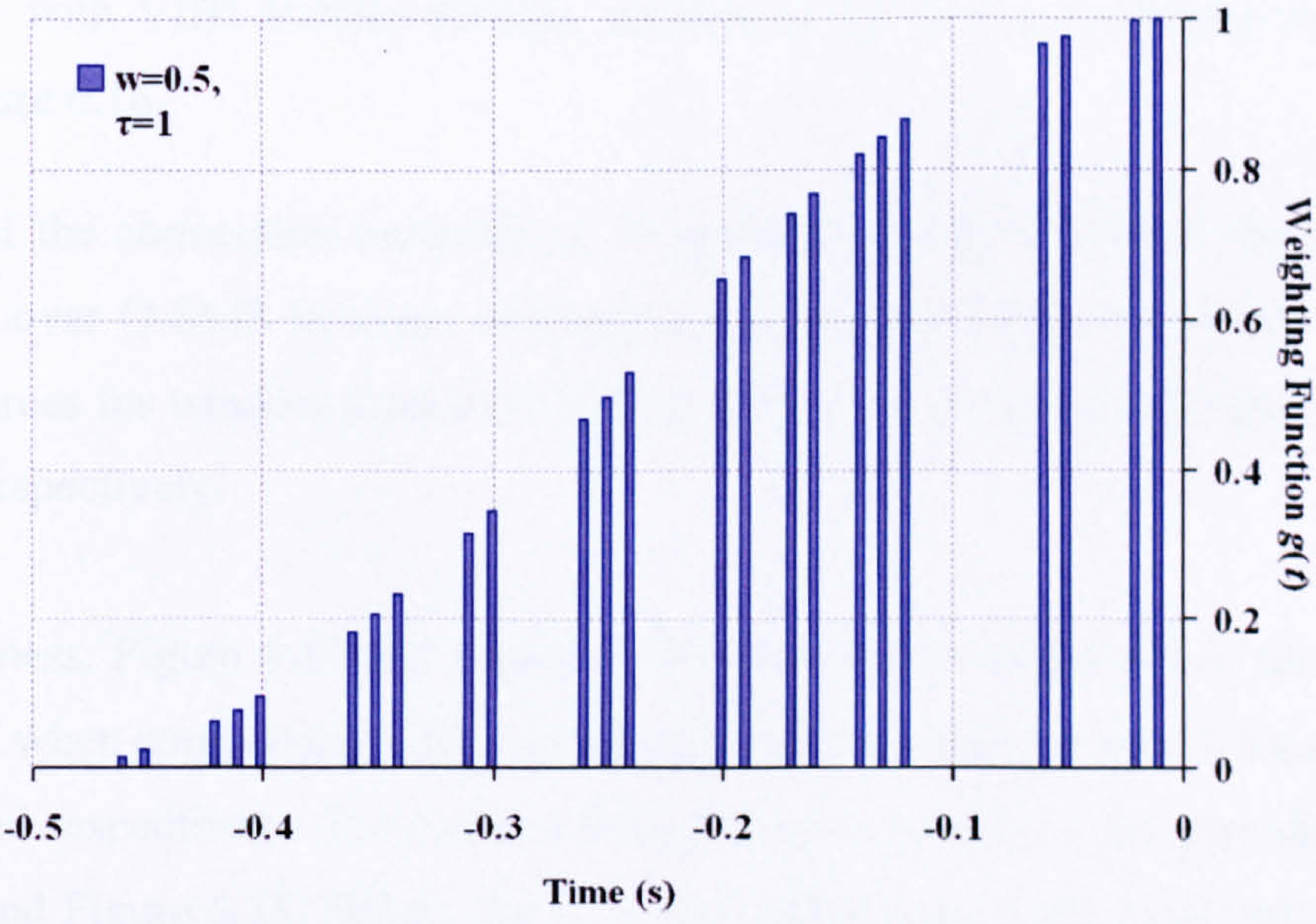


Figure 6.14 Example of the weighted channel activity for *w*=0.5 and τ=1

6.5 Simulation Results and Discussion

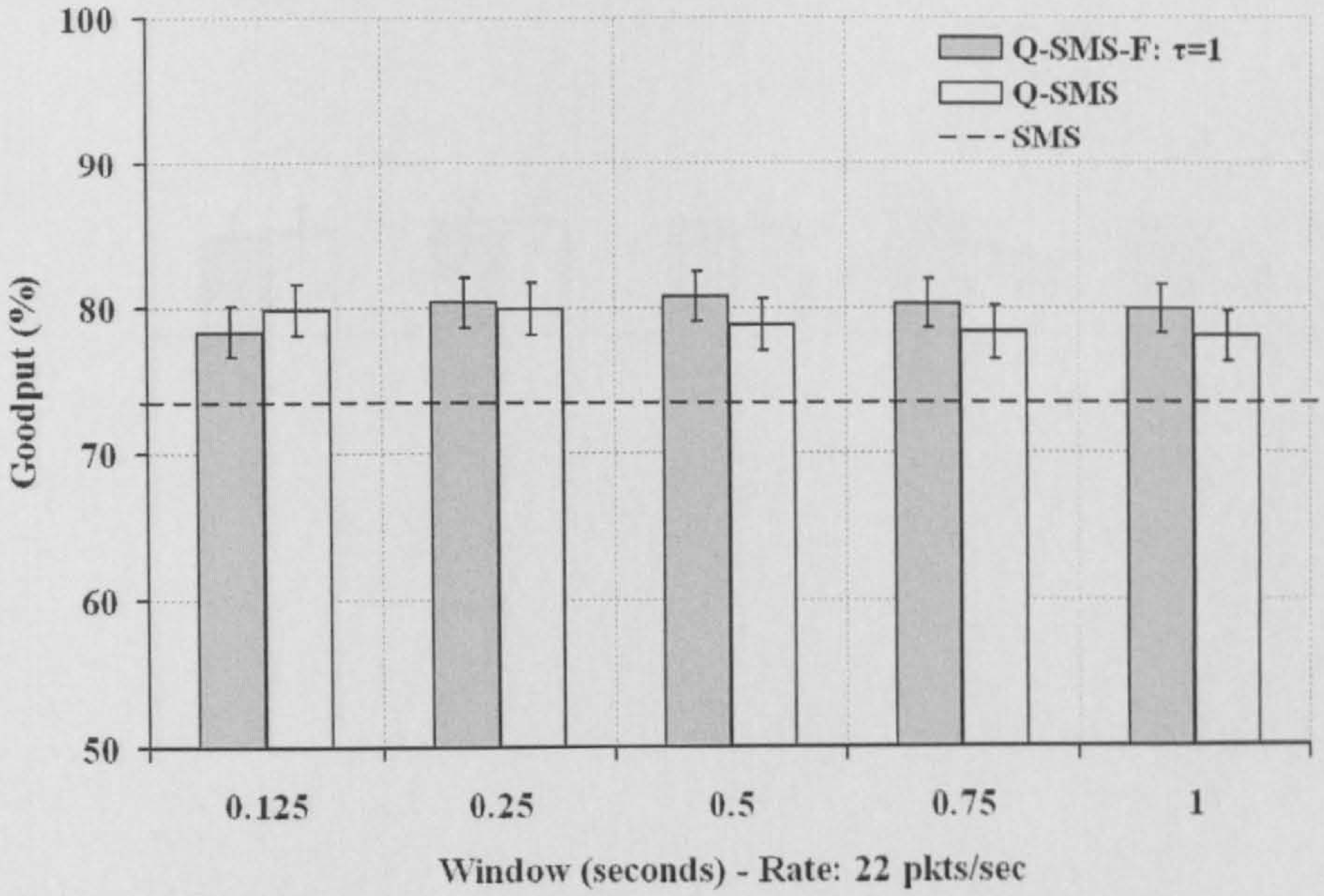
This part of the experimental work evaluates Q-SMS with 'forgiven' capacity estimation (here referred as Q-SMS-F) and compares with Q-SMS (which performs capacity estimation with no weighting function) to benchmark any difference that results from a change in the window size. To maintain uniformity, the parameters used with Q-SMS-F remain the same as in the Q-SMS case.

6.5.1 Impact of varying Window Size

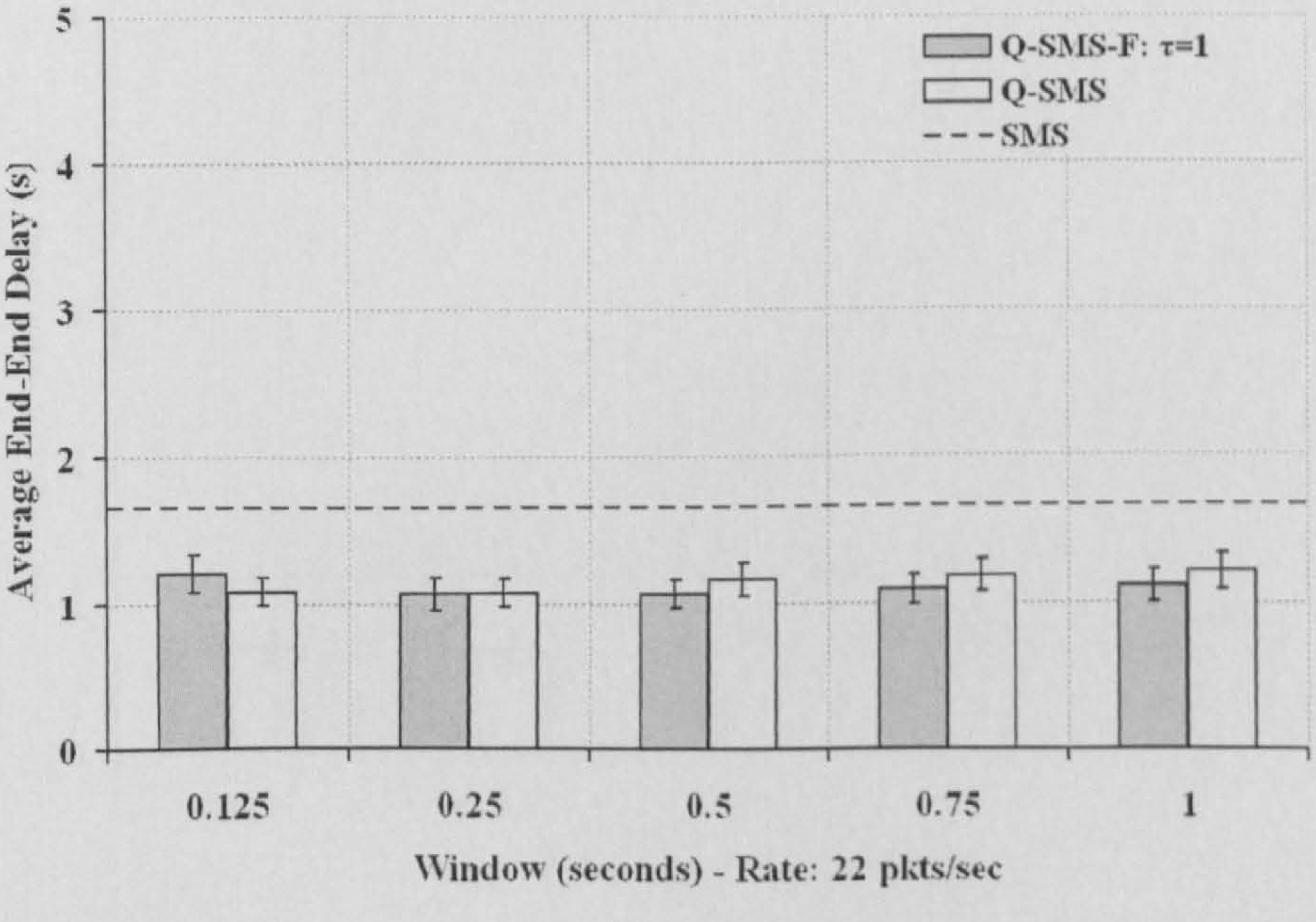
The window size is varied with the forgiveness factors (r) set at 1 to evaluate the potential of the proposed residual capacity estimation scheme (Q-SMS-F). Figure 6.15 shows goodput and average end-to-end delay verses window size when considering CBR sources sending packets at the rate of 22 packets/second. The horizontal line on the graph shows SMS as the reference scheme with rate=22 packets/second. Q-SMS-F shows improvement over Q-SMS in terms of goodput (up to 3%) and average end-to-end delay (up to 0.1 second) for window sizes of 0.5, 0.75 and 1. Similar are the results for Q-SMS-F with VBR sources sending packets at the rate of 22 packets/second as shown in Figure 6.16.

As the rate of the connection increases to 26 packets/second, Q-SMS-F shows further improvement over Q-SMS in terms of goodput and average end-to-end delay for CBR and VBR sources for window sizes of 0.25, 0.5, 0.75 and 1 as shown in Figure 6.17 and Figure 6.18 respectively.

For completeness, Figure 6.19 and Figure 6.20 shows results of Q-SMS-F for different window sizes when considering CBR and VBR sources sending packets at the rate of 30 packets/second respectively. The patterns followed are identical to the one observed in Figure 6.17 and Figure 6.18. Hence, the Q-SMS-F offers marginally better performance over Q-SMS but not in cases where window size is very small ($w = 0.125$).

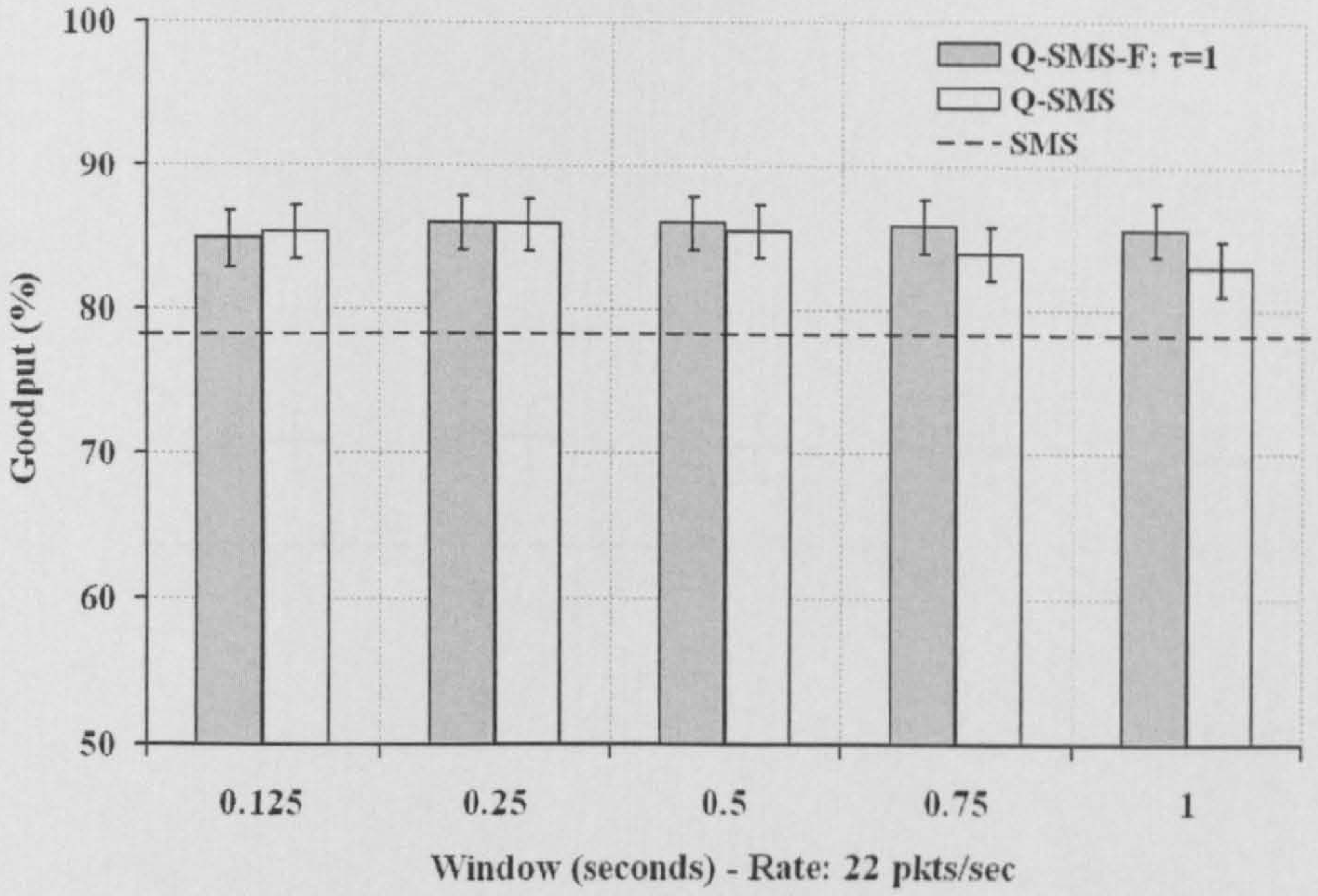


(a) Goodput versus window size

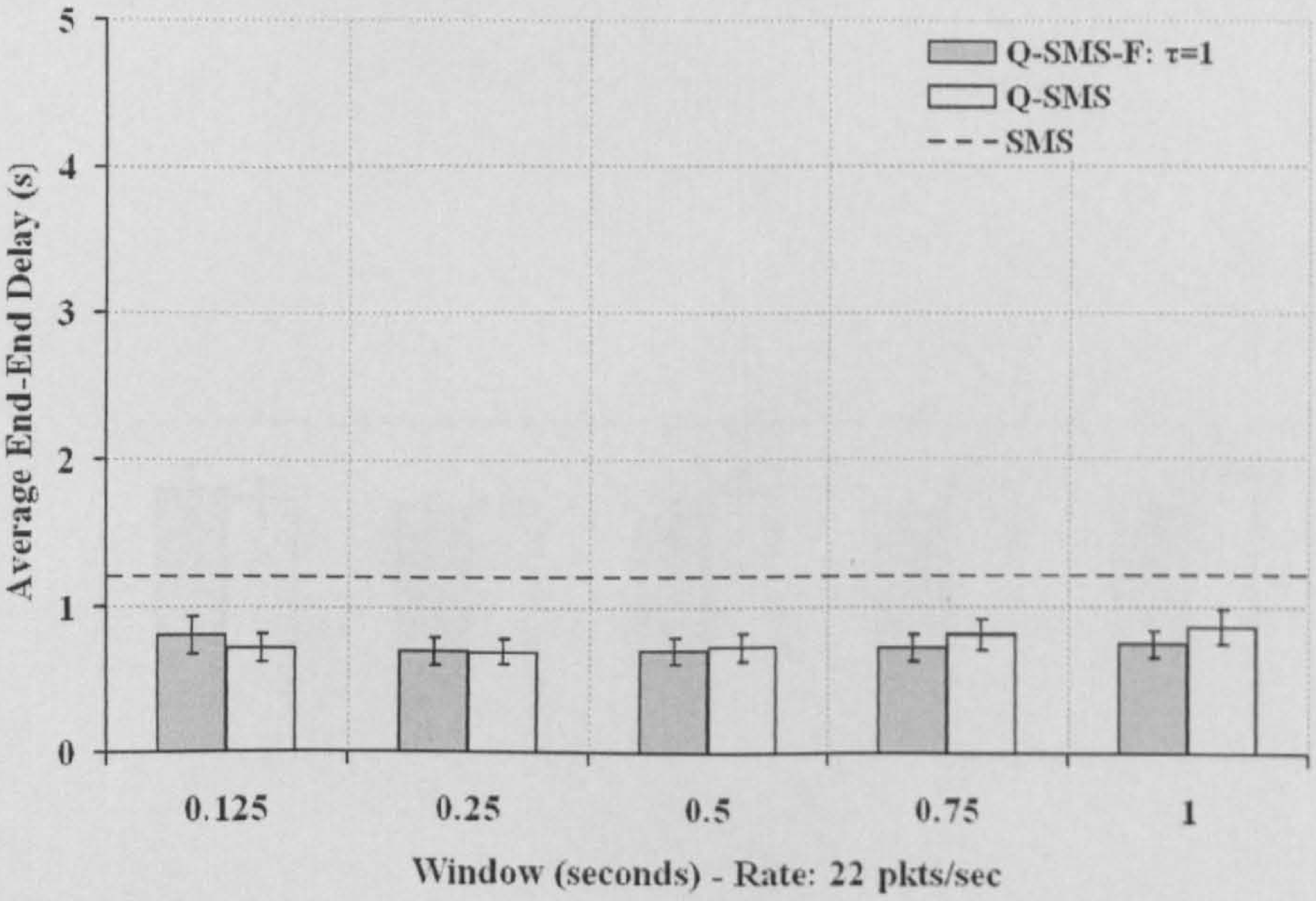


(b) Average end-to-end delay versus window size

Figure 6.15 Results for Q-SMS-F with different w – CBR sources at 22 pkts/sec

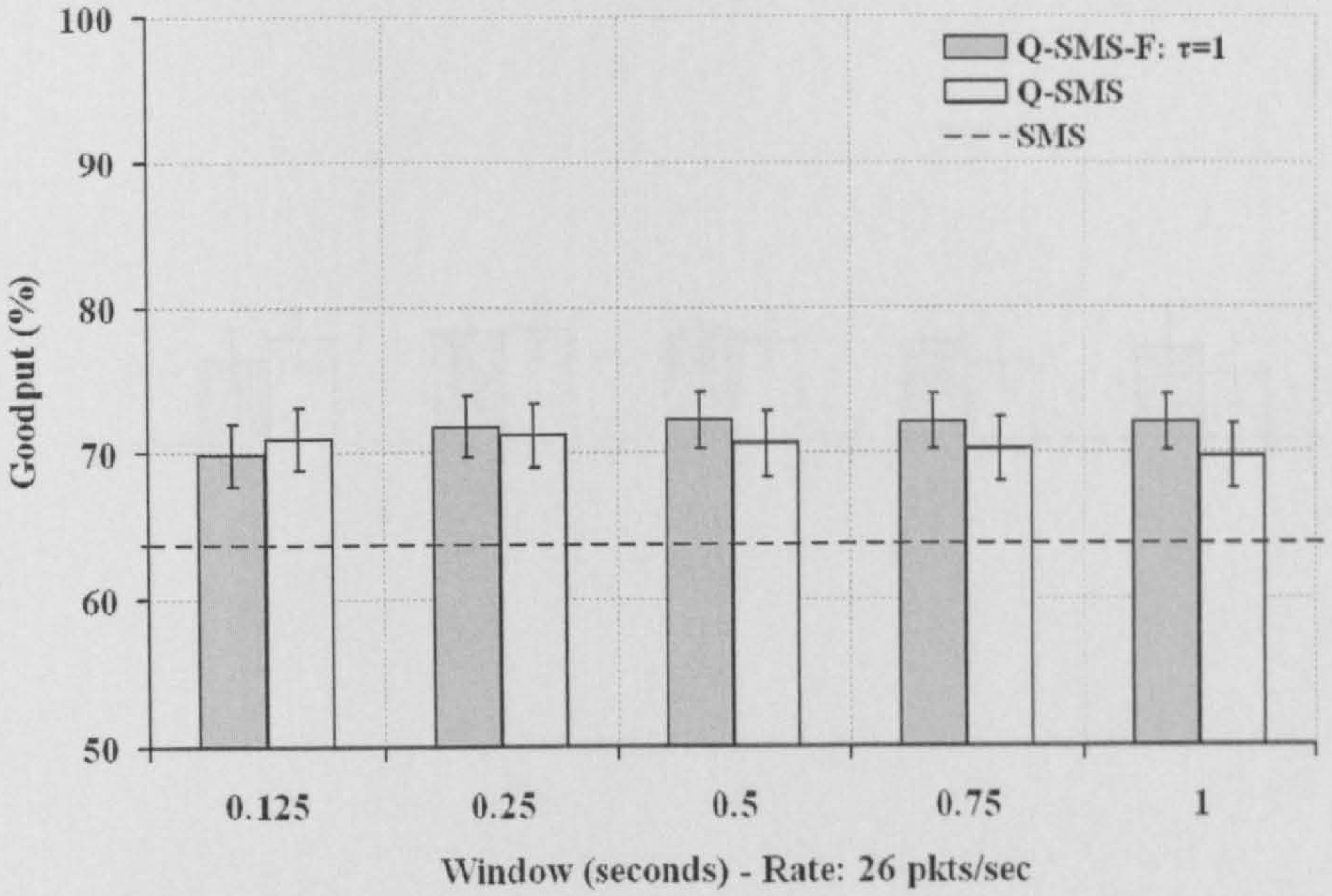


(a) Goodput versus window size

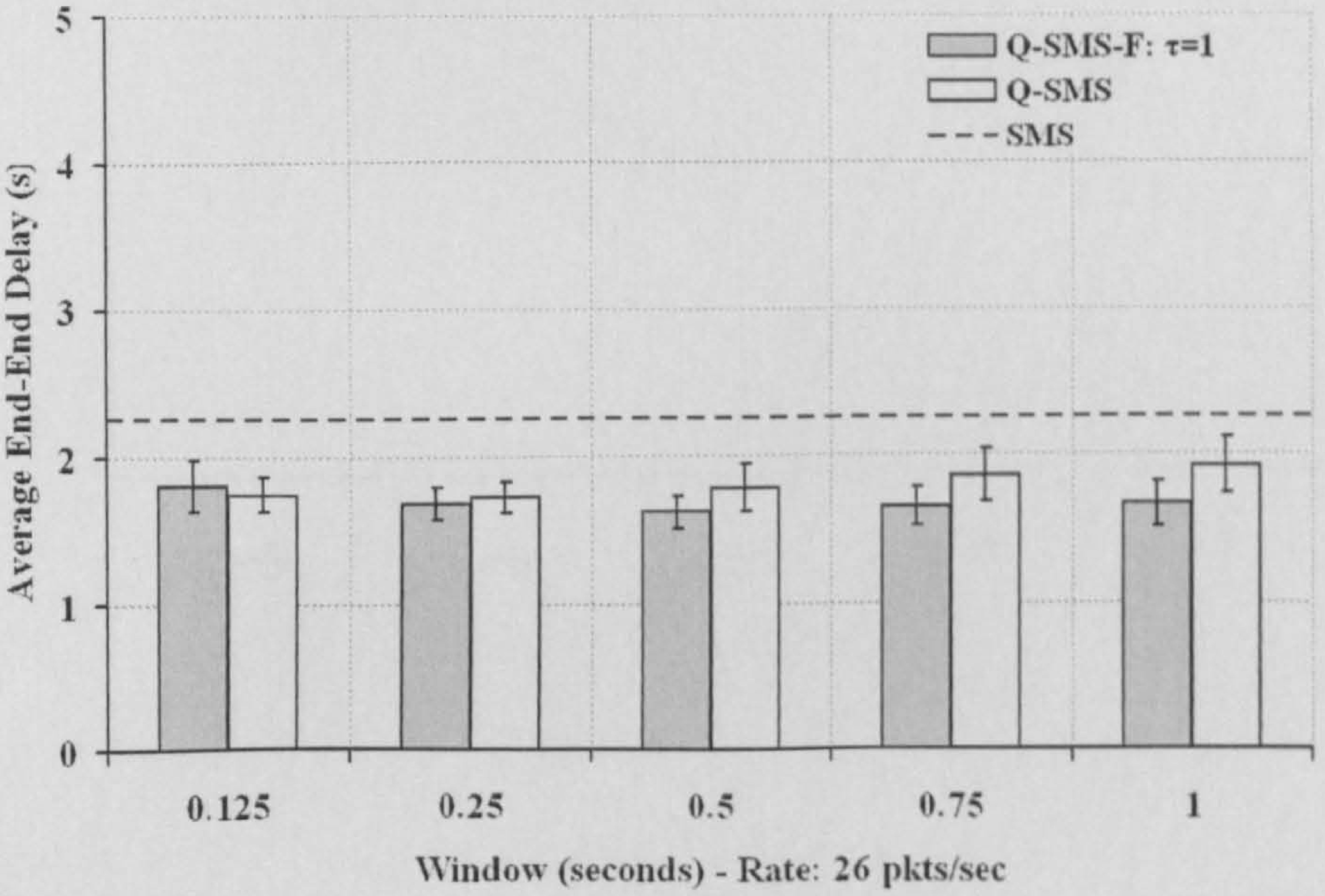


(b) Average end-to-end delay versus window size

Figure 6.16 Results for Q-SMS-F with different w – VBR sources at 22 pkts/sec

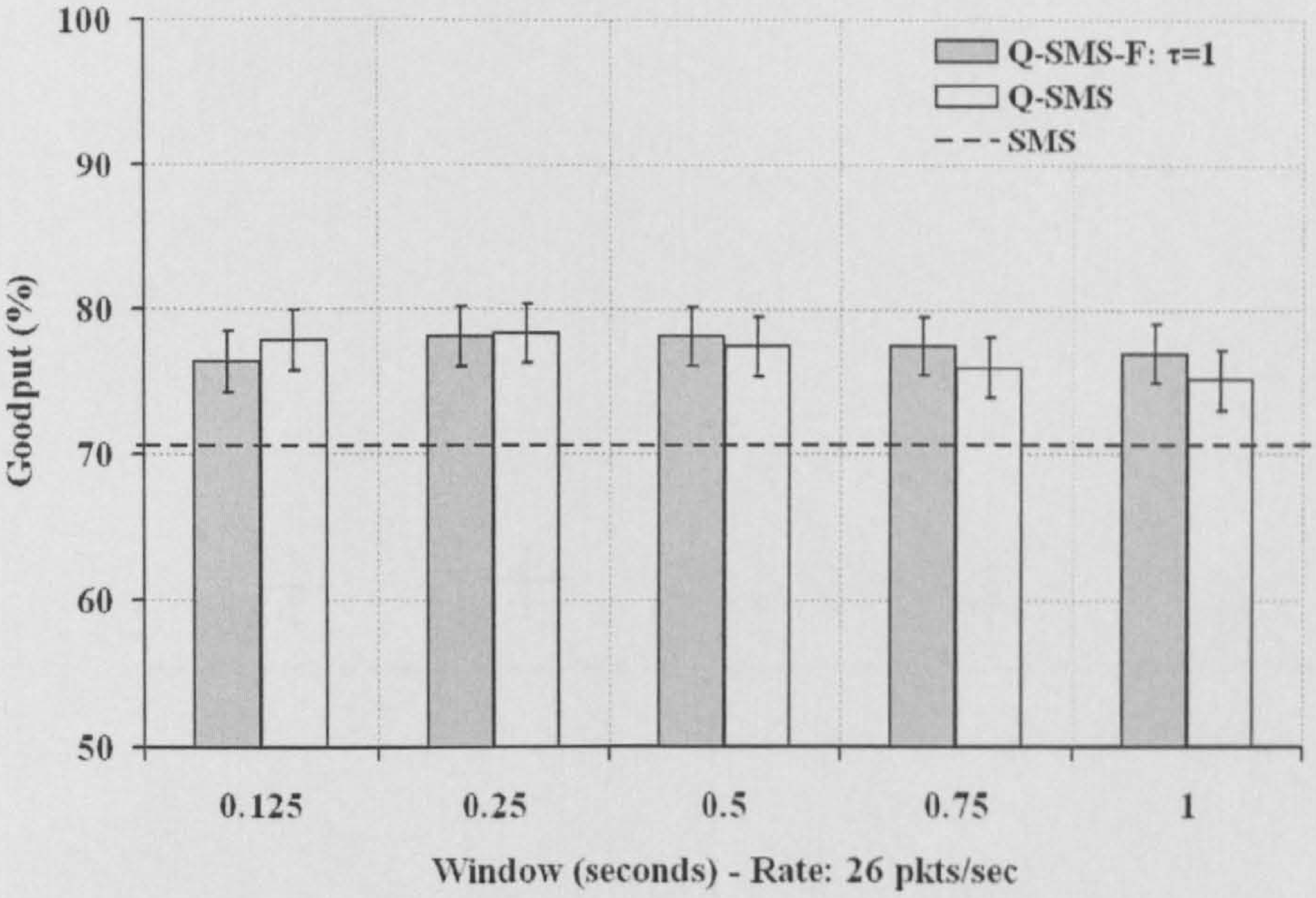


(a) Goodput versus window size

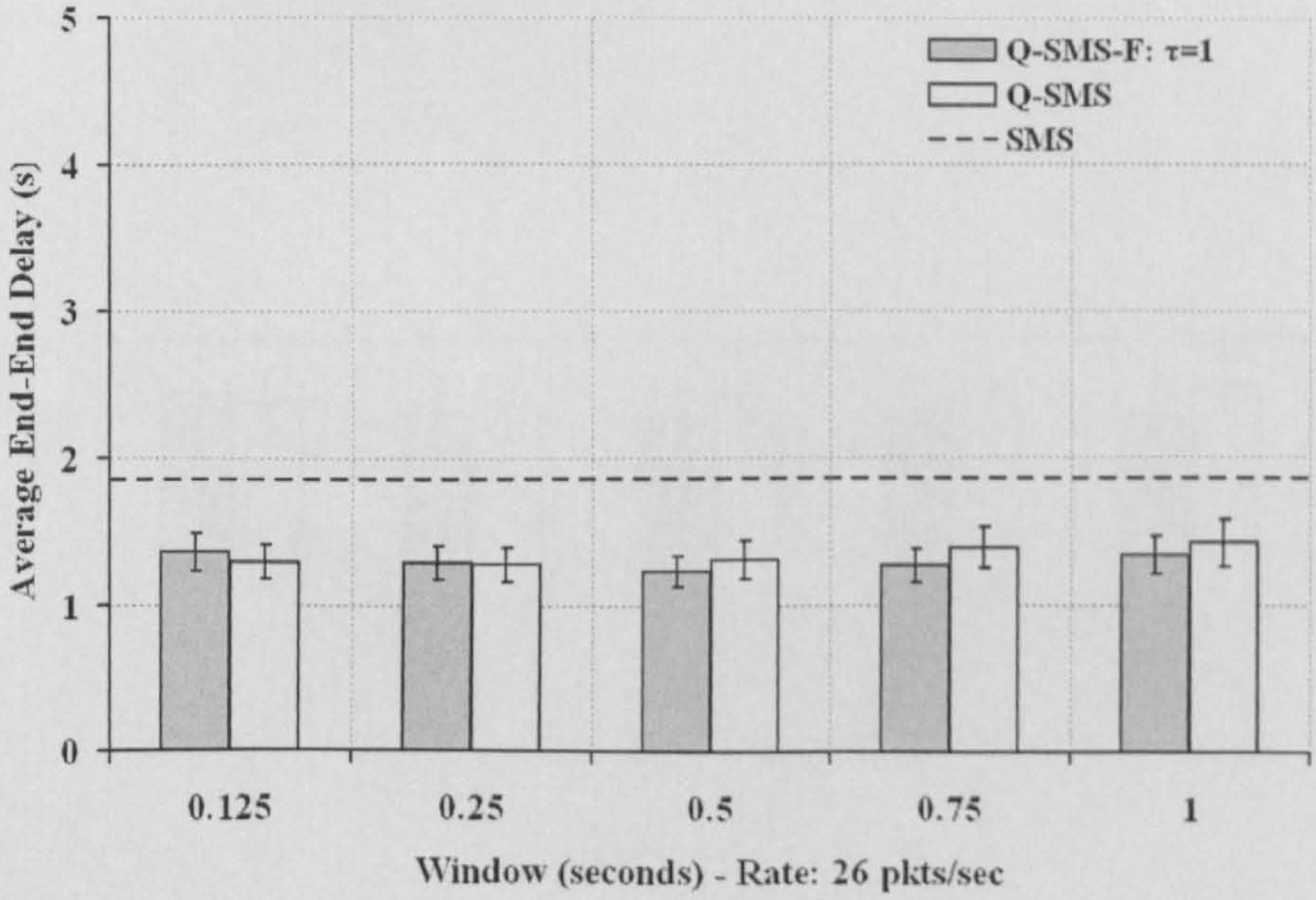


(b) Average end-to-end delay versus window size

Figure 6.17 Results for Q-SMS-F with different *w* – CBR sources at 26 pkts/sec

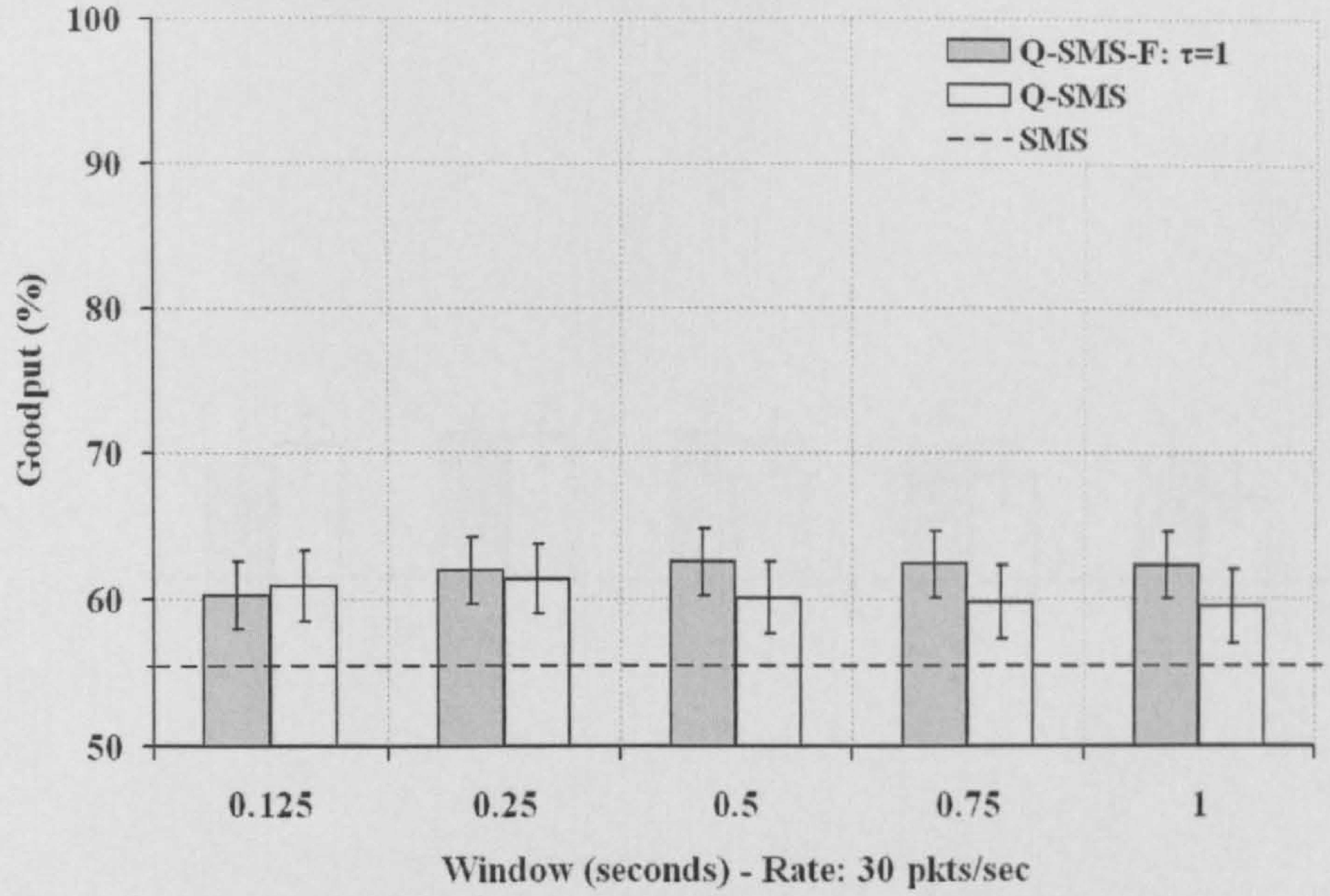


(a) Goodput versus window size

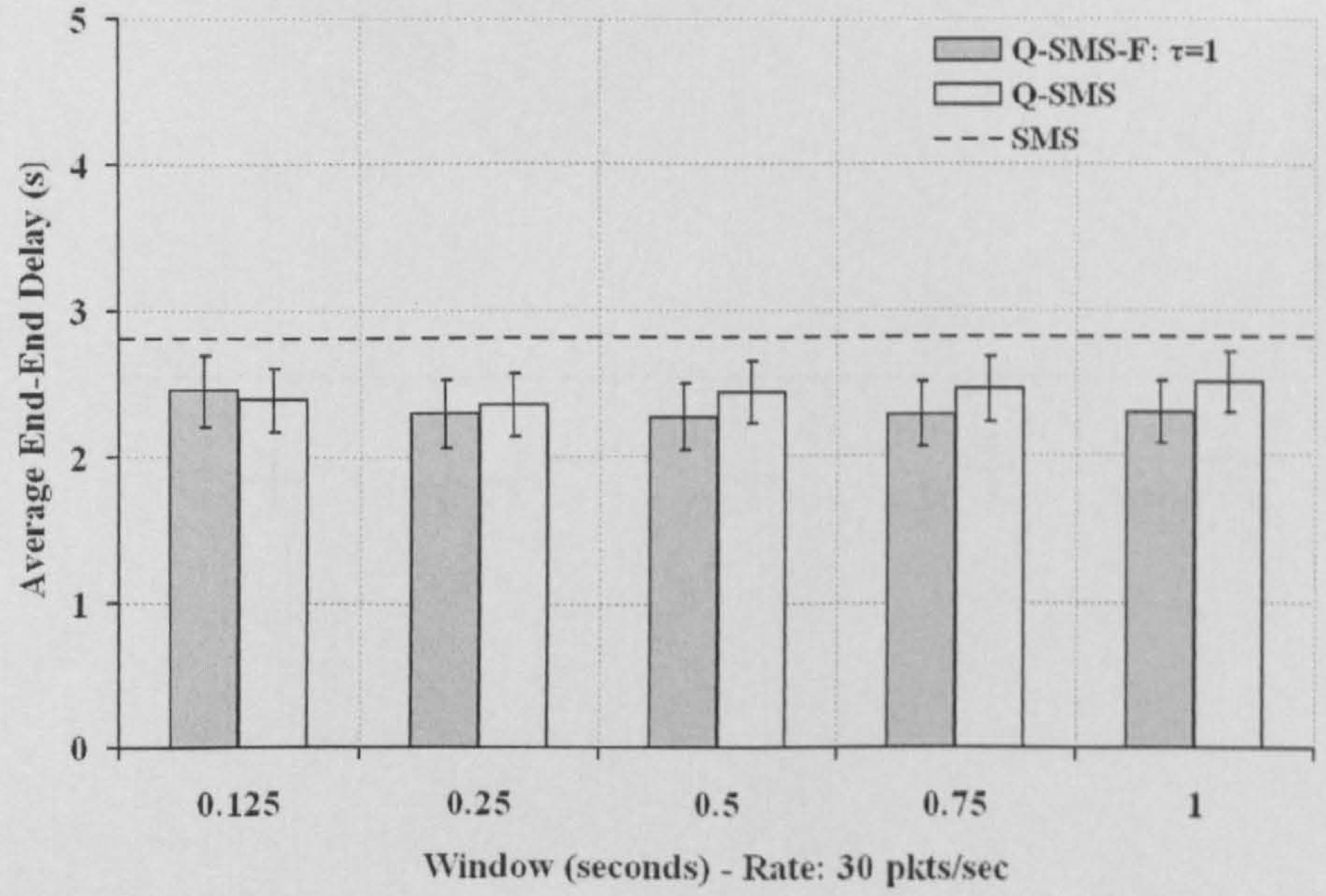


(b) Average end-to-end delay versus window size

Figure 6.18 Results for Q-SMS-F with different *w* – VBR sources at 26 pkts/sec

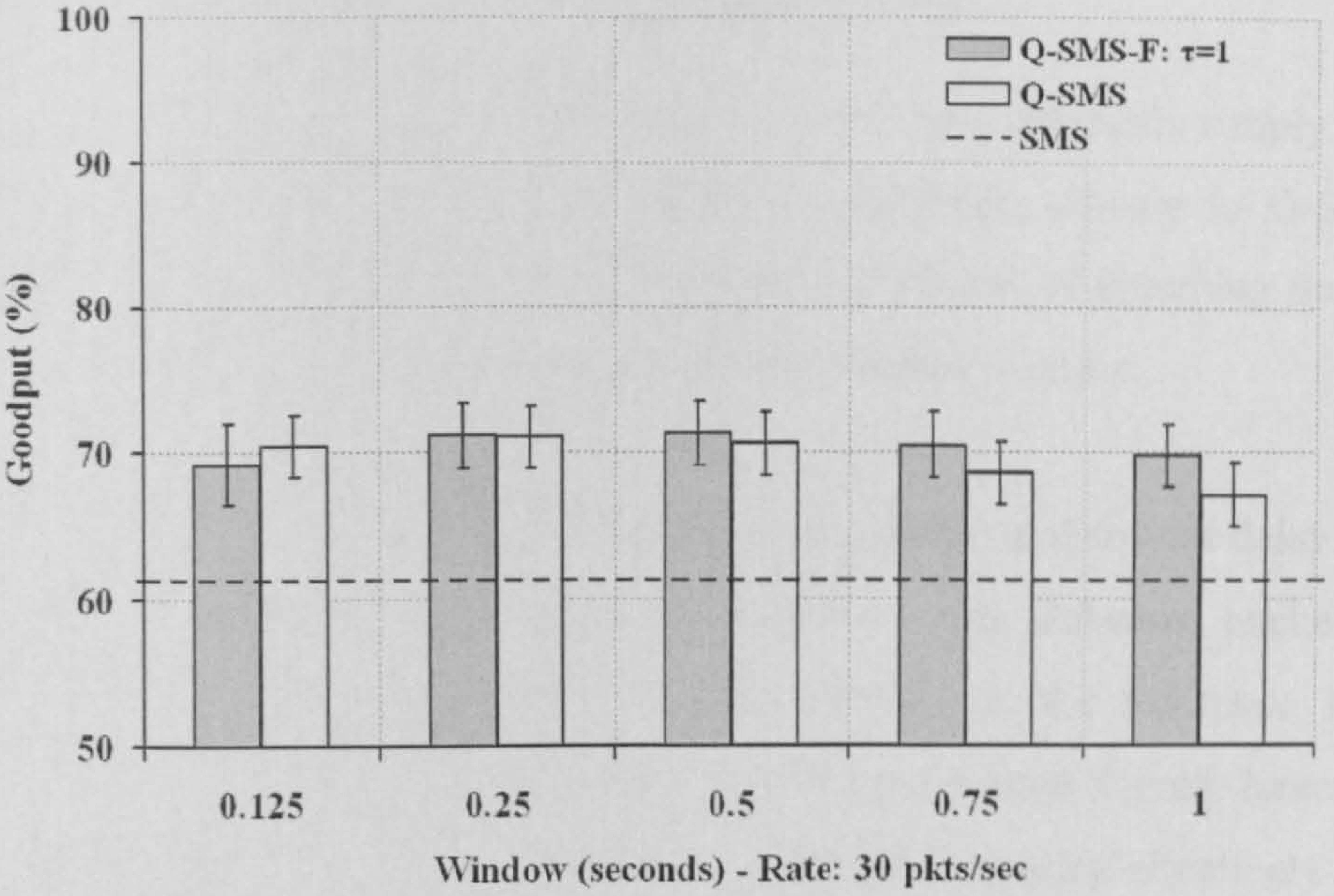


(a) Goodput versus window size

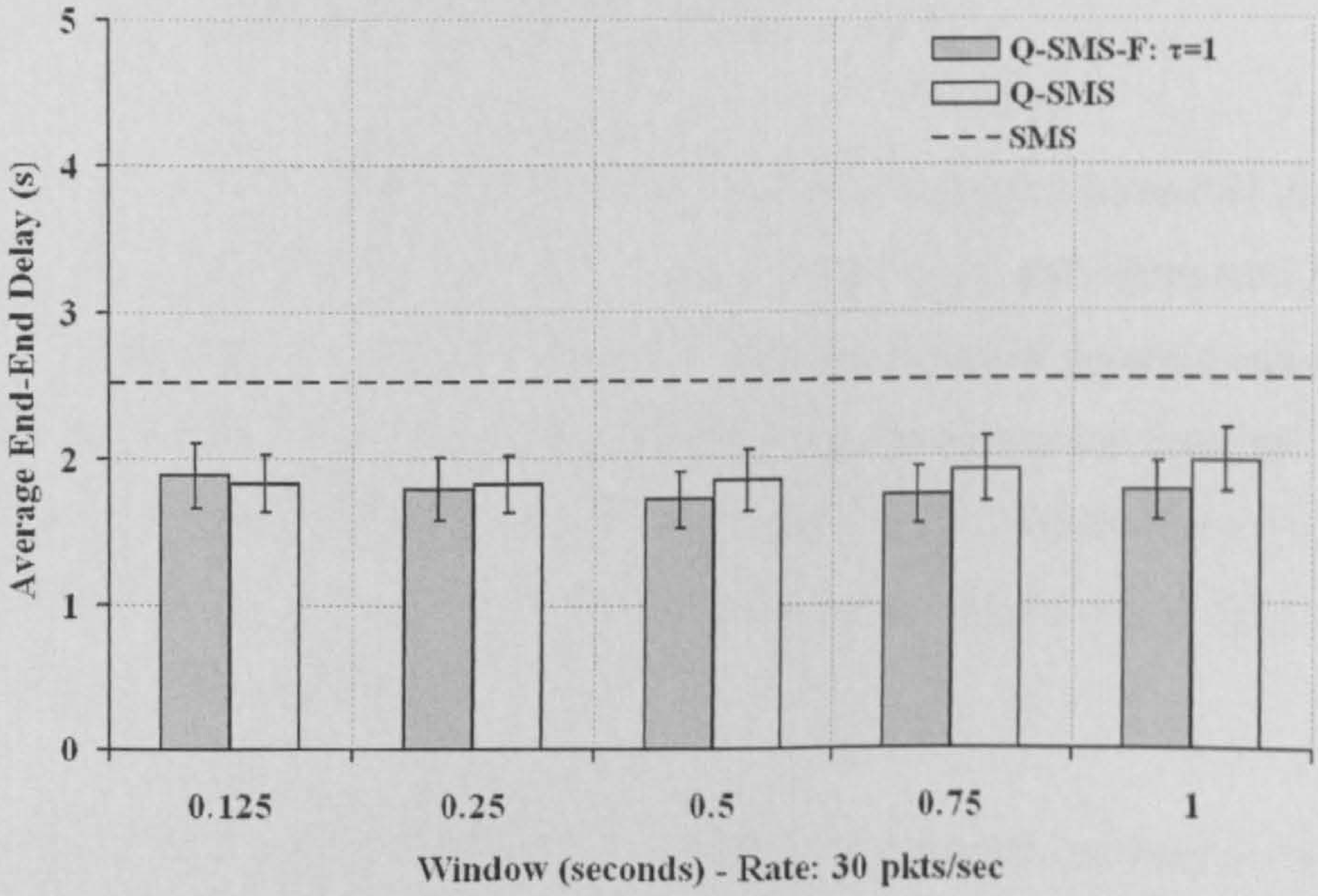


(b) Average end-to-end delay versus window size

Figure 6.19 Results for Q-SMS-F with different *w* – CBR sources at 30 pkts/sec



(a) Goodput versus window size



(b) Average end-to-end delay versus window size

Figure 6.20 Results for Q-SMS-F with different *w* – VBR sources at 30 pkts/sec

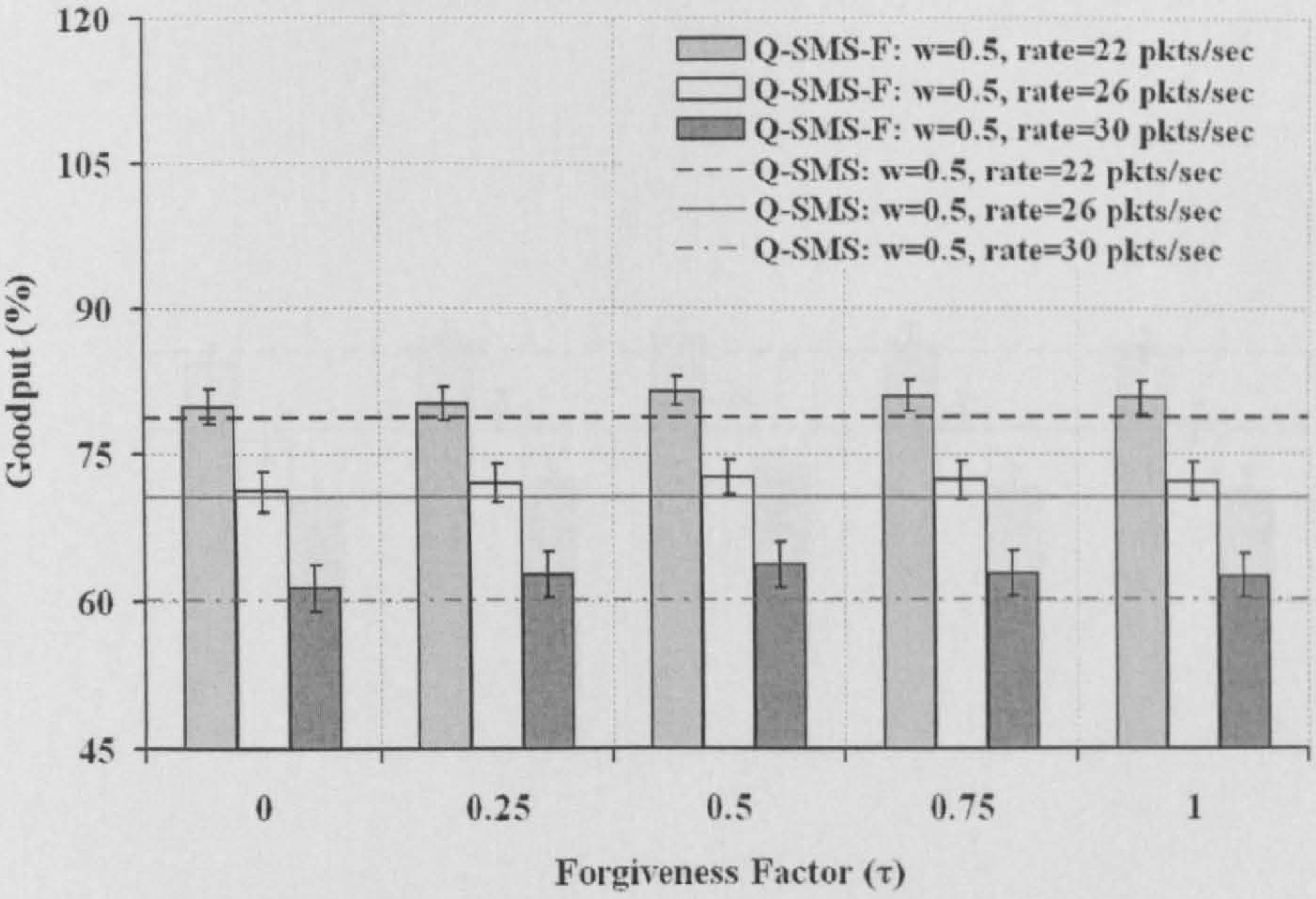
6.5.2 Impact of varying the Forgiveness Factor

In the previous section, a forgiveness factor (τ)=1 was adopted; simply to show the general applicability of applying the forgiven estimation scheme to Q-SMS. Having shown merit, it is then appropriate to consider the effects of changing the forgiveness factor. The aim is to find an appropriate balance between w and τ.

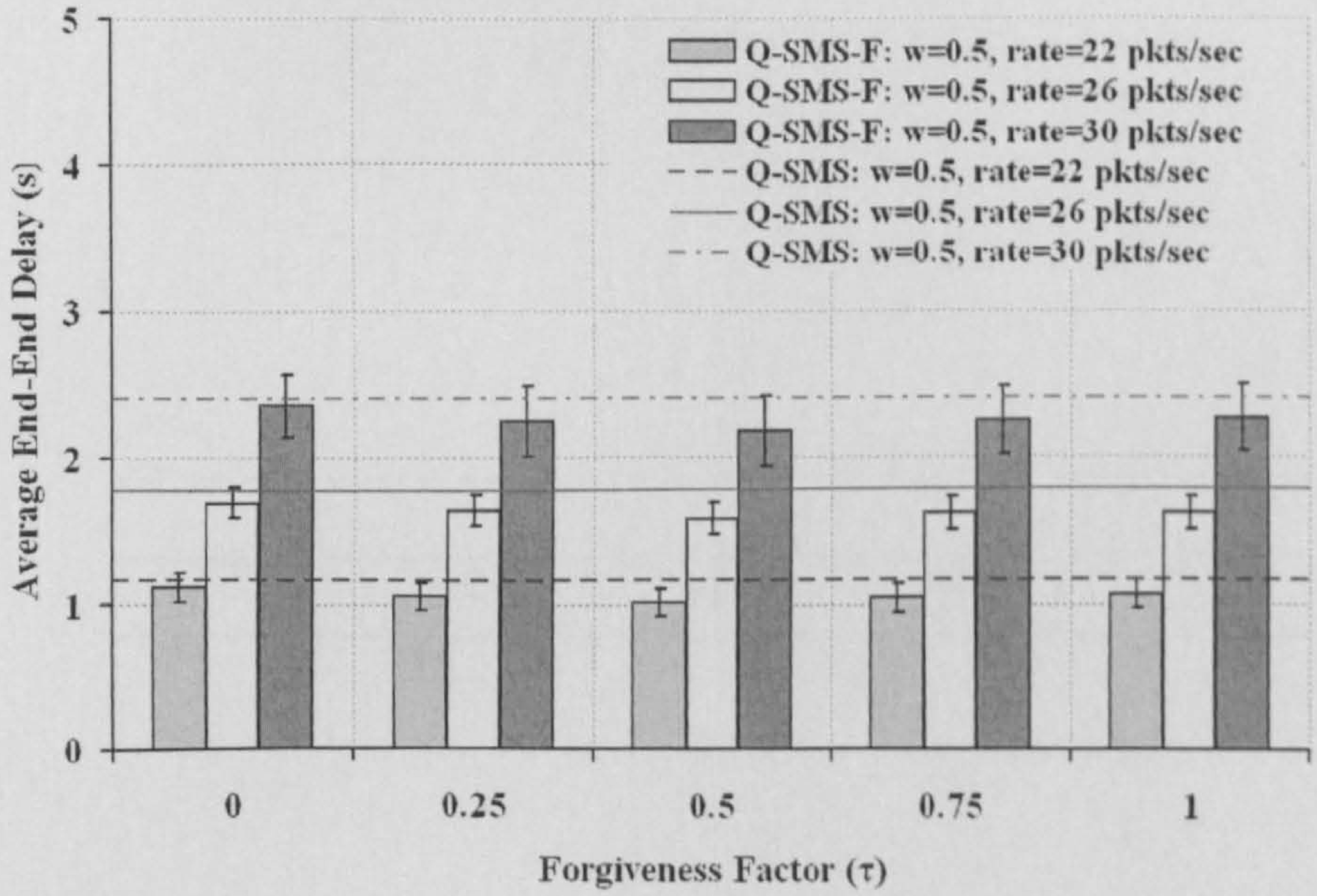
Figure 6.21 and Figure 6.22 shows goodput and average end-to-end delay verses τ for w=0.5 when considering CBR and VBR sources with different packet rates. The horizontal lines on the graph show Q-SMS (w=0.5) as the reference scheme with corresponding packet rates. The Q-SMS-F at τ=0.5 performed slightly better in terms of goodput and average end-to-end delay as the estimated capacity effectively allowed fast reactions to load variation and to nodes mobility. The performance of Q-SMS-F for $0.5< \tau \leq 1$ is slightly lower because it does not completely consider the current traffic for residual capacity estimation as illustrated in Figure 6.13.

Figure 6.23 and Figure 6.24 shows the goodput and average end-to-end delay against τ for w=1 with CBR and VBR traffic sources respectively and compared with Q-SMS (w=1). Here, Q-SMS-F at τ=0.75 performs better in terms of goodput and average end-to-end delay. There is no significant difference in performance between Q-SMS-F at w=0.5 and τ=0.5 and Q-SMS-F at w=1 and τ=0.75 with Q-SMS-F at w=0.5 and τ=0.5 having slightly better performance. Hence, there is an impact of the forgiveness factor with respect to the window size.

It is evident from the simulation results Q-SMS-F is a simple mechanism that has lower reliance upon window size and it can offer improvement when compared with Q-SMS in the changing environment of MANETs.

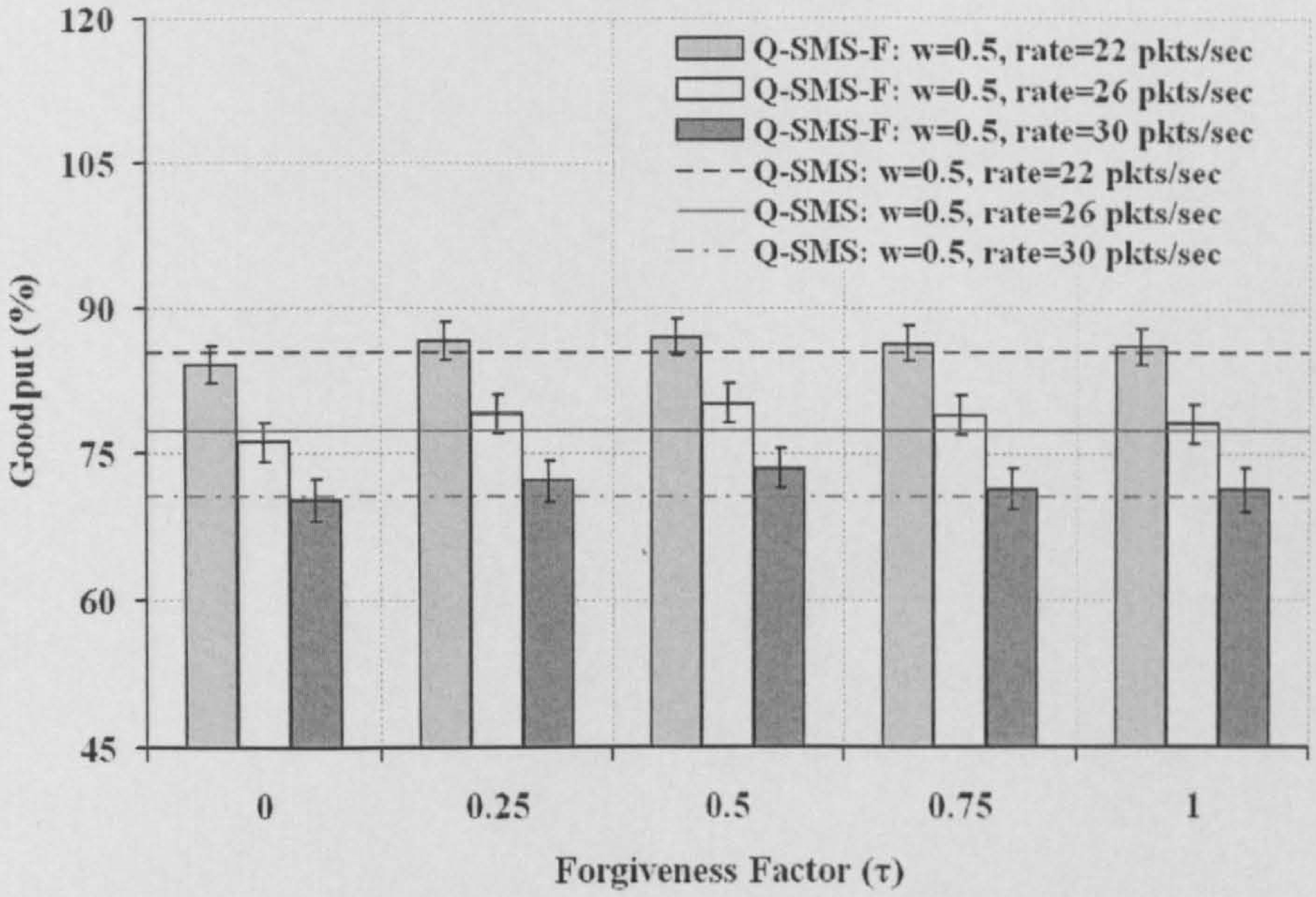


(a) Goodput versus forgiveness factor (*w*=0.5)

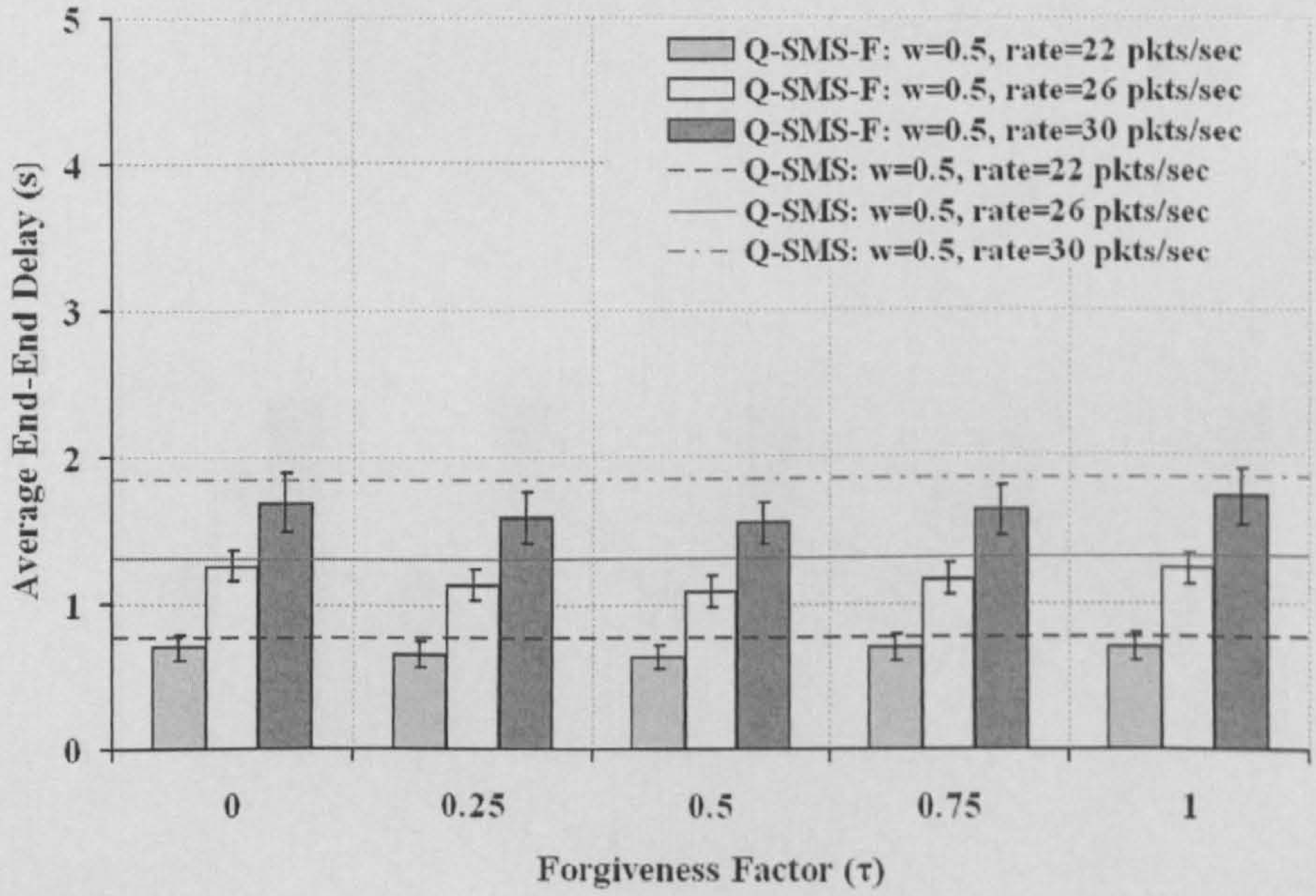


(b) Average end-to-end delay versus forgiveness factor (*w*=0.5)

Figure 6.21 Results for Q-SMS-F with different τ at w=0.5 – CBR sources

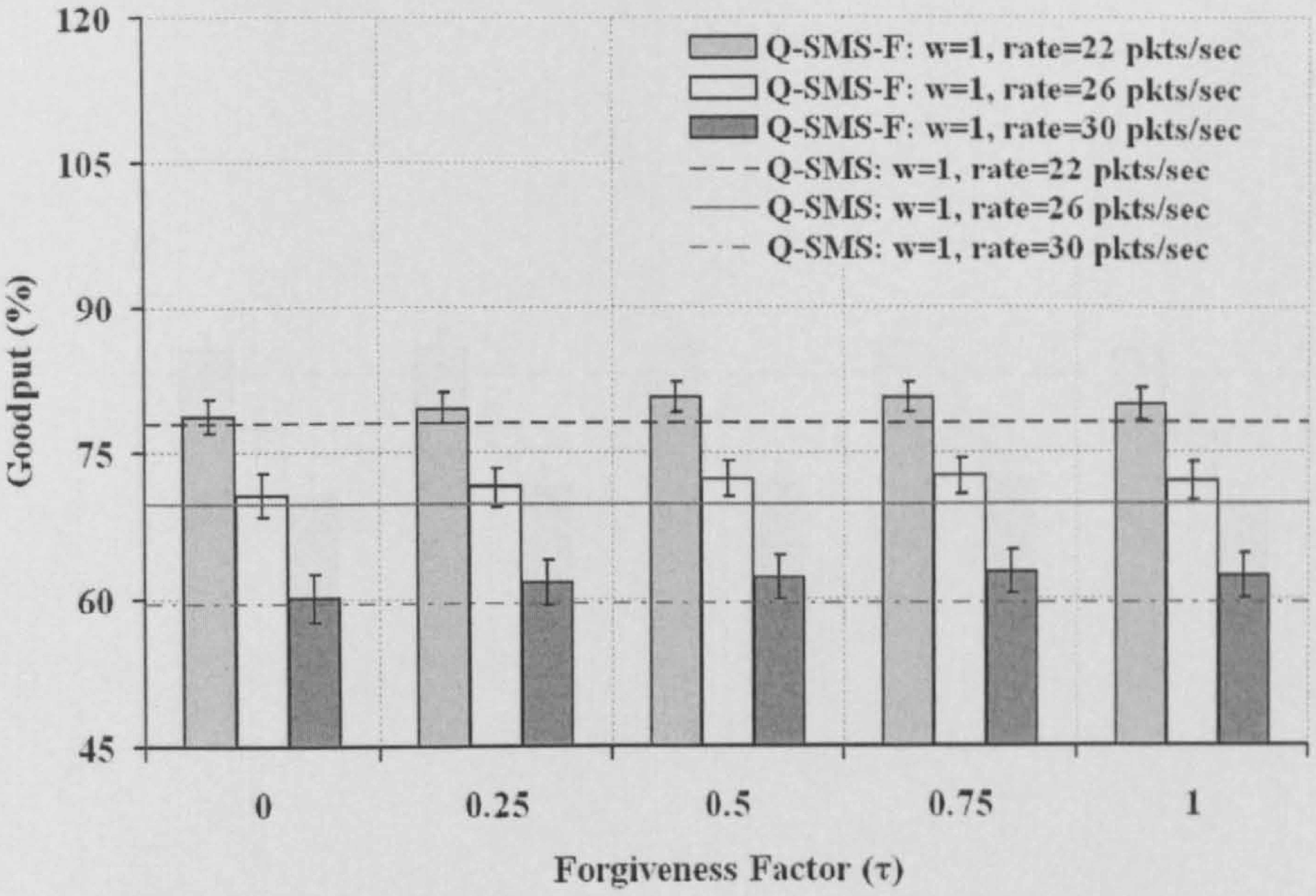


(a) Goodput versus forgiveness factor (*w*=0.5)

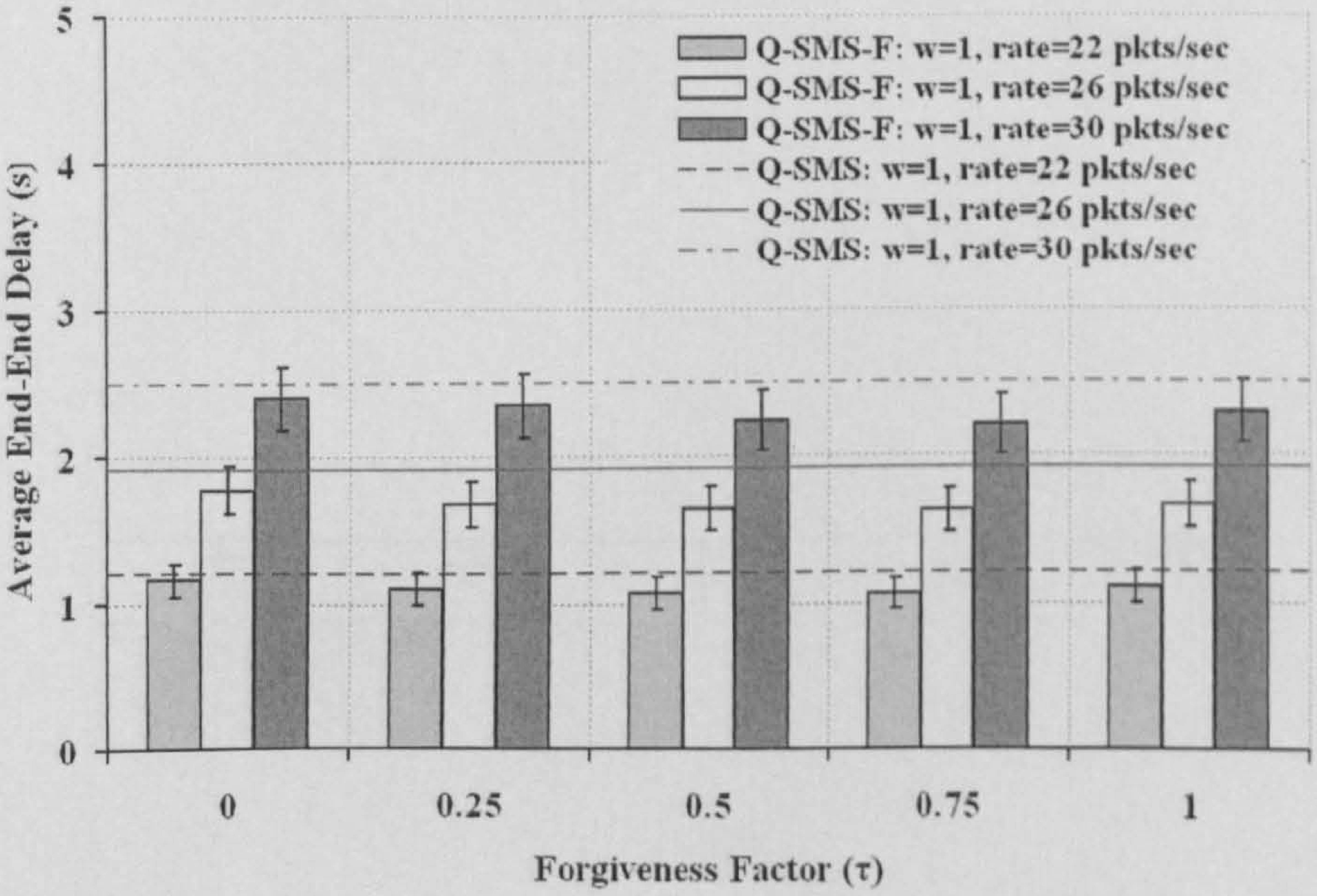


(b) Average end-to-end delay versus forgiveness factor (*w*=0.5)

Figure 6.22 Results for Q-SMS-F with different τ at w=0.5 – VBR sources

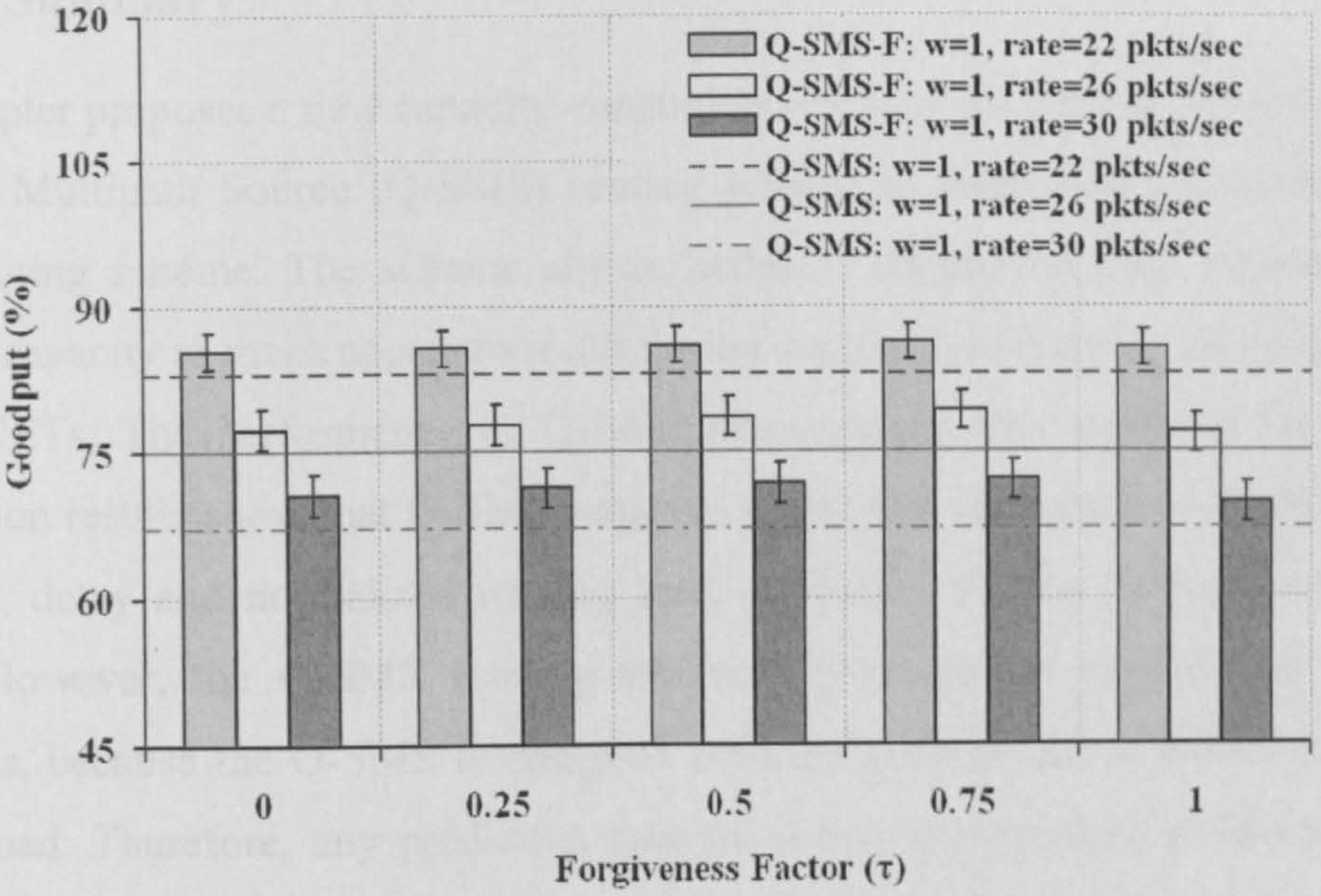


(a) Goodput versus forgiveness factor (*w*=1)

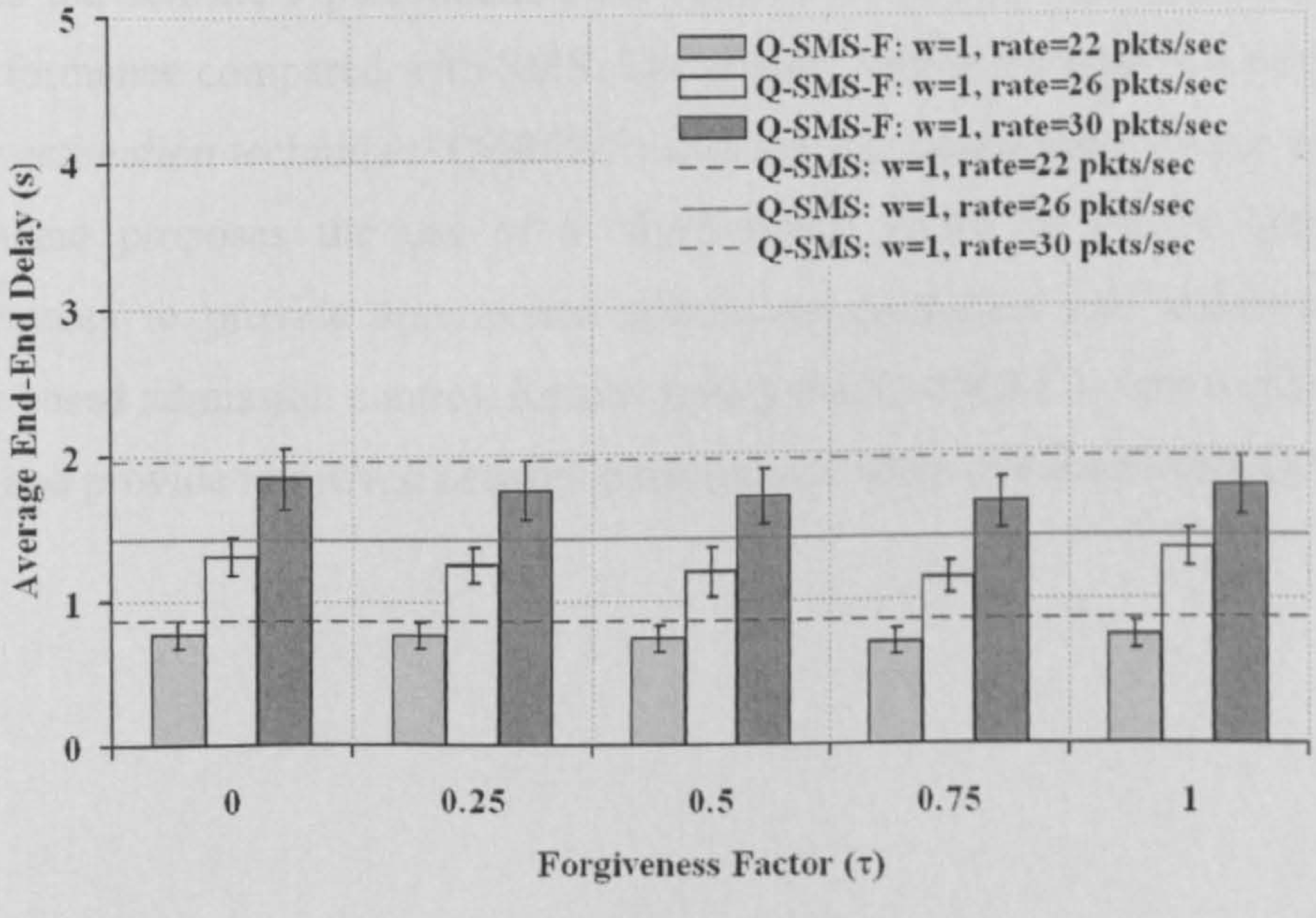


(b) Average end-to-end delay versus forgiveness factor (*w*=1)

Figure 6.23 Results for Q-SMS-F with different τ at w=1 – CBR sources



(a) Goodput versus forgiveness factor (*w*=1)



(b) Average end-to-end delay versus forgiveness factor (*w*=1)

Figure 6.24 Results for Q-SMS-F with different τ at *w*=1 – VBR sources

6.6 Summary

This chapter proposes a new capacity-constrained QoS-aware routing scheme referred as Shortest Multipath Source (Q-SMS) routing scheme to overcome the shortcomings of SMS routing scheme. The scheme allows nodes to depend on their estimation of the residual capacity to make appropriate admission control decisions to ensure QoS support in MANETs. The performance of Q-SMS demonstrates the merits of the algorithm. Simulation results show that Q-SMS achieves better performance than SMS in terms of goodput, delay and normalised routing load, especially at low to intermediate traffic loads. However, the Q-SMS routing scheme's performance degrades as the traffic increases, because the Q-SMS is designed with a simple model of low to intermediate traffic load. Therefore, any predictive scheme is not incorporated to find a new route before the previous route is broken. This results in a long transient time when the required QoS is not guaranteed, due to a route break or network partition, which decreases the scheme's performance. However, the Q-SMS routing scheme still gets high performance compared with SMS. The chapter further introduces a simple residual capacity estimation technique (Q-SMS-F) that has less dependence on the window size. The scheme proposes the use of a 'forgiveness' factor to weight these previous measurements to provide appropriate utilisations estimation and improved available capacity based admission control. Results justify that Q-SMS-F is less window sensitive scheme and provide improved network performance when compared with Q-SMS.

7. Conclusion and Future Work

Mobile ad hoc networks (MANETs) are rapidly evolving and the notion of ad hoc networking has become one of the most challenging research areas of wireless communications. This type of network offers unique benefits and greater flexibility for a range of applications and situations, and therefore is going to play a major role in future leisure, commercial and military scenarios. However, this comes at the cost of overcoming extremely difficult challenges posed mainly due to the absence of fixed infrastructure, capacity constrained operation and random mobility patterns of nodes, which are unique to such networks. Irrespective of the open challenges, the rising popularity of real-time applications among end users in MANETs has stimulated a surge of research in routing and providing QoS support in such networks. However, despite recent advances in the development of algorithms and protocols specifically designed to address the challenges in routing and QoS support, a large research space remains open for further exploration. The work described in this thesis has demonstrated the advantages of multipath routing and sharing information between MAC and routing layers of the traditional OSI protocol stack to support QoS in MANETs.

7.1 Summary and Conclusion

Chapter 3 is divided into two parts. The first part presented the concept of multipath routing and described a number of multipath routing schemes proposed for MANETs. Several areas in multipath routing were identified to improve the performance in terms of delay bounds or guarantees and lower routing overheads which form the basis of work presented in this thesis. The second part of the chapter discussed in detail the evaluation methodology and models adopted in this thesis to carry out the performance evaluation of the routing schemes with accurate results and conclusions. A number of simulation-based experiments for Dynamic Source Routing (DSR) and Split Multipath Routing (SMR) schemes are presented to validate the model behaviour and assumptions. This work ensured that such an implementation mirrored as far as possible the operation and performance of different models.

In Chapter 4, a novel and practical Shortest Multipath Source (SMS) routing scheme to overcome the limitations of existing on-demand unipath and multipath routing schemes described in Chapter 2 and 3 respectively is proposed. The advantages of the proposed routing scheme are a) its simplicity, b) it reduces the wireless broadcast storm by simple hop count mechanism and, c) it uses shortest alternative partial-disjoint paths to reduce delays, overhead of additional route discoveries and recover quickly in case of route breaks. A comprehensive performance evaluation of the SMS scheme is performed using NS-2 based model and its merits are illustrated over competing and directly compatible schemes. Simulation results reveal that SMS outperformed DSR and SMR under wide range of scenarios with significant improvements in terms of goodput, delay and routing efficiency, especially at intermediate loads. However, at high mobility or traffic loads, the performance of SMS degrades due to traffic congestion which in turn results in the bottleneck of the node. Hence, this thesis continues to propose an effective QoS provisioning mechanism to further improve the performance of the SMS scheme.

In Chapter 5, a survey on existing QoS-aware routing schemes in MANETs is presented. The chapter highlighted key routing design issues to support QoS in MANETs. Motivated by such issues, a novel and effective QoS-aware Shortest Multipath Source (Q-SMS) routing scheme is proposed in Chapter 6. The scheme modifies and extends SMS routing scheme to allow nodes to depend on their estimation of the residual capacity to make accurate admission control decisions to ensure QoS support in MANETs. The capacity estimation is based on the measurement of local channel utilisation over a time window. It was demonstrated that Q-SMS scheme achieves better performance over SMS in terms of goodput, delay and normalised routing load, especially at low to intermediate traffic loads. However, the Q-SMS routing scheme's performance degrades as the traffic increases which in turn results in a long transient time when the required QoS is not guaranteed, due to a route break or network partition, which decreases the scheme's performance.

The next section of the chapter studied how varying the window size affects the network performance. Figure 7.1 summarises the impact of varying window size depicting the ratio of the average delay of Q-SMS to SMS average delay for average packet rate (between 22 and 30 packets per second). Slightly better performance was observed for a small window (*w*=0.25) and relatively low performance was recorded for large windows for the reasons explained in Section 6.3.5.

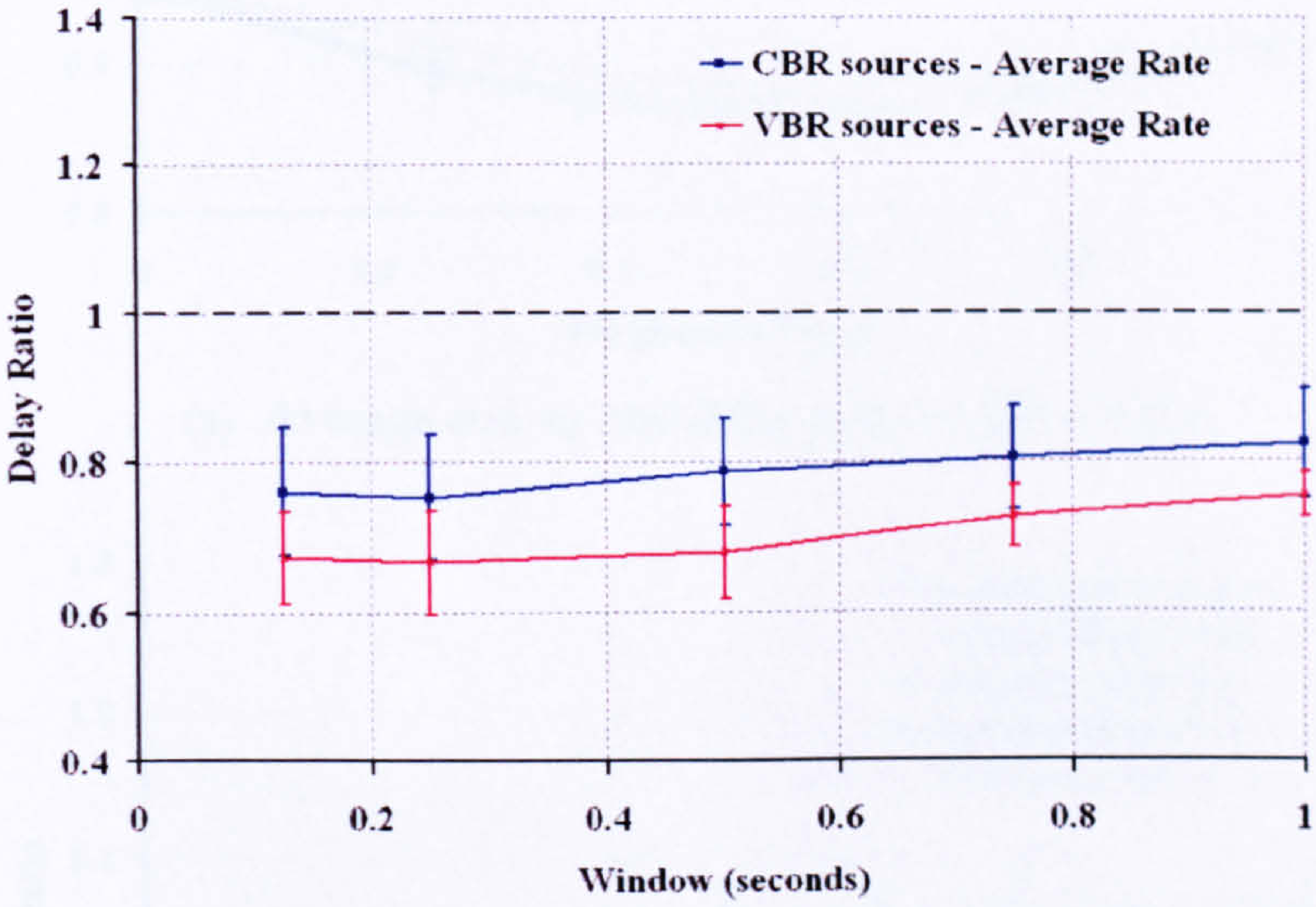
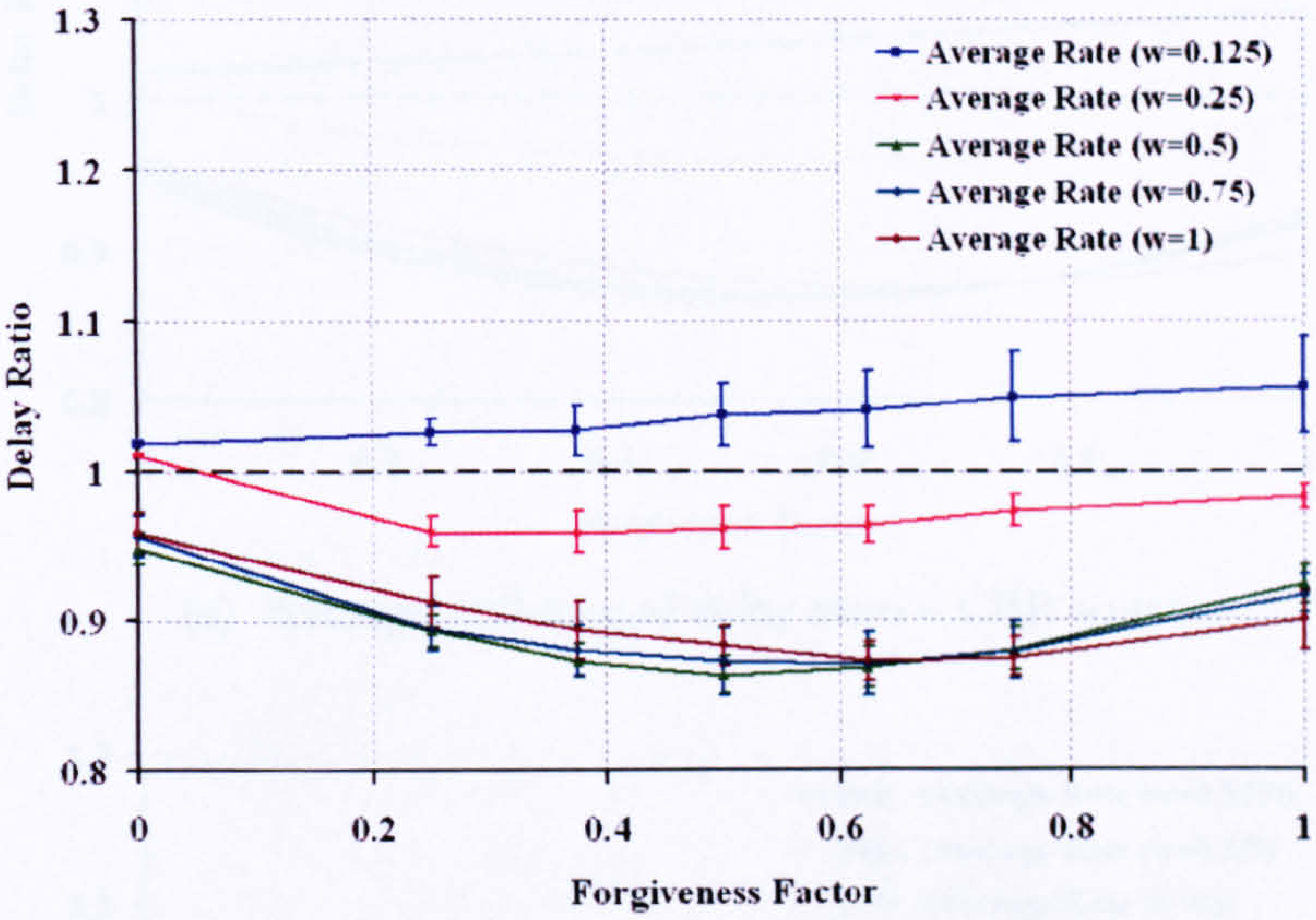
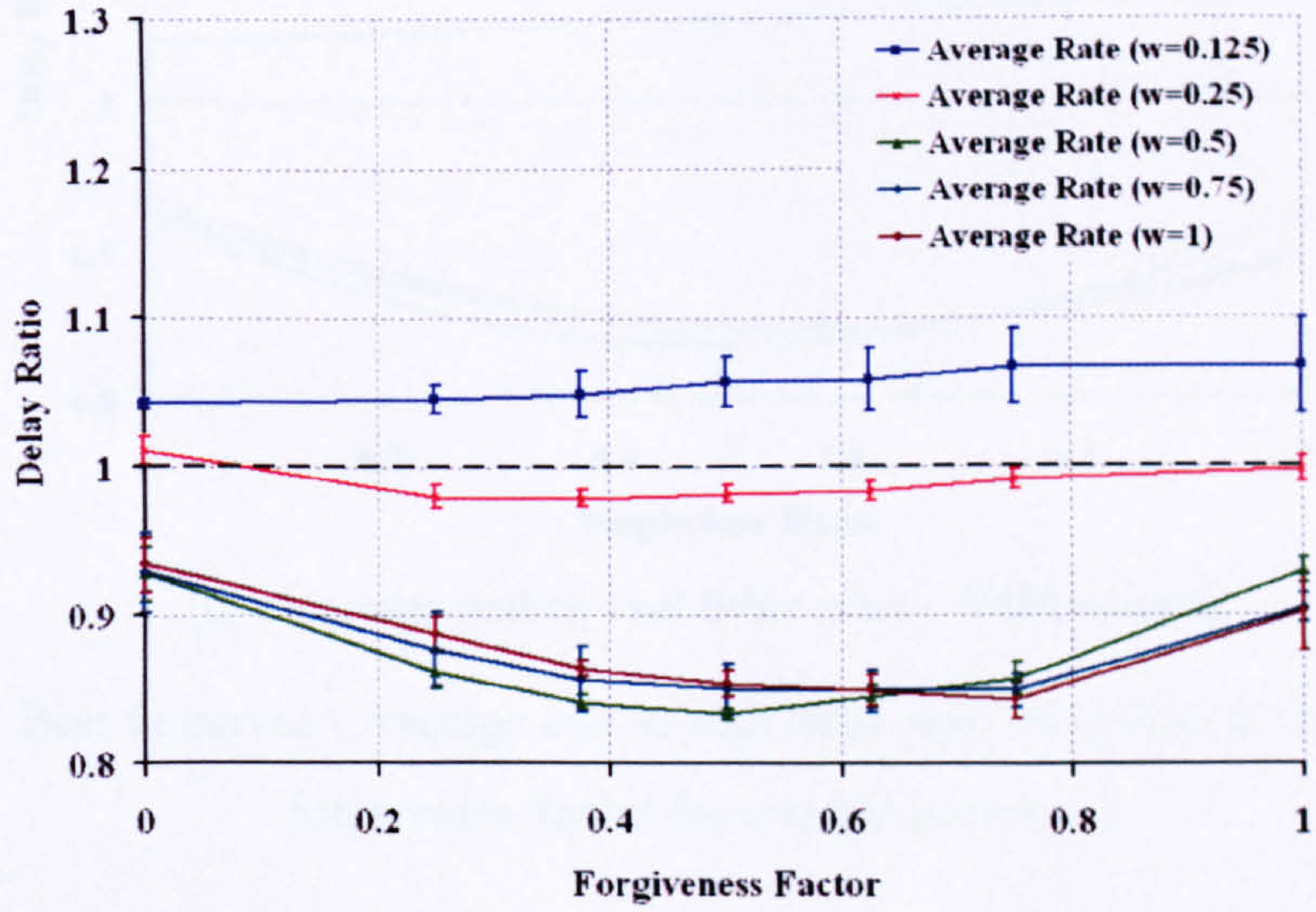


Figure 7.1 Average end-to-end delay ratio of Q-SMS to SMS versus window size for average packet rate

In the final part of the work, a 'forgiven' capacity estimation method (Q-SMS-F) was presented. The main idea of this scheme is the use of a 'forgiveness' factor to weight utilisation measurements so that greater emphasis can be placed on recent information rather than the past information within a time window in order to overcome the drawbacks of the previous window based schemes to improve the network performance. Figure 7.2 summarises the impact of varying the forgiveness factor for different window sizes discussed in Section 6.5.2 and shows the relative delay performance of the network of the two schemes whereby delay ratio is simply the ratio of the average delay of Q-SMS-F compared to Q-SMS average delay for average packet rate. The graph is further plotted in terms of the best fit line as shown in Figure 7.3, in order to look at the general trend.

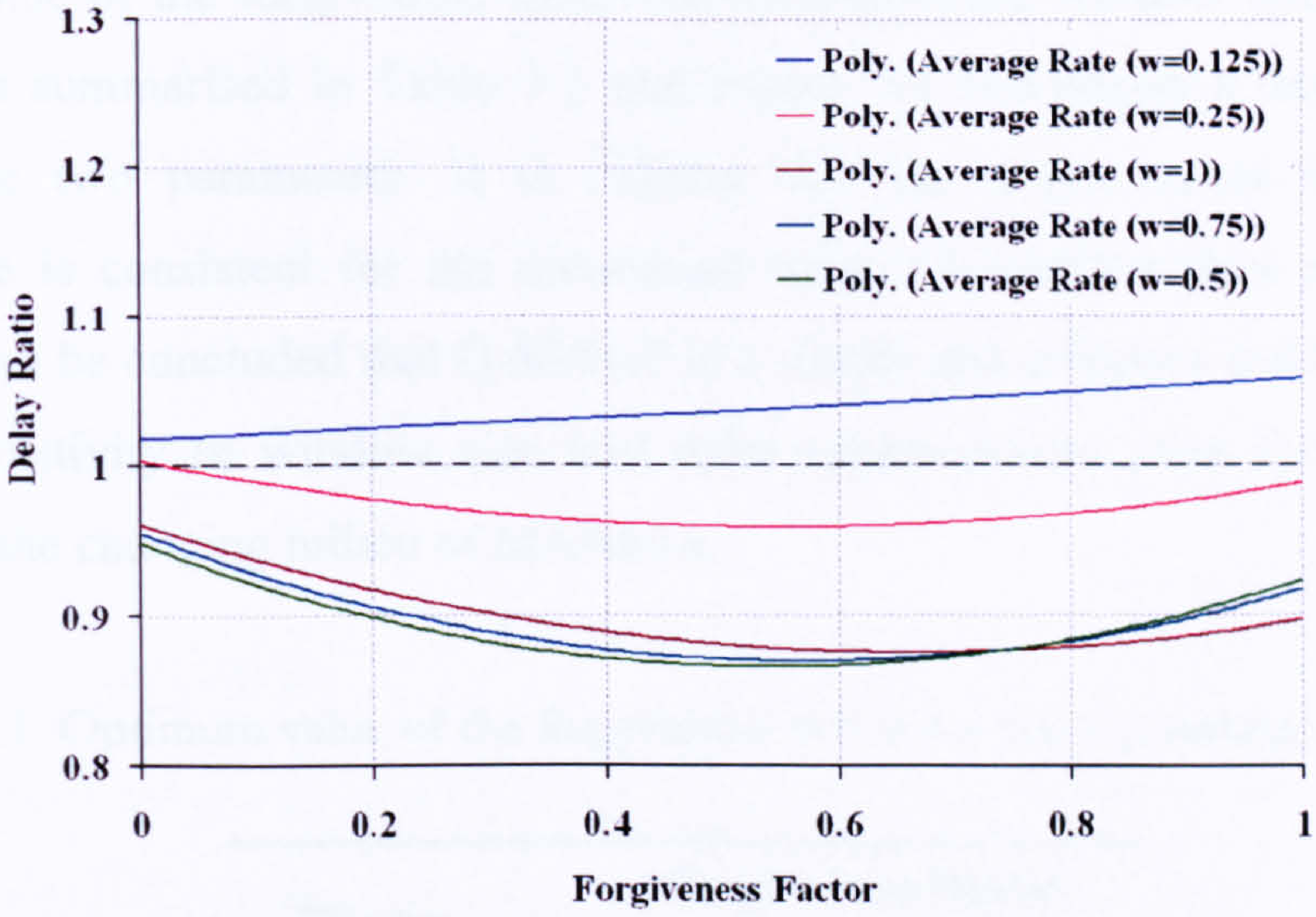


(a) Average end-to-end delay ratio – CBR sources

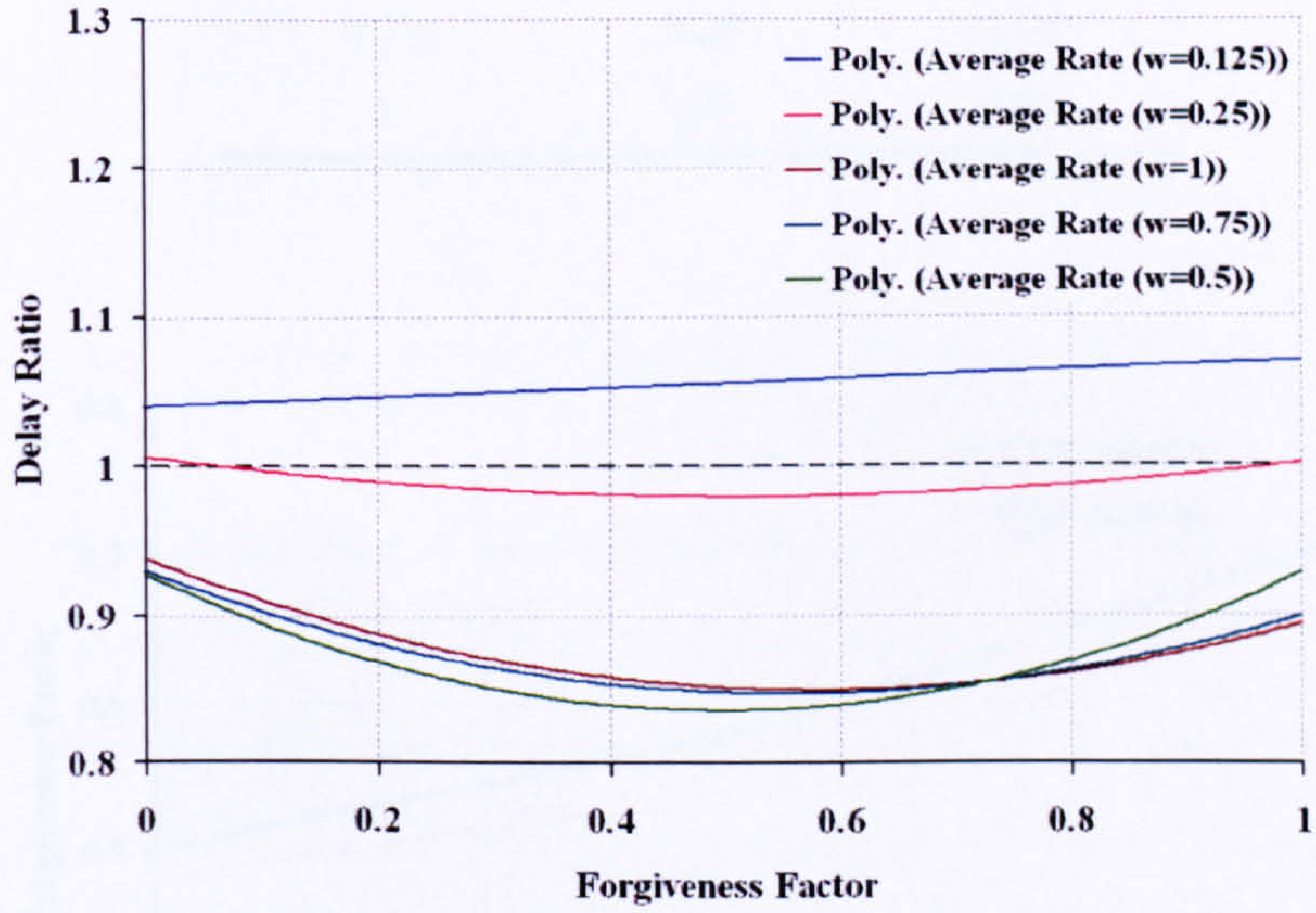


(b) Average end-to-end delay ratio – VBR sources

Figure 7.2 Average end-to-end delay ratio of Q-SMS-F to Q-SMS versus forgiveness factor for average packet rate



(a) Average end-to-end delay ratio – CBR sources



(b) Average end-to-end delay ratio – VBR sources

Figure 7.3 Best fit curves - Average end-to-end delay ratio of Q-SMS-F to Q-SMS versus forgiveness factor for average packet rate

From Figure 7.3, it can be clearly seen that the Q-SMS-F offers improvement for window sizes of *w*=0.5, 0.75 and 1 for both CBR and VBR sources. It was also observed that there is an impact of the forgiveness factor with respect to the window size. The optimum value of the forgiveness factor for corresponding window size depicted from the graph is summarised in Table 7.1 and Figure 7.4 and shows a linear relationship between the two parameters. It is evident that the improvement in the network performance is consistent for the mentioned range of window sizes and forgiveness factors. It can be concluded that Q-SMS-F is a simple and effective mechanism that has no high sensitivity to window size and offer improvements over Q-SMS and SMS schemes in the changing milieu of MANETs.

Table 7.1 Optimum value of the forgiveness factor for corresponding window size

Window	Forgiveness Factor	
	CBR sources	**VBR sources**
0.5	0.5	0.48
0.75	0.58	0.54
1	0.7	0.66

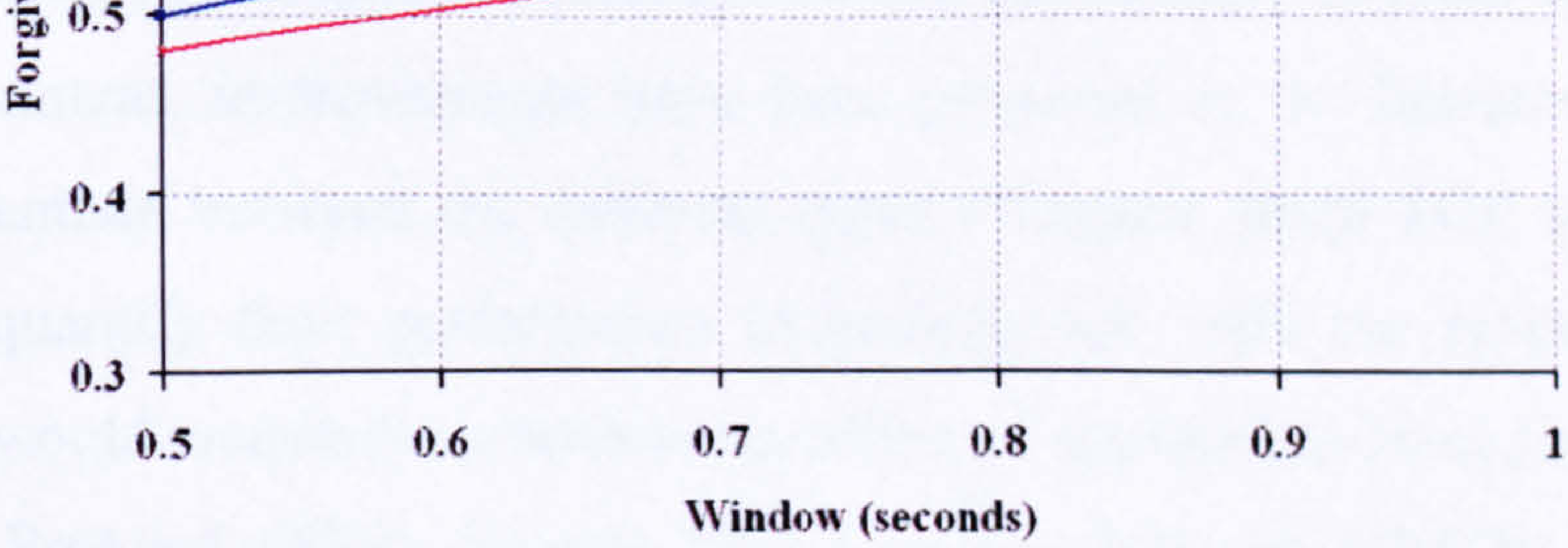


Figure 7.4 Relationship between window size and forgiveness factor

7.2 Future Work

As stated before, the thesis explored the problem of QoS provisioning in MANETs mainly from the perspective of the network layer and MAC sub-layer. However, it is insufficient to consider QoS merely at the network layer without considering higher and underlying layers. Overall, in order to support QoS for real-time applications in MANETs, each layer should provide appropriate QoS support [13][15][17]. The future directions for extensions in this area are:

7.2.1 Transport Layer

As real-time applications in MANETs typically use User Datagram Protocol (UDP), in this thesis only UDP performance is investigated. Although each scheme presented in this thesis works well with simple CBR and VBR traffic, it is appropriate to analyse their robustness with perhaps more complex Real-time Transport Protocol (RTP) for providing QoS in MANETs. RTP runs on top of the UDP and provides mechanisms for the sending and receiving applications to support streaming data.

It is important to also consider the performance of Transmission Control Protocol (TCP) in MANETs, since TCP is widely used today and the efficient integration of MANET with the Internet is paramount whenever possible. However, when TCP is applied to MANETs, it may result in degraded and suboptimal performance due to transmission related losses and hence invokes congestion control mechanism resulting in reduction of the link utilisation which in turn decrease network throughput. In order to adapt TCP to ad hoc environment, improvements have been proposed in the literature [15] to help TCP to differentiate between the different types of losses. Such TCP schemes can be examined to quantify their performance in combination with the proposed Q-SMS-F scheme. This would enable the seamless operation of application-level protocols such as File Transfer Protocol (FTP), Simple Mail Transfer Protocol (SMTP) and Hypertext Transfer Protocol (HTTP) [19] across the integrated MANETs and the Internet.

7.2.2 Network Layer

The multipath routing scheme proposed in this thesis does not exploit the discovered multiple paths simultaneously, but instead, the alternative paths are used for reliability purposes. A potential extension in this work is to distribute the load on multiple paths simultaneously. This can be done using a per-packet allocation scheme as it is known to work well in most networks [60][64][65][70][101][102][103].

Further, in Q-SMS scheme, there is no provisioning of any predictive way to anticipate a route break, which causes performance degradation particularly in mobile scenarios. Therefore, some pre-emptive maintenance mechanisms such as periodic capacity estimation during data transmission can help find a new route before the old route is broken so that the routing scheme can react much better to mobile scenarios.

7.2.3 MAC and PHY Layers

The IEEE 802.11 standard [2] is the predominantly used MAC mechanism. However, it cannot differentiate data traffic. All node and traffic classes have the same priority to access the wireless medium. To support applications with QoS, IEEE 802.11 working group developed IEEE 802.11e [22], which enhances the original IEEE 802.11 to provide a differentiated treatment of the traffic at the link layer. It will thus be interesting to see the actual performance of Q-SMS model based on IEEE 802.11e-based ad hoc network.

Although a limited account of the underlying channel impairments is taken into consideration in the capacity estimation process, it has been inevitable for it to assume that the underlying physical channel is perfect to some extent in most cases. On the contrary, this assumption is unrealistic. It is, therefore, necessary for every model the thesis has presented to consider the physical channel impairments and, in this respect, correct channel modelling becomes paramount. Hence, significant attention is required in this direction.

7.2.4 QoS Architecture

It is believed that a cross-layer design will be the key to providing QoS to real-time applications in MANETs [13][15][17]. An ultimate goal is to provide a model by considering all layers and their possible interactions for providing QoS in MANETs. Once these tasks have been completed successfully, it would be interesting to implement the complete model in an experimental testbed to see its practical viability enabling a real MANET to support QoS.

References

---

[1] T.S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, NJ, USA, 1996.

[2] "IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", *IEEE 802.11*, June 2007.

[3] "Bluetooth", http://www.bluetooth.com/bluetooth/

[4] ECMA International, "High Rate Ultra Wideband PHY and MAC Standard ECMA-368", 3rd Edition, December 2008, http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-368.pdf

[5] IETF Mobile Ad hoc Networks (MANET) Working Group, http://www.ietf.org/html.charters/manet-charter.html

[6] "IEEE Standard for Local and Metropolitan Area Networks", IEEE Computer Society and the IEEE Microwave Theory and Techniques Society Std. *IEEE 802.16*, 2004.

[7] S. Corson, J. Macker, "Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations", *RFC 2501*, January 1999.

[8] R.E. Kahn, S.A. Gronemeyer, J. Burchfiel, R.C. Kunzelman, "Advances in packet ratio technology", *Proc. IEEE*, 66, (11), November 1978, pp. 1468 1496.

[9] B. Leiner, R. Ruth, A. R. Sastry, "Goals and challenges of the DARPA GloMo program", *IEEE Personal Communications*, 3, (6), December 1996, pp. 34–43.

[10] R. Ruppe, S. Griswald, P.Walsh, R. Martin, "Near term digital radio (NTDR) system", *Proc. IEEE Military Communications Conference (MILCOM)*, 3, November 1997, pp. 1282–1287.

[11] C.K. Toh, W.K. Tsai, V. Li, G. Guichai, "Transporting audio over wireless ad hoc networks: experiments and new insights", *Proc. IEEE International Conference on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Beijing, China, September 2003, pp. 772–777.

[12] E. Setton, T. Yoo, X. Zhu, A. Goldsmith, B. Girod, "Cross-layer design of ad hoc networks for real-time video streaming", *IEEE Wireless Communications*, 12, (4), 2005, pp. 59–65.

[13] S. Sivavakeesar, "Quality of Service Support for Multimedia Applications in Mobile Ad hoc Networks", *PhD. thesis*, Centre for Communication Systems Research (CCSR), University of Surrey, Guildford, December 2005.

[14] C. Chowa, H. Ishiib, "Enhancing real-time video streaming over mobile ad hoc networks using multipoint-to-point communication", *Elsevier Computer Communications*, 30, (8), June 2007, pp. 1754–1764.

[15] C.S. Murthy, B.S. Manoj, *Ad hoc Wireless Networks Architectures and Protocols*, Prentice Hall, NJ, USA, 2004.

[16] E. Crawley, R. Nair, "A Framework for QoS-based Routing in the Internet", *RFC 2386*, August 1998.

[17] L. Chen, "Protocols for Supporting Quality of Service in Mobile Ad hoc Networks", *PhD. thesis*, School of Engineering and Applied Sciences, University of Rochester, New York, 2006.

[18] M. Drozda, "Performance Analysis of Wireless Ad hoc Networks", *PhD. thesis*, Universität Hannover, 2005.

[19] A. Leon-Garcia, I. Widjaja, *Communication Networks 2/e*, McGraw-Hill Science Engineering, July 2003.

[20] "Zigbee alliance", http://www.zigbee.org/en/index.asp

[21] A. Hassan, "IEEE 802.11 Technologies", *IEEE Communication Society Talks*, IEEE Seattle Communications (COM-19) Society Chapter, USA, February 2005.

[22] "IEEE 802 LAN/MAN Standards Committee", http://www.ieee802.org/

[23] "ETSI standards and technologies", http://www.etsi.org/WebSite/homepage.aspx

[24] "IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks--Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications", *IEEE 802.3*, 2005.

[25] Wikipedia, "IEEE 802.11", http://en.wikipedia.org/wiki/IEEE_802.11

[26] A. Chandra, V. Gummalla, J.O. Limb, "Wireless Medium Access Control Protocols", *IEEE Communications Surveys*, 3, (2), 2000, pp 2–15.

[27] J. Schiller, *Mobile Communications 2/e*, Addison-Wesley, August 2003.

[28] N. Abramson, "The ALOHA System - Another Alternative for Computer Communications", *Proc. Fall Joint Computer Conference*, 36, 1970, pp. 281 285.

[29] G. Malkin, "RIP Version 2", *RFC 2453*, November 1998.

[30] J. Moy, "OSPF Version 2", *RFC 2328*, April 1998.

[31] L.M. Feeney, "A Taxonomy for Routing Protocols in Mobile Ad hoc Networks", *SICS Technical Report T99/07*, October 1999.

[32] D. Lang, "A comprehensive overview about selected Ad hoc Networking Routing Protocols", *Technical Report*, Department of Computer Science, Technische Universität München, 2003.

[33] C.E. Perkins. P. Bhagwat. "Highly Dynamic Destination-Sequenced Distance Vector (DSDV) for Mobile Computers". *ACM SIGCOMM Computer Communication Review*. 24. (4). October 1994. pp 234–244.

[34] S. Murthy. J.J. Garcia-Luna-Aveces. "An Efficient Routing Protocol for Wireless Networks". *AACM/Baltzer Journal on Mobile Networks and Applications* Special Issue on Routing in Mobile Communication Networks. 1. (2). October 1996. . pp 183-197.

[35] C.E. Perkins. *Ad hoc Networking*. Addison-Wesley. 2001.

[36] C.E. Perkins. E.M. Royer. S. Das. "Ad hoc On-demand Distance Vector (AODV) Routing". *RFC 3561*. July 2003.

[37] V. Park. S. Corson. "Temporally-Ordered Routing Algorithm (TORA) Version 1". *Internet Draft* <draft-ietf-manet-tora-spec-04.txt>. June 2001.

[38] M. Gerla. T. Chen. "Global state routing: A new routing scheme for ad hoc wireless networks". *Technical Report*. Computer Science Department. University of California. Los Angeles. 1998.

[39] D.B. Johnson. D.A. Maltz. "Dynamic source routing in ad hoc wireless networks." *Technical Report*. Computer Science Department Carnegie Mellon University. 5000 Forbes Avenue Pittsburgh. PA. 1996.

[40] D.A. Maltz. J. Broch. D.B. Johnson. "Experiences designing and building a multi-hop wireless ad hoc network testbed" *Technical Report*. The CMU Monarch Project. Computer Science Department. Carnegie Mellon University. Pittsburgh. PA. 1999.

[41] D.B. Johnson. Y. Hu. D.A. Maltz. "Dynamic Source Routing in Ad hoc Wireless Networks". *RFC 4728*. February 2007.

[42] P. Jacquet. P. Muhlethaler. A. Qayyum. A. Laouiti. L. Viennot. T. Clausen. "Optimized link state routing protocol". *Internet Draft* <draft-ietf-manet-olsr-04.txt>. March 2001.

[43] T. Clausen, C. Dearlove, P. Jacquet, "The Optimized Link State Routing Protocol version 2", *Internet Draft* <draft-ietf-manet-olsrv2-07>, July 2008.

[44] Z.J. Haas, M.R. Pearlman, P. Samar, "The Zone Routing Protocol (ZRP) for Ad hoc Networks", *Internet Draft* <draft-ietf-manet-zone-zrp-04.txt>, July 2002.

[45] M. Jiang, J. Li, Y.C. Tay, "Cluster Based Routing Protocol (CBRP)", *Internet Draft* <draft-ietf-manet-cbrp.txt>, June 1999.

[46] R. Sivakumar, P. Sinha, V. Bharghavan, "CEDAR: A Core-Extraction Distributed Ad hoc Routing Algorithm", *IEEE Journal on Selected Areas in Communications*, 17, (8), August 1999, pp 1454–1465.

[47] I. Chakeres, C.E. Perkins, "Dynamic MANET On-demand Routing Protocol (DYMO)", *Internet Draft* <draft-ietf-manet-dymo-14.txt>, June 2008.

[48] C.K. Toh, E.M. Royer, "Review of current routing protocols for ad hoc mobile wireless networks", *Technical Report*, UCSB, 1999.

[49] T.D. Dyer, R.V. Boppana, "A comparison of TCP performance over three routing protocols for mobile ad hoc networks", *Proc. ACM Symposium on Mobile Ad hoc Networking & Computing (MOBIHOC)*, Long Beach, CA, October 2001.

[50] J. Broch, D.A. Maltz, D.B. Johnson, Y. Hu, J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols", *Proc. Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, Dallas, TX, October 1998, pp. 85–97.

[51] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, M. Degermark, "Scenario-based performance analysis of routing protocols for mobile ad hoc networks", *Proc. Fifth Annual ACM International Conference on Mobile Computing and Networking (MOBICOM)*, Seattle, Washington, August 1999, pp. 195–206.

[52] S.R. Das, C.E. Perkins, E.M. Royer, "Performance comparison of two on-demand routing protocols for ad hoc networks", *Proc. INFOCOM 2000 Conference*, Tel-Aviv, Israel, March 2000.

[53] S.R. Das, R. Castaneda, J. Yan, R. Sengupta, "Comparative performance evaluation of routing protocols for mobile ad hoc networks", *Proc. 7th Int. Conference on Computer Communications and Networks*, Lafayette, LA, University of Texas at San Antonio, October 1998, pp. 153–161.

[54] I.D. Aron, S.K. Gupta, "On the scalabty of on-demand routing protocols for mobile ad hoc networks: An analytical study", *Journal of Interconnection Networks*, 2, (1), March 2001, pp. 5–21.

[55] D.A. Maltz, "Resource management in multi-hop ad hoc networks", *Technical Report*, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, November 1999.

[56] P. Jacquet, A. Laouiti, "Analysis of mobile ad hoc network routing protocols in random graph models", *Technical Report* 3835, INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78 153 Le Chesany Cedex, France, 1999.

[57] L. Christensen, G. Hansen, "The optimized link state routing protocol", *M.S. thesis*, Department of Computer Science, Aalborg University, Fredrik Bajers Vej 7, 9220 Aalborg Ost, 2001.

[58] A. Nasipuri, R. Castaneda, S. Das, "Performance of Multipath Routing for On-demand Protocols in Mobile Ad hoc Networks", *ACM Kluwer Mobile Networks and Applications Journal*, 6, (4), 2001, pp. 339–349.

[59] L. Wang, L. Zhang, Y. Shu, M. Dong, "Multipath Source Routing in Wireless Ad hoc Networks", *Proc. Canadian Conference on Electrical and Computer Engineering*, 2000, pp. 479–483.

[60] S. Lee, M. Gerla, "Split Multipath Routing with Maximally Disjoint Paths in Ad hoc Networks", *Proc. IEEE International Conference on Communication*, Helsinki, June 2001, pp. 3201–3205.

[61] M. Marina, S. Das, "On-demand Multipath Distance Vector Routing in Ad hoc Networks", *Proc. 9th IEEE International Conference on Network Protocols*, California, 2001, pp.14–23.

[62] X. Li, L. Cuthbert, "A Reliable Node-Disjoint Multipath Routing with Low Overhead in Wireless Ad hoc Networks", *Proc. 7th ACM IEEE International Symposium on Modelling, Analysis and Simulation of Wireless and Mobile Systems*, Venezia, 2004, pp. 230–233.

[63] L.R. Reddy, S.V. Raghavan, "SMORT: Scalable Multipath On-demand Routing for Mobile Ad hoc Networks", *Elsevier Ad hoc Networks*, 5, (2), 2007, pp. 162–188.

[64] C. Tachtatzis, "Load distribution and energy awareness in MANETs using multipath routing", *PhD. thesis*, Department of Electrical & Electronic Engineering, University of Strathclyde, Glasgow, April 2008.

[65] C. Tachtatzis, D. Harle, "Performance Evaluation of Multi-path and Single-path Routing Protocols for Mobile Ad hoc Networks", *Proc. 2008 International Symposium on Performance Evaluation of Computer and Telecommunication Systems*, Edinburgh, June 2008.

[66] A. Wang, G.H. Kuenning, P. Reiher, "Multipath Routing for Ad hoc Networks", *Book Chapter: Mobile and Wireless Internet: Protocols, Algorithms, and Systems*, Kluwer Academic Publishers, July 2003.

[67] M.T. Toussaint, "Multipath Routing in Mobile Ad hoc Networks", *CACTUS Project*, August 2003.

[68] J. Tsai, T. Moors, "A Review of Multipath Routing Protocols: From Wireless Ad hoc to Mesh Networks", *Proc. ACoRN Early Career Researcher Workshop on Wireless Multihop Networking*, July 2006.

[69] S. Mueller, R.P. Tsang, D. Ghosal, "Multipath Routing in Mobile Ad hoc Networks: Issues and Challenges", *Lecture Notes in Computer Science*, 2965, April 2004, pp. 209–234.

[70] R. Krishnan, J.A. Silvester, "Choice of Allocation Granularity in Multipath Source Routing Schemes", *Proc. 12th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, San Francisco, CA, March 1993, pp. 322–329.

[71] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modelling*, Wiley-Interscience, New York, NY, April 1991.

[72] K. Watkins, *Discrete event simulation in C*, McGraw-Hill, England, 1993.

[73] A.M. Law, W.D. Kelton, *Simulation Modelling and Analysis 3 e*, McGraw-Hill, 2000.

[74] T.R. Andel, A. Yasinsac, "On the credibility of MANET simulations", *IEEE Computer*, 19, (7), July 2006, pp. 48–54.

[75] I. Stojmenovic, "Simulations in wireless sensor and ad hoc networks: matching and advancing models, metrics, and solutions", *IEEE Communications Magazine*, 46, (12), December 2008, pp. 102–107.

[76] The CMU Monarch Project's Wireless and Mobility Extensions to NS, August 1998, http://www.monarch.cs.rice.edu/cmu-ns.html

[77] K. Fall, K. Varadhan (Eds.), "The ns Manual (formerly ns Notes and Documentation)", December 2008, http://www.isi.edu/nsnam/ns/doc/ns_doc.pdf

[78] NS-2 (The Network Simulator 2), http://www.isi.edu/nsnam/ns/

[79] GloMoSim (Global Mobile Information Systems Simulation Library), http://pcl.cs.ucla.edu/projects/glomosim/

[80] QualNet, Scalable Network Technologies, http://www.scalable-networks.com/

[81] OPNET (Optimized Network Engineering Tool). http://www.opnet.com/

[82] OMNeT++ (Objective Modular Network Testbed in C++) Community Site. http://www.omnetpp.org/

[83] G.A. DiCaro. "Analysis of simulation environments for mobile ad hoc networks". *Technical Report*. Dalle Molle Institute for Artificial Intelligence. Manno. Switzerland. December 2003.

[84] M. Halvardsson. P. Lindberg. "Reliable Group Communication in Military Ad hoc Network". *Technical Report*. Vaxjo University. 2004.

[85] L. Begg. W. Liu. K. Pawlikowski. S. Perera. H. Sirisena. "Survey of Simulators of Next Generation Networks for Studying Service Availability and Resilience". *Technical Report*. University of Canterbury. New Zealand. February 2006.

[86] NS-3 (The Network Simulator 3). http://www.nsnam.org/

[87] OTcl - Object Tcl Extensions. http://bmrc.berkeley.edu/research/cmt/cmtdoc/otcl/

[88] B. Tuch. "Development of WaveLAN, an ISM Band Wireless LAN". *AT&T Technical Journal*. 72. (4). 1993. pp. 27–33.

[89] "Simulation Model for the DSR MANET Routing Protocol". National Institute of Standards and Technology Group (WCTG). National Institute of Standards and Technology. Gaithersburg. Maryland. http://www.antd.nist.gov/wctg/

[90] J. Yoon. M. Liu. B. Noble. "Random Waypoint Considered Harmful". *Proc. IEEE INFOCOM*. 2. San Francisco. CA. April 2003. pp 1312–1321.

[91] M.K. Nakayama. "Statistical analysis of simulation output". *Proc. Winter Simulation Conference*. 2008. pp. 62–72.

[92] O. Balci. "Verification. validation and accreditation of simulation models". *Proc. Winter Simulation Conference*. 1997. pp. 135–141.

[93] J.S. Carson. "Model verification and validation". *Proc. Winter Simulation Conference*. 2002. pp. 52–58.

[94] R.G. Sargent, "Verification and validation of simulation models", *Proc. Winter Simulation Conference*, 2003, pp. 37–48.

[95] S. Ni, Y. Tseng, Y. Chen, J. Sheu, "The Broadcast Storm Problem in a Mobile Ad hoc Network", *Proc. ACM IEEE International Conference on Mobile computing and networking (MOBICOM)*, Seattle, August 1999, pp. 151–162.

[96] W. Peng, X.C. Lu, "On the Reduction of Broadcast Redundancy in Mobile Ad hoc Networks", *Proc. First Annual Workshop Mobile and Ad hoc Networking and Computing (MOBIHOC)*, Boston, August 2000, pp. 129–130.

[97] W. Lou, J. Wu, "On Reducing Broadcast Redundancy in Ad hoc Wireless Networks", *IEEE Transactions on Mobile Computing*, 1, (2), 2002, pp. 111–123.

[98] M.I. Stojmenovic, "Comments and corrections to dominating sets and neighbour elimination-based broadcasting algorithms in wireless networks", *IEEE Transactions on Parallel and Distributed Systems*, 15, (11), 2004, pp. 1054–1055.

[99] GDB: The GNU Project Debugger, http://www.gnu.org/software/gdb/

[100] Data Display Debugger, http://www.gnu.org/software/ddd/

[101] W. Wei, A. Zakhor, "Robust Multipath Source Routing Protocol (RMPSR) for Video Communication over Wireless Ad hoc Networks", *Proc. International Conference on Multimedia and Expo*, Taipei, June 2004, pp.1379–1382.

[102] M. Takeshi, B. Masaki, S. Iwao, "Split Multipath Routing Protocol with Load Balancing policy (SMR-IB) to improve TCP performance in Mobile Ad hoc Networks", *IEICE Communications*, 89, (5), 2006, pp. 1517–1525.

[103] A. Chunsoo, S. Jitae, H. Eui-Nam, "Enhanced Multipath Routing Protocol using congestion metric in Wireless Ad hoc Networks", *Lecture Notes in Computer Science*, 4096, 2006, pp. 1089–1097.

[104] S.J. Lee, E.M. Belding-Royer, C.E. Perkins, "Scalability Study of the Ad hoc On-demand Distance Vector Routing Protocol", *International Journal of Network Management*, 13, (2), 2003, pp. 97–114.

[105] P. Mohapatra, J. Li, C. Gui, "Qos in Mobile Ad hoc Networks", *IEEE Wireless Communications Magazine* Special Issue on QoS in Next-Generation Wireless Multimedia Communications Systems, 10, (3), June 2003, pp. 44–52.

[106] H. Zhu, M. Li, I. Chlamtac, B. Prabhakaran., "A survey of quality of service in IEEE 802.11 networks", *IEEE Wireless Communications*, 11, (4), August 2004, pp. 6–14.

[107] L. Hanzo (II.), R. Tafazolli, "A Survey of QoS Routing Solutions for Mobile Ad hoc Networks", *IEEE Communications Surveys and Tutorials*, 9, (2), 2007, pp. 50–70.

[108] L. Chen; W.B. Heinzelman, "A Survey of Routing Protocols that Support QoS in Mobile Ad hoc Networks", *IEEE Network*, 21, (6), December 2007, pp. 30–38.

[109] X. Zhang, S.B. Lee, A. Gahng-Seop, A.T. Campbell, "INSIGNIA: An IPBased Quality of Service Framework for Mobile Ad hoc Networks", *Journal of Parallel and distributed Computing*, 60, (4), April 2000, pp. 374–406.

[110] R. Braden (ed.), L. Zhang, S. Berson, S. Herzog, S.Jamin, "Resource ReSerVation Protocol (RSVP) – Volume 1 Functional Specification", *RFC 2205*, Sept. 1997.

[111] H. Xiao, W. K. G., Seah, A. Lo, K. C. Chua, "A Flexible Quality of Service Model for Mobile Ad hoc networks", *Proc. IEEE Vehicular Technology Conference*, Tokyo, Japan, May 2000, pp. 445–449.

[112] R. Braden, D. Clark, S. Shenker, "Integrated services in the Internet architecture: an overview", *RFC 1633*, June 1994.

[113] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, "An Architecture for Differentiated Services", *RFC 2475*, Dec. 1998.

[114] G. Ahn, A.T. Campbell, A. Veres, L. Sun, "Supporting Service Differentiation for Real-Time and Best-effort Traffic in Stateless Wireless Ad hoc Networks (SWAN)", *IEEE Transactions on Mobile Computing*, 1, (3), September 2002, pp. 192–207.

[115] S. Chen, K. Nahrstedt, "Distributed Quality-of-Service Routing in Ad hoc Networks", *IEEE Journal on Selected Areas in Communications*, 17, (8), August 1999, pp. 1488–1505.

[116] C.R. Lin, "On-Demand QoS Routing in Multihop Mobile Networks", *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 1, April 2001, pp. 1735–1744.

[117] C. Zhu, M.S. Corson, "QoS Routing for Mobile Ad hoc Networks", *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, Anchorage, Alaska, 21, (1), June 2002, pp. 958–967.

[118] Y. Chen, Y. Tseng, J. Sheu, P. Kuo, "On-Demand, Link-State, Multi-Path QoS Routing in a Wireless Mobile Ad hoc Network", *Proc. European Wireless 2002*, Florence, Italy, February 2002, pp. 135–141.

[119] Y. Li, X. Chen, D. Yu, "Disjoint Multipath QoS Routing in Ad hoc Networks", *Proc. 2005 International Conference on Wireless Communications, Networking and Mobile Computing*, September 2005, pp. 739–742.

[120] Q. Xue, A. Ganz, "Ad hoc QoS On-demand Routing (AQOR) in Mobile Ad hoc Networks", *Journal of Parallel and Distributed Computing*, 63 (2), February 2003, pp.154–165.

[121] H. Sun, H.D. Hughes, "Adaptive QoS Routing based on Prediction of Local Performance in Ad hoc Networks", *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 4, (1), March 2003, pp. 1191–1195.

[122] L. Hanzo (II.), R. Tafazolli, "Throughput Assurances Through Admission Control for Multi-hop MANETs", *Proc. 18th IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Athens, Greece, September 2007, pp. 1-5.

[123] L. Chen, W.B. Heinzelman, "QoS-Aware Routing Based on Bandwidth Estimation for Mobile Ad hoc Networks", *IEEE Journal on Selected Areas in Communications*, 23, (3), March 2005, pp. 561-572.

[124] Y. Yang, R. Kravets, "Contention-aware admission control for ad hoc networks", *IEEE Transactions on Mobile Computing*, 4, (4), August 2005, pp. 363–377.

[125] I.D. Chakeres, E.M. Belding-Royer, J.P. Macker, "Perceptive Admission Control for Wireless Network Quality of Service", *Elsevier Ad hoc Networks*, 5, (7), September 2007, pp. 1129–1148.

[126] G. Anastasi, E. Borgia, M. Conti, E. Gregori, "IEEE 802.11b Ad hoc Networks: Performance Measurements", *Cluster Computing*, 8, (2–3), July 2005, pp. 135–145.

[127] R. Prasad, M. Murray, C. Dovrolis, K. Claffy, "Bandwidth estimation: metrics, measurement techniques, and tools", *IEEE Network*, 17, (6), 2003, pp. 27–35.

[128] J. Jun, P. Peddabachagari, M. Sichitiu, "Theoretical Maximum Throughput of IEEE 802.11 and its Applications", *Proc. 2nd IEEE International Symposium on Network computing and Applications*, Cambridge, USA, April 2003, pp. 249–256.

[129] Q. Chen, F. Schmidt-Eisenlohr, D. Jiang, M. Torrent-Moreno, L. Delgrossi, H. Hartenstein, "Overhaul of IEEE 802.11 modelling and simulation in ns-2", *Proc. 10th ACM Symposium on Modelling, Analysis, and Simulation of Wireless and Mobile Systems*, Chania, Greece, October 2007, pp. 159–168.

[130] I. Glover; P. Grant, *Digital Communications 2/e*, Pearson Education Ltd., 2004.