Modelling and Performance Analysis of Mobile Ad Hoc Networks

Thesis by

Osama Younes

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Newcastle University
Newcastle Upon Tyne, UK

(Submitted January, 2013)

Abstract

Mobile Ad hoc Networks (MANETs) are becoming very attractive and useful in many kinds of communication and networking applications. This is due to their efficiency, relatively low cost, and flexibility provided by their dynamic infrastructure. Performance evaluation of mobile ad hoc networks is needed to compare various architectures of the network for their performance, study the effect of varying certain network parameters and study the interaction between various parameters that characterise the network. It can help in the design and implementation of MANETs.

It is to be noted that most of the research that studies the performance of MANETs were evaluated using discrete event simulation (DES) utilising a broad band of network simulators. The principle drawback of DES models is the time and resources needed to run such models for large realistic systems, especially when results with a high accuracy are desired. In addition, studying typical problems such as the deadlock and concurrency in MANETs using DES is hard because network simulators implement the network at a low abstraction level and cannot support specifications at higher levels.

Due to the advantage of quick construction and numerical analysis, analytical modelling techniques, such as stochastic Petri nets and process algebra, have been used for performance analysis of communication systems. In addition, analytical modelling is a less costly and more efficient method. It generally provides the best insight into the effects of various parameters and their interactions. Hence, analytical modelling is the method of choice for a fast and cost effective evaluation of mobile ad hoc networks.

To the best of our knowledge, there is no analytical study that analyses the performance of multi-hop ad hoc networks, where mobile nodes move according to a random mobility model, in terms of the end-to-end delay and throughput. This work

presents a novel analytical framework developed using stochastic reward nets and mathematical modelling techniques for modelling and analysis of multi-hop ad hoc networks, based on the IEEE 802.11 DCF MAC protocol, where mobile nodes move according to the random waypoint mobility model. The proposed framework is used to analysis the performance of multi-hop ad hoc networks as a function of network parameters such as the transmission range, carrier sensing range, interference range, number of nodes, network area size, packet size, and packet generation rate.

The proposed framework is organized into several models to break up the complexity of modelling the complete network and make it easier to analyse each model as required. This is based on the idea of decomposition and fixed point iteration of stochastic reward nets. The proposed framework consists of a mathematical model and four stochastic reward nets models; the path analysis model, data link layer model, network layer model and transport layer model. These models are arranged in a way similar to the layers of the OSI protocol stack model.

The mathematical model is used to compute the expected number of hops between any source-destination pair; and the average number of carrier sensing, hidden, and interfering nodes. The path analysis model analyses the dynamic of paths in the network due to the node mobility in terms of the path connection availability and rate of failure and repair. The data link layer model describes the behaviour of the IEEE 802.11 DCF MAC protocol. The actions in the network layer are modelled by the network layer model. The transport layer model represents the behaviour of the transport layer protocols. The proposed models are validated using extensive simulations.

Publications

Journals:

- Osama Younes and Nigel Thomas "An SRN Model of the IEEE 802.11 DCF MAC Protocol in Multi-Hop Ad Hoc Networks with Hidden Nodes," *The Computer Journal*, Vol 54(6): pp. 875-893, 2011.

- Osama Younes and Nigel Thomas, "Modelling and Performance Analysis of Multi-hop Ad Hoc Networks," submitted to *Simulation Modelling Practice and Theory*.

Conferences and workshops:

- Osama Younes, Wail Elkilany, and Nigel Thomas, "SRN Model for Performance Evaluation of TCP Sessions Sharing Bottleneck Links in WAN," *European Symposium on Modeling and Simulation, October 26-28, 2009, Leicester, United* Kingdom.

- Osama Younes and Nigel Thomas, "SRN Model for IEEE 802.11 DCF MAC Protocol in Multi-hop Ad Hoc Networks with Hidden Nodes," *UK Performance Engineering Workshop* (*UKPEW 2010*), July 8-9, 2010, University of Warwick, United Kingdom.

- Osama Younes and Nigel Thomas, "Analysis of the Expected Number of Hops in Mobile Ad Hoc Networks with Random Waypoint Mobility," *in Fifth International Workshop on Practical Applications of Stochastic Modelling* (*PASM'11*), *Electronic Notes in Theoretical Computer Science*, Vol. 275, pp. 143-158, September 2011.

- Osama Younes and Nigel Thomas, "A Path Connection Availability Model for MANETs with Random Waypoint Mobility," *European Performance Engineering Workshop* (*EPEW*), Munich, Germany, July 2012.

- Osama Younes and Nigel Thomas, "SRN Models for Analysis of Multihop Wireless Ad Hoc Networks," *UK Performance Engineering Work*shop (*UKPEW*), Edinburgh, July 2012.

Acknowledgments

I am deeply grateful to many people for their help and support during the period of my PhD.

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Nigel Thomas, for his help, guidance and for the many valuable suggestions he made throughout the work for this thesis.

I would like to acknowledge the financial support of Egyptian government that provided the necessary financial support for my PhD.

I am grateful to Chris Ritson for helping to solve the technical problems in the Mill cluster which was used to do my simulation experiments.

I also would like to thank my friends from the 10th floor: Abubkr Abdelsadiq, Khaled Alekeish, Johari Abdullah, Ayad Keshlaf, and Kamarul Abdul Basit.

I am most grateful to Dr. Wail Shawki Elkilani, who introduced me to the world of Petri nets and performance engineering, and inspired me to start this work.

Last but not least, I would like to thank my mother and my wife for their love, support, encouragement, and patience. Without their constant support I would not have made it.

Contents

List of Figures

List of Tables

List of Abbreviations

AODV	Ad-hoc On-Demand Distance Vector
BA	Basic Access
CA	Collision Avoidance
CBK	Call Back
CBR	Constant Bit Rate
CCA	Clear Channel Assessment
CDF	Cumulative Distribution Function
CSMA	Carrier Sensing Multiple Access
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CTMC	Continuous Time Markov Chain
CTS	Clear To Send
CW	Contention Window
DBPSK	Differential Binary Phase Shift Keying
DCF	Distributed Coordination Function
DES	Discrete Event Simulation
DFIR	Diffused Infrared
DIFS	DCF Inter Frame Space
DQPSK	Quadruple Phase Shift Keying
DSDV	Destination-Sequenced Distance-Vector routing
DSR	Dynamic Source Routing
DSSS	Direct Sequence Spread Spectrum
FAMA	Floor Acquisition Multiple Access
FHSS	Frequency Hopping Spread Spectrum
FSK	Frequency Shift Keying
GOD	General Operations Director
LAR	Location-Aided Routing

LLC	Logical Link Control
LRC	Long Retry Counter
LRD	Least Remaining Distance
MAC	Medium Access Control
MACA	Multiple Access Collision Avoidance
MANETs	Mobile Ad hoc Networks
MHD	Maximum Hop Distance
MPDU	MAC Protocol Data Unit
MRL	Maximum Retry Limit.
NAV	Network Allocation Vector
OLSR	Optimized Link State Routing
PCF	Point Coordination Function
PDF	Probability Density Function
PDU	Protocol Data Unit
PPM	Pulse Position Modulation
QoS	Quality Of Service
RERR	Route Error
RREP	Route Reply
RREQ	Route Request
RTS	Request To Send
RTS/CTS	Request-To-Send/Clear-To-Send
RWPMM	Random Waypoint Mobility Model
SIFS	Short Inter Frame Space
SNR	Signal to Noise Ratio
SRC	Short Retry Counter
SRN	Stochastic Reward Net
TAB	Traffic-Analysis-Based
TBRPF	Topology Broadcast Based on Reverse-Path Forwarding

TCP	Transmission Control Protocol
TORA	Temporally-Ordered Routing Algorithm
TTL	Time To Live
UDP	User Datagram Protocol
WLAN	Wireless Local Area Networks

List of Symbols

A_{AB}	Intersection area between node *A* and *B*
A_h	Average size of the hidden area
A_i	Average size of the interfering area
A_{int}	Size of intersection area
A_s	Average number of backoff slots
B_1	Minimum transmission bit rate
B_2	Maximum transmission bit rate
CW	Contention backoff window
CW_{max}	Maximum contention window
CW_{min}	Minimum contention window
d	Distance between a source and destination node
d_r	Distance between 2-hop-apart nodes
$E(x)$	Expected value for a parameter x
$Ft(T_x)$	Average firing time of transition T_x
G_R	Antenna gains of the receiver
G_t	Antenna gains of the sender
L	Side length of the squared network area
N	Number of nodes in the network
N_B	Buffer size
N_{cs}	Average number of carrier sensing nodes
N_e	Average number of hops from any broken link in the path to the source node
N_H	Average number of hidden nodes
N_{HS}	Number of hops to the source of undeliverable packet
N_h	Average path length in hops

N_i	Average number of interfering nodes
N_{in}	Average number of nodes in the intersections areas
N_{LH}	Last known hop count to the destination
N_n	Number of neighbour nodes
N_r	Number of nodes located in the half circle towards of the destination
N_s	Number of traffic sources
N_t	Average number of nodes in a circle with radius equal to the average interference range
n_c	Average number of tries to transmit the MAC frame
P_i	Power level of received signals from any interfering node
PhH	Size of the physical layer header
P_i	Probability of a state *i*, where *i* = 0, 1, 2,.....
P_r	Power level of received signals from the desired sender
$Pr(E)$	Probability of an event *E*
$\#P_x$	Number of token in the place P_x
p_{BA}	Probability that an interfering node transmits a data or ACK frame in the case of a multi-hop ad hoc network with the basic access scheme
p_e	Probability that an interfering node transmits a RTS or CTS frame in the case of a multi-hop ad hoc network with the RTS/CTS scheme
p_R	Probability that an interfering node transmits a RTS, CTS, data or ACK frame in the case of a multi-hop ad hoc network with the RTS/CTS scheme
Q	Queue length
R	Transmission range
Rate(T_x)	Firing rate of the transition T_x
R_{cs}	Carrier sensing range
R_i	Interference range
R_{iA}	Average interference range
RNS	Random number of slots
r	Expected forward distance

T_{CCA}	Time required to recognise the signal
$Thr(T_x)$	Throughput of the transition T_x
T_p	Propagation time
T_{RxTx}	Time required to convert from receiving to transmitting state
T_s	Slot time
T_{SNR}	Signal to noise ratio threshold
t_{CTS}	Transmission time of the CTS frame
V_a	Average speed of a node
V_{max}	Maximum speed of a node
V_{min}	Minimum speed of a node
V_r	Relative speed between two nodes
V_B	Speed of a node *B*
X	Remaining distance to the destination
X_r	Expected value of the remaining distance to the destination
α	Average node utilisation
α_x	Utilisation of a node x
β_1	Probability that the channel is busy in the case of single hop ad hoc networks with the basic access scheme
β_2	Probability that the channel is busy in the case of single hop ad hoc networks with the RTS/CTS scheme
β_{BA}	Probability that the channel is idle in the case of a multi-hop ad hoc network with the basic access scheme
β_c	Backoff counter
β_{cMAX}	Maximum value of the retry counter
β_R	Probability that the channel is idle in the case of a multi-hop ad hoc network with the RTS/CTS scheme
δ	Packet delay per hop
δ_d	Packet delay in data link layer
δ_n	Average delay of packets in network layer model

ε	Packet loss probability in the data link layer model
ε_B	Packet loss probability due to buffer overflow
λ	Data transmission rate
λ_c	Number of control packets sent per unit time
λ_e	Entering rate to intersection area
λ_h	Packet generation rate of hidden nodes
λ_n	Throughput of the network layer model
λ_r	Average number of routed packets per node per unit time
λ_r^x	Packet forward rate for a node *x*
λ_{rt}	Average number of routed packets per unit time
λ_T	Throughput of the transport layer model
λ_t^x	Number of packets that is successfully transmitted by any node *x* per unit time
μ	Intersection area leaving time
μ_B	Probability of failure to transmit the data frame due to interference induced by interfering nodes in the case of a multi-hop ad hoc network with the basic access scheme
μ_c	Probability of failure to transmit the RTS frame due to interference induced by interfering nodes in the case of a multi-hop ad hoc network with the RTS/CTS scheme
μ_f	Average failure rate of paths
μ_R	Probability of success to complete the RTS/CTS handshake in the case of a multi-hop ad hoc network
μ_r	Average repair rate of paths
μ_s	Path failure and repairing frequency
μ_1	Probability of failure to send a data frame in the case of single hop ad hoc networks with the basic access scheme
μ_2	Probability of failure to send a data frame in the case of single hop ad hoc networks with the RTS/CTS scheme
ξ	Square of the distance between a source and destination node
ρ_{BA}	Probability that a hidden node transmits a data or ACK frame in the case of a multi-hop ad hoc network with the basic access scheme

ρ_i	Distance to a neighbour node number i, where i = 1, 2, ...
ρ_{max}	Maximum value of ρ_i, where i = 1, 2, ... , N_r
ρ_R	Probability that a hidden node transmits a RTS, CTS, data or ACK frame in the case of a multi-hop ad hoc network with the RTS/CTS scheme
τ_{LR}	Time required to finish the local repair process
τ_{NR}	Time required for establishing a new route
τ_{RERR}	Time required for broadcasting RERR message
τ_{SR}	Time required to finish the source repair process
Ψ	Average path connection availability
ω_R	Probability that a hidden node transmits a RTS or CTS frame in the case of a multi-hop ad hoc network with the RTS/CTS scheme

Chapter 1

Introduction

1.1 Mobile Ad Hoc Networks

Traditional wireless communication networks require a fixed infrastructure over which communication takes place. Therefore, considerable resources and effort are required to set up such networks, even before they can actually be used. In cases where setting up infrastructure is a difficult or even impossible task, such as in military applications, disaster relief, or emergency operations, other alternatives need to be developed.

Mobile Ad hoc Networks (MANETs) are stand-alone wireless networks which lack the service of a backbone infrastructure [1]. MANETs are formed dynamically by mobile nodes that are connected via wireless links without using existing network infrastructure or centralised administration. The nodes in MANETs are free to move at any time; thus the topology of the network may possibly change rapidly and unpredictably. In addition, the nodes in the network not only act as sources but also as routers that direct data to or from other nodes which cannot communicate directly with one another. A gateway node may be present in an ad hoc network which allows the nodes to communicate with an external network such as the Internet.

In MANETs, nodes are supplied with antennas which allow them to transmit and receive signals from the other nodes. The antenna can radiate and receive within a certain radius, which is called the *transmission range* (*R*). The radius is determined by the level of transmission power. When a node transmits to another node, its transmission can be heard by all nodes that lie within the transmission range, and these nodes are called *neighbour nodes*. The area covered by the transmission range is called the *capture area*. The higher the transmission power, the larger the size of the capture area and the number of neighbour nodes, but potentially also the higher the amount of interference that may be experienced.

The ad hoc network is formed as soon as one of the nodes expresses a wish to exchange information with one other node (*unicast* transmission) or with more than one node (*multicast* transmission). By using some nodes as relay points, a mobile node is able to send a packet to another node located outside its transmission range. This mode of communication is known as *wireless multi-hop*. Thus, MANETs are sometimes referred to as *multi-hop ad hoc networks*. MANETs were initially designed for use in emergency relief and military applications. Recently, the ad hoc network model has been proposed for many other applications [1], such as sensor networks, vehicular ad hoc networks (for intelligent transportation), and educational applications (such as virtual classes and conference rooms). An example of a mobile ad hoc network is shown in Figure 1.1.

Figure 1.1: A mobile ad hoc network

In addition to the transmission range, in wireless networks, nodes with multi-directional antennas also have other two radio ranges*:* the *carrier sensing range* (R_{cs}) and the *interference range* (R_i). The carrier sensing range is a physical parameter for a wireless radio which depends on the sensitivity of the antenna. Any transmission from other nodes in the carrier sensing range of node *S* will trigger the carrier sensing detection, and *S* then detects the channel as busy. If the channel is detected as busy, node *S* will wait for the channel to become idle before it starts trying to transmit a packet [2]. The area covered by the carrier sense range of a node is called the *carrier sensing area* for the node. The nodes located in the carrier sensing area are called *carrier sensing nodes*.

The interference range is a range around a receiver within which an unrelated transmission causes interference to any received signal at the receiver [3]. For example, if node *S* transmits to node *D*, any transmission from any node located within the interference range of *D* interferes with the signal received at *D*.

1.2 Characteristics and Challenges of MANETs

Mobile ad hoc networks share many properties in common with wired and infrastructure wireless networks, but also have certain unique features which arise from the characteristics of the wireless channel, the mobility of the nodes and the routing mechanisms used to establish and maintain communication paths. These features add more complexity and constraints that render the design or analysis of this type of network a challenge. These unique features are summarised as follows:

- **Node mobility**

Nodes in wireless ad hoc networks are free to move. Hence the network topology often changes rapidly and unpredictably. The dynamic nature of the network topology results in frequent path breaks. Therefore, nodes need to periodically collect connectivity information from other nodes. One implication of this is that the message overhead needed to collect topology information will increase. Mobility is a crucial factor affecting the design and analysis of MANETs.

- **Limited bandwidth**

In general, wireless networks are bandwidth limited. In MANETs, the bandwidth is even more limited because there is no backbone infrastructure to handle (or multiplex) higher bandwidth traffic. Therefore, MANETs usually operate in bandwidth-constrained and variable-capacity links. This results in high bit errors, low bandwidth, and unstable and asymmetric links, which result in congestion problems. Hence, the optimal usage of bandwidth is necessary to keep the overhead of any protocol designed for MANETs as low as possible.

- **Energy constrained operation**

Most ad hoc nodes rely on batteries of limited life. Therefore, the energy preservation and efficient use may be the most important criteria for designing protocols for MANETs. Thus, the protocols of MANETs must be developed to be power-aware.

- **Spatial contention and reuse**

In wireless networks, nodes contend with each other to access the communication channel. However, when a node starts to transmit, it reserves the area around it for the duration of the transmission, so that no other transmission can take place during that time interval as it would result in a collision and, consequently, a waste of bandwidth. Spatial reuse indicates the number of concurrent transmissions which may take place in a network without interfering with each other. Transmissions should be coordinated in such a way that maximizes the property of spatial reuse.

- **Security**

Securing mobile ad hoc networks is a greatly challenging issue. This is because ad hoc networks have to cope not only with the same kinds of vulnerability as their wired and wireless counterparts, but often also with new types of vulnerability specific to ad hoc networks resulting from their inherent mobility [4] and lack of physically secured infrastructure. A detailed analysis of security issues and solutions for mobile ad hoc networking can be found in [5] and are not considered further within this thesis.

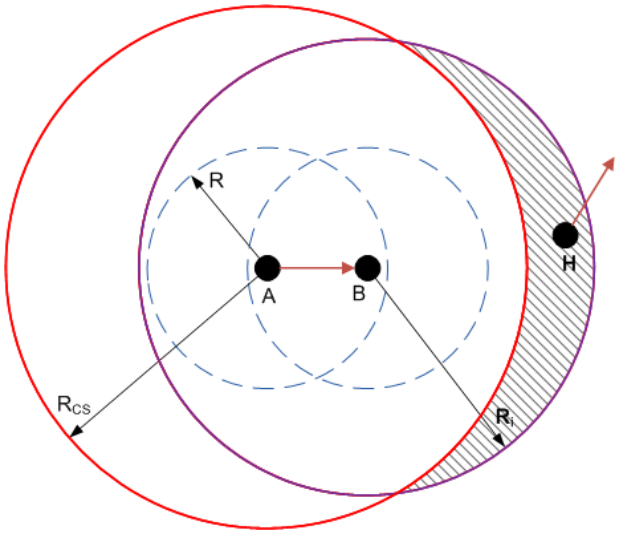

Figure 1.2: Hidden node problem

- **Hidden and exposed nodes**

The *hidden area* is the area covered by the interference range of the receiver which is not covered by the carrier sensing range of the sender. Nodes located in the hidden area are called *hidden nodes*. Figure 1.2 illustrates the hidden node problem, where circles with radii R, R_{cs} and R_i around any node respectively represent the transmission range, carrier sensing range and interference range of the node. Consider a case where node A is transmitting to node B. The dashed area is in the interference range of *B* and out of the carrier sensing range of *A*, as shown in Figure 1.2. Therefore, any node located in this area (e.g. the node *H*) is hidden from *A*. This means that *A* will not be able to detect an ongoing transmission from *H* to any other node. Consequently, if *A* and *H* send their packets at the same time, there will be a packet collision at node *B*.

The *exposed node* problem can be considered as the opposite of the hidden node problem. Instead of nodes transmitting when they should not, as happens with hidden nodes, exposed nodes are nodes that are prevented from transmitting when they could. The exposed node effect occurs when a node that needs to transmit a message senses a busy medium and defers the transmission even though it would not interfere with the other sender's transmission. Figure 1.3 shows an example of an exposed

node. In this case the transmission from node A to node B prevents node C, located in the carrier sensing node of A indicated by shaded area, from transmitting to any other node, although its transmission would not interfere with that between A and B. The exposed node problem prevents the full utilisation of the available bandwidth of the medium. The problems of hidden and exposed nodes are well-known in multi-hop ad hoc networks that can severely affect performance.

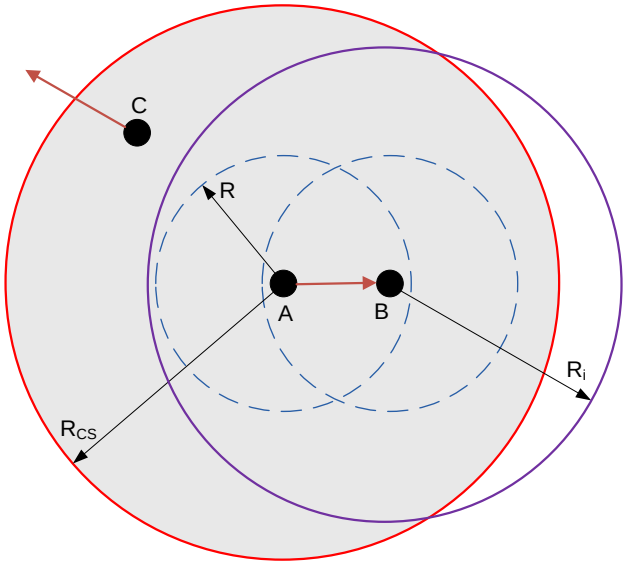

Figure 1.3: Exposed node problem

1.3 MANETs Protocol Stack

This section focuses on the traditional OSI protocol stack, depicted in Figure 1.4, for wireless networks [6]. The first layer in the protocol stack is the application and services layer, which occupies the top of the stack followed by the transport, network, data link, and physical layers. The application and service layer deals with the partitioning of tasks between fixed and mobile nodes as well as power management and *Quality Of Service* (QoS) management. Other layers in the protocol stack are discussed below.

1.3.1 Physical Layer

The first standard for *Wireless Local Area Networks* (WLAN), named *IEEE 802.11*, was released in 1997 by the IEEE 802.11 working group [2]. It gives specifications for the physical and media access control layers for WLAN. Following the success of the first standard, many IEEE 802.11 extensions have been released (i.e. 802.11a, 802.11b, 802.11e, 802.11g). These focused on achieving higher data rates and enhance QoS for real time applications [7]. The IEEE 802.11 standard supports two modes of operation for WLAN: infrastructure-based and infrastructure-less or ad hoc operation. Network interface cards can be set to work in either of these modes. Today, most wireless devices support the IEEE 802.11 standards, the most widely used standard in mobile ad hoc and infrastructure networks.

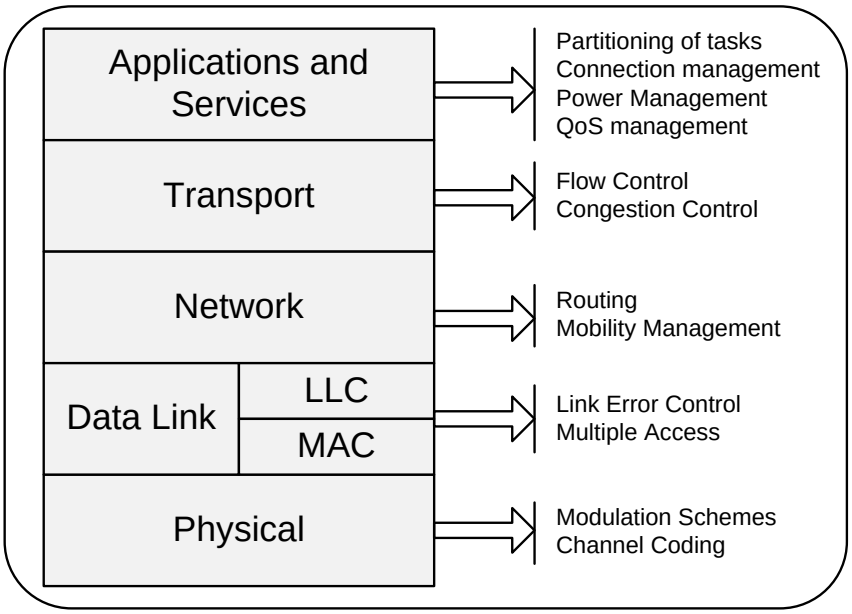

Figure 1.4: Protocol stack of wireless networks and corresponding main functions

The major functions and services performed by the *physical layer* are character encoding, modulation, transmission, reception and decoding. The IEEE 802.11 standard supports three physical layer technologies [1]: Direct Sequence Spread Spectrum (DSSS), Frequency Hopping Spread Spectrum (FHSS), and Diffused Infrared (DFIR). DSSS uses the radio frequencies ranging from 2.4 to 2.4835 MHz. It

uses a Differential Binary Phase Shift Keying (DBPSK) and Differential Quadruple Phase Shift Keying (DQPSK) modulation. FHSS uses the frequencies from 2.4 GHz to 2.4835 GHz, and a bandwidth of 83.5 MHz. It uses 2 and 4 levels Frequency Shift Keying (FSK) and divides the total bandwidth into 79 channels of 1 MHz each. It then hops between these channels in one of 78 orthogonal. DFIR is for indoor use only, and uses a wavelength range from 850 to 950 nm. The modulation technique used is Pulse Position Modulation (PPM).

1.3.2 Data Link Layer

The data link layer is divided into two sub-layers which are the *Logical Link Control* (LLC) and *Medium Access Control* (MAC) [7]. The LLC provides a way for the upper layers to deal with any type of MAC layer. It makes the IEEE 802.11 standard accessible to higher layers as a wired IEEE 802 LAN [8]. MAC layer protocols for wireless networks specify how nodes coordinate their communication over a common broadcast channel. They allow the wireless nodes to share their communication channel in a stable, fair, and efficient way. The typical tasks of MAC protocols are PDU (Protocol Data Unit) addressing, channel allocation, frame formatting, error checking, and fragmentation and reassembling. It is also the responsibility of the MAC layer to overcome the hidden and exposed node problems, resolve packet collisions between nodes, and conduct error corrections for packets experiencing corruptions in the physical layer [7].

The IEEE 802.11 standard [2] specifies the description of the MAC layer. The standard defines three frame types: management, control, and data. Management frames are used for timing, synchronization, authentication, and de-authentication. Control frames are used for handshaking and acknowledgments. For the transmission of data, data frames are used.

The MAC layer offers two different types of service [9]: a contention service (where any node that has a frame to transmit contends to access the channel) called the *Distributed Coordination Function* (DCF), and a contention-free service called the

Point Coordination Function (PCF). PCF is based on a polling scheme. It uses a point coordinator that regularly polls stations to give them the opportunity to transmit. The PCF cannot be used in the ad hoc mode of operation, and its setup in wireless devices is optional. DCF is the fundamental access method in the 802.11 MAC protocol for data transmission. It is based on a *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA) scheme. DCF is the only service operating in the ad hoc mode. Chapter 3 describes the IEEE 802.11 DCF MAC protocol in more detail.

1.3.3 Network Layer

In order to establish a connection between two nodes in MANET, the routing protocol in the network layer should first discover routes between them. Designing an efficient routing protocol for MANET is a challenge. This is due to the lack of infrastructure and frequent topology changes. Also, considering the limited power and bandwidth recourses in MANET, the routes should be constructed with minimum overhead and bandwidth.

Routing protocols for MANETs can be classified into two major categories [1]: proactive and reactive (or on-demand). The nodes using proactive routing protocols attempt to maintain up-to-date routing information to all nodes, regardless of the need for such information. They periodically propagate topology updates throughout the network to keep route tables up-to-date, consequently incurring a significant overhead. DSDV (Destination-Sequenced Distance-Vector routing) [10], OLSR (Optimized Link State Routing) [11], and TBRPF (Topology Broadcast Based on Reverse-Path Forwarding routing) [12] are examples of proactive routing protocols.

In contrast to proactive routing protocols, reactive routing protocols initiate a routing discovery only when a route is needed. They avoid the overhead due to the periodic updating of routing tables by adapting routing activities to traffic needs. Thus, they efficiently utilize the network bandwidth and reduce power consumption. In addition, they use route caches to store discovered routes for future use in order to reduce the overhead and latency of initiating a route discovery for each packet to be sent.

Several routing protocols use on-demand mechanisms, such as AODV (*Ad-hoc On-Demand Distance Vector*) [13], DSR (*Dynamic Source Routing*) [14], LAR (Location-Aided Routing) [15], and TORA (Temporally-Ordered Routing Algorithm) [16].

1.3.4 Transport Layer

The transport layer provides end-to-end communication services for applications. It provides convenient services such as connection-oriented data stream support, error control, flow control, congestion control and multiplexing. The two most common Transport layer protocols are *Transmission Control Protocol* (TCP) and *User Datagram Protocol* (UDP).

UDP is a simple transport layer protocol, which provides the best effort (or connectionless) service to transfer messages between nodes. UDP is not a reliable protocol because it does not provide either error or flow control. It is basically an interface between the network and application layers. Ports of the UDP protocol distinguish between multiple applications running on a single device. UDP was designed for applications for which extensive control features are not necessary, such as streaming audio and video applications.

TCP is a connection-oriented transport protocol that provides the essential flow and congestion control mechanisms required to ensure reliable packet delivery [17]. To use network bandwidth efficiently and control the flow of packets, TCP uses a mechanism known as a sliding window, which allows the sender to send multiple packets before waiting for an acknowledgment [18].

TCP congestion mechanisms prevent a sender from overrunning the capacity of the network. To avoid congestion, TCP maintains a limit called the congestion window, which restricts the amount of data sent. Several congestion control enhancements have been added to and suggested for TCP over the years. Congestion control mechanisms consist of four basic algorithms: slow start, congestion avoidance, fast retransmit, and fast recovery [17].

TCP provides reliable end-to-end data transfer through a technique known as positive acknowledgement with retransmission [18]. It assigns a sequence number to each byte transmitted, and expects a positive Acknowledgment (ACK) from the receiver. The sender starts a timer called the retransmission timer when it sends a packet. If the ACK is not received and the timer expires, then the data is retransmitted.

TCP was originally designed to work in wired networks where packet losses are mainly due to congestion. So, TCP uses packet loss as an indication of network congestion, and deals with this effectively by making a corresponding transmission adjustment to its congestion window. However, MANETs suffer from several other types of packet losses, such as those occurring due to excessive noise, interference, signal loss, lack of power, the collision of packets, and frequent route failures due to node mobility. Therefore, TCP is not well suited for mobile ad hoc networks [19]. Numerous enhancements and optimisations have been proposed to improve TCP performance for WLANs and MANETs [20-26].

1.4 Motivations, Objectives and Methodology

1.4.1 Motivations

Mobile ad hoc networks are becoming very attractive and useful in many kinds of communication and networking applications. This is due to their efficiency, simplicity of installation and use, low relative cost, and the flexibility provided by their dynamic infrastructure. High performance is a very important goal in designing communication systems such as MANETs. Therefore, the performance evaluation of ad hoc networks is needed to compare various network architectures for their performance, and to study both the effect of varying certain network parameters and the interaction between parameters.

It should be noted that most research into the performance of MANETs has been evaluated using *Discrete Event Simulation* (DES) utilising a broad band of simulators such as NS2 [27], OPNET [28], and GloMoSim [29]. The principle drawback of DES

models is the time and resources needed to run such models for large realistic systems, especially when highly accurate results (i.e., narrow confidence intervals) are desired. In other words, DES tends to be expensive because a large amount of computation time may be needed in order to obtain statistically significant results for MANETs.

In highly variable scenarios, with a number of nodes ranging from tens to hundreds, and node mobility varying from zero to tens of m/s, the simulation time of ad hoc networks will increase dramatically to unacceptable levels. For example, to run simulation experiments for an ad hoc network with five input factors, where each of these has only three values, would require $3^5 = 243$ experiments for all combinations of values. In addition, to obtain statistically reliable results with random node mobility, each experiment should be repeated many times with different mobility patterns. For ten repetitions, the total number of experimental runs would then be 2430. If each experiment is run sequentially for 60 minutes, the total time required to complete the experimental design would be about 101 days.

In addition to the large amount of computation time, it is difficult to study typical problems such as deadlock and concurrency in MANETs using DES because the network simulators implement the network at a low level of abstraction and specifications at a higher level cannot be supported. Due to the advantages of quick construction and numerical analysis, analytical modelling techniques, such as stochastic Petri nets and process algebra, have been used for the performance analysis of communication systems. In addition, analytical modelling is less costly and more efficient. It generally provides the best insight into the effects of various parameters and their interactions [30]. Hence, analytical modelling is the method of choice for fast and cost effective evaluation of ad hoc networks.

There are many challenges and characteristics associated with mobile ad hoc networks, as discussed in Section 2.1. Therefore, the performance of MANETs is affected by several factors, including traffic load, the number of nodes in the network, network area size, frequency of path failure and repair, mobility patterns, interactions

between protocols in different layers, and the effects of the wireless channel and wireless ranges (transmission, interference and carrier sensing ranges). Moreover, the behaviour of a node in a MANET depends not only on the behaviour of its neighbours, but also on the behaviour of other unseen nodes. Thus, mobile ad hoc networks are too complex to allow analytical study for explicit performance expressions. Consequently, in the literature, the number of analytical studies of this type of network is small [31-41]. In addition, most of these studies have many drawbacks, which can be summarized as follows:

(1) Most of analytical research in MANETs supposes that the nodes are stationary (no mobility) or the network is connected all the times to simplify the analytical analysis.

(2) In order to be mathematically tractable, most of analytical studies suppose that the nodes in the network area are uniformly or regularly distributed in the network.

(3) Some of the research is restricted to analysis of single hop ad hoc networks.

(4) The impact of the interference range on the performance of multi-hop ad hoc networks is either ignored or largely simplified.

(5) To simplify the analysis, most studies investigate MANETs in the case of a saturated traffic load (i.e. all the time every node has a packet to send) or finite load traffic.

(6) For computing the expected length (number of hops) of paths in multi-hop ad hoc networks, inaccurate methods were used.

(7) To reduce the state space of the analytical models of MANETs, most of the research is macroscopic (dynamics of actions are aggregated, motivated by limit theorems) and not scalable.

To the best of my knowledge, no analytical study so far has analysed the performance of multi-hop ad hoc networks based on the *IEEE 802.11 DCF MAC* protocol, where

nodes move according to a random mobility model, in terms of end-to-end delay and throughput. Therefore, this is the motivation of this thesis. This thesis presents an analytical framework, developed using the *Stochastic Reward Net* (SRN) [42, 43] and mathematical modelling techniques, for the modelling and analysis of multi-hop ad hoc networks based on the IEEE 802.11 DCF MAC protocol where nodes move according to the *random waypoint mobility model* (RWPMM). The proposed framework is used to analyse the performance of multi-hop ad hoc networks as a function of different parameters such as transmission range, carrier sensing range, interference range, node density, packet size, and packet generation rate.

The stochastic reward net modeling technique has been chosen because it allows the concise specification and automated generation of the underlying CTMC (Continuous Time Markov Chain). Moreover, compared to other Petri nets variants such as GSPN and SPN [42], it is the only technique that supports the specification of transition guards, transition rates, arc multiplicity, and the number of tokens as functions which are required to model complex communication systems such as MANETs.

1.4.2 Objectives and Methodology

The main objective of this work is to design an analytical framework that can be used to analyse the performance of multi-hop ad hoc networks in terms of throughput and end-to-end delay. Moreover, the proposed framework can be used to study the effects of various factors such as transmission range, carrier sensing range, interference range, the density of nodes, random access behaviour, packet size, mobility patterns, and traffic load on the performance of these networks. The proposed framework is validated via simulation using the network simulator NS2 [27].

To present an approach for the modelling and analysis of a scalable ad hoc network, there are two essential requirements. First, the model should be detailed enough to describe important network characteristics that have a significant impact on performance. Second, it should be simple enough to be scalable and analyzable. It is clear that these two requirements are potentially contradictory. Therefore, to model

multi-hop ad hoc networks using stochastic reward nets, we cannot construct a model for all nodes in the network by placing a model for each node into it one by one, because that would result in a state explosion problem. Alternatively, in the same way as introduced in previous analytical studies of multi-hop ad hoc networks [31-41], the large amount of symmetry in multi-hop ad hoc networks can be exploited in order to simplify the analysis, so that only the behaviour of a single hop communication between any two nodes in the network is modelled. Then, the single hop communication model is used to derive some parameters that are used to compute performance metrics such as delay and throughput for the whole path.

The single hop communication is modelled under the average workload computed for all possible instances of network topologies, taking into account the average effects of the random access behaviour of each node, the buffer overflow probabilities at each node, interference induced from neighbour and hidden nodes, and frequent path failure and redirection due to the random mobility of nodes. Because the underlying CTMC would be too large for numerical analysis, we cannot model the single hop communication using one SRN model. Therefore, in order to achieve this, a framework is proposed which is organized into several models to limit complexity.

The proposed framework consists of one mathematical model (called the network parameters model), and four SRN models (called the path analysis model, data link layer model, network layer model and transport layer model). The proposed framework is based on the idea of decomposition and fixed point iteration [44, 45] of stochastic reward nets. Thus, to derive any network performance metric, the SRN models are solved iteratively until the convergence of that performance metric is reached. The proposed framework describing the behaviour of a single hop communication is used to evaluate the delay and throughput per hop, which are then used to compute the end-to-end delay and throughput per path, as explained in Chapter 6. The next section describes the proposed framework in more detail.

1.4.3 Network Model and Assumptions

To develop a stochastic reward net model for MANET, we consider a network consisting of N nodes that are randomly distributed in a square area of dimension $L \times L$ and move according to the random waypoint mobility model [46]. All nodes are independent and behave identically. Each node is equipped with an omni-directional antenna and has a fixed transmission range R. Each node in the network is a source of traffic, where it generates packets with a Constant Bit Rate (CBR) λ. The packets are transmitted over a channel which is assumed to be noiseless. So, the error in packet reception caused by noise is not considered, whereas errors due to interference are taken into consideration. The destination of any source is chosen randomly from all other nodes.

The random waypoint mobility model is chosen as a mobility model because it is one of the most commonly used mobility models in MANET studies. In RWPMM, a node chooses a uniform random destination anywhere in the network area. Then, the node moves towards the destination point with a speed that is chosen uniformly from 1 to maximum speed (V_{max}). When the node reaches the destination, it may stop for a duration defined by the 'pause time' parameter. Then, it chooses and moves towards a new destination in a similar manner. To increase the mobility of nodes, the pause time is considered to be zero.

Path loss is the reduction in power density (or attenuation) of a signal as it propagates through space. Path loss may be due to many factors, such as free-space loss, refraction, diffraction, reflection, and absorption. To model signal propagation, different path loss models have been proposed in the literature. Free space, plane earth, diffraction, and the *two-ray ground* are examples of path loss models [47]. The two-ray ground path loss model is a simple model which considers both the direct path and the ground reflection path.

The transmission and carrier sensing ranges are determined by the transmission and reception power threshold and the path loss model of signal power. To simplify the analysis, the two-ray ground path loss model is adopted because we assume that the

ad hoc network is in an open space environment. In addition, both the carrier sensing and transmission ranges of all nodes are assumed to be fixed and identical. Compared to the transmission and carrier sensing ranges, the interference range is not fixed. It depends on the distance between the sender and receiver and the power of the sent and received signal [3].

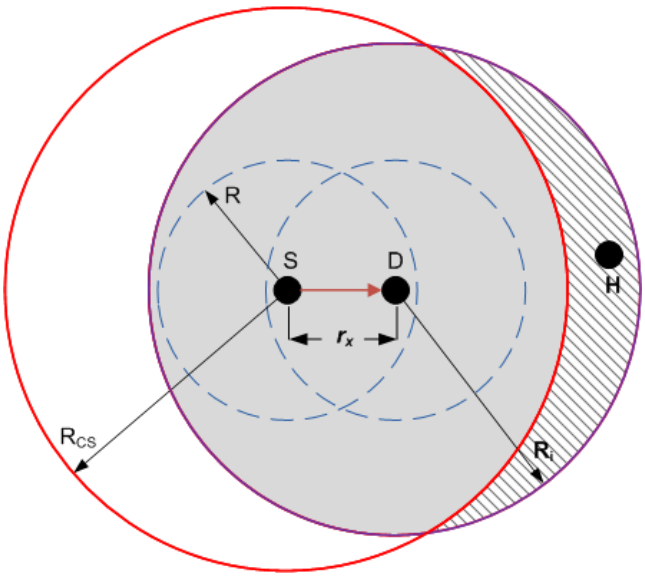

Figure 1.5: Single hop communication illustrating interfering and hidden nodes

1.5 Proposed Framework

Figure 1.5 shows a single hop communication between any sender node *S* and receiver node *D*, where the distance between *S* and *D* is r_x and the dashed area is the hidden area of the sender *S*. The area of intersection between the carrier sensing range of the sender and the interference range of the receiver is called the *interfering area*. The nodes located in this area are called *interfering nodes*. For example, for the sender *S* and receiver *D* shown in Figure 1.5, the interfering nodes are located in the shaded area. Any transmission from these nodes is sensed by *S*, but can interfere with simultaneous transmissions from *S* to *D*. As explained in Sections 1.1 and 1.2, the nodes located in the hidden and carrier sensing areas are called hidden and carrier sensing nodes respectively. In a single hop communication in multi-hop ad hoc

networks (such as that shown in Figure 1.5); hidden, carrier sensing, and interfering nodes have considerable effects on the transmission between the sender and receiver.

Modelling a single hop communication between any two nodes in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model is a multi-layer problem. The physical layer must adapt to rapid changes in link characteristics. The multiple access control layer should allow fair access, minimise collisions and transport data reliably over the shared wireless links in the presence of hidden or exposed nodes and rapid changes. The network layer protocols should determine and distribute information used to calculate paths in an efficient way. The transport layer should be able to handle frequent packet losses and delays that are very different from those in wired networks. In addition, the topology of MANETs is highly dynamic because of frequent node motion and so the effect of frequent path failure and redirection should be taken into account. Moreover, the single hop communication model should capture the different effects of hidden, carrier sensing, and interfering nodes.

From the above, it can be concluded that there are many interacting parameters, mechanisms, and phenomena in any single hop communication. Therefore, to limit the complexity of modelling a single hop communication in multi-hop ad hoc networks, and to avoid the state explosion problem, we propose a framework which is structured into several models. The proposed framework organises these models and the interactions between them in a way similar to the layers of the OSI protocol stack model and their interactions, explained in Section 1.3. The proposed framework defines the function of each model and the parameters which need to be computed in each model which are required in order to solve other models. This section describes the proposed framework.

Figure 1.6 illustrates the analytical framework used for modelling a single hop communication in multi-hop ad hoc networks. The meanings of the symbols used in Figure 1.6 are shown in Table 1.1. The proposed framework consists of five models which are divided into two groups; *mobility models* and *layer models*. The five

models interact with each other by exporting and importing different parameters, as shown in Figure 1.6. The mobility models are used to perform the analysis of the path between any source and destination. It consists of two models; the *network parameters model* and the *path analysis model*.

According to the number of nodes in the network (N), the mobility pattern (such as random waypoint, random walk point, free way, etc.), and the size of the network area (L^2), the network parameters model is used to compute the expected number of hops between any source-destination pair (N_h), and the average number of hidden, interfering and carrier sensing nodes. The network parameters model is a mathematical model and Chapter 4 introduces the first part of this model which is used to calculate the expected number of hops between any source and destination in MANETs where nodes move according to the random waypoint mobility model. The second part of this model, which extends the results introduced in Chapter 4 to compute N_H, N_i, and N_{CS}, is presented in Section 6.3.

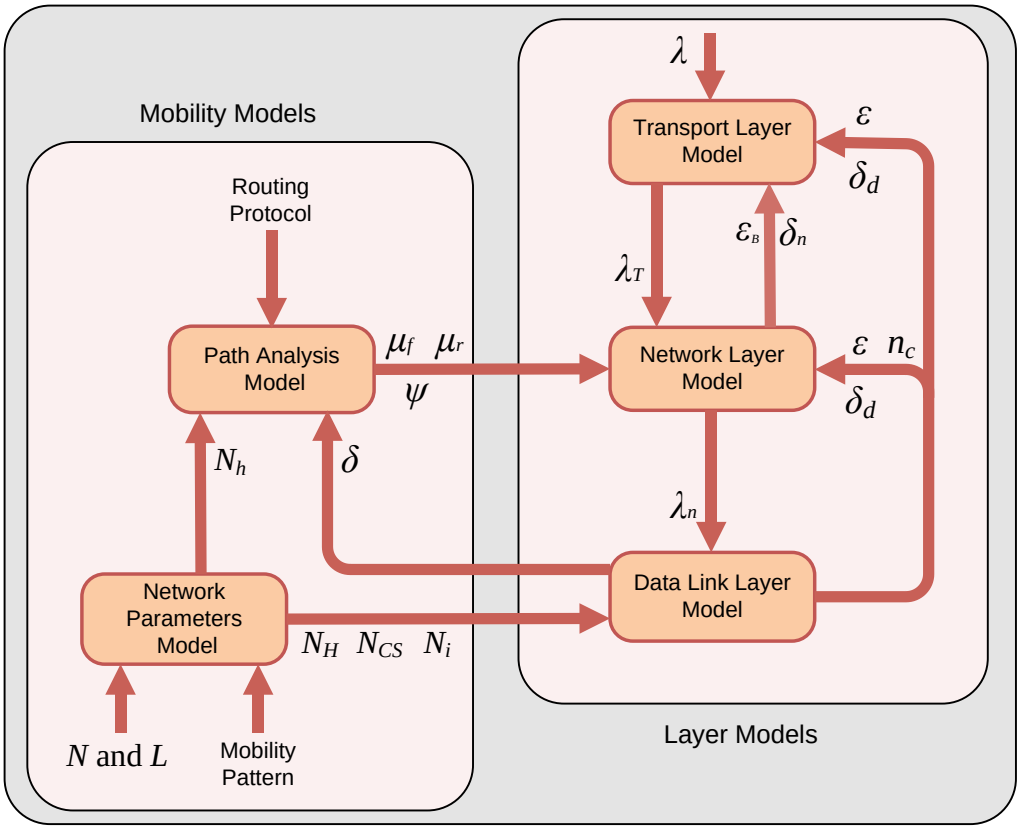

Figure 1.6: Proposed framework for modelling MANETs

Due to the mobility of nodes, mobile ad hoc networks have inherently dynamic topologies. Therefore the routes are prone to frequent breaks. Consequently, the routes followed by packets to reach their destinations vary frequently. This is a crucial factor that affects the performance of the network. The path analysis model is used to analyse the dynamics of paths in the network due to the mobility of nodes in terms of three measures: path connection availability (the probability that the path is available at any time) (ψ); average rate of failure (μ_f); and average rate of repair (μ_r). According to the routing protocol (such as AODV, DSR, or LAR), average packet delay per hop (δ); and N_h, the path analysis model is used to study the connection availability of paths and to calculate the average rate of failure and repair of the path between any source and destination. This model is a stochastic reward net model, and is described in Chapter 5.

Table 1.1: Meaning of symbols in Figure 1.6

Symbol	Meaning
λ	Packets generation rate
ε_B	Packet loss probability due to buffer overflow
δ_n	Average delay of packets in the network layer
λ_T	Throughput of the transport layer model
λ_n	Throughput of the network layer model
δ_d	Average packet delay in the data link layer
δ	Average packet delay per hop
ε	Packet loss probability in the data link layer
n_c	Average number of tries to transmit a packet
N	Number of nodes in the network
μ_f	Average failure rate of paths
μ_r	Average repair rate of paths
ψ	Average path availability
L	Side length of the squared network area
N_h	Average path length in hops
N_H, N_i, N_{cs}	Average number of hidden nodes, interfering nodes, and carrier sensing nodes

The layers model group consists of three models: the *data link layer model*, *network layer model, and transport layer model*. Data link layer protocols are modelled by the data link layer model. As explained in Section 1.3, the data link layer is divided into two sub-layers, which are the LLC and MAC. In wireless networks, the packet processing time in the LLC layer is negligible compared to that in the MAC layer [31-41]. Hence, the data link layer model only describes the behaviour of MAC layer protocols.

The data link layer model is an SRN model which uses the throughput of the network layer model (λ_n) to compute the average number of tries to transmit a packet (n_c), packet loss probability (ε), average packet delay in the data link layer (δ_d), and average packet delay per hop (δ). The data link layer model for a single hop ad hoc network based on the IEEE 802.11 DCF MAC protocol is presented in Chapter 3. This model is extended for multi-hop ad hoc networks in Chapter 6.

The actions in the network layer are modelled by the network layer model which uses the parameters ψ, μ_r, μ_f, λ_T (the throughput of the transport layer model), n_c, δ_d and ε to calculate the average number of packets per unit time that is sent to the data link layer model (λ_n), the probability of packet loss due to buffer overflow (ε_B), and the average delay of packets in the network layer model (δ_n). The network layer SRN model is introduced in Chapter 6. The transport layer model represents the analytical model for any of the transport layer protocols such as TCP or UDP. The inputs of the transport layer model are λ, ε_B, δ_n, δ_d, and ε, and the output is λ_T. To simplify the analysis, only the UDP protocol is adopted as a transport layer protocol. Because of its simplicity, the modelling of UDP is included in the network layer model introduced in Chapter 6.

The proposed framework is based on the idea of decomposition and fixed point iteration [44, 45] of stochastic reward nets. Therefore, the proposed SRN models are solved iteratively using the fixed point iteration technique to compute the required performance indices, such as the average delay and throughput per hop. This is explained in Chapter 6 which in addition shows how the performance indices per hop are used to compute the performance indices per path.

An important feature of this framework is that the only dependencies between the different models are the input and output parameters. Therefore it is clear that future researchers could adopt this framework and the underlying models, substituting their own models as and where they choose.

1.6 Contributions

The main contributions of this work are summarised as follows:

- For the first time, a structured analytical study is presented for the modelling and performance analysis of mobile ad hoc networks under a random mobility environment.
- A new stochastic reward net model is presented for the IEEE 802.11 DCF MAC protocol, for both BA and RTS/CTS methods, in single hop ad hoc networks in the presence of hidden nodes. Unlike previous studies, that adopt simplified assumptions to reduce the complexity of the proposed models which deviate from the IEEE 802.11 standard, the proposed model captures most of the features of the IEEE 802.11 DCF MAC protocol. The proposed model is used to demonstrate the effects of network parameters such as traffic load, packet size, and number of nodes.
- For the first time, an expression for the expected Euclidean distance between any source and destination nodes moving according to the random waypoint mobility model is derived in Chapter 4.
- A novel analytical approach called *Maximum Hop Distance* (MHD) is proposed in Chapter 4. This is used to compute the expected hop count between any random source-destination pair in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model. In addition, MHD is used to investigate the effect on the expected hop count of the number of nodes, network area size, and transmission range.

- To analyse path connection availability in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model, a new stochastic reward net model is proposed in Chapter 5.
- A closed form solution for path connection availability using the path analysis model is introduced in Chapter 5. Moreover, two performance metrics in the analysis of paths in MANETs are suggested: the path failure and repairing frequency.
- A stochastic reward net model is developed for the IEEE 802.11 DCF MAC protocol, for both BA and RTS/CTS methods, in multi-hop hop ad hoc networks with the random waypoint mobility model. The proposed SRN model takes into account the effects of hidden nodes, exposed nodes and interference from other nodes.
- A stochastic reward net model for the actions in the network layer is proposed in Chapter 6. The proposed model captures the effects of buffer overflow, packet receiving and forwarding, and dropping packets due to the unavailability of paths.
- The framework explained in Section 1.5 is proposed to model multi-hop ad hoc networks with the random waypoint mobility model. Also, an analytical procedure is presented in Chapter 6 that shows the sequence in which the proposed models are solved.

1.7 Thesis Overview

The remainder of this thesis is organized as follows. First, the related work is discussed in Chapter 2. This chapter highlights the limitations and merits of analytical studies which are directly related to the proposed models.

The data link layer model for single hop ad hoc networks is introduced in Chapter 3. The model represents the behaviour of the IEEE 802.11 DCF MAC protocol, for both BA and RTS/CTS methods, in a single hop ad hoc network with hidden nodes. The

effects of traffic load, packet size, and number of nodes on network performance in terms of throughput and delay are investigated.

In Chapter 4, we develop a simple closed form analytical approach to estimate the expected hop count between any random source-destination pair in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model. This represents the first part of the network parameters model.

The path analysis model explained in Section 1.5 is presented in Chapter 5. In this chapter, the proposed SRN model for analysing paths in multi-hop ad hoc networks with the random waypoint mobility model is described in detail. Also, a closed form solution is proposed for path connection availability and failure and repairing frequency. In addition, the influences of different factors on the path connection availability are investigated such as the number of nodes, transmission range, network area size, data transmission rate, and routing protocol.

In Chapter 6, the model introduced in Chapter 3 is first extended to model the IEEE 802.11 DCF MAC protocol for both BA and RTS/CTS methods in multi-hop hop ad hoc networks with the random waypoint mobility model. Then, the second part of the network parameters model is introduced which is used to compute the average number of hidden, carrier sensing and interfering nodes. Next, the network layer model explained in Section 1.5 is described. After that, the analytical procedure is presented that shows the sequence in which the proposed models are solved. The proposed framework is then validated using the network simulator NS2, and the analytical and simulation results are discussed in detail.

Finally, the thesis is concluded in Chapter 7 with a summary of the work and key results, and suggestions are made for future work.

Chapter 2

Related Work

This chapter introduces a brief summary of previous studies that are directly related to the proposed work in this thesis. First, previous studies investigating the performance of single hop ad hoc networks based on IEEE 802.1 DCF MAC protocol are discussed. Then, Section 2.2 discuses some relevant work that has been proposed to study the expected number of hops of paths in multi-hop ad hoc networks. The analytical models that have been developed to analyse the path connection availability and path life time are summarised in Section 2.3. Section 2.4 outlines the analytical studies that consider multi-hop ad hoc networks with random access MAC protocols. Finally, the analytical studies that have been proposed to investigate performance of transmission control protocol (TCP) in MANETs are discussed.

2.1 Single Hop Ad Hoc Networks

Since its development for WLAN, the IEEE 802.11 standard [2] has been widely used for various wireless networks due to its low cost and effectiveness in reducing collisions with simple and decentralised mechanisms. Many analytical studies have appeared in the literature investigating the performance of infrastructure and single hop ad hoc networks based on the IEEE 802.11 DCF MAC protocol. Bianchi [48]

proposed a Markov chain model to compute the saturation throughput and the probability that a packet transmission fails due to collision. The backoff mechanism of the IEEE 802.11 DCF protocol was studied under heavy traffic conditions. In addition, the proposed analytical model was a simplified version of the IEEE 802.11 DCF MAC protocol. The proposed model in [48] has been extended in [49] by including the discarding of the MAC frame when it reaches the maximum retransmission limit. In [50] the authors analysed the throughput and delay of CSMA/CA protocol under maximum load conditions by using a bi-dimensional discrete Markov chain. Also, the proposed model extend the model introduced in [48] by taking into account the busy medium conditions when invoking the backoff procedure. An additional transition state was introduced to Bianchi's model in order to model the freezing of the backoff counter. To simplify the analysis of the proposed model it was assumed that the access probability and station collision probability are independent of channel status.

Foh and Tantra [51] proposed an analytical model that improves the model introduced in [50] by relaxing its assumptions. The effect of post-DIFS (the time slot immediately following the DIFS guard time after a successful transmission) was modelled and the representation of the backoff freezing mechanism and maximum retry limit specified by the IEEE 802.11 standard were improved. This model assumes that the medium access probability depends on whether the previous period is busy or idle which makes the model more complicated. All these previous studies assumed that all stations in the network work in heavy traffic conditions (saturated traffic) where every station always has a data frame to transmit, which is rarely found in real-life applications. In addition, the proposed models only consider cellular networks where every station can communicate directly with all others.

Because of the complexity, few studies have been proposed to investigate the performance of the IEEE 802.11 DCF protocol under general traffic conditions [52-54]. In [52], the model was based on the presentation of the system with a pair of one-dimensional state diagrams which accommodate different input parameters. The model deviated from the 802.11 protocol standard because it assumed that all stations

collide or succeed at the same time. In [53], the authors modified Bianchi's Markov model to calculate the transmission probability of a station that may have different traffic loads, but the proposed model failed to capture some aspects of the standard, e.g. the station enters the backoff state if it receives a frame when the channel is busy.

Tickoo and Sikdar [54] proposed an analytical model based on a discrete time G/G/1 queue to study the performance of IEEE 802.11 MAC based wireless networks. A different approach was introduced to model the unsaturated traffic using a probability generating function that allows the computation of the probability distribution function of the packet delay. A unified analytical model for IEEE 802.11 MAC protocol in ad hoc networks with unsaturated conditions was presented in [55]. The proposed model was a combination of a 2D Markov chain model and an M/G/1/K queuing model. The optimal value of the total load and the optimal achievable performance metrics for the network were driven. The Markov chain model, which was based on the Bianchi's model [48], did not take into account the busy medium conditions when the backoff procedure was invoked.

Unfortunately, most of the previous studies have not addressed the problem of hidden nodes, despite of its importance in wireless networks. This is because it significantly complicates the mathematical analysis of IEEE 802.11 based systems. A small number of analytical studies [56-59] have been proposed considering the effect of hidden nodes on the performance of the IEEE 802.11 DCF protocol. Hou *et al* [56] presented an analytical study to compute the normalised throughput of the IEEE 802.11 DCF protocol with hidden nodes in a multi-hop ad hoc network. The drawback of this work is that it does not consider the state of the retransmission counter in obtaining the collision probability. In [57] the throughput of the IEEE 802.11 DCF scheme with hidden nodes in single hop ad hoc networks was analysed assuming that the carrier sensing range is equal to the transmission range, which is not generally applicable in the real world.

A simple analytical model was presented in [58] to derive the saturation throughput of MAC protocols in single hop ad hoc networks, although the model was only

validated under a heavy traffic assumption. The work in [59] introduced an analytical model for IEEE 802.11 DCF function in symmetric networks in the presence of the hidden node problem and unsaturated traffic. The model had inaccuracies, especially in high traffic load, because it assumes the collision probability is constant regardless of the state retransmission counter.

All previous studies evaluated the performance of IEEE 802.11 DCF protocol using mathematical and Markov chains models. The main drawback of these types of models is that if you need to modify or add a new feature to the operation of the protocol, you usually have to redesign the models from scratch. Petri nets and its variants (SPN, GSPN, SRN) [42] are a graphical tool used for formal depiction of systems whose dynamics are characterised by synchronisation, concurrency, conflict, and mutual exclusion, which are features of communication protocols, such as IEEE 802.11 DCF. They are a high-level formalism used for modelling very large and complex Markov chains. Compared to mathematical and Markov chains models, stochastic Petri nets models can generally be easily modified to cope with changes in the modelled system. Although the effectiveness of stochastic Petri nets has been demonstrated for modelling complex communications protocols, there are few studies that evaluate the functions of IEEE 802.11 DCF protocol using stochastic Petri nets [60, 61].

In [60] the authors modelled all stations in an IEEE 802.11 based WLAN in one SPN model. The complete model was solved using simulation because it was too large for direct analytical analysis, due to state space explosion. Although the authors introduced two compact analytical models, they did not include some aspects of the IEEE 802.11 DCF protocol, e.g. the effect of NAV on freezing and continuing of the backoff counter. Unfortunately, the results were not validated using network simulations. Jayaparvath *et al* [61] introduced an SRN model to evaluate the average system throughput and delay of the IEEE 802.11 DCF protocol. Although they succeeded in modelling the effect of freezing the backoff counter, they failed to model the retransmission retry counter. In addition, the proposed model did not take the RTS/CTS (Request-To-Send/Clear-To-Send) handshake into account and was

only verified for light load conditions. Neither [60] nor [61] model the effect of the hidden node problem.

2.2 Path Length in MANETs

In MANETs, the route or path is the sequence of mobile nodes which data packets pass through in order to reach the intended destination node from a given source node. The path length (number of hops) between the source and destination nodes is a key parameter in performance analysis of MANETs. Many studies have been issued to analyse how the performance of MANETs is affected by the hop count of paths [62-64]. The impact of hop count on searching cost and delay in ad hoc routing protocols has been investigated in [62]. Jinyang *et al* [63] have simulated the impact of different traffic patterns on the scalability of per node throughput. They showed that the network throughput deteriorates when the number of hops of the path increases due to interference between nodes. In [64], Gamal *et al* introduced a scheme to analyse the impact of the transmission range, degree of node mobility and number of hops on the trade-off between the delay and throughput in fixed and mobile ad hoc networks.

Although the impacts of the hop count of multi-hop paths on the performance of MANETs have been well recognized, there have been a very limited number of studies that focussed on the theoretical analysis of the expected number of hops in multi-hop paths in MANETs [65-68]. In [65], Jia-Chun and Wanjiun modelled the behaviour of packet forwarding on a multi-hop path for mobile ad hoc networks with high node density as circles centred at the initial location of the destination node. However, the results are not accurate because it is assumed that the progress per hop is equal to the transmission range. The relation between source-to-destination Euclidean distance and the hop count has been examined in [66]. The authors considered a greedy routing approach called Least Remaining Distance (LRD) which attempts to minimize the remaining distance to the destination in each hop. An analytical model for LRD and bounds on the number of hops for a given Euclidean

distance between source and destination has been developed. Unfortunately, the accuracy of the LRD approach is good only when the node density is very high.

In [67] an analytical model describing the hop count distribution for each source and destination pair in multi-hop wireless networks has been developed. Also, the trade-off between flooding cost and search latency for target location discovery, used in most ad hoc routing protocols, has been evaluated. The drawback of this work is that it supposed that the distance between the source and destination nodes is uniformly distributed, and the impact of the size of the network area is neglected. A mathematical model for the expected number of hops based on a Poisson randomly distributed network has been presented in [68]. The probability of n-hop count is derived and used to compute the expected number of hops. Unfortunately, all of these previous studies suppose that the nodes are stationary (no mobility) and are either uniformly or exponentially distributed over the network area.

2.3 Path Analysis in MANETs

Understanding the factors that affect the path connection availability in multi-hop ad hoc networks can help to understand the path stability under various degrees of system dynamics. In addition, the connection availability of paths can be used as a global measure for the performance of ad hoc networks. There are several works in the literature that have analytically studied the path connection availability and path life time in multi-hop ad hoc networks. In [69] Gruber and Hui investigated the average link expiration time for two-hop wireless ad hoc networks, where the source and destination are fixed. However, the influence of node density and routing protocol is not included. Based on a probabilistic model, the probability distribution of the lifetime of a routing path has been derived using a discrete-time analysis for the random walk mobility model in [70]. Expressions for broken link probabilities are derived by partitioning the area covered by the ad hoc network into a number of hexagonal cells where nodes roam around in cell-to-cell basis.

In [71], Yu-Chee *et al* used a two–state Markov model to characterize the wireless link lifetime in MANETs as a function of node mobility, where nodes move according to the random walk mobility model within a constrained area. A mathematical model has been proposed by Xianren *et al* [72] to estimate the route duration in MANETs when nodes move according to the random walk or random waypoint mobility models. This work extended the work introduced in [70] and [71] by relaxing their limiting conditions. The authors analysed the route duration in multi-hop paths by computing the minimum route duration of two-hop routes. The drawback of this work is that the authors assume that the probability density function (PDF) of the route duration for a two-hop route is known.

Pascoe-Chalke *et al* [73] derived statistical results of link and path availability properties using a mathematical model. They described a probability distribution function for link available time of one-hop link, assuming that nodes move according to the random walk mobility model, which has been used to investigate multi-hop cases. However, they did not take into account the effect of node density, routing protocol, and the size and shape of the intersection regions.

Markov chain models for a two-hop ad hoc network that incorporate three types of router failures were investigated by Dongyan *et al* in [74]. The proposed models were used to study the survivability of ad hoc networks where the excess packet loss and delay due to failures are evaluated as the survivability performance metric. Network survivability was also evaluated by John *et al* [75] using a generalized Markov chain model including more types of node failure than [74].

The path connection availability of a two-hop ad hoc network was presented in [76]. Analytical expressions for the leaving and returning rate in the intersection area between the source and destination were proposed. The authors tried to include the effect of routing protocols to the proposed Markov chain model, but they failed. In [77] Georgios, and Ruijie introduced a path connection availability model for wireless networks. They extended the proposed Markov model introduced in [74] by

combining it to a MAC buffer survivability model which has the properties of leaky buckets.

Unfortunately, none of the previous work provide a closed form solution for analytical analysis of the path connection availability or path life time for multi-hop ad hoc networks. Also, to simplify the analysis, most studies suppose that the source and destination nodes are static and two-hop apart and other intermediate nodes move according to the random walk (direction) mobility model. Random walk was chosen as the mobility model because its spatial node distribution is uniform, making the analytical analysis simpler. In addition, there is no investigation of the impact of different ad hoc routing protocols on the path connection availability.

2.4 Multi-hop Ad Hoc Networks

Many analytical studies have appeared in the literature investigating the performance of wireless single hop ad hoc networks with a random access MAC protocol, as discussed in Section 2.1. However, because performance modelling and analysis of multi-hop ad hoc networks is much more challenging, few papers addressed this issue [31-41].

The first attempt studied the performance of wireless multi-hop ad hoc network with a random access MAC protocol [31]. To analyse the saturation throughput in wireless multi-hop ad hoc networks, a simple analytical model was proposed. The transmission probability for a single hop was derived which was used to investigate multi-hop scenarios. To simplify the analysis, nodes were distributed in the network according to the Poisson distribution. Moreover, the status of the channel and backoff behaviour of the MAC protocol were simplified into limiting probabilities.

The performance of the IEEE 802.11 DCF protocol in multi-hop network scenario was investigated in [32] using an analytical model. The proposed model used a two-dimension Markov Chain model introduced in [48] to derive an expression for the transmission probability which was used to compute the packet collision probability

taken into account the impact of the hidden node problem. Although the proposed model takes into account the effects of hidden and interfering nodes, the nodes in the network were regularly placed in a grid topology in order to simplify the analysis.

In [33], an approximate analytical model for the performance analysis of a single hop and multi-hop ad hoc network was presented. The behaviour of the DCF MAC layer protocol was modelled using the Markov chain model introduced in [48]. For single hop scenarios, to derive an expression for the queuing delay, and distribution function and first moment of the service time, the M/G/1 queuing system has been adopted. The authors extended the analytical model for a single hop network to model a multi-hop network. They derived expressions for the probabilities of collision occurrence due to the hidden node problem. However, in multi-hop scenarios, they only addressed the approximate throughput and the end-to-end delay has not been considered.

In [34] Wang *et al* presented an analytical model for the performance analysis of wireless ad hoc network with the 802.11 DCF MAC protocol under finite load conditions in terms of the network throughput and delay. The model is limited for a chain network topology and hidden node problem was not considered. A model called Traffic-Analysis-Based (TAB) for the throughput analysis of wireless ad hoc networks with a chain topology was proposed in [35]. The TAB model is used to analyse the state transition process of the wireless nodes with increasing traffic load. The backoff states of wireless nodes have been presented using the approximate model introduced in [48].

Ali *et al* [36] presented approximate analytical models to estimate the throughput and delay per node in wireless multi-hop ad hoc networks. They used the Markov chain model introduced in [48] to model the channel access and backoff behaviour of the MAC protocol. In addition, the random network topologies are generated using a two-dimension Poisson distribution for the node location in the network. The authors did not derive an expression for either delay or throughput per path. Also, they did not consider the traffic load induced by the routed packets received from neighbour

nodes. Kumar *et al* [37] proposed an analytical model for estimating the average end-to-end delay of multi-hop ad hoc networks in which the IEEE 802.11 DCF protocol are used at the MAC layer. This work has not considered the packet queuing delay and has not been validated using random network topologies.

An analytical model for random access MAC based wireless ad hoc networks using open G/G/1 queuing networks has been introduced in [38]. The performance of single and multi hop scenarios were investigated in terms of the throughput and end-to-end delay. The proposed model is used to derive a closed form expressions for the maximum achievable throughput and end-to-end delay. The single hop communication was modelled as an open queuing network which is used to evaluate the mean and second moment of the packet service time per hop. Then, to derive expression for end-to-end delay, the diffusion approximation was adopted to solve the open queuing network. Also, the average service time per hop was used to obtain the expression for the maximum achievable throughput. However, although the main target of the proposed queuing model was gaining insights into the queuing delay, dropping of packets due to the buffer overflow has not been considered. In addition, effects of hidden and interfering nodes which increase in multi-hop networks have not been taken into account.

In [39], Ghadimi *et al* extended the work introduced in [33] to address the end-to-end delay analysis in multi-hop wireless ad hoc network under unsaturated traffic condition considering the hidden and exposed terminal problem. Each single wireless node was modelled as an M/G/1 queue which is used to compute service time distribution function. Using the service time distribution function for a single hop, the probability distribution function of a single hop delay and its first and second moment were obtained. In addition, the probabilities of collisions in both hidden and exposed node conditions were calculated using the single node media access delay distribution, which was used to extend the modelling approach to investigate the delay in multi-hop scenarios. This work used the Markov chain model introduced in [48] to model the transmission state of each node that follows the 802.11 DCF MAC protocol. This model deviates from the standard because it much simplifies the IEEE

802.11 MAC protocol [49]. Moreover, in multi-hop scenarios, the method used to compute the expected number of hops is not accurate.

An approximate stochastic Petri net model for ad hoc network was presented in [40]. The proposed model tried to exploit the symmetry between nodes by describing the behaviour of one node under a workload that is generated by the whole network. The SPN model consists of two subnets; incoming and outgoing subnets. The incoming subnet represents the processing of packets received from other nodes, whereas the outgoing subnet models the transmission of packets generated in the current node. Fixed point iteration was used to solve the proposed model. Lin *et al* [41] modified the work introduced in [40] to be suitable for a heavily loaded network. To model the sending and receiving process in ad hoc network, they adopted independent and receiving buffers. Also, they introduced a more accurate method for calculating the packet dropping probability. The main drawbacks of the work introduced in [40] and [41] are (1) although the MAC protocol plays a prominent role in the performance of ad hoc networks, the proposed SPN models in both [40] and [41] did not capture the behaviour of any MAC protocol, (2) the effects of hidden and interfering nodes on the performance of the network have not been considered.

2.5 Analysis of TCP in MANETs

TCP (Transmission Control Protocol) is a transport layer protocol designed for reliable end-to-end communication in wired networks. Because TCP is the most common transport protocol, the majority of wireless networks run TCP. However, because TCP was not created for wireless networks, the interaction between the MAC and TCP protocol causes serious performance issues in wireless networks [78, 79]. Due to the complexity of the transmission control protocol (TCP), a few analytical studies have been proposed to investigate its performance in MANETs [80-84]. The authors in [80] studied the performance of TCP traffic over a multi-hop wireless network where all nodes share the same physical channel and use the IEEE 802.11 MAC protocol. They tried to get an upper bound of the throughput of TCP over a

multi-hop network with a string topology. This work is based on many simplifying assumptions such as the instantaneous ACK delivery and constant contention window size for the IEEE 802.11 MAC protocol, and constant TCP congestion window size.

In [81], Kherani and Shorey presented a mathematical model of TCP over IEEE 802.11 two-hop networks using a simple topology consisting of a linear chain of nodes. To simplify the analysis, many assumptions have been imposed, e.g. the congestion window size was considered to be constant, nodes can directly transmit to each other and the backoff timer of the IEEE802.11 MAC protocol has been assumed to be geometrically distributed. A multi-dimensional Markov chain model for analysing TCP performance in ad hoc networks was presented in [82]. The authors attempted to provide more accurate model for TCP by considering the main phases of TCP (the slow start and congestion avoidance). They modelled the effect of changing the congestion window size with changing the state of the system.

To analyse TCP performance in multi-hop ad hoc networks with a string topology, a Markov chain model was proposed in [83]. The proposed model considers the spatial reuse of the wireless channel, contention of nodes to access the wireless channel, and packet buffering in intermediate nodes. A Markov chain model for a single hop is used to predict the throughput of multi-hop scenarios. The results show that the throughput is independent of the TCP congestion window size if the TCP session crosses a fixed number of hops. This work did not consider the effect of packet dropping due to the buffer overflow and packet loss due to collision or link layer contention.

In [84], the authors presented an analytical model developed using the stochastic reward net (SRN) modeling technique for the behaviour of a TCP variant called TCP Reno in wireless local area networks. This work evaluates the behaviour of the stationary TCP flow and investigates the fairness problem in WLANs between the upload user and download user. The work introduced in [82-84] did not consider the behaviour of MAC or routing protocols.

Chapter 3

Performance Analysis of the IEEE 802.11 DCF MAC Protocol

3.1 Introduction

MAC layer protocols for wireless networks specify how nodes coordinate their communication over a common broadcast channel. This allows the wireless nodes to share their communication channel in a stable, fair, and efficient way. MAC layer protocols should address several problems such as hidden and exposed nodes, and higher error rates. They can be broadly classified as contention and contention-free (schedule) based protocols. Contention-free based MAC protocols require coordination between nodes where they are following some particular schedule which prevents collision of packets. In contention-based MAC protocols, the nodes do not need any coordination between themselves to access the channel. Consequently, there is still a possibility of packet collision.

Contention-based MAC protocols, also known as random access protocols, have been widely used in wireless networks because of their simplicity and ease of implementation. Pure ALOHA [85] and Slotted ALOHA [86] were the first contention-based MAC protocols; many other protocols have been proposed

subsequently. Carrier Sensing Multiple Access (CSMA) [87] significantly enhanced the throughput of ALOHA-like protocols. It requires sensing the channel for the ongoing transmission before sending a packet. If the channel is busy, the node defers its transmission for a random period of time before retrying the transmission. Hence, CSMA reduces the possibility of collisions at the sender-side. Multiple Access Collision Avoidance (MACA) [88] and its variant MACAW (Multiple Access with Collision Avoidance for Wireless) [89] are alternative medium access control protocols for wireless networks that improve CSMA by taking steps to avoid the hidden node problem. They attempt to reduce the possibility of collisions at the receiver side.

The Floor Acquisition Multiple Access (FAMA) [90] protocol consists of both non-persistent carrier sensing and a collision avoidance handshake between the source and destination of a packet. It provides another solution for the hidden node problem. Before sending any frame, the node has to acquire the control of the channel to avoid the collision with any other packet. Carrier Sensing Multiple Access with Collision Avoidance (CSMA/CA) is a variant of the FAMA protocol that combines properties of CSMA and MACA. It uses two small control packets to mitigate the hidden node problem. CSMA and its enhancements with Collision Avoidance (CA) and Request To Send (RTS) and Clear To Send (CTS) mechanisms have led to the IEEE 802.11 standard for Wireless Local Area Networks [2].

Since its development for WLAN, the IEEE 802.11 standard has been widely used for various wireless networks due to its low cost, effectiveness in reducing collisions with simple and decentralised mechanisms and the wide availability of IEEE 802.11 hardware. It has been widely deployed in many electronic devices such as personal computers, laptops, and mobile phones.

The IEEE 802.11 MAC protocol defines two different access methods, the Distributed Coordination Function (DCF) for traffic without quality of service, and the Point Coordination Function (PCF) for traffic with QoS requirements. PCF, which is only used on infrastructure networks, is built on top of DCF. PCF uses a point

coordinator (access point) to determine which node has the right to transmit. The PCF mode is not widely implemented and its setup in warless devices is optional. DCF is the fundamental access method of the IEEE 802.11 MAC protocol for data transmission and it is the only service operating in the ad hoc mode. It is described in depth in Section 3.2.

There are numerous analytical studies that evaluated the performance of *IEEE 802.11 DCF MAC* protocol in WLAN [48-54, 56-59]. The studies introduced in [48-51] do not consider finite load situations which are important practical conditions in real-life applications. A few studies have been proposed to investigate the performance of the IEEE 802.11 DCF protocol under general traffic conditions [52-54]. However the proposed models did not consider the hidden node problem. The effect of the hidden node problem on the performance of the IEEE 802.11 DCF protocol has been discussed in [56-59], but not modelled precisely.

Most analytical studies have used mathematical and Markov chains models to evaluate the IEEE 802.11 DCF protocol. If the protocol is modified, these models are generally difficult to modify and they need to be redesigned from scratch. Petri nets are a high-level formalism used for modelling very large and complex Markov chains that can be easily modified to cope with many changes in the modelled system. A few Petri nets models have been proposed to evaluate the function of IEEE 802.11 DCF protocol in WLAN [60, 61], but the protocol has not been modelled accurately and the hidden node problem has not been addressed.

This chapter presents a novel SRN model for performance evaluation of IEEE 802.11 DCF MAC protocol in single hop ad hoc networks in the presence of hidden nodes, taking into account the characteristics of the physical layer, different traffic loads, packet size, and carrier sensing range. The proposed model captures most features of the protocol. It consists of two interacting SRN models: the *one node detailed model* and *abstract model*. All of the detailed activities in any mobile node in the network are represented in the one node detailed model. The abstract model describes interactions between all nodes in the network. The two models are solved iteratively

until the convergence of the performance measures is reached. Performance measures such as the goodput and packet delay for various network configurations are computed.

3.2 IEEE 802.11 DCF MAC Protocol

The IEEE 802.11 DCF MAC protocol is basically a Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol. The carrier sensing function is performed at both the MAC and physical layers. Physical carrier-sensing functions are provided by the physical layer by using a channel sensing function called Clear Channel Assessment (CCA). CCA analyses all detected packets from other nodes and detects activities in the channel by analysing relative signal strength. Virtual carrier sensing functions are provided by the MAC layer by using the Network Allocation Vector (NAV). NAV is a timer that decrements irrespectively of the status of the medium and is updated by frames transmitted on the medium. Any node considers the channel is busy if the carrier sensing indicates the medium is busy or the NAV is set to a value greater than zero. As long as the NAV is set to a non-zero value or the node senses the channel as being busy, the node is not allowed to initiate transmissions. The collision avoidance portion of CSMA/CA is performed through a random backoff procedure which is illustrated below.

According to the IEEE 802.11 WLAN media access control standard [2], DCF uses one of two access methods depending on the packet size: Basic Access (BA) and Request-To-Send/Clear-To-Send (RTS/CTS). If the size of packet is less than or equal a configurable parameter called RTS-threshold, DCF uses the BA method. However if the size of the packet is greater than the RTS-threshold, DCF uses the RTS/CTS method. As shown in Figure 3.1, BA is a two-way handshake method because it uses only data and ACK frames. However, RTS/CTS is a four-way handshake because it uses RTS, CTS, data, and ACK frames. Only the first frame in both cases contends to access the medium.

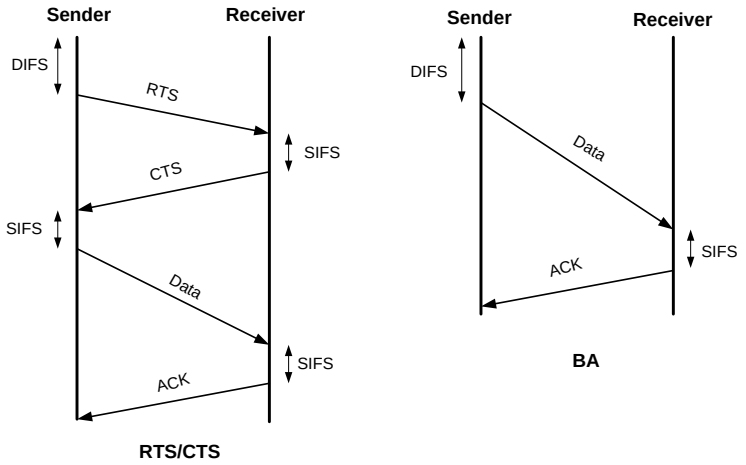

Figure 3.1: The BA and RTS/CTS methods handshake

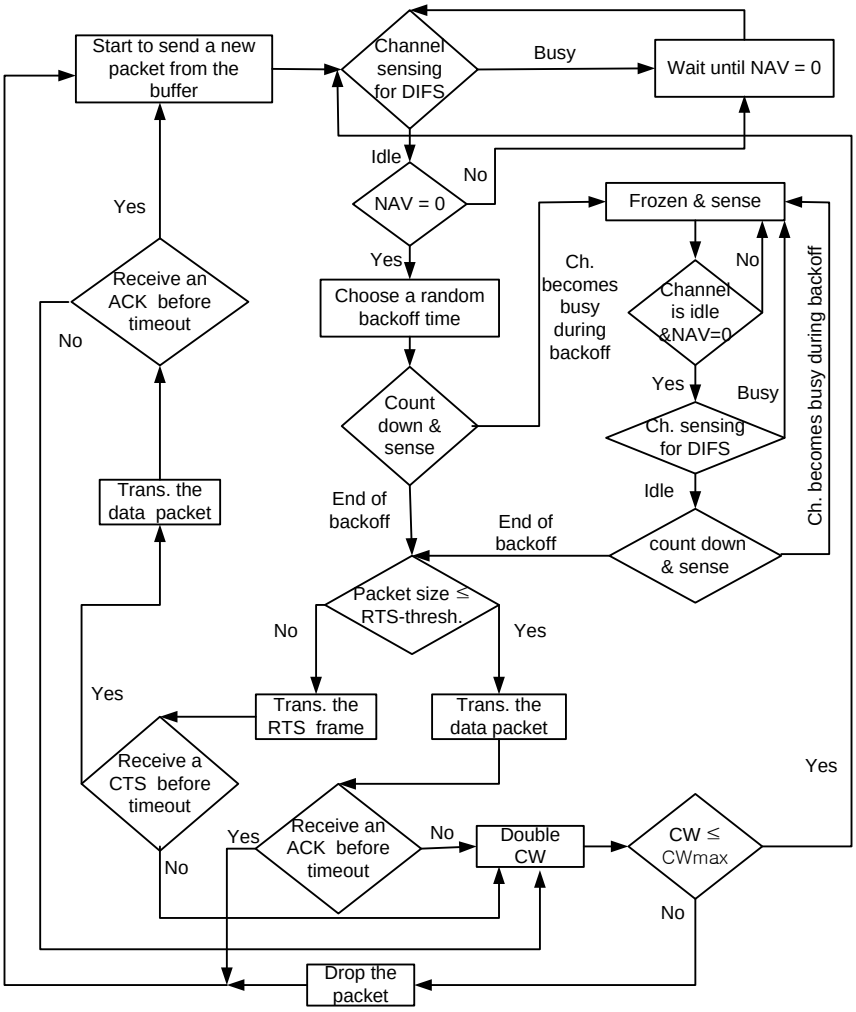

Figure 3.2: Block diagram for the operation of the BA and RTS/CTS methods

Figure 3.2 illustrates the operation of the BA and RTS/CTS methods. Also, Figure 3.3 shows the timing diagram for the operation of the RTS/CTS method. To send a new data packet, the node first has to sense the channel. If the channel is idle for a specific amount of time, known as DCF Inter Frame Space (DIFS), and the network allocation vector (NAV) equals zero, the node proceeds to transmit the packet. During sensing the channel for the DIFS interval, if the channel becomes busy (the NAV of the node is set to a non-zero value) the node waits until the NAV is reset to zero and starts again to sense the channel for a DIFS interval.

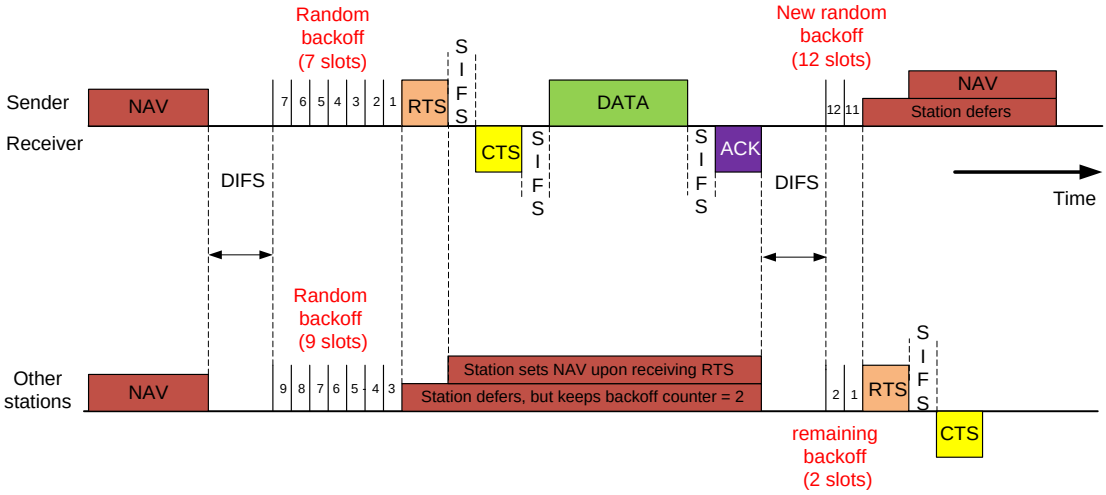

Figure 3.3: Timing diagram for the operation of the RTS/CTS method

If two or more nodes try to send a MAC frame at the same time, and they detect the channel as being idle for the DIFS interval, a collision occurs when these nodes start to transmit their frames. DCF defines a Collision Avoidance (CA) mechanism to reduce the probability of such collisions. Any node has to defer for a random backoff time before starting a transmission in order to resolve medium contention conflicts. The backoff time is slotted in time periods called the slot time (T_s) which depends on the physical layer standard. Any node is permitted to transmit only at the beginning of each slot time. The random backoff time equals $K \cdot T_s$, where K is an integer number that is uniformly chosen from the range [0, CW], and CW is the contention or backoff window. CW is calculated from the following equation:

$$CW = (CW_{min} + 1) \cdot 2^{\beta_c} - 1$$

where CW_{min} is the minimum contention window and β_c is the backoff counter (retry counter) that counts the number of failures of sending a packet; β_c increases by one each time a transmission fails. At the first attempt to transmit a packet, β_c is initialised with zero and then it is incremented by one at each retransmission for the same packet. β_c increases to its maximum value, called Maximum Retry Limit (MRL), corresponding to the maximum contention window (*CWmax*). After successful transmission of any packet, β_c is reset to zero.

During the backoff stage, the node uses the physical and virtual carrier sensing mechanisms to determine whether the channel is idle or busy. As long as the channel is idle and NAV = 0, the backoff timer decreases (counts down) by a slot time, as shown in Figure 3.3. At the beginning of any slot, if the channel is sensed busy or NAV > 0, the backoff timer is frozen. If NAV is reset to zero and the channel is sensed idle for a time greater than DIFS, the backoff timer resumes decreasing. In the case of the RTS/CTS method, if the channel is sensed idle for a period greater than 2·SIFS + t_{CTS} + 2·T_s (where t_{CTS} is the transmission time of CTS frame and SIFS (Short Inter Frame Space) is a time interval defined by the standard) then NAV is reset and the backoff timer resumes decreasing. Finally, depending on the packet size, the data frame or RTS frame is transmitted when the backoff timer reaches zero. If the packet size greater than RTS-Threshold, then the RTS/CTS method is used; otherwise, the BA method is used.

In the case of the BA method, when the receiver receives the data frame sent by the source it waits for a SIFS interval, then it sends the ACK frame. The SIFS interval is less than the DIFS interval and the slot time; so the channel will not be free for a period greater than or equal to the DIFS interval. Consequently, all other nodes wait until the end of transmission of the ACK frame. Because the CSMA/CA does not depend on physical collision detection, it uses the ACK frame as logical collision detection. If the source node does not receive the ACK frame within the timeout

period, it increases the retry count by one, which doubles the CW, and starts retransmission of the same packet.

In the case of the RTS/CTS method, when the receiver node receives the RTS frame, it responds after the SIFS interval with a CTS frame. The source node sends the data frame after the SIFS interval if it correctly received the CTS frame. Also, the receiver node sends an ACK frame after the SIFS interval if it correctly received the data frame. If the source does not receive the CTS or ACK frame within a specified timeout, it increases the retry count by one, which doubles CW, and starts retransmission of the same packet. According to the standard, for all MAC frames the physical header is transmitted with minimum bit rate (B_1), whereas the MAC Protocol Data Unit (MPDU) is transmitted with a higher rate (B_2).

Each MAC frame is associated with a single retry counter. Depending on the size of the MAC frame, there are two retry counters that can be associated with frames: the Short Retry Counter (SRC) and the Long Retry Counter (LRC). If the size of the frame is less than or equal to the RTS-threshold (short frame), the frame is associated with SRC. Otherwise, the frame is associated with LRC. The retry counter is increased every time the transmission of MAC frames fails. However, when the transmission of a MAC frame succeeds, the retry counter is reset to zero. Retries for failed transmission attempts continue until the short or long retry counter reaches the maximum retry limit. When any of these maximum retry limits is reached, retry attempts will stop, the retry counter is reset to zero and the MAC frame is discarded.

After transmitting the data (or RTS) frame, all nodes in the transmission range of the sender receive the data frame. According to the duration field value in the data (or RTS) frame, all nodes hearing the frame set their NAV. The duration field defines how long the subsequent frames exchange may take. As long as NAV is set to a value greater than zero, the node is not allowed to initiate transmissions, thus reducing collisions in subsequent frames.

3.3 Network model and assumptions

For performance modelling of the IEEE 802.11 DCF MAC protocol in single hop ad hoc networks with hidden nodes, we consider the network architecture shown in Figure 3.4. The network consists of M independent stationary nodes distributed in a square area. There are N neighbour nodes (e.g. S_1 to S_8 in Figure 3.4) where each node can transmit to all of the other nodes, i.e. they are in the transmission range of each other. We call the area where the N neighbour nodes are distributed as the *active area*. Each node in the active area generates packets with the rate λ and sends them to a destination D_x, which has N_h nodes in its interference range that are hidden from the source. For example, in Figure 3.4, the nodes S_1 and S_2 send their packets to D_1. The nodes S_{h1} and S_{h2} are hidden from the nodes S_1 and S_2 because they are in the interference range of D_1 and are not in the carrier sensing range of either S_1 or S_2.

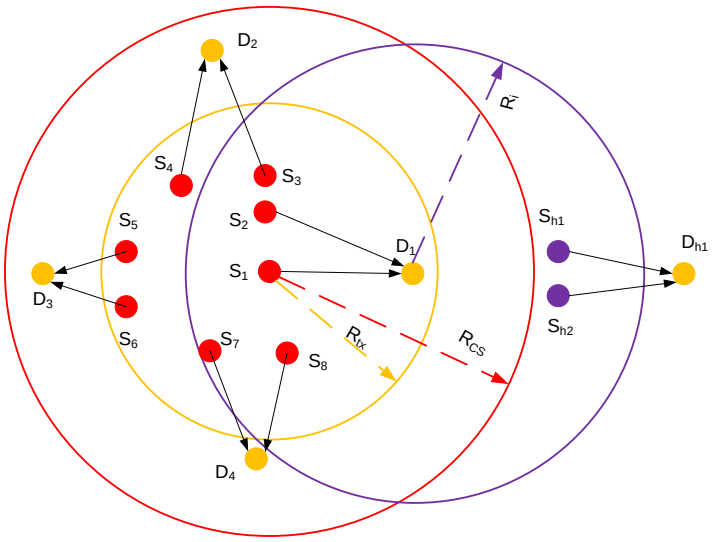

Figure 3.4: The Network architecture for a single hop ad hoc network with hidden nodes

Each of the hidden nodes generates packets at a rate λ_h and sends them to a different destination D_{hx}, e.g. the nodes S_{h1} and S_{h2} send their packets to the destination D_{h1} as shown in Figure 3.4. The nodes that are hidden from a source S can sense each other. N, N_h, and λ_h are the model parameters that are varied to different values, as explained in Section 3.6. Also, λ is a parameter of the model that is varied through a wide range

of values, from a small to a large value, in order to represent conditions of light and heavy load. To eliminate the effect of network layer protocols, because we are interested in modelling the effect of hidden nodes on the performance of MAC layer protocols, any destination node is located in the transmission range of the source. All nodes have multi-directional antennas. A two-way path loss propagation model is used for simulation and analysis. The radio channel is assumed to be free of noise errors. Also, it is supposed that the MAC protocol does not use fragmentation and management frames (such as beacon frames).

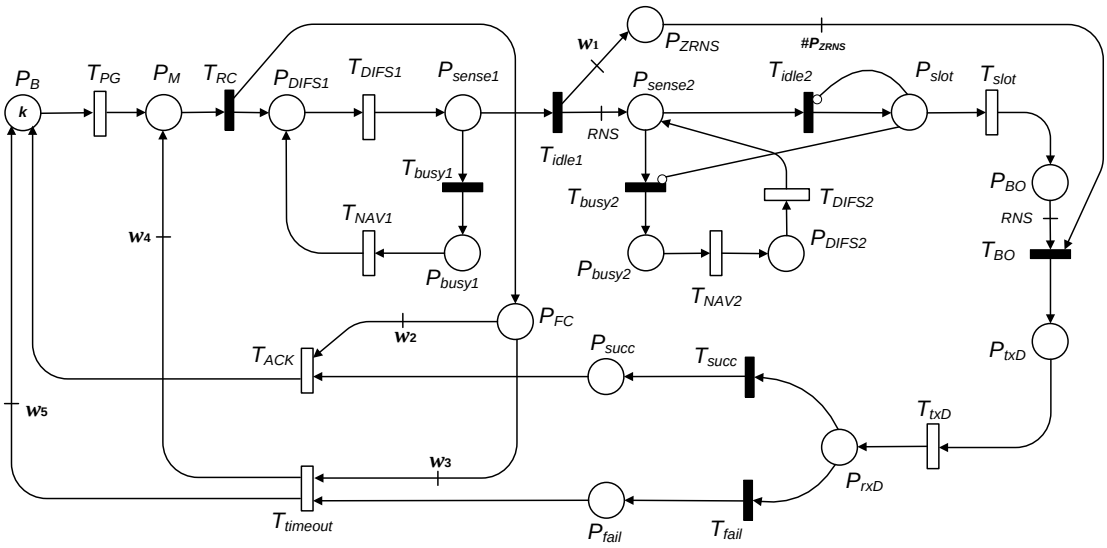

Figure 3.5: One node detailed model for the BA method

3.4 Model Description

To model the system shown in Figure 3.4 using stochastic reward nets, the model should capture the behaviour of the IEEE 802.11 DCF MAC protocol, interaction between the nodes in the active area, interaction between nodes in the hidden area and how hidden nodes affect the nodes in the active area. If we modelled all these actions in one model it would be prohibitively difficult to solve due to the state space explosion problem. So, to model the system we propose two interacting SRN models which depend on lumping and decomposition techniques. The two models are solved iteratively until convergence of the performance measures. The two models are (1)

the one node detailed model (Figures 3.5 and 3.6) which describes all detailed activities in one node in either active or hidden area, (2) the abstract model (Figures 3.7 and 3.8) which describes the interaction between the nodes within the active area, between the nodes within the hidden area and between the nodes in the active and hidden area. The two models are described below.

3.4.1 One Node Detailed Model for the BA Method

In this subsection the one node detailed model for the BA method is described. Figure 3.5 shows the SRN model of the one node detailed model for the BA method. The number of tokens in the place P_B represents the free places that are available for frames in the buffer of the MAC layer of the node. The number of tokens in the place P_B is k. Because the MAC layer transmits only one packet (the packet at the head of the queue) at each time, k is set to 1. The generation of packets from upper layer is modelled by the transition T_{PG}. The firing of T_{PG} deposits a token in the place P_M, which represents that the MAC layer has received the packet and started the transmission process.

Each MAC frame is associated with a single retry counter (backoff counter) that counts the number of failures to transmit it. The place P_{FC} models the retransmission retry counter. The number of tokens in this place represents the number of tries to transmit the MAC frame. The firing of the transition T_{RC} deposits a token in the places P_{DIFS1} and P_{FC}. The place P_{DIFS1} represents that the node is sensing the channel for a DIFS period. The firing of the transition T_{DIFS1} represents the end of sensing the channel after the DIFS period, and so it deposits one token in the place P_{sense1} that models the end of sensing the channel. At this point there are two probabilities:

- The channel is idle during sensing the channel for the DIFS period which is modelled by the firing of the immediate transition T_{idle1}.
- Any of other neighbour nodes is using the channel (channel is busy) when the node try to sense it for the DIFS period which is modelled by the firing of the immediate transition T_{busy1}.

If the channel becomes busy during sensing it for the DIFS interval, this means that one of the neighbours is sending a packet. So, the node has to wait until the neighbour node finishes sending the packet to start again to sense the channel for the DIFS interval. This is represented by depositing a token in the place P_{busy1} after the firing of T_{busy1}, which is returned back to the place P_{DIFS1} after the firing of T_{NAV1}. The probabilities of firing of the transitions T_{busy1} (β_1) and T_{idle1} $(1-\beta_1)$ are the probabilities that the channel is busy and idle respectively. The parameter β_1 is computed from the abstract model, as explained in Section 3.4.3.

As shown in Table 3.1, the average firing time of transition T_{DIFS1} is the DIFS interval. The average firing time of the transition T_{NAV1} equals the time required to send a data frame and receive an ACK frame. In all tables, $Ft(T_x)$ is the average firing time of transition T_x.

The firing of the transition T_{idle1} deposits a number of tokens in the place P_{sense2} (start of backoff procedure) depending on *CW*, where the weight of arc between T_{idle1} and P_{sense2} equals *RNS*. *RNS* is a random number which is uniformly distributed in the range [0, *CW*], where *CW* is computed as:

$$CW = (CWmin + 1) \cdot 2^{\#P_{FC}-1} - 1$$

The number of tokens in P_{sense2} represents the number of time slots that the node has to wait before transmitting the data frame. During any slot time, the channel may be busy, which is modelled by the transition T_{busy2}, or idle, which is modelled by the transition T_{idle2}. If the channel became busy, the backoff timer is frozen for a time equals to the time of transmitting the data frame and receiving the ACK frame. This is modelled by the transition T_{busy2}, place P_{busy2}, and transition T_{NAV2}. The end of frozen time is represented by the firing of T_{NAV2} which deposits a token in the place P_{DIFS2}. Sensing the channel for a DIFS interval before the backoff timer resumes decreasing is modelled by P_{DIFS2} and T_{DIFS2}. The probability that the channel is idle at the end of the current slot time is represented by the firing of T_{idle2} which moves a token from P_{sense2} to P_{slot}. The firing of the transition T_{slot} moves a token from P_{slot} to P_{BO} which represents the decrement of the backoff timer by one slot.

Table 3.1: The average firing time of timed transitions of SRN models shown in Figures 3.5 and 3.7

Transition	Average firing time
T_{PG}	λ^{-1}
T_{DIFS1}, T_{DIFS2}	$DIFS$
T_{NAV1}, T_{NAV2}	$Ft(T_{txD}) + Ft(T_{ACK})$
T_{slot}	T_s
T_{BO}	$A_s \cdot T_s$
T_{txD}	$\frac{PhH}{B1} + \frac{MPDU}{B2} + T_p$
T_{ACK}	$\frac{PhH}{B1} + \frac{ACK}{B2} + SIFS + T_p + T_{CCA} + T_{RxTx}$
$T_{timeout}$	$Ft(T_{txD}) + Ft(T_{ACK})$

Table 3.2: Transitions guard functions for SRN models shown in Figures 3.5 and 3.7

Transition	Guard Function
T_{busy2}	$\#P_{busy2}+\#P_{DIFS2} = 0$
T_{idle2}	$\#P_{busy2}+\#P_{DIFS2} = 0$
T_{DIFS}	$\#P_{ch}+\#P_{succ}+\#P_{fail} = 0$
T_{BO}	$\#P_{ch}+\#P_{succ}+\#P_{fail} = 0$
T_{coll}	$\#P_{ch}+\#P_{succ}+\#P_{fail} = 0$ and $\#P_{FBO} > 1$
T_{Ncoll}	$\#P_{ch}+\#P_{succ}+\#P_{fail} = 0$ and $\#P_{FBO} > 0$
T_{succ}	$\#P_{ch}=1$
T_{fail}	$\#P_{ch} + \#P_{ch\text{-}h} > 1$
$T_{DIFS\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{succ\text{-}h}+\#P_{fail\text{-}h} = 0$
$T_{BO\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{succ\text{-}h}+\#P_{fail\text{-}h} = 0$
$T_{coll\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{succ\text{-}h}+\#P_{fail\text{-}h} = 0$ and $\#P_{FBO\text{-}h} > 1$
$T_{Ncoll\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{succ\text{-}h}+\#P_{fail\text{-}h} = 0$ and $\#P_{FBO\text{-}h} > 0$
$T_{succ\text{-}h}$	$\#P_{ch\text{-}h} = 1$
$T_{fail\text{-}h}$	$\#P_{ch\text{-}h} > 1$
$T_{txD\text{-}h}$	$\#P_{succ} = 0$

The average firing time of the timed transition T_{slot} equals to the slot time T_s. The firing probabilities of the transitions T_{busy2} and T_{idle2} are β_1 and $(1-\beta_1)$ respectively. The average firing times of the transitions T_{DIFS2} and T_{NAV2} are equal to that of transitions T_{DIFS1} and T_{NAV1} respectively. The guard function of the transition T_{idle2}, shown in Table 3.2, prevents its firing when there are any tokens in places P_{busy2} and P_{DIFS2} to prevent the decrement of the backoff timer when the channel is busy. The guard function of the transition T_{busy2} and the inhibitor arcs between the place P_{slot} and transitions T_{idle2} and T_{busy2} ensure that the processing of the next slot will not start before the end of the current slot (i.e. moving the token from P_{slot} to P_{BO}).

Table 3.3: Arcs weight functions for SRN models shown in Figures 3.5 and 3.6

Arc name	**Arc weight function**
W_1	0 if $RNS > 0$ 1 if $RNS = 0$
W_2	$\#P_{FC}$
W_3	$\#P_{FC}$ if $\#P_{FC} = MRL$ 0 if $\#P_{FC} < MRL$
W_4	1 if $\#P_{FC} < MRL$ 0 if $\#P_{FC} = MRL$
W_5	0 if $\#P_{FC} < MRL$ 1 if $\#P_{FC} = MRL$

The firing of the transition T_{BO} represents the end of the backoff period. Because of the weight of arc between P_{BO} and T_{BO}, T_{BO} is enabled if the number of tokens in P_{BO} is greater than or equal to *RNS* which means that the backoff timer reached zero. If the *RNS* is equal zero, the node has to transmit the MAC frame immediately without backeoff delay. This means that the transition T_{BO} must be enabled if $RNS = 0$. So, the place P_{ZRNS} is added, where the transition T_{idle1} deposits a token in it if $RNS = 0$. This is controlled by the arc weight function W_1 shown in Table 3.3.

The firing of T_{BO} deposits a token in P_{txD} which represents the start of transmission of the data frame by the physical layer. The end of transmission of the data frame is

represented by the firing of T_{txD} which moves the token to P_{rxD} which models the delivery of the frame to the receiver.

If any other node starts to transmit any data frame at the same slot time, a collision occurs and the transmission fails; otherwise, the frame is transmitted successfully. Therefore, the token in P_{rxD} may move to P_{succ} due to the firing of T_{succ}, representing the success of transmitting the data frame, or it may move to P_{fail} due to the firing of T_{fail}, representing the failure to transmit the frame because of collision. The average firing time of the transition T_{txD} is the transmission time of MPDU, the transmission time of the physical header (*PhH*), and the propagation time (T_p), as shown in Table 3.1. The probabilities of firing of T_{fail} and T_{succ} are μ_1 and $(1-\mu_1)$ respectively, where μ_1 is the probability of failure to transmit the data frame due to interference induced by neighbour or hidden nodes. The parameters $\beta_1$and μ_1 are computed from the abstract model as explained in Section 3.4.3.

Once the receiver has successfully received the data frame (the token in P_{succ}), it sends the ACK frame after a SIFS interval which is represented by firing T_{ACK}. The transition T_{ACK} flushes the place P_{FC}, which models resetting the backoff counter to zero, and deposits a token in P_B which lets a new packet to be transmitted. The firing of the transition $T_{timeout}$ models the ACK frame timeout. Depending on the number of tokens in P_{FC}, the transition $T_{timeout}$ may deposit a token in P_B or P_M. If $\#P_{FC}$ is less than the maximum retry limit, $T_{timeout}$ deposits a token in P_M and does not remove any tokens from P_{FC}. Otherwise it deposits a token to P_B and flushes P_{FC} which models dropping the packet after reaching the maximum retry limit. This is controlled by the arc weight functions w_2, w_4, and w_5 shown in Table 3.3.

As shown in Table 3.1, the average firing time of the transition T_{ACK} is the transmission time of the ACK frame, the transmission time of the physical header, the propagation time, the time required to recognise the signal (T_{CCA}), the time required to convert from receiving to transmitting state (T_{RxTx}), and the SIFS interval. The average firing time of the transition $T_{timeout}$ reflects the time required to send a data

frame and to receive an ACK frame, so it is equal to the average firing times of the transitions T_{txD} and T_{ACK}.

3.4.2 One Node Detailed Model for the RTS/CTS Method

Figure 3.6 shows the SRN model of the one node detailed model for the RTS/CTS method. Compared with that for the BA method shown in Figure 3.5, the one node detailed model for RTS/CTS method has a few differences. The token in P_{busy1} represents that the channel is busy and the node has to wait till the end of the ongoing transmission from any other node. If the node sensed the channel busy, it sets the NAV and wait till the end of transmitting the ACK frame (modelled by T_{s1}, P_{s1}, and T_{NAV1}). Otherwise, the channel will be sensed free for a period greater than (2·SIFS + t_{CTS} + 2·T_s) that lets the node to reset the NAV to zero and start again to sense the channel for the DIFS interval (modelled by T_{f1}, P_{f1}, and T_{RNAV1}).

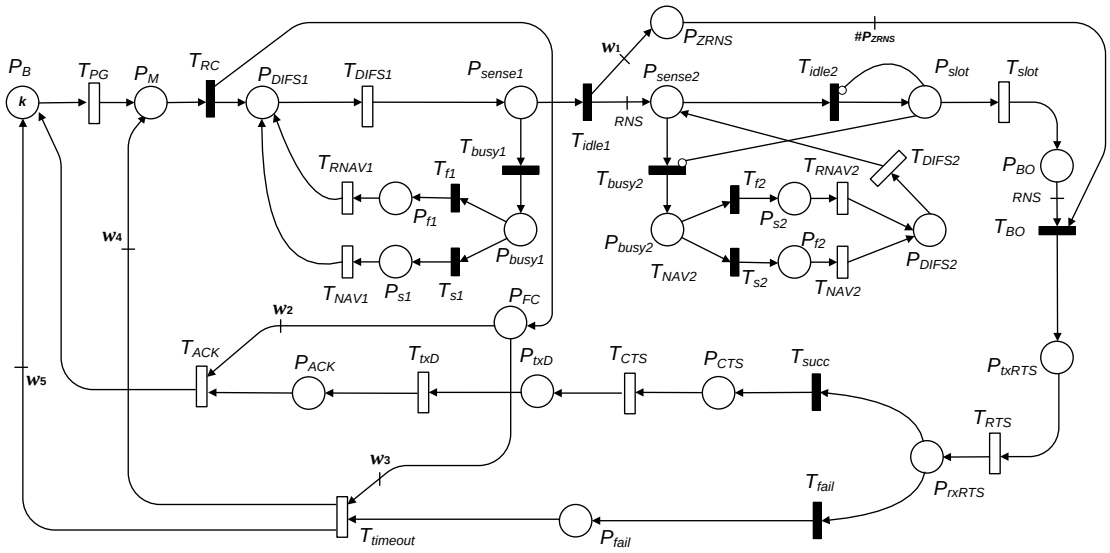

Figure 3.6: One node detailed model for the RTS/CTS method

The firing probabilities of the transitions T_{busy1} (β_2) and T_{idle1} $(1-\beta_2)$ are the probabilities that the channel is busy and idle respectively. The firing probabilities of the conflicted transitions T_{f1} (μ_2) and T_{s1} $(1-\mu_2)$ are the probabilities of failure and success to complete the RTS/CTS handshake respectively. The parameter β_2 and μ_2

are computed from the abstract model, as explained in Section 3.4.3. As shown in Table 3.4, the average firing time of the timed transition T_{NAV1} reflects the time needed to complete the RTS/CTS handshake. The average firing time of T_{RNAV1} is $2{\cdot}\text{SIFS} + Ft(T_{CTS}) + 2{\cdot}T_s$.

The token in P_{busy2} represents that the channel is busy and the backoff timer stopped till the end of the ongoing transmission. The function of transitions and places T_{s2}, T_{f2}, P_{s2}, P_{f2}, T_{NAV2}, and T_{RNAV2} are the same as T_{s1}, T_{f1}, P_{s1}, P_{f1}, T_{NAV1}, and T_{RNAV1}, respectively, but with the backoff procedure. Also the average firing time and probability of the corresponding transitions are the same, as shown in Table 3.4. P_{txRTS}, T_{RTS}, and P_{rxRTS} model transmitting and receiving the RTS frame by the sender and receiver. The average firing time of T_{RTS} equals the transmission and propagation time of the RTS frame. If the receiver received the RTS frame without any errors, T_{succ} fires depositing a token in P_{CTS}; otherwise, T_{fail} fires. The firing probabilities of T_{fail} and T_{succ} are μ_2 and $(1-\mu_2)$ respectively.

P_{CTS} and T_{CTS} represent transmission of the CTS frame from the receiver to the sender. Receiving the CTS frame and transmitting the data frame are represented by P_{txD} and T_{txD}. The firing of the transition T_{txD} moves the token from P_{txD} to P_{ACK}. The receiver sends the ACK frame after receiving the data frame; this is represented by P_{ACK} and T_{ACK}. The firing of the transition T_{ACK}, which moves the token from P_{ACK} to P_B, represents the successful transmission and reception of the ACK frame. The firing of the transition $T_{timeout}$ models the CTS and ACK frame timeout. The average firing times of the transitions T_{CTS}, T_{txD}, T_{ACK}, and $T_{timeout}$ are shown in Table 3.4.

3.4.3 Abstract Model for the BA and RTS/CTS Methods

Figure 3.7 shows the abstract model for the BA method. It consists of two parts with a similar structure; the active and hidden parts. The active and hidden parts represent the abstracted model for the nodes in the active and hidden areas respectively. The arcs between the active and hidden parts illustrate the interaction between nodes in the active and hidden areas. To derive an abstract model for the nodes in either the

active or hidden area, the backoff procedure and retry count in the one node detailed model are folded. Then, to exploit the identical behaviour of all nodes, the models of all nodes in the same area are combined together using the lumping technique. The meaning of the places and transitions are explained below.

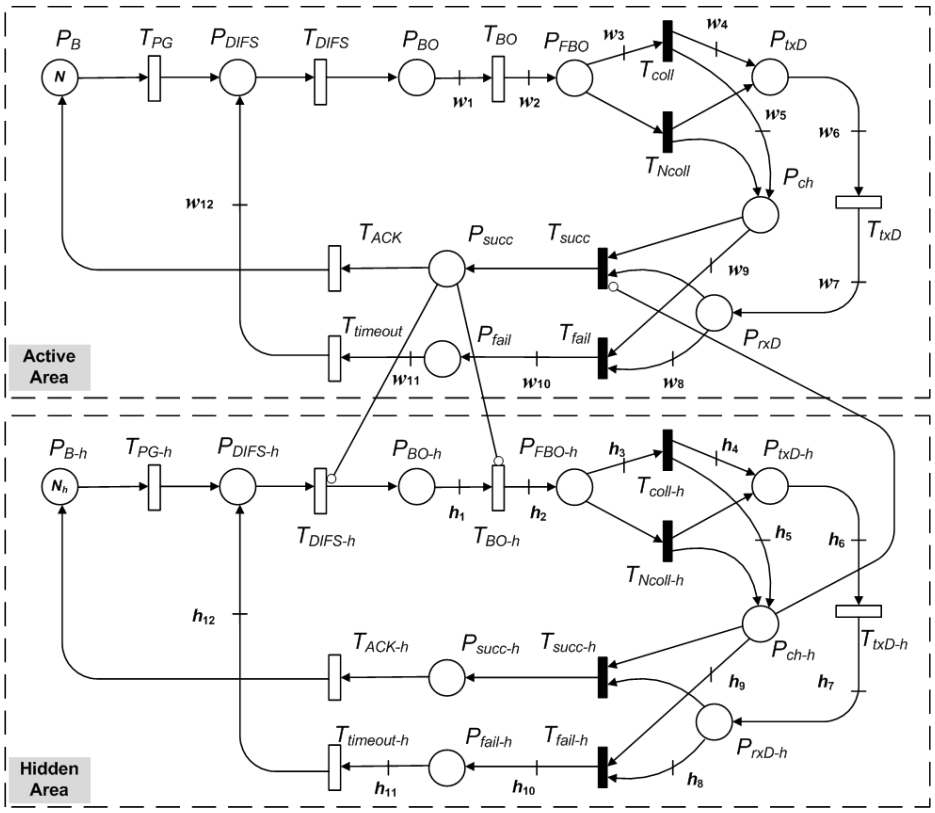

Figure 3.7: The abstract model for the BA method

The number of nodes that do not have a packet to transmit is represented by the number of tokens in P_B. The transition T_{PG} models the generation of packets from the upper layer. The place P_{DIFS1} represents that the node is sensing the channel for a DIFS period. If the channel is free for the DIFS interval, the transition T_{DIFS} fires moving a token from P_{DIFS} to P_{BO}. The state of the channel is represented by the place P_{ch}. If the number of tokens in P_{ch} is zero, the channel is idle; otherwise, the channel is busy. As shown in Table 3.2, the transition T_{DIFS} is assigned a guard that disables it if the channel is busy ($\#P_{ch} > 0$).

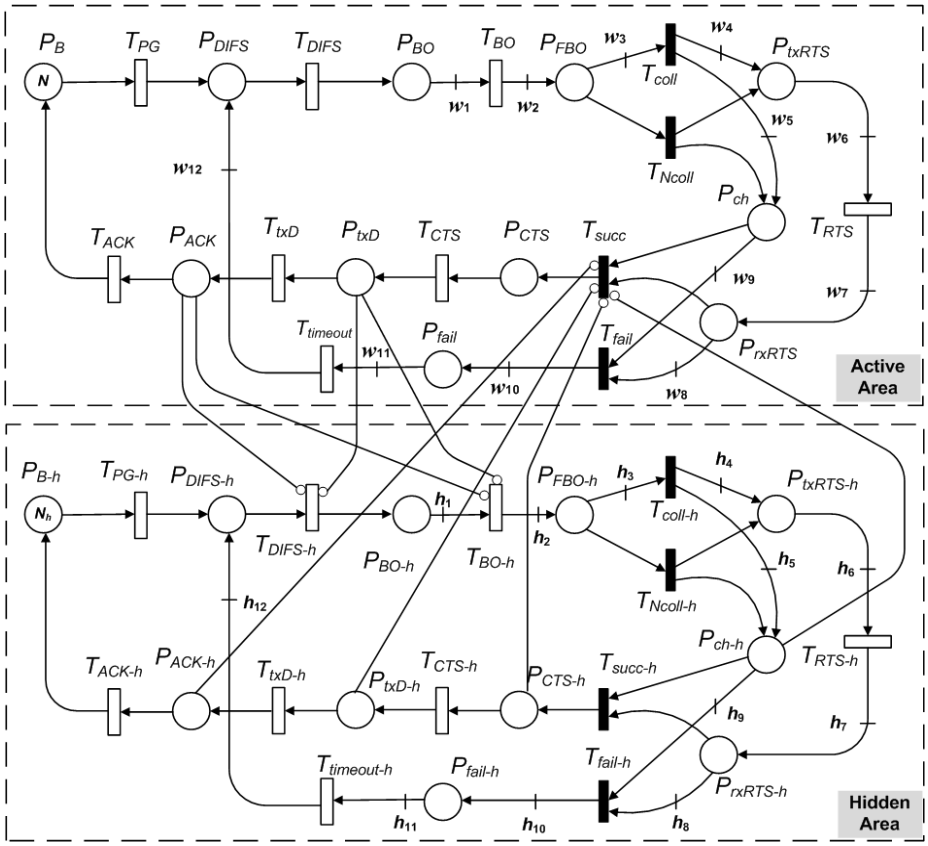

Figure 3.8: The abstract model for RTS/CTS method

The number of tokens in P_{BO} represents the number of nodes in the backoff state. The firing of the transition T_{BO} represents the end of the backoff procedure for all nodes that entered the backoff state (moving all tokens from P_{BO} to P_{FBO}). A guard is assigned to the transition T_{BO} to disable it when the channel is busy. The average firing time of T_{BO} is $(A_s \cdot T_s)$, where A_s is the average number of backoff slots of nodes in the active area. A_s is computed using the one node detailed model, shown in Figure 3.5, where A_s equals to the average number of tokens in the place P_{slot}. To compute A_s, the one node detailed model is solved using parameters β_1 and μ_1 derived from the abstract model as follows:

$$\beta_1 = Pr(\#P_{ch} > 0 \text{ OR } \#P_{ACK} > 0)$$

$$\mu_1 = Pr(\#P_{ch} > 1 \text{ OR } \#P_{ch\text{-}h} \geq 1)$$

where $Pr(E)$ and $\#P_x$ are the probability of an event E and the number of tokens in the place P_x, respectively.

Table 3.4: The average firing time of timed transitions of SRN models shown in Figures 3.6 and 3.8

Transition	Average firing time
T_{PG}	λ^{-1}
T_{DIFS1}, T_{DIFS2}	$DIFS$
T_{NAV1}, T_{NAV2}	$Ft(T_{RTS}) + Ft(T_{CTS}) + Ft(T_{txD}) + Ft(T_{ACK})$
T_{RNAV1}, T_{RNAV2}	$Ft(T_{CTS}) + 2\,SIFS + 2\,T_s$
T_{slot}	T_s
T_{BO}	$A_s \cdot T_s$
T_{RTS}	$\frac{PhH}{B1} + \frac{RTS}{B2} + T_p$
T_{CTS}	$\frac{PhH}{B1} + \frac{CTS}{B2} + SIFS + T_p + T_{CCA} + T_{RxTx}$
T_{txD}	$\frac{PhH}{B1} + \frac{MPDU}{B2} + SIFS + T_p + T_{CCA} + T_{RxTx}$
T_{ACK}	$\frac{PhH}{B1} + \frac{ACK}{B2} + SIFS + T_p + T_{CCA} + T_{RxTx}$
$T_{timeout}$	$Ft(T_{RTS}) + Ft(T_{CTS})$

The tokens in P_{FBO} enable the conflicted transitions T_{coll} and T_{Ncoll}. The transition T_{coll} represents the probability that the backoff timer of two or more nodes reached zero at the same time making packets collide, whereas the probability of no collision is represented by T_{Ncoll}. If $\#P_{ch} > 0$ (the channel is busy), the guards of T_{coll} and T_{Ncoll} disable them. The collision probability increases with increasing $\#P_{BO}$ and decreasing A_s. So, the firing probabilities of T_{coll} and T_{Ncoll} depend on $\#P_{FBO}$ and A_s, as shown in Table 3.5 [54, 91, 92].

The firing of T_{coll} moves all tokens in P_{FBO} to P_{txD} and P_{ch}, while the firing of T_{Ncoll} moves one token from P_{FBO} to P_{txD} and P_{ch}. Places P_{txD} and P_{rxD} and the transition T_{txD} represent the transmitting and receiving the data frame. Depending on the number of tokens in P_{ch} either the immediate transition T_{succ} or T_{fail} is enabled. If the number of tokens in P_{ch} equals one (only one node uses the channel), the transition T_{succ} is enabled; otherwise, T_{fail} is enabled. The firing of the transition T_{succ} deposits a token in P_{succ} which represents the success of receiving the data frame.

Table 3.5: The firing probabilities of immediate transitions of SRN models shown in Figures 3.7 and 3.8

Transition	Firing Probability
T_{coll}	$1-\left(1-\frac{1}{A_s}\right)^{\#P_{FBO}}$
T_{Ncoll}	$\left(1-\frac{1}{A_s}\right)^{\#P_{FBO}}$
$T_{coll\text{-}h}$	$1-\left(1-\frac{1}{A_s}\right)^{\#P_{FBO-h}}$
$T_{Ncoll\text{-}h}$	$\left(1-\frac{1}{A_s}\right)^{\#P_{FBO-h}}$

Table 3.6: Arcs weight functions for SRN model shown in Figure 3.7

Arc name	Arc weight function
W_1	$\#P_{BO}$ if $\#P_{ch} = 0$ and $\#P_{BO} > 0$ 1 if $\#P_{BO} = 0$
W_2	$\#P_{BO}$
W_3	$\#P_{BO}$ if $\#P_{FBO} > 1$ 1 if $\#P_{FBO} = 0$
W_4	$\#P_{FBO}$
W_5	$\#P_{FBO}$
W_6	$\#P_{txD}$ if $\#P_{txD} > 0$ 1 if $\#P_{txD} = 0$
W_7	$\#P_{txD}$
W_8	$\#P_{rxD}$
W_9	$\#P_{ch}$ if $\#P_{ch} > 0$ 0 if $\#P_{ch} = 0$
W_{10}	$\#P_{rxD}$
W_{11}	$\#P_{fail}$ if $\#P_{fail} > 0$ 1 if $\#P_{fail} = 0$
W_{12}	$\#P_{fail}$

Transmitting the ACK frame is represented by T_{ACK}. Tokens in P_{fail} represent failure to receive the data frame. The ACK frame timeout is modelled by the transition $T_{timeout}$. To model the synchronisation between collided packets, the same number of tokens moves from P_{FBO} to P_{DIFS} through T_{coll}, P_{txD}, T_{txD}, P_{rxD}, T_{fail}, P_{fail}, and $T_{timeout}$. This is controlled by the arc weight functions w_4, w_5, w_6, w_7, w_8, w_{10}, shown in Table 3.6.

Table 3.7: Arcs weight functions for SRN model shown in Figure 3.8

Arc name	**Arc weight function**
W_1	$\#P_{BO}$ if $\#P_{ch} = 0$ and $\#P_{BO} > 0$ 1 if $\#P_{BO} = 0$
W_2	$\#P_{BO}$
W_3	$\#P_{BO}$ if $\#P_{FBO} > 1$ 1 if $\#P_{FBO} = 0$
W_4	$\#P_{FBO}$
W_5	$\#P_{FBO}$
W_6	$\#P_{txRTS}$ if $\#P_{txRTS} > 0$ 1 if $\#P_{txRTS} = 0$
W_7	$\#P_{txRTS}$
W_8	$\#P_{rxRTS}$
W_9	$\#P_{ch}$ if $\#P_{ch} > 0$ 0 if $\#P_{ch} = 0$
W_{10}	$\#P_{rxRTS}$
W_{11}	$\#P_{fail}$ if $\#P_{fail} > 0$ 1 if $\#P_{fail} = 0$
W_{12}	$\#P_{fail}$

As shown in Figure 3.7, the structure of the abstracted SRN model for nodes in the active area is similar to that of nodes in the hidden area. The place $P_{x\text{-}h}$, the transition $T_{x\text{-}h}$, and the arc weight function h_x correspond to P_x, T_x, and w_x, respectively, where x is the name of the identifier (place, transition, or arc weight function). The meaning and function of all corresponding identifiers are the same. The only difference between the two models is the rate of the transition $T_{BO\text{-}h}$, which is equal to $\left(A_{s\text{-}h} \cdot T_s\right)$, where $A_{s\text{-}h}$ is the average number of backoff slots of nods in the hidden area. The

parameter $A_{s\text{-}h}$ is computed in the same way as A_s using the one node detailed model, where its parameters β_1 and μ_1 are computed from the abstract model as follows:

$$\beta_1 = Pr(\#P_{ch\text{-}h} > 0 \text{ OR } \#P_{ACK\text{-}h} > 0 \text{ OR } \#P_{ACK} > 0)$$

$$\mu_1 = Pr(\#P_{ch\text{-}h} > 1)$$

If a node *S* in the active area is transmitting a data frame to a node *D* that overlaps with a transmission of another data frame in the hidden area, the collision occurs at the destination *D*. So, the inhibitor arc between $P_{ch\text{-}h}$ and T_{succ} is added to disable T_{succ} and enable T_{fail} when the number of tokens in $P_{ch\text{-}h}$ is greater than zero. The receiver *D* sends the ACK frame if it received the data frame successfully. During sending the ACK frame the hidden nodes sense the channel busy which make them stop sensing the channel for the DIFS interval and stop the backoff counter. This is represented by the inhibitor arcs from P_{succ} to transitions $T_{DIFS\text{-}h}$ and $T_{BO\text{-}h}$.

The abstract model for the RTS/CTS method is shown in Figure 3.8. Compared with the corresponding SRN model for the BA method shown in Figure 3.7, there are a few differences. As explained for the one node detailed model shown in Figure 3.6, places P_{txRTS} and P_{rxRTS} and the transition T_{txRTS} represent the transmission and reception of an RTS frame. Receiving the RTS frame, transmitting the CTS frame and receiving the CTS frame are modelled by P_{CTS}, T_{CTS}, and P_{txD} respectively. Once the source has received the CTS frame, it transmits the data frame. When the receiver receives the data frame successfully, it sends the ACK frame. This is modelled by the places P_{txD} and P_{ACK}, and transitions T_{txD} and T_{ACK}.

The average firing time of T_{BO} is $(A_s \cdot T_s)$, where A_s is the average number of backoff slots of nodes in the active area. A_s is computed using the one node detailed model, shown in Figure 3.6, where A_s is equal to the average number of tokens in the place P_{slot}. To compute A_s, the one node detailed model is solved using parameters β_2 and μ_2 derived from the abstract model shown Figure 3.8 as follows:

$$\beta_2 = Pr(\#P_{ch} > 0 \; OR \; \#P_{CTS} > 0 \; OR \; \#P_{txD} > 0 \; OR \; \#P_{ACK} > 0)$$

$$\mu_2 = Pr(\#P_{ch} > 1 \ \ OR \ \ \#P_{ch\text{-}h} > 0 \ \ OR \ \ \#P_{CTS\text{-}h} > 0$$

$$\text{OR} \ \ \#P_{TxD\text{-}h} > 0 \ \ OR \ \ \#P_{ACK\text{-}h} > 0)$$

As shown in Table 3.4, the average firing times of the transitions T_{RST}, T_{CTS}, T_{txD}, and ACK are the transmission, sensing and interframe spacing time of RTS, CTS, data, and ACK frames respectively. The arcs weight functions and transition guard functions are shown in Table 3.7 and 3.8 respectively.

Table 3.8: Transition guard functions for SRN models shown in Figures 3.6 and 3.8

Transition	**Guard Function**
T_{busy2}	$\#P_{busy2}+\#P_{s2}+\#P_{f2}+\#P_{DIFS2} = 0$
T_{idle2}	$\#P_{busy2}+\#P_{s2}+\#P_{f2}+\#P_{DIFS2} = 0$
T_{DIFS}	$\#P_{ch}+\#P_{CTS}+\#P_{txD}+\#P_{ACK}+\#P_{fail} = 0$
T_{BO}	$\#P_{ch}+\#P_{CTS}+\#P_{txD}+\#P_{ACK}+\#P_{fail} = 0$
T_{coll}	$\#P_{ch}+\#P_{CTS}+\#P_{txD}+\#P_{ACK}+\#P_{fail} = 0$ and $\#P_{FBO} > 1$
T_{Ncoll}	$\#P_{ch}+\#P_{CTS}+\#P_{txD}+\#P_{ACK}+\#P_{fail} = 0$ and $\#P_{FBO} > 0$
T_{succ}	$\#P_{ch} = 1$
T_{fail}	$\#P_{ch}+\#P_{ch\text{-}h}+\#P_{CTS\text{-}h}+\#P_{txD\text{-}h}+\#P_{ACK\text{-}h} > 1$
$T_{DIFS\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{CTS\text{-}h}+\#P_{txD\text{-}h}+\#P_{ACK\text{-}h}+\#P_{fail\text{-}h}+\#P_{CTS} = 0$
$T_{BO\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{CTS\text{-}h}+\#P_{txD\text{-}h}+\#P_{ACK\text{-}h}+\#P_{fail\text{-}h}+\#P_{CTS} = 0$
$T_{coll\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{CTS\text{-}h}+\#P_{txD\text{-}h}+\#P_{ACK\text{-}h}+\#P_{fail\text{-}h} = 0$ and $\#P_{FBO\text{-}h} > 1$
$T_{Ncoll\text{-}h}$	$\#P_{ch\text{-}h}+\#P_{CTS\text{-}h}+\#P_{txD\text{-}h}+\#P_{ACK\text{-}h}+\#P_{fail\text{-}h} = 0$ and $\#P_{FBO\text{-}h} > 0$
$T_{succ\text{-}h}$	$\#P_{ch\text{-}h} = 1$
$T_{fail\text{-}h}$	$\#P_{ch\text{-}h} > 1$
$T_{RTS\text{-}h}$	$\#P_{CTS}+\#P_{txD}+\#P_{ACK} = 0$

As shown in Figure 3.8, the structure of the abstracted SRN model for the nodes in the hidden area is similar to that of the nodes in the active area. The place $P_{x\text{-}h}$, the transition $T_{x\text{-}h}$, and the arc weight function h_x correspond to P_x, T_x, and w_x, respectively, where x is the name of the identifier. The meaning and function of all corresponding identifiers are the same. The average firing time of the transition $T_{BO\text{-}h}$ is $\left(A_{s\text{-}h} \cdot T_s\right)$, where $A_{s\text{-}h}$ is the average number of tokens in the place P_{slot} in the one

node detailed model which its parameters β_2 and μ_2 are computed from the abstract model as follows:

$$\beta_2 = Pr\ (\#P_{ch\_h} > 0\ OR\ \#P_{CTS\_h} > 0\ OR\ \#P_{txD\_h} > 0\ OR\ \#P_{ACK\_h} > 0$$
$$OR\ \#P_{CTS} > 0\ OR\ \#P_{txD} > 0\ OR\ \#P_{ACK} > 0)$$

$$\mu_2 = Pr(\#P_{ch\_h} > 1)$$

For the RTS/CTS method, there are more interactions between the nodes in the active area and hidden nodes than the BA method. In Figure 3.4, if the hidden node S_{h1} sent an RTS frame to the destination D_{h1}, the destination D_1 of the source node S_1 will receive it. Consequently, D_1 sets its NAV to a value that prevents it from sending any CTS or ACK frames until S_{h1} receives the ACK frame from D_{h1}. Therefore, if D_1 received a RTS frame from S_1, it will not send any response which produces a timeout error for the RTS frame. This is modelled by adding inhibitor arcs from places $P_{CTS\text{-}h}$, $P_{txD\text{-}h}$, and $P_{ACK\text{-}h}$ to the transition T_{succ}, as shown in Figure 3.8. In addition, if the nodes S_{h1} and S_1 send a RTS frame at the same time, a collision occurs at the destination D_1 which also produces a timeout error. So, an inhibitor arc is added between the place $P_{ch\text{-}h}$ and transition T_{succ} which disables it and enables T_{fail}. When any destination in the active area (e.g. D_1) sends a CTS frame to the source (e.g. S_1), the hidden nodes will receive it, thus they stop all activities until the destination receives the data frame and sends the ACK frame. This situation is modelled by adding the inhibitor arcs between transitions $T_{DIFS\text{-}h}$ and $T_{BO\text{-}h}$ and places P_{txD} and P_{ACK}, as depicted in Figure 3.8, and assigning a guard function for each of these transitions.

3.5 Analytical Procedure

As explained in Section 3.4, the proposed models depend on the decomposition technique. Thus, to compute the performance metrics, the two models for either the BA or RTS/CTS method are solved iteratively. For the BA or RTS/CTS method, the one node detailed model is used to derive the average size of the backoff window of

nodes in the active area (A_s) and nodes in the hidden area ($A_{s.h}$), which equals to the average number of tokens in the place P_{slot}. Using A_s and $A_{s.h}$, the abstract model is solved to derive the performance metric and the parameters β_1, μ_1, β_2, and μ_2, which are used to solve the one node detailed model. According to the following procedure, the two models for either the BA or RTS/CTS method are solved iteratively until the convergence of the performance metrics:

Step 1: The number of iterations *n* is initialised to 1 and the initial value of the average size of the backoff window is computed using the following equation:

$$A_s = A_{s\text{-}h} = \frac{\sum_{x=0}^{MRL-1}(CW_{min} + 1) \cdot 2^x - 1}{MRL}$$

Step 2: The abstract model is solved using the initial value of the backoff window to get the initial values of a performance metric τ^0 (e.g. throughput) and parameters β_1 and μ_1 (or β_2 and μ_2 in the case of RTS/CTS method).

Step 3: The one node detailed model is solved using the last computed values of parameters β_1 and μ_1 (or β_2 and μ_2) to compute the new values for A_s and $A_{s\text{-}h}$

Step 4: The abstract model is solved to get the performance metric τ^n and parameters β_1 and μ_1 (or β_2 and μ_2).

Step 5: The error of the performance metric is computed using the following equation

$$err(\tau) = \frac{\tau^n - \tau^{n-1}}{\tau^n}$$

Step 6: If $err(\tau)$ is less than a specified threshold, stop the iteration process; otherwise, increase *n* by one and go to Step 3.

The number of iterations depends on the error threshold. In all validation scenarios introduced in the next section, the error threshold is set to 0.01. In all cases the convergence of the performance metric is achieved in only a few iterations.

Table 3.9: Parameters of the MAC and Physical layers

Parameter	Value
CW_{min}	31
CW_{max}	1023
T_s	20 μs
T_{RxTx}	5 μs
T_{CCA}	15 μs
DIFS	50 μs
SIFS	10 μs
PhH	192 bit
MAC Header	292 bit
RTS	160 bit + *PhH*
CTS	112 bit + *PhH*
ACK	112 bit + *PhH*
B_1	1 Mbps
B_2	2 Mbps
SRC	6
LRC	4

3.6 Model Validation

In this section, the proposed SRN models for both BA and RTS/CTS methods are validated by conducting extensive comparisons of their results with those of simulation experiments. The simulation and analytical results were obtained by using the NS2 simulator [27] and SPNP [93], respectively. Table 3.9 shows the parameters of the physical and MAC layers used in the simulation and analysis. The capacity of the wireless channel is set to 2 Mbps. All simulation results are obtained with 95% confidence interval and a maximum relative error of 1%. Simulation time is set to 1000s. The first 100s are discarded in order to be sure that the network has reached steady state.

The performance metrics obtained from both analytical models and simulations are the packet delay (δ) and goodput. The packet delay is the time needed to transmit a

packet which is the time from the packet generation until the ACK frame is correctly received. Goodput is the number of data bits, not including protocol overhead and retransmitted bits, which are received correctly per unit time. Thus, goodput represents the application level throughput. Goodput and packet delay can be calculated from the abstract model using the following equations:

$$\text{Goodput} = Thr(T_{PG}) \tag{3.1}$$

$$\delta = \frac{N - M(P_B)}{Thr(T_{PG})} \tag{3.2}$$

where $Thr(T_x)$ is the throughput of transition T_x and $M(P_x)$ is the expected number of tokens in the place P_x. For simulation experiments, the packet delay is obtained by averaging the delay of all packets produced during the simulation time.

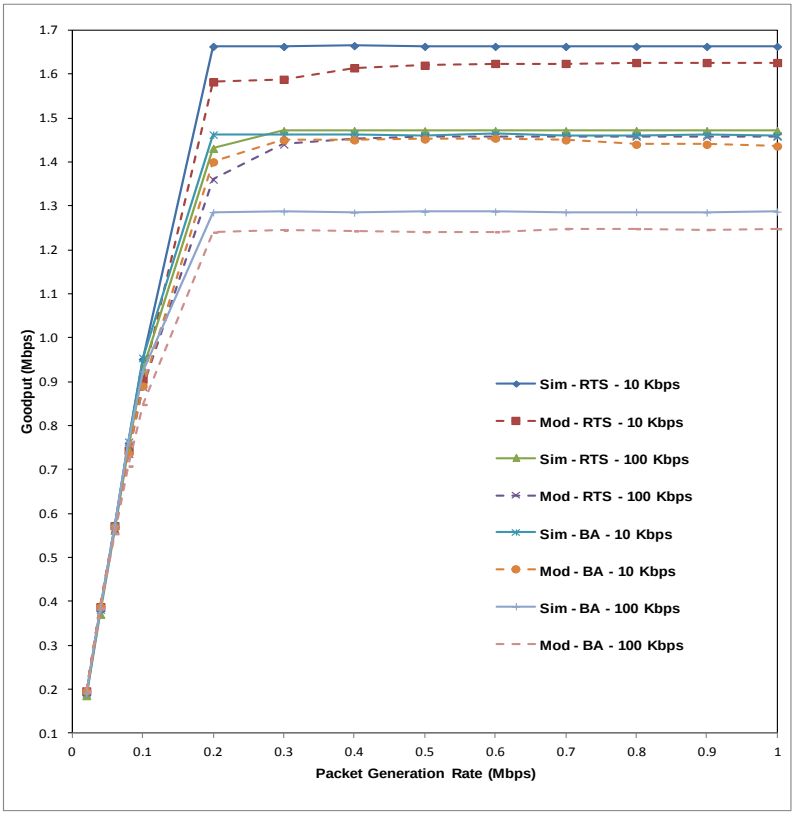

Figure 3.9: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of N = 10, N_h = 2, Packet Size = 2 KB, λ_h = 10 and 100 Kbps

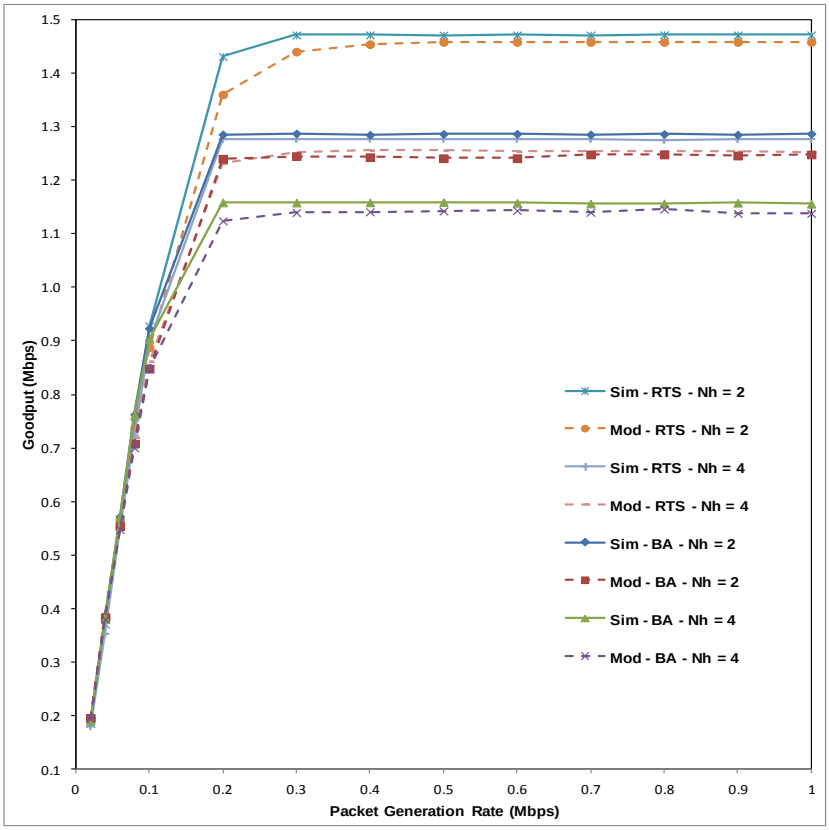

Figure 3.10: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of N = 10, N_h = 2 and 4, Packet Size = 2 KB, λ_h = 100 Kbps

To validate the proposed SRN models for the BA and RTS/CTS methods, several simulation scenarios have been considered. In Figures 3.9–3.12, solid lines refer to simulation results (labelled Sim), while dashed lines represent SRN model results (labelled Mod). First consider a scenario with 10 nodes in the active area where the packet generation rate at each node changes from 0.01 to 1 Mbps, two hidden nodes, and the packet size equals 2 KB. To investigate the effect of the packet generation rate of hidden nodes on goodput of nodes in the active area, we set it to either 10 or 100 Kbps. In Figure 3.9, goodput is plotted against increasing values of packet generation rate for the BA and RTS/CTS methods.

From Figure 3.9, it can be seen that the RTS/CTS method outperforms the BA method especially in heavy load conditions. In addition, in light load conditions increasing the packet generation rate of hidden nodes does not affect the performance of either the BA or RTS/CTS method, whereas in heavy load conditions it has a

considerable effect. Increasing the packet generation rate of hidden nodes from 10 to 100 Kbps decreases the saturated goodput by about 19% and 11% in the case of the BA and RTS/CTS methods respectively. This is because the collision probability increases rapidly when the packet generation rate of hidden nodes increases. As shown in Figure 3.9, we can notice the accuracy of the analytical results of the proposed model compared to simulation results in conditions of either light or heavy load.

To illustrate the influence of the number of hidden nodes on goodput of the nodes in the active area, in Figure 3.10 we plot goodput versus the packet generation rate at nodes in the active area for both BA and RTS/CTS methods, where N = 10, N_h = 2 or 4, λ_h = 100 Kbps, and the packet size is 2 KB. The figure shows that, in both the BA and RTS/CTS methods with high traffic load, goodput deteriorates when the number of hidden nodes increases. This is due to the increase of interference and the collision probability.

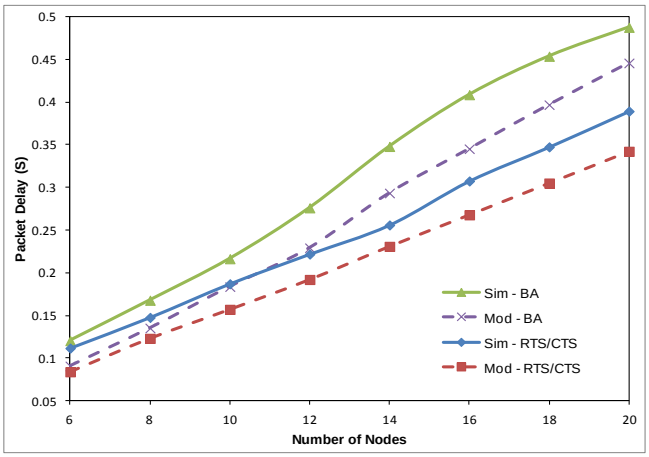

Figure 3.11: Packet delay versus number of nodes for both BA and RTS/CTS methods, in the case of λ = 2 Mbps, N_h = 2, Packet Size = 2 KB, λ_h = 10 Kbps

Figure 3.11 shows how the packet delay is affected by varying the number of nodes in active area (N) from 6 to 20 for the BA and RTS/CTS methods, where N_h = 2, λ_h = 10 Kbps, λ = 2 Mbps, and the packet size is 2 KB. It is clear that the performance of the RTS/CTS method is better than the BA method, especially with a large number of nodes.

In the last scenario, the effect of the packet size on the performance of the BA and RTS/CTS methods is investigated. We consider the case where the number of nodes in the active area is fixed to 10 nodes, the number of hidden nodes is set to $N_h = 2$, the packet generation rate at hidden nodes is set to 0.01 Mbps, and the packet size is set to 2 KB or 0.5 KB. The packet generation rate at nodes in the active area is varied from 0.01 to 1 Mbps. Figure 3.12 shows goodput of nodes in the active area versus the packet generation rate. The following can be observed from this figure:

- With light load conditions, the packet size has no significant effect on the performance of the network either in the case of the BA or RTS/CTS method.
- With heavy load conditions, the packet size has significant effect on the performance of the network, where increasing the packet size from 0.5 to 2 KB increased goodput with about 20% and 37% in the case of the BA and RTS/CTS methods respectively.
- The performance of the BA method is a little better than RTS/CTS method when the packet size is small.
- For large packet size, the performance of RTS/CTS method is much better than the BA method.
- In all cases, the results of the proposed models are accurate compared to simulation results.

In order to analytically solve the proposed model, some deterministic events such as the DIFS interval, backoff slot time, and packets transmission time, are approximated to be exponentially distributed. Therefore, all simulation results have an additional overhead over analytical results. As shown in Figures 3.9–3.12, it is clear that this overhead is very small.

The number of states in the abstract model and the one node detailed model depend on the number of nodes and *RNS*, respectively. The solution time (the time needed to generate the Markov chains model and compute the required performance metric) of these models depends on their state space size and specifications of the used machine.

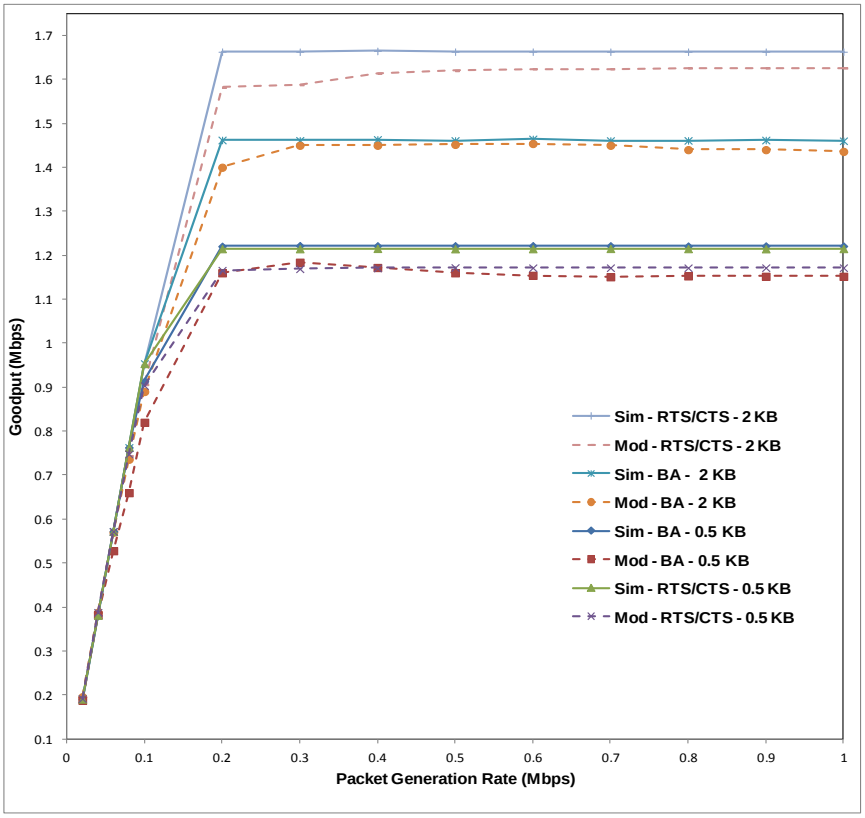

Figure 3.12: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of N = 10, N_h = 2, Packet Size = 2 and 0.5 KB, λ_h = 10 Kbps

For simulation and analytical analysis, a desktop workstation was used which was equipped with a 2.6 GHz (Intel Q9400 Core 2 Quad) processor, 4 GB of RAM and Ubuntu Linux version 8.10. The maximum number of states of the one node detailed model is 1173 and 5905 for the BA and RTS/CTS methods, respectively, and the solution time is about two seconds.

The number of states and solution time of the abstract model depend on the number of nodes in the active and hidden area. In the case of N = 20 and N_h =2, for the BA and RTS/CTS methods, the number of states of the abstract model is 16534 and 32323, and solution time is about 2 and 4 seconds respectively. In all scenarios, the analytical analysis is much faster than simulation. The time need for analytical analysis is in order of seconds, whereas the simulation time is in order of hours.

3.7 Summary

In this chapter we have investigated the performance of the IEEE 802.11 DCF MAC protocol, for both BA and RTS/CTS methods, in single hop ad hoc networks in the presence of hidden nodes using SRN models. The proposed models capture most features of this MAC protocol. The influences of network parameters, such as the traffic load, packet size, and number of nodes, have been demonstrated.

The proposed SRN models for both BA and RTS/CTS methods have been validated through extensive comparisons between analytical and simulation results. Comparisons showed that the proposed models succeeded in providing an accurate representation of the dynamic behaviour of the IEEE 802.11 DCF MAC protocol under several different settings of network parameters.

Chapter 4

Expected Path Length in Mobile Ad Hoc Networks with Random Waypoint Mobility

4.1 Introduction

In MANETs, a route or path is the sequence of mobile nodes which data packets pass through in order to reach the intended destination node from a given source node. Due to the mobility of nodes, mobile ad hoc networks have inherently dynamic topologies. Therefore, the routes are prone to frequent breaks which reduce the throughput of the network compared to wired or cellular networks. Consequently, the route followed by packets to reach the destination varies frequently. This is a crucial factor that affects the performance of the network.

The hop count specifies the number of hops on the path between source and destination nodes. The analysis of the hop count in multi-hop networks is very important because it can provide design guidelines for ad hoc networks. It can be used in many applications which are summarized as follows:

- Estimation of the packet delivery ratio
- With the packet delay per hop, the packet end-to-end delay can be estimated

- Given the number of simultaneous communications in the network, the network traffic load can be estimated
- Performance comparison between different multi-hop routing protocols
- Evaluation of the flooding cost and search latency for on-demand routing protocols and determination of the optimal flooding strategy [94]
- Studying the connectivity and estimation of the capacity of the network

In addition to the above, the hop count is a key parameter in the performance analysis of multi-hop ad hoc networks using analytical methods.

Random mobility models such as the random waypoint, random walk (random direction), free way, and Manhattan models, play an important role in the simulation of mobile ad hoc networks. However, no analytical study has yet investigated the expected hop count of paths in MANETs in a random mobility environment. This is the motivation for the work introduced in this chapter, in which a simple closed form analytical approach is developed to estimate the expected number of hops between any source-destination pair in MANETs where nodes are scattered in a square area and move according to the random waypoint mobility model. The RWPMM is selected because it is one of the most commonly used mobility models in MANET studies. The hop count of paths for other mobility models can also be investigated using the proposed approach.

For a given distance between the source and destination, in order to analytically compute the expected hop count, a packet forwarding algorithm is needed which uses an optimization criteria to choose a relay node from neighbour nodes that minimizes the number of hops a packet has to traverse in order to reach the destination. A new packet forwarding strategy is proposed called *Maximum Hop Distance* (MHD) that attempts to minimize the number of hops needed for a packet to reach its destination by forwarding the packet to a neighbour node with the maximum forward distance in the direction of the destination.

To calculate the average number of hops analytically using MHD without the need to run time-consuming simulations, the probability density function of the distance between the source (or a relay node) and its neighbour nodes is derived using geometric

probability. Then, this is used to compute the expected value for the maximum forward distance toward the destination, which is essential to compute the expected value for the remaining distance to the destination. By recursively computing the remaining distance to the destination, the expected hop counts can be computed.

The MHD approach is a greedy routing approach which is inspired by the LRD approach introduced in [66], but it is simpler and more accurate, as is clear from the comparison between the two approaches in Section 4.4. In addition, MHD can be used for networks with low node density. The proposed process, which uses the MHD approach to analytically compute the expected hop count between source and destination nodes moving according to the RWPMM, can be summarized as follows:

(1) With a given network size, the expected distance between any source-destination pair is computed

(2) The maximum expected distance (maximum forward distance) between any two nodes in the route for a given transmission range is derived

(3) With a given node density, the per-hop progress is calculated

(4) By recursive computation, the expected number of hops for each packet to traverse from a source to a destination is derived

The number of hops between the source and destination in multi-hop ad hoc networks is jointly affected by many network factors, such as node density, the transmission range of nodes, the pattern of mobility, and the size of the network area. The proposed approach is used to analyse the effect of these factors on the expected number of hops of paths in MANETs.

4.2 Euclidean Distance between a Source and Destination Node

This section derives an expression for the expected Euclidean distance between any random source and destination nodes moving according to RWPMM. First, it is derived for one dimension and then the square area is considered.

4.2.1 Expected Distance on One Dimension

We first consider the distance between two nodes in a line segment. Suppose that two random points X_1 and X_2 are located in a line segment with length L. The distance between X_1 and X_2 is S. X_1 and X_2 are independent, identically distributed (i.i.d.), random variables. According to [46], for the random waypoint mobility model the distribution of X_1 or X_2 is non-uniform at the long run. The probability distribution function of the location of a point x_n moving on a line with length L according to the RWPMM is [46]:

$$f_{X_n}(x_n) = \frac{6}{L^2}x_n + \frac{6}{L^3}x_n^2 \qquad 0 \le x_n \le L$$

Because X_1 and X_2 are i.i.d., the Probability Distribution Function (PDF) of the location of the two points is

$$f_{X_1X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

Where $0 \le x_1 \le L$ and $0 \le x_2 \le L$.

The Cumulative Distribution Function (CDF) of the distance $S = |x_2 - x_1|$ between the two points (i.e. the probability that S is smaller than a given value d) can be obtained by the integration of $f_{X_1X_2}(x_1, x_2)$ over the bounds of S as follows:

$$\begin{aligned} P(s \le d) &= \iint f_{X_1X_2}(x_1, x_2)\, dx_2 dx_1 = \iint f_{X_1}(x_1) \cdot f_{X_2}(x_2)\, dx_2 dx_1 \\ &= \int_0^d \int_0^{d+x_1} f_{X_1}(x_1) \cdot f_{X_2}(x_2)\, dx_2 dx_1 \\ &\quad + \int_d^{L-d} \int_{x_1-d}^{d+x_1} f_{X_1}(x_1) \cdot f_{X_2}(x_2)\, dx_2 dx_1 \\ &\quad + \int_{L-d}^{L} \int_{x_1-d}^{L} f_{X_1}(x_1) \cdot f_{X_2}(x_2)\, dx_2 dx_1 \end{aligned}$$

The integrations in the foregoing equation can be evaluated, yielding the following result:

$$P(s \leq d) = \frac{12d}{5L} - \frac{4d^3}{L^3} + \frac{3d^4}{L^4} - \frac{2}{5}\frac{d^6}{L^6} \tag{4.1}$$

By definition, the probability density function $f(d)$ of the distance d is given by the derivative of CDF shown in Equation (4.1), which yields the following:

$$f(d) = \frac{12}{5L} - \frac{12d^2}{L^3} + \frac{12d^3}{L^4} - \frac{12}{5}\frac{d^5}{L^6}$$

4.2.2 Expected Distance in Two Dimensions

Now, consider two random points X_1 and X_2 located in a square area of size $L \times L$ with coordinates (x_1, y_1) and (x_2, y_2), respectively. If ξ is the square of the distance d between X_1 and X_2, ξ is given by:

$$\xi = d^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

If $f_{\xi_x}$and f_{ξ_y} are the PDF of the events $(x_1 - x_2)^2$ and $(y_1 - y_2)^2$ respectively, then, the PDF of the distance ξ is given by the convolution of $f_{\xi_x}$and f_{ξ_y}as follows:

$$f_{xy}(\xi) = \int f_{\xi_x}(z) \cdot f_{\xi_y}(\xi - z)\, dz \tag{4.2}$$

Let $(x_1 - x_2)^2 = \xi_x$, then the CDF of ξ_x can be obtained by using Equation (4.1) by substituting ξ by $\sqrt{\xi_x}$. We then get the following:

$$F_{\xi_x}(\xi_x) = \frac{12}{5}\frac{\sqrt{\xi_x}}{L} - \frac{4\sqrt{\xi_x^3}}{L^3} + \frac{3\,\xi_x^2}{L^4} - \frac{2}{5}\frac{\xi_x^3}{L^6}$$

The PDF of ξ_x is obtained as follows:

$$f_{\xi_x}(\xi_x) = \frac{\partial}{\partial \xi_x} F_{\xi_x}(\xi_x) = \frac{6\,\xi_x}{L^4} - \frac{6}{5}\frac{\xi_x^2}{L^6} + \frac{6}{5\,L\sqrt{\xi_x}} - \frac{6\sqrt{\xi_x}}{L^3} \tag{4.3}$$

In the same way, $f_{\xi_y}(\xi_y)$ can be obtained. Because the domain of ξ is divided into two parts, $0 < \xi \leq L^2$ and $L^2 < \xi \leq 2L^2$, there are two cases for Equation (4.2) which are:

$$f_{xy}(\xi) = \begin{cases} I_1(\xi) = \int_0^{\xi} f_{\xi_x}(z) \cdot f_{\xi_y}(\xi - z)\, dz & 0 < \xi \le L^2 \\ I_2(\xi) = \int_{\xi - L^2}^{L^2} f_{\xi_x}(z) \cdot f_{\xi_y}(\xi - z)\, dz & L^2 < \xi \le 2L^2 \end{cases} \tag{4.4}$$

By substituting Equation (4.3) into Equation (4.4), the integrals I_1 and I_2 can be evaluated, and with some simplification and reduction of their terms we get the following:

$$I_1(\xi) = \frac{6\,\xi^3}{L^8} - \frac{6\,\xi^4}{5L^{10}} + \frac{6\,\xi^5}{125L^{12}} + \frac{96\sqrt{\xi^3}}{5L^5} - \frac{1584\sqrt{\xi^5}}{125L^7} + \frac{36\pi}{25L^2} + \frac{192\,\xi^2\sqrt{\xi^3}}{175L^9} - \frac{36\pi\xi}{5L^4} - \frac{48\xi\sqrt{\xi^3}}{5L^7} + \frac{9\pi\xi^2}{2L^6} \tag{4.5}$$

$$I_2(\xi) = \frac{312\xi}{25L^4} - \frac{1104}{875\,L^2} - \frac{12\,\xi^2}{5\,L^6} - \frac{6\,\xi^3}{L^8} + \frac{6\,\xi^4}{5\,L^{10}} - \frac{6\,\xi^5}{125\,L^{12}} + \frac{9}{25L^6}(8L^4 - 40L^2\xi + 25\,\xi^2)\arctan\left(\frac{L^2 - \frac{\xi}{2}}{L\sqrt{\xi - L^2}}\right) - \frac{6}{175L^9}(165L^4 - 232L^2\xi + 32\,\xi^2)\sqrt{(\xi - L^2)^3} - \frac{6}{875L^9}(407L^6 + 1936L^4\xi - 1768L^2\xi^2 + 160\,\xi^3)\sqrt{\xi - L^2} \tag{4.6}$$

Because ξ is the square of the distance between X_1 and X_2, the expected distance between the two nodes $E(d)$ is given by:

$$E(d) = \int_0^{2L^2} \sqrt{\xi}\, f_{xy}(\xi)\, d\xi = \int_0^{L^2} \sqrt{\xi}\; I_1(\xi)\, d\xi + \int_{L^2}^{2L^2} \sqrt{\xi}\; I_2(\xi)\, d\xi$$

$$= \int_0^{\sqrt{2}L} 2\,d^2\, f_{xy}(d)\, dd = \int_0^{L} 2\,d^2\; I_1(d)\, dd + \int_L^{\sqrt{2}L} 2\,d^2\; I_2(d)\, dd \tag{4.7}$$

where $d = \sqrt{\xi}$ and $d\xi = 2d\,dd$. The expected distance between the two nodes can be evaluated by plugging Equations (4.5) and (4.6) into Equation (4.7) which yields:

$$E(d) = \left(\frac{111}{350}\ln(\sqrt{2}+1) + \frac{28083}{750750}\sqrt{2} + \frac{19064}{375375}\right) \cdot L \tag{4.8}$$

For uniformly distributed nodes in a square area of size $L \times L$, the expected distance between two random nodes is [95]:

$$E(d) = 0.5214054\,L$$

Figure 4.1 shows the expected Euclidian distance between any random source and destination nodes ($E(d)$) that are uniformly scattered or moving according to the RWPMM in a square area, plotted against different values of the side length of the square area (L). It is clear that the expected distance between the two nodes in the case of the RWPMM is much less than that of uniform distributed nodes, especially for large values of L. This is because the spatial distribution of nodes moving according to the RWPMM at long run is non-uniform, since the probability that a node is located at the centre of the square area is high, and it reaches zero at the border of the area [96].

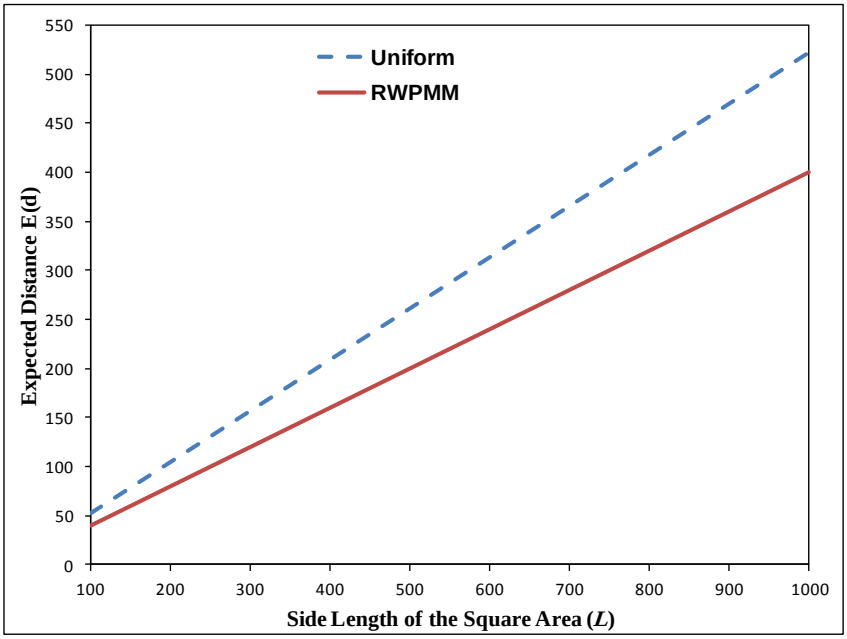

Figure 4.1: The expected Euclidian distance between any random source and destination nodes

4.3 Expected Hop Count

To analyse the expected hop count in MANETs where nodes move according to the RWPMM, we consider any source node *S* that tries to send its packet to a destination node *D*, as shown in Figure 4.2, where the circle with radius *R* around any node indicates the transmission area. The expected distance between any source and destination nodes is *d*. If *d* is greater than the transmission range *R*, which is equal for all nodes in the network, the source uses the intermediate nodes to forward the packets to the destination through two or more hops. The routing protocol searches all routes to the destination and chooses the shortest one. If the source has N_n neighbour nodes within its transmission range, the routing protocol in *S* will choose the closest neighbour node to the destination (e.g. the node *A* in Figure 4.2) to work as the next relaying node to forward the packet in the path. The number of hops in the path depends on the distance between the source and destination nodes (*d*) and the remaining distance to the destination per hop (per hop progress).

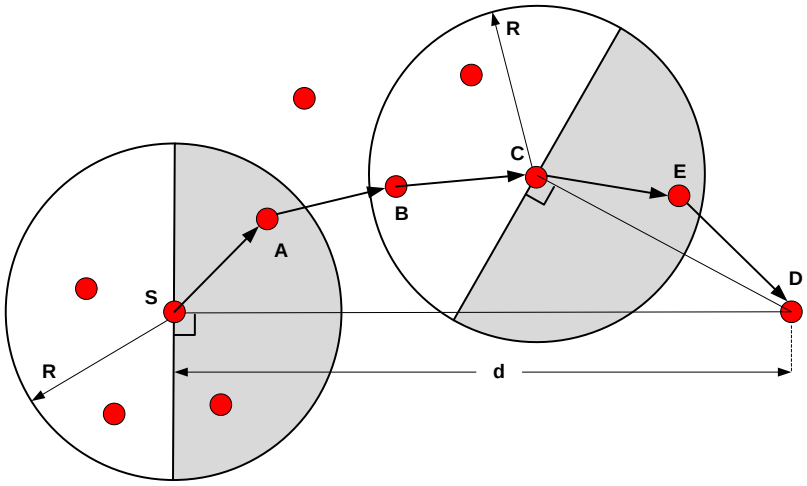

Figure 4.2: Packet forwarding in a multi-hop path

To compute the expected hop count analytically, a greedy routing approach called Maximum Hop Distance (MHD) is proposed. MHD is a packet forwarding algorithm that uses the maximum forward distance toward the destination as the optimization criterion to choose the relay node among neighbour nodes that minimizes the number of hops a packet has to traverse in order to reach the destination. In the MHD approach, the geometric probability is used to derive the PDF of the distance between the source (or a

relay node) and its neighbour nodes which is used to compute the expected value for the maximum forward distance toward the destination. Also, the expected remaining distance to the destination, which is used to calculate the expected hop count, is computed using the geometric probability.

The MHD approach succeeds if at least one router is located towards the destination (shaded regions shown in Figure 4.2) in each hop to prevent back-forwarding of packets; otherwise, it fails. For example, as shown in Figure 4.2, for nodes *S* and *C*, nodes *A* and *E* are located in the grey half circle towards the destination *D* to forward the packets from *S* and *C*, respectively, to the destination. Intuitively, to keep the connectivity of the route, each node needs at least two neighbour nodes; one is for the previous hop and the other is for the next hop. Therefore, the node density must exceed a certain threshold in order to ensure the route and network connectivity. In [97, 98], it has been shown that the average number of neighbour nodes required to ensure one-connectivity is eight. Hence, to use MHD to analyse the expected hop count in MANETs, the number of neighbour nodes might be greater than or equal eight to ensure the network connectivity. Factors that affect the average number of neighbour nodes in MANETs are the number of nodes in the network, the size of the network area, mobility patterns, and transmission range [99].

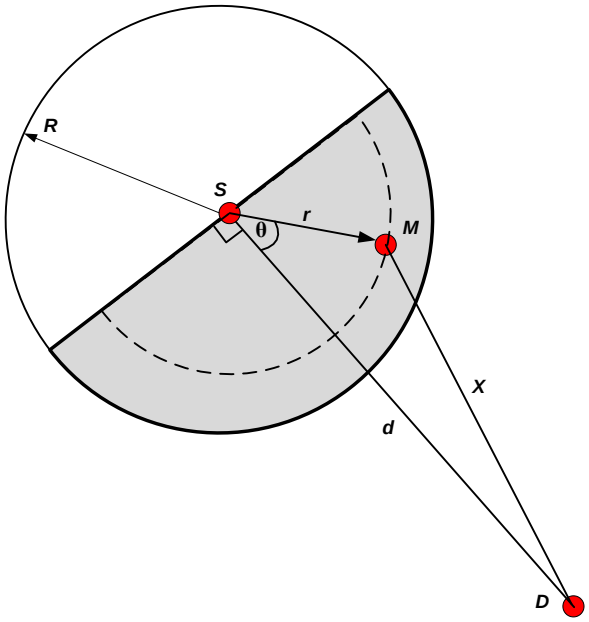

Figure 4.3: Least remaining distance for the first hop

Let *M* be the potential router used to forward the packets from *S* to *D* for the first hop, as shown in Figure 4.3. Also, let *r* and *X* respectively be the distance between the source and the router *M* (the maximum forward distance) and the remaining distance from *M* to *D*. The PDF and expected value for *r* and *X* must be derived in order to compute the expected hop count.

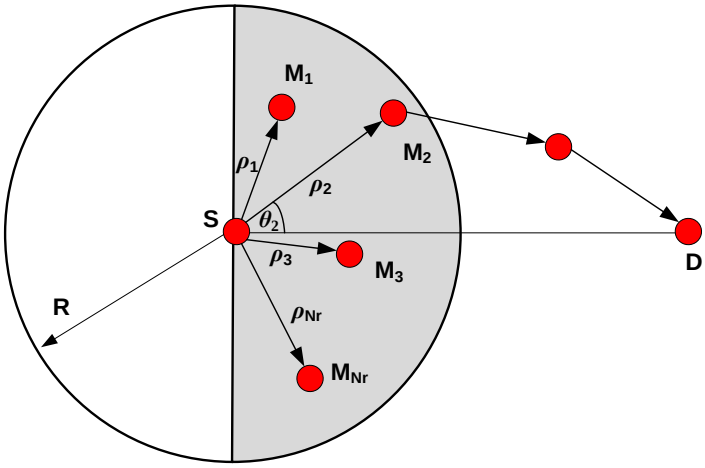

Figure 4.4: The distances between *S* and neighbour nodes

First, we derive the PDF of the maximum forward distance *r* that is used by the MHD approach as the optimization criterion to minimize the hop count. Suppose that there are N_r forwarding neighbour nodes ($M_1, \ldots, M_{N_r}$) distributed over the half circle towards the destination *D*. The distances and angles from the source *S* to the neighbour nodes are ρ_i and θ_i, where $i = 1 \ldots N_r$, as shown in Figure 4.4. For simplicity of the analysis, the neighbour nodes are assumed to be uniformly distributed around *S*. So, the expected value of N_r equals half of the expected number of neighbour nodes (N_n). For the RWPMM, the value of N_n can be computed using the methods introduced in [99]. The PDF of the distance (ρ) between *S* and the neighbour nodes is:

$$f_{\rho\theta}(\rho, \theta) = \frac{2\rho}{\pi R^2}$$

where $0 \leq \rho \leq R$ and $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. Integrating this Equation over θ gives the PDF of ρ as:

$$f_{\rho}(\rho) = \frac{2\,\rho}{R^2}$$

To minimize the hop count to the destination, the neighbour node with the maximum distance (ρ_{max}) from the source S is chosen to forward the packets. According to [100], because ρ_1, .. , ρ_{Nr} are i.i.d. random variables each with PDF $f_\rho(\rho)$, the PDF of ρ_{max} is:

$$f_{\rho_{max}}(\rho) = N_r \left(F_\rho(\rho)\right)^{N_r} f_\rho(\rho) = 2N_r \frac{\rho^{2N_r-1}}{R^{2N_r}} \tag{4.9}$$

where $F_\rho(\rho)$ is the CDF of ρ. By definition, the expected value of ρ_{max} is:

$$E(\rho_{max}) = \int_0^R \rho \, f_{\rho_{max}}(\rho) \, d\rho = \frac{2N_r}{2N_r + 1} R$$

Therefore, the expected distance r between the source S and router M, shown in Figure 4.3, is given by:

$$r = E(\rho_{max}) = \frac{2N_r}{2N_r + 1} R \tag{4.10}$$

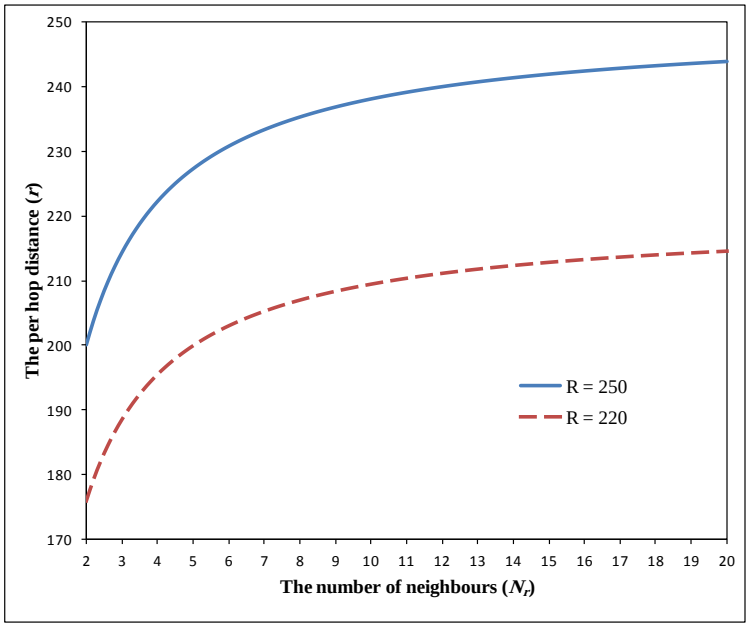

Figure 4.5: The per hop distance for different values for n and R

The resulting function for r for a given R = 250 or 220 m and increasing values of N_r is shown in Figure 4.5. Clearly, for a given transmission range R, for small values of N_r, r increases rapidly. For large values of N_r, r may reach R. Therefore, increasing the node density decreases the expected hop count, but it increases the interference between

neighbour nodes. Equation (4.10) can be used for the analysis of the distance between the source and other nodes in the path, which is important in investigating the survivability of the path.

RWPMM significantly increases the average number of neighbour nodes compared to uniformly distributed nodes [99]. As shown in Figure 4.5, an increase in the number of neighbour nodes (N_r) increases the maximum forward distance (r), which decreases the expected hop count. Therefore, the expected number of hops for nodes moving according to RWPMM is less than that for uniformly distributed nodes.

To derive an expression for the remaining distance, we consider that the router *M* may be located at any point on the circumference of a half circle with a radius *r* (the dashed half circle shown in Figure 4.3) computed using Equation (4.10), as shown in Figure 4.3. Let *M* be located at a random angle θ. So, the domain of θ is $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. The remaining distance *X* can be described using a PDF as follows:

$$f_\theta(\theta) = \frac{1}{\pi} \qquad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

The probability that θ is smaller than a given value a can be computed by the integral of the last equation as:

$$P_\theta(\theta \leq a) = \int_{-a}^{a} f_\theta(\theta)\, d\theta = \frac{2}{\pi}\,\theta \tag{4.11}$$

From geometry, $d^2 + r^2 - 2\,d\,r\cos\theta = X^2$. Therefore, by substitution in Equation (4.11), we get the CDF of *X* as:

$$F_X(X) = P_x(X \leq x) = \frac{2}{\pi} Arccos\left(\frac{d^2 + r^2 - X^2}{2\,d\,r}\right)$$

The last equation is differentiated to get its PDF of *X* as:

$$f_X(X) = \frac{2X}{\pi\,d\,r\sqrt{1 - \left(\frac{d^2 + r^2 - X^2}{2\,d\,r}\right)^2}} \tag{4.12}$$

By definition, the expected value of *X* can be deduced from Equation (4.12) as follows:

$$X_r = E(X) = \int_{d-r}^{\sqrt{d^2+r^2}} X \cdot f_X(X)\, dX \tag{4.13}$$

The last equation can be easily evaluated numerically.

Having derived the remaining distance X_r from the router M to the destination for the first hop, in order to get the expected number of hops the current distance to the destination d in the next hop is replaced by the remaining distance X_r obtained using Equation (4.13). Then, the process is repeated and the hops are counted until X_r falls below the transmission range R. The following procedure summarizes this process:

Step 1: Set the inputs N, R, and L

Step 2: Set the number of hop count *to* $hop\_count = 0$

Step 3: Compute the expected distance between the source and destination (d) using Equation (4.7)

Step 4: Compute the expected maximum distance between the source and router (r) using Equation (4.10)

Step 5: If $d \leq R$, then $hop\_count = \frac{d}{R}$, go to the End

Step 6: Set $hop\_count = hop\_count + 1$

Step 7: The remaining distance between the router and destination (X_r) is computed using Equation (4.13)

Step 8: If $X_r \leq$ R, then $hop\_count = hop\_count + \frac{X_r}{R}$

Step 9: If $X_r > R$, then set $d = X_r$ and go to step 6

Step 10: End

4.4 Validation

In this section, the proposed approach is validated by comparing the theoretical and simulation results. The theoretical analysis of the expected Euclidean distance between any random source and destination nodes, introduced in Section 4.2, is first validated by network simulation. For this validation the MobiSim tool [101] was used. This employs

topological characteristics to analyse and manage the mobility scenarios for ad hoc networks. A simulation scenario is considered which consists of a square system area of a side length L that varies from 400 to 1000 m. A set of 200 nodes are uniformly scattered in the square area and move according to RWPMM. Every node moves towards the destination point with a velocity chosen uniformly from 1 to maximum speed (V_{max}). When it reaches the destination it chooses and moves towards a new destination in a similar manner. The maximum moving speed is set to 20 m/s. A zero pause time was chosen so that the nodes are constantly moving. All nodes have a transmission range of 250 m.

For each mobility scenario, the expected distance between any source-destination pair is computed by taking the average of the distances between every pair of nodes. Many different mobility scenarios (with different random seeds) are generated until the expected distance between nodes is within a 95% confidence interval with 1% relative error. Figure 4.6 shows the simulation and analysis results for the expected distance between any two nodes for varying values of the side length of the square area. The comparison between analytical and simulation results shows the accuracy of the proposed analysis.

To validate the proposed theoretical analysis and the procedure used to compute the expected number of hops for a packet transmission in ad hoc networks, a series of simulation tests was performed using NS2 [27]. The simulation settings consist of a network with a square area. The side length of the square area varies from 700 to 1600 m. The maximum speed of a node is set to 20 m/s. The simulation time is set to 1500 seconds. To be sure that the average number of neighbour nodes is greater than or equal to 8 nodes, the node density is varied depending on the size of the system area. To illustrate its effect on the expected number of hops, the transmission range is considered to be 200 or 250 m. The random waypoint mobility patterns used in all simulation tests are generated using the setdest tool, which is a node movement generator tool implemented by the current NS2 version.

The number of hops between nodes can be computed on the fly during simulation runs. But this method takes a long time (possibly lasting days) especially with a large number of nodes and network area size. Alternatively, an object called the General Operations Director (GOD) was used, which is implemented with the setdest tool and used to

manage the shortest path information between nodes. For the whole simulation period, GOD is aware of any changes in the mobile wireless network topology. GOD is an omniscient observer which is used to store global information about the topology of the network. This global information is not completely available to any node, but partial information is provided to each node when needed. GOD is used to store an array of the optimal path length in hops between every pair of nodes. This information is used to analyse and develop ad hoc network routing protocols.

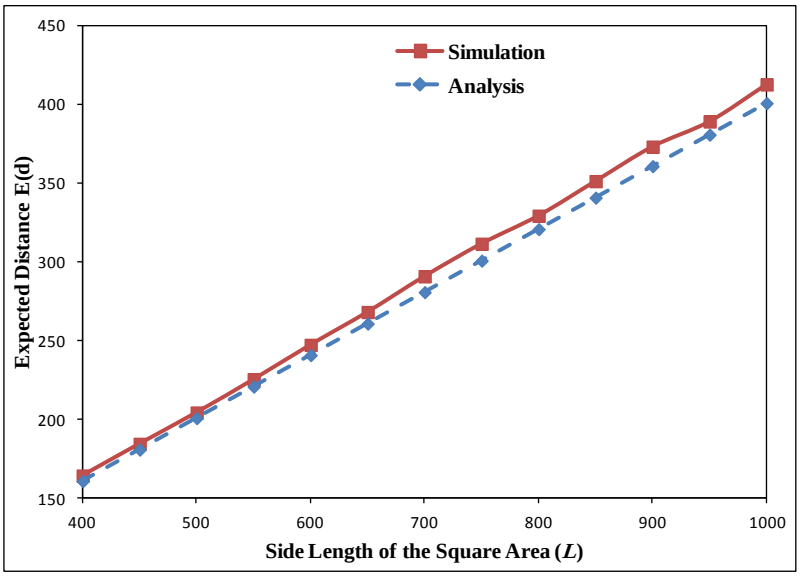

Figure 4.6: Expected distance for different sizes of the network area

For the same network settings, the expected number of hops is computed by averaging the number of hops between every pair of nodes. Many mobility patterns for the same network settings were generated and the expected number of hops was computed with a confidence level of 95% and a relative error threshold of 2%. Table 4.1 shows the simulation and theoretical results for the expected number of hops for two different values of transmission range (R = 200 or 250 m) and increasing values of the side length of the square area of the simulated network. As shown in Table 4.1, for a given transmission range, the expected number of hops increases significantly as the network size increases because of the increasing expected distance between the source and destination. In addition, as expected, the expected number of hops decreases with increasing transmission range because of corresponding increase in the per hop progress. As can be seen in Table 4.1, the theoretical results are accurate compared to the simulation results.

Table 4.1: Analytical and simulation results for expected hop count for increasing values of the side length of the network area where R = 200 or 250 m

L	Expected Number of Hops			
	R = 200		R = 250	
	Sim	Ana	Sim	Ana
700	2.61	2.73	2.01	1.82
800	2.93	2.84	2.14	1.91
900	3.19	2.97	2.69	2.75
1000	3.65	3.8	2.94	2.83
1100	4.10	3.92	3.18	2.93
1200	4.58	4.76	3.57	3.76
1300	4.93	4.86	3.86	3.83
1400	5.28	4.99	4.16	3.92
1500	5.79	5.82	4.52	4.75
1600	6.19	5.93	4.63	4.81

To compare the LRD and MHD approaches, Table 4.2 shows the expected number of hops computed using the two approaches and simulations for the same network settings used to validate the MHD approach where R = 150 m. Compared with the simulation results, it is clear that the accuracy of the MHD approach is much better than the LRD approach, as shown in Table 4.2. The expected number of hops computed using the LRD approach is much lower than in the simulation, especially for long routes. This is because the LRD approach supposes that the routers with the minimum remaining distance to the destination constitute the shortest path to the destination, which is only true when node density is very high.

Compared with simulation, the computation time required for theoretical analysis is trivial. For example, in the case of N = 250, R = 150 m, L = 1600 m, the simulation time = 1500s, and 95% confidence interval and 2% relative error, the time required for generating the mobility patterns and computing the expected hop count is about 28.2 hours, whereas the time required for theoretical analysis is less than 2 seconds, when the simulation and theoretical analysis was conducted on a desktop workstation equipped with a 2.6 GHz (Intel Q9400 Core 2 Quad) processor, 4 GB of RAM and Ubuntu Linux version 8.10.

Table 4.2: Comparison of simulation and LRD and MHD results for expected hop count for increasing values for L where R = 150 m

L	**Sim**	**LRD**	**MHD**
700	3.37	2.39	3.74
800	3.83	2.62	3.86
900	4.40	2.86	4.75
1000	4.84	3.40	4.86
1100	5.42	3.63	5.75
1200	5.83	3.88	5.86
1300	6.35	4.4	6.74
1400	6.72	4.64	6.85
1500	7.06	4.89	6.99
1600	7.35	5.42	7.83

4.5 Summary

This chapter presents a theoretical analysis of the expected number of hops in mobile ad hoc networks where nodes move according to the random waypoint mobility model. The proposed approach can be used to analyse the hop count for other mobility models. It depends on computing the expected distance between the source and destination nodes, per hop distance, and per hop progress, which are used to compute the expected hop count. The proposed approach has been validated using network simulations for different network parameters. The impact of the transmission range, node density, and size of the network area on the hop count have been investigated. Compared to other methods proposed in the literature, the accuracy of the proposed approach is much better.

Chapter 5

A Path Analysis Model for Mobile Ad Hoc Networks with Random Waypoint Mobility

5.1 Introduction

Mobile nodes in MANETs cannot communicate directly with all other nodes in the network via the wireless channel. When a node sends a message to another node beyond its transmission range it uses the other nodes as relay points. As a result, mobile nodes work as both sources and routers for other mobile nodes in the network. Due to the dynamic topology of MANETs, the route followed by packets to reach a destination varies frequently. Thus, the routes in MANETs are prone to frequent breaks (called mobility failures) which strongly affect the performance of MANETs compared to wired or cellular networks.

A network failure is any condition that does not achieve a normal network operation. There are two types of network failures: node failure and link failure [102]. Any failure that makes a node unavailable due to hardware or software faults is a node failure. A link failure involves the disconnection of radio links due to excessive noise, interference, signal loss, or mobility. The ability of the network to efficiently deliver

a service to users is significantly affected by network failures. Therefore, the network performance under failures is an issue of great concern.

The ability of the network to avoid or cope with failure is measured in the form of performance metrics such as survivability, reliability, and availability [103]. Path connection availability is the probability that a link or route exists between any source-destination pair at a given time. This is an essential reliability performance characteristic of MANETs because of the need for multi-hop communication and it can be used as a global measure for the performance of MANETs. Also, understanding the factors that affect path connection availability can help in understanding path stability under various degrees of system dynamics.

Although the random waypoint mobility model is one of the most commonly used mobility models in MANET studies, to the best of our knowledge, no analytical study has yet investigated path connection availability, and path failure and repairing frequency in multi-hop ad hoc networks with this mobility model. This is because the spatial distribution of nodes moving according to the random waypoint mobility model is non-uniform, which significantly complicates the analytical analysis of the network. In addition, no previous work in the literature provides a closed form solution for the analytical analysis of path connection availability for multi-hop ad hoc networks.

In this chapter, a closed form solution is proposed using a new stochastic reward net model to analyse path connection availability in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model. The effects of link failure due to the mobility of nodes on path connection availability in MANETs are analytically investigated using the proposed model. In addition, the effect on path connection availability of different factors are investigated, such as the number of nodes in the network, transmission range, network area size, data transmission rate, and routing protocol. The proposed model incorporates the characteristics of reactive routing protocols such as dynamic source routing and ad hoc on-demand distance vector.

To choose the best route to the destination, different criteria may be used such as hop counts, path quality, and path bandwidth. The most widely used criterion is hop counts. Thus, the proposed model can be exploited to evaluate paths based on their connection stability. It can also be used to study the relationship between path length and path connection availability, which can help in determining the appropriate network size and node density to achieve high connection availability.

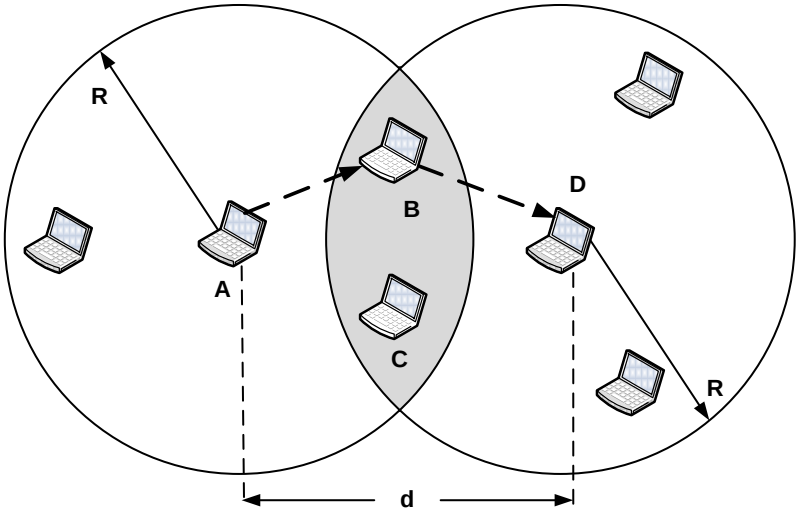

Figure 5.1: Two hops communication path

5.2 Ad Hoc Network Model Description

To develop a path connection availability model, we consider a network consisting of N nodes distributed in a square area of dimension $L \times L$ which move according to the random waypoint mobility model. All nodes are independent and behave identically. Each node is equipped with an omni-directional antenna and has a fixed transmission range R. The destination of any source is chosen randomly from other nodes. For the end-to-end connection, if the destination is not in the transmission range of the source, the packets are routed via N_h hops through neighbour nodes.

Figure 5.1 shows a two hops communication path between the source node A and destination node D, where the transmission area of each node is represented by a

circle with radius *R* (the transmission range). To be able to establish a connection with node *D*, node *A* has to choose one of the nodes (*B* or *C*) located in the intersection area between the area covered by the transmission ranges of *A* and *D* (A_{AD}) (shaded area), which is simply called the intersection area. As shown in Figure 5.1, node *A* uses node *B* as a router to forward its packets to node *D*. If there are N_{in} nodes in the intersection area, one of them is used as a router (called the active router) and $N_{in} - 1$ nodes are considered as backup routers. When the active router fails, one of the backup routers is used to forward the packets.

Faults in nodes in the intersection area can be classified into two categories: node and link faults. A node fault is the failure of a node due to hardware, software, or power faults, where the latter is caused by the insufficient battery power to send the packets. The sources of link faults are errors in the wireless channel caused by signal attenuation, signal loss, multipath fading, excessive noise and interference, and obstacles between nodes. At any instant, because of their mobility, either the active or any backup router may leave the intersection area (and becomes unavailable) which is considered to be a link fault. In this chapter, we are interested in studying the effect of node mobility on path connection availability in MANETs. Therefore, the proposed model only considers the effect of link faults due to mobility, but it can be easily modified to cope with other types of faults.

At any instant of time, any node can enter the intersection area A_{AD} and leave it after an average period of time of μ seconds (called the leaving time). It is supposed that one of the nodes located outside the intersection area enters the intersection area A_{AD} every average period of time of $1/\lambda_e$ seconds. μ and λ_e are the model parameters which are directly affected by many other network parameters such as the number of nodes, size of network area, mobility pattern, speed of nodes, type of routing protocol, and transmission range.

A three hop communication path between a source node *A* and destination node *D* is shown in Figure 5.2. It is clear that there are two intersection areas (shaded areas) in the route between nodes *A* and *D*. In general, the number of intersection areas in N_h

hops route is $N_h - 1$. When the active router fails due to any type of fault in any intersection area of the path, the connection between the source and destination becomes unavailable. The routing protocol then tries to re-establish the connection by starting the route recovery (or maintenance) process in which one of the backup routers in the intersection area is chosen to forward the packets. During route recovery, queued packets are delayed until the route is established. The time required for route recovery depends on many parameters, such as node density, transmission range, type of fault, distance (or number of hops) between the source and destination, and type of routing protocol. During the search for a new route, the connection will be completely unavailable.

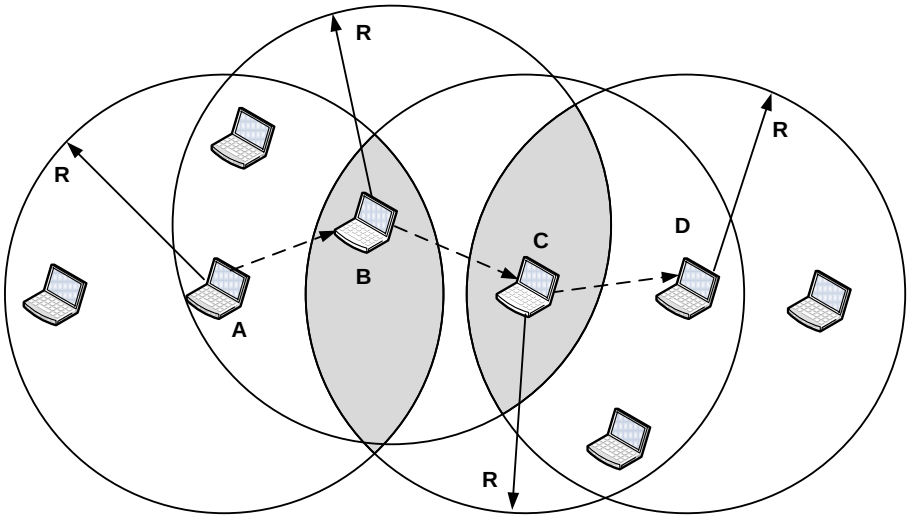

Figure 5.2: Three hops communication path

For reactive routing protocols, the route recovery mechanism differs from one routing protocol to another. For the AODV protocol, there are two route recovery mechanisms: local and source recovery. In the local recovery mechanism, if the node that detected the link failure (called the upstream node) is nearer to the destination than the source, it tries to repair the link locally itself. The upstream node sends a Route Request (RREQ) message where the Time To Live (TTL) of the message is set to (Max (N_{LH}, N_{HS}/2) + 2), where N_{LH} is the last known hop count to the destination, and N_{HS} is the number of hops to the source of the undeliverable packet [13]. When

the local repair fails, or the upstream node is nearer to the source than the destination, the upstream node starts the source repair process by sending back a Route Error (RERR) message to the source, which then initiates a new route discovery.

For the DSR protocol, the route maintenance mechanism does not locally repair a broken link [9]. If a link failure is detected, the upstream node returns a RERR message to the source of the packet, identifying the link over which the packet could not be forward. Then, once the source node receives the RERR message, it removes the broken link from its cache and searches within it for another route to the same destination. If there exists a cached route to the same destination, the source immediately sends the packet using the new route. Otherwise, it may perform a new route discovery for this destination after an exponential backoff delay.

5.3 SRN Model Description

Figure 5.3 shows the proposed SRN model for the connection availability of a path with N_h hops. The model consists of $(N_h - 1)$ parts with similar structure (Figure 5.3 shows a dashed rectangle around each part), where each part models one of the intersection areas in the path. The following describes the model structure of the intersection area number k in the proposed SRN model for N_h hops.

The number of tokens in the place P_{ik} (N_{in}) represents the number of nodes in the intersection area, which represents the number of available routers to the next hop. One of the nodes in the intersection area is used as a router (active router) in the current route between the source and destination, and the other $(N_{in} - 1)$ nodes work as backup routers. For the random waypoint mobility model, the method introduced in [99] can be used to compute the average number of neighbour nodes which can be used to compute N_{in}, as illustrated in Section 5.5. At any time, there is a probability that any of the backup routers can leave the intersection area, representing the failure of one of the backup routers, which reduces the number of available routers by one. This is modelled by the arc between the place P_{ik} and transition T_{Lk} which moves one token from the place P_{ik} to P_{Lk} after the firing of the transition T_{Lk}. On the other hand,

there is a possibility that the active router may leave the intersection area, which renders the route to the destination unavailable. The arc between the place P_{ik} and transition T_{PFk} represents this action. The firing of the transition T_{PFk} moves one token from the place P_{ik} to the place P_{Fk} which represents the failure of the path.

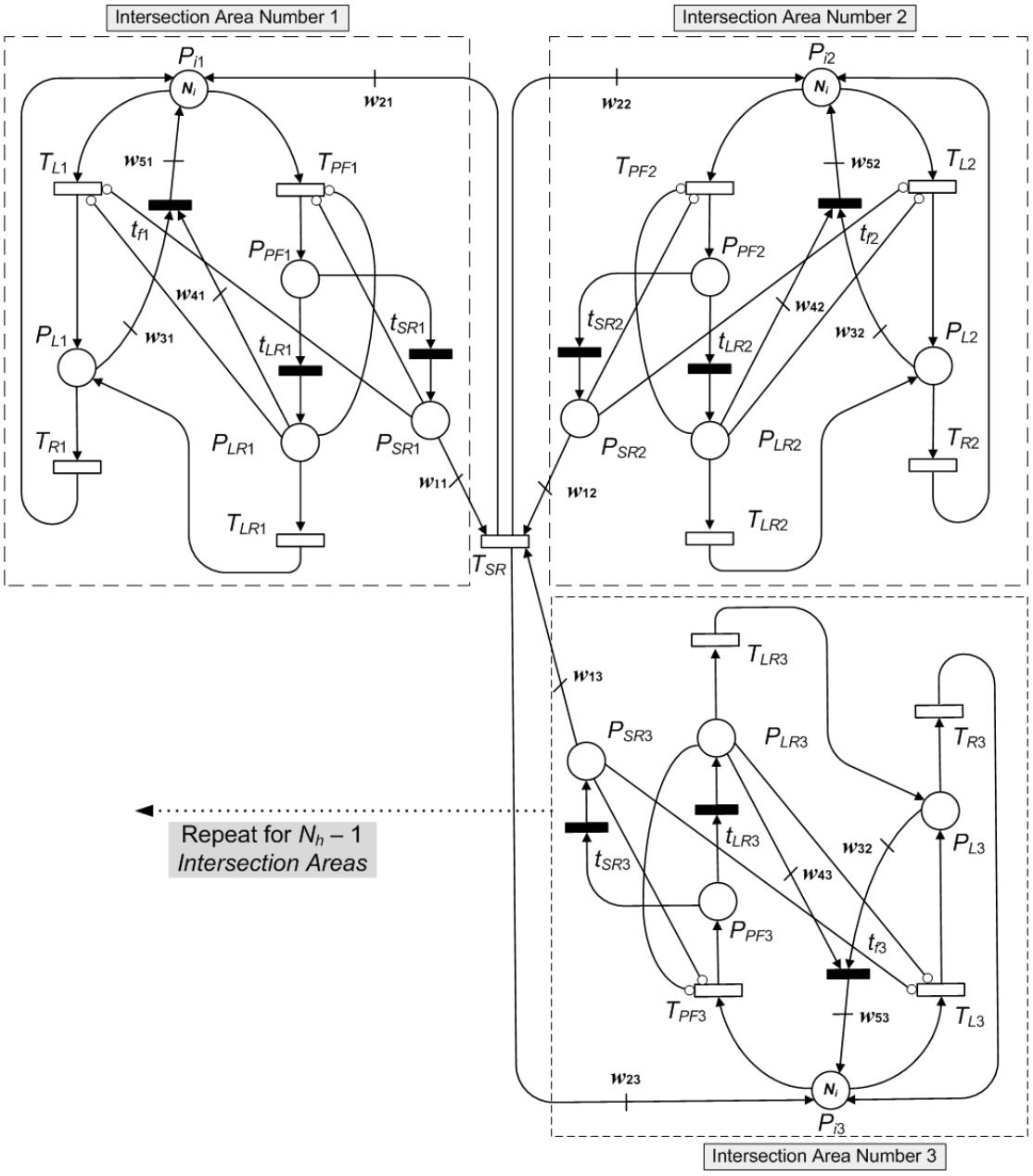

Figure 5.3: SRN model for connection availability

The average firing rate of the transition T_{PFk} depends on the leaving time μ (the average time that a node spends in the intersection area), whereas the average firing rate of the transition T_{Lk} not only depends on μ but also on the number of nodes in the intersection area. The average firing rates of T_{Lk} and T_{PFk} are $1/\mu$ and $(\#P_{ik}/\mu)$, respectively, where $\#P_x$ is the number of tokens in the place P_x. The leaving time μ depends on the size of the intersection area and the relative speed between any router in the intersection area and the source or destination node. Increasing the maximum limit of the node speed decreases the leaving time μ, whereas increasing the size of the intersection area increases it. Section 5.4 derives an expression for the leaving time μ.

The number of tokens in the place P_{Lk} represents the number of backup routers that have left the intersection area. These nodes that left the intersection area or any other node in the network may enter the intersection area. This is represented by the firing of the transition T_{Rk} which moves a token from P_{Lk} to P_{ik}. The average firing time of the transition T_{Rk} is the frequency with which the nodes in the network enter into an intersection area (called the entering rate λ_e). The entering rate depends on network parameters such as the node density, speed of nodes, pause time, and transmission range. The larger the node density, speed of nodes, or transmission range, the greater the entering rate. An expression for the entering rate λ_e is derived in Section 5.4.3.

The place P_{PFk} represents the failure of the active router in the intersection area and consequently the whole route. After the failure of the active router (e.g. node *C* in Figure 5.2), the node that detected the failure (e.g. node *B* in Figure 5.2) will try to recover the route. For some routing protocols such as AODV, in order to avoid the route discovery by the source, which takes a long time, if the upstream node is closer to the destination than the source, it starts the local repair which is presented by the firing of the immediate transition t_{LRk}. If the local repair process is not supported or the upstream node is nearer to the source than the destination, the upstream node sends a route error message to the source node indicating the failed link. Then the source node initiates another route search process to find a new path to the destination, which is modelled by the firing of the transition t_{SRk}.

The firing of the transition t_{LRk} deposits a token in the place P_{LRk} whose marking represents the success of the local repair process, whereas the firing of the transition t_{SRk} deposits a token in the place P_{SRk} representing the initiation of the source repair process. Because the local repairing of the route needs at least one node to be in the intersection area and cannot be started if the upstream node is closer to the source than destination, a guard function is set to disable the transition t_{LRk} when ($\#P_{ik} = 0$) or ($k > 1 + N_h/2$). The firing of the transition T_{LRk} deposits a token in the place P_{Lk} to represent that the active router has left the intersection area. As illustrated in Section 5.2, in some cases local repair is not supported. So, for these cases, the transitions t_{LRk} and T_{LRk} should be removed from the model.

Table 5.1: Arcs weight functions for SRN model of intersection area number k

Arc name	Arc weight function
W_{1k}, W_{2k}	1 IF $\#P_{SRk} > 0$ 0 ELSE
W_{3k}	$\#P_{Lk}$ IF $\#P_{SR\Delta} > 0$, $\Delta = 1, 2, \ldots$, or $N_h - 1$ 0 ELSE
W_{4k}	1 IF $\#P_{SR\Delta} > 0$, $\Delta = 1, 2, \ldots$, or $N_h - 1$ 0 ELSE
W_{5k}	$\#P_{Lk} + 1$ IF $\#P_{LRk} = 1$ $\#P_{Lk}$ ELSE

The end of the source repair process is represented by the firing of the transition T_{SR}, which moves the token from the place P_{SRk} to P_{ik}. During the source repair process, the source node tries to find new routers in new intersection areas. So, the failure of the nodes in the old intersection area is not a concern. Thus, an inhibitor arc between the place P_{SRk} and transition T_{Lk} is added to disable it when $\#P_{SRk} > 0$. Also, to disable the transition T_{PFk} during the local or source repair ($\#P_{LRk} > 0$ or $\#P_{SRk} > 0$), the inhibitor arcs from places P_{LRk} and P_{SRk} to the transition T_{PFk} are added. During searching for the new route, it is supposed that there will be N_{in} routers in new intersection areas. So, the immediate transition t_{fk} is added, which flushes P_{Lk} and P_{LRk} and moves all tokens back to P_{ik} when $\#P_{SRx} > 0$, where $x = 1, 2, \ldots$, or $N_h - 1$. This is

controlled by the arc weight functions w_{3k}, w_{4k} and w_{5k}, shown in Table 5.1, and a guard function for the transition t_{fk}. If $\#P_{SRk} = 0$, the arc weight functions w_{1k} and w_{2k} prevent depositing a token to the place P_{ik} when T_{SR} fires. To disable the transition T_{Lk} and enable the transition T_{PFk} when all backup routers fail ($\#P_{Lk} = N_{in} - 1$) and only the active router is in the intersection area ($\#P_{ik} = 1$), a guard function is set to the transition T_{Lk}.

For the AODV route maintenance mechanism, if the upstream node is far from the source node, it broadcasts RREQ with TTL set to $\left(\text{Max}\left(N_{LH}, \frac{N_{HS}}{2}\right) + 2\right)$ in order to repair the broken link locally, as illustrated in Section 5.2. Therefore, the average firing time of the transition T_{LR} (the time required to finish the local repair process τ_{LR}) is:

$$\tau_{LR} = 2\,\delta\left(\text{Max}\left(N_{LH}, \frac{N_{HS}}{2}\right) + 2\right)$$

where δ is the packet delay per hop. δ can be computed using another analytical model or simulation.

The average firing time of the transition T_{SR} (τ_{sr}) is the average time needed to complete the source repair process. It can be computed as:

$$\tau_{sr} = \tau_{LR} + \tau_{RERR} + \tau_{NR}$$

where τ_{RERR} and τ_{NR} are the times required for broadcasting the RERR message and establishing a new route, respectively. Hence,

$$\tau_{sr} = 2\,\delta\left(\text{Max}\left(N_{LH}, \frac{N_{HS}}{2}\right) + 2\right) + \delta \cdot N_{HS} + 2\,\delta \cdot N_h$$

$$= 2\,\delta \cdot \left[\left(\text{Max}\left(N_{LH}, \frac{N_{HS}}{2}\right) + 2\right) + \frac{N_{HS}}{2} + N_h\right]$$

In cases where local repair is not supported, the local repairing time is equal to zero ($\tau_{LR} = 0$) and τ_{sr} is given by:

$$\tau_{sr} = \delta \cdot (N_{HS} + 2N_h)$$

In DSR, when the source receives the RERR message, but before starting a new route discovery process, it tries to use all other alternative routes in the cache to send the packet. So, to compute the average repairing time, the caching mechanism of DSR with random waypoint mobility should be modelled, which is beyond the scope of this work. Therefore, it is measured by simulation.

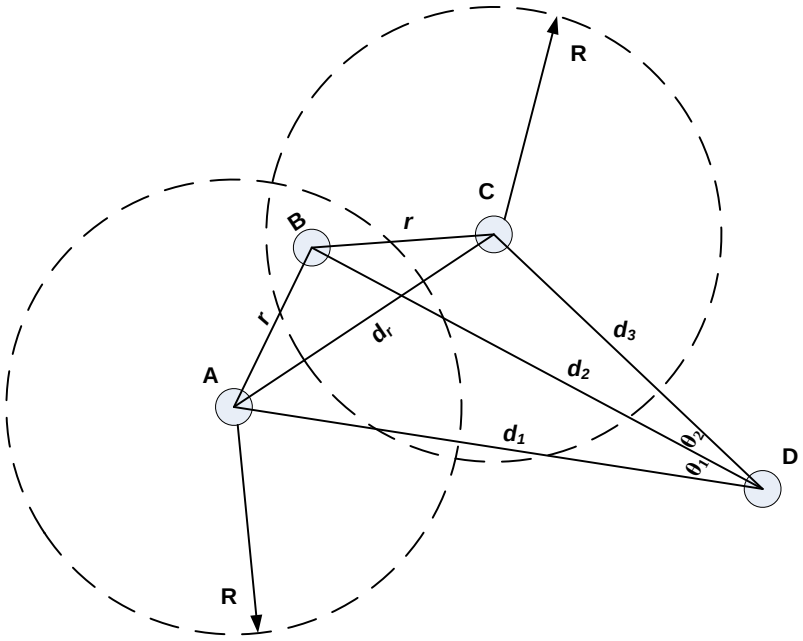

Figure 5.4: Distance between nodes

5.4 Model Parameters

As illustrated above, to solve the proposed model, two important parameters should be known: the average time needed for a node to pass through the intersection area (leaving time μ), and the frequency with which the nodes in the network enter into an intersection area (entering rate λ_e). To compute μ and λ_e, the distance between 2-hop-apart nodes d_r in the path must be derived. This section derives expressions for these parameters.

5.4.1 Distance Between Nodes

To derive an expression for the distance d_r between 2-hop-apart nodes in the path, suppose that a source node *A* tries to send its packets to a destination node *D*, and the first two routers in the path are nodes *B* and *C*, as shown in Figure 5.4. The distances from *D* to *A*, *B*, and *C* are d_1, d_2, and d_3, respectively, which are called the remaining distances to the destination. The distance between any two nodes in the path (*r*) is called the forward distance. Chapter 4 derived expressions for computing the expected values for the remaining distances and forward distance in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model.

From the geometry of Figure 5.4, the distance d_r between nodes *A* and *C* is:

$$d_r^2 = d_1^2 + d_3^2 - 2d_1 d_3 Cos(\theta_1 + \theta_2) \tag{5.1}$$

Also, from geometry it is to be noted that:

$$Cos\ \theta_1 = \frac{T_1}{2d_1 d_2} \qquad\qquad Cos\ \theta_2 = \frac{T_2}{2d_2 d_3} \tag{5.2}$$

where $T_1 = d_1^2 + d_2^2 - r^2$ and $T_2 = d_2^2 + d_3^2 - r^2$. It is known that:

$$Cos(\theta_1 + \theta_2) = Cos\ \theta_1\ Cos\ \theta_2 - Sin\ \theta_1\ Sin\ \theta_2 \tag{5.3}$$

By substituting from Equations (5.2) and (5.3) into Equation (4.1), the distance between nodes is derived as:

$$d_r = \sqrt{d_1^2 + d_3^2 - \frac{T_1 \cdot T_2}{2\ d_2^2} - \frac{1}{2\ d_2^2} \cdot \sqrt{(4d_1^2\ d_2^2 - T_1^2) \cdot (4d_2^2\ d_3^2 - T_2^2)}} \tag{5.4}$$

5.4.2 Leaving Time

The leaving time (μ) is the average time needed to pass through the intersection area. Consider the two hops communication path between the source node *A* and

destination node *D* shown in Figure 5.1. The leaving time of the nodes *B* or *C* depends on the distance between 2-hop-apart nodes in the path (d_r), transmission range, and the speed of nodes. In random waypoint mobility, the speed of nodes is uniformly randomly chosen from the predefined range [V_{min}, V_{max}]. Therefore, the average speed of nodes (V_a) is given by [99]:

$$V_a = \frac{V_{max} - V_{min}}{\ln(V_{max}) - \ln(V_{min})}$$

Because the intersection region is very small compared to the other network area, to simplify the analysis it is assumed that nodes do not change their direction and speed when they cross the intersection area. Thus, the average leaving time is given by:

$$\mu = \frac{L_i}{E(V_r)} \tag{5.5}$$

where L_i (called the expected intersection area path length) is the average length of the path that a node passes through in the intersection area, and $E(V_r)$ is the expected value of the relative speed between the router (e.g. node *B* in Figure 5.1) and the preceding or following node in the path (e.g. nodes *A* or *D*). L_i depends on the distance between nodes and the angle of entry into the intersection area. The average value of the intersection area path length is given by [77]:

$$L_i = \frac{\pi R}{2}\left[1 - \frac{d_r \cdot \sqrt{4R^2 - {d_r}^2}}{4R^2 \; ArcCos\left(\frac{2}{d_r}\right)}\right] \tag{5.6}$$

In order to compute the average leaving time, $E(V_r)$ should be known first. According to the law of cosine, the relative speed (V_r) between nodes *A* and *B* is given by:

$$V_r = \sqrt{V_A^2 + V_B^2 - 2V_A V_B \cos(\theta)} \tag{5.7}$$

where V_A and V_B are the speeds of node *A* and *B* respectively, and θ is the angle between V_A and V_B. The angle θ can vary from 0 to π. Because all nodes move with a

speed that is uniformly distributed in the range [V_{min}, V_{max}], it is assumed that the velocities of node A and B are equal to the average speed of nodes (V_A = V_B = V_a). Hence Equation (5.7) is expressed as follows:

$$V_r = 2V_a \sin\left(\frac{\theta}{2}\right)$$

Therefore, the angle θ is expressed using V_r and V_a as:

$$\theta = 2\ ArcSi\,n\left(\frac{V_r}{2V_a}\right) \tag{5.8}$$

Assuming that θ is uniformly distributed in the range [0, π], the probability density function of θ can be described as follows:

$$f_\theta(\theta) = \frac{1}{\pi}$$

The probability that θ is less than Ω is given by:

$$F_\theta(\Omega) = P(\theta \leq \Omega) = \int_0^{\Omega} f_\theta(\theta)\ d\theta = \frac{\Omega}{\pi}$$

The cumulative distribution function of V_r can be obtained by substituting θ from Equation (5.8) into the last equation as follows:

$$F_{V_r}(v) = P(V_r \leq v) = \frac{2}{\pi} ArcSi\,n\left(\frac{v}{2V_a}\right) \tag{5.9}$$

where $0 \leq v \leq 2V_a$. By definition the PDF of V_r ($f_{V_r}(v)$) is given by the differentiation of Equation (5.9):

$$f_{V_r}(v) = \frac{2}{\pi\sqrt{4\,V_a^2 - v^2}}$$

The expected value for V_r is:

$$E(V_r) = \int_0^{2V_a} v \; f_{V_r}(v)\, dv \;\; = \frac{4V_a}{\pi} \tag{5.10}$$

By substituting from Equations (5.6) and (5.10) into Equation (5.5), the expected leaving time is:

$$\mu = \frac{\pi^2 R}{8\, V_a}\left[1 - \frac{d_r \cdot \sqrt{4R^2 - d_r^2}}{4\, R^2 \;\; ArcCos\left(\frac{2}{d_r}\right)}\right]$$

5.4.3 Entering Rate

At any time, any node located outside the intersection area can enter it to be used as a backup router. The frequency with which the nodes in the network enter into an intersection area is called the entering rate (λ_e). The entering rate depends on many parameters such as the mobility pattern, node density, speed of nodes, pause time of node, transmission range, and distance between nodes. An approximate method has been introduced to compute the entering rate [76, 77], but this method does not take into account the effect of mobility model or node density. This section thus introduces a more accurate method to compute the entering rate.

To simplify the analysis, it is assumed that no more than one node enters the intersection area at the same time. In addition, the path length of any node crossing the intersection area is equal to the average path length computed using Equation (5.6). So, it is assumed that only one node leaves the intersection area at a time. Therefore, the intersection area can be approximately modelled as a simple *M*/*M*/1/*K* queue model where the intersection area and nodes present the queue and jobs. Thus, the arrival rate of jobs equals the entering rate λ_e and *K* is the queue size which equals the number of nodes *N*. The queue service rate equals the rate at which the nodes leave the intersection area, which depends on the number of nodes in the intersection area and μ.

The steady state probabilities of the *M*/*M*/1/*K* queue with state dependent service rates are [104]:

$$P_n = \frac{P_0\ \omega^n}{n!} \tag{5.11}$$

$$P_0 + \sum_{n=1}^{K} P_n = 1 \tag{5.12}$$

where P_0 and P_n are the probabilities of initial state and state number n, respectively, and $\omega = \lambda_e \cdot \mu$. By substituting from Equation (5.11) into Equation (5.12), we obtain:

$$1 + \sum_{n=1}^{K} \frac{\omega^n}{n!} = \frac{1}{P_0} \tag{5.13}$$

The expected length of the queue $E(Q)$ can be derived as follows:

$$\begin{aligned} E(Q) &= \sum_{n=1}^{K} n\ P_n = \sum_{n=1}^{K} n\ \frac{P_0\ \omega^n}{n!} \\ &= P_0\ \omega \sum_{n=1}^{K} \frac{\omega^{n-1}}{(n-1)!} = P_0\ \omega \sum_{n=0}^{K-1} \frac{\omega^n}{n!} \\ &= P_0\ \omega \left[1 + \sum_{n=1}^{K} \frac{\omega^n}{n!} - \frac{\omega^K}{K!}\right] \end{aligned} \tag{5.14}$$

Substituting Equation (5.13) into (5.14) gives:

$$E(Q) = \omega - P_0\ \frac{\omega^K}{K!} \tag{5.15}$$

For large values of K, the second term in the last equation ($P_0\ \frac{\omega^K}{K!}$) is very small (less than 10^{-7} in the case of $P_0 < 1$, $\omega < 10$, $K > 40$) compared to the first term (ω), so it can be neglected. Hence, the expected length of the queue can be approximately evaluated to:

$$E(Q) = \omega = \lambda_e \cdot \mu \tag{5.16}$$

To compute the entering rate λ_e, the expected number of nodes in the intersection area (the expected queue size $E(Q)$) must be known. For random waypoint mobility, the author in [99] derived an expression for the expected number of neighbour nodes N_n (node degree) using a complex geometric probability analysis taking into account the speed of nodes, pause time, node density, border effects and the non-uniformity of node distribution of the mobility model. Using the average number of neighbour nodes N_n computed using the method introduced in [99], the expected number of nodes in the intersection area can be computed as follows:

$$E(Q) = \frac{A_{int}}{\pi R^2} N_n \tag{5.17}$$

where A_{int} is the size of intersection area, which can be evaluated as [77]:

$$A_{int} = R^2 \left[2\, ArcCos\left(\frac{d_r}{2R}\right) - \frac{d_r}{R}\sqrt{1 - \left(\frac{d_r}{2R}\right)^2} \right] \tag{5.18}$$

From Equations (5.16), (5.17) and (5.18), the entering rate can be evaluated as:

$$\lambda_e = \frac{N_n}{\mu}\frac{A_i}{\pi R^2} = \frac{N_n}{\pi \cdot \mu}\left[2\, ArcCos\left(\frac{d_r}{2R}\right) - \frac{d_r}{R}\sqrt{1 - \left(\frac{d_r}{2R}\right)^2} \right]$$

5.5 Validation

In this section, the proposed model is validated by comparing the analytical results obtained from the solution of the proposed SRN model using SPNP [93] with the simulation results obtained using the network simulator NS2 [27].

Two performance metrics have been used to validate the proposed model: path connection availability Ψ, and path failure and repair frequency μ_s. Path connection

availability is the probability that a route exists between a source-destination pair. It can be computed from the proposed SRN model shown in Figure 5.3 using the following equation:

$$\Psi = Pr\big((\#P_{LR1} = 0 \ \& \ldots \& \#P_{LRn} = 0) \,\&\, (\#P_{SR1} = 0 \ \& \ldots \& \#P_{SRn} = 0)\big) \quad (5.19)$$

where $Pr(E)$ is the probability of the event E and $n = N_h - 1$. The path failure and repair frequency (μ_s) is the frequency with which the path failure and repair occur, which is computed as follows [74]:

$$\mu_s = \frac{1}{\frac{1}{\mu_f} + \frac{1}{\mu_r}} = \frac{\mu_f \cdot \mu_r}{\mu_f + \mu_r} \quad (5.20)$$

$$\mu_f = A_v \cdot Rate\big(T_{pfk}\big) \quad (5.21)$$

$$\mu_r = (1 - A_v) \cdot Rate(T_{SR}) \quad (5.22)$$

where μ_f, μ_r, and $Rate(T_x)$ are path failure frequency, path repair frequency and firing rate of the transition T_x respectively. A series of simulation scenarios have been adopted to validate the proposed model and to study the effect on path connection availability of network parameters such as the number of nodes, size of simulated area, transmission range, routing protocol, and packet generation rate.

The settings of the simulation scenarios consist of a network in a square area with the side length L varying from 800 to 1500 m, number of nodes N = 60 or 100, transmission range R = 250 or 200 m, the routing protocol used is AODV or DSR, and packet generation rate (λ) is 10 or 40 Kbps. All nodes move according to random waypoint mobility where the velocity of nodes is chosen uniformly from 5 to 20 m/s and the pause time is set to zero to increase the mobility of nodes. For all mobility scenarios, nodes start to move at the start of the simulation and do not stop until the end of the simulation.

The source-destination pairs are chosen randomly over the network where constant bit rate traffic sources are used. For all scenarios, the number of sources is half of the number of nodes and the packet size is 512 bytes. Identical mobility scenarios and traffic patterns are used across simulation scenarios in order to achieve a fair comparison. The simulation time is set to 1100s and the first 100s are discarded to be sure that the network has reached steady state. All simulation results are obtained with a 95% confidence interval and a maximum relative error of 2%. In Figures 5.5−5.9, solid lines refer to simulation results (labelled Sim), while dashed lines represent SRN model results (labelled Mod).

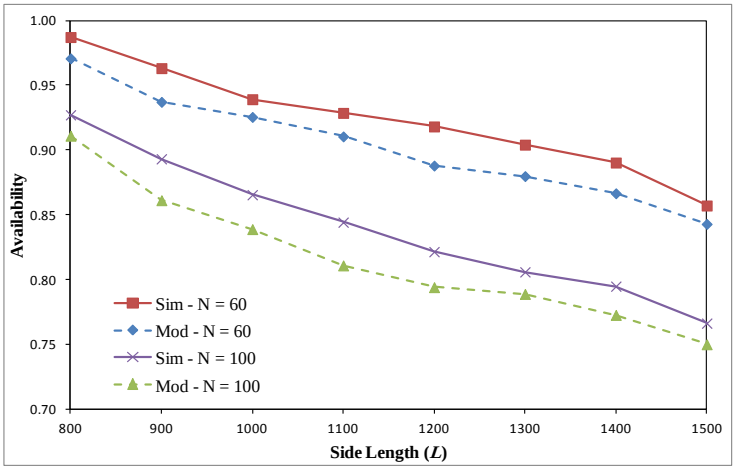

Figure 5.5: Path connection availability versus the side length of the network area, where R = 250 m, N = 60 or 100, and λ = 10 kbps

First, the effects on path availability of increasing the network area size and number of nodes are investigated. The side length of the simulated area is increased from 800 to 1500 m, while the number of nodes is observed for constant values (60 and 100 nodes) where R = 250 m, λ = 10 Kbps, and the routing protocol is AODV. Figure 5.5 shows the numerical results for this scenario.

Figure 5.5 shows interesting results. Although increasing the number of nodes in the network increases the expected number of nodes in the intersection areas (backup routers), which increases path availability, Figure 5.5 shows that the larger the number of nodes the smaller path connection availability. This is because increasing

the number of nodes has another contradictory effect on path connection availability. Increasing the number of nodes increases the number of sources and number of control/management packets, which then increases the interference between neighbour nodes and consequently increases the per hop delay (δ). Increasing the per hop delay increases the time needed to repair path breaks, which decreases path availability. For this network scenario, increasing the per hop delay due to increasing the number of nodes in the network has a greater effect on path availability compared to increasing the number of backup routers, as shown in Figure 5.5.

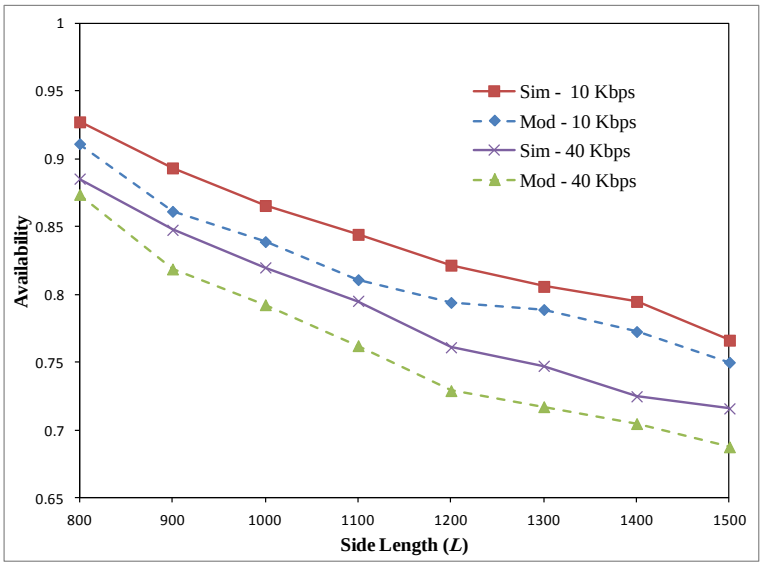

Figure 5.6: Path connection availability versus the side length of the network area, where R = 250 m, N = 100, and λ = 10 or 40 kbps

Also, Figure 5.5 verifies that the network area size has a significant impact on the path availability. For a fixed number of nodes, increasing the network area size may reduce path availability. Although increasing the network area size reduces the node density and interference between nodes, which reduces the per hop delay and increases path availability, it also increases the average number of hops of the paths (as explained in Chapter 4), which thus reduces path availability due to increasing end-to-end delay, path repairing time and the probability of path breaks.

To analyse the impact of the packet generation rate on path availability, two data transmission rates are considered: 10 and 40 Kbps where N = 100, R = 250 m, with AODV used as a routing protocol, and the side length of the network area varying from 800 to 1500 m. Figure 5.6 shows the numerical results for this scenario. Figure 5.6 shows that the larger the packet generation rate, the smaller the path availability. This is because increasing the packet generation rate increases contention and interference between neighbour nodes, which increases the per hop delay and the time needed for path repair.

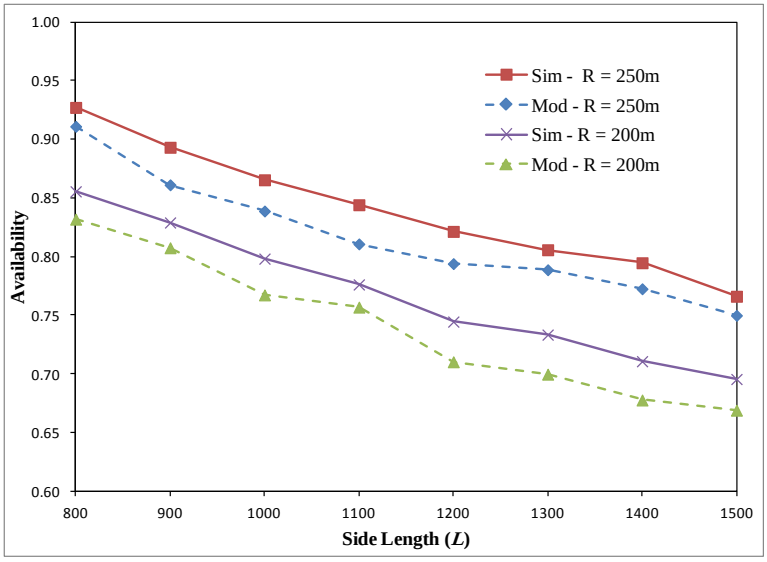

Figure 5.7: Path connection availability versus the side length of the network area, where R = 250 or 200 m, N = 100, and λ = 10 kbps.

Figure 5.7 shows that path connection availability is affected by decreasing the transmission range from 250 to 200 m; in the case of N = 100, λ =10 Kbps, where AODV is the routing protocol, and the side length of the network area varies from 800 to 1500 m. Decreasing the transmission range has two contradictory effects on path availability. The first is to increase path availability due to reducing the interference between nodes. The second is decreasing path availability because of an increase in the average number of hops of the paths, which increases the end-to-end delay and path break probability. For this network scenario, as is clear from Figure 5.7, compared to reducing interference due to decreasing transmission range,

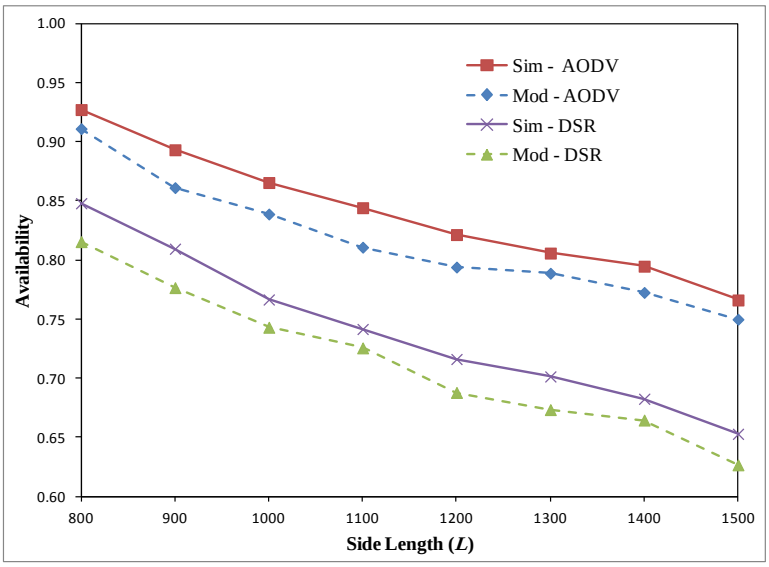

Figure 5.8: Path connection availability versus the side length of the network area, where R = 250 m, N = 100, λ = 10 Kbps and routing protocol is AODV or DSR

increasing the path length has a greater effect on path connection availability. So, path connection availability drops when the transmission range decreases.

Figure 5.8 shows the effect of using DSR as a routing protocol instead of AODV, where the other network parameters are N = 100, R =250 m, and λ =10 Kbps. It is clear that using AODV as a routing protocol provides better path availability than in the case of using DSR. This is because, for high mobility scenarios (mobile nodes move with high speed and low pause time), DSR has a larger end-to-end delay than AODV, which increases the path repairing time. This can be attributed to the aggressive caching strategy used by DSR. Before starting a new route discovery, DSR tries to use all cached routes. With high mobility, the route changes rapidly, which makes all cached routes invalid. Thus, route discovery is delayed until all cached routes fail, which decreases path availability.

The path failure and repair frequency versus the side length of the network area are shown in Figure 5.9 for N = 100, R = 250 m, λ =10 Kbps, and with AODV as the routing protocol. For this network scenario, it is clear that the greater the network area size, the greater the path failure and repair frequency because of the increase in

the number of hops required to reach the destination which increases the probability of path breaks. As shown in Figures 5.5−5.9, the analytical results are close to the simulation results.

In order to solve the proposed model analytically, the time interval of link failure, entering the intersection area and path recovery are approximated to be exponentially distributed. In addition, the approximate value for the number of neighbour nodes, computed using the method introduced in [99], is used to compute the model parameters. Therefore, the simulation results have an additional overhead compared to the analytical results. For all scenarios, computation time of simulations is in the order of hours, whereas the analytical results take a few seconds.

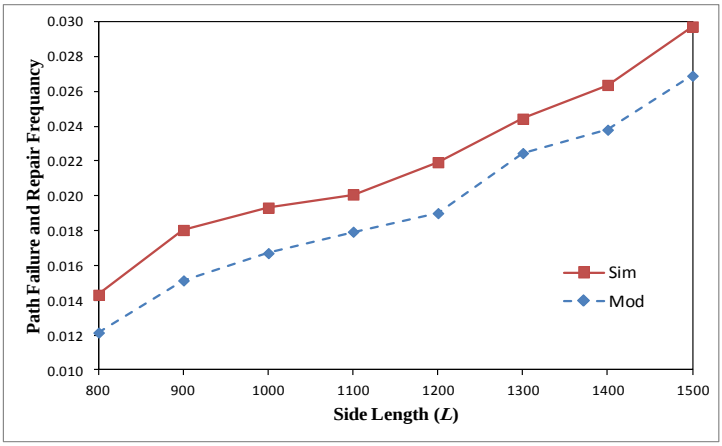

Figure 5.9: Path failure and repair frequency versus the side length of the network area, where R = 250 m, N = 100, and λ = 10 Kbs.

5.6 Summary

In this chapter, a closed form solution is presented for the analytical analysis of paths in multi-hop ad hoc networks with random waypoint mobility. An SRN model is proposed to study path connection availability and path failure and repairing frequency in multi-hop paths. Analytical expressions for leaving time and entering rate for the intersection area, which are the model parameters, are derived. The proposed model is validated by extensive simulations. Compared to simulation results

obtained using NS2, the analytical results are accurate. The impacts of different network parameters on path connection availability, such as the number of nodes, packet generation rate, network size, transmission range and routing protocol, are investigated.

Chapter 6

Performance Modelling of Multi-hop Ad Hoc Networks

6.1 Introduction

To develop an analytical model for multi-hop ad hoc networks based on the IEEE 802.11 DCF MAC protocol, the network model and assumptions explained in Section 1.4.3 are adopted. As explained in Chapter 1, to construct a scalable analytical model for multi-hop ad hoc networks, and to avoid having to model each node in the network, the similarities between nodes in the network are once again exploited. Therefore, in order to do this, a single hop communication between any two nodes in the network under the average workload computed for all possible instances of network topologies is modelled. The proposed single hop communication model captures the average effects of the random access behaviour of each node, the buffer overflow probabilities at each node, interference induced from neighbour and hidden nodes, and frequent path failure and redirection (repairing) due to the random mobility of nodes. Using the single hop communication model, the average node utilisation, which is the ratio of the throughput to the packet generation and forward rate, as well as the packet delay per hop are derived, which are then used to compute the throughput and delay per path, as explained in Section 6.2.

There are many interacting parameters, mechanisms, and phenomena involved in any single hop communication in multi-hop ad hoc networks based on the IEEE 802.11 MAC protocol. Therefore, to model a single hop communication, and to break up the complexity of modelling and avoid the state explosion problem, a framework is proposed which is structured in several models, as explained in Chapter 1. The proposed framework consists of one mathematical model (the network parameters model) and three SRN models (the path analysis, data link layer, and network layer models). These models are solved iteratively, as explained in Section 6.6, in order to compute the average node utilisation and delay per hop which are then used to compute the throughput and delay per path.

In Chapter 3, an SRN model for the IEEE 802.11 DCF MAC protocol in single hop ad hoc networks in the presence of hidden nodes has been presented. This chapter extends this model to capture the behaviour of the IEEE 802.11 DCF MAC protocol in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model. This represents the data link layer model in the proposed framework. The network layer model, which models actions in the network layer, is also presented in this chapter.

For end-to-end connection in a multi-hop ad hoc network, the packets are routed via N_h hops through neighbour nodes if the destination is not in the transmission range of the source. These intermediate nodes are used as connection relays in forwarding packets to their destinations. The average number of routed (or forwarded) packets per node per unit time (λ_r) is a significant parameter in the network layer model, and thus an expression for it is derived in this chapter. In addition, Section 6.3 introduces the second part of the network parameters model, which is used to compute the average number of hidden, carrier sensing and interfering nodes. Moreover, this chapter presents the analytical procedure that shows the sequence in which the proposed models are solved to compute the average throughput and delay per path.

Furthermore, the proposed models are validated by extensive simulations using the network simulator NS2.

Firstly, the path traffic load is analysed in order to derive expressions for the average packet forward rate per node and throughput per path. Then, expressions for the expected number of interfering and hidden nodes are derived. Next, the data link layer and network layer models for multi-hop ad hoc networks are presented. After that, the analytical procedure used to solve the proposed models is introduced. Finally, the results obtained from the analytical models and simulations are discussed.

6.2 Analysis of Paths Traffic Load

When the destination is out of the transmission range of the source node, other nodes are used in MANETs as relays to forward packets to their destinations. So, the route or path is the sequence of mobile nodes which data packets pass through in order to reach the intended destination node from a given source node. The traffic load of any path in a multi-hop ad hoc network depends on the packet generation rate (λ) and packet forward rate (the number of received packets to be forwarded per unit time) per node. The packet generation rate is a network parameter; whereas the packet forward rate depends on λ and other network parameters such as the network area size, the number of nodes, and the mobility model. This section analyses the path traffic load to derive expressions for the average packet forward rate per node (λ_r) and the throughput per path using average node utilisation (α).

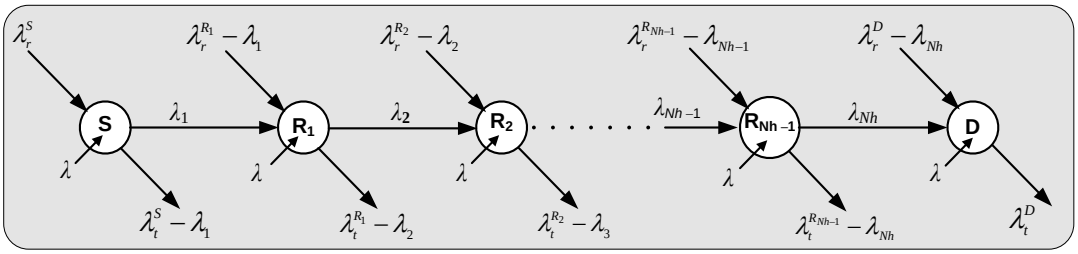

Figure 6.1: A network communication path

Figure 6.1 shows a communication path with N_h hops between the source node *S* and destination node *D*, where λ_t^x is the number of packets that are successfully transmitted by any node *x* per unit time, λ_r^x is the packet forward rate for a node *x*, and λ_1, λ_2, …., and λ_{Nh} are the number of packets sent by the source *S* and received by the nodes R_1, R_2, …., and *D*, respectively. For a node *k*, the node utilisation (α_k) is the ratio of the number of packets that are successfully transmitted to the number of packets that are received or generated by the node *k* to be forwarded to other neighbour nodes per unit time. The utilisation for a node *k* can be expressed as follows:

$$\alpha_k = \frac{\lambda_t^k}{\lambda_r^k + \lambda} \tag{6.1}$$

Therefore, for node *S*, shown in Figure 6.1, the node utilisation is computed as follows:

$$\alpha_S = \frac{\lambda_t^S}{\lambda_r^S + \lambda} = \frac{\lambda_t^S - \lambda_1}{\lambda_r^S} = \frac{\lambda_1}{\lambda}$$

So, the average number of packets that node R_1 successfully receives from the source node *S* per unit time to be sent to the destination node *D* is:

$$\lambda_1 = \alpha_S \cdot \lambda \tag{6.2}$$

For node R_1, the node utilisation is computed as follows:

$$\alpha_{R_1} = \frac{\lambda_t^{R_1}}{\lambda_r^{R_1} + \lambda} = \frac{\lambda_t^{R_1} - \lambda_2}{\left(\lambda_r^{R_1} - \lambda_1\right) + \lambda} = \frac{\lambda_2}{\lambda_1}$$

So, the number of packets sent by *S* and successfully received by R_2 is:

$$\lambda_2 = \alpha_{R_1} \cdot \lambda_1 \tag{6.3}$$

By substitution from Equation (6.2) into Equation (6.3), we get:

$$\lambda_2 = (\alpha_S \cdot \alpha_{R_1}) \cdot \lambda$$

In the same way, it can be deduced that the number of packets that a node R_k receives from the source *S* per unit time is:

$$\lambda_k = (\alpha_S \cdot \alpha_{R_1} \cdot \alpha_{R_2} \cdot \dots \dots \cdot \alpha_{R_k}) \cdot \lambda$$

Consequently, the number of packets received by the destination node *D* per unit time, which represents the throughput per path, is:

$$\text{Throughput} = \lambda_{N_h} = \left(\alpha_S \cdot \alpha_{R_1} \cdot \alpha_{R_2} \cdot \dots \dots \cdot \alpha_{R_{Nh-1}}\right) \cdot \lambda \tag{6.4}$$

To simplify the analysis, the node utilisation of any node in the path is considered to be equal to the average node utilisation of all nodes in the network (α). Therefore, Equation (6.4) can be simplified to:

$$\text{Throughput} = \lambda_{N_h} = \alpha^{N_h} \lambda \tag{6.5}$$

If δ is the average delay per hop, the end-to-end delay of the path is computed as follows:

$$\text{End-to-End Delay} = N_h \cdot \delta \tag{6.6}$$

The average node utilisation (α) and delay per hop (δ) are used to compute the average throughput and end-to-end delay of the path between any source-destination pair in the network, and are calculated using the proposed framework for a single hop communication, as explained in Section 6.6.

The number of data packets sent per unit time by a source node *S* and forwarded (routed) by the intermediate nodes (routers) between the source *S* and destination *D* in the path can be computed as follows:

$$\lambda_x = \lambda_1 + \lambda_2 + \dots . + \lambda_{N_h - 1}$$

From Equation (6.5), λ_x can be computed as:

$$\lambda_x = (\alpha + \alpha^2 + \cdots + \alpha^{N_h - 1}) \cdot \lambda \tag{6.7}$$

To compute the traffic load in the path, control packets, such as RREQ (Route Request), RREP (Route Reply), and RERR (Route Error), which are used by reactive routing protocols such as AODV and DSR, should be taken into account. The routing protocol in the source node broadcasts RREQ to search for the shortest route to the destination. When the destination receives an RREQ message, it sends back an RREP message to the source. When the link between any two nodes in the path between the source and destination is broken, an RERR message is sent back to the source to show the route breakage. As explained in [105], the number of control packets of the reactive routing protocols sent per unit time (λ_c) in the network can be computed using the following equation:

$$\lambda_c = \mu_s \cdot N_s \cdot (N + N_h + N_e) \tag{6.8}$$

where μ_s, N_s, and N_e are the route discovery (failure and repairing) frequency, the number of sources in the network, and the average length of the path from any broken link in the path to the source node respectively. N_e is considered to be half of the average number of hops between any source-destination pair, and μ_s is computed using the path analysis model introduced in Chapter 5.

The average number of routed packets (either data or control packets) per unit time for all nodes in the network (λ_{rt}) is:

$$\lambda_{rt} = N_s \cdot \lambda_x + \lambda_c$$

Therefore, the average number of routed packets per node per unit time (λ_r) is computed as follows:

$$\lambda_r = \frac{\lambda_{rt}}{N} = \frac{N_s \cdot \lambda_x + \lambda_c}{N} \tag{6.9}$$

By substitution from Equations (6.7) and (6.8) into Equation (6.9), the average number of routed packets per node per unit time is as follows:

$$\lambda_r = \frac{N_s}{N} \cdot [(\alpha + \alpha^2 + \cdots + \alpha^{N_h - 1}) \cdot \lambda + \mu_s \cdot (N + N_h + N_e)] \tag{6.10}$$

6.3 Expected Number of Interfering and Hidden Nodes

As explained in Chapter 1, the hidden area is the area covered by the interference range of the receiver but not covered by the carrier sensing range of the sender. The interfering area is the area of intersection between the carrier sensing range of the sender and the interference range of the receiver. Hidden and interfering nodes are these located in the hidden and interfering areas respectively. The dashed and shaded areas shown in Figure 6.2 illustrate the hidden and interfering areas of a sender *S*. As is clear from the figure, the size of the interfering or hidden areas depends on the carrier sensing range R_{cs} and interference range R_i.

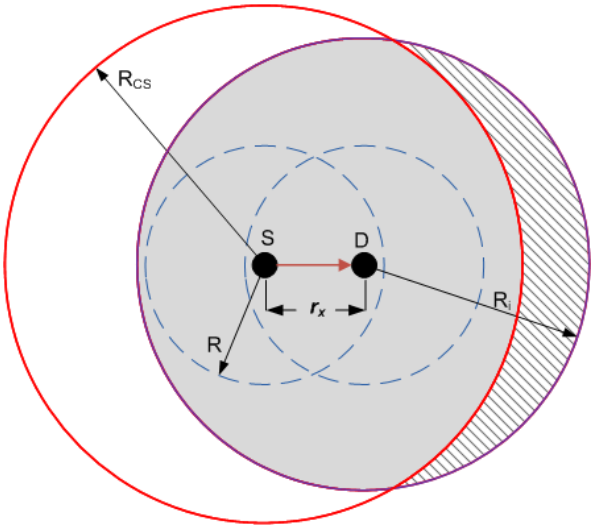

Figure 6.2: Hidden and interfering area

For any sender and receiver, the received signal is considered to be valid and the receiver can recognise it if the Signal to Noise Ratio (SNR) of the received signal is greater than a certain threshold (T_{SNR}). Therefore, to prevent packet collision, the power level of the signals received from the desired sender (P_r) and from any interfering node (P_i) at the receiver must satisfy the following equation:

$$\frac{P_r}{P_i} \geq T_{SNR}$$

The power of the received signal either from the desired sender or an interfering node depends on the signal path loss model which models the signal attenuation over the distance between the sender and receiver. In an open space environment where the two-ray ground path loss model is adopted, the power of the received signal (P_r) from a sender that is r_x metre away can be computed as:

$$P_r = P_t G_t G_r \frac{h_t^2 h_r^2}{r_x^k}$$

where G_t and G_R are the antenna gains of the sender and receiver, h_t and h_r are the heights of the sender and receiver antennas, and k depends on the speed of signal decay. For the two-ray ground path loss model, k is equal to 4 and T_{SNR} is usually set to 10 [47]. Hence, compared to the transmission range and carrier sensing range, the interference range is not fixed, but depends on the distance between the transmitter and receiver [3]. For the two-ray ground path loss model, a simple method has been introduced in [3] to compute the interference range as follows:

$$R_i(x) = \sqrt[k]{T_{SNR}} \cdot x \tag{6.11}$$

where x is the distance between the desired sender and receiver. From Equation (4.9), the probability density function of the distance between two subsequent nodes in the path $\left(f_{r_x}(x)\right)$ is given by:

$$f_{r_x}(x) = 2N_r \cdot \frac{x^{2N_r-1}}{R^{2N_r}} \tag{6.12}$$

From Equations (6.11) and (6.12), the PDF of the interference range can be derived as:

$$f_i(y) = \frac{2N_r}{\sqrt[k]{T_{SNR}}} \cdot \frac{\left(\frac{y}{\sqrt[k]{T_{SNR}}}\right)^{2N_r-1}}{R^{2N_r}} \tag{6.13}$$

By definition, the average interference range is computed as follows:

$$R_{iA} = \int_0^{\sqrt[k]{T_{SNR}} \cdot R} y \cdot f_i(y)\, dy = \frac{2N_r}{2N_r + 1} \cdot \sqrt[k]{T_{SNR}} \cdot R \tag{6.14}$$

Because the sizes of the interfering and hidden areas depend on the interference range, they also depend on the distance between the sender and receiver. For two circles with radii of R_1 and R_2, and where the distance between the centres of the two circles is r_x, the intersection area between the two circles $A_{int}(R_1, R_2, r_x)$ can be computed as follows:

$$P1(R_1, R_2, r_x) = R_1^2 \cdot Arccos\left(\frac{r_x^2 + R_1^2 - R_2^2}{2\, r_x \cdot R_1}\right)$$

$$P2(R_1, R_2, r_x) = R_2^2 \cdot Arccos\left(\frac{r_x^2 + R_2^2 - R_1^2}{2\, r_x \cdot R_2}\right)$$

$$P3(R_1, R_2, r_x) = \frac{1}{2}\sqrt{(-r_x + R_1 + R_2) \cdot (r_x + R_1 - R_2) \cdot (r_x + R_1 + R_2)}$$

$$A_{int}(R_1, R_2, r_x) = P1(R_1, R_2, r_x) + P2(R_1, R_2, r_x) + P3(R_1, R_2, r_x) \tag{6.15}$$

If the distance between the sender and receiver is *x*, the sizes of the hidden area $A_h(x)$ and interfering area $A_i(x)$ are computed using Equations (6.11) and (6.15) as follows:

$$A_h(x) = \pi \cdot \left(R_i(x)\right)^2 - A_{int}(R_{cs}, R_i(x), x) \tag{6.16}$$

$$A_i(x) = A_{int}(R_{cs}, R_i(x), x) \tag{6.17}$$

Therefore, from Equations (6.12), (6.16), and (6.17), the average size of the hidden area A_h and interfering area A_i are computed using the following equations:

$$A_h = \int_0^R 2N_r \cdot \frac{x^{2N_r - 1}}{R^{2N_r}} \cdot A_h(x)\, dx$$

$$A_i = \int_0^R 2N_r \cdot \frac{x^{2N_r - 1}}{R^{2N_r}} \cdot A_i(x)\, dx$$

The average number of hidden nodes N_H and interfering nodes N_i are numerically computed as:

$$N_H = \frac{A_h}{A_h + A_i} \cdot N_t \tag{6.18}$$

$$N_i = \frac{A_i}{A_h + A_i} \cdot N_t \tag{6.19}$$

where N_t is the average number of nodes in a circle with radius R_{iA}. For the random waypoint mobility model, a mathematical analysis for the number of nodes in a circle with radius R in a network using geometric probability has been introduced in [99]. Thus, N_t is computed using the method introduced in [99]. Also, because the carrier sensing range is fixed, the average number of carrier sensing nodes (N_{cs}) is also computed using the method introduced in [99].

6.4 Data Link Layer Model

In the proposed framework introduced in Section 1.5, data link layer protocols are modelled by the data link layer model. As explained in Chapter 1, the data link layer is divided into two sub-layers which are LLC and MAC. In wireless networks, the packet processing time in the LLC layer is negligible compared to that in the MAC layer [31-41]. Hence, the data link layer model only describes the behaviour of the MAC layer protocols. In Chapter 3, SRN models have been introduced for the IEEE 802.11 DCF MAC protocol for both BA and RTS/CTS methods, in single hop ad hoc networks with hidden nodes where all nodes in the network are stationary (i.e. have no mobility). In this section, these models are extended to model the IEEE 802.11 DCF MAC protocol, for both BA and RTS/CTS methods, in multi-hop hop ad hoc networks where nodes move according to the random waypoint mobility model.

An SRN model describing the behaviour of the IEEE 802.11 DCF MAC protocol in a single hop communication between any two nodes in a multi-hop ad hoc network should capture all the dynamics of the protocol, interaction with the network layer protocols, and interaction between the sender (or receiver) node and the carrier sensing, interfering, and hidden nodes. To meet these requirements and to avoid the state explosion problem, the modelling technique introduced in Chapter 3 has been

adopted. The proposed SRN model for either the BA or RTS/CTS method is divided into two interactive SRN models; the one node detailed model and the abstract model, which depend on lumping and decomposition techniques. The one node detailed model describes all of the detailed activities in one node; whereas the abstract model describes interactions between any node and interfering nodes taking into account the effects of hidden and carrier sensing nodes. The two models are solved iteratively until convergence of the performance measures is reached. The one node detailed and abstract models for both BA and RTS/CTS methods are shown in Figures 6.3−6.6 and are explained below.

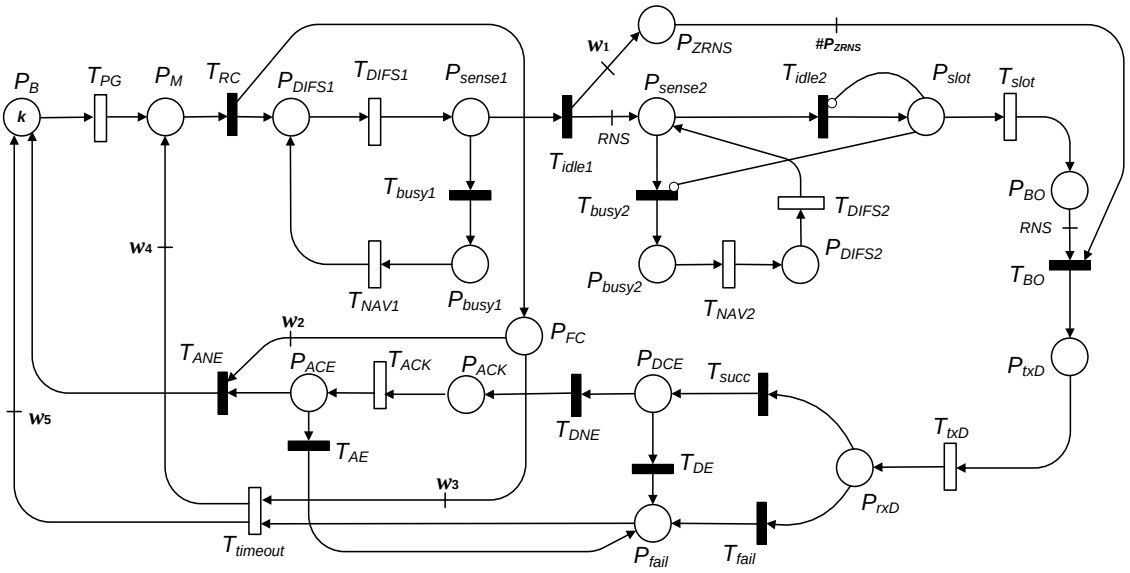

Figure 6.3: One Node detailed model for the BA method

6.4.1 SRN Models of the BA Method

The one node detailed model and the abstract model for the BA method are shown in Figures 6.3 and 6.4. When the structures of these models are compared with those of the one node detailed model and active area abstract model shown in Figures 3.5 and 3.7, which are described in detail in Chapter 3, several differences are apparent which are explained in this section. All corresponding identifiers (places, transitions, or arc weight functions) in these models which have the same names also have the same meaning, function, firing rate, firing probability, or firing guard.

The number of interfering nodes that do not have a packet to transmit is represented by the number of tokens in P_B (N_i) in the abstract model. The transition T_{PG} models the number of packets received from the upper layer per unit time λ_n (the throughput of the network layer). The place P_{txD} represents the start of transmission of the data frame by the physical layer. The transmission of the data frame is represented by the firing of T_{txD} which moves the token to P_{rxD} which models the delivery of the data frame to the receiver. If any interfering nodes start to transmit any data frame at the same slot time, a collision occurs and transmission fails; otherwise, the frame is transmitted successfully. Therefore, the token in P_{rxD} may move to P_{DCE} due to the firing of T_{succ}, representing the success of transmitting the data frame, or it may move to P_{fail} due to the firing of T_{fail}, representing the failure to transmit the frame because of collision. In addition, the data fame will not be correctly received if either of the following two events hold at any system slot time during the receiving of the data frame:

- Any node hidden to the sender starts to transmit a data frame.
- Any node hidden to the sender which has already received a data frame starts to transmit an ACK frame.

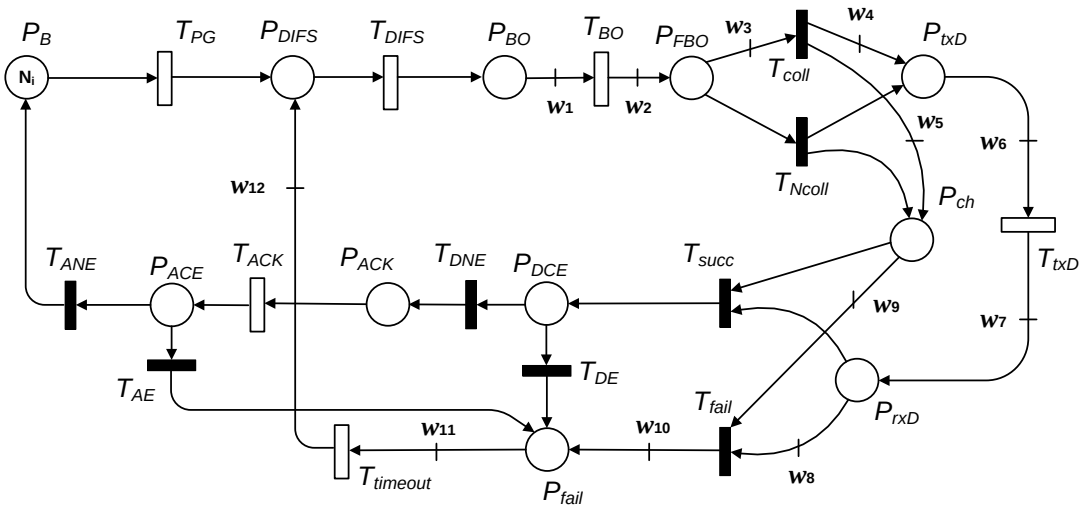

Figure 6.4: The abstract model for the BA method

Moving the token from P_{DCE} to P_{fail} due to the firing of T_{DE} represents failure to receive the data frame due to interference induced by hidden nodes. On the other hand, firing of the transition T_{DNE} deposits the token to P_{ACK} which represents the case when there is no interference from hidden nodes and the receiver has successfully received the data frame.

Once the receiver has successfully received the data frame, it sends the ACK frame after an SIFS interval which is represented by the firing of T_{ACK} that deposits a token in P_{ACE}. The sender does not correctly receive the ACK frame if any nodes hidden to the receiver which are located in the interference range of the sender did not send a data or ACK frame at any system slot time during the reception of the ACK frame. Success and failure to receive the ACK frame are represented by the firing of T_{ANE} and T_{AE}, respectively.

For any node, if the probability of transmitting the data or ACK frame is ρ_{BA}, the firing probabilities of T_{DE} (T_{AE}) and T_{DNE} (T_{ANE}) are (ρ_{BA}) and (1– ρ_{BA}), respectively. The parameter ρ_{BA} is computed from the abstract model shown in Figure 6.4 using the following equations:

$$p_{BA} = 1 - \left(1 - Pr(\#P_{txD} > 0 \text{ OR } \#P_{ACK} > 0\ )\right)^{\frac{1}{N_i}}$$

$$\rho_{BA} = 1 - (1 - p_{BA})^{N_H}$$

In the one node detailed model shown in Figure 6.3, the firing of the transition T_{ANE} flushes the place P_{FC}, which models the resetting of the backoff counter to zero, and deposits a token in P_B which allows a new packet be transmitted. The firing of the transition $T_{timeout}$ removes all tokens from P_{fail}, representing the ACK frame timeout. Depending on the number of tokens in P_{FC}, the transition $T_{timeout}$ may deposit a token in P_B or P_M. If $\#P_{FC}$ is less than MRL, $T_{timeout}$ deposits a token in P_M and does not remove any tokens from P_{FC}; otherwise, it deposits a token to P_B and flushes P_{FC}. This models the dropping of the packet after reaching the maximum retry limit and is managed by the arc weight functions w_3, w_4, and w_5.

For the one node detailed model, the firing probabilities of the transitions T_{idle1} (β_{BA}) and T_{busy1} (1 − β_{BA}) represent the probabilities that the channel is idle and busy respectively, during the sensing of the channel in the DIFS interval, which is computed from the abstract model shown in Figure 6.4 as follows:

$$\beta_{BA} = (1 - p_{BA})^{Ncs}$$

The firing probabilities of the transitions T_{idle2} and T_{busy2} are equal to those of T_{idle1} and T_{busy1} respectively. The firing probabilities of T_{fail} and T_{succ} are (μ_B) and (1 – μ_B) respectively, where μ_B is the probability of failure to transmit the data frame due to interference induced by interfering nodes. μ_B is computed from the abstract model as:

$$\mu_B = Pr(\#P_{ch} > 1)$$

In the abstract model shown in Figure 6.4, transmitting the ACK frame successfully is represented by the firing of T_{ACK} and T_{ANE}, whereas failure to transmit the ACK frame is represented by the firing of T_{AE}. Firing of the transition T_{ANE} returns the token back to the place P_B. The tokens in the place P_{fail} represent failure to receive the data or ACK frame. The ACK frame timeout is modelled by the transition $T_{timeout}$, and its firing moves all tokens from P_{fail} to P_{DIFS}.

6.4.2 SRN Models of the RTS/CTS Method

The one node detailed and abstract models for the RTS/CTS method are shown in Figures 6.5 and 6.6. The differences in structure between these models and those shown in Figures 3.6 and 3.8 and described in detail in Chapter 3 are explained in this section. As the SRN models of the BA method, all corresponding identifiers with the same names have the same meaning, function, firing rate, firing probability, or firing guard.

The transition T_{PG} models the number of packets received from the upper layer per unit time (the throughput of the network layer). The firing of the transition T_{RTS} and moving the token from place P_{txRTS} to place P_{rxRTS} represent transmitting the RTS frame. If a sender *S* transmits the RTS frame to a receiver *D* (see Figure 6.2) without

any errors due to simultaneous transmission from at least one interfering node, T_{succ} fires and deposits a token in P_{REC}; otherwise T_{fail} fires. The RTS frame is received successfully if any of the following events do not occur at any system slot time during the receiving of the RTS frame [32, 33, 39]:

- Any nodes hidden to the sender initiate a transmission by sending an RTS frame to any other node.
- The transmission of the RTS frame is started at any system slot time during an ongoing transmission between any of the nodes hidden to the sender and any other node, where the hidden node sends a data, CTS, or ACK frame.

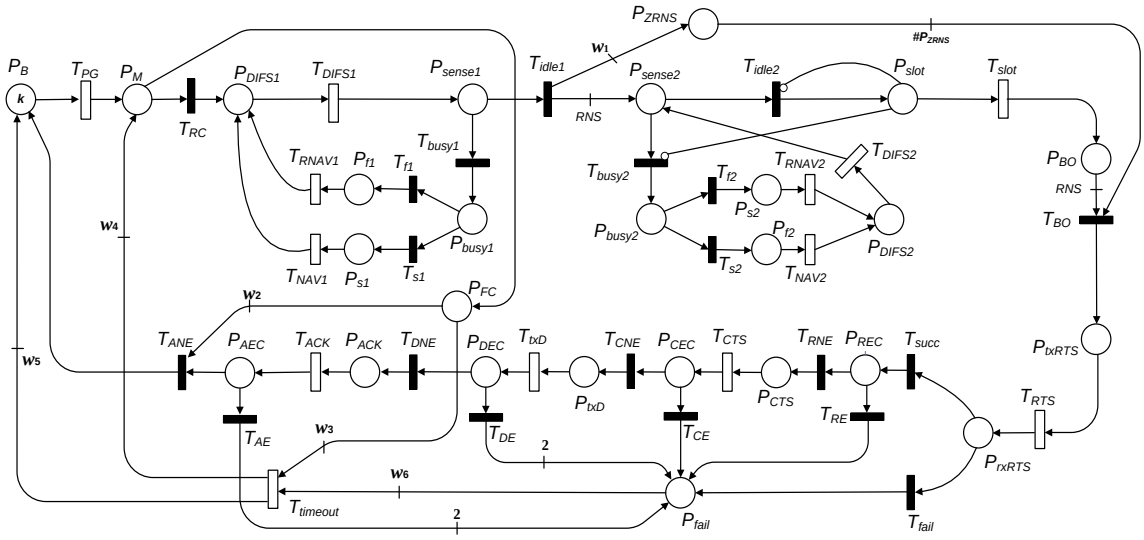

Figure 6.5: One node detailed model for the RTS/CTS method

The successful reception of the RTS frame is represented by the firing of the transition T_{RNE}. Otherwise, the transition T_{RE} fires representing the receiving of the RTS frame with errors. P_{CTS} and T_{CTS} represent the transmission of the CTS frame from the receiver *D* to the sender *S*. The successful reception of the CTS frame by the sender *S* is represented by the firing of the transition T_{CNE} which moves the token from P_{CEC} to P_{txD}. Once *S* has received the CTS frame, it sends the data frame to the receiver *D* after an SIFS period. The successful transmission and reception of the data frame is represented by the firing of the transitions T_{txD} and T_{DNE} which move the

token from P_{txD} to P_{ACK} through P_{DEC}. The receiver D sends the ACK frame to S immediately after receiving the data frame. The firing of the transitions T_{ACK} and T_{ANE} which move the token from P_{ACK} to P_B through P_{AEC} represents the successful transmission and reception of the ACK frame.

The transmission of the CTS and ACK frame from the receiver D to the sender S may fail if any node hidden to D located in the interference range of S transmits a RTS or CTS frame at any slot time during transmission of the CTS or ACK frame [32, 33, 39]. Also, D will not correctly receive the data frame sent by S if any node hidden to S started to transmit a RTS or CTS frame at any slot time during reception of the data frame. Failure to receive the CTS, data, and ACK frames is represented by the firing of the transitions T_{CE}, T_{DE}, and T_{AE}, respectively, which remove the token from P_{CEC}, P_{DEC}, or P_{AEC}. The firing of the transition $T_{timeout}$ removes all tokens from P_{fail} and models the CTS or ACK frame timeout.

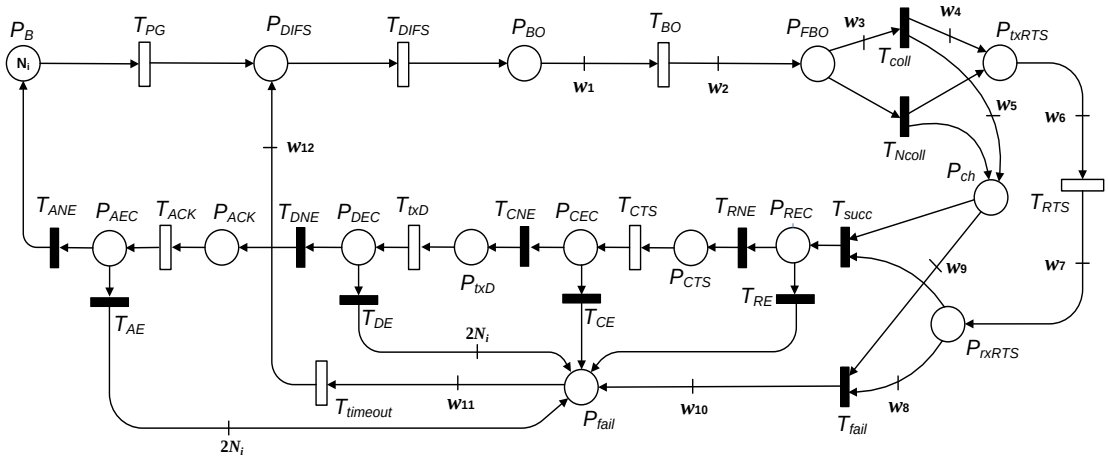

Figure 6.6: The abstract model for the RTS/CTS method

If the transmission probability of any RTS, CTS, data or ACK frame for any node is ρ_R, (ρ_R) and (1– ρ_R) are the firing probabilities of the transitions T_{RE} and T_{RNE}, respectively. The parameter ρ_R is computed from the abstract model as:

$$p_R = 1 - \left(1 - Pr(\#P_{txRTS} > 0 \text{ OR } \#P_{CTS} > 0 \text{ OR } \#P_{txD} > 0 \text{ OR } \#P_{ACK} > 0)\right)^{\frac{1}{N_i}}$$

$$\rho_R = 1 - (1 - p_R)^{N_H}$$

The firing probability of the transitions T_{CE}, T_{DE}, and T_{AE} is ω_R, whereas that of T_{CNE}, T_{DNE}, and T_{ANE} is $(1 - \omega_R)$. The parameter ω_R is computed from the abstract model as follows:

$$p_e = 1 - \left(1 - Pr(\#P_{txRTS} > 0 \text{ OR } \#P_{CTS} > 0\ )\right)^{\frac{1}{N_i}}$$

$$\omega_R = 1 - (1 - p_e)^{N_H}$$

In the one node detailed model shown in Figure 6.5, the firing of the transitions T_{CE}, T_{DE}, and T_{AE} removes the token from P_{CEC}, P_{DEC}, or P_{AEC} respectively and deposits one or 2 tokens to P_{fail}. The firing of the transition $T_{timeout}$ models the CTS and ACK frame timeout. Thus, the average firing time of $T_{timeout}$ depends on the number of tokens in P_{fail}. If $\#P_{fail} = 1$, $T_{timeout}$ represents the CTS frame timeout interval ($Ft(T_{txRTS}$) + $Ft(T_{CTS})$). Otherwise, it represents the ACK frame timeout interval ($Ft(T_{txD}$) + $Ft(T_{ACK})$). The firing probabilities of the conflicted transitions T_{s1} (μ_R) and T_{f1} $(1 - \mu_R)$ represent the probability of success and failure in completing the RTS/CTS handshake respectively, where μ_R is computed from the abstract model shown in Figure 6.6 as:

$$\mu_R = 1 - Pr\left(\#P_{fail} > 0\ \right)$$

The probability that the channel is idle β_R is the firing probability of T_{idle1} and $T_{idle2,}$ computed from the abstract model as follows:

$$\beta_R = (1 - p_R)^{Ncs}$$

The firing probability of the transition T_{fail} (μ_C) is the probability of the failure to transmit the RTS frame due to interference induced by interfering nodes, where μ_C is computed from the abstract model. The firing probability of the transition T_{succ} is $(1 - \mu_C)$.

In the abstract model shown in Figure 6.6, the failure to transmit the RTS or CTS frame is represented by the firing of the transitions T_{RE} and T_{CE} which deposit one token to P_{fail}. However, the firing of T_{DE} or T_{AE}, which deposits $2N_i$ tokens in P_{fail}, represents the failure to transmit the data or ACK frame. The firing of the transition

T_{ANE} returns the token back to the place P_B, and the firing of the transition $T_{timeout}$ models the CTS or ACK frame timeout. Thus, the average firing time of $T_{timeout}$ depends on the number of tokens in P_{fail}. If $\#P_{fail} = 2N_i$, $T_{timeout}$ represents the ACK frame timeout interval; otherwise, it represents the CTS frame timeout interval. If $\#P_{fail} = 2N_i$, the firing of T_{timout} removes all tokens in P_{fail} and deposits one token in P_{DIFS}; otherwise it moves all tokens from P_{fail} to P_{DIFS}. This is managed by the arc weight functions w_{11} and w_{12}.

6.5 Network Layer Model

In MANETs, nodes are free to move and organise themselves arbitrarily. Thus, the network topology may change rapidly and unpredictably. So, the destination node is usually out of the transmission range of the source node. Therefore, the packets reach the destination after a number of hops via intermediate nodes between the source and destination which are used as connection relays. As a result, mobile nodes work as both sources and routers for other mobile nodes in the network.

The main goal of network layer protocols (or routing protocols) is the correct and efficient establishment and maintenance of the route between a pair of nodes in order that messages are sent or forwarded reliably and in a timely manner. In addition, because each node works as a router, the routing protocols maintain information about the routes in the network to be used to forward any received packets. The design of MANET routing protocols is a challenge, because they operate in resource-constrained devices and networks with highly dynamic topologies.

The proposed network layer model is shown in Figure 6.7. It is an SRN model for network layer events in MANETs. The transition T_{GP} models the generation of packets in the transport layer. The firing of the transition T_{GP} deposits a token in the place P_{GP}. The mean firing time of T_{GP} is the mean time of the generation of UDP packets in transport layer. The place P_{Buffer} contains tokens corresponding to free buffer spaces in the current node. The initial number of tokens in P_{Buffer} (N_B) is the total number of free buffer spaces in the node. The firing of the immediate transition

T_B reserves a buffer space for the outgoing packet by removing a token from P_{Buffer} and depositing a token into the place P_s which represents the reception of packets by the network layer.

When a token arrives in the place P_s, there are two possibilities at this point. The first is that the path to the destination is available, and so the transition T_{YPS} fires and thus moves the token from the place P_s to the place P_{BP1}. The firing of the transition T_{Frd1} moves the token from P_{BP1} to P_{MAC}, which represents forwarding the packet from the network layer to the MAC layer. The second possibility is that the path to the destination is not available, and therefore the transition T_{NPS} fires depositing the token to the place P_{BP2}. Once the route has been recovered or re-established, the transition T_{Frd2} fires to forward the packet to the MAC layer.

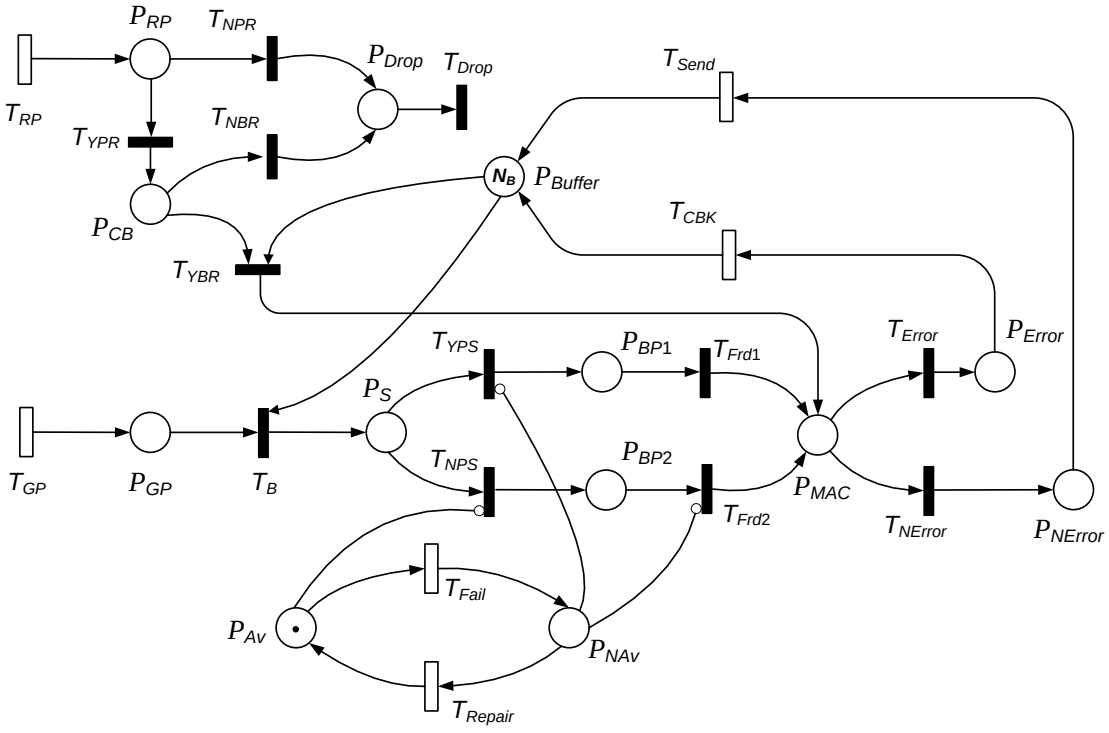

Figure 6.7: Network layer model

The places and transitions P_{MAC}, P_{Error}, P_{NError}, T_{Error}, T_{NError}, T_{CBK}, and T_{send} represent interaction with the data link layer model presented in the last section. The token in the place P_{MAC} shows that the MAC layer received the packet and started to send it. If

the MAC protocol failed to transmit the packet due to packet collision or interference, it drops the packet and sends a CBK (Call Back) error message to the network layer. This is represented by the place P_{Error} and the firing of transitions T_{Error} and T_{CBK}. The firing of the timed transition T_{CBK} represents the completion of error detection and dropping the packet, after which a place in the buffer of the current node is released by returning a token to the place P_{Buffer}. On the other hand, successful transmission and reception of the packet are modelled by the firing of the transition T_{NError} which moves the token from P_{MAC} to P_{NError}, and the firing of the transition T_{Send} that returns the token back to the place P_{Buffer} to represent an increase in the free buffer space by one.

The firing probability of T_{Error} (ε) is the probability of CBK error (packet dropping probability), whereas the firing probability of T_{NError} is (1– ε). The one node detailed model in the data link layer model is used to compute ε using the following equation:

$$\varepsilon = Pr\ (\#P_{FC} = \text{MRL and } \#P_{fail} > 0) \tag{6.20}$$

The average firing times of the timed transitions T_{Send} ($Ft(T_{Send})$) and T_{CBK} ($Ft(T_{CBK})$) are the average time needed to send and drop a packet in the data link layer, which are computed from the abstract model in the data link layer model as follows:

$$Ft(T_{Send}) = \delta_d = \frac{N_i - M(P_B)}{Thr(T_{PG})} \tag{6.21}$$

$$Ft(T_{CBK}) = \frac{N_i - M(P_B)}{Thr(T_{PG})} \cdot \frac{MRL}{n_c}$$

where n_c is the average number of attempts to transmit the MAC frame, which is computed as $n_c = M(P_{FC})$, and δ_d is the packet delay in the data link layer.

Each node in a MANET has a routing table that indicates which is the next hop and the number of hops for each destination in the network. The main function of the routing protocols which work in the network layer is to build and update the routing table. For any packet processed by the network layer, the routing protocol checks all

available paths to the destination and chooses the best one according to defined criteria. Due to node mobility, there are frequent failures in paths between sources and destinations, which has a considerable effect on network performance. The average time to failure of any path between any source and destination depends on the density distribution of nodes, network area size, transmission range and the type of mobility pattern. For any path failure, the routing protocol tries to recover the path to the destination. The average time of path recovery depends on the type of routing protocols used, the density of nodes, mobility patterns, transmission range, and path length, as explained in Chapter 5. The behaviour of path failure and recovery should be captured by the network layer model.

The places P_{Av} and P_{NAv} and transitions T_{Fail} and T_{Repair} model the effect of path failure and repair process. The token in the places P_{Av} and P_{NAv} represent cases that the path between the source and destination is available and unavailable respectively. The timed transitions T_{Fail} and T_{Repair} represent the process of failure and repair of the path between the source and destination respectively. The firing rates of transitions T_{Fail} (μ_f) and T_{Repair} (μ_r) are the average rate of failure and repair of any path between any source-destination pair respectively, which are computed using the path analysis model (using Equations (5.21) and (5.22)) as explained in Chapter 5. The inhibiter arcs from places P_{AV} and P_{NAv} to transitions T_{Frd2}, T_{NPS} and T_{YPS} ensure that, if there is no path to the destination in the routing table (i.e. the path is not available), the packet (or token) will not be forwarded from the routing layer (or place P_s) to the data link layer (or place P_{MAC}).

Any MANET node may work as a source, destination or router. A node may receive packets from neighbour nodes to be forwarded to another node (working as a router) or to absorb them (working as a destination). Thus, the network layer model should capture how a node deals with packets received from neighbour nodes. The firing of the timed transition T_{RP} and depositing a token in the place P_{RP} represent the completion of the reception of a packet from a neighbour node. The firing rate of T_{RP} is the average number of received packets to be forwarded per unit time (λ_r). Section 6.2 derives an expression for λ_r. If the path to the destination of the received packet

sent by a neighbour node is not available, the node drops the packet immediately. This is modelled by the place P_{Drop} and transitions T_{NPR} and T_{Drop}. Otherwise, the node tries to save the packet in the buffer, which is represented by the transition T_{YPR} and place P_{CB}.

The firing of the transition T_{NBR} means that the buffer is full ($\#P_{Buffer} = 0$) and the node is unable to forward the packet, which is dropped. A guard function is assigned to T_{NBR} to disable it when $\#P_{Buffer} > 0$. If the buffer can accommodate a packet ($\#P_{Buffer} > 0$), the packet enters a queue and waits in order to be processed by the MAC protocol. This is represented by the firing of the transition T_{YBR} that moves a token from P_{CB} to P_{MAC}. Transitions T_{GP} and T_{RP} are assigned with guard functions that prevent their firing when the buffer is full. Also, to prevent the forwarding of packets to the MAC layer during an attempt to send a packet, transitions T_{Error} and T_{NError} are assigned with guards that disable them when the transitions T_{send} or T_{CBK} are enabled. If ψ is the average probability that any path in the network is available, the firing probabilities of the transitions T_{YPR} and T_{YPS} are ψ, whereas the firing probabilities of transitions T_{NPR} and T_{NPS} are $(1- \psi)$. The probability of path availability ψ is computed using the path analysis model and Equation (5.19), as explained in Chapter 5.

6.6 Analytical Procedure

As explained in Chapter 1, the proposed framework for the modelling and analysis of multi-hop ad hoc networks consists of a mathematical model (the network parameters model) and three SRN models (the data link layer, path analysis, and network layer models). To compute the required performance indices such as delay and throughput, the three SRN models are solved iteratively using the fixed point iteration technique. First, for any network setting (number of nodes, network size, transmission range, carrier sensing range, etc.), the network parameters model introduced in Chapter 4 and Section 6.3 is used to compute the expected number of hops between any source-destination pair (N_h), the expected number of nodes in the carrier sensing range (N_{cs}),

the expected number of interfering nodes (N_i), and the expected number of hidden nodes (N_H). Then, the data link layer model introduced in Section 6.4 is solved (i.e. generating the Markov chains model and computing the required performance metric) in order to compute the packet dropping probability (ε), packet delay in the data link layer (δ_d) and packet delay per hop (δ) using Equations (6.20), (6.21) and (6.22) respectively. The data link layer model consists of two models: the one node detailed model and abstract model. As explained in Chapter 3, these two models are solved iteratively using the procedure introduced in Section 3.5 until the convergence of any performance metric such as the packet delay in the data link layer.

$$\delta = \frac{N_B - M(P_{Buffer})}{Thr(T_{PG})} \cdot M(P_B) \tag{6.22}$$

Next, the path analysis model is solved to compute the path availability (ψ), the path failure rate (μ_f), and the path repairing rate (μ_r), which are all required in order to solve the network layer model. At the end of the first iteration, using ε and δ_d computed using data link layer model and ψ, μ_f and μ_r computed using the path analysis model, the network layer model can be solved to compute the node utilisation (α) and network layer throughput (λ_n) using the following equations:

$$\alpha = \frac{Thr(T_{GP})}{\lambda} \cdot (1 - \varepsilon) \tag{6.23}$$

$$\lambda_n = Thr(T_{GP}) + Thr(T_{RP}) \tag{6.24}$$

The iterative process continues by solving the models until the convergence of any performance metric such as α or λ_n. The following procedure and Figure 1.6 summarise the iterative process for solving the proposed models to compute the delay per hop and the node utilisation, which are used to compute the end-to-end delay and throughput per path:

Step 1: Parameters N_H, N_{cs}, N_h, and N_i, are computed using the network parameters model. Also, the number of iterations n is initialised to 1 and the

probabilities of events required to solve SRN models for the initial iteration (ρ_R, ω_R and α) are considered to be equal to 0.5.

Step 2: The data link layer model is solved to compute ε, δ and δ_d using the following sub-procedure:

Step 2.1: The number of iterations m in the sub-procedure is initialised to 1 and the initial value of the average size of the backoff window A_s is computed using the following equation:

$$A_s = \frac{\sum_{x=0}^{MRL-1} \left[(CW_{min} + 1) \cdot 2^x - 1\right]}{MRL}$$

Step 2.2: If n = 1, the initial value of the backoff window is used to solve the abstract model; otherwise, the last computed value of the backoff window is used to solve the model to compute the initial value of the packet delay in the data link layer δ_d^0 (or any other performance metric) and the parameters μ_R, μ_C, β_R, ρ_R, and ω_R for the RTS/CTS method (μ_B, β_{BA}, and ρ_{BA} for the BA method).

Step 2.3: The one node detailed model is solved using the last computed values of parameters μ_R, μ_C, β_R, ρ_R, and ω_R to compute ε and the new value for A_s.

Step 2.4: The abstract model is solved to obtain the packet delay in the data link layer δ_d^m and the other parameters μ_R, μ_C, β_R, ρ_R, and ω_R.

Step 2.5: The relative error of the packet delay in data link layer $err(\delta_d)$ is computed as follows:

$$err(\delta_d) = \frac{|\delta_d^m - \delta_d^{m-1}|}{\delta_d^m}$$

Step 2.6: If $err(\delta_d)$ is less than a specified threshold, stop the iteration process; otherwise increase m by one and go to Step 2.3.

Step 3: Using δ (or δ_d in the case of n = 1), the path analysis model is solved to compute ψ, μ_f and μ_r.

Step 4: Using the last computed values for α, ε, and δ_d, the network layer model is solved to compute new values for α and λ_n. Also, any of the performance metrics τ^n, such as throughput per hop, can be computed.

Step 5: If $n = 1$, increase n by one and go to Step 2.

Step 6: The relative error of the performance metric $err(\tau)$ is computed as follows:

$$err(\tau) = \frac{|\tau^n - \tau^{n-1}|}{\tau^n}$$

Step 7: If $err(\tau)$ is less than a specified threshold, stop the iteration process; otherwise, increase n by one and go to Step 2.

At the end of the iterations, the last computed values for node utilisation α and delay per hop δ are used to compute the throughput and end-to-end delay per path using Equations (6.5) and (6.6) respectively. The number of iterations mainly depends on the error threshold. In all validation scenarios introduced in the next section, the error threshold is set to 0.05. In all cases the convergence of the performance metric is achieved after only a few iterations. Compared to the time needed for simulation, the proposed models are solved using the procedure described above for different network settings very quickly as explained in the next section.

6.7 Validation and Results

In this section, the proposed models are validated by conducting extensive comparisons of their results with those of a series of simulation experiments. The simulation results are obtained by using the NS2 simulator [27], whereas the analytical results derived from the proposed models are obtained using SPNP [93].

Two fundamental performance metrics are used to evaluate the proposed SRN models: the goodput and end-to-end delay. Goodput is the number of data bits, not including the protocol overhead and retransmitted bits, which are received correctly at a destination per unit time. It is computed from the network layer model using

Equations (6.5) and (6.23). The end-to-end delay of data packets is the average time that a packet takes from the initiation of its transmission at a source node until delivery to a destination. This includes the delay time caused by the buffering of data packets during route discovery, queuing at the interface queue for transmission at the MAC layer, retransmission delays at the MAC layer, and propagation and transmission delay. Using the data link layer and network layer models, the end-to-end delay is computed using Equations (6.6) and (6.22).

For network simulations with any mobility scenario, goodput is computed by dividing the total number of packets received at all destinations by the simulation time, whereas the end-to-end delay is obtained by summing individual packet delays at all destinations and dividing the sum by the total number of received packets. The average goodput and packet end-to-end delay per source-destination pair are obtained by averaging over the goodput and end-to-end delay for all mobility scenarios.

Table 6.1: The key network simulation parameters

Parameter	**Value**
Number of nodes	60, 80, …, 240
Side length of the network area	600, 800, 1000 m
Packet size	2, 6 kB
Packet generation rate	100, 200, …., 2200 kB
Queue Length	30
Transmission range	150, 250 m
Carrier sensing range	150, 250, 350, 450 m
Routing protocol	AODV
Pause time	0 sec
Maximum speed of nodes	20 m/s
Antenna type	Omni-directional
Propagation path loss model	Two-ray ground
Simulation time	1100 sec

For all simulation scenarios, nodes move according to the random waypoint mobility model, and their velocity is chosen uniformly from 1 to 20 m/s and the pause time is set to zero. For all mobility scenarios, nodes start to move at the start of the simulation and do not stop until the simulation ends. Source-destination pairs are chosen randomly over the network where constant bit rate traffic sources are used. The number of sources is equal to the number of nodes, where the destinations are randomly chosen. Identical mobility scenarios and traffic patterns are used across simulation scenarios in order to achieve a fair comparison. The simulation time is set to 1100s and the first 100s are discarded in order to be sure that the network has reached the steady state. All simulation results are obtained with 95% confidence interval and relative error less than 5%.

To validate the proposed models, many network simulation scenarios were conducted. The settings of the simulation scenarios consist of a network in a square area with a side length L, where the number of nodes varies from 60 to 240, the packet generation rate varies from 100 to 2200 kb/s, transmission range R equals 150 or 250 m. The key simulation parameters for all scenarios are summarised in Table 6.1. Also, all simulation and analytical results have been obtained assuming the same values of MAC and physical layer parameters shown in Table 3.9. The simulation and analytical results are shown in Figures 6.8−6.18, where solid lines refer to simulation results (labeled *Sim*), while dashed lines represent the results of the SRN models (labeled *Mod*).

The first scenario is based on varying the packet generation rate in each source node from 100 to 2200 kb/s, where the number of nodes N = 60, the size of the network area is 600x600 m^2, and the transmission range is 150 m. To investigate the effect of increasing the carrier sensing range and packet size on the performance of the network, R_{cs} is set to 150, 250, or 350 m and the packet size is set to 2 or 6 kB. For this scenario, Figures 6.8 and 6.9 show the average goodput per source-destination pair versus the increasing values of packet generation rate for the BA and RTS/CTS methods respectively.

As is clear from Figures 6.8 and 6.9, in the case of light load conditions (small packet generation rates) the greater the packet generation rate the greater the goodput. However, in heavy load conditions, increasing the packet generation rate does not have much effect on goodput. This is because, in the later conditions when every node has a packet to send at all times, the contention to access the channel increases, which increases the probability of packet collision, interference between nodes, and buffer overflow. Thus, the number of packet losses increases, so that make any further increase in packet generation rate has no significant effect on goodput. Also, Figures 6.8 and 6.9 show the effect of increasing the carrier sensing range and packet size on the average goodput per source-destination pair under various channel traffic loads.

Increasing the carrier sensing range decreases the size of the hidden area and the number of hidden nodes, which consequently decreases the probability of packet collision. However, the greater the carrier sensing range, the greater the size of the interference area and the number of interfering nodes, which subsequently increases the probability of packet collision and contention between nodes, thus decreasing channel availability. Therefore, from Figures 6.8 and 6.9 it can be observed that a larger carrier sensing range results in a smaller goodput for both the BA and RTS/CTS schemes.

Figures 6.8 and 6.9 show the relationship between the network goodput and packet size for the BA and RTS/CTS methods. Although increasing packet size increases the probability of packet collision as a result of hidden nodes due to the increase in transmission time, the number of data packets sent per unit time is also reduced, which thus leads to reduced contention between nodes, exponential backoff time, and the probability of packet collision as a result of interfering nodes. In addition, although the number of packets received per unit time is smaller when the packet size is larger, the number of bits received per unit time is larger. Thus, as is clear from Figures 6.8 and 6.9, larger packet size improves the performance of networks with different carrier sensing ranges in both the BA and RTS/CTS schemes.

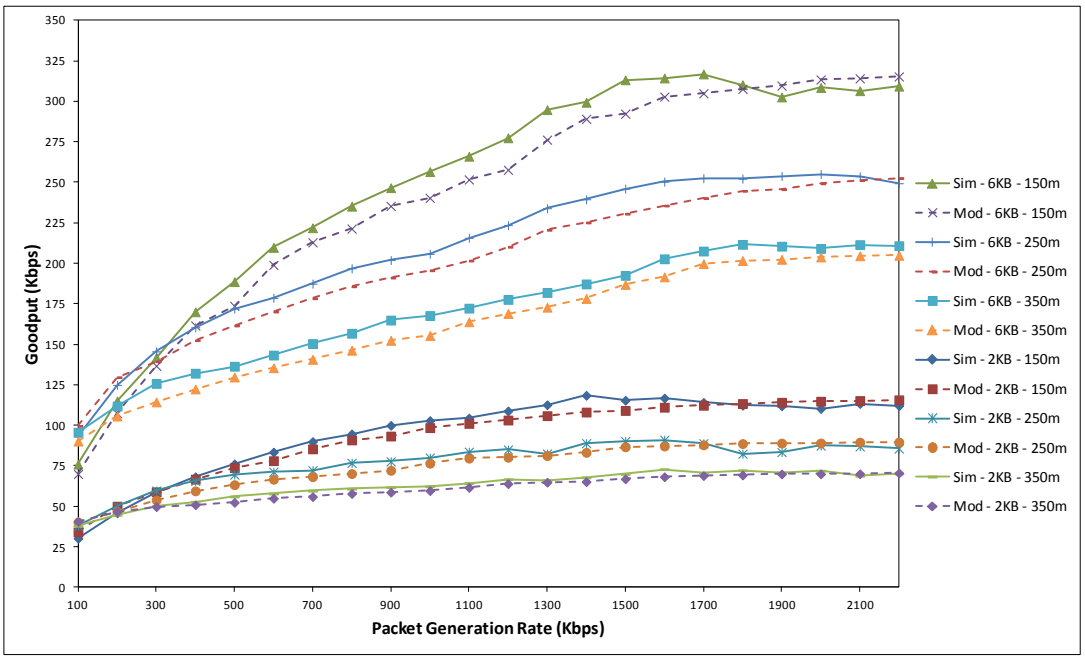

Figure 6.8: Goodput versus packet generation rate for the BA method, in the case of packet size = 2 or 6 kB, R_{cs} = 150, 250 or 350 m, L = 600 m, and R = 150 m

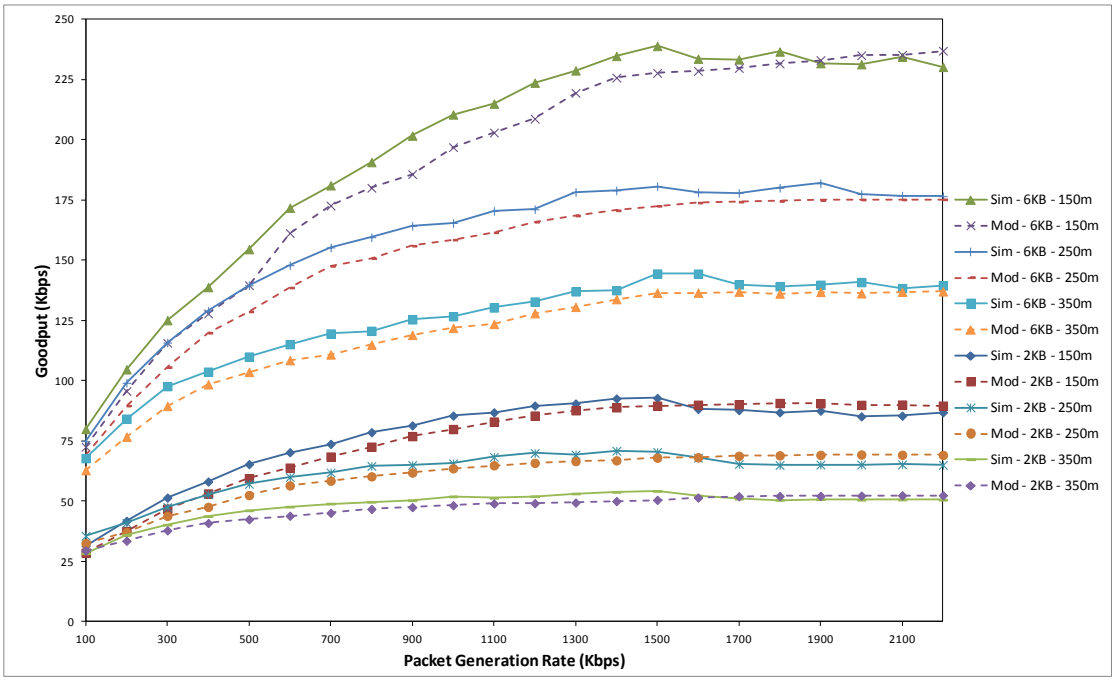

Figure 6.9: Goodput versus packet generation rate for the RTS/CTS method, in the case of packet size = 2 kB or 6 kB, R_{cs} = 150, 250 or 350 m, L = 600 m, and R = 150 m

As shown in Figures 6.8 and 6.9, with a light traffic load, increasing the packet size or carrier sensing range has no significant effect on the performance of the network. This is because when the network load is very low, most packet arrivals can be serviced successfully. In addition, it is to be noted that decreasing the carrier sensing range has more effect on the network goodput when packets are larger but overall traffic is the same. This is because the smaller packets increase interference and contention between nodes that makes the goodput saturate fast with increasing the traffic load.

The results in Figures 6.10 and 6.11 allow a comparison of the BA and RTS/CTS methods according to the packet size and carrier sensing range, showing goodput versus the packet generation rate. In Figure 6.10 the packet size is 2 or 6 kB and R_{cs} = 150 m, and in Figure 6.11 the packet size is 6 kB and R_{cs} = 150 or 350 m. The comparison reveals that in multi-hop ad hoc networks, as opposed to single hop ad hoc networks, the BA method outperforms the RTS/CTS method especially in conditions of heavy load and large packet size.

In the case of a packet size of 6 kB, increasing the carrier sensing range from 150 to 350 m decreases the saturated goodput by about 32.1% and 40.2% for the BA and RTS/CTS methods respectively. In addition, the levels of saturated goodput for the BA method are about 32.4% and 51.4% higher than for the RTS/CTS method in cases where R_{cs} is 150 and 350 m respectively. This is because of the blocking problem which arises in the RTS/CTS mechanism [106] which occurs because any node receiving an RTS or CTS frame defers its transmission. This leads all neighbour nodes of the sender and receiver to be unable to transmit until the sender finishes transmitting the data packet and receiving the ACK frame.

In Figures 6.12 and 6.13, the effect of increasing the size of the network area on network performance is investigated for the BA and RTS/CTS methods under various traffic loads, where N = 60, packet size = 2 kB, R_{cs} = 150, 250 or 350 m, and the side length of the network area is 600 or 1000 m. It can be observed that, with lighter loads, decreasing the size of the network area has no significant effect on the network

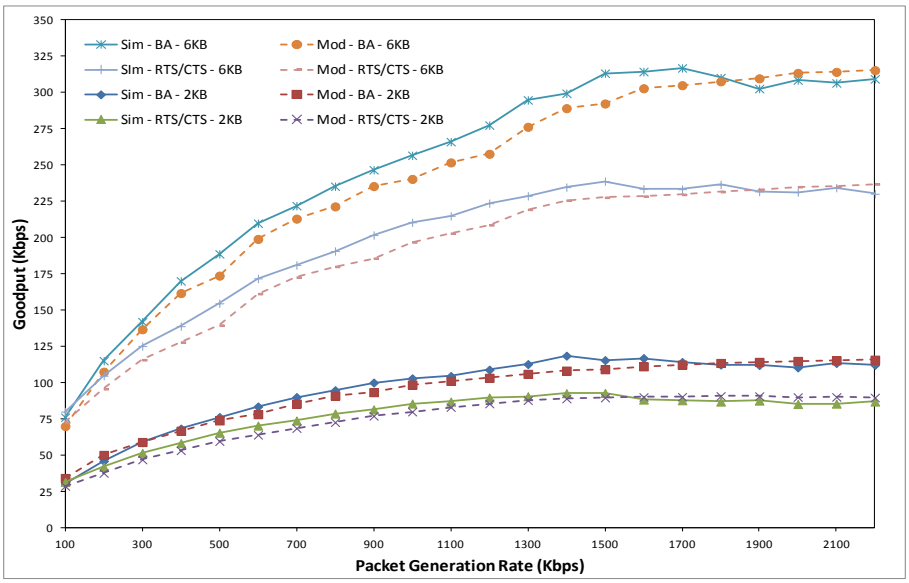

Figure 6.10: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of packet size = 2 or 6 kB, R_{cs} = 150 m, L = 600 m, and R = 150 m

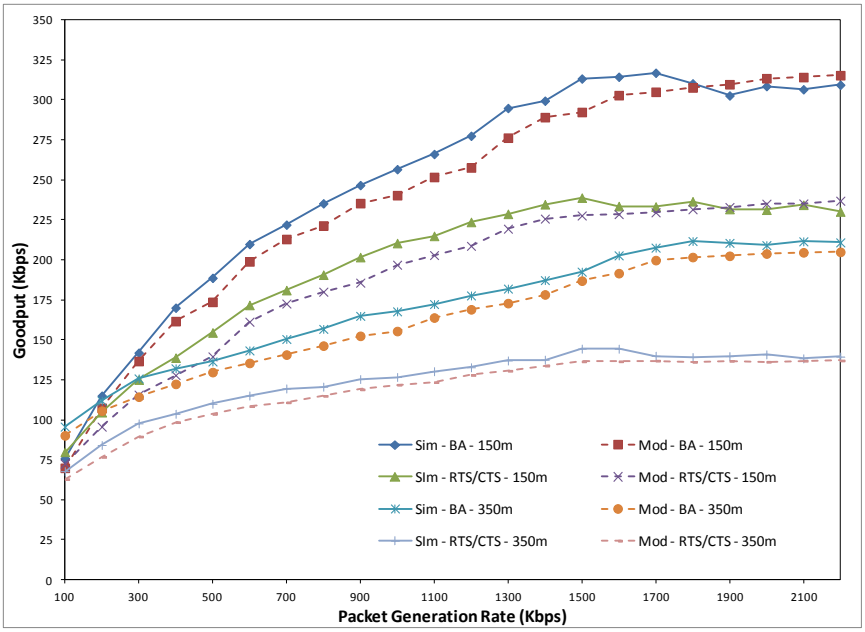

Figure 6.11: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of packet size = 6 kB, R_{cs} = 150 or 350 m, L = 600 m, and R = 150 m

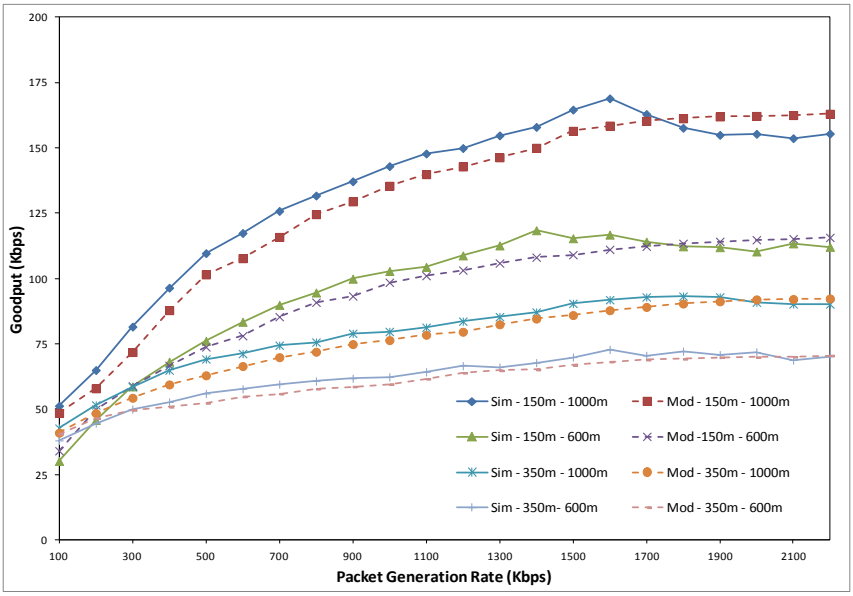

Figure 6.12: Goodput versus packet generation rate for the BA method, in the case of packet size = 2 kB, R_{cs} = 150 or 350 m, L = 600 or 1000 m, and R = 150 m

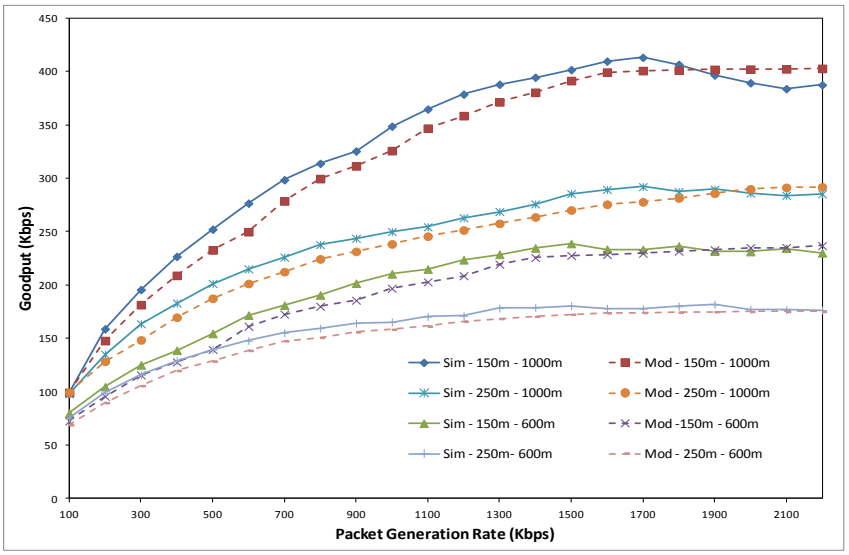

Figure 6.13: Goodput versus packet generation rate for the RTS/CTS method, in the case of packet size = 6kB, R_{cs} = 150m or 250, L = 600m or 1000m, and R = 150m

goodput. However, for higher load conditions, the performance of the network degrades with decreasing network area size. Decreasing network area size has two contradictory effects on network performance. On the one hand, it decreases the path length (the number of hops between the source and destination) which may improve performance. Nevertheless, on the other hand, it increases node density and hence interference and contention between nodes. The probability of packet collision as a result of either hidden nodes or interfering nodes correspondingly increases, which worsens network performance. In this case, increasing the interference induced by hidden and interfering nodes has a greater effect on network performance than decreasing the path length.

To compare the BA and RTS/CTS methods with respect to the network area size, Figure 6.14 shows the relationship between goodput and traffic load when the side length of the network area decreases from 1000 to 600 m, where N = 60, packet size = 2 kB, and Rcs = 150 m. The results demonstrate that the performance of the BA scheme is better than that of the RTS/CTS scheme especially with a small network area (with high node density) because the effect of the blocking problem in the RTS/CTS scheme increases with node density. The network goodput of the BA method is higher than that of the RTS method by 30.2% and 10.8% respectively when the side length of the network area is 600 and 1000 m. So, it is recommended that the BA method should be used in multi-hop ad hoc networks with high node density.

To investigate the dependency of the network goodput on transmission range, Figures 6.15 and 6.16 show the impact of increasing it from 150 to 250 m under various traffic load conditions for the BA and RTS/CTS methods, where the other network parameters are N = 60, R_{cs} = 250 or 350 m, packet size = 2 kB, and L = 800 m. Figures 6.15 and 6.16 show interesting results. For both BA and RTS/CTS schemes, when R_{cs} = 250 m, increasing the transmission range from 150 to 250 m leads to the performance of the network deteriorating, whereas in the case of R_{cs} = 350 m performance is enhanced. On the one hand, increasing the transmission range decreases the number of hops between sources and destinations, and thereby increases the path availability which has a considerable effect on increasing the

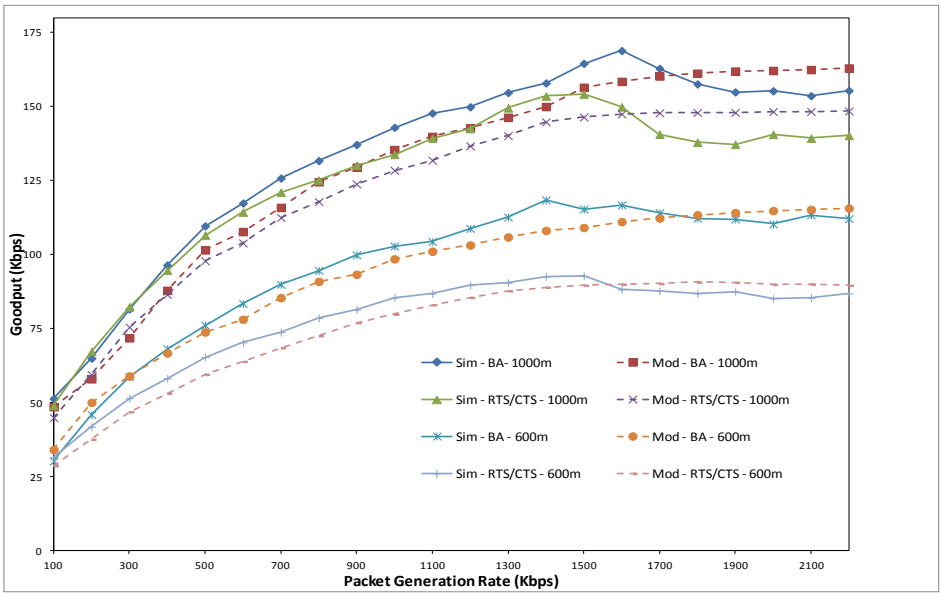

Figure 6.14: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of packet size = 2 kB, R_{cs} = 150 m, L = 600 or 1000 m, and R = 150 m

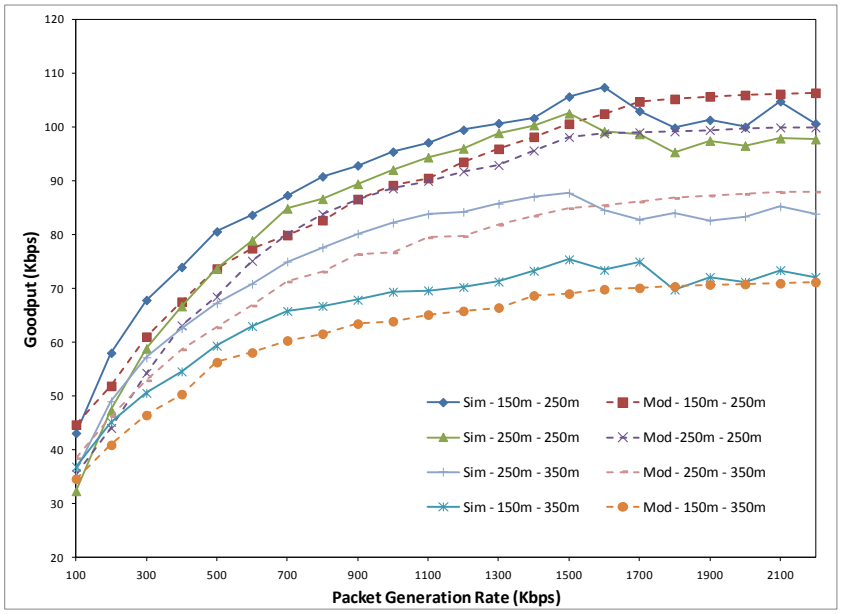

Figure 6.15: Goodput versus packet generation rate for the BA method, in the case of packet size = 2 kB, R_{cs} = 250 or 350 m, L = 800 m, and R = 150 or 250 m

network goodput. However, the interference range is increased as well. This increases the size of the interference and hidden areas and consequently increases the number of interfering and hidden nodes, thus reducing the network goodput. Therefore, because of these contradictory effects on the network goodput, increasing the transmission range does not usually enhance the performance of the network, and this depends on other network parameters such as the carrier sensing range and node density.

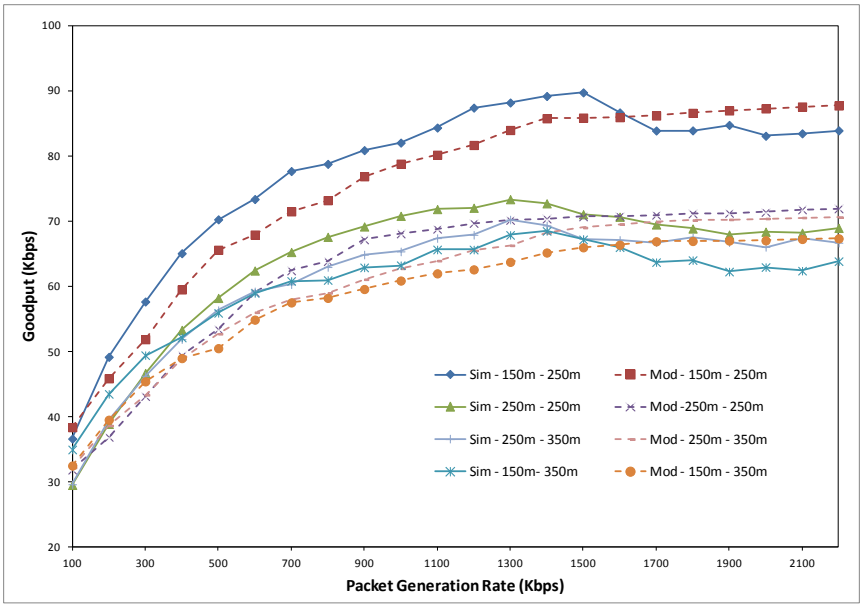

Figure 6.16: Goodput versus packet generation rate for the RTS/CTS method, in the case of packet size = 2 kB, R_{cs} = 250 or 350 m, L = 800 m, and R = 150 or 250 m

Figure 6.17 shows a performance comparison of the BA and RTS/CRS methods with different values of transmission range (150 and 250m), where other network parameter are set to N = 60, R_{cs} = 250 m, packet size = 2 kB, and L = 800 m. The results reveal that, with either a small or large transmission range, the performance of the BA method is much better than that of the RTS/CTS method. In cases where R = 150 or 250, the network goodput of the BA method is greater than that of the RTS/CTS method by 20.2% and 43.2% respectively. This is because increasing the

transmission range increases the effect of the node blocking problem in the RTS/CTS method which reduces goodput of the network.

To investigate the influence of the number of nodes on the end-to-end delay, Figure 6.18 shows the effect on the end-to-end delay of increasing the number of nodes in the network from 80 to 240 for the BA method, where packet size = 2 kB, R_{cs} = 250 or 450 m, L = 1200 m, packet generation rate = 1000 kB/s, and R = 250 m. It can be seen that, for a small number of nodes (less than 180), the greater the number of nodes the greater the end-to-end delay, due to an increased probability of collision and contention between nodes which thus increase the random exponential backoff time which increases the end-to-end delay. However, for larger numbers of nodes, the end-to-end delay only increases slightly when the number of nodes increases. This is because the system starts to become saturated and unable to serve any more packets. The overestimation of the end-to-end delay, as shown in Figure 6.18, is due to the overestimation of the expected number of hops in paths in the network computed using the network parameters model.

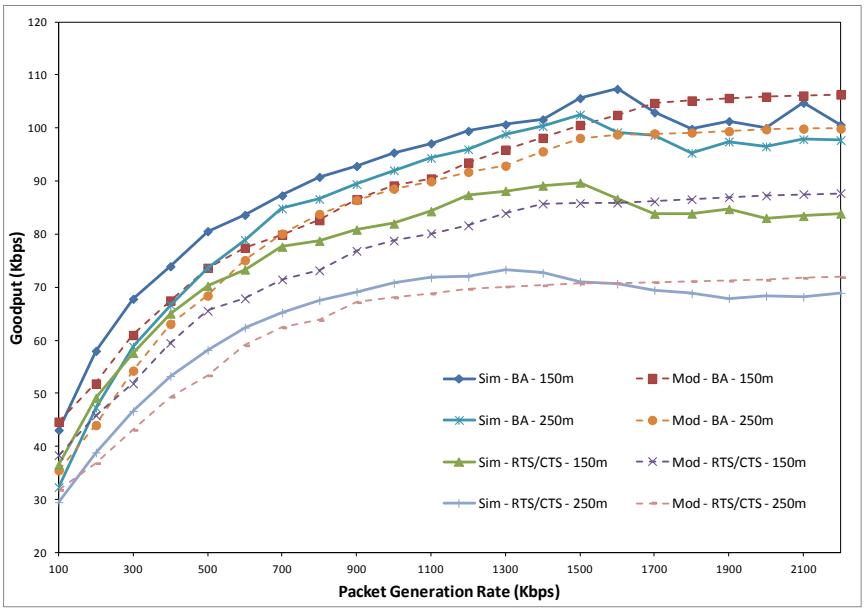

Figure 6.17: Goodput versus packet generation rate for the BA and RTS/CTS methods, in the case of packet size = 2 kB, R_{cs} = 250 m, L = 800 m, and R = 150 or 250 m

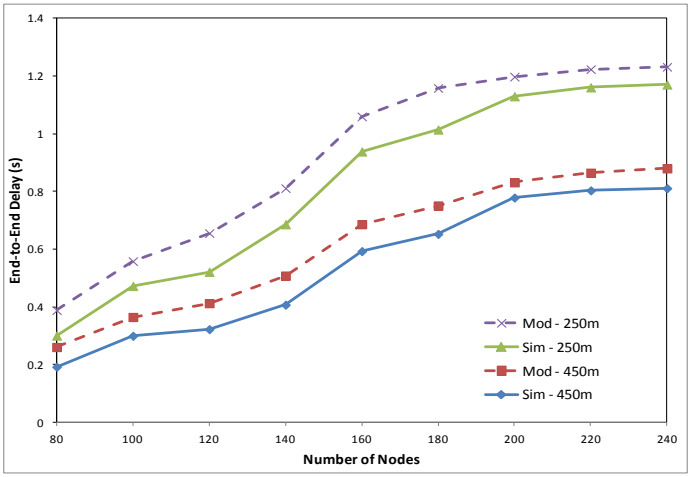

Figure 6.18: End-to-end delay versus number of nodes for the BA method, in the case of packet size = 2 kB, R_{cs} = 250 or 450 m, L = 1200 m, and R = 250 m

The number of states of the abstract and path analysis models depends on the node density, carrier sensing range and interference range, whereas that of the network layer and one node detailed models depends on buffer size and *MRL* respectively. The solution time needed to generate the Markov chains model and compute the required performance metric for these models depends on their state space size and the specifications of the machine used. The maximum numbers of states of the one node detailed model are 1173 and 5905 for the BA and RTS/CTS methods respectively. For the BA method, for N = 240, R_{cs} = 250 m, R = 250 m, L =1200 m, and N_B = 30, the numbers of states of the abstract, path analysis, and network layer models are 1122, 735, and 1022 respectively.

Table 6.2 compares the computation time needed for the analytical analysis and simulation for different values of N for the BA method, where packet size = 2 kB, R_{cs} = 250 m, L = 1200 m, packet generation rate = 1000 kB/s, and R = 250 m. For the simulation and analytical analysis, a desktop workstation was used which was equipped with a 2.6 GHz (Intel Q9400 Core 2 Quad) processor, 4 GB of RAM and Ubuntu Linux version 8.10. Table 6.2 illustrates the scalability of the proposed analytical model compared to the simulation model. It is clear that the time needed for the analytical analysis of the proposed models is very small, even with a large

number of nodes. On the other hand, the simulation time increases exponentially with the number of nodes. When the number of nodes N = 240, the computation time for the simulation analysis is 232 hours and 12 minutes (9 days and 16 hours). For numbers of nodes greater than 240, network simulation becomes so computationally expensive or unfeasible.

Table 6.2: The time needed for analytical analysis and simulation of the network for different number of nodes

N	Mod. (s)	Sim. (HH:MM)
80	7	3:39
100	7	6:50
120	10	13:20
140	11	28:39
160	13	49:09
180	14	77:19
200	14	113:49
220	15	172:41
240	18	232:12

As shown in Figures 6.8−6.18, the analytical results agree closely with those of the simulations. The difference between analytical and simulation results is due to the following approximations:

(1) The time intervals of some events in the data link layer, network layer, and path analysis models have been approximated so as to be exponentially distributed in order to be able to analytically solve the proposed SRN models.

(2) The approximate method introduced in [99] has been used to compute the number of neighbour nodes which is used to derive the number of hidden, carrier sensing, and interfering nodes.

(3) The number of interfering nodes, which is used to solve the proposed data link layer model, must be rounded to the nearest integer.

(4) The average number of hops, which is computed using the method introduced in Chapter 5, usually overestimates the actual value which underestimates the throughput and overestimates the end-to-end delay.

6.8 Summary

In this chapter, the second part of the network parameters model is introduced, which is used to compute the average number of hidden, carrier sensing and interfering nodes. Then, the model introduced in Chapter 3 is extended to model the IEEE 802.11 DCF MAC protocol for both BA and RTS/CTS methods in multi-hop ad hoc networks with the random waypoint mobility model. Next, the network layer model is described. After that, the analytical procedure which shows the sequence in which the proposed models are solved, is presented. At the end, the proposed framework and models are validated using the network simulator NS2.

The effects of various network factors such as communication range, density of nodes, random access behaviour, interference range, carrier sensing range, and traffic load on the performance of multi-hop ad hoc networks, have been analysed in terms of the end-to-end delay and throughput. The results show a close match between the analytical and simulation results. The computation time needed to solve the proposed analytical models is negligible compared to that required for the simulations. The computation time of the network simulation also increases exponentially with the number of nodes in the network. With a large number of nodes, network simulation is computationally expensive and ultimately infeasible.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis has presented a novel analytical framework developed using stochastic reward nets and mathematical modelling techniques for the modelling and analysis of multi-hop ad hoc networks, based on the IEEE 802.11 DCF MAC protocol, where mobile nodes move according to the random waypoint mobility model. The proposed framework has been used to analyse the performance of multi-hop ad hoc networks as a function of network parameters such as the transmission range, carrier sensing range, interference range, number of nodes, network area size, packet size, and packet generation rate.

To break up the complexity, the proposed framework has been organized into several models, based on the ideas of decomposition and fixed point iteration of stochastic reward nets. The proposed framework consists of a mathematical model (called the network parameters model) and four SRN models (the path analysis, data link layer, network layer, and transport layer models). The framework arranges these models and their interactions in a way similar to the layers of the OSI protocol stack model.

The data link layer model for single hop ad hoc networks has been introduced in Chapter 3. This model represents the behaviour of the IEEE 802.11 DCF MAC protocol for both BA and RTS/CTS methods in single hop ad hoc networks in the presence of hidden nodes. Compared to previous studies that have adopted simplified assumptions to reduce the complexity of proposed models which deviate from the IEEE 802.11 standard, the data link layer model captures most of the features of the IEEE 802.11 DCF MAC protocol. The proposed model has been used to demonstrate the effects of network parameters such as traffic load, packet size, and number of nodes.

The data link layer model introduced in Chapter 3 has been validated through extensive comparisons between analytical and simulation results, which show that the proposed model succeeds in providing an accurate representation of the dynamic behaviour of the IEEE 802.11 DCF MAC protocol under several different settings of the network parameters. The analytical results show that in conditions of light load there is not much difference in the performance of the BA and RTS/CTS methods. Conversely, in conditions of heavy load the performance of RTS/CTS method is much better than that of the BA method. Furthermore, the packet size, number of neighbour nodes, and number of hidden nodes have considerable effects on the performance of single hop ad hoc networks, especially in the case of the BA method under saturated load conditions.

The network parameters model is used to compute the expected number of hops between any source-destination pair and the average numbers of carrier sensing, hidden, and interfering nodes. The first part of the network parameters model has been introduced in Chapter 4 in which an approach called the maximum hop distance is presented for the theoretical analysis of the expected number of hops in mobile ad hoc networks where nodes move according to the random waypoint mobility model in a square area. First, an expression for the expected Euclidean distance between any source-destination pair has been derived using geometric probability. Then, expressions have been derived for the probability density function of the distance between any node and its neighbour nodes, and the expected remaining distance to

the destination. By recursive computing of the remaining distance to the destination, the expected hop count has been computed. The results of the proposed approach illustrate the following:

(1) For the random waypoint mobility model, the expected distance between any two nodes in the network is much lower than is the case with uniformly distributed nodes, especially for large network area sizes. This is because the spatial distribution of nodes moving according to the RWPMM is non-uniform, due to the increased probability that a node will be located at the centre of the network area rather than near the borders.

(2) For a given transmission range *R*, increasing the node density increases the maximum forward distance and consequently decreases the expected hop count. The RWPMM significantly increases the average number of neighbour nodes compared to uniformly distributed nodes. Therefore, the expected number of hops for nodes moving according to RWPMM is lower than when nodes are uniformly distributed.

(3) Compared to other methods in the literature, the accuracy of the proposed approach is much better.

To analyse the dynamic nature of paths in multi-hop ad hoc networks where nodes move according to the random waypoint mobility model, a path analysis model has been proposed in Chapter 5. This is a stochastic reward net model which has been used to investigate the instability of paths due to the mobility of nodes. The proposed model presents a closed-form solution for the analytical analysis of paths in multi-hop ad hoc networks in terms of three measures: path connection availability, and the average rates of failure and repair. To solve the proposed SRN model, expressions have been derived for the expected distance between 2-hop-apart nodes, the average time a node needs to pass the intersection area, and the frequency with which nodes in the network enter into an intersection area.

Using the path analysis model, the impact on path connection availability has been investigated of different network parameters, such as the number of nodes, packet

generation rate, network size, transmission range and routing protocol. The results presented in Chapter 5 lead to the following conclusions:

(1) The larger the number of nodes or data transmission rate, the smaller the path connection availability because of increased interference between neighbour nodes which increases end-to-end delay and route recovery delay.

(2) Increasing the network size or decreasing the transmission range may decrease path connection availability. This is because of increases in the end-to-end delay and path break probability due to the increasing number of intersection areas and numbers of hops in paths.

(3) The routing protocol has a significant effect on path connection availability. For example, with high mobility patterns, the DSR protocol decreases path connection availability compared to AODV.

In Chapter 6, the data link layer model introduced in Chapter 3 has been extended to model the IEEE 802.11 DCF MAC protocol for both BA and RTS/CTS methods in multi-hop hop ad hoc networks with the random waypoint mobility model. In addition, the second part of the network parameters model, which is used to compute the average numbers of hidden, carrier sensing and interfering nodes, has been introduced. Moreover, the network layer model has been described. Because the proposed framework is based on the idea of decomposition and fixed point iteration of stochastic reward nets, the proposed SRN models are solved iteratively to compute the required performance indices. The analytical procedure used for the iterative process of solving the proposed models has been presented in Chapter 6. The proposed framework is used to derive the average node utilisation and delay per hop which are then used to compute the throughput (or goodput) and end-to-end delay per path, as explained in Section 6.2.

In addition, Chapter 6 has validated the proposed models using extensive simulations. For various network settings, the results show a close match between the analytical results and those obtained from network simulation using NS2. The computation time needed to solve the proposed analytical models is negligible compared to that

required for the simulations. Moreover, the computation time of network simulation increases exponentially with the number of nodes in the network. With a large number of nodes, the network simulation is very computationally expensive and ultimately infeasible.

The effects of various network factors, such as transmission range, number of nodes, network area size, random access behaviour, interference range, carrier sensing range, packet size, and traffic load, on the performance of the multi-hop ad hoc networks have been analysed in terms of the end-to-end delay and throughput. From the analytical and simulation results, the following conclusions are drawn:

(1) In multi-hop ad hoc networks, as opposed to single hop ad hoc networks, the BA method outperforms the RTS/CTS method especially in conditions of heavy load, high node density, large packet size, and large carrier sensing or transmission ranges. This is because of the blocking problem which arises in the RTS/CTS mechanism.

(2) With light load conditions, changing network parameters such as packet size, carrier sensing range, transmission range, and network area size has no significant effect on network performance because the network load is very low, and so most packet arrivals can be serviced successfully.

(3) For both the BA and RTS/CTS schemes, the performance of multi-hop ad hoc networks deteriorates with increasing the carrier sensing range. This is because of increasing number of interfering nodes which increases the probability of packet collision and contention between nodes, and decreases the channel availability.

(4) For both BA and RTS/CTS methods, increasing packet size enhances the performance of multi-hop ad hoc networks because the probability of packet collision and contention between nodes are reduced, and the numbers of bits received per unit time are increased.

(5) Decreasing the size of network area has two contradictory effects on network performance. Although reducing it decreases the path length, which improves

network performance, it also increases node density which increases interference and contention between nodes and thus packet collision probability. This causes network performance for either BA or RTS/CTS methods to degrade.

(6) Increasing the transmission range may improve the performance of the network due to decreasing path length. However, it also increases the interference range. The greater the interference range, the greater the interference induced by hidden and interfering nodes, which causes deterioration in network performance. Therefore, for both BA and RTS/CTS methods, due to these two contradictory effects, increasing the transmission range does not usually enhance the performance of the network, although this depends on other network parameters.

7.2 Future Work

As the next step of this work, there is a large scope for further work. Future work can be classified into two categories: addressing the limitations of the proposed models, and extending the proposed framework. These are summarised as follows:

- **Addressing the limitations of the proposed models**

(1) To improve the proposed models, phase-type distributions can be applied for the time delays of non-Markovian (deterministic or nondeterministic) events and actions (transitions) that have been approximated with exponential distributions. However, this will increase the complexity of and computation time required for the analytical analysis of the models.

(2) Packet fragmentation allows packets to be broken into smaller pieces before transmission over the wireless medium. It may help to improve reliability in the presence of interference. The data link layer model proposed in Chapter 6 can be extended to support packet fragmentation and reassembly.

(3) The Gilbert–Elliott model [107] is a simple channel model which is widely used for describing error patterns in transmission channels. This model can be used to extend the data link layer model in order to model errors in wireless channels.

(4) Using the approach introduced in Chapter 4, an expression for the expected number of hops in MANETs with a rectangular network area can be derived.

(5) To analyse the expected number of hops in MANETs with low node density, a new approach can be developed or the MHD approach introduced in Chapter 4 should be extended.

(6) The expected number of hops in MANETs with different mobility models, such as random walk (random direction), free way, and Manhattan, can be analysed using the approach introduced Chapter 4.

(7) The path analysis model can also be extended to investigate path connection availability and path failure and repairing frequency in multi-hop ad hoc networks with different mobility models, such as random walk, free way, and Manhattan.

- **Extending the proposed framework**

(1) We aim to extend our work introduced in [108] to develop the transport layer model which captures the behaviour of the transmission control protocol in mutli-hop ad hoc networks.

(2) The performance of wireless networks is inevitably influenced by intrinsic interference effects of wireless channel, such as pathloss, fading, shadowing, and multipath propagation. These effects are dynamic, random, and relevant to environment. The proposed framework can be extended by adding a physical layer model representing the physical layer in the OSI protocol stack, which would model the wireless channel and capture these interference effects.

(3) The proposed framework only supports AODV as a routing protocol because of its simplicity. To adopt other routing protocols such as DSR, the proposed framework should be extended by developing a model for each routing mechanism. This model might be integrated with the path analysis model to analysis the path connection availability and path failure and repairing frequency.

(4) Using the proposed framework, the optimal frame size, transmission range, carrier sensing range, and node density can be computed according to different network parameters and channel conditions so as to maximise network throughput and minimise end-to-end delay.

References

[1] S. K. Sarkar, T. G. Basavaraju, and C. Puttamadappa, *Ad Hoc Mobile Wireless Networks: Principles, Protocols and Applications*: Auerbach Publications, 2004.

[2] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE standards 802.11, IEEE-SA and industry support, (1997–2007)

[3] K. Xu, M. Gerla, and S. Bae, "Effectiveness of RTS/CTS handshake in IEEE 802.11 based ad hoc networks," *Ad Hoc Networks*, vol. 1, no. 1, pp. 107-123, 2003.

[4] R. Sheikh, M. Singh Chande, and D. Kumar Mishra, "Security issues in MANET: A review," in Proceedings of the International Conference On Wireless And Optical Communications Networks (WOCN), 2010, pp. 1-4.

[5] C. Hakima, and L. Maryline, *Wireless and Mobile Networks Security*: John Wiley and Sons, 2010.

[6] M. Barbeau, and E. Kranakis, *Principles of ad hoc networking*: John Wiley and Sons, 2007.

[7] S. G. Matthew, *802.11 Wireless Networks: The Definitive Guide, Second Edition*: O'Reilly Media, Inc., 2005.

[8] W. Stallings, *Wireless Communications and Networks*: Prentice Hall, 2005.

[9] C. S. R. Murthy, and B. S. Manoj, *Ad Hoc Wireless Networks: Architectures and Protocols*: Prentice Hall, 2004.

[10] E. P. Charles, and B. Pravin, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," in Proceedings of the ACM Conference on Communications Architectures, Protocols and Applications, 1994, pp. 234-244.

[11] P. Jacquet, P. Muhlethaler, A. Qayyum, A. Lanouiti, L. Viennot, and T. Clausen, *Optimized Link State Routing Protocol (OLSR)*: RFC Editor, 2003.

[12] B. Bellur, and R. G. Ogier, "A reliable, efficient topology broadcast protocol for dynamic networks," in Proceedings of the Annual Joint Conference of the

IEEE Computer and Communications Societies (INFOCOM), 1999, pp. 178-186.

[13] C. Perkins, E. Belding-Royer, and S. Das, *Ad hoc On-Demand Distance Vector (AODV) Routing*: RFC Editor (RFC 3561), 2003.

[14] D. Johnson, D. Maltz, and Y. C. Hu, *The Dynamic Source Routing for mobile ad hoc networks*: RFC Editor, 2004.

[15] K. Young-Bae, and H. V. Nitin, "Location-aided routing (LAR) in mobile ad hoc networks," *Wireles Networks*, vol. 6, no. 4, pp. 307-321, 2000.

[16] V. Park, and S. Corson, *Temporally-Ordered Routing Algorithm (TORA)*: RFC Editor, 2001.

[17] W. R. Stevens, *TCP/IP Illustrated*: Addison-Wesley, 1994.

[18] D. Comer, *Internetworking with TCP/IP: principles, protocols, and architectures*: Prentice Hall, 2000.

[19] A. Al Hanbali, E. Altman, and P. Nain, "A survey of TCP over ad hoc networks," *IEEE Communications Surveys & Tutorials*, vol. 7, no. 3, pp. 22-36, 2005.

[20] H. Gavin, and V. Nitin, "Analysis of TCP performance over mobile ad hoc networks," *Wireless Networks*, vol. 8, no. 2/3, pp. 275-288, 2002.

[21] J. Liu, and S. Singh, "ATCP: TCP for mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications,* vol. 19, no. 7, pp. 1300-1315, 2001.

[22] W. Feng, and Z. Yongguang, "Improving TCP performance over mobile ad-hoc networks with out-of-order detection and response," in the ACM International Symposium on Mobile Ad Hoc Networking, Lausanne, Switzerland, 2002.

[23] K. Dongkyun, C. K. Toh, and C. Yanghee, "TCP-BuS: improving TCP performance in wireless ad hoc networks," in Proceedings of the IEEE International Conference on Communications, 2000, pp. 1707-1713 vol.3.

[24] A. Vaidyanathan, P. Seung-Jong, S. Karthikeyan, and S. Raghupathy, "TCP performance over mobile ad hoc networks: a quantitative study," *Wireless Communications and Mobile Computing*, vol. 4, no. 2, pp. 203-222, 2004.

[25] H. Lim, "TCP performance over multipath routing in mobile ad hoc networks," in Proceedings of the IEEE International Conference on Communications, Anchorage, Alaska, 2003, pp. 1064-068.

[26] M.Tayade, and S. Sharma, “Performance Improvement of Tcp in MANETs,” *international journal of engineering science and technology*, vol. 3, no. 3, pp. 1914-1920, 2011.

[27] "The Network Simulator ns2," available at http://www.isi.edu/nsnam/ns/.

[28] "OPNET Modeler," available at http://www.opnet.com.

[29] X. Zeng, R. Bagrodia, and M. Gerla, “GloMoSim: a library for parallel simulation of large-scale wireless networks,” in Proceedings of the Workshop on Parallel and Distributed Simulation, 1998, pp. 154-161.

[30] R. Jain, *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modelling*: John Wiley and Sons, 1991.

[31] Y. Wang, and J. J. Garcia-Luna-Aceves, “Collision avoidance in multi-hop ad hoc networks,” in Proceedings of the IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems, 2002, pp. 145-154.

[32] H. Jianhua, K. Dritan, M. Alistair, W. Yiming, D. Angel, M. Joe, and F. Zhong, “Performance investigation of IEEE 802.11 MAC in multihop wireless networks,” in ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Montrial, Quebec, Canada, 2005.

[33] F. Alizadeh-Shabdiz, and S. Subramaniam, “Analytical Models for Single-Hop and Multi-Hop Ad Hoc Networks,” *Mobile Networks and Applications*, vol. 11, no. 1, pp. 75-90, 2006.

[34] B. Wang, F. Song, S. Zhang, and H. Zhang, “Throughput modeling analysis of IEEE 802.11 DCF mechanism in multi-hop non-saturated wireless ad-hoc networks,” in Proceedings of the International Conference on Communications, Circuits and Systems, 2008, pp. 383-387.

[35] L. Fangqin, L. Chuang, W. Hao, and U. Peter, “Throughput Analysis of Wireless Multi-hop Chain Networks,” in Proceedings of the IEEE/ACIS International Conference on Computer and Information Science, 2009, pp. 834-839.

[36] A. Ahmad Ali, G. Fayez, and C. Lin, “Modeling the throughput and delay in wireless multihop ad hoc networks,” in the IEEE conference on Global telecommunications, Honolulu, Hawaii, USA, 2009.

[37] R. Kumar, M. Misra, and A. K. Sarje, “A Simplified Analytical Model for End-To-End Delay Analysis in MANET,” *IJCA Special Issue on MANETs*, vol. 4, pp. 195--199, 2010.

[38] N. Bisnik, and A. Abouzeid, "Queuing network models for delay analysis of multihop wireless ad hoc networks," *Ad Hoc Networks,* vol. 7, no. 1, pp. 79-97, 2009.

[39] E. Ghadimi, A. Khonsari, A. Diyanat, M. Farmani, and N. Yazdani, "An analytical model of delay in multi-hop wireless ad hoc networks," *Wireless Networks,* vol. 17, no. 7, pp. 1679-1697.

[40] C. Zhang, and M. Zhou, "A stochastic Petri net approach to modeling and analysis of ad hoc network," in Proceedings of the International Conference on Information Technology: Research and Education 2003, pp. 152-156.

[41] C. Lin, J. Chang-jun, F. Yu, and L. Fei, "Performance evaluation of ad hoc networks based on SPN," in Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing, 2005, pp. 816-819.

[42] D. Kartson, G. Balbo, S. Donatelli, G. Franceschinis, and C. Giuseppe, *Modelling with Generalized Stochastic Petri Nets*: John Wiley and Sons, 1994.

[43] K. Trivedi, *Probability and Statistics with Reliability, Queueing, and Computer Science Applications*: Wiley-Interscience, 2001.

[44] G. Ciardo, and K. S. Trivedi, "A decomposition approach for stochastic Petri net models," in Proceedings of the International Workshop on Petri Nets and Performance Models, 1991, pp. 74-83.

[45] G. Ciardo, L. Cherkasova, V. Kotov, and T. Rokicki, "Modeling a scalable high-speed interconnect with stochastic Petri nets," in the International Workshop on Petri Nets and Performance Models, 1995.

[46] C. Bettstetter, G. Resta, and P. Santi, "The node distribution of the random waypoint mobility model for wireless ad hoc networks," *IEEE Transactions on Mobile Computing,* vol. 2, no. 3, pp. 257-269, 2003.

[47] R. Theodore, *Wireless Communications: Principles and Practice*: Prentice Hall, 2001.

[48] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications,* vol. 18, no. 3, pp. 535-547, 2000.

[49] H. Wu, Y. Peng, K. Long, S. Cheng, and J. Ma, "Performance of reliable transport protocol over IEEE 802.11 wireless LAN: analysis and enhancement," in Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies, 2002, pp. 599-607.

[50] E. Ziouva, and T. Antonakopoulos, "CSMA/CA performance under high traffic conditions: Throughput and delay analysis," *Computer Communications*, vol. 25, no. 3, pp. 313-321, 2002.

[51] C. H. Foh, and J. W. Tantra, "Comments on IEEE 802.11 saturation throughput analysis with freezing of backoff counters," *IEEE Communications Letters*, vol. 9, no. 2, pp. 130-132, 2005.

[52] F. Eshghi, and A. K. Elhakeem, "Performance analysis of ad hoc wireless LANs for real-time traffic," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 2, pp. 204-215, 2003.

[53] E. Ziouva, and T. Antonakopoulos, "The IEEE 802.11 Distributed Coordination Function in Small-Scale Ad-Hoc Wireless LANs," *International Journal of Wireless Information Networks*, vol. 10, no. 1, pp. 1-15, 2003.

[54] O. Tickoo, and B. Sikdar, "Queueing analysis and delay mitigation in IEEE 802.11 random access MAC based wireless networks," in Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (NFOCOM), 2004, pp. 1404-1413.

[55] C. Xu, J. Gao, Y. Xu, and J. He, "Unified Model for Performance Analysis of IEEE 802.11 Ad Hoc Networks in Unsaturated Conditions," *Transactions On Internet And Information Systems*, vol. 6, no. 2, pp. 683-701, 2012.

[56] T. C. Hou, L. F. Tsao, and H. C. Liu, "Throughput Analysis of the IEEE 802.11 DCF Scheme in Multi-hop Ad Hoc Networks," in Proceedings of the International Conference on Wireless Networks, 2003, pp. 653-659.

[57] H. Ting-Chao, T. Ling-Fan, and L. Hsin-Chiao, "Analyzing the throughput of IEEE 802.11 DCF scheme with hidden nodes," in Proceedings of the IEEE Vehicular Technology Conference (VTC), 2003, pp. 2870-2874.

[58] Y. Wang, and J. J. Garcia-Luna-Aceves, "Modeling of collision avoidance protocols in single-channel multihop wireless networks," *Wireless Networks*, vol. 10, no. 5, pp. 495-506, 2004.

[59] O. Ekici, and A. Yongacoglu, "IEEE 802.11a throughput performance with hidden nodes," *IEEE Communications Letters*, vol. 12, no. 6, pp. 465-467, 2008.

[60] R. German, and A. Heindl, "Performance evaluation of IEEE 802.11 wireless LANs with stochastic Petri nets," in Proceedings of the International Workshop on Petri Nets and Performance Models, 1999, pp. 44-53.

[61] R. Jayaparvathy, S. Anand, S. Dharmaraja, and S. Srikanth, "Performance analysis of IEEE 802.11 DCF with stochastic reward nets," *International Journal of Communication Systems*, vol. 20, no. 3, pp. 273-296, 2007.

[62] C. Zhao, and B. H. Wendi, "Flooding strategy for target discovery in wireless networks," *Wireless Networks,* vol. 11, no. 5, pp. 607-618, 2005.

[63] L. Jinyang, B. Charles, S. J. D. C. Douglas, L. Hu Imm, and M. Robert, "Capacity of Ad Hoc wireless networks," in the International Conference on Mobile Computing and Networking, Rome, Italy, 2001.

[64] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-delay trade-off in wireless networks," in Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), 2004, pp. 464-475.

[65] K. Jia-Chun, and L. Wanjiun, "Modeling the behavior of flooding on target location discovery in mobile ad hoc networks," in Proceedings of the IEEE International Conference on Communications, 2005, pp. 3015-3019.

[66] S. De, "On hop count and euclidean distance in greedy forwarding in wireless ad hoc networks," *IEEE Communications Letters,* vol. 9, no. 11, pp. 1000-1002, 2005.

[67] K. Jia-Chun, and L. Wanjiun, "Hop Count Distribution of Multihop Paths in Wireless Networks With Arbitrary Node Density: Modeling and Its Applications," *IEEE Transactions on Vehicular Technology,* vol. 56, no. 4, pp. 2321-2331, 2007.

[68] M. H. Shadi, and M. Janise, "Analytical Study of the Expected Number of Hops in Wireless Ad Hoc Network," in the International Conference on Wireless Algorithms, Systems, and Applications, Dallas, Texas, 2008.

[69] I. Gruber, and L. Hui, "Link expiration times in mobile ad hoc networks," in Proceedings of the Annual IEEE Conference on Local Computer Networks 2002, pp. 743-750.

[70] Y. Dan, L. Hui, and I. Gruber, "Path availability in ad hoc network," in Proceedings of the International Conference on Telecommunications, 2003, pp. 383-387 vol.1.

[71] T. Yu-Chee, L. Yueh-Feng, and C. Yu-Chia, "On route lifetime in multihop mobile ad hoc networks," *IEEE Transactions on Mobile Computing,* vol. 2, no. 4, pp. 366-376, 2003.

[72] W. Xianren, H. R. Sadjadpour, and J. J. Garcia-Luna-Aceves, "Link Lifetime as a Function of Node Mobility in MANETs with Restricted Mobility: Modeling and Applications," in Proceedings of the International Symposium on Modeling and Optimization in Mobile, 2007, pp. 1-10.

[73] M. Pascoe-Chalke, J. Gomez, V. Rangel, and M. Lopez-Guerrero, "Route duration modeling for mobile ad-hoc networks," *Wireless Networks*, vol. 16, no. 3, pp. 743-757, 2010.

[74] C. Dongyan, G. Sachin, and S. T. Kishor, "Network survivability performance evaluation: a quantitative approach with applications in wireless ad-hoc networks," in the ACM International Workshop on Modeling analysis and simulation of wireless and mobile systems, Atlanta, Georgia, USA, 2002.

[75] K. John, L. Wei, and K. Demetrios, "A generalized model for network survivability," in Proceedings of the Conference on Diversity in Computing, Atlanta, Georgia, USA, 2003, pp. 47-51.

[76] T. Dimitar, F. Sonja, E. Marija, and G. Aksenti, "Ad hoc networks connection availability modeling," in Proceedings of the ACM International Workshop on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks, Venezia, Italy, 2004, pp. 56 - 60.

[77] K. Georgios, and L. Ruijie, "Connection availability and transient survivability analysis in wireless ad-hoc networks," in the ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, Tenerife, Canary Islands, Spain, 2009.

[78] N. Kitae, H. Ahmed, and C. C. J. Kuo, "TCP over multihop 802.11 networks: issues and performance enhancement," in the ACM International Symposium on Mobile Ad Hoc Networking and Computing, Urbana-Champaign, USA, 2005.

[79] S. Xu, and T. Saadawi, "Does the IEEE 802.11 MAC protocol work well in multihop wireless ad hoc networks?," *IEEE Communications Magazine*, vol. 39, no. 6, pp. 130-137, 2001.

[80] M. Prabakaran, A. Mahasenan, and A. Kumar, "Analysis and enhancement of TCP performance over an IEEE 802.11 multi-hop wireless network: single session case," in Proceedings of the IEEE International Conference on Personal Wireless Communications (ICPWC), 2005, pp. 29-33.

[81] A. A. Kherani, and R. Shorey, "Throughput analysis of TCP in multi-hop wireless networks with IEEE 802.11 MAC," in Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), 2004, pp. 237-242 Vol.1.

[82] S. Gollakota, B. V. Ramana, and C. S. R. Murthy, "Modeling TCP over Ad hoc Wireless Networks using Multi-dimensional Markov Chains," in Proceedings of the International Conference on Broadband Communications, Networks and Systems, 2006, pp. 1-10.

[83] H. Xiao, Y. Zhang, J. Malcolm, B. Christianson, and K. Chua, "Modelling and Analysis of TCP Performance in Wireless Multihop Networks," *Wireless Sensor Network,* vol. 2, no. 7, pp. 493-503, 2010.

[84] G. Vandana, S. Dharmaraja, and G. Mingwei, "Analytical modeling of TCP flow in wireless LANs," *Mathematical and Computer Modelling,* vol. 53, no. 5-6, pp. 684-693, 2011.

[85] A. Norman, "THE ALOHA SYSTEM: another alternative for computer communications," in the Fall Joint Computer Conference, Houston, Texas, 1970.

[86] G. R. Lawrence, "ALOHA packet system with and without slots and capture," *ACM SIGCOMM Computer Communication Review,* vol. 5, no. 2, pp. 28-42, 1975.

[87] L. Kleinrock, and F. Tobagi, "Packet Switching in Radio Channels: Part I--Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics," *IEEE Transactions on Communications,* vol. 23, no. 12, pp. 1400-1416, 1975.

[88] P. Karn, "MACA - a new channel access method for packet radio," in Proceedings of the ARRL Computer Networking Conference, 1990, pp. 134 – 140.

[89] A. D. V. Bharghavan, S. Shenker, and L. Zhang, "MACAW: A media access protocol for wireless LAN's," in Proceedings of the ACM conference on Communications Architectures, Protocols and Applications, 1994, pp. 221–225.

[90] L. F. Chane, and J. J. Garcia-Luna-Aceves, "Floor acquisition multiple access (FAMA) for packet-radio networks," in the conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, Cambridge, Massachusetts, United States, 1995.

[91] Y. Jihwang, and A. Agrawala, "Packet error model for the IEEE 802.11 MAC protocol," in Proceedings on Personal, Indoor and Mobile Radio Communications, 2003, pp. 1722-1726 vol.2.

[92] Z. Mustafa, and A. B. McDonald, "On the performance of ad hoc wireless LANs: a practical queuing theoretic model," *Performance Evaluation,* vol. 63, no. 11, pp. 1127-1156, 2006.

[93] G. Ciardo, J. Muppala, and K. Trivedi, "SPNP: Stochastic Petri Net Package," in Proceedings of the Third International Workshop on Petri Nets and Performance Models, 1989, pp. 142-151.

[94] C. Nicholas, and L. Mingyan, "Revisiting the TTL-based controlled flooding search: optimality and randomization," in the International Conference on Mobile Computing and Networking, Philadelphia, PA, USA, 2004.

[95] L. A. Santalo, *Integral geometry and geometric probability*: Addison-Wesley, Advanced Book Program, 1976.

[96] B. Christian, and W. Christian, "The Spatial Node Distribution of the Random Waypoint Mobility Model," in Proceedings of the Workshop on Mobile Ad hoc Networks (WMAN), 2002, pp. 41-58.

[97] H. Takagi, and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals," *IEEE Transactions on Communications*, vol. 32, no. 3, pp. 246-257, 1984.

[98] E. M. Royer, P. M. Melliar-Smith, and L. E. Moser, "An analysis of the optimum node density for ad hoc mobile networks," in Proceedings of the IEEE International Conference on Communications, 2001, pp. 857-861.

[99] C. Bettstetter, "On the Connectivity of Ad Hoc Networks," *Computer Journal*, vol. 47, no. 4, pp. 432-447, 2004.

[100] W. Ewens, and G. Grant, *Statistical Methods in Bioinformatics: An Introduction*: Springer-Verlag, 2004.

[101] S. M. Mousavi, H. R. Rabiee, M. Moshref, and A. Dabirmoghaddam, "MobiSim: A Framework for Simulation of Mobility Models in Mobile Ad-Hoc Networks," in Proceedings of the IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMOB), 2007, pp. 82-82.

[102] F. E. de Deus, R. S. Puttini, L. F. Molinaro, and J. Kabara, "On survivability of IEEE 802.11 WLAN," in Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, 2006, pp. 8-15.

[103] A. P. Snow, U. Varshney, and A. D. Malloy, "Reliability and survivability of wireless and mobile networks," *IEEE Computer*, vol. 33, no. 7, pp. 49-55, 2000.

[104] L. Kleinrock, *Queuing Systems*: John Wiley and Sons, 1975.

[105] P. P. Pham, and S. Perreau, "Performance analysis of reactive shortest path and multipath routing mechanism with load balance," in Proceedings of the Annual Joint Conference of the IEEE Computer and Communications, 2003, pp. 251-259.

[106] S. Ray, and D. Starobinski, "On False Blocking in RTS/CTS-Based Multihop Wireless Networks," *IEEE Transactions on Vehicular Technology,* vol. 56, no. 2, pp. 849-862, 2007.

[107] J. P. Ebert, and A. Willig, *A Gilbert-Elliot Bit Error Model and the Efficient Use in Packet Level Simulation*, TKN Technical Report TKN-99-002, Technical University Berlin, Telecomm. Networks Group, Berlin, Germany, 1999.

[108] O. Younes, W. Elkilany, and N. Thomas, "SRN Model for Performance Evaluation of TCP Sessions Sharing Bottleneck Links in WAN," in the European Symposium on Modeling and Simulation, Leicester, United Kingdom, 2009.