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:


Conferences and workshops:


- 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, “A Path Connection Availability Model for MANETs with Random Waypoint Mobility,” European Performance Engineering Workshop (EPEW), Munich, Germany, 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 ......................................................................................................................... ix

List of Tables .......................................................................................................................... xiv

List of Abbreviations .............................................................................................................. xvi

List of Symbols ....................................................................................................................... xix

Chapter 1 Introduction .............................................................................................................. 1

1.1 Mobile Ad Hoc Networks ................................................................................................. 1

1.2 Characteristics and Challenges of MANETs ................................................................. 3

1.3 MANETs Protocol Stack ................................................................................................. 6

1.3.1 Physical Layer .............................................................................................................. 7

1.3.2 Data Link Layer .......................................................................................................... 8

1.3.3 Network Layer ........................................................................................................... 9

1.3.4 Transport Layer ........................................................................................................ 10

1.4 Motivations, Objectives and Methodology .................................................................... 11

1.4.1 Motivations ................................................................................................................ 11

1.4.2 Objectives and Methodology ..................................................................................... 14

1.4.3 Network Model and Assumptions ............................................................................. 16

1.5 Proposed Framework ....................................................................................................... 17

1.6 Contributions .................................................................................................................. 22
1.7 Thesis Overview ............................................................................................................. 23

Chapter 2 Related Work ........................................................................................................ 25

2.1 Single Hop Ad Hoc Networks ...................................................................................... 25

2.2 Path Length in MANETs ............................................................................................... 29

2.3 Path Analysis in MANETs ............................................................................................ 30

2.4 Multi-hop Ad Hoc Networks ......................................................................................... 32

2.5 Analysis of TCP in MANETs ......................................................................................... 35

Chapter 3 Performance Analysis of the IEEE 802.11 DCF MAC Protocol ............ 37

3.1 Introduction ...................................................................................................................... 37

3.2 IEEE 802.11 DCF MAC Protocol .................................................................................. 40

3.3 Network model and assumptions ................................................................................... 45

3.4 Model Description .......................................................................................................... 46

3.4.1 One Node Detailed Model for the BA Method ......................................................... 47

3.4.2 One Node Detailed Model for the RTS/CTS Method ............................................. 52

3.4.3 Abstract Model for the BA and RTS/CTS Methods .............................................. 53

3.5 Analytical Procedure ..................................................................................................... 61

3.6 Model Validation ............................................................................................................ 63

3.7 Summary ....................................................................................................................... 69

Chapter 4 Expected Path Length in Mobile Ad Hoc Networks with Random Waypoint Mobility ............................................................................................................... 70
4.1 Introduction........................................................................................................................................... 70

4.2 Euclidean Distance between a Source and Destination Node .......................................................... 72

4.2.1 Expected Distance on One Dimension ............................................................................................ 73

4.2.2 Expected Distance in Two Dimensions ........................................................................................... 74

4.3 Expected Hop Count................................................................................................................................... 77

4.4 Validation................................................................................................................................................. 82

4.5 Summary................................................................................................................................................. 86

Chapter 5 A Path Analysis Model for Mobile Ad Hoc Networks with Random Waypoint Mobility........................................................................................................................................ 87

5.1 Introduction............................................................................................................................................. 87

5.2 Ad Hoc Network Model Description...................................................................................................... 89

5.3 SRN Model Description............................................................................................................................ 92

5.4 Model Parameters.................................................................................................................................. 97

5.4.1 Distance Between Nodes .................................................................................................................... 98

5.4.2 Leaving Time....................................................................................................................................... 98

5.4.3 Entering Rate..................................................................................................................................... 101

5.5 Validation............................................................................................................................................... 103

5.6 Summary............................................................................................................................................... 109

Chapter 6 Performance Modelling of Multi-hop Ad Hoc Networks............................................................. 111

6.1 Introduction............................................................................................................................................ 111
6.2 Analysis of Paths Traffic Load ................................................................. 113
6.3 Expected Number of Interfering and Hidden Nodes ................................. 117
6.4 Data Link Layer Model ........................................................................... 120
  6.4.1 SRN Models of the BA Method .............................................................. 121
  6.4.2 SRN Models of the RTS/CTS Method .................................................... 124
6.5 Network Layer Model ............................................................................ 128
6.6 Analytical Procedure ............................................................................. 132
6.7 Validation and Results ........................................................................... 135
6.8 Summary ................................................................................................. 149
Chapter 7 Conclusion and Future Work .......................................................... 150
  7.1 Conclusion ............................................................................................. 150
  7.2 Future Work ........................................................................................... 155
References ...................................................................................................... 158
List of Figures

Figure 1.1: A mobile ad hoc network ................................................................. 2

Figure 1.2: Hidden node problem ................................................................. 5

Figure 1.3: Exposed node problem ................................................................. 6

Figure 1.4: Protocols stack of wireless networks and corresponding main functions .. 7

Figure 1.5: Single hop communication illustrating interfering and hidden nodes...... 17

Figure 1.6: Proposed framework for modelling MANETs .............................. 19

Figure 3.1: The BA and RTS/CTS methods handshake .................................... 41

Figure 3.2: Block diagram for the operation of the BA and RTS/CTS methods ...... 41

Figure 3.3: Timing diagram for the operation of the RTS/CTS method ............. 42

Figure 3.4: The Network architecture for a single hop ad hoc network with hidden
nodes .............................................................................................................. 45

Figure 3.5: One node detailed model for the BA method ............................... 46

Figure 3.6: One node detailed model for the RTS/CTS method ....................... 52

Figure 3.7: The abstract model for the BA method ......................................... 54

Figure 3.8: The abstract model for RTS/CTS method .................................... 55

Figure 3.9: Goodput versus packet generation rate for the BA and RTS/CTS methods,
in the case of N = 10, Nh = 2, Packet Size = 2 KB, λh =10 and 100 Kbps 64
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, $\lambda_h = 100$ Kbps ................................................................. 65

Figure 3.11: Packet delay versus number of nodes for both BA and RTS/CTS methods, in the case of $\lambda = 2$ Mbps, $N_h = 2$, Packet Size = 2 KB, $\lambda_h = 10$ Kbps .................................................................................. 66

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, $\lambda_h = 10$ Kbps .................................................................................. 68

Figure 4.1: The expected Euclidian distance between any random source and destination nodes ........................................................................... 76

Figure 4.2: Packet forwarding in a multi-hop path ............................................. 77

Figure 4.3: Least remaining distance for the first hop ........................................ 78

Figure 4.4: The distances between $S$ and neighbour nodes ............................... 79

Figure 4.5: The per hop distance for different values for $n$ and $R$ ....................... 80

Figure 4.6: Expected distance for different sizes of the network area ................. 84

Figure 5.1: Two hops communication path .......................................................... 89

Figure 5.2: Three hops communication path ......................................................... 91

Figure 5.3: SRN model for connection availability ............................................. 93

Figure 5.4: Distance between nodes ................................................................. 97

Figure 5.5: Path connection availability versus the side length of the network area, where $R = 250$ m, $N = 60$ or 100, and $\lambda = 10$ kbps ......................... 105
Figure 5.6: Path connection availability versus the side length of the network area, where \( R = 250 \text{ m}, N = 100, \) and \( \lambda = 10 \) or 40 kbps........................................ 106

Figure 5.7: Path connection availability versus the side length of the network area, where \( R = 250 \) or 200 m, \( N = 100, \) and \( \lambda = 10 \) kbps................................. 107

Figure 5.8: Path connection availability versus the side length of the network area, where \( R = 250 \text{ m}, N = 100, \lambda = 10 \text{ Kbps} \) and routing protocol is AODV or DSR.......................................................................................................................... 108

Figure 5.9: Path failure frequency versus the side length of the network area, where \( R = 250 \text{ m}, N = 100, \) and \( \lambda = 10 \text{ Kbs.} \) ................................................................................ 109

Figure 6.1: A network communication path ................................................................ 113

Figure 6.2: Hidden and interfering area...................................................................... 117

Figure 6.3: One Node detailed model for the BA method......................................... 121

Figure 6.4: The abstract model for the BA method................................................... 122

Figure 6.5: One node detailed model for the RTS/CTS method.............................. 125

Figure 6.6: The abstract model for the RTS/CTS method......................................... 126

Figure 6.7: Network layer model.............................................................................. 129

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, \( L = 600 \text{ m}, \) and \( R = 150 \text{ m} \)......................................................................................................................... 139

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, \( L = 600 \text{ m}, \) and \( R = 150 \text{ m} \)......................................................................................................................... 139
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.......................................................... 141

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.......................................................... 141

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......................................................................... 142

Figure 6.13: Goodput versus packet generation rate for the RTS/CTS method, in the case of packet size = 6 kB, $R_{cs} = 150$ m or 250, $L = 600$ m or 1000 m, and $R = 150$ m......................................................................... 142

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......................................................................... 144

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......................................................................... 144

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 .......................................................... 145

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 .......................................................... 146
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
## List of Tables

Table 1.1: Meaning of symbols in Figure 1.6 ................................................................. 20

Table 3.1: The average firing time of timed transitions of SRN models shown in Figures 3.5 and 3.7 .................................................................................................................. 49

Table 3.2: Transitions guard functions for SRN models shown in Figures 3.5 and 3.7 ................................................................................................................................. 49

Table 3.3: Arcs weight functions for SRN models shown in Figures 3.5 and 3.6...... 50

Table 3.4: The average firing time of timed transitions of SRN models shown in Figures 3.6 and 3.8 .................................................................................................................. 56

Table 3.5: The firing probabilities of immediate transitions of SRN models shown in Figures 3.7 and 3.8 ............................................................................................................ 57

Table 3.6: Arcs weight functions for SRN model shown in Figure 3.7 ....................... 57

Table 3.7: Arcs weight functions for SRN model shown in Figure 3.8 ....................... 58

Table 3.8: Transition guard functions for SRN models shown in Figures 3.6 and 3.8 ................................................................................................................................. 60

Table 3.9: Parameters of the MAC and Physical layers .............................................. 63

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... 85

Table 4.2: Comparison of simulation and LRD and MHD results for expected hop count for increasing values for \( L \) where \( R = 150 \) m ............................... 86

Table 5.1: Arcs weight functions for SRN model of intersection area number \( k \) ...... 95
Table 6.1: The key network simulation parameters................................. 136

Table 6.2: The time needed for analytical analysis and simulation of the network for different number of nodes................................................................. 148
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AODV</td>
<td>Ad-hoc On-Demand Distance Vector</td>
</tr>
<tr>
<td>BA</td>
<td>Basic Access</td>
</tr>
<tr>
<td>CA</td>
<td>Collision Avoidance</td>
</tr>
<tr>
<td>CBK</td>
<td>Call Back</td>
</tr>
<tr>
<td>CBR</td>
<td>Constant Bit Rate</td>
</tr>
<tr>
<td>CCA</td>
<td>Clear Channel Assessment</td>
</tr>
<tr>
<td>CDF</td>
<td>Cumulative Distribution Function</td>
</tr>
<tr>
<td>CSMA</td>
<td>Carrier Sensing Multiple Access</td>
</tr>
<tr>
<td>CSMA/CA</td>
<td>Carrier Sense Multiple Access with Collision Avoidance</td>
</tr>
<tr>
<td>CTMC</td>
<td>Continuous Time Markov Chain</td>
</tr>
<tr>
<td>CTS</td>
<td>Clear To Send</td>
</tr>
<tr>
<td>CW</td>
<td>Contention Window</td>
</tr>
<tr>
<td>DBPSK</td>
<td>Differential Binary Phase Shift Keying</td>
</tr>
<tr>
<td>DCF</td>
<td>Distributed Coordination Function</td>
</tr>
<tr>
<td>DES</td>
<td>Discrete Event Simulation</td>
</tr>
<tr>
<td>DFIR</td>
<td>Diffused Infrared</td>
</tr>
<tr>
<td>DIFS</td>
<td>DCF Inter Frame Space</td>
</tr>
<tr>
<td>DQPSK</td>
<td>Quadruple Phase Shift Keying</td>
</tr>
<tr>
<td>DSDV</td>
<td>Destination-Sequenced Distance-Vector routing</td>
</tr>
<tr>
<td>DSR</td>
<td>Dynamic Source Routing</td>
</tr>
<tr>
<td>DSSS</td>
<td>Direct Sequence Spread Spectrum</td>
</tr>
<tr>
<td>FAMA</td>
<td>Floor Acquisition Multiple Access</td>
</tr>
<tr>
<td>FHSS</td>
<td>Frequency Hopping Spread Spectrum</td>
</tr>
<tr>
<td>FSK</td>
<td>Frequency Shift Keying</td>
</tr>
<tr>
<td>GOD</td>
<td>General Operations Director</td>
</tr>
<tr>
<td>LAR</td>
<td>Location-Aided Routing</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>LLC</td>
<td>Logical Link Control</td>
</tr>
<tr>
<td>LRC</td>
<td>Long Retry Counter</td>
</tr>
<tr>
<td>LRD</td>
<td>Least Remaining Distance</td>
</tr>
<tr>
<td>MAC</td>
<td>Medium Access Control</td>
</tr>
<tr>
<td>MACA</td>
<td>Multiple Access Collision Avoidance</td>
</tr>
<tr>
<td>MANETs</td>
<td>Mobile Ad hoc Networks</td>
</tr>
<tr>
<td>MHD</td>
<td>Maximum Hop Distance</td>
</tr>
<tr>
<td>MPDU</td>
<td>MAC Protocol Data Unit</td>
</tr>
<tr>
<td>MRL</td>
<td>Maximum Retry Limit</td>
</tr>
<tr>
<td>NAV</td>
<td>Network Allocation Vector</td>
</tr>
<tr>
<td>OLSR</td>
<td>Optimized Link State Routing</td>
</tr>
<tr>
<td>PCF</td>
<td>Point Coordination Function</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PDU</td>
<td>Protocol Data Unit</td>
</tr>
<tr>
<td>PPM</td>
<td>Pulse Position Modulation</td>
</tr>
<tr>
<td>QoS</td>
<td>Quality Of Service</td>
</tr>
<tr>
<td>RERR</td>
<td>Route Error</td>
</tr>
<tr>
<td>RREP</td>
<td>Route Reply</td>
</tr>
<tr>
<td>RREQ</td>
<td>Route Request</td>
</tr>
<tr>
<td>RTS</td>
<td>Request To Send</td>
</tr>
<tr>
<td>RTS/CTS</td>
<td>Request-To-Send/Clear-To-Send</td>
</tr>
<tr>
<td>RWPPPMM</td>
<td>Random Waypoint Mobility Model</td>
</tr>
<tr>
<td>SIFS</td>
<td>Short Inter Frame Space</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise Ratio</td>
</tr>
<tr>
<td>SRC</td>
<td>Short Retry Counter</td>
</tr>
<tr>
<td>SRN</td>
<td>Stochastic Reward Net</td>
</tr>
<tr>
<td>TAB</td>
<td>Traffic-Analysis-Based</td>
</tr>
<tr>
<td>TBRPF</td>
<td>Topology Broadcast Based on Reverse-Path Forwarding</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TORA</td>
<td>Temporally-Ordered Routing Algorithm</td>
</tr>
<tr>
<td>TTL</td>
<td>Time To Live</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>WLAN</td>
<td>Wireless Local Area Networks</td>
</tr>
</tbody>
</table>


List of Symbols

\( A_{AB} \) \hspace{1cm} \text{Intersection area between node } A \text{ and } B

\( A_h \) \hspace{1cm} \text{Average size of the hidden area}

\( A_i \) \hspace{1cm} \text{Average size of the interfering area}

\( A_{int} \) \hspace{1cm} \text{Size of intersection area}

\( A_s \) \hspace{1cm} \text{Average number of backoff slots}

\( B_1 \) \hspace{1cm} \text{Minimum transmission bit rate}

\( B_2 \) \hspace{1cm} \text{Maximum transmission bit rate}

\( CW \) \hspace{1cm} \text{Contention backoff window}

\( CW_{\text{max}} \) \hspace{1cm} \text{Maximum contention window}

\( CW_{\text{min}} \) \hspace{1cm} \text{Minimum contention window}

\( d \) \hspace{1cm} \text{Distance between a source and destination node}

\( d_r \) \hspace{1cm} \text{Distance between 2-hop-apart nodes}

\( E(x) \) \hspace{1cm} \text{Expected value for a parameter } x

\( Ft(T_x) \) \hspace{1cm} \text{Average firing time of transition } T_x

\( G_R \) \hspace{1cm} \text{Antenna gains of the receiver}

\( G_t \) \hspace{1cm} \text{Antenna gains of the sender}

\( L \) \hspace{1cm} \text{Side length of the squared network area}

\( N \) \hspace{1cm} \text{Number of nodes in the network}

\( N_B \) \hspace{1cm} \text{Buffer size}

\( N_{cs} \) \hspace{1cm} \text{Average number of carrier sensing nodes}

\( N_e \) \hspace{1cm} \text{Average number of hops from any broken link in the path to the source node}

\( N_H \) \hspace{1cm} \text{Average number of hidden nodes}

\( N_{HS} \) \hspace{1cm} \text{Number of hops to the source of undeliverable packet}

\( N_h \) \hspace{1cm} \text{Average path length in hops}
\( N_i \) Average number of interfering nodes
\( N_{in} \) Average number of nodes in the intersections areas
\( N_{L,H} \) 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 \( i = 0, 1, 2, \ldots \)
\( P_r \) 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

$\alpha$  Average node utilisation

$\alpha_x$  Utilisation of a node $x$

$\beta_1$  Probability that the channel is busy in the case of single hop ad hoc networks with the basic access scheme

$\beta_2$  Probability that the channel is busy in the case of single hop ad hoc networks with the RTS/CTS scheme

$\beta_{BA}$  Probability that the channel is idle in the case of a multi-hop ad hoc network with the basic access scheme

$\beta_c$  Backoff counter

$\beta_{cMAX}$  Maximum value of the retry counter

$\beta_R$  Probability that the channel is idle in the case of a multi-hop ad hoc network with the RTS/CTS scheme

$\delta$  Packet delay per hop

$\delta_d$  Packet delay in data link layer

$\delta_n$  Average delay of packets in network layer model
\( \varepsilon \)  
Packet loss probability in the data link layer model

\( \varepsilon_B \)  
Packet loss probability due to buffer overflow

\( \lambda \)  
Data transmission rate

\( \lambda_c \)  
Number of control packets sent per unit time

\( \lambda_e \)  
Entering rate to intersection area

\( \lambda_h \)  
Packet generation rate of hidden nodes

\( \lambda_n \)  
Throughput of the network layer model

\( \lambda_r \)  
Average number of routed packets per node per unit time

\( \lambda_{rx} \)  
Packet forward rate for a node \( x \)

\( \lambda_{rt} \)  
Average number of routed packets per unit time

\( \lambda_T \)  
Throughput of the transport layer model

\( \lambda_{tx} \)  
Number of packets that is successfully transmitted by any node \( x \) per unit time

\( \mu \)  
Intersection area leaving time

\( \mu_B \)  
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

\( \mu_c \)  
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

\( \mu_f \)  
Average failure rate of paths

\( \mu_R \)  
Probability of success to complete the RTS/CTS handshake in the case of a multi-hop ad hoc network

\( \mu_r \)  
Average repair rate of paths

\( \mu_s \)  
Path failure and repairing frequency

\( \mu_1 \)  
Probability of failure to send a data frame in the case of single hop ad hoc networks with the basic access scheme

\( \mu_2 \)  
Probability of failure to send a data frame in the case of single hop ad hoc networks with the RTS/CTS scheme

\( \xi \)  
Square of the distance between a source and destination node

\( \rho_{BA} \)  
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
\( \rho_i \) Distance to a neighbour node number \( i \), where \( i = 1, 2, \ldots \)

\( \rho_{\text{max}} \) Maximum value of \( \rho_i \), where \( i = 1, 2, \ldots, N_r \)

\( \rho_R \) 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

\( \tau_{\text{LR}} \) Time required to finish the local repair process

\( \tau_{\text{NR}} \) Time required for establishing a new route

\( \tau_{\text{RERR}} \) Time required for broadcasting RERR message

\( \tau_{\text{SR}} \) Time required to finish the source repair process

\( \Psi \) Average path connection availability

\( \omega_R \) 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](image)
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](image)

### 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.

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) $\lambda$. 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 ($V_{\text{max}}$). 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](image)

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.
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) ($\psi$); average rate of failure ($\mu_f$); and average rate of repair ($\mu_r$). According to the routing protocol (such as AODV, DSR, or LAR), average packet delay per hop ($\delta$); 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**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Packets generation rate</td>
</tr>
<tr>
<td>$\varepsilon_B$</td>
<td>Packet loss probability due to buffer overflow</td>
</tr>
<tr>
<td>$\delta_n$</td>
<td>Average delay of packets in the network layer</td>
</tr>
<tr>
<td>$\lambda_T$</td>
<td>Throughput of the transport layer model</td>
</tr>
<tr>
<td>$\lambda_n$</td>
<td>Throughput of the network layer model</td>
</tr>
<tr>
<td>$\delta_d$</td>
<td>Average packet delay in the data link layer</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Average packet delay per hop</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>Packet loss probability in the data link layer</td>
</tr>
<tr>
<td>$n_c$</td>
<td>Average number of tries to transmit a packet</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of nodes in the network</td>
</tr>
<tr>
<td>$\mu_f$</td>
<td>Average failure rate of paths</td>
</tr>
<tr>
<td>$\mu_r$</td>
<td>Average repair rate of paths</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Average path availability</td>
</tr>
<tr>
<td>$L$</td>
<td>Side length of the squared network area</td>
</tr>
<tr>
<td>$N_h$</td>
<td>Average path length in hops</td>
</tr>
<tr>
<td>$N_{H}, N_{I}, N_{cs}$</td>
<td>Average number of hidden nodes, interfering nodes, and carrier sensing nodes</td>
</tr>
</tbody>
</table>
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 ($\lambda_n$) to compute the average number of tries to transmit a packet ($n_c$), packet loss probability ($\epsilon$), average packet delay in the data link layer ($\delta_d$), and average packet delay per hop ($\delta$). 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 $\psi, \mu_r, \mu_f, \lambda_T$ (the throughput of the transport layer model), $n_c, \delta_d$ and $\epsilon$ to calculate the average number of packets per unit time that is sent to the data link layer model ($\lambda_n$), the probability of packet loss due to buffer overflow ($\epsilon_B$), and the average delay of packets in the network layer model ($\delta_n$). 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 $\lambda, \epsilon_B, \delta_n, \delta_d$, and $\epsilon$, and the output is $\lambda_T$. 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 discusses 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 queueing 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
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 wireless 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.

![Diagram of BA and RTS/CTS methods](image)

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 ∙ 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:
where \( CW_{\text{min}} \) is the minimum contention window and \( \beta_c \) is the backoff counter (retry counter) that counts the number of failures of sending a packet; \( \beta_c \) increases by one each time a transmission fails. At the first attempt to transmit a packet, \( \beta_c \) is initialised with zero and then it is incremented by one at each retransmission for the same packet. \( \beta_c \) increases to its maximum value, called Maximum Retry Limit (MRL), corresponding to the maximum contention window \( (CW_{\text{max}}) \). After successful transmission of any packet, \( \beta_c \) 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 \( 2 \cdot \text{SIFS} + t_{\text{CTS}} + 2 \cdot T_s \) (where \( t_{\text{CTS}} \) 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 $\lambda$ 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](image)

Each of the hidden nodes generates packets at a rate $\lambda_h$ 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 $\lambda_h$ are the model parameters that are varied to different values, as explained in Section 3.6. Also, $\lambda$ 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](image)

### 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 $T_{busy1}$. 

47
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 \( P_{busyl} \) after the firing of \( T_{busyl} \), which is returned back to the place \( P_{DIFS1} \) after the firing of \( T_{NAV1} \). The probabilities of firing of the transitions \( T_{busyl} (\beta_1) \) and \( T_{idle1} (1 - \beta_1) \) are the probabilities that the channel is busy and idle respectively. The parameter \( \beta_1 \) 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:

\[
CW = (CWmin + 1) \cdot 2^{\#PFC-1} - 1
\]

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 \( T_{busyl2} \), 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 \( T_{busyl2} \), place \( P_{busyl2} \), 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

<table>
<thead>
<tr>
<th>Transition</th>
<th>Average firing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{PG}$</td>
<td>$\lambda^{-1}$</td>
</tr>
<tr>
<td>$T_{DIFS1}, T_{DIFS2}$</td>
<td>$DIFS$</td>
</tr>
<tr>
<td>$T_{NAV1}, T_{NAV2}$</td>
<td>$Ft(T_{txD}) + Ft(T_{ACK})$</td>
</tr>
<tr>
<td>$T_{slot}$</td>
<td>$T_s$</td>
</tr>
<tr>
<td>$T_{BO}$</td>
<td>$A_s \cdot T_s$</td>
</tr>
<tr>
<td>$T_{txD}$</td>
<td>$P_h H + \frac{MPDU}{B1} + \frac{MPDU}{B2} + T_p$</td>
</tr>
<tr>
<td>$T_{ACK}$</td>
<td>$P_h H + \frac{ACK}{B1} + \frac{SIFS}{B2} + T_p + T_{CCA} + T_{RTx}$</td>
</tr>
<tr>
<td>$T_{timeout}$</td>
<td>$Ft(T_{txD}) + Ft(T_{ACK})$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Transition</th>
<th>Guard Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{busy2}$</td>
<td>$#P_{busy2} + #P_{DIFS2} = 0$</td>
</tr>
<tr>
<td>$T_{idle2}$</td>
<td>$#P_{busy2} + #P_{DIFS2} = 0$</td>
</tr>
<tr>
<td>$T_{DIFS}$</td>
<td>$#P_{ch} + #P_{succ} + #P_{fail} = 0$</td>
</tr>
<tr>
<td>$T_{BO}$</td>
<td>$#P_{ch} + #P_{succ} + #P_{fail} = 0$</td>
</tr>
<tr>
<td>$T_{coll}$</td>
<td>$#P_{ch} + #P_{succ} + #P_{fail} = 0$ and $#P_{FBO} &gt; 1$</td>
</tr>
<tr>
<td>$T_{Ncoll}$</td>
<td>$#P_{ch} + #P_{succ} + #P_{fail} = 0$ and $#P_{FBO} &gt; 0$</td>
</tr>
<tr>
<td>$T_{succ}$</td>
<td>$#P_{ch} = 1$</td>
</tr>
<tr>
<td>$T_{fail}$</td>
<td>$#P_{ch} + #P_{ch-h} &gt; 1$</td>
</tr>
<tr>
<td>$T_{DIFS-h}$</td>
<td>$#P_{ch-h} + #P_{succ-h} + #P_{fail-h} = 0$</td>
</tr>
<tr>
<td>$T_{BO-h}$</td>
<td>$#P_{cht-h} + #P_{succ-h} + #P_{fail-h} = 0$</td>
</tr>
<tr>
<td>$T_{coll-h}$</td>
<td>$#P_{ch-h} + #P_{succ-h} + #P_{fail-h} = 0$ and $#P_{FBO-h} &gt; 1$</td>
</tr>
<tr>
<td>$T_{Ncoll-h}$</td>
<td>$#P_{ch-h} + #P_{succ-h} + #P_{fail-h} = 0$ and $#P_{FBO-h} &gt; 0$</td>
</tr>
<tr>
<td>$T_{succ-h}$</td>
<td>$#P_{ch-h} = 1$</td>
</tr>
<tr>
<td>$T_{fail-h}$</td>
<td>$#P_{ch-h} &gt; 1$</td>
</tr>
<tr>
<td>$T_{txD-h}$</td>
<td>$#P_{succ} = 0$</td>
</tr>
</tbody>
</table>
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 $\beta_1$ 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

<table>
<thead>
<tr>
<th>Arc name</th>
<th>Arc weight function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_1$</td>
<td>0 if $RNS &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>1 if $RNS = 0$</td>
</tr>
<tr>
<td>$W_2$</td>
<td>$#P_{FC}$</td>
</tr>
<tr>
<td>$W_3$</td>
<td>$#P_{FC}$ if $#P_{FC} = MRL$</td>
</tr>
<tr>
<td></td>
<td>0 if $#P_{FC} &lt; MRL$</td>
</tr>
<tr>
<td>$W_4$</td>
<td>1 if $#P_{FC} &lt; MRL$</td>
</tr>
<tr>
<td></td>
<td>0 if $#P_{FC} = MRL$</td>
</tr>
<tr>
<td>$W_5$</td>
<td>0 if $#P_{FC} &lt; MRL$</td>
</tr>
<tr>
<td></td>
<td>1 if $#P_{FC} = MRL$</td>
</tr>
</tbody>
</table>

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 backoff 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 $P_{txD}$ 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 $P_{txD}$ 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 $\mu_1$ and $(1 - \mu_1)$ respectively, where $\mu_1$ is the probability of failure to transmit the data frame due to interference induced by neighbour or hidden nodes. The parameters $\beta_1$ and $\mu_1$ 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 $(2\cdot\text{SIFS} + t_{CTS} + 2\cdot T_s)$ 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](image)

The firing probabilities of the transitions $T_{busy1}$ ($\beta_2$) and $T_{idle1}$ ($1 - \beta_2$) are the probabilities that the channel is busy and idle respectively. The firing probabilities of the conflicted transitions $T_{f1}$ ($\mu_2$) and $T_{s1}$ ($1 - \mu_2$) are the probabilities of failure and success to complete the RTS/CTS handshake respectively. The parameter $\beta_2$ and $\mu_2$
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 $T_{NAV}$ reflects the time needed to complete the RTS/CTS handshake. The average firing time of $T_{RNAV}$ is $2 \cdot \text{SIFS} + \text{Ft}(T_{CTS}) + 2 \cdot T_s$.

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 $T_s, T_f, P_s, P_f, T_{NAV2},$ 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 $\mu_2$ 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.

![Diagram](image)

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 $P_{DIFS}$ 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 ($#P_{ch} > 0$).
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 $\beta_1$ and $\mu_1$ derived from the abstract model as follows:

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

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

where $Pr(E)$ and $#P_s$ are the probability of an event $E$ and the number of tokens in the place $P_s$, respectively.
Table 3.4: The average firing time of timed transitions of SRN models shown in Figures 3.6 and 3.8

<table>
<thead>
<tr>
<th>Transition</th>
<th>Average firing time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{FBO}$</td>
<td>$\lambda^{-1}$</td>
</tr>
<tr>
<td>$T_{DIFS}$, $T_{DIFS2}$</td>
<td>$DIFS$</td>
</tr>
<tr>
<td>$T_{NAV1}$, $T_{NAV2}$</td>
<td>$F(t_{RTS}) + F(t_{CTS}) + F(t_{txD}) + F(t_{ACK})$</td>
</tr>
<tr>
<td>$T_{RNAV1}$, $T_{RNAV2}$</td>
<td>$F(t_{CTS}) + 2SIFS + 2T_s$</td>
</tr>
<tr>
<td>$T_{slot}$</td>
<td>$T_s$</td>
</tr>
<tr>
<td>$T_{BO}$</td>
<td>$A_s \cdot T_s$</td>
</tr>
<tr>
<td>$T_{RTS}$</td>
<td>$\frac{PH}{B_1} + \frac{RTS}{B_2} + T_p$</td>
</tr>
<tr>
<td>$T_{CTS}$</td>
<td>$\frac{PH}{B_1} + \frac{CTS}{B_2} + SIFS + T_p + T_{CCA} + T_{RXTX}$</td>
</tr>
<tr>
<td>$T_{txD}$</td>
<td>$\frac{PH}{B_1} + \frac{MPDU}{B_2} + SIFS + T_p + T_{CCA} + T_{RXTX}$</td>
</tr>
<tr>
<td>$T_{ACK}$</td>
<td>$\frac{PH}{B_1} + \frac{ACK}{B_2} + SIFS + T_p + T_{CCA} + T_{RXTX}$</td>
</tr>
<tr>
<td>$T_{timeout}$</td>
<td>$F(t_{RTS}) + F(t_{CTS})$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Transition</th>
<th>Firing Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{\text{coll}}$</td>
<td>$1 - \left(1 - \frac{1}{A_s}\right)^{#P_{\text{FBO}}}$</td>
</tr>
<tr>
<td>$T_{\text{Ncoll}}$</td>
<td>$\left(1 - \frac{1}{A_s}\right)^{#P_{\text{FBO}}}$</td>
</tr>
<tr>
<td>$T_{\text{coll-h}}$</td>
<td>$1 - \left(1 - \frac{1}{A_s}\right)^{#P_{\text{FBO-h}}}$</td>
</tr>
<tr>
<td>$T_{\text{Ncoll-h}}$</td>
<td>$\left(1 - \frac{1}{A_s}\right)^{#P_{\text{FBO-h}}}$</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Arc name</th>
<th>Arc weight function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_1$</td>
<td>#P_{BO} if #P_{ch} = 0 and #P_{BO} &gt; 0  &lt;br&gt;1 if #P_{BO} = 0</td>
</tr>
<tr>
<td>$W_2$</td>
<td>#P_{BO}</td>
</tr>
<tr>
<td>$W_3$</td>
<td>#P_{BO} if #P_{FBO} &gt; 1  &lt;br&gt;1 if #P_{FBO} = 0</td>
</tr>
<tr>
<td>$W_4$</td>
<td>#P_{FBO}</td>
</tr>
<tr>
<td>$W_5$</td>
<td>#P_{FBO}</td>
</tr>
<tr>
<td>$W_6$</td>
<td>#P_{txD} if #P_{txD} &gt; 0  &lt;br&gt;1 if #P_{txD} = 0</td>
</tr>
<tr>
<td>$W_7$</td>
<td>#P_{txD}</td>
</tr>
<tr>
<td>$W_8$</td>
<td>#P_{rsD}</td>
</tr>
<tr>
<td>$W_9$</td>
<td>#P_{ch} if #P_{ch} &gt; 0  &lt;br&gt;0 if #P_{ch} = 0</td>
</tr>
<tr>
<td>$W_{10}$</td>
<td>#P_{rsD}</td>
</tr>
<tr>
<td>$W_{11}$</td>
<td>#P_{fail} if #P_{fail} &gt; 0  &lt;br&gt;1 if #P_{fail} = 0</td>
</tr>
<tr>
<td>$W_{12}$</td>
<td>#P_{fail}</td>
</tr>
</tbody>
</table>
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 $T_{coll}$, $P_{txD}$, $T_{txD}$, $P_{fail}$, $T_{fail}$, 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

<table>
<thead>
<tr>
<th>Arc name</th>
<th>Arc weight function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_1$</td>
<td>$#P_{BO}$ if $#P_{ch} = 0$ and $#P_{BO} &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>1 if $#P_{BO} = 0$</td>
</tr>
<tr>
<td>$W_2$</td>
<td>$#P_{BO}$</td>
</tr>
<tr>
<td>$W_3$</td>
<td>$#P_{BO}$ if $#P_{FBO} &gt; 1$</td>
</tr>
<tr>
<td></td>
<td>1 if $#P_{FBO} = 0$</td>
</tr>
<tr>
<td>$W_4$</td>
<td>$#P_{FBO}$</td>
</tr>
<tr>
<td>$W_5$</td>
<td>$#P_{FBO}$</td>
</tr>
<tr>
<td>$W_6$</td>
<td>$#P_{txRTS}$ if $#P_{txRTS} &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>1 if $#P_{txRTS} = 0$</td>
</tr>
<tr>
<td>$W_7$</td>
<td>$#P_{txRTS}$</td>
</tr>
<tr>
<td>$W_8$</td>
<td>$#P_{rxRTS}$</td>
</tr>
<tr>
<td>$W_9$</td>
<td>$#P_{ch}$ if $#P_{ch} &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>0 if $#P_{ch} = 0$</td>
</tr>
<tr>
<td>$W_{10}$</td>
<td>$#P_{rxRTS}$</td>
</tr>
<tr>
<td>$W_{11}$</td>
<td>$#P_{fail}$ if $#P_{fail} &gt; 0$</td>
</tr>
<tr>
<td></td>
<td>1 if $#P_{fail} = 0$</td>
</tr>
<tr>
<td>$W_{12}$</td>
<td>$#P_{fail}$</td>
</tr>
</tbody>
</table>

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 $P_{x-h}$, the transition $T_{x-h}$, 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 $T_{BO-h}$, which is equal to $\left( A_{x-h} \cdot T_x \right)$, where $A_{x-h}$ is the average number of backoff slots of nodes in the hidden area. The
parameter $A_{s,h}$ is computed in the same way as $A_s$ using the one node detailed model, where its parameters $\beta_1$ and $\mu_1$ are computed from the abstract model as follows:

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

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

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 $P_{ch,h}$ and $T_{succ}$ is added to disable $T_{succ}$ and enable $T_{fail}$ when the number of tokens in $P_{ch,h}$ 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 $T_{DIFS,h}$ and $T_{BO,h}$.

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 $\beta_2$ and $\mu_2$ derived from the abstract model shown Figure 3.8 as follows:

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

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

<table>
<thead>
<tr>
<th>Transition</th>
<th>Guard Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{busy2}$</td>
<td>$P_{busy2}+P_{x2}+P_{DIFS2} = 0$</td>
</tr>
<tr>
<td>$T_{idle2}$</td>
<td>$P_{idle2}+P_{x2}+P_{DIFS2} = 0$</td>
</tr>
<tr>
<td>$T_{DIFS}$</td>
<td>$P_{ch}+P_{CTS}+P_{txD}+P_{ACK}+P_{fail} = 0$</td>
</tr>
<tr>
<td>$T_{BO}$</td>
<td>$P_{ch}+P_{CTS}+P_{txD}+P_{ACK}+P_{fail} = 0$</td>
</tr>
<tr>
<td>$T_{coll}$</td>
<td>$P_{ch}+P_{CTS}+P_{txD}+P_{ACK}+P_{fail} = 0$ and $P_{FBO} &gt; 1$</td>
</tr>
<tr>
<td>$T_{Ncoll}$</td>
<td>$P_{ch}+P_{CTS}+P_{txD}+P_{ACK}+P_{fail} = 0$ and $P_{FBO} &gt; 0$</td>
</tr>
<tr>
<td>$T_{succ}$</td>
<td>$P_{ch} = 1$</td>
</tr>
<tr>
<td>$T_{fail}$</td>
<td>$P_{ch}+P_{ch-1}+P_{CTS-h}+P_{txD-h}+P_{ACK-h} &gt; 1$</td>
</tr>
<tr>
<td>$T_{DIFS-h}$</td>
<td>$P_{ch-h}+P_{CTS-h}+P_{txD-h}+P_{ACK-h}+P_{fail-h}+P_{CTS} = 0$</td>
</tr>
<tr>
<td>$T_{BO-h}$</td>
<td>$P_{ch-h}+P_{CTS-h}+P_{txD-h}+P_{ACK-h}+P_{fail-h}+P_{CTS} = 0$</td>
</tr>
<tr>
<td>$T_{coll-h}$</td>
<td>$P_{ch-h}+P_{CTS-h}+P_{txD-h}+P_{ACK-h}+P_{fail-h} = 0$ and $P_{FBO-h} &gt; 1$</td>
</tr>
<tr>
<td>$T_{Ncoll-h}$</td>
<td>$P_{ch-h}+P_{CTS-h}+P_{txD-h}+P_{ACK-h}+P_{fail-h} = 0$ and $P_{FBO-h} &gt; 0$</td>
</tr>
<tr>
<td>$T_{succ-h}$</td>
<td>$P_{ch-h} = 1$</td>
</tr>
<tr>
<td>$T_{fail-h}$</td>
<td>$P_{ch-h} &gt; 1$</td>
</tr>
<tr>
<td>$T_{RTS}$</td>
<td>$P_{CTS}+P_{txD}+P_{ACK} = 0$</td>
</tr>
</tbody>
</table>

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 $P_{x-h}$, the transition $T_{x-h}$, 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 $T_{BO-h}$ is $\left(A_{x-h} \cdot T_S\right)$, where $A_{x-h}$ is the average number of tokens in the place $P_\text{slot}$ in the one
node detailed model which its parameters $\beta_2$ and $\mu_2$ are computed from the abstract model as follows:

\[
\beta_2 = Pr \left( \#P_{\text{ch.h}} > 0 \ OR \ #P_{\text{CTS.h}} > 0 \ OR \ #P_{\text{txD.h}} > 0 \ OR \ #P_{\text{ACK.h}} > 0 \\
OR \ #P_{\text{CTS}} > 0 \ OR \ #P_{\text{txD}} > 0 \ OR \ #P_{\text{ACK}} > 0 \right)
\]

\[
\mu_2 = Pr(\#P_{\text{ch.h}} > 1)
\]

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 $P_{\text{CTS.h}}, P_{\text{txD.h}},$ and $P_{\text{ACK.h}}$ to the transition $T_{\text{succ}}$, 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 $P_{\text{ch.h}}$ and transition $T_{\text{succ}}$ which disables it and enables $T_{\text{fail}}$. 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 $T_{\text{DIFS-h}}$ and $T_{\text{BO-h}}$ and places $P_{\text{txD}}$ and $P_{\text{ACK}}$, 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 ($A_{s\cdot h}$), which equals to the average number of tokens in the place $P_{slot}$. Using $A_s$ and $A_{s\cdot h}$, the abstract model is solved to derive the performance metric and the parameters $\beta_1$, $\mu_1$, $\beta_2$, and $\mu_2$, 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:

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

**Step 2:** The abstract model is solved using the initial value of the backoff window to get the initial values of a performance metric ($\tau^0$ e.g. throughput) and parameters $\beta_1$ and $\mu_1$ (or $\beta_2$ and $\mu_2$ in the case of RTS/CTS method).

**Step 3:** The one node detailed model is solved using the last computed values of parameters $\beta_1$ and $\mu_1$ (or $\beta_2$ and $\mu_2$) to compute the new values for $A_s$ and $A_{s\cdot h}$.

**Step 4:** The abstract model is solved to get the performance metric $\tau^n$ and parameters $\beta_1$ and $\mu_1$ (or $\beta_2$ and $\mu_2$).

**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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CW_{\text{min}}$</td>
<td>31</td>
</tr>
<tr>
<td>$CW_{\text{max}}$</td>
<td>1023</td>
</tr>
<tr>
<td>$T_s$</td>
<td>20 $\mu$s</td>
</tr>
<tr>
<td>$T_{\text{RxTx}}$</td>
<td>5 $\mu$s</td>
</tr>
<tr>
<td>$T_{\text{CCA}}$</td>
<td>15 $\mu$s</td>
</tr>
<tr>
<td>DIFS</td>
<td>50 $\mu$s</td>
</tr>
<tr>
<td>SIFS</td>
<td>10 $\mu$s</td>
</tr>
<tr>
<td>$PhH$</td>
<td>192 bit</td>
</tr>
<tr>
<td>MAC Header</td>
<td>292 bit</td>
</tr>
<tr>
<td>RTS</td>
<td>160 bit + $PhH$</td>
</tr>
<tr>
<td>CTS</td>
<td>112 bit + $PhH$</td>
</tr>
<tr>
<td>ACK</td>
<td>112 bit + $PhH$</td>
</tr>
<tr>
<td>$B_1$</td>
<td>1 Mbps</td>
</tr>
<tr>
<td>$B_2$</td>
<td>2 Mbps</td>
</tr>
<tr>
<td>SRC</td>
<td>6</td>
</tr>
<tr>
<td>LRC</td>
<td>4</td>
</tr>
</tbody>
</table>

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 ($\delta$) 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:

\[
\text{Goodput} = \text{Thr}(T_{PG})
\]

\[
\delta = \frac{N - M(P_B)}{\text{Thr}(T_{PG})}
\]

where \(\text{Thr}(T_x)\) 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, \(\lambda_h = 10\) 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, $\lambda_h = 100$ 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, \( \lambda_h = 100 \) 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.

![Graph](image)

Figure 3.11: Packet delay versus number of nodes for both BA and RTS/CTS methods, in the case of \( \lambda = 2 \) Mbps, \( N_h = 2 \), Packet Size = 2 KB, \( \lambda_h = 10 \) 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 \( N_h = 2, \lambda_h = 10 \) Kbps, \( \lambda = 2 \) 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.
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.

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, $\lambda_h = 10$ Kbps
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]:

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

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 $0 \leq x_1 \leq L$ and $0 \leq x_2 \leq L$.

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:

$$P(s \leq d) = \int \int f_{X_1X_2}(x_1, x_2) \, dx_2 \, dx_1 = \int \int f_{X_1}(x_1) \cdot f_{X_2}(x_2) \, dx_2 \, dx_1$$

$$= \int_0^d \int_0^{d+x_1} f_{X_1}(x_1) \cdot f_{X_2}(x_2) \, dx_2 \, dx_1$$

$$+ \int_d^{L-d} \int_{x_1-d}^{d+x_1} f_{X_1}(x_1) \cdot f_{X_2}(x_2) \, dx_2 \, dx_1$$

$$+ \int_{L-d}^L \int_{x_1-L}^L f_{X_1}(x_1) \cdot f_{X_2}(x_2) \, dx_2 \, dx_1$$

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

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

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 \( \xi \) is the square of the distance \( d \) between \( X_1 \) and \( X_2 \), \( \xi \) is given by:

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

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

\[
f_{\xi y}(\xi) = \int f_{\xi x}(z) \cdot f_{\xi y}(\xi - z) \, dz \tag{4.2}
\]

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

\[
F_{\xi x}(\xi_x) = \frac{12}{5} \sqrt{\frac{\xi_x}{L}} - \frac{4}{L^3} \sqrt{\frac{\xi_x^3}{L^3}} + \frac{3 \xi_x^2}{L^4} - \frac{2 \xi_x^3}{5L^6}
\]

The PDF of \( \xi_x \) is obtained as follows:

\[
f_{\xi x}(\xi_x) = \frac{\partial}{\partial \xi_x} F_{\xi x}(\xi_x) = \frac{6 \xi_x}{L^4} - \frac{6 \xi_x^2}{5L^6} + \frac{6}{5L \sqrt{\xi_x}} - \frac{6 \sqrt{\xi_x}}{L^3} \tag{4.3}
\]

In the same way, \( f_{\xi y}(\xi_y) \) can be obtained. Because the domain of \( \xi \) 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:
\[ f_{xy}(\xi) = \begin{cases} 
I_1(\xi) = \int_0^\xi f_{\xi x}(z) \cdot f_{\xi y}(\xi - z) \, dz & 0 < \xi \leq L^2 \\
I_2(\xi) = \int_{L^2}^{\xi} f_{\xi x}(z) \cdot f_{\xi y}(\xi - z) \, dz & L^2 < \xi \leq 2L^2 
\end{cases} \]

(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:

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

(4.5)

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

(4.6)

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

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

(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](image)
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](image)

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 $\rho_i$ and $\theta_i$, 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 ($\rho$) 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 $\theta$ gives the PDF of $\rho$ as:

$$f_{\rho}(\rho) = \frac{2\rho}{R^2}$$
To minimize the hop count to the destination, the neighbour node with the maximum distance ($\rho_{\text{max}}$) from the source $S$ is chosen to forward the packets. According to [100], because $\rho_1, \ldots, \rho_{N_r}$ are i.i.d. random variables each with PDF $f_\rho(\rho)$, the PDF of $\rho_{\text{max}}$ is:

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

(4.9)

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

$$E(\rho_{\text{max}}) = \int_{0}^{R} \rho f_{\rho_{\text{max}}}(\rho) d\rho = \frac{2N_r}{2N_r + 1} R$$

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

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

(4.10)

![Figure 4.5: The per hop distance for different values for $n$ and $R$](image)

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 (Nv) 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:

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

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

$$P_\theta(\theta \leq a) = \int_{-\alpha}^{a} f_\theta(\theta) \, d\theta = \frac{2}{\pi} \alpha$$  \hspace{1cm} (4.11)

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

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

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

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

By definition, the expected value of X can be deduced from Equation (4.12) as follows:
\[ X_r = E(X) = \int_{d-r}^{\sqrt{d^2+r^2}} X \cdot f_X(X) \, dX \quad (4.13) \]

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 \( \text{hop\_count} = 0 \)

**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 \( \text{hop\_count} = \frac{d}{R} \), go to the End

**Step 6:** Set \( \text{hop\_count} = \text{hop\_count} + 1 \)

**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 \( \text{hop\_count} = \text{hop\_count} + \frac{X_r}{R} \)

**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 ($V_{\text{max}}$). 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.

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

<table>
<thead>
<tr>
<th>$L$</th>
<th>Expected Number of Hops</th>
<th>$R = 200$</th>
<th>$R = 250$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sim</td>
<td>Ana</td>
<td>Sim</td>
</tr>
<tr>
<td>700</td>
<td>2.61</td>
<td>2.73</td>
<td>2.01</td>
</tr>
<tr>
<td>800</td>
<td>2.93</td>
<td>2.84</td>
<td>2.14</td>
</tr>
<tr>
<td>900</td>
<td>3.19</td>
<td>2.97</td>
<td>2.69</td>
</tr>
<tr>
<td>1000</td>
<td>3.65</td>
<td>3.8</td>
<td>2.94</td>
</tr>
<tr>
<td>1100</td>
<td>4.10</td>
<td>3.92</td>
<td>3.18</td>
</tr>
<tr>
<td>1200</td>
<td>4.58</td>
<td>4.76</td>
<td>3.57</td>
</tr>
<tr>
<td>1300</td>
<td>4.93</td>
<td>4.86</td>
<td>3.86</td>
</tr>
<tr>
<td>1400</td>
<td>5.28</td>
<td>4.99</td>
<td>4.16</td>
</tr>
<tr>
<td>1500</td>
<td>5.79</td>
<td>5.82</td>
<td>4.52</td>
</tr>
<tr>
<td>1600</td>
<td>6.19</td>
<td>5.93</td>
<td>4.63</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>$L$</th>
<th>Sim</th>
<th>LRD</th>
<th>MHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>700</td>
<td>3.37</td>
<td>2.39</td>
<td>3.74</td>
</tr>
<tr>
<td>800</td>
<td>3.83</td>
<td>2.62</td>
<td>3.86</td>
</tr>
<tr>
<td>900</td>
<td>4.40</td>
<td>2.86</td>
<td>4.75</td>
</tr>
<tr>
<td>1000</td>
<td>4.84</td>
<td>3.40</td>
<td>4.86</td>
</tr>
<tr>
<td>1100</td>
<td>5.42</td>
<td>3.63</td>
<td>5.75</td>
</tr>
<tr>
<td>1200</td>
<td>5.83</td>
<td>3.88</td>
<td>5.86</td>
</tr>
<tr>
<td>1300</td>
<td>6.35</td>
<td>4.4</td>
<td>6.74</td>
</tr>
<tr>
<td>1400</td>
<td>6.72</td>
<td>4.64</td>
<td>6.85</td>
</tr>
<tr>
<td>1500</td>
<td>7.06</td>
<td>4.89</td>
<td>6.99</td>
</tr>
<tr>
<td>1600</td>
<td>7.35</td>
<td>5.42</td>
<td>7.83</td>
</tr>
</tbody>
</table>

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.

![Two hops communication path](image)

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 $\mu$ 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. $\mu$ and $\lambda_e$ 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](image)

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 \( \text{Max} (N_{LH}, N_{HS}/2) + 2 \), 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 forwarded. 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 $\mu$ (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 $\mu$ 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 $#P_i$ is the number of tokens in the place $P_i$. The leaving time $\mu$ 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 $\mu$, whereas increasing the size of the intersection area increases it. Section 5.4 derives an expression for the leaving time $\mu$.

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 $\lambda_e$). 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 $\lambda_e$ 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 $k$

<table>
<thead>
<tr>
<th>Arc name</th>
<th>Arc weight function</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{1k}, W_{2k}$</td>
<td>1 IF $#P_{SRk} &gt; 0$ ELSE 0</td>
</tr>
<tr>
<td>$W_{3k}$</td>
<td>$#P_{Lk} 0$ IF $#P_{SRk} &gt; 0$, $\Delta = 1, 2, ..., N_h - 1$ ELSE 0</td>
</tr>
<tr>
<td>$W_{4k}$</td>
<td>1 IF $#P_{SRk} &gt; 0$, $\Delta = 1, 2, ..., N_h - 1$ ELSE 0</td>
</tr>
<tr>
<td>$W_{5k}$</td>
<td>$#P_{Lk} + 1$ IF $#P_{LRk} = 1$ ELSE $#P_{Lk}$</td>
</tr>
</tbody>
</table>

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 $t_{fik}$ is added, which flushes $P_{Lk}$ and $P_{LRk}$ and moves all tokens back to $P_{ik}$ when $#P_{SRx} > 0$, where $x = 1, 2, ..., N_h - 1$. 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 $t_{ik}$. 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 $\left(\text{Max}(N_{LH}, N_{HS}/2) + 2\right)$ 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 $\tau_{LR}$) is:

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

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

The average firing time of the transition $T_{SR}$ ($\tau_{SR}$) 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 $\tau_{RERR}$ and $\tau_{NR}$ are the times required for broadcasting the RERR message and establishing a new route, respectively. Hence,

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

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

In cases where local repair is not supported, the local repairing time is equal to zero ($\tau_{LR} = 0$) and $\tau_{SR}$ 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 \( \mu \)), and the frequency with which the nodes in the network enter into an intersection area (entering rate \( \lambda_e \)). To compute \( \mu \) and \( \lambda_e \), 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:

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

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 ($\mu$) 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 $[V_{\text{min}}, V_{\text{max}}]$. Therefore, the average speed of nodes ($V_a$) is given by [99]:

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

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:

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

(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]:

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

(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:

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

(5.7)

where $V_A$ and $V_B$ are the speeds of node $A$ and $B$ respectively, and $\theta$ is the angle between $V_A$ and $V_B$. The angle $\theta$ can vary from 0 to $\pi$. Because all nodes move with a
speed that is uniformly distributed in the range \([V_{\text{min}}, V_{\text{max}}]\), 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 \(\theta\) is expressed using \(V_r\) and \(V_a\) as:

\[
\theta = 2 \text{ArcSi} n \left( \frac{V_r}{2V_a} \right)
\]

(5.8)

Assuming that \(\theta\) is uniformly distributed in the range \([0, \pi]\), the probability density function of \(\theta\) can be described as follows:

\[
f_{\theta}(\theta) = \frac{1}{\pi}
\]

The probability that \(\theta\) is less than \(\Omega\) 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 \(\theta\) from Equation (5.8) into the last equation as follows:

\[
F_{V_r}(v) = P(V_r \leq v) = \frac{2}{\pi} \text{ArcSi} n \left( \frac{v}{2V_a} \right)
\]

(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):

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

The expected value for \(V_r\) is:
By substituting from Equations (5.6) and (5.10) into Equation (5.5), the expected leaving time is:

$$E(V_r) = \int_0^{2V_a} v f_{V_a}(v) dv = \frac{4V_a}{\pi}$$  \hspace{1cm} (5.10)$$

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 ($\lambda_e$). 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 $\lambda_e$ 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 $\mu$. 

$$\mu = \frac{\pi^2 R}{8V_a} \left[ 1 - \frac{d_r \cdot \sqrt{4R^2 - d_r^2}}{4R^2 \text{ArcCos} \left( \frac{2}{d_r} \right)} \right]$$
The steady state probabilities of the $M/M/1/K$ queue with state dependent service rates are [104]:

$$p_n = \frac{p_0 \omega^n}{n!}$$  \hspace{1cm} (5.11)

$$p_0 + \sum_{n=1}^{K} p_n = 1$$  \hspace{1cm} (5.12)

where $p_0$ and $p_n$ are the probabilities of initial state and state number $n$, respectively, and $\omega = \lambda_0 \cdot \mu$. By substituting from Equation (5.11) into Equation (5.12), we obtain:

$$1 + \sum_{n=1}^{K} \frac{\omega^n}{n!} = \frac{1}{p_0}$$  \hspace{1cm} (5.13)

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

$$E(Q) = \sum_{n=1}^{K} n \cdot p_n = \sum_{n=1}^{K} n \cdot \frac{p_0 \omega^n}{n!}$$

$$= p_0 \omega \sum_{n=1}^{K} \frac{\omega^{n-1}}{(n-1)!} = p_0 \omega \sum_{n=0}^{K-1} \frac{\omega^n}{n!}$$

$$= p_0 \omega \left[ 1 + \sum_{n=1}^{K} \frac{\omega^n}{n!} - \frac{\omega^K}{K!} \right]$$  \hspace{1cm} (5.14)

Substituting Equation (5.13) into (5.14) gives:

$$E(Q) = \omega - p_0 \frac{\omega^K}{K!}$$  \hspace{1cm} (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 ($\omega$), so it can be neglected. Hence, the expected length of the queue can be approximately evaluated to:
To compute the entering rate $\lambda_e$, 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:

$$E(Q) = \omega = \lambda_e \cdot \mu$$

(5.16)

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

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

(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 $\Psi$, and path failure and repair frequency $\mu_s$. 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:

$$\Psi = Pr\left((\#P_{LR1} = 0 & \ldots & \#P_{LRn} = 0) \& (\#P_{SR1} = 0 & \ldots & \#P_{SRn} = 0)\right)$$  \hspace{1cm} (5.19)

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

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

$$\mu_f = A_v \cdot \text{Rate}(T_{pfk})$$  \hspace{1cm} (5.21)

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

where $\mu_f$, $\mu_r$, and $\text{Rate}(T_s)$ are path failure frequency, path repair frequency and firing rate of the transition $T_s$ 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 ($\lambda$) 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).

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, $\lambda = 10$ 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 ($\delta$). 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 $\lambda = 10$ or 40 kbps](image)

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 5.6: Numerical results for different packet generation rates.](image)

Figure 5.6: Path connection availability versus the side length of the network area, where $R = 250$ or 200 m, $N = 100$, and $\lambda = 10$ 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 $N = 100$, $\lambda =10$ 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,
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 $\lambda = 10$ 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, $\lambda = 10$ 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.

![Graph showing path failure and repair frequency versus the side length of the network area](image)

**Figure 5.9:** Path failure and repair frequency versus the side length of the network area, where $R = 250$ m, $N = 100$, and $\lambda = 10$ 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 ($\lambda_r$) 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 ($\lambda$) 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 $\lambda$ 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 ($\lambda_{r}$) and the throughput per path using average node utilisation ($\alpha$).

![Figure 6.1: A network communication path](image-url)
Figure 6.1 shows a communication path with \( N_h \) hops between the source node \( S \) and destination node \( D \), where \( \lambda^x_t \) is the number of packets that are successfully transmitted by any node \( x \) per unit time, \( \lambda^x_p \) is the packet forward rate for a node \( x \), and \( \lambda_1, \lambda_2, \ldots, \) and \( \lambda_{N_h} \) are the number of packets sent by the source \( S \) and received by the nodes \( R_1, R_2, \ldots, \) and \( D \), respectively. For a node \( k \), the node utilisation (\( \alpha_k \)) 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:

\[
\alpha_k = \frac{\lambda^k_t}{\lambda^k_p + \lambda}
\]  

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

\[
\alpha_S = \frac{\lambda^S_t}{\lambda^S_p + \lambda} = \frac{\lambda^S_t - \lambda_1}{\lambda^S_p} = \frac{\lambda_1}{\lambda}
\]

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:

\[
\lambda_1 = \alpha_S \cdot \lambda
\]

For node \( R_1 \), the node utilisation is computed as follows:

\[
\alpha_{R_1} = \frac{\lambda^{R_1}_t}{\lambda^{R_1}_p + \lambda} = \frac{\lambda^{R_1}_t - \lambda_2}{\left(\lambda^{R_1}_p - \lambda_1\right) + \lambda} = \frac{\lambda_2}{\lambda_1}
\]

So, the number of packets sent by \( S \) and successfully received by \( R_2 \) is:

\[
\lambda_2 = \alpha_{R_1} \cdot \lambda_1
\]

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:

\[
\lambda_k = (\alpha_s \cdot \alpha_{R_1} \cdot \alpha_{R_2} \cdot \ldots \cdot \alpha_{R_k}) \cdot \lambda
\]

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

\[
\text{Throughput} = \lambda_{N_h} = (\alpha_s \cdot \alpha_{R_1} \cdot \alpha_{R_2} \cdot \ldots \cdot \alpha_{R_{N_h-1}}) \cdot \lambda \quad (6.4)
\]

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 \( \alpha \). Therefore, Equation (6.4) can be simplified to:

\[
\text{Throughput} = \lambda_{N_h} = \alpha^{N_h} \cdot \lambda \quad (6.5)
\]

If \( \delta \) is the average delay per hop, the end-to-end delay of the path is computed as follows:

\[
\text{End-to-End Delay} = N_h \cdot \delta \quad (6.6)
\]

The average node utilisation \( \alpha \) and delay per hop \( \delta \) 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:

\[
\lambda_x = \lambda_1 + \lambda_2 + \ldots + \lambda_{N_{h-1}}
\]

From Equation (6.5), \( \lambda_x \) can be computed as:

\[
\lambda_x = (\alpha + \alpha^2 + \ldots + \alpha^{N_h-1}) \cdot \lambda \quad (6.7)
\]
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 \( \lambda_c \) in the network can be computed using the following equation:

\[
\lambda_c = \mu_s \cdot N_s \cdot (N + N_h + N_e)
\]

where \( \mu_s \), \( N_s \), 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 \( \mu_s \) 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 \( \lambda_{rt} \) is:

\[
\lambda_{rt} = N_s \cdot \lambda_x + \lambda_c
\]

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

\[
\lambda_r = \frac{\lambda_{rt}}{N} = \frac{N_s \cdot \lambda_x + \lambda_c}{N}
\]

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:

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

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:

\[
R_i(x) = \frac{k}{\sqrt{T_{SNR}}} \cdot x
\]  

(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 \( f_{r_x}(x) \) is given by:

\[
f_{r_x}(x) = 2N_r \cdot \frac{x^{2N_r-1}}{R^{2N_r}}
\]  

(6.12)

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

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

(6.13)

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

\[
P1(R_1, R_2, r_x) = R_1^2 \cdot \text{arccos} \left( \frac{r_x^2 + R_1^2 - R_2^2}{2 \cdot r_x \cdot R_1} \right)
\]

\[
P2(R_1, R_2, r_x) = R_2^2 \cdot \text{arccos} \left( \frac{r_x^2 + R_2^2 - R_1^2}{2 \cdot r_x \cdot R_2} \right)
\]

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

\[
A_{int}(R_1, R_2, r_x) = P1(R_1, R_2, r_x) + P2(R_1, R_2, r_x) + P3(R_1, R_2, r_x)
\]  

(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:

\[
A_h(x) = \pi \cdot (R_l(x))^2 - A_{int}(R_{cs}, R_l(x), x)
\]

(6.16)

\[
A_i(x) = A_{int}(R_{cs}, R_l(x), x)
\]

(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:
where $N_i$ 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, $N_i$ 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 used.
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 $\lambda_n$ (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 frame 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.

![Diagram](https://example.com/diagram.png)

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 $\rho_{BA}$, the firing probabilities of $T_{DE}$ ($T_{AE}$) and $T_{DNE}$ ($T_{ANE}$) are ($\rho_{BA}$) and ($1 - \rho_{BA}$), respectively. The parameter $\rho_{BA}$ is computed from the abstract model shown in Figure 6.4 using the following equations:

$$p_{BA} = 1 - (1 - Pr(#_txD > 0 \ OR #_P_{ACK} > 0))^N_i$$

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

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 $T_{idle1}$ ($\beta_{BA}$) and $T_{busy1}$ ($1 - \beta_{BA}$) 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 ($\mu_B$) and ($1 - \mu_B$) respectively, where $\mu_B$ is the probability of failure to transmit the data frame due to interference induced by interfering nodes. $\mu_B$ 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, $T_{\text{succ}}$ fires and deposits a token in $P_{\text{REC}}$; otherwise $T_{\text{fail}}$ 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 $T_{\text{RNE}}$. Otherwise, the transition $T_{\text{RE}}$ fires representing the receiving of the RTS frame with errors. $P_{\text{CTS}}$ and $T_{\text{CTS}}$ 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 $T_{\text{CNE}}$ which moves the token from $P_{\text{CEC}}$ to $P_{\text{txD}}$. 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 $T_{\text{txD}}$ and $T_{\text{DNE}}$ 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.

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

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

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

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

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

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 $T_{s1}$ ($\mu_R$) and $T_{f1}$ ($1 - \mu_R$) represent the probability of success and failure in completing the RTS/CTS handshake respectively, where $\mu_R$ is computed from the abstract model shown in Figure 6.6 as:

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

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

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

The firing probability of the transition $T_{fail}$ ($\mu_C$) is the probability of the failure to transmit the RTS frame due to interference induced by interfering nodes, where $\mu_C$ 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
\( T_{\text{ANE}} \) returns the token back to the place \( P_B \), and the firing of the transition \( T_{\text{timeout}} \) models the CTS or ACK frame timeout. Thus, the average firing time of \( T_{\text{timeout}} \) depends on the number of tokens in \( P_{\text{fail}} \). If \( \#P_{\text{fail}} = 2N_i \), \( T_{\text{timeout}} \) represents the ACK frame timeout interval; otherwise, it represents the CTS frame timeout interval. If \( \#P_{\text{fail}} = 2N_i \), the firing of \( T_{\text{timeout}} \) removes all tokens in \( P_{\text{fail}} \) and deposits one token in \( P_{\text{DIFS}} \); otherwise it moves all tokens from \( P_{\text{fail}} \) to \( P_{\text{DIFS}} \). 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 \( T_{\text{GP}} \) models the generation of packets in the transport layer. The firing of the transition \( T_{\text{GP}} \) deposits a token in the place \( P_{\text{GP}} \). The mean firing time of \( T_{\text{GP}} \) is the mean time of the generation of UDP packets in transport layer. The place \( P_{\text{Buffer}} \) contains tokens corresponding to free buffer spaces in the current node. The initial number of tokens in \( P_{\text{Buffer}} (N_B) \) 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.

![Network layer model diagram]

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

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

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:

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

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

where $n_c$ is the average number of attempts to transmit the MAC frame, which is computed as $n_c = M(P_{FC})$, and $\delta_d$ 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
$T_{Fail} (\mu_f)$ and $T_{Repair} (\mu_r)$ 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 inhibitor
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 ($\lambda_r$). Section
6.2 derives an expression for $\lambda_r$. 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 $\psi$ is the average probability that any path in the network is available, the firing probabilities of the transitions $T_{YPR}$ and $T_{YPS}$ are $\psi$, whereas the firing probabilities of transitions $T_{NPR}$ and $T_{NPS}$ are $(1 - \psi)$. The probability of path availability $\psi$ 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 \((\varepsilon)\), packet delay in the data link layer \((\delta_d)\) and packet delay per hop \((\delta)\) 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 \((\psi)\), the path failure rate \((\mu_f)\), and the path repairing rate \((\mu_r)\), which are all required in order to solve the network layer model. At the end of the first iteration, using \(\varepsilon\) and \(\delta_d\) computed using data link layer model and \(\psi, \mu_f\) and \(\mu_r\) computed using the path analysis model, the network layer model can be solved to compute the node utilisation \((\alpha)\) and network layer throughput \((\lambda_n)\) 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 \(\alpha\) or \(\lambda_n\). 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 ($\rho_R, \omega_R$ and $\alpha$) are considered to be equal to 0.5.

**Step 2:** The data link layer model is solved to compute $\varepsilon$, $\delta$ and $\delta_d$ 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:

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

**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 $\delta_d^{0}$ (or any other performance metric) and the parameters $\mu_R, \mu_C, \beta_R, \rho_R$, and $\omega_R$ for the RTS/CTS method ($\mu_B, \beta_{BA}$, and $\rho_{BA}$ for the BA method).

**Step 2.3:** The one node detailed model is solved using the last computed values of parameters $\mu_R, \mu_C, \beta_R, \rho_R$, and $\omega_R$ to compute $\varepsilon$ 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 $\delta_d^{m}$ and the other parameters $\mu_R, \mu_C, \beta_R, \rho_R$, and $\omega_R$.

**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 $\delta$ (or $\delta_d$ in the case of $n = 1$), the path analysis model is solved to compute $\psi, \mu_f$ and $\mu_r$. 
**Step 4:** Using the last computed values for $\alpha$, $\epsilon$, and $\delta_d$, the network layer model is solved to compute new values for $\alpha$ and $\lambda_n$. Also, any of the performance metrics $\tau^n$, 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 $\alpha$ and delay per hop $\delta$ 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

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>60, 80, …., 240</td>
</tr>
<tr>
<td>Side length of the network area</td>
<td>600, 800, 1000 m</td>
</tr>
<tr>
<td>Packet size</td>
<td>2, 6 kB</td>
</tr>
<tr>
<td>Packet generation rate</td>
<td>100, 200, …. , 2200 kB</td>
</tr>
<tr>
<td>Queue Length</td>
<td>30</td>
</tr>
<tr>
<td>Transmission range</td>
<td>150, 250 m</td>
</tr>
<tr>
<td>Carrier sensing range</td>
<td>150, 250, 350, 450 m</td>
</tr>
<tr>
<td>Routing protocol</td>
<td>AODV</td>
</tr>
<tr>
<td>Pause time</td>
<td>0 sec</td>
</tr>
<tr>
<td>Maximum speed of nodes</td>
<td>20 m/s</td>
</tr>
<tr>
<td>Antenna type</td>
<td>Omni-directional</td>
</tr>
<tr>
<td>Propagation path loss model</td>
<td>Two-ray ground</td>
</tr>
<tr>
<td>Simulation time</td>
<td>1100 sec</td>
</tr>
</tbody>
</table>
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, $R_{cs} = 150$ or 250 m, $L = 600$ m or 1000 m, and $R = 150$ m.
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 \( R_{cs} = 150 \) 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](image)
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

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

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 multi-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


