Performance metrics and routing in vehicular ad hoc networks

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University.

Metadata Record: [https://dspace.lboro.ac.uk/2134/9136](https://dspace.lboro.ac.uk/2134/9136)

Publisher: © by Zulkefli M Yusof

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
Performance Metrics and Routing in Vehicular Ad hoc Networks

By

Zulkefli Muhammed Yusof

A Doctoral Thesis submitted in partial fulfilment of the requirements for the award of Doctor Philosophy of Loughborough University

November 2011

© by Zulkefli M Yusof 2011
To my mum, wife and children
ABSTRACTS

The aim of this thesis is to propose a method for enhancing the performance of Vehicular Ad hoc Networks (VANETs). The focus is on a routing protocol where performance metrics are used to inform the routing decisions made. The thesis begins by analysing routing protocols in a random mobility scenario with a wide range of node densities. A Cellular Automata algorithm is subsequently applied in order to create a mobility model of a highway, and wide range of density and transmission range are tested. Performance metrics are introduced to assist the prediction of likely route failure. The Good Link Availability (GLA) and Good Route Availability (GRA) metrics are proposed which can be used for a pre-emptive action that has the potential to give better performance. The implementation framework for this method using the AODV routing protocol is also discussed. The main outcomes of this research can be summarised as identifying and formulating methods for pre-emptive actions using a Cellular Automata with NS-2 to simulate VANETs, and the implementation method within the AODV routing protocol.
ACKNOWLEDGEMENTS

All praise to Allah S.W.T. for that His willing has allowed me to complete the research and produce this thesis.

First and foremost I would like to give my utmost gratitude and thanks to my primary supervisor, Dr James Flint who has provided me the support, knowledge and guidance which enable me to complete my research. I would also like to make a special mention of my other supervisor now retired Dr Sekharjit Datta, whom with his vast experience, expertise and advice have successfully guided me throughout the studies.

Thanks must also be passed on to fellow research students who have assisted me with various aspects of the studies, especially to Simon Dible.

On a personal note, I would like to thank my mum and siblings for their support and prayers. My greatest appreciation and thanks go to my wife Hasruniza, children Naim, Yasmin, and Najmi who are the pillar of my life - for their incomparable love, encouragement and support.

Final and important acknowledgement must go to the Dean for the Kulliyah of Information and Communication Technology, International Islamic University, Prof. Dr. Adam Suhaimi, and the Government of Malaysia for the support and sponsorship of this PhD study.
TABLE OF CONTENTS

CHAPTER 1: .............................................................................................................. 1
INTRODUCTION .................................................................................................. 1
1.1 Research Background ................................................................................. 1
1.2 Objectives .................................................................................................. 2
1.3 Research Methodology .............................................................................. 3
1.4 Thesis Overview .......................................................................................... 4
1.5 Thesis Contributions .................................................................................... 5

References ........................................................................................................... 6

CHAPTER 2: .......................................................................................................... 7
VEHICULAR AD HOC NETWORKs (VANETs) AND PROTOCOLs .................. 7
2.1 The Functional Layer of the OSI Models and Ad hoc Networks ................. 7
2.2 Wireless Ad Hoc Networks ......................................................................... 10
2.3 Mobile Ad hoc Networks (MANETs) ........................................................ 12
2.4 Vehicular Ad Hoc Networks (VANETs) ..................................................... 15
2.5 Transport Layer Protocols ......................................................................... 17
2.5.1 UDP Fundamentals .............................................................................. 17
2.5.2 TCP Operation ..................................................................................... 17
2.6 File Transfer Protocol (FTP) Application ................................................. 24
2.7 Routing Protocols in MANETs ................................................................. 25
2.7.1 DSDV .................................................................................................. 25
2.7.2 DSR .................................................................................................... 27
2.7.3 AODV .................................................................................................. 28
2.7.3.1 AODV Control Packets and the Routing Table ............................... 30
2.7.4 Other Routing Protocols ...................................................................... 33
2.7.5 Routing Protocols Summary ................................................................. 34
2.8 Transmission Range and VANETs ........................................................... 35
2.9 Performance Metrics ............................................................................... 36
2.9.1 Field Measurement ............................................................................. 37
2.9.2 Simulation ............................................................................................ 37
2.10 Discussions and Summary ....................................................................... 38
5.4 Link Availability................................. ................................................... .................... 88
5.5 Route Availability.................................. ................................................... .................... 89
5.6 Proposed Approach: Good Link and Route Availability................................. 90
  5.6.1 Location Prediction.................................. ................................................... ............... 90
  5.6.2 Good Link Availability.................................. ................................................... ............ 94
  5.6.3 Good Route Availability.................................. ................................................... ........... 95
5.7 Discussions and Conclusions.................................. ................................................... ........... 97
  5.7.1 Discussions ....................................... ................................................... ................... 97
  5.7.2 Conclusion ........................................ ................................................... ................... 98
References:.................................................. ................................................... ....................... 99

CHAPTER  6:.................................................. ................................................... ............... 102
MODELLING AND EVALUATION OF PERFORMANCE METRICS ........................................ 102
  6.1 Introduction ...................................... ................................................... ..................... 102
  6.2 Simulation Set up and Parameters ................................................... ....... 103
  6.3 Re-routing: Measurement and Analysis ................................................... 104
  6.4 Selecting a Path for Re-route ................................................... ............ 105
  6.5 Implementation Framework ................................................... ........ 107
    6.5.1 Main Processes .................................. ................................................... .............. 110
  6.6 Discussions and Conclusions.................................. ................................................... ......... 114

CHAPTER  7:.................................................. ................................................... ............... 115
CONCLUSION AND FURTHER WORK.................................................. .................................................. 115
  7.1 Contribution of this thesis.................................. ................................................... .............. 115
    7.1.1 Identified link and route availability as the performance metrics ......... 115
    7.1.2 Provide framework for the implementation ................................................ 116
    7.1.3 Develop a mobility model using Cellular Automata (CA) for a highway scenario..................................................... 116
    7.1.4 Approach in comparing MANET and VANET ....................................... 116
  7.2 Suggestions for further work ................................................... ........... 117
    7.2.1 Development work .................................. ................................................... .............. 117
    7.2.2 Test bed .......................................... ................................................... .................... 117
    7.2.3 Further effects from density and transmission range ............................................ 118
    7.2.4 Enhancement of Mobility model ................................................... ........... 118
  7.3 Closing remarks ................................... ................................................... .................. 118

Page viii
## LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACK</td>
<td>Acknowledgement</td>
</tr>
<tr>
<td>AGT</td>
<td>Agent</td>
</tr>
<tr>
<td>AODV</td>
<td>Ad hoc On-Demand Distance-Vector</td>
</tr>
<tr>
<td>ARQ</td>
<td>Automatic repeat request</td>
</tr>
<tr>
<td>BER</td>
<td>Bit error rate</td>
</tr>
<tr>
<td>CA</td>
<td>Cellular Automata</td>
</tr>
<tr>
<td>CBR</td>
<td>Continuous Bit Rate</td>
</tr>
<tr>
<td>DSDV</td>
<td>Destination-sequenced Distance Vector</td>
</tr>
<tr>
<td>DSR</td>
<td>Dynamic Source Routing</td>
</tr>
<tr>
<td>DSRC</td>
<td>Dedicated Short Range Communications</td>
</tr>
<tr>
<td>FRIIS</td>
<td>Free Space</td>
</tr>
<tr>
<td>FTP</td>
<td>File transfer protocol</td>
</tr>
<tr>
<td>GLA</td>
<td>Good Link Availability</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>GRA</td>
<td>Good Route Availability</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>LOS</td>
<td>Line of Sight</td>
</tr>
<tr>
<td>MAC</td>
<td>Media Access Control</td>
</tr>
<tr>
<td>MANETs</td>
<td>Mobile Ad hoc Networks</td>
</tr>
<tr>
<td>NLOS</td>
<td>No Line of Sight</td>
</tr>
<tr>
<td>NS2</td>
<td>Network Simulator 2</td>
</tr>
<tr>
<td>OSI</td>
<td>Open System Interconnection</td>
</tr>
<tr>
<td>RERR</td>
<td>Route Error</td>
</tr>
<tr>
<td>RMEA</td>
<td>Route Measurement</td>
</tr>
<tr>
<td>RREP</td>
<td>Route Reply Packet</td>
</tr>
<tr>
<td>RREQ</td>
<td>Route Request</td>
</tr>
<tr>
<td>RRP</td>
<td>Route Request Packet</td>
</tr>
<tr>
<td>RSSI</td>
<td>Received signal strength indication</td>
</tr>
<tr>
<td>RTR</td>
<td>Router</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>TTL</td>
<td>Time-to-live</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>VANETs</td>
<td>Vehicular Ad hoc Networks</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Research Topic Sequence Overview</td>
<td>2</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>The OSI Model</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Schematic Diagram of a MANET</td>
<td>14</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Simple simulation with 3 nodes</td>
<td>32</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>NS2 Simulation Parameter Definitions</td>
<td>52</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>The NS-2 Wireless Trace File</td>
<td>53</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Sample of cars positions during simulations according to lanes</td>
<td>61</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Initial node positions before the simulation starts</td>
<td>62</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>CBR delivery rate for various speeds and number of nodes</td>
<td>71</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>CBR packets delivery rate for various scenarios using the Rayleigh Fading Model</td>
<td>74</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Routing and MAC overhead for CBR traffic</td>
<td>75</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Routing and MAC Overhead for TCP traffic</td>
<td>75</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Routing and MAC Overhead for CBR traffic with Fading</td>
<td>76</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>Effect of transmission range on packet delivery</td>
<td>79</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>Dropped packets against all packets</td>
<td>80</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Packet delivered and received ratio</td>
<td>81</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>Overhead packets vs. received packets ratio</td>
<td>82</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Performance metrics</td>
<td>90</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Projection of node movement</td>
<td>91</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Comparison overview between calculated and simulation of the time the first packet dropped</td>
<td>94</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Routing example to illustrate Good Route Availability</td>
<td>96</td>
</tr>
<tr>
<td>Figure 6.1</td>
<td>AODV Process Flow Chart</td>
<td>108</td>
</tr>
<tr>
<td>Figure 6.2</td>
<td>Scenario illustrating the AODV process</td>
<td>109</td>
</tr>
<tr>
<td>Figure 6.3</td>
<td>RMEA packet format</td>
<td>111</td>
</tr>
<tr>
<td>Figure 6.4</td>
<td>RMEA packet process from destination to source node</td>
<td>111</td>
</tr>
<tr>
<td>Figure 6.5</td>
<td>A comparison of the standard and new process</td>
<td>112</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>An extract of Routing Tables</th>
<th>33</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.2</td>
<td>MANET Routing Protocol Summary</td>
<td>34</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Comparison between OPNET Modeller and NS-2</td>
<td>50</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>Simulation Details</td>
<td>63</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Scenario Details</td>
<td>70</td>
</tr>
<tr>
<td>Table 4.2</td>
<td>TCP Packet Delivery Rate for Various Speed and Nodes</td>
<td>73</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Comparison details between calculated and simulation of the time the first packet dropped</td>
<td>93</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>Examples of Calculating Good Route Availability</td>
<td>96</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Simulation Parameters</td>
<td>103</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Re-routing Measurements</td>
<td>104</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Route selection comparison</td>
<td>106</td>
</tr>
</tbody>
</table>
AUTHOR’S PUBLICATIONS

A number of publications have resulted from the work in this thesis. These are as follows:


CHAPTER 1:
INTRODUCTION

Mobile Ad hoc Networks (MANETs) is a research area which addresses the issues of wireless network access without a base station. With applications extending to vehicular scenarios known as Vehicular Ad hoc Networks (VANETs), the pertinent problems generally centre on the effects of node mobility which leads to the issue of link breakages eventually degrading the performance of the network. This thesis focuses on the maintenance of links in an active route or path by predicting the link and route availability which can be used proactively, executing a reroute in the case of possible failure. The objective is to improve the network performance of protocols such as Transmission Control Protocol (TCP)-based connections in a highway scenario.

1.1 Research Background

MANETs have unique characteristics associated only with them such as dynamic topologies, decentralized operation, and particular bandwidth constraints [1.1]. MANETs are also subject to issues which relate to their characteristics [1.2] such as high bit error rates, the need to rapidly recalculate the routes, network partitioning problems and MAC failure.

The consideration of TCP traffic in this thesis is mainly motivated by the common usage of this protocol and the fact that it is subject to more complex and interesting behaviour when compared to User Datagram Protocol (UDP) traffic, which is a simpler broadcast method [1.3, 1.4]. Also taken into consideration are the applications applied in VANETs, presented in Chapter 2.

In VANETs, communication between a sending and receiving node would usually involve routing through several nodes or multihops. Ideally each link in the path should be in a stable or
‘good’ condition to ensure smooth transmission and acceptable throughput. This motivates the research on this thesis. The overview of the sequence from the broad scope narrowed to the research area is depicted in Figure 1.1.

![Research Topic Sequence Overview](image)

Note that the shaded boxes show the focus area in this thesis. From the general topic of wireless networks, the research is narrowed to consideration of the link and route availability as the performance metrics in VANETs with TCP as the transport control protocol.

### 1.2 Objectives

The objectives of this research are as follow:

i. to explore the protocols that can handle a robust MANET environment and to justify the choice of a routing protocol for the remainder of the thesis,

ii. to develop a novel mobility model representing a highway scenario with its own characteristic which is different from the randomly moving nodes found in the general theory of MANETs,
iii. to develop a probabilistic model which can predict link breakage using transmission range and calculations of the future position of each node to allow proactive corrective actions to be taken,

iv. to produce a novel reroute method based on the probabilistic value of the link and route reliability so that the transmission can remain uninterrupted, hence reducing dropped packets and improving performance,

v. to review the literature and discuss the use of TCP with a focus on the performance issues in a vehicular environment,

vi. to discuss and justify the suitability of simulation software in providing accurate results in the study of VANETs in a highway scenario.

1.3 Research Methodology

The research begins by reviewing proactive and reactive routing protocols, and presenting the most suitable protocols for dynamic networks. Computer simulations are used with a wide range of parameters on the movement, density and speed. Subsequently the thesis presents a mobility model to reflect the highway scenario, which is the main focus of this thesis.

The highway model, along with the routing protocol is then used as a starting place for development. The strategy is to ensure that each link in a route or path is maintained as long as possible in order that the packets can be safely sent. It is assumed that the distance, speed, and transmission range distance information is available or can be derived from the node or neighbour table in the case of the simulation software used in this thesis. A method is proposed to calculate and predict when the link would finally become disconnected so that preemptive rerouting action can be taken.
1.4 Thesis Overview

Chapter 2 provides a literature survey of the different elements of this study, including the functional layers of the OSI model, routing protocols and performance metrics. The chapter also discusses TCP, giving a brief introduction to its operating fundamentals and variants. It also relates the current literature of the MANET and VANET scenario with the aim to support the main ideas in the thesis.

Chapter 3 sees the development of a highway mobility model using the Cellular Automata (CA) method. Discussions on simulation software and specifying the measurement parameters are also included.

Chapter 4 presents a series of simulations which illustrates the effect of transmission range and the performance analysis demonstrates how each of the protocols reacts when different node densities are applied. This exercise has also shown the suitability of the proposed model to represent the VANETs scenario. The Ad hoc On-Demand Distance-Vector (AODV) protocol has been chosen for the subsequent work in this thesis based on the justifications from the analysis. VANET simulation is discussed with reference to the NS-2 program [1.5, 1.6] and the pertinent characteristics, constraints and parameters are discussed.

Chapter 5 presents the proposed performance metrics in theory involving the availability and reliability of links and routes.

Chapter 6 focuses on the implementation of the algorithms and in addition provides further simulation work for the methods described in chapter 5.

Finally, chapter 7 concludes the thesis with discussions on what has been achieved and also proposes future works.
1.5 Thesis Contributions

The summary of the contributions from this thesis are:

i. Performance metrics have been proposed to assist the prediction of likely route failure. This would benefit the pre-emptive actions taken by routing protocols.

ii. Demonstrating that Cellular Automata (CA) Models can be used within NS-2 and applied to VANETs

iii. Implementation of the AODV routing protocol using the proposed metrics

Further discussions on the contributions can be seen in the concluding chapter
References


CHAPTER 2:

VEHICULAR AD HOC NETWORKs (VANETs) AND PROTOCOLS

This chapter discusses the context of the network components involved in the research. The OSI model is introduced and its role is explained especially in the wireless environment such as in VANETs. The purpose of the chapter is to understand the functions of each layer and specify which layers are involved in the research. Wireless, MANET, and VANET are the network types introduced here, along with the Transport Protocols, File Transfer Protocol applications. Specific information on VANETs Routing Protocols, transmission range and measurements are also discussed.

2.1 The Functional Layer of the OSI Models and Ad hoc Networks

The OSI model [2.1] is a layered framework produced by the International Standards Organization (ISO), designed to allow communications between all types of computer systems. It consists of interactive modules each of which provides a specific function. The focus here is to provide an understanding on how it operates on its respective layers. The layers are shown in Figure 2.1.

The first layer of the OSI model is the Physical Layer which deals with the medium specifications for the data transmission between devices. It specifies the parameters, among others, like operating frequency, the transmission range, data rate, channelling scheme, timing, and power requirements. It works closely with the MAC sub layer to ensure smooth performance of the network.
In a wireless network, there are additional considerations such as the fact that the medium is error-prone due to interference from other wireless and RF systems in proximity. Multipath effects can also be problematic as they change dynamically with time, as can frequent disconnections of wireless links. The layer specifications depend on the wireless system used such as IEEE 802.11 which consists of sets of protocols for Wireless Local Area Networks (WLANs), and the IEEE 802.15.3 for Wireless Personal Area Networks (PANs).

The Data Link Layer is the second layer consists of logical link control (LLC) and medium access control (MAC) sublayers. LLC is responsible for maintaining the link, framing data unit, synchronization, error detection, and recovery and flow control while MAC looks into the channel access which tries to gain access to the shared channel so that the transmitted frame will not collide and get distorted with other frames from other nodes sharing the same channel. Some protocols are designed to have a controller while others are distributed in nature. The former is a reservation-based protocol which has a controller that allocates time slots or frequencies to modes requesting channel access while the latter is purely based on contention or non reservation-based, which is more suitable for MANETs where it is impossible to designate a
leader that might be moving all the time. Carrier Sense Multiple Access (CSMA) is used where the MAC senses the carrier for any other traffic and sends the frame immediately if the medium is free. When a channel is busy the MAC layer waits for a random time period which grows exponentially and then tries to resend the frame after sending the medium. In a MANET, this method has its weaknesses as in situations that create the Hidden Terminal and Exposed Terminal problems [2.2].

The third layer is the Network layer which in the case of MANET has two main challenges which are finding a route from the source to destination node and maintaining routes when there is at least one session during the route. There are three types of algorithms in MANETs which are the reactive, proactive, and hybrid where each of them possesses its own strengths and weaknesses. In a MANET situation, certain routes are more stable than others because links on those routes are not severed for a period of time. The routing algorithm needs to determine the routes. Generally, each node advertises its presence to its neighbours periodically. In the case of determining the stability of a link, Associativity-based Routing (ABR) [2.3] and Signal-Stability-Based Adaptive Routing (SSA) [2.4] are two good examples where in ABR the nodes receiving the advertisement increment a counter associated with the sending node and the degree of stability is based on the value of the counter. While in SSA the concept is similar but instead based on the relative signal strength between nodes.

Next is the Transport Layer which supports the integrity of data packets from the source to the destination node and can be either connection-oriented as in TCP, or connectionless as in UDP. It ensures reliable delivery by the use of acknowledgements and retransmissions, and also sequencing of data packets. Congestion avoidance and controls are also taken in here. The throughput is increased if acknowledgement arrives within time and if late it assumes that the network is overloaded and reduces the throughput to avoid congesting the network. In MANETs, the mobility of nodes has the high tendency to cause the packets to be out of order and a significant delay in the acknowledgements. Packet losses could also possibly caused by errors in the wireless channel.
The final layers are the Session, Presentation, and Application layers which provide network access to applications and protocols commonly used by end users such as FTP, SMTP, HTTP and so on. In the case of MANETs, it is also responsible for providing location-based services.

The main work in this thesis focuses on the Network layer which is where the routing functions are handled. The following sections will discuss the types of networks involved in the thesis.

2.2 Wireless Ad Hoc Networks

As in any wireless networks, there are distinctive differences between wireless and wired services. Some of these are listed below:

i. Address may not always refer to particular geographical location because a node does not need to be stationary

ii. Restricted connectivity since mobile nodes may go out of range of each other in a dynamic topology

iii. The exact range of wireless coverage cannot be determined accurately as it would depend on various factors such as the signal strength and noise levels

iv. It is an error prone medium because transmissions by a wireless node are affected by simultaneous transmission by neighbouring nodes that are located within direct transmission range of the transmitting node. As a comparison, typical bit error rates are $10^{-4}$ in a wireless channel against $10^{-9}$ in fibre optic cables [2.5].

Due to the constraints of wireless networks, there is a need to examine the design consideration in the offering of wireless services. Some of the factors are listed below:

i. Able to support communications with quality approximating that of wired networks for various types of information such as speech, text, data, images, and video

ii. Capable of expanding its geographical coverage

iii. Scalable by allowing virtually limitless growth in the number of users

iv. Support moving nodes as in this case of the thesis at vehicular speed [2.6]
The wireless devices need to be able to work with each other and this requires some standards that can be adhered to. At the time of writing this thesis, the two most commonly used standards used by industry are IEEE 802.11x for shorter distance and IEEE 802.16 for a wider area [2.7].

Other technologies are also briefly discussed in the following paragraphs.

**IEEE 802.11** is a technology that refers to the family of IEEE standards on which most of the wireless local networks are built. The main standards under this are 802.11a which offers higher speed, smaller coverage and more channels; 802.11b where most of the chips used in the earlier devices were installed with but still available in hot spots for Internet access; 802.11g for higher data rates using 2.4 GHz bands.

**Bluetooth** [2.8] applications are simple to use and supported by many devices. It usually allows short distance communication such as accessories connections to devices like headsets and keyboards to mobile phone, notebooks, PDAs, etc. It falls under the IEEE 802.15 Personal Area Network (PAN) standard.

**Radio frequency Identification (RFID)** usually comes with inexpensive antennas and is widely used in automated toll collections and also for tracking purposes such as inventory control, livestock, and for tracking people.

**Ultra-wideband (UWB)** supports transmission over a large bandwidth of more than 500 MHz which is a much wider frequency range than the conventional systems and results in having much higher data rates than RFID. Due to its high data rate and short distance it is suitable for wireless monitors, transferring data from digital camcorders, direct wireless printing and file transfers without the need for intervening devices. UWB is used as a part of location systems and real time location detection such as finding people trapped under rubble. It is also good for use in sensitive environments such as hospitals due to its precision capabilities combined with very low power consumption.
**Zigbee** is a protocol that currently operates on slower technology than UWB. It is meant more for RF applications that require low data rate and long battery life like in the sensor networks and automation in the home.

**IEEE 802.16** is a standard for the WiMAX service [2.9], also termed Broadband Wireless Access which provides fixed and fully mobile internet access. As in the other standards, it also comes with alphabets each having its own features. 802.16a is intended to be used in carriers’ backhaul networks intermediate links between the network backbone and the small sub-network at the edge of the entire network. 802.16d is an enhancement of 802.16a meant for fixed wireless applications and lower-cost applications. 802.16e is a proposed standard for mobile applications in licensed frequency bands and supports handling over traffic between base stations for mobility. WiMAX can provide broadband wireless access up to 50 km for fixed stations, and up to 15 km for mobile stations. In contrast, the Wi-Fi/802.11 wireless local area network standard is limited in most cases to only 100 - 300 feet (30 - 100m).

The standards are available to enable interoperability of devices produced by different manufacturers. They also reflect the various types of applications with each requires one particular standards that fits the environment. For example, Wi-Fi devices are relatively cheap and viable for mobile type of network such as MANET but require highly powered devices for applications where the environment requires higher transmission range. On the other hand, WiMax infrastructure has the potential of supporting VANETs applications due to its broader coverage and higher bandwidth.

### 2.3 Mobile Ad hoc Networks (MANETs)

Mobile Ad hoc Networks (MANETs) are defined as a collection of mobile platforms or nodes where each node is free to move about arbitrarily [2.10]. This relatively new term introduces the concept of mobile packet radio networks where every node in the network is mobile and ‘store-and-forward’ wireless multihop routing is utilized. It describes distributed, mobile, wireless, multihop networks that operate without any existing infrastructure except for the nodes themselves.
All nodes in MANETs behave as routers and take part in discovery and maintenance of routes to other nodes in the network using ad hoc routing protocols. An ad hoc protocol is a convention or standard that controls how nodes come to agree which way packets are routed between computing devices in a MANET. In an ad hoc network, nodes do not have prior knowledge of the topology of network around them. Instead, they have to discover the properties of their own local network by announcing their presence and then listening to broadcast announcements from their own neighbours. In turn, each node is able to communicate its own connectivity to its neighbours and ultimately it becomes possible to communicate over large distances, far beyond the signal range of an individual node.

Initially developed for military use [2.11], MANETs now have numerous civil applications due to the advance in use of mobile personal communication devices and Global Positioning Systems (GPS). Their development has been rapid with the standardization by the MANET working group which is a part of the Internet Engineering Task Force (IETF) [2.12]. The words ad hoc themselves imply self-organizing or adaptive which means nodes should be able to detect the presence of other similar devices and perform the necessary handshaking to allow communications, sharing of information and services. In MANETs, the distance between nodes can sometimes exceed the transmission range. In another situation, at any one time two nodes can be within range initially, and can be instantly be out of range in moments later. Regardless of the situation, nodes must always try to ‘assist’ each other in the process of delivering packet data by establishing re-establishing links.

Figure 2.2 shows an example network. In this MANET, the route consists of the source (N1), destination (N4) and intermediate nodes (N2, N3). Movements by any of these nodes will affect the validity of routes. N1 has a downstream link, and when it moves out of its downstream neighbours’ N2 and N6 radio coverage, the existing route immediately becomes invalid.
All subsequent downstream nodes may also have to be informed so that they can erase invalid routes from their routing tables. The same situation applies with the destination node where the upstream route becomes invalid and consequently intermediate nodes need to be informed. Similarly, any movements of intermediate nodes supporting an existing route may cause the route to be invalidated.

There are numerous examples of applications for MANETs. Just a few examples of those proposed to date are given here. A very simple scenario is the use by conference delegates or meeting participants of MANET principles. Here there are a number of nodes, each participant may own a laptop computer mobile and they are likely to be close together. It is perfectly feasible to create an ad-hoc network to provide connectivity with minimal fixed network support. The use of MANETs in the battlefield can also see the existence of non-homogenous collections of hosts such as the tanks, infantry, ground assault soldiers and transport vehicles. In theory at least there is no reason why an ad-hoc network could not be established between them, and clearly they could benefit since there is unlikely to be any established infrastructure where they operate. The vehicular environment, which is the prime focus of this thesis, offers some very attractive applications. Examples of vehicle-deployed systems utilising ad-hoc networks include the possibility to provide collision warning (by communicating the occurrence of accidents directly to other vehicles), road sign alarms, and enhanced navigational aids that can make use of data from other vehicles in order to select the optimum road routes. In vehicle applications it has
to be assumed, in simulation at least, that each vehicle is equipped with devices that can receive and relay messages in the network. In reality, deployment of such a system would require widespread adoption by drivers, so there are some obstacles to its application.

2.4 Vehicular Ad Hoc Networks (VANETs)

Vehicular Ad hoc Networks are in a specific class or subset of MANETs. These are characterized by being self organizing communication networks built up by moving vehicles with very high node mobility and limited degrees of freedom in the mobility process. In other words, rather than moving completely at random, vehicles tend to move in an organized fashion, though some random processes do take place. Most vehicles are restricted in their range of motion, for example by being constrained to follow single or multi-lane highways.

2.4.1 VANET Applications

Initial applications of VANETs were mainly in Dedicated Short Range Communications (DSRC) systems. DSRC is a short to medium range wireless protocol specifically designed for automotive use such as electronic toll collection. Currently there are a broad range of information and electrical technologies which are based on wireless and wired communications. The applications of wireless technologies are more readily available and explored in the market such as Intelligent Transportation Systems (ITS) applications [2.13]. When integrated into the transportation system's infrastructure, and in vehicles themselves, these technologies have the potential to relieve congestion, improve safety and enhance productivity [2.14, 2.15].

The use of Floating Car Data (FCD) technology [2.16] also provides an advantage over the need to use or rely only on GPS devices. Virtually every car would at least have a mobile telephone and as they move all these mobile phones will act as anonymous traffic probes by routinely transmitting their location information to the network even when no voice connection is established. By measuring and analyzing the triangulation of network data, accurate traffic flow
information can be calculated. No infrastructure needs to be built along the road as only the mobile phone network is being used.

There are also efforts to standardize VANETs by the Car-to-Car Communication Consortium (C2C-CC) [2.17] which aims to ensure interoperability on network side in the way that data is transmitted and forwarded, and also ensure that nodes become cooperative by following a specific protocol.

An educational project called Programmable Open Mobile Internet [2.18] has been carried out to reach the poor or underserved children living in areas where infrastructures for schools are lacking. The children use devices called PocketSchool Interactive Learning Ad hoc Network (PSILAN) which allow them to easily connect for group activities.

Recent projects have also involved large companies such as the NEC Corporation in collaboration with vehicle manufacturers DaimlerChrysler and BMW [2.19]. A prototype system has been developed that turns each car into a communication hub able to transmit safety information to nearby vehicles based on sensor data such as speed, sudden braking, road surface temperature, or airbag deployment. Using smart cars, the car-to-car systems managed to make each car able to transmit and receive data over a range of 500 metres. Another research group in ULM University in Germany has also developed an application called the Emergency Vehicle Warning System that uses radio communication to warn other vehicles or to pre-empt traffic lights with the aim being to reduce accident risks during emergency response trips and also help save valuable time [2.20].

Examples of other VANETs related applications include: Management systems for Incidents, Emergency services contact, Electronic Payment, Roadway Operations and Maintenance, Crash Prevention, and information on Safety, Highway, Weather, and Traveller.
2.5 Transport Layer Protocols

The transport layer provides the mechanism for exchange of data between two end systems. In TCP/IP there are two common protocols which are the User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). Each protocol has its own advantages and disadvantages depending on the applications. The protocols are also sometimes used to measure performance of a network.

2.5.1 UDP Fundamentals

This protocol provides a connectionless service which means it does not do any handshaking between the communicating nodes and as a consequence it becomes unreliable since it cannot guarantee delivery and duplicate protection. It is a simple protocol where the datagram is small with limited functions. Nevertheless, it is more appropriate in certain contexts due to its simplicity for example at the lower layers where UDP is more robust, easily adapting higher layer services, and producing a smaller amount of overheads due to its small packet header size and no handshaking. UDP works better in a real-time applications which can tolerate a small amount of packet loss and does not react well with controls such as congestion, for example Internet phone and video conferencing [2.21].

2.5.2 TCP Operation

TCP is well known for its connection-oriented protocol which ensures a reliable connectivity. It is used along with Internet Protocol (IP) which handles the actual delivery of data to the destination while TCP is responsible for ensuring that a message is divided into the packets that IP manages and for reassembling the packets back into the complete message at the other end. In order to establish the required reliable connection, a TCP connection goes into a 3-way handshake involving the connection establishment, data transfer, and connection termination phases.
In a wired network, packet losses are assumed to indicate network congestion but this is not the case in wireless networks as factors such as high Bit Error Rate (BER), unstable channel characteristics, and user mobility may all contribute to packet losses. Many studies [2.22] have shown that standard TCP performs poorly in a wireless environment due to its inability to distinguish packet losses caused by network congestion from those attributed to transmission errors.

The main characteristic of TCP is that it provides reliable connections between nodes. [2.23]. There are 3 elements of TCP which exist to provide this feature:

i. **Reliability**

There is always a high tendency for packets to be damaged, lost, duplicated, or delivered out of order during transmission especially in a wireless environment. In order to recover from these problems, TCP assigns a sequence number to each transmitted packet and requires a positive acknowledgment through an ACK packet from the receiving node. The data is retransmitted if ACK is not received within a timeout interval. At the receiver, the sequence numbers are used to correctly order segments that may be received out of order and to eliminate duplicates. Damage control is performed by adding a checksum to each segment transmitted, checking it at the receiver, and discarding damaged segments.

ii. **Flow Control**

TCP provides a mechanism to control the amount of data to be sent in order to provide a seamless transmission and to avoid traffic congestion. A concept called Sliding Window is applied where buffers are used at each end of the TCP connection to speed up data flow when the network is busy. It is used to keep a record of the packet sequences sent and their respective acknowledgements received by both sender and receiver, and allows a sender to transmit a specified number of data units before an acknowledgment is received or before a specified event occurs. There is also a Congestion Window (CWND) which determines the number of bytes that can be prepared to be accepted in its buffer at any time. This is a means of stopping the link between two places from getting overloaded with too much traffic. The size of this window is calculated by estimating how much congestion there is between the two places. The calculated
size is the maximum number of bytes that can be transmitted without acknowledgment that they have been received. Basically the size of the window usually controls the speed of transmission as transmission pauses until there is acknowledgment. In an ad hoc network, since routes may change during the lifetime of a connection, the relationship is lost between the CWND size and the tolerable data rate for the route. In other words, the CWND as computed for one route may be too large for a newer route, resulting in network congestion when the sender transmits at the full rate allowed by the old CWND.

iii. Connection

The reliability and flow control mechanisms require TCP to initialize and maintain certain status information for each data stream. The combination of this information including sockets, sequence numbers, and window sizes is called a connection. Each connection is uniquely specified by a pair of sockets identifying its two sides. When two nodes wish to communicate, the first thing to do is to establish a connection that is initializing the status information on each side. When the communication is complete, the connection is terminated or closed to free the resources for other uses. Since connections must be established between unreliable hosts and over the unreliable communication system, a 3-way handshake mechanism with clock-based sequence numbers is used to avoid error in the initialization of connections.

The characteristics which support the connection reliability of TCP provide the opportunities for researchers to manipulate TCP for further enhancing the network performance. This is one of the reasons why it is selected in this thesis.

A. TCP Variants

The evolution of network applications, especially involving wireless, requires the evolution of TCP as well. Many variants of TCP have been produced to reduce specific weakness by complementing the existing TCP. Features of some of the well known variants are listed and discussed below:
i. Vegas

Vegas was introduced by Brakmo et al. [2.24] and includes a modified retransmission strategy based on measurements of the round-trip time (RTT) with also new mechanisms for congestion detection during slow-start and congestion avoidance. The Vegas congestion detection mechanism is proactive because it tries to sense early signs of congestion by observing changes in the throughput rate. By referring to the congestion window adjustment policy from such throughput measurements, TCP Vegas may be able to reduce the sending rate before the connection experiences losses. Further details of the method can be found in [2.25].

ii. Tahoe

Tahoe refers to the TCP congestion control algorithm proposed by Van Jacobson [2.26] in order to improve TCP congestion control in the Internet environment. Two main approaches of this algorithm are to avoid congestion collapse and congestion avoidance. Congestion collapse is a condition when more packets were sent than could be handled by intermediate routers, the intermediate routers discarded many packets, expecting the end points of the network to retransmit the information. Instead of a successful re-transmission, more packets were being sent, causing more congestion and subsequently collapsing the whole route. In avoiding a collapse, when the connection is initialized and after a timeout, TCP initiates a Slow Start where it initiates a window with a low rate but increases this rapidly. When the congestion window exceeds a threshold, or a packet is lost, the algorithm enters a state of congestion avoidance As long as non-duplicate ACKs are received the congestion window is additively increased every round trip time. When a packet is lost, duplicate ACKs will be received. Tahoe’s solution is to reduce the congestion window to 1 MSS, and reset to the slow-start state

iii. Reno

This applies the same basic principles of Tahoe such as slow starts and the use of a re-transmit timer. However, it has a mechanism to enable the lost packets to be detected earlier and the pipeline is not emptied every time a packet is lost. Reno requires that the acknowledgement is received immediately whenever a segment is received, whereas in Tahoe loss is detected when a
timeout expires before an ACK is received. Reno also performs a fast retransmit when three
duplicate ACKs are received, and then goes into the Fast recovery phase. In this phase, TCP
retransmits the missing packet that was signaled by 3 duplicate ACKs, and waits for an
acknowledgment of the entire transmit window before returning to congestion avoidance. If there
is no acknowledgment, TCP Reno experiences a timeout and enters the slow-start state.

iv. New Reno

New Reno is a slight modification over Reno where it is able to detect multiple packet losses. It
also enters into a fast-retransmit mode when it receives multiple duplicate packets. The
difference is that it does not exit Fast recovery until all the data which was outstanding at the
time it entered Fast recovery are acknowledged.

A more detailed discussion on the performance of TCP variants in MANETs environment can be
found in [2.27].

B. Protocol Performance Issues

Generally, the main problems experienced by ad hoc networks particularly VANETs are caused
by its highly mobile nature which leads to the following issues:

i. High Bit Error Rate: packets get corrupted resulting in lost data segments or
acknowledgements. When the sender does not receive the acknowledgement within the
retransmit timeout (RTO), it retransmits the segment and exponentially backs off the
retransmit timer for the next retransmission thus reducing the congestion control window
threshold. Repeated errors will ensure that the congestion window at the sender remains
small resulting in low throughput. Error correction may be used but it will only waste
valuable wireless bandwidth when the correction is not necessary.

ii. Route Recalculation: The network layer at the sender starts to find a new route to the
destination when an old route is no longer available. In proactive protocols, table
exchanges are triggered that eventually result in a new route being found, while in
reactive protocols, this is done via route discovery messages. The time it takes to discover a new route could possibly be longer than the RTO at the sender causing the sender to timeout, then retransmit a packet and invoke congestion control. Thus, when a new route is discovered, this becomes less efficient because the throughput will continue to be small for some time as the TCP at the sender increases its congestion window using the slow start and congestion avoidance algorithm. In high node mobility situation as found in MANETs the route computations are done frequently causing the TCP connection to have difficulty in transmitting at the maximum negotiated rate. This is because the congestion window will always be significantly smaller than the advertised window size from the receiver.

iii. Network Partitions: The mobile nature of MANETs is likely to cause the network to be periodically partitioned for a few seconds at a time. If the sender and the receiver of a TCP connection are located in different partitions, all of the sender’s packets will get dropped by the network resulting in the sender invoking congestion control. If the partition lasts for a significant amount of time longer than the RTO, the situation gets even worse because multiple consecutive retransmissions of the same segment are transmitted to the receiver while it is disconnected from the sender causing all these retransmissions to be lost. Since the retransmission timer at the sender is doubled with each unsuccessful retransmission attempt, several consecutive failures can lead to inactivity lasting possibly up to one or two minutes even when the sender and receiver get reconnected.

iv. Multipath Routing: Some routing protocols maintain multiple routes between destination pairs in order to minimize the frequency of re-computing the route. Unfortunately, this sometimes results in a significant number of out-of-sequence packets arriving at the receiver. Consequently, the receiver generates duplicate acknowledgements (ACKs) which cause the sender on receiving the three duplicate ACKs to invoke congestion control.
MAC Failure Detection Issues: the time taken to detect link failure at this stage takes a longer time as it has to go through multiple retransmissions before concluding link failure. This increase of time contributes to the increase of load in the network. The link failure is sent to the resource packet and if another node is using the same path link the upstream node will have to wait longer before receiving the information about link failure.

These issues are very important and VANETs are usually affected by them. Studies on enhancing the performance need to have these considered carefully. What now follows is a review of some of the relevant work on the improvements of TCP.

C. Protocols Improvement Reviews

TCP performance improvements are usually focused on the flow and congestion control which have produced the many established variants mentioned earlier. Even these variants are continuously improved sometimes with new and refined versions to accommodate wireless-based applications. For example, in [2.28] additions were made to Vegas so that the sending node knows how to distinguish between congestion and transmission errors by efficiently adjusting the TCP CWND and transmission rate while in [2.29] Vegas_M was introduced where fuzzy logic theory was used to distinguish network congestion states and wireless channel states in order to optimize ad hoc network throughput and achieve better performance.

Other examples on the flow and congestion control are mentioned in [2.30]. Flow control mechanisms were proposed which provide link failure feedback in order to avoid the misinterpretation of route failures as congestion. Early Packet Loss Notification (EPLN) seeks to inform TCP senders about lost packet, and Best Effort ACK Delivery (BEAD) attempts to retransmit lost ACKs at either intermediate nodes or receivers. These methods extensively use cache routes which make it suitable for the DSR protocol. Congestion control was proposed where congestion can be predicted and avoided by using active packets to look around the network and collect state information about source routes and queue length. The former provides better updates on routes hence reduce unnecessary retransmission while the latter identifies the potential congestion nodes and computes alternative routes.
Another approach utilizes network layer feedback from intermediate hops to put the TCP sender into either a ‘persist’, ‘congestion control’, or ‘retransmit’ state. Thus, when the network is partitioned, the TCP sender is put into persist mode so that it does not needlessly transmit and retransmit packets. On the other hand, when packets are lost due to errors as opposed to congestion, the TCP sender retransmits packets without invoking congestion control. When the network is truly congested, the TCP sender invokes congestion control normally. This is done by implementing a thin layer called Ad hoc TCP (ATCP) [2.31] inserted between Internet Protocol and the standard TCP and listening to the network state information provided by ECN (explicit congestion notification) messages, and by ICMP “Destination Unreachable” messages. This then puts TCP at the sender into an appropriate state. Upon receiving the ICMP, the sender enters the persistent or frozen state until a new route is found ensuring that the sender does not invoke congestion control. Another technique which works by supplementing TCP is TCP Smart-Framing [2.32] which uses a segmentation algorithm and can be implemented on top of the TCP variants.

One important approach is the Cross-layer design philosophy. In this case the transport and link layer are working together in order to achieve an optimal solution. The introduction of an ARQ Snoop agent within the protocol stack of a wireless host called Link Layer ARQ Exploitation TCP (LLE-TCP) creates the possibility for the exploitation of link layer ARQ messages within TCP acknowledgement scheme [2.33]. Another promising approach is the Interlayer Collaboration Protocol (ILC-TCP) [2.34] where the main concept is the introduction of a State manager in parallel to the protocol stack used to gather information from TCP, IP and the Link-Physical layers and provide the information to the other layers if necessary.

2.6 File Transfer Protocol (FTP) Application

File transfer is one of the most common actions taken in the network environment and FTP is the standard mechanism provided by TCP/IP in order to copy a file from one host to another. It is a client/server application where it establishes one connection for data transfer and another separate connection for control information which makes it efficient. The latter uses a simple
rule that is it allows transfer only a line of command or a line of response at a time. Data transfer or connection requires more complex rules due to the variety of data types.

FTP works both on UDP and TCP where it works well with elastic traffic due to its ability to adjust to changes. During a file transfer, FTP supports the expectation of users who would want the delay to be proportional to the file size that eventually affect the changes in throughput. The main work in this thesis focuses on the Network layer which is where the routing functions are handled. FTP is the tool at the application layer used to send and receive the data. Both UDP and TCP are initially used with the latter being the selected transport protocol for further works throughout the thesis.

2.7 Routing Protocols in MANETs

There are a variety of MANET protocols and ways to classify them. The most popular classifications are the proactive (or table-driven), and reactive (or on-demand) routing method which is becoming one of the most commonly used routing strategies [2.35]. The strategy of proactive routing protocols is for each node to maintain consistent and up-to-date routing information while the reactive routing protocols create routes only when required by the network traffic. In the following section a description of the three main routing protocols will be presented, namely DSDV, DSR and AODV.

2.7.1 DSDV

Destination-Sequenced Distance Vector (DSDV) is a proactive protocol [2.36] based on the Distributed Bellman Ford (DBF) [2.37] protocol and sequence number tagging. It uses the shortest path computation method like in the link state approach but with more efficiency by using a vector with which each node maintains the view of networks with each cost of each link. Estimation is also more accurate, easy to implement and also uses less storage space compared to the link state method. The routing loop is a potential problem when there is a long propagation delay because it could make the cost inaccurate. Coordinating connections between nodes is the mechanism to eliminate this but does not really work well in a rapidly changing topology.
Packets are transmitted between nodes by utilizing routing tables stored at each node. The routing table lists all the available destinations, and the number of hops to reach each one is also stored. Route table entries are tagged with a sequence number that originates from the destination node. The consistency of routing tables in a dynamically varying topology is maintained with each node periodically transmitting updates, doing so immediately when significant new information about the link becomes available. Routing information is advertised frequently to each of the node’s current neighbours through broadcasts, transmitted periodically and incrementally as topological changes are detected. Data are also maintained about the length of time between the arrival of the first and the arrival of the best route for each particular destination. Decisions based on this information may be made to prevent fluctuations in the route tables.

For each new route, the routing table for each data broadcast contains a new sequence number, a destination address, the total number of hops required to reach the destination, and the sequence number of information received regarding the destination as originally stamped by the destination. As for the packet header of a transmitting node, the transmitted route table also contains the hardware and network address, and the sequence number of the transmitter. Routes with more recent sequence numbers are always preferred (as the basis for forwarding decisions), but not necessarily advertised. For a path with the same sequence number, the one with the lowest number of hops will be used. As the routing tables are being propagated through the network the sequence number is sent to all nodes which may then decide to maintain a routing entry for the originating node. At the receiving node, routes received in broadcast are also advertised by the receiver after it adds an increment to the metric. No node may insert routing information received from a neighbour unless that neighbour shows that it can receive packets from the node. This is done in order to avoid a one-way link.

Topology changes and broken links are detected either at layer 2 (the MAC layer), or inferred if no broadcast is received for a period of time from a former neighbour. When a link to the next hop is broken, any route through that next hop is immediately assigned with a number that is greater than the allowed number of hops and an updated sequence number. It is a substantial
route change so the modified routes are immediately disclosed in a broadcast routing information packet. This is the only situation when sequence numbers are being generated by nodes other than the destination node. In a large population of nodes, the time between broadcasts of routing information packets needs to be adjusted to reduce the amount of information carried by the packets. Full dumps carry all available routing information while incremental dumps carry information concerning only changed information from the last full dump. The former can be transmitted infrequently when the topology is stable while the latter is used more when there are active topology changes. In selecting a route, a new packet is compared with the one available previously. A route with a more recent sequence number is always preferred while older ones are discarded.

The drawback of this protocol is the update procedure as it increases the volume of control traffic and adversely affects the network. It becomes difficult to maintain the routing table properly when the number of nodes in a network gets larger and when the mobile nodes move around quickly. The consequences of fast topological changes are typically broken links and heavy usage of the network bandwidth. Nevertheless, DSDV has uses in ad hoc networks with static nodes or where the nodes do not move excessively such as in indoor or office situations.

2.7.2 DSR

Dynamic Source Routing (DSR) is a reactive routing protocol [2.38] where each node sends packets to a destination according to the routing information contained in its route cache. It uses the route discovery and route maintenance mechanisms in order to allow nodes to discover and maintain source route to arbitrary destinations within the network. These mechanisms work entirely on a demand basis. Route discovery takes place when a source node wishes to send a packet to a destination node and does not already know a route to it. Route maintenance is used when a source node detects changes to the network topology such that its existing route information can no longer be used. When a source route is known to be broken, the source node can then attempt to use any other route to the destination route it happens to know, or it can initiate a Route discovery once again to find a new route. Route maintenance is used only when the source node is actually sending packets to the destination node.
In DSR each data packet sent carries in its header the complete, ordered list of nodes through which the packet must pass, loop free avoiding the need for up-to-date routing information in the intermediate nodes through which the packet is forwarded. Other nodes forwarding or overhearing the packets may also cache the routing information for future use. This routing protocol is designed to have a low overhead and will be able to react quickly to changes in the network, providing a highly reactive service to help ensure successful delivery of data packets despite the movement of nodes and other changes in the network.

Unlike DSDV, it does not require periodic broadcast packets to be transmitted within the network. It does not use any periodic routing advertisement, link status sensing, or neighbour detection packets. It only broadcasts, as it initiates Route Discovery, a special data packet known as the Route Request Packet (RRP) which contains the address of the destination along with the source node address and a unique identification number. If a node that does not know the route to the destination receives the RRP, it adds its own address to the route record of the packet and then forwards the packet to the next node. Eventually, providing the network topology allows, the RRP reaches a node that does not know the route to the destination and is able to forward the complete route back to the source node. There is no overhead packet when nodes are stationary, and as nodes begin to move the routing protocol overhead automatically scales to only that needed to track the routes currently in use.

With DSR, a node may learn and cache multiple routes to any destination via Route Discovery. It allows the reaction to routing changes to be more rapid because a node with multiple routes to a destination can try another cached route if the one it has been using fails. The caching of multiple routes also avoids the overhead incurred by performing a new Route Discovery each time a route in use breaks.

2.7.3 AODV

Ad Hoc On-Demand Distance-Vector Routing (AODV) [2.39] is a source-initiated on-demand-driven protocol which provides communication between mobile nodes with minimal control overhead and minimal route acquisition latency. It does not maintain routes from every node to every other node in the network. Routes are discovered on an ad hoc basis and maintained only
as long as they are necessary. It is loop free because it increases the sequence number each time it learns of any change in the topology of its neighbours. The sequence number ensures that the most recent route is selected for a route discovery.

AODV has the benefit of having the ability to provide unicast, multicast, and broadcast communication. The route information for both unicast and multicast can be shared and in a mobile environment any reduction in control overhead is a significant benefit. Route tables are used to store relevant routing information. AODV utilizes both tables each for unicast and multicast routes respectively. It is used to store the destination and next-hop address as well as the destination sequence number. For each destination the node maintains a list of so called precursor nodes which contains the route needed in order to reach the destination. The list is modified for the purpose of route maintenance if a link breaks. Each route table entry incorporates the concept of route ‘life-time’ which is updated whenever a route is used. A route is made to be invalid when it is not being used within its life time, and is not maintained when the nodes along the route are likely to have moved.

This protocol maintains both unicast and multicast routes even for nodes in constant movement. It also quickly deletes invalid routes through the use of the Route Error (RERR) message and responds to topological changes that affect active routes in a timely manner. When a source wants to send a message to some destination node and does not already have a valid route to that destination, it initiates a path-discovery process or route discovery in order to locate the other node which follows a request/route reply discovery cycle. Requests are sent using a route request (RREQ) packet by broadcasting to its neighbours. The nodes that receive the RREQ packet then forward the request to their neighbours, and this process repeats until the RREQ packet reaches either the destination or an intermediate node that knows the route to the destination. Information enabling the creation of a route is sent back in a Route Reply (RREP) packet. Sequence numbers are used to eliminate the stale routes (inactive routes after a predefined time) and routes with old sequence numbers are discarded from the system.

Routes in AODV are built with only small amount of overhead from routing control messages and no additional network overhead. The storage requirement at each node can be minimal as
each node is required to maintain only next-hop routing information. AODV also does not place any additional overhead on data packets because it does not utilize source routing.

Unlike DSDV, AODV minimizes the number of required broadcast by creating routes on an on-demand basis and not maintaining a complete list of routes. It will not broadcast if a particular link status does not affect the ongoing communication. Only nodes in the broken route will have to be informed of the link’s changes status. Compared with DSR, routes not in use are expired and consequently discarded which reduces the effect of stale routes as well as the need for route maintenance for unused routes.

2.7.3.1 AODV Control Packets and the Routing Table

Packets are used in the AODV protocol in order to establish links, provide data transmission, and also maintenance. The descriptions of the packets are as below:

i. Route Request (RREQ)

This packet is used to initiate the route discovery process. It has the control flags to indicate the specific status of the packet like repair for multicast, set as a gratuitous RREP which facilitates the routing process by allowing intermediate nodes to respond to RREQ of the messages if the route to the requires destination is known. The hop count is also recorded to indicate the number of hops from the SN up to the current node which is the element of a distant vector. A SN may potentially issue more than one RREQ packets during the route discovery process, and to uniquely identify each packet the RREQ ID is used which is a unique sequence number. The IP addresses and sequence numbers for both SN and DN are also part of the packet.
ii. Route Reply (RREP)

RREP is used to finalize a route. As in other packets it also has control flags, DN sequence number, hop count, SN and DN IP addresses. In addition, it also has the prefix size field for subnet routing, and lifetime to consider if the route is still valid.

iii. Route Error (RERR)

This packet is sent when there is a link break that could cause the destination of any node in the active route to become unreachable. It has the flag control elements and one or more IP address and sequence number of the unreachable destination. Also included is the number of unreachable destinations.

iv. Hello

Hello packets are used to detect if a link is still alive. They are broadcast periodically and can only be used if the node is part of an active route. A failure in getting a reply leads to the assumption that the link is broken. The hello message is a RREP with the Time To Live (TTL) set to 1 with the following fields are set accordingly:

i. Destination IP address
ii. Destination sequence number
iii. Hop count
iv. lifetime

Each node will have its own routing table entries. Whenever a node receives control packets it will check the routing table and respond accordingly, while if there is no entry is available then a new entry is created.
The fields of a routing table are as follows:

i. Destination IP address  
ii. Destination sequence number  
iii. Valid destination sequence number flag  
iv. Other states and routing flags  
v. Network interface  
vi. Hop count  
vii. Next hop  
viii. List of precursors  
ix. Lifetime

A simple simulation shown in Figure 2.3 was run in order to demonstrate the process where there are three nodes in a network with node 0 is the SN, 1 is intermediate, and 2 is the DN. The extraction of routing tables from the NS2 trace file for each node consisting of destination, next hops, number of hops, flag, and lifetime fields are shown in Table 2.1.

![Simple simulation with 3 nodes](image)

Figure 2.3 Simple simulation with 3 nodes
### Table 2.1  An extract of Routing Tables

<table>
<thead>
<tr>
<th>Destination</th>
<th>Next Hop</th>
<th>No. of hops</th>
<th>Flag</th>
<th>Lifetime</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>30.968288</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Destination</th>
<th>Next Hop</th>
<th>No. of hops</th>
<th>Flag</th>
<th>Lifetime</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>29.988465</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Destination</th>
<th>Next Hop</th>
<th>No. of hops</th>
<th>Flag</th>
<th>Lifetime</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>29.96838</td>
</tr>
</tbody>
</table>

Each node updates its own table and as in the case of node 1 shows a multiple entries consist of the forwarding route and also the reverse route entries.

### 2.7.4 Other Routing Protocols

Temporally-Ordered Routing Algorithm (TORA) [2.40] is a reactive routing protocol originally developed for possible use in military networks which is basically large-scale, dynamic networks with a mixture of physical layer and multiple access technologies. It is designed to discover routes on demand, provide multiple routes to a destination, establish routes quickly, and minimize communication overhead by localizing algorithmic reaction to topological changes when possible [2.41]. TORA only maintains information about adjacent node and maintains state on per-destination basis.

The dynamics of route changes require the updating of routing information in the remaining nodes so that consistency can be maintained. However, this process involves broadcasting over the wireless medium which reduces the available bandwidth for data and increases the overall network control traffic. This explains the need for new or enhanced routing protocols. For example, the Geocasting [2.42] and Location-aided Routing (LAR) [2.43] are protocols which utilize physical locations to improve performance. Meanwhile, the works of [2.44] and [2.45] to produce Ad hoc Positioning Systems (APS) and Adapting to Route-demand and Mobility (ARM)
respectively show the enhancement of an existing routing protocol which is DSDV, where the former adds the GPS positioning information, and the latter uses a control mechanism to dynamically adapt to changes in node mobility and route demands.

2.7.5 Routing Protocol Summary

A comparison of the different routing protocols considered is shown in Table 2.2:

<table>
<thead>
<tr>
<th></th>
<th>DSDV</th>
<th>DSR</th>
<th>AODV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>Proactive</td>
<td>Reactive/On-demand</td>
<td>Reactive/On-demand</td>
</tr>
<tr>
<td><strong>Establish route and update</strong></td>
<td>Each node periodically broadcasts for each change in the network.</td>
<td>Route Discovery and Maintenance applied on the affected route or on demand only.</td>
<td>Route Discovery and Maintenance applied on the affected route. Sequence numbers are used to discard stale routes.</td>
</tr>
<tr>
<td><strong>Route table contents</strong></td>
<td>List of all available destinations.</td>
<td>Complete, ordered list of nodes through which the packet must pass (source route)</td>
<td>Stores destination and next-hop address, sequence number, life time</td>
</tr>
<tr>
<td><strong>Overhead control</strong></td>
<td>Uses full dump when stable, partial dump when link active.</td>
<td>Allowed to use cache containing information on the previous route.</td>
<td>Does not add any overheads when link is functional. Maintains only next-hop routing information.</td>
</tr>
</tbody>
</table>
2.8 Transmission Range and VANETs

In VANETs the transmission power is an important issue [2.46, 2.47]. There are two contradictory requirements which must be balanced. Firstly the requirement that the transmission range is sufficient to allow communication with adjacent nodes. In areas of low node density this dictates a large range. One characteristic of transmission range is that longer range does not necessarily equate to better network performance [2.48]. Even though a longer transmission range reduces the number of hops in a route, it also increases the number of nodes that compete on the shared channel. Secondly the requirement is that the communications band does not get congested by network traffic. This dictates a short range so that in areas of large node density it is still possible to get sufficient throughput. A shorter transmission range allows better reuse of bandwidth and minimal use of power and also throughput can be improved because more simultaneous transmission are allowed at the same time in different areas of the network. The disadvantage of having shorter range is that it creates network partitioning which could prevent some nodes from communicating with their destinations. This should also be avoided, thus a balanced range is the choice according to the required situation.

Another pertinent issue for transmission power is that MANET deployments are usually powered by batteries with finite capacity. Studies such as [2.49] have proposed a transmission strategy that can determine the optimum transmission range that gives maximum efficiency of energy consumption. In VANETs, however, the nodes can usually be assumed to have a stable and ample power source that comes from the vehicle. Therefore, in this thesis the issue of power supply is assumed to be non critical and it is assumed the nodes can provide a sufficiently high coverage area when needed.

In the practical wireless environment, the optimal range cannot be represented by a fixed number as the value changes as it responds to interference and constraints from the environment. In addition, the coverage may not be truly isotropic although that is a necessary assumption in most network simulations. In the simulations in this thesis, the assumption is that any node can be successfully heard by any other node providing a distance between them is to be less than certain
Euclidean distance in flat two-dimensional space without any real propagation obstacles that could potentially block or reflect signals.

2.9 Performance Metrics

There are many aspects or metrics that can be looked into or used when it comes to measuring the performance of a network and data can be obtained through simulations or actual measurements. In this thesis, the focus is on the measurements of performance that would reflect the effects of mobility, density and transmission range. In this case it involves the transmissions of TCP data packets via FTP using specific mobility models. The evaluations in the thesis were made using the following metrics:

i. **Packet delivered, received and the ratio**: is the number of packets successfully sent and received. The ratio is of the data packets sent by the source node to those actually being received by the destination node. Delivery ratio shows the fraction of successful data transfer which is an indicator of the overall performance. This is determined by counting the number of sent and received packets at the routing agent.

ii. **Overhead packet**: is the number of routing packets transmitted reaching the router and the MAC layer. This is determined by counting the packets that reached the router and the MAC layer of the receiving nodes. The number of overhead packets shows the efficient use of control in the routing functions, with lower number shows minimal and efficient usage as they would not contribute to the congestion of the traffic channel.

iii. **Dropped packet**: is the number of dropped packets in the duration of the simulation. This is to be used as one of the indicators for studying the link effect on the performance where efforts to reduce the dropped packets are to be looked into. It indicates data transfer failure, and this is used to identify broken links which is the focus of the study in this thesis.
These parameters are widely used in measurements of network performances and can be found in many experiments involving wireless networks especially ad hoc networks with examples from [2.50-2.53]. Measurements can also be made in 2 modes namely live or field measurements, and through simulations which are discussed in the following subsections.

2.9.1 Field Measurement

This technique measures the actual scenario and hence produces the exact result from a particular scenario. Ideally it is the best way to measure since it is made in real world setting hence does not require any inaccurate assumptions especially on the external influences. The disadvantages of this method are the lack of repeatability and scalability which can lead to high cost [2.54]. Example of field measurement on MANET routing protocol can be found in [2.55].

Another method which is similar to field measurement is the network test bed. It provides a platform for experimentation and unlike field measurement it allows rigorous, transparent, and replicable testing. The results are then compared with the one obtained from simulation, and test bed are known to be commonly used in many research findings such as by [2.56-2.58] for both MANETs and VANETs environment.

2.9.2 Simulation

An alternative to field measurement is simulation. It is an imitation of the actual thing to measure and it also allows the simulation to be made on a particular abstract model of a particular system. Through simulations researchers are able to do measurements which are sometimes very difficult and costly to do in the real environment especially in a vehicular environment as there are a large number of variables and traffic situations which are difficult to measure. The downside is that usually the setting of parameters are often managed in a simplistic way thus producing results that are not as much realistic as they could be. For instance, heavy impacts on bit-error-rate and packet-error-rate may come from physical layer phenomena which are either neglected to make the overall simulation process faster, or have to be managed by computationally heavy
simulators [2.59]. The challenge therefore is to ensure that the parameters and defined as much as possible. There are many known simulation works on MANET and its routing protocol with examples given in [2.60, 2.61].

2.10 Discussions and Summary

Simulations produce useful data of VANETs which are comparable with experimental results. However, they are very dependent on the accuracy in defining the parameters such as of the physical layer parameters and propagation model [2.62]. The technical challenges for MANETs have encouraged various efforts to resolve them through research. Examples are issues on flow and error control, maintaining accurate network topology information, handling node as router mobility, shared channel access, and power requirements. As the work is being focused further from MANETs to VANETs, more specific challenges can be seen for example such as the effects of variable speed of static, low and high speed vehicles, traffic density and flow, antenna direction, electromagnetic effects, and heterogeneous devices. Some of these factors will be investigated in the following chapters.

This thesis considers a solution through enhancement of TCP via link management in order to improve the performance of routing protocol in VANETs.
References


CHAPTER 3:
SIMULATION TECHNIQUES AND MEASUREMENTS FOR VANETs

Simulation is the approach used in this research to study the network performance. This chapter covers the relevant simulation topics, such as the setup and parameters, the wireless propagation models used in the thesis. It also discusses and introduces the mobility models and the proposed model using Cellular Automata.

3.1 Network Simulation Software

The increase in research work covered with VANETs has seen the existence of several communication network simulation tools made available to provide a platform to test and evaluate the network protocols such as NS-2, OPNET, GloMoSim, etc [3.1, 3.2,]. These tools are primarily designed to support generic simulations and they are not specifically designed for complex dynamic scenarios such as the VANET environment. Even though many simulation tools are available for the purpose of study on transportation sciences, they do not integrate network communication features and techniques for the use of network research work. This subsequently leaves researchers with the only option of extending the available network simulation tools to incorporate the environment of MANETs and VANETs as in the case of this research study.

For the purpose of comparison, the OPNET and NS2 simulators are introduced and discussed in the following subsections.
3.1.1 OPNET

This is a discrete event simulation software package used to provide performance analysis in a wide range of computer network analyses including ad hoc networks, which is part of the OPNET Modeler Wireless Suite [3.3]. It supports routing protocols which includes AODV, DSR and TORA and also integrates 802.11 devices [3.4] and propagation models for the outdoor environment. OPNET is a hierarchical model consisting of the top level which defines the nodes and topology, then constructed by processes represented by block-diagrams in the software component layer; eventually each of the processes can be constructed using finite state machine (FSM) models. Finally, all these will be executed by the discrete event simulator of OPNET.

OPNET also provides schemes called the IT Guru Academic Edition GURU [3.5] where various levels of software are provided to educational institutions with minimal fee or made available free for a limited version of the software.

3.1.2 NS2

NS2 is an open-source software which runs on various platforms including UNIX, LINUX, Windows, and Mac systems. It provides users with an executable command, ns, which takes as its input the file name of a short program written in the high-level computer language Tcl. The input file contains various directives for the simulation and configures the output. In most cases, the simulation outputs a trace file which is used to plot various graph and/or to create animations using the NS-2 animation tool NAM. The NS-2 source code consists of two key languages: C++ and Object-oriented Tool Command Language (OTcl). While the C++ defines the internal mechanism of the simulation objects, the OTcl sets up simulation by assembling and configuring the objects as well as scheduling discrete events. The C++ code and the OTcl are linked together using Tcl with classes (TclCL) [3.6].

Being open source, the kernel engine is available to users for modifications and development. It has built-in propagation models which allow a wide range of network configurations supporting the networking stacks from the physical layer all the way up to the transport layer.
For VANET simulations, NS-2 does not work well for very large topologies which are better served by microscopic models. It simulates networks at the packet level only, meaning that large models with many nodes eventually become intractable. In NS-2 graphical representation is limited to its NAM animator which is very basic, although users can develop their own animation. In terms of accuracy, [3.7] found that performance, such as the speed assigned to a node, is dependent on the model source code, and it would therefore be unfair to compare two models that are not coded identically. Some of the functions in NS-2 are also sometimes performed on assumptions which may not precisely represent the actual scenario such as in a case of allocating channels. Nevertheless, coding is accessible and users may have the ability to modify the software according to their needs thus producing more accurate results. Therefore it may not be right to claim that two simulations will definitely produce the same results as each of them can be modified to suit its own need.

3.1.3 Comparison

The issue of simulation acceptability was considered in depth in [3.8]. A total of 151 papers from the ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc) conferences from the years 2000–2005 were considered, which showed that 75.5% of the papers had used simulations in their research. From this, 95.6% of the simulations conducted MANET protocol studies, with 43.8% of the MANET studies used the NS-2 simulator. Further breakdown showed that 38.5% of the NS-2 simulation work was on mobility, with 10.5% of the mobility studies focused on road mobility model.

A brief comparison between OPNET Modeller and NS-2 is listed in Table 3.1. Even though OPNET may have certain advantages over NS-2, the software was not available for this research at the point of time when the simulation work was done.
Table 3.1 Comparison between OPNET Modeller and NS-2

<table>
<thead>
<tr>
<th>OPNET Modeller</th>
<th>Network Simulator 2 (NS-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part of OPNET Technologies suite</td>
<td>Developed by Virtual InterNetwork Testbed (VINT) project</td>
</tr>
<tr>
<td>Finite state machine plus analytical (mathematical) model: high processing speed and accuracy</td>
<td>Object oriented with C++ and OTcl: performance and user friendly</td>
</tr>
<tr>
<td>Ability to model protocols, devices and behaviours with many built-in modelling functions</td>
<td>Numerous models of common Internet protocols</td>
</tr>
<tr>
<td>Good GUI for all tools</td>
<td>Network animator NAM for visualization of simulation</td>
</tr>
<tr>
<td>Good documentation and study cases</td>
<td>Limited documentation</td>
</tr>
<tr>
<td>Proprietary software: not open source</td>
<td>Open-source</td>
</tr>
<tr>
<td>Extensive component available: less development effort, time and errors</td>
<td>High learning curve</td>
</tr>
<tr>
<td>Good memory manipulation and highly scalable</td>
<td>Lack in scalability due to constraint in large memory usage</td>
</tr>
<tr>
<td>License fee</td>
<td>Free</td>
</tr>
</tbody>
</table>

Therefore, NS-2 has been chosen as the simulation software in this thesis because it is a fully free open source software and has been shown to produce reliable and comparable results to OPNET. This is because since the behaviour of both simulators is quite similar in general [3.9].

3.2 Simulation of Wireless Channels: Fading and Ricean Model

Multipath propagation can cause fast fading to occur when a transmitter and receiver are surrounded by objects which reflect and scatter the transmitted energy causing several waves to arrive at the receiver via different routes. Both Rayleigh and Ricean distributions are the statistical models which provide a good approximation to the effect of a propagation environment for mobile fading channel for the No Line of Sight (NLOS) and Line of Sight (LOS) situations respectively [3.10].
According to Bai et al. [3.11] an acceptable simulation model for a dynamic scenario such as a highway is the Rayleigh fading propagation model. In the simulations presented in this section the Rayleigh and Ricean fading extension module [3.12] is used as the propagation model. The formula for Rayleigh distribution is very much similar to the Ricean and if the Rice factor, $k$, is set to zero the two distributions are identical. This module uses the Ricean distribution by considering Rayleigh fading as a case where the magnitude component is zero. The Rice factor $k$ is defined as:

$$ k = \frac{\text{Power in constant part}}{\text{Power in random part}} $$

(3.1)

This modelling uses a pre-computed dataset containing the components of a time-sequenced fading envelope. It is used as a lookup table during the simulation run to model a wide range of parameters. Adjusted parameters are the time-average power, the maximum Doppler frequency, and the Ricean $k$ factor. It is also assumed that the small scale fading envelope is used to modulate the calculations of a large scale propagation model like two-ray ground or some other deterministic model.

### 3.3 NS2 Simulation Parameters

NS-2 is used in all the simulations contained in this thesis. The actual setups and parameters are defined in the respective chapters. In this section, the description will be on the common details of the parameters which are listed below in Figure 3.1. This figure is part of the NS-2 codes used to define the parameters.

The first parameter channel simulates the actual transmission of the packet at the physical layer. The basic channel model in the simulation implements a shared medium with support for contention mechanisms. It allows the MAC to carry out carrier sense, contention, and collision detection. If more than one transmission overlaps in time, a channel raises the collision flag. By checking this flag, the MAC object can implement collision detection and handling.
The second parameter defines the propagation models used to predict the received signal power of each packet. At the physical layer of each wireless node, there is a receiving threshold. When a packet is received, if its signal power is below the receiving threshold, it is dropped by the MAC layer. There are three propagation models in NS-2, which are the free space model, two-ray ground reflection model and the shadowing model.

The following parameter is the antenna which defines the type of antenna attached to the network interface card. In most cases, an omni-directional antenna is used in the simulations. This is a reasonable assumption because although real antennas exhibit lobes and other non ideal behaviours, the design objective in automotive is typically to achieve all-around coverage.

The link layer parameter is responsible for simulating the data link protocols. Many protocols can be implemented within this layer such as packet fragmentation and reassembly, and establishing reliable link.

The Interface queue type parameter specifies the queue type to be used such as DropTail or PriQueue where the queue size can also be specified.

The next parameter is the Medium Access Controls which allows the choice of MAC to be used such as 802.11. For sending, the MAC follows a certain medium access protocol before
transmitting the packet on the channel and upon receiving, the MAC layer is responsible for delivering the packet to the link layer.

There are four different ad-hoc routing protocols currently implemented in NS-2 and they are DSDV, DSR, AODV and TORA. Of those the first 3 have been used in the thesis.

The list of parameters is not limited to this, for example, the topology size and the number of nodes can also be specified. For the analysis of simulation work, NS-2 provides the details of the results via trace file. It is in a form of text file which may capture the following information:

1. Action
2. Time
3. Node ID
4. X Coordinate (If Logging Position)
5. Y Coordinate (If Logging Position)
6. Trace Name
7. Reason
8. Event Identifier
9. Packet Type
10. Packet Size
11. Time To Send Data
12. Destination MAC Address
13. Source MAC Address
14. Type (ARP, IP)

Figure 3.2 The NS-2 Wireless Trace File

The value of action can be set to send, received, dropped, or forward. This is usually referred to when making the selection for the extraction of any particular metrics. The data is usually large and for the purpose of analysis software with good query abilities can be used such as Perl, Awk, and even Excel. The trace file can also interact with other software such as JAVA, which for example, can produce visualization of the simulations and graphs. NS-2 can generate trace files for visualization via its own NAM tool. Apart from the action, this thesis also look among others at the nodes involved, and where at that particular time the packet is when it happens such as the application layer, the router, MAC layer, etc.
3.4 Mobility Models

Mobility plays an important key role in the deployment of wireless networks especially in VANETs. When it comes to the development of protocols and their related applications, full system simulation is often required and within that simulation the mobility model becomes one of the key elements that determine its accuracy and success.

In a simulation, a mobility model is used to represent the physical movement of the nodes within the network. A typical mobility model first places the mobile nodes in a pseudo random initial location and defines the way the nodes move within the network. It tries to emulate as close as possible the real node movement by including the critical movement factors such as direction, speed and destination. Defining a realistic model is a challenge because no model to date has successfully represented a real large scale VANET or even MANET environment [3.13]. Researchers have used a wide variety of models ranging from simple to complex algorithms in implementing the generations of movement patterns [3.14]. Basically, a comprehensive model is developed to make it more realistic according to the specific scenario enabling more accurate simulation and evaluation of the required network parameters while a simple model is used for more general purposes like understanding an unpredictable movement pattern.

There are two generally accepted classes of mobility models; trace-driven and synthetic models. Trace-driven models are based on the histories of real nodes or users traces. They assume that by recording and using traces of individual user’s historical position data the overall mobility can be determined. An example of a trace-based method can be found in [3.15]. However, in a dynamic situation such as those found in VANETs, or if traces cannot be obtained, a synthetic model is often used. It simulates the movement behaviours of mobile users using particular constraints, assumptions, or mathematical equations. Of the two classes the latter is much more convenient to study and avoids the need to collect large (and in most cases, infeasible) quantities of trace data.
3.4.1 Reviews of Existing Mobility Models

When formulating a mobility model it is possible to view the vehicle traffic flow at two levels, namely macroscopic and microscopic. This is analogous to the flow of electrical current where a microscopic model would attempt to consider the movement of individual electrons (vehicles) and a macroscopic would just look at the bulk flow (traffic). Of the two, the former microscopic model typically has to be used since communication issues between specific individuals tend to be the primary concern. Macroscopic models also have their places in VANET modelling in the area of networking planning, however this is not the concern of the current thesis.

In a microscopic model of vehicle movement on the highway there are a number of quite complex issues that can affect it. For example, road topology, speed limits, the number of lanes and car movement constraints. There are also the differing effects such as road conditions, the performance of individual drivers and the weather to name but a few. A more detailed discussion of these issues is included in [3.16].

The performance of a variety of mobility models was evaluated in [3.17]. Simple models like Column, Pursue, and Reference Point Group Mobility, Random Way point were shown to produce different performance results.

Random Walk is the basic type of mobility model where a node chooses a random direction and a speed from a given speed distribution and moves in that direction for a specified time. When the time ends, the node repeats this motion.

The Random Way Point model is a widely used alternative model in which each node individually chooses a random destination within the simulated boundary and also chooses a random speed between a preset minimum and maximum speed limit with which the node moves towards the destination. Once the node reaches the destination, the node pauses at the destination for a pre-specified pause time period. After the pause time, the node repeats the process, choosing another random destination and random speed. These two models are frequently used for basic simulations mainly due to their simplicity of implementation and analysis, and it has
been noted that Random Way Point model has been able to give a good approximation for simulating the motion of vehicles on a road [3.18]. Limitations of this model include inability to represent geographic restrictions, temporal dependence of movement of a node over time, and spatial dependency of movement among node [3.19]. Random Way Point allows nodes to move randomly with each node moving independently of time and without considering collisions with other nodes whereas in a specific environment like VANET, the nodes are arranged and movements are made according to time, and sequence (through queuing) or location of the neighbouring nodes.

The development of comprehensive models has been the topic of much research with some studies focused on the VANET purposed environments. The Highway Mobility model and Manhattan Mobility model are two examples from the many available highway specific models [3.20]. In the Highway model, the node moves according to a certain path in a certain direction. Nodes are not allowed to change direction. The velocity in the case of the Highway mobility model is varies in three lanes which are classed as slow, medium and fast. The Manhattan model emulates the movement pattern of mobile nodes on streets defined by maps. It can be useful in modelling movement in an urban area where a pervasive computing service between portable services is provided. Other synthetic entity models for ad hoc networks include the Random Direction, Boundless Simulation Area, Gauss-Markov, City Section, Exponential Correlated Random, Column Mobility, Nomadic Community, Pursue and Reference Point Group which are discussed in detail in [3.21].

### 3.4.2 Traffic Flow with Cellular Automata (CA)

A very powerful concept in traffic modelling is the idea of Cellular Automata (CA). Cellular automata are collections of cells on a grid of specified shape that evolve through a number of discrete time steps according to a set of rules based on the states of neighbouring cells. The rules are then applied iteratively for as many time steps as desired. This creates models that are discrete in space, time and state variables [3.22]. The discrete nature of CAs makes them much easier and more efficient to implement in a computerized simulation environment. The basic one-dimensional Cellular Automata Model for highway traffic flow can be adapted from CA rule
184 [3.23]. It describes a one-lane traffic road with sequence of grid points with each being an identical square representing one vehicle. The model considers the effects of acceleration and delay of vehicles which and captures the realistic situations where the car accelerates and decelerates.

Example of other implementations of Cellular Automata in traffic flow models can also be seen in many publications. A simulator was developed to evaluate dynamic route selection methods using standard map where traffic congestions occurs frequently [3.24]. In Artificial Intelligence research, a modelling language for modular design of interacting traffic system was built based on multi-agent technology [3.25]. A street network model integrating one and two dimension CA was used as a heuristic tool for analyzing and observing the shape of state of street networks traffic flow in a whole city [3.26]. A traffic simulation environment and test development of intelligent transportation systems called GESTRAF uses an approach where the individual uniqueness has been stressed in the cell model. Each vehicle-driver pair has the same dynamic structure based on basic kinematics and structural behaviour of human-in-the-loop as a controller for speed and relative distance to other cars. The model includes also the decision making process from the driver in changing lanes [3.27].

3.5 The Proposed Traffic Flow Highway Modelling with CA

The implementation of this model is by translating the CA traffic flow rules into the computer simulation package NS-2. In the CA, the state of a street is described by firstly dividing the streets into cells of length 7.5 m. This corresponds to the typical amount of space occupied by a car length plus the distance to the preceding car in a dense jam. Each cell can now either be empty or occupied by exactly one car. Each vehicle is characterized by its current velocity $v$ which can take the values $v=0,1,2,..., v_{\text{max}}$. In the simplest case, $v_{\text{max}}$ corresponds to the speed limit and is therefore the same for all cars on the same lane.
A set of clearly defined rules is needed for this behaviour and Nagel and Schreckenberg [3.28] have suggested the following 4 steps:

i. **Acceleration**: if the velocity $v$ of a vehicle is lower than $v_{\text{max}}$ and if distance to the next car ahead is larger than $v+1$, the speed is advanced by one $\{v \rightarrow v+1\}$. This describes the drivers desire to drive as fast as possible within the speed limit. This is termed as free driving where it is assumed that there is no influence from preceding vehicles on the same lane.

ii. **Slowing down or braking due to other cars**: if a vehicle at position $i$ sees the next vehicle at site $i+j$ with $j \leq v$, it reduces its speed to $j-1$ $\{v \rightarrow j-1\}$. This means that there is a slower preceding car right in front. This step represents the interaction between cars and avoids collisions.

iii. **Randomization**: with probability $p$, the velocity of each vehicle (if greater than zero) is decreased by one $\{v \rightarrow v-1\}$. It is essential in simulating realistic traffic flow since the dynamic is completely deterministic otherwise. It takes into account the natural velocity fluctuations due to human behaviour or due to varying external conditions. Without randomness, every initial configuration of vehicles and corresponding velocities very quickly reaches a stationary pattern.

iv. **Car motion**: each vehicle is advanced $v$ sites. It means all cars move according to their new velocity.

The above rule set can be considered minimal and it is possible to produce more complex models. However, for the purpose of this thesis this has been assumed adequate.
Listed below is the summary of the algorithm which describes the implementation of CA in NS-2:

**Loading Velocity**

**Loading Node**

*Initialize Simulation time (SimTime)*

*For SimTime < 30*

**Define lanes**

**Compare gap between nodes (For movement of first node)**

If {First node of each lane then assign destination}

**Calculate gap**

**Do movement of each node (Iterate each node)**

**Calculate gap**

If { the Gap > Speed of the adjacent node (\(V_{a-1}\))}

Set \(V_{a-1} = \text{Assigned Speed}\)

else

Set \(V_{a-1} = \text{Gap only}\)

Set \(X_{a-1} = X_{a-1} + V_{a-1}\)

Set destination for each node for each second (Iteration)

*Produce trace file for performance analysis*

*Produce trace file for visualization*

In NS-2, the velocity and initial position of each node are created and put in separate files. The main consideration of doing this to apply the rule on randomization where the random command is used to position the nodes and assigning velocities so that the probability \(p\) is maintained.

The rule of acceleration and deceleration is shown where the gap is calculated compared if enough space is available based on the speed of both comparing nodes. The next movement is then defined according to the car motion rule.

All the nodes in each lane will go through the same process through iteration and the movements of nodes happen simultaneously for all lanes.

The final products are the trace file where visualization through NAM program of NS-2 provide the means of simulation validity by showing no nodes would over-ride each other and move according to the rules.
As for the simulation involving the highway scenario, there are assumed to be multiple vehicles travelling on both sides of the highway with 3 lanes on each side. All nodes are assumed to be equipped with sensing, communication, computation and storage capabilities and GPS, and have the ability to act as routers that can forward packets. Even though the transmission range is assumed to be the same for all nodes in the simulations, the proposed method mentioned in chapter 5 can be applied to any scenario with different coverage for each node because the transmission range is defined as a variable and that any permissible number can be assigned accordingly. The highway is assumed to be straight and the measurements are taken when the road is straight. There is also no change of lane and the node is moving according to the Cellular Automata approach where movement is within a single cell, where a cell is typically defined to measure one metre in length.

The proposed model has been developed using the traffic flow based on the CA method mentioned earlier. It represents a realistic movement of nodes on a 2-direction 3-lane highway with each lane parameterised for slow, medium and high speed and each set moving in the opposite direction. Each lane consists of 30 randomly distributed nodes or cars moving according to the lane speed without changing lanes for a simulated period of 30 s. The length of each lane is 1500 m with the width of 5 m. The nodes in the outer lane move at the speed of 20-24 m/s (72-86.4 km/h), middle lane at 25-29 m/s (90-104.4 km/h), and the inner lane at the speed of 30-34 m/s (108-122.4 km/h). There is an equal probability of the speed taking any value of the specified ranges ($\min \leq v \leq \max$). Figure 3.3 illustrates the settings of lane of the road according to the parameters mentioned in this paragraph and also how the positions would probably illustrates a typical snapshot of how the traffic would look during simulations.
There are two sets of simulations presented here that show 2 types of node connection. The first one is called Single Flow, and the conditions are:

i. Sending and receiving nodes are on the same lane (moving at the same direction)

ii. The sending node always following the receiving node at the front

Another set is termed Random where there is no restriction on how the sending and receiving nodes can communicate. The intention is to create a dynamic scenario which is similar to a VANET. In contrast, Single Flow is an ideal situation due to the location and movement direction of the nodes which should allow the connection to last longer. The aim of having the two sets is to show the results between an ideal situation and the dynamic situation so that network performances of these 2 characteristics can be studied. Figure 3.4 shows the initial position of the nodes before the simulation starts.
Other constraints are the ones provided by the NS-2 simulation software which includes the support of only flat or non 3D movement of nodes, and the antenna heights do not change throughout the simulation. Other assumptions are that all nodes have the same wireless interface and that the maximum distance of Carrier Sense Threshold at maximum transmission power level are the same for all nodes.

The transmission ranges used in the simulation are 20m, 50m, 100m, 150m, 200m, 300m, 500m, and 750m. The range is obtained by setting the power based on the circular distance in meters. Depending on the range, either the Friis or the Two-Ray formula is used. This is adopted from a method which has been proved to be successfully implemented in [3.29].
Table 3.2  Simulation Details

<table>
<thead>
<tr>
<th>No.</th>
<th>Items</th>
<th>Values</th>
<th>No.</th>
<th>Items</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>Outer Lane</td>
<td>20-24 m/s</td>
<td>v</td>
<td>Node distribution</td>
<td>Random</td>
</tr>
<tr>
<td>ii</td>
<td>Middle Lane</td>
<td>25-29 m/s</td>
<td>vi</td>
<td>Changing lane</td>
<td>No</td>
</tr>
<tr>
<td>iii</td>
<td>Inner Lane</td>
<td>30-34 m/s</td>
<td>vii</td>
<td>Length of each lane</td>
<td>1500m x 5 m</td>
</tr>
<tr>
<td>iv</td>
<td>No. of nodes in each lane</td>
<td>30</td>
<td>viii</td>
<td>Transmission Range</td>
<td>20m, 50m, 100m, 150m, 200m, 300m, 500m, 750m</td>
</tr>
</tbody>
</table>

The summary for all the aforementioned parameters are described in Table 3.2. The formulas for the propagation models are shown below:

Friis,

\[ P_r = P_t G_t G_r \left( \frac{\lambda}{4\pi R} \right)^2 \]  \hspace{1cm} (3.2)

Two-Ray,

\[ P_r = \frac{P_t G_t G_r h_t^2 h_r^2}{d^4 L} \]  \hspace{1cm} (3.3)

where \( G_t \) and \( G_r \) are the antenna gain of the transmitting and receiving antennas, \( \lambda \) is the wavelength, and \( R \) is the distance, \( P_r \) is the available power at the receive antenna, \( P_t \) is the power delivered to the transmit antenna, \( d \) is the distance, and \( L=1 \).

3.6  Conclusions

This chapter presents the 3 main elements required in order to understand the simulations that followed in the following chapters. They are the simulators, the details of NS-2, and the mobility models which include the transmission range issues. The reviews on the available simulators have been made and although OPNET provides good features it was not made available to the
research, which made NS-2 chosen because it is an easily accessible open source software package and free. Open source software provides flexibility and encourages creativity whereby implementations of new ideas can be implemented with acceptable accuracy of results. Both software packages are acknowledged for their abilities to perform network simulations despite their differences [3.30, 3.31].

The following chapters describe the execution of the relevant simulations based on the setup and parameters mentioned here.
References


CHAPTER 4:
PERFORMANCE ANALYSIS OF MANETs AND VANETs WITH NS2

4.1 Introduction

The purpose of the simulations in this section is to compare the performance of ad hoc routing protocols in various conditions where the nodes can be in a stable, moderately stable and highly unstable scenario based on the node density and mobility. AODV and DSR, both on-demand protocols, and DSDV the proactive protocol are the three protocols being simulated. The results enable the theories of the ad hoc networks to be examined practically and also to be used as baseline reference for the simulation studies in the current thesis.

Both Constant Bit Rate (CBR) and Transmission Control Protocol (TCP) traffic sources were applied using the same parameters throughout the simulations. This approach allows comparisons to be made of the performance of the routing protocols in various conditions.

As mentioned in the earlier chapter, Network Simulator 2 or NS-2 [4.1] is the main simulation software used in this thesis. It is a discrete event driven simulator targeted at networking research, which provides support for simulation of TCP, routing, and multicast protocols over both wired and wireless networks. The Rice University Monarch Project [4.2] has made extensions to the NS-2 network simulator that enable it to accurately simulate mobile nodes connected by wireless network interfaces, including the ability to simulate multi-hop wireless ad hoc networks.
4.2 Routing Protocol Performance Analysis in MANETs

The classifications of scenarios are based on the number of nodes which are 20, 50, and 100 for low, medium and high number of nodes respectively, and the speeds are 5, 15, and 25 metres per second (ms\(^{-1}\)) for low, medium and high respectively which covers a range of simulation conditions. These are equivalent to vehicle speeds of 18, 54, and 90 km per hour. Nine scenarios are created for the simulations with the combination of number of nodes and speeds. The combination details are listed in Table 4.1.

Table 4.1 Scenario Details

<table>
<thead>
<tr>
<th>Scenario</th>
<th>No. of Nodes</th>
<th>Node Speed (m/s)</th>
<th>Node Speed (km/h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Node/Low Speed (LNLS)</td>
<td>20</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>Low Node/Med Speed (LNMS)</td>
<td>20</td>
<td>15</td>
<td>54</td>
</tr>
<tr>
<td>Low Node/High Speed (LNHS)</td>
<td>20</td>
<td>25</td>
<td>90</td>
</tr>
<tr>
<td>Medium Node/Low Speed (MNLS)</td>
<td>50</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>Medium Node/Med Speed (MNMS)</td>
<td>50</td>
<td>15</td>
<td>54</td>
</tr>
<tr>
<td>Medium Node/High Speed (MNHS)</td>
<td>50</td>
<td>25</td>
<td>90</td>
</tr>
<tr>
<td>High Node/Low Speed (HNLS)</td>
<td>100</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>High Node/Med Speed (HNMS)</td>
<td>100</td>
<td>15</td>
<td>54</td>
</tr>
<tr>
<td>High Node/High Speed (HNHS)</td>
<td>100</td>
<td>25</td>
<td>90</td>
</tr>
</tbody>
</table>
The mobility model uses a random waypoint model in a rectangular field with a size of 1500 m x 1000m for the CBR and TCP traffic. The transmission range for each node is assumed to be uniform and is limited to 250 m. The radio wave propagation is assumed to be ideal and no channel fading is considered. Each packet starts moving from a random location to a random destination with the defined speeds. Once it reaches the destination, it goes to another random targeted node after a pause of 1.00 second. Each simulation runs for 900 simulated seconds.

### 4.2.1 CBR-Traffic Simulation Results

In this simulation, the data packet size is fixed at 512 bytes and is transmitted at the rate of 4 packets per second. The number of active connections is set to be at half the number of total nodes. This configuration is made by choice to represent a moderately active network. The communicating nodes are selected randomly.

The simulation results are plotted in Figure 4.1:
While DSR and AODV share the on-demand behaviour (in that they initiate routing activities only in the presence of data packets in need of routing), many of their basic routing strategies are different. In particular, DSR uses source routing, whereas AODV uses a table-driven routing framework and destination sequence number [4.3]. The simulation results show that AODV and DSR have almost identical performance when the nodes and sources are low, with DSR slightly outperformed AODV. By using source routing, DSR has access to a significantly greater amount of routing information than AODV through caching. Also, in DSR, using a single request-reply cycle, the source can learn routes to each intermediate node on the route in addition to the intended destination. Each intermediate node can also learn routes to every node on the route [4.4].

As the network traffic gets increasingly heavy, AODV maintains its performance while DSR begins to decline as it turned to be at the most strained condition. This is due to the DSR caching becoming less effective at higher speeds where the cached information became stale much faster [4.5].

The proactive protocol DSDV was unable to proceed in a heavily loaded scenario where it could only achieve partial functionality in the low-node high-speed scenario and was unable to function at all in the more challenging scenario. The nature of the proactive protocol means that it does not work well in a dynamic scenario since the routing table cannot be updated quickly enough, thus making the entries in the table stale, causing the packets to be forwarded over broken links. Since DSDV maintains only one route per destination, each packet that the MAC layer was unable to deliver was being dropped due to there being no alternative routes.

4.2.2 TCP-Traffic Simulation Results

TCP is a protocol which is designed to give highly reliable and in-order delivery of sender to receiver data and which is why the simulation results show a very high delivery rate. In the scenario of minimum nodes and lower speed the delivery rates are almost 100%. The reason for this is that dropped packets are resent, so any deviation from 100% is due to packets which have not arrived when the simulation has ended. The results would vary for different simulation times.
Table 4.2   TCP Packet Delivery Rate for Various Speed and Nodes

<table>
<thead>
<tr>
<th></th>
<th>AODV</th>
<th>DSDV</th>
<th>DSR</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNLS</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>LNMS</td>
<td>0.98</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>LNHS</td>
<td>0.98</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>MNLS</td>
<td>0.98</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>MNMS</td>
<td>0.97</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>MNHS</td>
<td>0.97</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>HNLS</td>
<td>0.98</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>HNMS</td>
<td>0.96</td>
<td>0.0</td>
<td>0.97</td>
</tr>
<tr>
<td>HNHS</td>
<td>0.95</td>
<td>0.0</td>
<td>0.98</td>
</tr>
</tbody>
</table>

This is a valid performance metric, providing the simulation time is constant for the cases being compared. The results for applying TCP to the same scenarios discussed in the last section are in Table 4.2.

4.2.3 Simulations with Fading Channel

The simulation shows a consistent set of results with the earlier simulations without the fading. It shows slightly lower delivery rates reflecting a more realistic result. The set of results labelled rcAODV and rcDSR are shown alongside the previous results for comparison in Figure 4.2.
4.2.4 Overhead Performance

The measurements of overhead show the efficiency of the routing and the effective use of wireless medium by the data traffic. This section provides the overhead analysis from all the simulation results. The actual results are presented in Figures 4.3 – 4.5.

All the results show a similar overall pattern with the overheads for both routing and MAC packets increased as the number of nodes and speed increased. TCP produced a lower overhead when compared to CBR packets.
Figure 4.3  Routing and MAC overhead for CBR traffic

Figure 4.4  Routing and MAC Overhead for TCP traffic
AODV produces more overhead than DSR because each of its route discoveries typically propagates to every node in the network. DSR has the lowest number of packets but higher than AODV if measured in bytes. DSDV has an approximately constant overhead regardless the speed due to its proactive nature.

The simulation study demonstrates that overall AODV performed better in the majority of the scenarios. The CBR traffic shows more variable results in the packet delivery ratio when compared to TCP. DSDV managed only to perform well in a more predictable physical arrangement of nodes. TCP in general produces a lower MAC and routing overhead when compared to CBR. Of the on-demand routing protocols the experiments demonstrate that the MAC and routing overheads for AODV are much higher, however much better performance in delivery route can be achieved.
The incorporation of the Rayleigh Fading channel in the simulation is expected to give some insights on the effect of fading for future work which is the performance of routing protocols in a vehicular environment.

With these results, the AODV routing protocol has been chosen for more complex simulation for the work in chapter 6 of this thesis.

4.3 Effects of Transmission Range on VANETs

In VANETs the transmission power is an important issue [4.6, 4.7]. There are two contradictory requirements which must be balanced. Firstly the requirement that the transmission range is sufficient to allow communication with adjacent nodes. In areas of low node density this dictates a large range. Secondly the requirement is that the communications band does not get congested by network traffic. This dictates a short range so that in areas of large node density it is still possible to get sufficient throughput. This section examines the effects of various ranges of transmission in which the findings can be used as a reference in a later chapter involving manipulating the coverage area to improve performance.

Another pertinent issue for transmission power is that MANET deployments are usually powered by batteries with finite capacity. Studies such as [4.8] have proposed a transmission strategy that can determine the optimum transmission range that gives maximum efficiency of energy consumption. In VANETs, however, the nodes can be assumed to have a stable and ample power source that comes from the vehicle. Therefore, in this thesis the issue of power supply is assumed to be non critical and it is assumed the nodes can provide sufficiently high coverage area when needed.

In the practical wireless environment, the optimal range cannot be represented by a fixed number as the value changes as it responds to interference and constraints from the environment. In addition, the coverage may not be truly isotropic although that is the necessary assumption in the simulations. One characteristic of transmission range is that longer range does not necessarily
equate to better network performance [4.9]. Even though a longer transmission range reduces the number of hops in a route, it also increases the number of nodes that compete on the shared channel. On the other hand, a shorter transmission range allows better reuse of bandwidth and minimal use of power and also throughput can be improved because more simultaneous transmissions are allowed at the same time in different areas of the network. The disadvantage of having a shorter range is that it creates network partitioning which could prevent some nodes from communicating with their destinations. This should also be avoided, thus a balanced range is the choice according to the required situation.

In the simulations in this thesis, the assumption is that any node can be successfully heard by any other node providing a distance between them is to be less than a certain Euclidean distance in flat two-dimensional space without any real propagation obstacles that could potentially block or reflect signals.

The following graphs have been produced from the simulation exercise. The interpretation and analysis will be based on the performance metrics mentioned earlier.

4.3.1 Effects on the Packet Delivery

The simulation model can be categorized as in the mid-range for nodes density. Each node is being distributed fairly randomly with 2 neighbouring nodes in the same lane can be separated to the maximum of 125 m and as low as 10 m.

With this setup, it is shown in Figure 4.6 that a low transmission range of 20 m produces lower packet throughput because of the broken links or isolated clusters it causes. It becomes better as the transmissions enter the range of 100 – 200 m as it covers most of the nodes and supports better multi-path route to the nodes. As the transmission range increases exceeds 300 m, the performance also declines. Although the coverage area becomes bigger this situation also makes the nodes compete for channel bandwidth, making the average amount of radio capacity per user very small. This supports the assertion made earlier in this chapter regarding the optimum transmission range.
Another set of simulation results which uses randomly connected nodes is also shown with a poor performance of packet transfers. It reflects a similar situation as in the unorganized movement as in the Random Walk model. The routing path can easily break when the nodes are moving especially in opposite direction which is obvious in this random model. This exercise highlighted the differences between the random and properly arranged connections of nodes scenarios.

Figure 4.6  Effect of transmission range on packet delivery

It can also be seen that the number of dropped packets consistently increased with the higher transmission range as shown in Figure 4.7. This supports the earlier argument where the consequence of competing for resource could only lead to more packets being dropped instead. All packets here include every type of packet recorded in the NS-2 trace file.
In Figure 4.8, the single flow connection produced consistent performance due to the mobility and connection pattern allowing the overall packet transmission to be stable. It only shows a slight decrease in the range between 300m to 700m. The random connections show a low delivery ratio at the shorter ranges and start to stabilize from 100m onwards. High density conditions and high mobility are the possible contributing factors to this performance. The dropped packet ratio does not change after 100m as this could possibly be caused by the increasing distance and thus making the higher transmission range perform well.

Figure 4.7  Dropped packets against all packets
4.3.2 Effects on Overhead packets

The measurement of routing overhead evaluates the efficiency of the routing in the protocols while MAC overhead measures the effective use of wireless medium by the data traffic. In Figure 4.9, the lower transmission ranges have the best results but it was a situation where the number of nodes in the neighbourhood is low and that did not leave many chances for hopping to occur.
The optimal range of transmission still produced a reasonable number of overhead packets which is less than half of the successfully received packets. As the transmission range became higher, nodes were able to have more opportunity to try and repair the route. These efforts are the possible causes of making the number and the ratio of overhead packets high.

### 4.4 Discussions and Summary

The technical challenges of MANETs have been approached by numerous research studies. Determining a ‘typical’ use of a MANET is extremely difficult due to the wide variability of the problem. Strategies have been proposed which approach the problem from a variety of different angles. Examples are issues on flow and error control, maintaining accurate network topology information, handling node as router mobility, shared channel access, and power requirements. As the work focuses further from MANETs to VANETs, more specific challenges can be seen, for example the effects of variable speed of static, low and high speed vehicles, traffic density.
and flow, antenna direction, electromagnetic effects, and heterogeneous devices. Some of these will be looked into in the following chapter.

This research will also look into the solution through enhancement of TCP via link management in order to improve the performance of routing protocol in VANETs.

4.5 Conclusion

In this chapter MANETs and VANETs in general, and the routing protocols have been reviewed. The literature also appears to support the conclusion that the AODV routing protocol is one of the better approaches to be used on the study of VANETs environment as compared to the other two routing protocols. A series of simulations have been carried out to support the information already available from the theory. This was done using the packet delivery ratio and the routing overhead as the performance metrics. Although there are other challenges in MANETs and ways to resolve them the assumption made in this thesis that things can be done to the routing protocol in order to improve its overall performance. The same can be said on the selection of the routing protocol as each routing protocol has its own merit and advantages, and this chapter shows the justification of AODV for the purpose of this research on the VANET environment.
References


CHAPTER 5:

NOVEL PERFORMANCE METRICS

This chapter discusses the core components of the proposed method for the enhancement of performance which are the reliability and availability of the path and links. It starts with an introduction as to why this approach is taken, followed by a discussion of the literature, finally the chapter concludes with the presentation of theories for the proposed method.

5.1 Introduction

There are numerous opportunities to enhance performance in VANETs by exploiting their specific behaviour. There are also many levels within the OSI model where this can be done, including the possibility to make use of multiple layers, or cross-layers. The focus here is on the network layer where a routing protocol can be enhanced for improving performance by better maintaining the links in a route. The main concept introduced is the proposed use of link availability as a metric, which is estimated from positional and velocity vector data at the nodes, rather than being determined absolutely. This metric can be used to more accurately determine a good link than conventional techniques.

There are many other approaches that can be used to enhance a performance involving ad hoc networks. For examples in VANETs, [5.1] proposed a cooperative communications approach where packets are filtered to allow only those from the participating nodes to go through and forwarded, and in [5.2] a congestion control mechanism was implemented using enhanced packet forwarding and congestion control on top of the IEEE 802.11 MAC protocol. The aim is so that it can avoid the starvation of some nodes in the network producing information dissemination and fairness.
In this thesis however, the focus will be on the link availability and reliability. In a network, a communication between a sending and receiving nodes will at least consist of a link if not multiple links forming a route. It is logical that if a route can be maintained for as long as possible the performance is better as more data can be transferred without interruption. In the case of VANETs, nodes are actively moving and the movement is somewhat uniform. This makes the proposed approach even more viable to implement since it becomes more predictable to determine the locations.

5.2 Availability

Availability has various meanings and ways of being computed depending upon its use. In communications it is usually defined as a degree to which a system or task is in an operable state at the start of a mission, when the mission is called for at an unknown, random time. It is in effect the proportion of time a system is likely to remain in a functioning condition. In this study, ‘operable’ is defined as when a link in a path can be considered to be connected. A link is considered available as long as the ratio between distance and transmission range is greater than 0, which is discussed further in the later section.

In the VANETs scenario, a route usually consists of many links formed by the nodes on the streets or highways, and in many instances there is more than one route at a time. The availability for each of these links reflects the reliability of the particular route. The ability to identify the availability and reliability of the two elements would allow the routing protocol to choose the best route to maintain the communication between the sending and receiving nodes when there is a need to re-route.

5.3 Reliability

Widely used in engineering, reliability is used in many areas such as manufacturing, electronics, and also computer systems [5.3]. Reliability is generally defined as representing the probability of systems to perform their required functions for a desired period of time [5.4], [5.5]. In this context, probability shows that reliability deals with the rules of random chance or expected
events. By being able to determine the frequency of occurrences, a particular target can be achieved to determine the reliability of the system [5.6].

Even though probability theory is involved, it is important to have a complete understanding of the technical implications of the system in order to evaluate the reliability [5.7]. In this research, communication in VANETs will involve many nodes and routes or path in which a failure to perform is when any link connecting two nodes gets disconnected as the distance goes out of the transmission range. A failure in even one link in a path can effectively determine the reliability of the path. Eventually, a model will be derived to represent the characteristics and an algorithm will be proposed for the solution of ensuring reliability hence enhancing the overall network performance.

5.4 Link Availability

In ad hoc networks link availability is usually used to enable the selection of the best available route where better link availability means the route itself can last longer [5.8]. Nevertheless, in most cases the purpose of identifying link availability is so that the best route can be selected from the available alternative routes in the case of pre-emptive action [5.9].

In order to do this, there is a need to know the location of each node and with the mobile characteristics of ad hoc networks, prediction methods based on the mobility model are required. One popular approach is by calculation over the epoch where each node’s movement consists of a sequence of random length intervals during which a node moves in a constant direction at a constant speed which varies randomly from epoch to epoch [5.10]. The literature has shown that this method is usually used in MANETs with the Random Walk Model representing the movement [5.11]. Other approaches are similar in [5.12] where the received signal strength indication (RSSI) at the lower layer is used to estimate the locations and the link availability is calculated using the difference between the node time of arrival and time differential of arrival. [5.13] utilizes a Markov chain model to describe the relative random between two nodes In addition of using epoch to predict locations [5.14] considered channel capacity to determine link availability.
Also based on RSSI, [5.15] added a packet called the Physical Layer Convergence Protocol as the performance metric. One case in VANETs [5.16] uses the differential of RSSI on a proactive protocol. Other approaches are [5.17] where a specific topology allowing formation of clustering is used to predict mobility, [5.18] that estimates the rate a route expands or contract to indicate increasing or decreasing distance where less dynamic route lasts longer, and [5.19] which proposes a Steady State Connection Availability which is the summation of all no-faults states using the Continuous Time Markov Chain to show the state.

A link is considered available when the distance between nodes is less than the transmission range. Determining if a node is available is therefore trivial: if the radio signal strength is adequate then the link is available. What is not trivial, however, is estimating the probability that the link is available at some future point in time. This future availability cannot be known for certain in a VANET as it is impossible to know for sure where the nodes will be located in the future or indeed how long in the future it must be considered when calculating the availability.

Thus, this thesis applies various models which take into account factors such as the direction of nodes, their positions, speed etc. in order to form an estimate of the future availability of a link. It will be shown that estimates of availability can be used to determine the ‘health’ of a particular route, and hence form better evidence when selecting the optimum initial route.

5.5 Route Availability

Link availability is usually a subset of Route Availability as mentioned in [5.20], although there are other approaches which put more emphasis directly on the route itself. Clustering using the DSR routing protocol [5.21] enables the prediction of node movement restricted to the last movement legs of the node to ensure higher accuracy. The Reliable Probabilistic algorithm [5.22] is an extension to the GRID routing protocol which defines a new parameter called grid head stability to select stable grid head. For each grid, one mobile host will be elected as the grid head and others can go into sleep mode. The grid head utilises probability and is responsible for forwarding routing information and propagating data. The element of hop has also been proposed as in [5.23] where the power of nodes is increased to reduce the number of hops hence reducing...
the probability of having the tendency of each link to break. A mathematical model has also been
developed [5.24] to estimate path duration in a generic number of hop paths when nodes follow a
modified Random Way Point mobility model.

5.6 Proposed Approach: Good Link and Route Availability

The proposed approach consists of 3 main steps as shown below in Figure 5.1:

![Figure 5.1 Performance metrics]

This section presents the theory of each process which integrates into one practical model.

The location of the individual nodes are required and for this purpose the assumptions are: (a)
all nodes are equipped with GPS units or similar and can determine their position at any given
time, (b) the two nodes in a given potential link are aware of each others’ coordinates and
velocity vectors. These assumptions are made because the focus of this method is on the routing
protocol procedures.

5.6.1 Location Prediction

This process provides an estimate for how long a two-node link takes to go out of range in
seconds. Referring to the information from the routing table the location is then applied on a
Cartesian coordinate system where the nodes move in a 2-D vector direction in x and y. Using
the Pythagorean Theorem, the Euclidean distance can be derived as shown in Figure 5.2 where
\((x_0, y_0)\) and \((x_1, y_1)\) are points in the plane with the inclusion of time and velocity in the equation
so that the time when the distance between the two nodes greater than the transmission range can
be computed.
Node A starts from location \(x_A, y_A\) and node B from position \(x_B, y_B\). Each node moves with a known velocity vector which is assumed constant during the calculation.

The position of A and B after the time \(n\Delta t\), where \(n\) is step number and \(\Delta t\) is step length are as follows,

\[
A_n = A_0 + v_A n\Delta t, \quad B_n = B_0 + v_B n\Delta t
\]

for \(|A_n B_n| = D\), which is the distance between two nodes (used to determine if the link is out of the transmission range). The link life time \(t\) is defined as the time it takes from the link being first initiated, \(t_0\), until it reaches the location where the distance between the two nodes \((D)\) reaches the transmission range \((R)\) \(t_1\).

\[
A_n = \begin{pmatrix} x_{n0} + v_{nx} n\Delta t \\ y_{n0} + v_{ny} n\Delta t \end{pmatrix}, \quad B_n = \begin{pmatrix} x_{n0} + v_{nx} n\Delta t \\ y_{n0} + v_{ny} n\Delta t \end{pmatrix}
\]

\[
|B_n A_n| = \sqrt{A_n^2 + B_n^2}
\]
\[
\begin{align*}
&= \sqrt{(x_{t_2} + v_n \Delta t) - (x_{t_1} + v_n \Delta t)^2 + (0, y_{t_2} + v_n \Delta t) - (0, y_{t_1} + v_n \Delta t))^2}
\end{align*}
\] (5.1)

The time at which the point of time the distance \( |A_nB_n| \geq D \) is defined as the Link Life Time. This will later be used to calculate the route reliability mentioned in the following section.

As mentioned earlier, the nodes are moving in a rather uniform way in VANETs. There are four possible instances when the nodes are connected:

i. **Each node is moving in the same direction**

This is a common scenario in VANETs but it can produce two different outcomes. If an approaching node is moving much faster it creates a situation where the connection is strong when both nodes are initially approaching each other, and gradually weakens with possibility of getting disconnected as they are moving away. It could also be an infinite connectivity if the speed is not fast enough for a node or the same for both nodes.

ii. **Each node is moving away from each other in different directions**

This is a straightforward case where the connection will be disconnected after some time regardless of the speed. An example is when each of the two connecting nodes is travelling on the different side of the lane and going to different directions.

iii. **Each node is connected to the same axis or lane**

This causes the value for one of the axis to be zero when calculating using the Pythagorean Theorem because the link cannot create a triangle. Otherwise the situation is the similar to in (i). The approach is still applicable in this case which is
mostly probable in the road case scenario. It is possible for infinitely long connection to occur if the speed is the same or very similar.

iv. Only one node is moving

The link will get disconnected as the distance between the two nodes gets out of the transmission range.

The above instances can happen in VANETs such as in a highway scenario. A simple set of simulations referring to the above scenarios has been carried out in order to validate with the parameters as mentioned below in Table 5.1.

Table 5.1 Comparison details between calculated and simulation of the time the first packet dropped

<table>
<thead>
<tr>
<th>Sim.</th>
<th>Node A</th>
<th>Speed</th>
<th>Dest.</th>
<th>Node B</th>
<th>Speed</th>
<th>Dest.</th>
<th>Calculation</th>
<th>Simulation</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>i.</td>
<td>50,50</td>
<td>10</td>
<td>200,50</td>
<td>80,80</td>
<td>15</td>
<td>300,80</td>
<td>23.39387</td>
<td>23.463837701</td>
<td>0.069968</td>
</tr>
<tr>
<td>ii.</td>
<td>50,50</td>
<td>10</td>
<td>200,50</td>
<td>80,80</td>
<td>15</td>
<td>200,80</td>
<td>7.078774</td>
<td>7.119820598</td>
<td>0.041047</td>
</tr>
<tr>
<td>iii.</td>
<td>50,50</td>
<td>10</td>
<td>300,50</td>
<td>100,50</td>
<td>15</td>
<td>700,50</td>
<td>20.0</td>
<td>20.055375414</td>
<td>0.055375</td>
</tr>
<tr>
<td>iv.</td>
<td>50,50</td>
<td>0</td>
<td>none</td>
<td>75,75</td>
<td>15</td>
<td>300,75</td>
<td>8.193463</td>
<td>8.232692541</td>
<td>0.03923</td>
</tr>
</tbody>
</table>
The results in Table 5.1 and in Figure 5.3 have shown that the estimated calculations are very close with the results in the simulations which referred to the time of the first dropped packet. The difference average is 0.05 seconds. The reasons of this are the fact that the simulations take into considerations the effects and restrictions as much as possible in its parameters while the calculations is absolute.

5.6.2 Good Link Availability

Once link life time has been determined, the next step is to get the availability of each link which will be termed as Good Link Availability (GLA).

Availability of a link to be good, or GLA,

\[
P(G) = 1 - \frac{D_{\alpha}}{T_{\text{Xrange}}}\]

(5.2)
where $D(t)$ is the distance of the two nodes at the time of $t$, while $T_{\text{range}}$ is the transmission range of the two nodes, assumed to be the same. This equation is valid for $D<T$ and that $P(G)=0$ for $D\geq T$.

GLA indicates the availability based on the degree of stability by referring to the proportion of distance at time ($t$) against the transmission range. A higher value shows a possible higher degree of availability since the position between the two nodes is closer while lesser value shows the opposite.

5.6.3 Good Route Availability

When the need to re-route occurs, the alternative route is selected based on the next available route with a good reliability of sustaining the route and this is called Good Route Availability (GRA). Before this happens, the existing route should have already been able to predict which other route has the highest availability of being able to sustain route life time over a certain period of time.

Figure 5.4 shows an example of links and routes in VANETs. Assuming that node L2-1 is the source node and L2-4 is the destination, there are many possible routes available for this link, some of these possible routes are listed down below:

route 1: L2-1 > L2-2 > L2-3 > L2-4
route 2: L2-1 > L1-1 > L1-2 > L1-3 > L1-4 > L2-4
route 3: L2-1 > L2-2 > L3-2 > L2-3 > L3-3 > L2-4
For the purpose of describing the theory, it is assumed that route 1 is the current route. When travelling at a medium speed the vehicles are moving and there is a strong tendency that link L2-3 > L2-4 will break. Each link in a route knows its GLA and calculates the availability of the link to be alive within the stipulated speed (P(A)). Therefore the GRA for each route can be obtained by getting the product of the availability of each link in the route.

\[
P(GRA) = \prod_{i=0}^{n} P(GLAnode \ i)
\]  

Table 5.2 Examples of Calculating Good Route Availability

<table>
<thead>
<tr>
<th>Route</th>
<th>Links in Route</th>
<th>Sum</th>
<th>P(Gra)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>L2-1 &gt; L1-1 &gt; L1-2 &gt; L1-3 &gt; L1-4 &gt; L2-4</td>
<td>0.95 * 0.88 * 0.77 * 0.65 * 0.89</td>
<td>0.3724</td>
</tr>
<tr>
<td>3</td>
<td>L2-1 &gt; L2-2 &gt; L3-2 &gt; L2-3 &gt; L3-3 &gt; L2-4</td>
<td>0.99 * 0.79 * 0.88 * 0.77 * 0.90</td>
<td>0.4769</td>
</tr>
</tbody>
</table>

The source node has the information of GRA for each neighbouring node and chooses the route with the highest GRA for re-routing. In this example, route 3 has the highest availability 0.4769 and this then can be chosen for re-routing. Note that the number of hops is the same and hence the two routes cannot be distinguished in this way.
5.7 Discussions and Conclusions

5.7.1 Discussions

The literature review of the other work has shown that focus on the link status has been one of the established approaches to enhance performance. This is usually be used for pre-emptive actions to be taken before a link is broken. Although the works have been based on similar principals to that described here there are some differences from the proposed method. The characteristics of the work in the literature can be listed as follows:

i. The approach focuses on link availability only

ii. Solutions are based on purely theoretical mathematical models

iii. Route availability does not integrate directly with link availability

iv. Use of mobility models that are more suitable for MANETs such as the use of Random Walk Mobility model

In comparison, listed below are the highlights of the advantages of the proposed method:

i. Integrates link and route availability to produce GLA and GRA

ii. Use of location prediction that has the potential to work well in the VANETs environment and also suitable with the Cellular Automata (CA) applied for the node movement in the simulations

iii. Practical mathematical and statistical models used in the location prediction and calculation of GRA respectively

iv. Provides simple, efficient, and practical solution to be implemented with the aim to reduce controls overheads so that it can help to increase the performance
Link availability here shows the proportion of the transmission range against the current distance which is the probability of likeliness that a link will break or otherwise. In ensuring the reliability of the path or route, the probability is used where the rule of simultaneous occurrence of independent events is applied.

In the proposed method, the location prediction and distance measurement can be continuously calculated where they can be known at any time \( t \). The use of product to identify GRA is consistent with the simultaneous occurrence of independent events probability theory which produces acceptable chance of success in selecting the best route.

### 5.7.2 Conclusion

In this chapter metrics are produced which may be used for performance enhancement in Vehicular ad-hoc networks (VANETs) protocols. The method described makes use of the concept of link availability estimates which are produced on an individual link basis and which may be combined by multiplication to produce a quality metric for the whole route. The technique is discussed in the context of the Ad hoc On-demand Distance Vector (AODV) routing method and this will be discussed further in the next chapter along with the simulations and implementation framework. Work on validations of GRA in selected cases will also be presented.
References:


CHAPTER 6:

MODELLING AND EVALUATION OF PERFORMANCE METRICS

This chapter presents the simulation and analysis on the Good Link Availability (GLA) and Good Route Availability (GRA) metrics discussed in the previous chapter. It also describes how such metrics could be incorporated into a protocol.

6.1 Introduction

The chapter continues the discussion of GRA and GLA with simulations and analysis. The first series of simulations attempts to show the importance of focusing on re-routing by calculating the estimates of how much data could be lost, and the expected saving by applying GLA and GRA. The second group of simulations involves the selection path pattern for a re-route. It shows the routes taken and compares the results between the simulations and the proposed method using GLA and GRA.

The remaining part of the chapter lays out the framework for the implementation of the proposed method. It starts with the AODV process review which also mentions the functions of control packets. Following is the description of the main process which consists of 3 phases namely Identifying and Passing Node Information, Calculating Link and Route Probabilities, and Assigning Pre-emptive Actions.
6.2 Simulation Set up and Parameters

There are 5 sets of simulations with nodes in the range of 3 to 7 which are randomly distributed. The aim is to show the results in various situations, with only one node moving in one pair of transmissions using FTP. In each simulation only one node is moving horizontally to indicate that there is no change of lane in the case of VANETs. The details of the simulation parameters are shown in table 6.1. The initial position of the nodes is shown in table 6.2.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wireless channel</td>
<td>Two-Ray Ground propagation model</td>
</tr>
<tr>
<td>Media Access Control</td>
<td>MAC802.11</td>
</tr>
<tr>
<td>Antenna model</td>
<td>Omni-directional</td>
</tr>
<tr>
<td>Topography dimension</td>
<td>3500x670 (in m)</td>
</tr>
<tr>
<td>Number of packets</td>
<td>3-7</td>
</tr>
<tr>
<td>Routing protocol</td>
<td>AODV</td>
</tr>
<tr>
<td>Transmission range</td>
<td>150 m</td>
</tr>
<tr>
<td>Node Speed</td>
<td>25 m/s</td>
</tr>
<tr>
<td>Simulation time</td>
<td>7 seconds</td>
</tr>
<tr>
<td>Node movement</td>
<td>Cellular Automata</td>
</tr>
<tr>
<td>Traffic Protocol</td>
<td>TCP</td>
</tr>
</tbody>
</table>

The simulations cannot be claimed as a complete validation of the method, however they do provide some confidence of the validity of the technique.
6.3 Re-routing: Measurement and Analysis

One of the benefits of having the ability to predict locations and the time a link would break is to be able to calculate the time taken to do a re-route and the amount of data lost during the process. The data is obtained from the NS2 trace file where by referring to the calculated time the measurement is taken exactly with the difference when the first packet dropped and the first packet resumed. The data involved during that period is then counted where there 2 types of packets involved which are the TCP data packets and the control packets.

Table 6.2 Re-routing Measurements

<table>
<thead>
<tr>
<th>Simulations</th>
<th>Nodes Initial Positions</th>
<th>Re-route Time (ms)</th>
<th>Data Equivalent (Kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-node</td>
<td><img src="3-node_graph.png" alt="Graph" /></td>
<td>4.472</td>
<td>17.911</td>
</tr>
<tr>
<td>4-node</td>
<td><img src="4-node_graph.png" alt="Graph" /></td>
<td>8.032</td>
<td>22.302</td>
</tr>
<tr>
<td>5-node</td>
<td><img src="5-node_graph.png" alt="Graph" /></td>
<td>16.472</td>
<td>18.973</td>
</tr>
<tr>
<td>6-node</td>
<td><img src="6-node_graph.png" alt="Graph" /></td>
<td>9.012</td>
<td>24.434</td>
</tr>
</tbody>
</table>
The latter type consists of TCP acknowledgement packets (ACK), the AODV control packets, and the CBK packets. In NS2, CBK is referred to as the MAC call-back which indicates that the MAC layer was not able to transmit the packet and is therefore informing the upper routing layer about the transmission failure which usually is due to link failure.

Referring to the results in Table 6.2, the time taken for the AODV routing protocol to do a re-route ranges from 2–16 ms for a 7-s simulation and the data equivalent that could have been involved is from 13–24 Kbits. Although the results may not be linear with time it still shows quite a substantial amount of data can be lost without any pre-emptive actions.

The fact is that AODV uses Hello packets to identify alternative links and routes. The method proposed in this thesis would be able to avoid this as the function of GLA and GRA is to identify the best links and routes along with the time to do the pre-emptive re-routing. The simulations have successfully indicates the data lost that can be avoided.

### 6.4 Selecting a Path for Re-route

This section presents a comparison between the proposed method and the actual AODV routing protocol route selection for re-route from the simulations.

It appears that AODV has its own ways of selecting the path in which sometimes may and may not select the route selected by the proposed method as shown in Table 6.3.

The 4-node simulation shows the same length and distance hence does not make any difference to the proposed method in which case the choice of selection is return back to the default routing
protocol way. The 5-node simulation shows that the selected route is the same for both methods. The other two sets of simulation shows different selections of routes. The AODV routing protocol depends on the earliest hello packet received from its nearest neighbours, and in this case the node which is the nearest will be the preferred choice.

Table 6.3 Route selection comparison

<table>
<thead>
<tr>
<th>Simulations</th>
<th>Possible Routes</th>
<th>GRA for each route</th>
<th>Chosen GRA</th>
<th>Route in Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-node</td>
<td>0→2→1</td>
<td>0.082</td>
<td>Any</td>
<td>0→3→1</td>
</tr>
<tr>
<td></td>
<td>0→3→1</td>
<td>0.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-node</td>
<td>0→2→1</td>
<td>0.015</td>
<td>0→2→1</td>
<td>0→2→1</td>
</tr>
<tr>
<td></td>
<td>0→3→1</td>
<td>0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0→4→1</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-node</td>
<td>0→2→1</td>
<td>0.235</td>
<td>0→4→1</td>
<td>0→3→1</td>
</tr>
<tr>
<td></td>
<td>0→3→1</td>
<td>0.169</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0→4→1</td>
<td>0.249</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0→5→1</td>
<td>0.109</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7-node</td>
<td>0→2→1</td>
<td>0.143</td>
<td>0→2→1</td>
<td>0→6→1</td>
</tr>
<tr>
<td></td>
<td>0→3→1</td>
<td>0.122</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0→4→1</td>
<td>0.117</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0→5→1</td>
<td>0.082</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0→6→1</td>
<td>0.021</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Meanwhile, in the proposed method the calculation shows a route with a well balanced links where the transmission strength of each link is taken into consideration.
6.5 Implementation Framework

Before proceeding with the discussion of the implementation, it is necessary to understand how AODV works.

Ad Hoc On-Demand Distance-Vector Routing (AODV) is the routing protocol chosen by means of an example in this study. It provides communication between mobile nodes with minimal control overhead and does not maintain routes from every node to every other node in the network. Routes are discovered on an ad hoc basis and maintained only for as long as they are necessary. It is loop-free because it increases the sequence number each time it learns of any change in the topology of its neighbours. The sequence number ensures that the most recent route is selected for a route discovery.

Control packets are used in the AODV protocol to establish the link, provide data transmission, and also maintenance. A Route Request (RREQ) packet is used to initiate the route discovery process. Route Reply (RREP) is used to finalize a route. A Route Error (RERR) packet is used when there is a link break.

The AODV process description starts with a node, called the source node (SN) which begins to execute packet transmission to a destination node (DN). Any node involved in creating a link will first look at its own routing table to see if the DN is in the routing table. Referring to the flow chart in Figure 6.1, it is assumed that the DN is not in the table and therefore starts with SN broadcasting RREQ packets to its neighbouring nodes.

The node which received the RREQ packet will see if it is the destination by checking the IP address in the packet. If it is not the DN, the node adds the route reverse entry in its routing table to be used as a route for the reply packet from DN and data packet by the SN. The forwarding of packet continues until it reaches the DN, and if it fails to be delivered after a specified time the packet will be dropped.
As the RREQ packet reaches the destination, the DN will acknowledge by replying with a RREP packet which will go to the SN via the reverse route. Upon receiving the RREP packet, the SN starts the transmission by forwarding the data packets via the reverse route entry in the routing table of each node in the route.

![Figure 6.1 AODV Process Flow Chart](image-url)
A ‘hello’ packet continues to be periodically broadcasted even when a route has been established and communication between SN and DN is in progress to ensure that a link is still available or active. If a transmission is completed and the route is sensed to be idle after a specified time then the link is dropped. Whenever a link is broken by the way of idling, or by the detection of the use of Hello packets, or by the notification from the MAC layer, RREQ packets will be issued by the surviving node to its neighbours downstream and eventually reaching the SN to inform that the link is broken. In a situation if the link break is near the DN, a local repair will be carried out. The node will create a buffer to store the incoming data packets will try to re-establish link and resume if it is successful, otherwise the link is dropped and RREQ is issued.

Illustrating the process, as an example Figure 6.2 shows a source node that initiates the route discovery by broadcasting RREQ packets to its neighbours. Each of nodes A₁, B₁, and E₁ receives the packet and look up at each own routing table to see if it is the destination node.

As it is not, each node updates its routing table with a reverse route entry to be used for both RREP to finalize the discovery process and also for the data transmission. The node forwards the
RREQ packet to its neighbour, repeating the same process described earlier until it reaches the destination. In this example, the scenario applies to the route with the $B_n$ nodes.

Apart from the straightforward route discovery mentioned earlier, there are also other instances that need to be considered:

A node receiving 2 or more RREQ packets: $B_1$ receives packets directly from SN and followed by $E_1$. $B_1$ refers to its routing table and updates it with the information from SN. Subsequently, as it receives the packet from $E_1$ with the same destination and source it compares the number of hops and sequence number which then node $B_1$ decides that packet from $E_1$ is stale and thus drops the packet.

Destination node not found: Node $A_1$ forwards the packet with the same process as any forwarding node would do. As it reaches node $A_3$, one packet will be forwarded to $D_1$ and eventually reaches $D_3$ with no further node to forward to. In this case the node drops the packet and the route becomes void. Also assuming that there are many nodes the $D_n$ nodes route, the packet will be dropped after the forwarded packet still cannot reach the destination after certain period of time.

6.5.1 Main Processes

The implementation of the proposed functions can be classified into 3 main processes:

A. Process I: Identifying and Passing Node Information

Firstly it is necessary to pass information between nodes regarding their speed and velocity. The calculations of location predictions and probabilities for link and route stabilities are calculated once for each route based on the initial position and speed. Therefore, the information is gathered, passed and processed following the issuance of the first REPP packet from the destination node as it finalizes the route and this is done with a new packet called a Route Measurement packet (RMEA) as shown in Figure 6.3.
The packet adapts the format of RREP packet as it also navigates the same reverse route. The additional fields are the position, speed, and availability.

**B. Process II: Calculating Link and Route Probabilities**

At each intermediate node, the receiving node calculates the Good Link Availability (GLA) of the two nodes in that link and subsequently updates the accumulated availability value as a product until it reaches the SN with the total product for availability of each link in the route in which at SN the product value is the Good Routing Availability (GRA). This process is shown in Figure 6.4.

![Figure 6.3 RMEA packet format](image)

![Figure 6.4 RMEA packet process from destination to source node](image)
C. Process III: Assigning Pre-emptive Actions

As the RREPs traverse their way back to the source node, the values of initial positions and speeds are collected and received by the source node for the computation of when a link is predicted to break and thus affecting the route. Instead of the standard process of RERR being issued after a link breaks, a pre-emptive action is taken before a link is predicted to break.

In the proposed system, the process is targeted for a case where there is a tendency of a link within a route to break while in transmission due to the movement of nodes, and also there are a few alternative routes available for a pre-emptive re-routing. The general process is shown in Figure 6.5 which also differentiates the new and the standard process, and also the possible integration of the new to the existing process.

Figure 6.5 A comparison of the standard and new process
Listed below is the summary of the algorithm for the proposed method:

*SN sends RREQ and each Intermediate node (IN) receives incoming RREQ packet*

If this DSN > DSN of DN it indicates new Route Discovery (RD)

then add reverse route in routing table and forward packet to neighbouring nodes

else

indicates that a route has already been selected and the node is not part of the active packet

raise flag as inactive but still forward the packet to search for alternative route one time only

set Reserve=1

If the RREQ reaches the DN

DN send RREP and RMEA packet to SN

Normal propagation and process for RREP

RMEA follows RREP path and updates on each IN with flag reserve = 1

RMEA adapts RREP format with additional fields of position, speed and availability

Calculates each link availability (GLA) and the product of each link in the path (GRA)

\[ GLA = 1 - \frac{D(t)}{TxRange} \]

\[ D \text{ is distance at the time link breaks (out of transmission range)} \]

Based on Pythagorean Theorem

\[ TxRange \text{ is assumed to be the same for all nodes} \]

\[ GRA = \sum GLA_{node1-n} \]

If the active link is very unlikely to drop then drop RMEA

The next time for a link to break (NTLB) is predicted based on calculation

If the run time is nearing NTLB then initiates re-routing procedure based on route in RMEA

Choose the highest GRA as the new route

Finishes transmission

Restart procedure
6.6 Discussions and Conclusions

With a bigger picture in hand, for example in a highway situation, many link breakages can occur in a route. The approximate seconds lost for each link break in recurring events can be avoided, reducing the risk of dropped packets occurring. These are the two elements that can improve the network performance in VANETs.

By using the availability of each link (GLA) and combining by multiplication (GRA) provides a useful performance metric. In this chapter the concept of GLA and GRA has been discussed and evaluated in a basic simulation. The other contribution is the inclusion of a suggested method of implementing this metric for the improvement of the AODV protocol.
CHAPTER 7:
CONCLUSIONS AND FURTHER WORK

The aim of the work in this thesis is to propose a solution for the improvement of the VANET performance. The focus was on the routing protocol, and the performance metrics for this purpose have been explored and indentified. Simulation works using NS2 and calculations have been used to validate the proposed theory.

7.1 Contribution of this thesis

7.1.1 Identified link and route availability as the performance metrics

The work provides an alternative approach of estimating the availability of a link using the position vectors of the nodes and their velocities, and also the transmission range. The actual link lifetime between re-routes is clearly one of the main things that has been improved. Basically it should mean that a more suitable route is selected, and that route should last longer. A longer-lasting link means fewer failures from breaks during reroutes and hence better overall performance for the network. Since the routing is based on availability it ought to be vastly better for the highway scenario as it is likely to select vehicles which are travelling at more or less the same speed and going in the same direction.

Another point is that routes selected by their availability were superior to those based only on hop count. Unlike a number-of-hops based scheme which could select almost anything, the Good Route Availability (GRA) identifies the most probable alternative route by using a practical statistical approach of calculating the product of each Good Link Availability (GLA).
7.1.2 Provide framework for the implementation

Some may argue that enhancing the AODV routing protocol with the introduction of location prediction and the calculations of GLA and GRA have somehow turned the original reactive characteristics of AODV to a slightly hybrid protocol. Nevertheless, AODV is still highly applicable since the default functions of the routing protocol will still be applied if the situation does not require any selection of alternative route such as in the pre-emptive action. The proposed solution can be adapted to work in the original routing protocol while initiated when required with minimal changes.

A new Route Measurement Packet (RMEA) is introduced to support the implementation of new features mentioned earlier. The packet is adequate in size to provide the relevant information and trigger tasks but at the same time is not expected to increase overall overheads during the communication transactions.

7.1.3 Develop a mobility model using Cellular Automata (CA) for a highway scenario

The Mobility model is an important component in any VANET simulation because it must as realistic as possible in order to produce more accurate results. Cellular Automata are chosen because the cell can represent the node and its movement well, and it can be easily applied to NS2. In this thesis, the approach is to create a 6-lane highway with each 3 moving different ways as in a highway. The positions and speed are randomly generated and the simulation assumes flat and straight nodes when it is being run.

7.1.4 Approach in comparing MANET and VANET

The initial work starts with the exploration and analysis of how routing protocols would work and these two networks. There are many ways of doing this and this thesis chooses to have a wide variety of scenarios and parameters. In MANETs tests are made on the effect of TCP and UDP transport protocol packets, and also the effects from wide range if density and also speed. In VANETs the work involves the effect of transmission range with the metrics looking at the
effect on throughput, overhead and dropped packets. All this has provided an understanding of the behaviours of routing protocols in specific scenarios and help to identify the routing protocol to be used in this research.

7.2 Suggestions for further work

7.2.1 Development work

The core engine of AODV in this simulation software needs to be modified to introduce the RMEA packet function. This would involve a thorough understanding on how the routing table works to support the route discovery and forwarding methods. Following that would be the running of simulations that would compare the differences from the AODV and the method being proposed. The intention is to conclude that the results from the proposed method can show that it can produce a longer actual and not predicted link life time.

7.2.2 Test bed

The simulations have been successful in producing baseline and good estimations of the outcomes. The next stage is to apply this on a test bed where actual physical testing can be done. This way, the results are not influenced by assumptions concerning the parameters and constraints from the software, which are always the case in simulations.

There are two types of test beds that can be implemented:

i. Static is where the fundamental aspects are being tested to ensure working ad hoc network set up like the routing protocol functions and also the physical devices and software.

ii. Mobile is where the work would involve cars equipped with devices and software physically moving around. Mobility models can be tested on this which is the main feature of VANETs.
The assumption of getting location information of nodes in simulations can actually be implemented and tested using devices such as GPS where actual data can be extracted. Applications using different transport protocols like UDP and TCP can be tested as well.

7.2.3 Further effects from density and transmission range

In VANETs the effect from these two elements needs to be studied further. The known general idea is that different transmission ranges may respond differently in the context of performance to different size of density. There is a need to focus on the effects as this can advice the implementation aspects and very relevant in the case of VANETs. The electromagnetic and other effects from the propagation of signal can also be studied here.

7.2.4 Enhancement of Mobility model

In this thesis the model relies heavily on the CA to describe the nodes movement. Although the approach is adequate to show the flow of the movement, it can be expanded by applying other rules such as the driver behaviours, changing of lanes, approaching junctions, etc. This effort will make the model even more realistic for future works on VANETs.

7.3 Closing remarks

The underlying motivation of this work is for the development of performance metrics of a routing protocol in Vehicular networks. Researchers in this field need to be aware that adaptation of link and route availability in the routing protocol is a sound and practical way to improve the performance of VANETs. Simulation is proving to be an invaluable tool as it is not practical to have measurements of high density of nodes especially on a highway. Future implementations of applications on VANETs would benefit from the performance metrics mentioned in this thesis.