On the Performance of Traffic-Aware Reactive Routing in MANETs

A Thesis Submitted

by

Raad S. Al-Qassas

for

The Degree of Doctor of Philosophy

to

The Faculty of Information and Mathematical Sciences
University of Glasgow

Abstract

Research on mobile ad hoc networks (MANETs) has intensified over recent years, motivated by advances in wireless technology and also by the range of potential applications that might be realised with such infrastructure-less networks. Much work has been devoted to developing reactive routing algorithms for MANETs which generally try to find the shortest path from source to destination. However, this approach can lead to some nodes being loaded much more than others in the network. As resources, such as node power and channel bandwidth, are often at a premium in MANETs, it is important to optimise their usage as far as possible. Incorporating traffic aware techniques into routing protocols in order to distribute load among the network nodes would helps to ensure fair utilisation of nodes' resources, and prevent the creation of congested regions in the network.

A number of such traffic aware techniques have been proposed. These can be classified into two main categories, namely *end-to-end* and *on-the-spot*, based on the method of establishing and maintaining routes between source and destination. In the first category, end-to-end information is collected along the path with intermediate nodes participating in building routes by adding information about their current load status. However the decision as to which path to select is taken at one of the endpoints. In the second category, the collected information does not have to be passed to an endpoint to make a path selection decision as intermediate nodes can do this job. Consequently, the decision of selecting a path is made locally, generally by intermediate nodes.

Existing end-to-end traffic aware techniques use some estimation of the traffic load. For instance, in the *traffic density* technique, this estimation is based on the status of the MAC layer interface queue, whereas in the *degree of nodal activity* technique it is based on the number of active flows transiting a node. To date, there has been no performance study that evaluates and

compares the relative performance merits of these approaches and, in the first part of this research, we conduct such a comparative study of the traffic density and nodal activity approaches under a variety of network configurations and traffic conditions. The results reveal that each technique has performance advantages under some working environments. However, when the background traffic increases significantly, the degree of nodal activity technique demonstrates clear superiority over traffic density.

In the second part of this research, we develop and evaluate a new traffic aware technique, referred to here as *load density*, that can overcome the limitations of the existing techniques. In order to make a good estimation of the load, it may not be *sufficient* to capture only the number of active paths as in the degree of nodal activity technique or estimate the number of packets at the interface queue over a short period of time as in the traffic density technique. This is due to the lack of accuracy in measuring the real traffic load experienced by the nodes in the network, since these estimations represent only the current traffic, and as a result it might not be sufficient to represent the load experienced by the node over time which has consumed part of its battery and thus reduced its operational lifetime. The new technique attempts to obtain a more accurate picture of traffic by using a combination of the packet length history at the node and the averaged number of packets waiting at node's interface queue. The rationale behind using packets sizes rather than just the number of packets is that it provides a more precise estimation of the volume of traffic forwarded by a given node. Our performance evaluation shows that the new technique makes better decisions than existing ones in route selection as it preferentially selects less frequently used nodes, which indeed improves *throughput* and *end-to-end delay*, and distributes load more, while maintaining a low routing overhead.

In the final part of this thesis, we conduct a comparative performance study between the *end-to-end* and *on-the-spot* approaches to traffic aware routing. To this end, our new *load density* technique has been adapted to suggest a new "on-the-spot" traffic aware technique. The adaptation is intended to ensure that the comparison between the two approaches is fair and realistic. Our study shows that in most realistic traffic and network scenarios, the end-to-end performs better than the local approach. The analysis also reveals that relying on local decisions might not be always good especially if all the potential paths to a destination pass through nodes with an overload condition in which case an optimal selection of a path may not be feasible. In contrast, there is most often a chance in the end-to-end approach to select the path with lower load.

To my parents, my brothers and my sisters
for their support, encouragement and trust

Acknowledgments

I would like to express my sincere gratitude to my supervisors Dr. M. Ould-Khaoua and Dr. L.M. Mackenzie for their support and guidance through the course of this work. Their constructive suggestions and insightful comments have been very valuable, enabled me to progress in this research work.

I extend my thanks to Dr. A. Al-Ayyoub for his help and advice at the early stages of my PhD program. I would like to thank my officemates, S. A. Wahab and S. Papanastasiou, for their valuable comments and advice. My deepest thanks are also due to my colleagues: J. Abdulai, M. Al-Awadi, N. Alzeidi, M. Aminu, S. Bani Mohammad, and S. Manaseer.

I also thank the Overseas Research Students Awards Scheme (ORSAS), Universities UK, for their scholarship. My thanks are also due to all the staff of the Department of Computing Science, University of Glasgow, for their kind and friendly support.

Finally, I must thank my parents, my brothers and sisters for their unlimited trust and great support during the course of this research. I owe special gratitude to my father and my mother for their never-ending patience, understanding and moral support. My family's continuous encouragement and limitless support have made accomplishing this work much easier.

Contents

List of figures

List of tables

Acronyms

ACK	Acknowledgement
AODV	Ad hoc On-demand Distance Vector
AOMDV	Ad hoc On-demand Multi-path Distance Vector
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CBR	Constant Bit Rate
CTS	Clear To Send
DARPA	Defense Advanced Research Projects Agency
DCF	Distributed Coordination Function
DLAR	Dynamic Load Aware Routing
DSDV	Destination-Sequenced Distance Vector
DSR	Dynamic Source routing
HMP	Hotspot Mitigation Protocol
ID	Identification
IEEE	Institute of Electrical and Electronics Engineers
LARA	Load Aware Routing in Ad hoc
LBAR	Load-Balanced Ad hoc Routing
MAC	Medium Access Control
MANET	Mobile Ad hoc Network
NSF	National Science Foundation
OLSR	Optimized Link State Routing
OSI	Open Systems Interconnection
PDA	Personal Data Assistance

PCF Point Coordination Function

RFC Request For Comments

RTS Request To Send

SHARP Sharp Hybrid Adaptive Routing Protocol

Wi-Fi Wireless-Fidelity

ZRP Zone Routing Protocol

Chapter 1

Introduction

In recent years, mobile wireless networks have witnessed a tremendous rise in popularity [4, 19, 29, 30, 65, 78, 83, 89]. This has been chiefly motivated by the technological advances in wireless communication technology and its wide range practical applications [65, 83]. Mobile devices, such as laptops, PDAs and mobile phones, have become lightweight and portable enough to be conveniently carried by a user. Also, the advances in battery technology have enabled such devices to be used for increasingly long periods of time away from electrical power sources.

Wireless networks have a number of advantages over their traditional wired counterparts [65]. Wireless networks allow anywhere/anytime connectivity. They can be deployed in areas without pre-existing wired infrastructure. In addition, installing wireless infrastructure is much cheaper than wired infrastructure; making wireless networks an attractive option especially in less-developed world regions. Furthermore, wireless systems allow instantaneous network setup. For instance, conference attendees can turn on their laptops and immediately connect to the Internet or even network and share files with the other attendees.

Mobile Ad hoc Networks (*MANETs*) have been a hot research area in the last few years [26, 35, 36, 57]. This has been motivated by the attractive potential applications of MANETs and the challenging issues that they pose. A MANET is a collection of mobile wireless nodes (e.g. laptops, PDAs, etc.) which not only can act as hosts but also can form a temporary and arbitrary network without the need of any infrastructure. A major challenge in MANETs is the design of an efficient routing protocol that can accommodate their dynamic nature and frequent topology changes; the topology can change unpredictably, so the routing protocol should be able to adapt automatically. Furthermore, it has also to cope with the inherent constraints from which these networks suffer, notably limited bandwidth and power [5, 12, 16, 19, 52].

There has been a lot research effort devoted to developing routing algorithms for MANETs over the past few years [20, 22, 34, 43, 46, 58, 62, 63, 67, 68, 71, 81]. Most of these algorithms have considered finding the shortest path, as their prime target, when building a route from source to destination. However, this could lead to some nodes in the network being more overloaded with traffic than others [28]. In MANETs, resources, such as node power and channel bandwidth are often at a premium, and therefore it is important to optimise their use as much as possible. Consequently, developing a traffic aware technique that can efficiently distribute traffic among network nodes would be highly desirable so as to make good utilisation of nodes' resources. To this end, a number of traffic aware techniques [28, 37, 48, 64, 75, 84] have recently been suggested for MANETs to deal with this and they can be classified into two categories: *end-to-end* and *on-the-spot*, depending on the way they establish and maintain routes between source and destination. The two categories will be discussed in detail in the subsequent sections below.

The remainder of this chapter is organised as follows. Section 1.1 provides a description of MANETs, their characteristics, constraints and applications. Section 1.2 provides an overview of the existing routing protocols suggested for MANETs. Section 1.3 presents the motivations and contributions of our research study. Section 1.4 presents the thesis statement. Finally, section 1.5 provides the outline of the rest of the thesis.

1.1 Mobile Ad hoc Networks (MANETs)

A MANET is a set of wireless nodes that can communicate directly with one another forming a temporary network without the need of any infrastructure or centralized administration. The communication in ad hoc networks is peer-to-peer as nodes communicate directly with one another. However due to the limited propagation range of each mobile node wireless transmissions, it may be necessary for a node to enlist the aid of others in forwarding a packet to its destination, by forming a path to the desired destination using multi-hop wireless links.

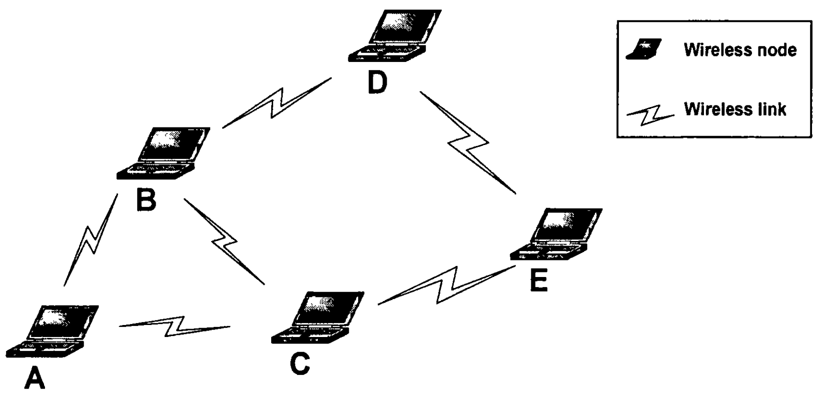


Figure 1.1: Example of a MANET.

Figure 1.1 illustrates an example of a MANET. Nodes that are within transmission range of one another such as nodes A, B and C can establish a direct communication with one another. However, for nodes that are not within transmission range of each other, such as node A and D, will require the aid of other nodes in order to establish a communication with one another, e.g. through node B. The establishment of communication requires the use of a routing protocol to build a path between the two communicating nodes.

Characteristics and constraints:

A MANET is a self-configured network. It could be deployed very quickly since it does not require the establishment of any infrastructure [85]. Thus, it could be used in situations where there is no infrastructure or such infrastructure cannot be used or deployed because of cost, security, or installation difficulty reasons. The self-configuring nature of the network often results in the creation of an arbitrary topology and hence the shape of the topology is unlikely to be determined beforehand.

Typically, nodes in a MANET run on batteries. However, since a battery's lifetime is limited, the power resource is at a premium [5, 12, 16, 19, 52]. As a consequence, it is important to minimise power usage where possible. Although a source of power consumption is from processing to generate data, most of the consumption is due to the transmission and reception of such data [19].

In MANETs, nodes could move freely within the network. However, this often leads to frequent topology changes and may cause frequent path breakages between communicating ends. Keeping track of network connectivity requires periodical collection of connectivity information, which implies an increased communication overhead [3].

In MANETs, each node operates as a router in addition to generating and processing packets; it forwards packets for other nodes that may not be within direct wireless transmission range of each other. A routing protocol enables each node to discover multi-hop paths through the network to any other destination node [10]. The primary goal of a routing protocol is to efficiently establish a route between a pair of nodes so that packets may be delivered in a timely manner.

The presence of wireless connectivity and mobility pose new technical challenges in maintaining communication routes among network nodes [65]. In addition, MANETs are limited in channel bandwidth, battery power and storage capacity [5, 12, 16, 19, 52]. Hence, minimizing the usage of such resources is an important technical challenge. Therefore, any routing protocol developed for MANETs should ensure that route establishment and maintenance incur minimum bandwidth consumption and communication overhead [73].

Applications of MANETs:

Potential applications of MANET can range from simple civil applications to complicated high-risk military applications [57, 65, 83]. We will list below some useful applications for MANETs including conferencing, emergency services and tactical operations; the interested might refer to [65] for further details and examples of other applications.

Conferencing: A MANET is a viable choice when there is a demand for temporary collaborative computing between a group of users as it offers a quick communication platform with minimal configuration [65, 85]. A class of students may need to interact during a lecture, friends or businessmen may wish to share some files [43]. Furthermore, the establishment of a MANET for collaborative mobile computer users could still be useful and cost effective even when there is an available Internet infrastructure. This results from the likely overhead required when utilizing infrastructure links [65].

Emergency Services: For emergency services such as search and rescue and crowd control, it is critical to find ways to enable the operations of a communication network even when infrastructure elements are missing or have been disabled as part of the effects of a natural disaster like an earthquake or hurricane. MANETs could be deployed to overcome network loss and would be a good solution for coordinating rescue efforts. For instance, police squad cars and fire brigade can remain in touch and exchange information more quickly if they can cooperate to form an ad hoc network.

Tactical Operations: A beneficial use of MANET is for military purposes by deploying it in a hostile area. The infrastructure nature of a MANET makes it an important option from military perspective as it does not require infrastructure establishment. Therefore it can be used during the deployment of forces in an unknown and hostile area, for fast establishment of military communication.

1.2 Routing algorithms for MANETs

A crucial component for the efficient operation of a MANET is the routing protocol. The purpose of routing protocol in MANETs is to establish and maintain paths between nodes in order to deliver a packet from source to destination. Due to the frequent changes in MANET topology, connectivity information is often required to be collected periodically in order to get a consistent view of the network, which in turn increases the bandwidth and power consumption resulting from collecting this information. MANETs have limited bandwidth, and therefore need an efficient routing protocol that could establish and maintain routes for both stable and dynamic topologies with minimum bandwidth and power consumption.

1.2.1 General classification of routing algorithms in MANET

A lot of research effort has been to developing routing protocols for MANETs [20, 22, 34, 43, 46, 58, 62, 63, 67, 68, 71, 81]. These protocols can be classified into three main categories based on the routing information update mechanism: *proactive* (or table driven), *reactive* (or source initiated on-demand driven), and *hybrid.*

Proactive routing:

Proactive protocols, such as those described in [20, 22, 58, 67], attempt to maintain consistent and up-to-date routing information (routes) from each node to every other node in the network. Topology updates are propagated throughout the network in order to maintain a consistent view of the network. Keeping routes for all destinations has the advantage that communication with arbitrary destinations experience minimal initial delay. Furthermore, a route could be immediately selected from the route table. However, these protocols have the disadvantage of generating additional control traffic that is needed to continually update stale route entries [65]. DSDV [67] and OLSR [22] are well known example of proactive protocols.

Reactive routing:

Reactive protocols, such as those proposed in [43, 63, 68, 81], establish routes only when network nodes request them. When a source node requires a route to a destination, it initiates a route discovery process within the network. Once a route has been established, some form of route maintenance procedure is used to maintain it, until either the destination become inaccessible or the route is no longer desired. These protocols tend to use less bandwidth for maintaining the route tables at every node. However, latency could drastically increase leading to long delays before a communication can start. This is because a route to the destination has to be acquired first. AODV [68] and DSR [43] are well known example of reactive routing protocols.

Hybrid routing:

A hybrid protocol [34, 41, 61, 71] is a combination of proactive and reactive protocols. It often consists of two routing protocols: one is proactive and the other is reactive. Hybrid protocols divide the network into areas called zones which could be overlapping or non-overlapping depending on the hybrid protocol. The proactive routing protocol operates inside the zone, and is responsible for establishing and maintaining routes to the destinations located within the zone. On the other hand, the reactive protocol is responsible for establishing and maintaining routes to destinations that are located outside the zone. ZRP [34] and SHARP [71] are well known example of hybrid routing protocols.

1.2.2 Traffic aware routing

The main goal in traffic aware routing is to avoid the condition that could occur in traditional routing [43, 68] where some nodes have to carry excessive traffic loads compared to the rest of the network [28, 37]. The excessive load would result in high battery consumption at these nodes and regions in the network being more congested than others. Achieving a good load distribution in MANETs, where nodes in the network have comparable loads, is a challenging task. The issue is complicated not only because a full knowledge of the network status is required in order to make good decisions, but also because it has to deal with balancing the load in the presence of nodes mobility. A number of traffic aware techniques [28, 37, 48, 64, 75] have recently been suggested.

Traffic aware techniques can be classified into two main categories, namely *end-to-end* and *on-the-spot*, based on the way of establishing and maintaining routes between any source and destination. The first category is based on using end-to-end information collected along the path from source to destination. Intermediate nodes participate in building the route by adding some information about their status. However the decision for selecting the path is taken at one of the ends. The second method does not require information to be passed to one of the ends in order to make a path selection decision; it is most likely that intermediate node will do this job. Consequently the decision of selecting a path is made on-the-spot at intermediate nodes. These techniques will be discussed in more detail later in Chapter 2.

Typical end-to-end techniques are represented by two recently proposed methods namely *Degree of nodal activity* and *traffic density*. Degree of nodal activity has been proposed by Hassanein and Zhou [37]. The main idea of this technique is to compute the number of traffic flows transiting a given node and that in its neighbourhood, which represent the cost at the node, and then it calculates the total cost of the path, by accumulating the costs at intermediate nodes. The path with the minimum cost is chosen as the route between the source and destination. On the other hand, traffic density has been proposed by Saigal *et. al.* [75] as a method for selecting the route with the minimum traffic load during the route setup. This method requires that each node maintains a record of the latest estimation of the traffic queue at each of its neighbours. The traffic queue is defined as the average value of the MAC layer interface queue length measured over a period of time. Traffic density of a node is defined as the sum of traffic queue of that node plus the traffic queues of all its neighbours. The path with the minimum accumulated traffic densities is chosen as the route.

On-the-spot techniques [28, 48] use local information and takes local actions in forming a path. A representative of this category is the *hotspot mitigation protocol* proposed in [48]. In this method, hotspots are a group of nodes that are characterised by excessive contention, congestion, and resource exhaustion in the network. Mobile nodes independently monitor local conditions including buffer occupancy, packet loss, and MAC contention and delay conditions, and take local actions on the emergence of a hotspot such as the suppressing new route. Another recent work on this category has been presented in [28]. The proposed method attempts to achieve good load balance while using short paths for minimizing communication latency. The basic idea of the method is to maintain for each node a set of nodes, called *bridges*. Then every time the node chooses the lightest bridge to relay a packet.

Results from the studies of [37, 48, 75] have shown that traffic aware routing improves performance in terms of, e.g., throughput and delay compared to traffic non-aware routing protocols, such as AODV and DSR. In the study of Saigal *et al* [75], a routing protocol with traffic density has been compared against DSR [43]. Their study has shown that traffic density achieves improvement in throughput as well as end-to-end delay over DSR. Moreover, in the study of Hassanein and Zhou [37], where nodal activity has been proposed, a routing protocol with nodal activity has outperformed both AODV and DSR in terms of both end-to-end delay and packet delivery ratio.

1.3 Motivations and contributions

There has been a lot of work on developing reactive routing algorithms for MANETs [43, 63, 68]. Most of these algorithms have considered finding the shortest path from source to destination to build a route. However, this can lead to some nodes being overloaded more than others in the network. Therefore, a traffic aware technique to distribute the load is highly desirable in order to make good and fair utilisation of nodes' scarce resources. In addition it is often useful to prevent the creation of congested areas in the network, which is likely to result in improved network performance.

Indeed, traffic aware routing protocols have recently received considerable attention [28, 37, 75, 84, 96] as they could play a major role in increasing network life time as they target the reservation of nodes' resources and improve system performance. The main objective of our research is to propose, study, and analyse traffic aware routing protocols for MANETs. This thesis focuses primarily on reactive routing mechanisms as a starting point in studying traffic aware techniques.

The existing end-to-end traffic aware techniques use an estimation of the load. This estimation is based either on the status of the MAC layer interface queue as in the case of traffic density [75] or based on the number of active traffic flows as it is in degree of nodal activity [37]. The performance merits of the existing end-to-end traffic aware techniques have been analysed and compared against traditional routing algorithms, such as AODV and DSR [37, 75]. There has also been a performance comparison among the existing on-the-spot techniques in the study of [48]. However, there has not been so far a similar study that evaluates and compares the relative performance merits of two approaches of end-to-end techniques.

Existing end-to-end traffic aware techniques, e.g. [37, 75], suffer from a number of serious limitation in their judgment of the path load. In order to make a good judgment about the load of given path, it may not be *sufficient* to only capture the number of active paths as is in the case of the nodal activity technique [37] or estimate the number of packets at the interface queue over a period of time as in the traffic density technique [75]. The number of active paths could be useful if nodes keep the correct estimation of number of active paths. However mobility might affect the accuracy of number of active paths at the intermediate nodes of a given path leading to inaccurate cost of that path, unless these changes were detected on time. Furthermore, since the number of flows passing through an intermediate node changes over time, even an accurate estimation might not be sufficient to represent the load experienced by that node over time, but it could be useful to represent the current traffic load. The number of packets at the interface queue is useful for capturing contention at the MAC layer. However, this is might not be sufficient to represent the load experienced by the node over time which has consumed part of its battery and reduced its operational lifetime. One of the most limited resources of a mobile node is its power, and since most of the power consumption of a node is due to sending and receiving packets [87], reducing the amount of packets transferred by a node helps to reduce its power consumption, leading to an increase in its lifetime. A more detailed discussion on the limitations of these techniques will undertaken in Chapter 4 (please see Section 4.2).

On-the-spot techniques use local information that capture a small region of the network, and do not require the collection of information about the total cost of path between the end nodes; source and destination. In such techniques nodes actively monitor load conditions, and could interactively detect overload conditions in that region of the network and act upon it. This might distribute the load among neighbours. However it might not provide a good solution for distributing traffic as it is an inherently end-to-end problem [48]. On the other hand, end-to-end techniques require knowing the information from one end to the other to get better knowledge of the network in order to make the decision of selecting a path. This could make better judgment for distributing the traffic. However in these techniques, nodes avoid interactively distributing the load, and tend to wait until route maintenance is initiated to do so. Despite the fact that there has been some research on traffic aware routing [28, 37, 48, 75], there has not been so far a research study that compares the relative performance merits of end-to-end and their on-the-spot counterparts.

Motivated by the above observations, the first goal of this research is to carry out an extensive performance analysis by means of simulations of the important techniques suggested for end-to-end traffic ware routing, notably, traffic density and degree of nodal activity proposed in [75] and [37], respectively. Such a study is the first to be reported in the literature and will help to identify the strengths and shortcomings of the proposed techniques, and their domain of applicability under various network operating conditions.

The second contribution of this research is to develop and evaluate a new traffic aware metrics, based on end-to-end technique, referred to here as *load density*, that can overcome the limitations of the existing techniques. To achieve this, the new proposed technique takes more accurate information compared to the existing ones about the traffic experienced by the network nodes. This is computed by using the history of lengths of packets passed over nodes and the number of packets waiting at the nodes' interface queue. The use of packets sizes in the calculations, rather than just using the number of packets, provides a more precise estimation of the amount of traffic forwarded by nodes, and it covers any variation in packets' sizes while the number of packets at the interface queue provides an indication of the degree of contention over the wireless communication medium. One of the main features of the suggested *load density* technique is that the amount of exchanges of load information among neighbouring nodes is lower compared to the other similar methods, which results in reduced communication overhead.

As a third contribution, the last part of this research, presents a comparative performance study of the two main approaches of traffic aware routing, namely end-to-end and on-the-spot. To the best of our knowledge, this is the first work that carries out such a performance comparison. To this end, we have adapted our new traffic aware technique to suggest a new "on-the-spot" traffic aware technique. The main reason for doing this adaptation is to ensure that the comparison between the two approaches is fair and realistic. The study will reveal the main performance characteristics of the two approaches under various traffic and network conditions.

1.4 Thesis statement

The focal focus of this thesis is to study and develop traffic aware techniques to avoid the problem where some nodes take part in the routing activity more actively than others, which might exhaust more quickly their limited resources of channel bandwidth and battery lifetime. The techniques herein are embedded in reactive routing protocols that operate at the network layer. Although the implementation is independent from the implementation of the underlying MAC layer, the techniques use some feedback from the MAC layer regarding the status of the interface queue.

In this research, I assert that:

T1: The existing end-to-end traffic aware techniques use an estimation of the load. This estimation is based either on the MAC layer interface queue status as in the case of traffic density [75] or based on the number of active flows as it is in degree of nodal activity [37]. However, to the best of our knowledge, there has not been in the literature a study that compares these two approaches. Therefore, I conduct an extensive performance evaluation of the two representatives of end-to-end techniques namely nodal activity and traffic density.

T2: After analysing the behaviour of the existing techniques, in T1, a limitation has been identified. This limitation arises from the lack of accuracy of measuring the real traffic load experienced by nodes in the network. Therefore I develop a new traffic aware technique that can overcome the limitation of the existing methods by using load history, which captures the amount of data passed over each node. I also conduct extensive evaluation of the new technique and compare its performance against the existing techniques.

T3: Existing traffic aware techniques has been classified into two categories: end-to-end and on-the-spot. However, to the best of our knowledge, there has not been a study that compares the two categories. In an effort towards filling this gap, in this research, I investigate the relative performance merits of the two categories.

1.5 Outline of the thesis

The rest of this thesis is organised as follows. Chapter 2 introduces traffic aware techniques, both end-to-end and on-the-spot, and describes related routing algorithms. The chapter also justifies the selection of simulation as a tool of study, and outlines the list of assumptions used in this research.

Chapter 3 conducts an extensive performance comparison between two existing end-to-end traffic aware techniques by means ns-2 simulations.

Chapter 4 proposes a new end-to-end traffic aware technique. The chapter describes the operation of the suggested technique, discusses its main features, and then evaluates its performance considering various mobility and traffic conditions, and compare it against that of existing end-to-end traffic aware techniques.

Chapter 5 carries out a comparative analysis between the two main approaches of traffic aware routing, end-to-end and on-the-spot. To do so, the chapter presents first a new on-the-spot traffic aware technique adapted from our new traffic aware technique.

Finally, chapter 6 summarises the main results of this research, and then outlines some possible directions for further work in the future.

Chapter 2

Preliminaries and related work

2.1 Introduction

A key challenge that has accompanied the design of MANET is the design of an efficient routing protocol that can deal with its inherent constraints, such as mobility. Mobility and thus frequent topology changes have proven to be one of the obstacles to applying the wide range of routing protocols that have been developed for wired networks [27, 49, 56] to MANETs without resorting to modifications. As consequence, various routing algorithms [20, 22, 34, 43, 46, 58, 62, 63, 67, 68, 71, 81] have been proposed to tackle the mobility issue. Any routing protocol should also optimise as much as possible the usage of the scarce resources in MANETs, including channel bandwidth and node power. This is because the optimisation of channel bandwidth, for instance, can lead to performance improvement in terms of delivered data throughput and end-to-end delay [37, 75]. Optimising the usage of battery could extend the life time of the network. Consequently, a traffic aware technique that aims to distribute the load among network nodes during the routing of the packets would be very highly desirable in order to make good utilisation of the nodes' resources. Indeed, research efforts on developing traffic aware routing protocols have recently intensified over the past a few years [28, 37, 48, 64, 75, 84, 95] as they can extend network life time and improve performance.

It is worth noting that traffic aware routing protocols [28, 37, 48, 64, 75, 84, 95] are different from energy aware routing protocols [17, 18, 24, 51, 82, 88, 90]. While both types of protocols attempt to increase the network's life time, the main difference lies on the way they achieve this objective. Energy aware routing algorithms aim at maximizing the life time of the network by minimizing the per packet power consumed by optimising the transmission power as in [17, 24, 51] or by routing packets through energy-rich nodes based on their residual power and a certain threshold as in [82, 88]. However, such an approach does not necessarily provide an end-to-end route with the lowest energy [94]. Other solutions target minimising power during inactive communication through sleep/down mode as in [18, 90]. In contrast, traffic aware routing focuses on optimising the usage of a given node's resources; both channel bandwidth and battery, for forwarding packets inside the network. This class of routing protocols attempt to distribute the task of routing among as many network nodes as possible without, of course, scanting factors that affects network performance, e.g. without increasing considerably the number of hop counts of a typical packet path.

The main objective of this chapter is to describe well-known traffic aware techniques that have been proposed in the literature [37, 48, 75] as well as describe the operation of traditional routing algorithms that have been proposed in [43, 55, 68]. Such background information is necessary for understanding the subsequent chapters. The remainder of this chapter is organised as follows. Section 2.2 Presents MANETs enabling technologies. Section 2.3 describes traditional routing algorithms. Section 2.4 describes traffic aware techniques. Section 2.5 illustrates the routing framework used to implement end-to-end traffic aware techniques in this research work as it will be useful for the following chapters. Section 2.6 describes the used simulator. Section 2.7 justifies the selection of simulation as a tool of study while Section 2.8 outlines the list of assumptions used in this research. Finally, Section 2.9 summarises this chapter.

2.2 Enabling technologies

The enabling technology for MANET is emerged from the available technology for the wireless networks such as Bluetooth [57] and IEEE 802.11 [39]. The Bluetooth technology could be used to build a personal area wireless network that connects devices placed around a person. Large scale networks could be built by exploiting the IEEE 802.11. These technologies could allow building a MANET for different scales, although there are many challenges remain to be solved [21].

The IEEE 802.11 standard defines two operational modes for wireless networks: the *infrastructure* mode and the *ad hoc* mode. The infrastructure-based network consists of some fixed nodes called *access points* and mobile wireless nodes. The access points can interact with the other wireless nodes as well as with the existing wired network. The rest of the wireless nodes communicate through these access points. The infrastructure-based mode is commonly used to construct Wi-Fi (Wireless-Fidelity) hotspot which provides wireless access to the Internet. In the ad hoc mode wireless nodes can communicate directly with each other without using access points. Access to the Internet could be granted through nodes that are connected to the service where these nodes act as gateways for the other nodes in the network.

The IEEE 802.11 standard which is also known as IEEE 802.11 legacy operates at data rate up to 2Mbps. However, since the adoption of this standard several task groups have been created to extend the IEEE 802.11 standard, of these extensions: 802.11a, 802.11b, and 802.11g. The 802.11a supports a rate up to 54Mbps, and operates in the 5 GHz band. The 802.11b supports a rate up to 11Mbps and operates in the 2.4GHz band same as the original standard. The 802.11g operates also in the 2.4GHz band and supports a rate up to 54Mbps. A more detailed description of these extensions can be found in [57].

The IEEE 802.11 MAC layer provides two access methods to the wireless media: the Distributed Coordination Function (DCF) and the Point Coordination Function (PCF). The DCF is the fundamental MAC access method and it works in a distributed way. The PCF is an optional access method built on top of the DCF. It relies on a central node, which is typically the access point, and hence it is not suitable for MANET.

The DCF is based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) scheme. The CSMA is a contention-based algorithm ensures that each station first sense the medium if it is idle before transmitting in order to avoid having stations transmitting at the same time, which may results in collisions and retransmissions. The collision avoidance aspect of the protocol is related to the use of acknowledgment sent from the receiver to the sender to ensure that packet has been received correctly. In addition to the *physical carrier sensing* the DCF has a *virtual carrier sensing* that exchanges a Request-to-Send(RTS)/Clear-to-Send(CTS) control packets between neighbouring nodes preceding the transmission of each unicast packet to reduce the probability of collisions due to hidden terminals problem [39].

2.3 Related routing protocols

Before presenting traffic aware techniques, this section describes some traditional routing protocols for MANETs, namely the *Ad hoc On-demand Distance Vector* (AODV) [68], *Ad hoc On-demand Multipath Distance Vector* (AOMDV) [55] and *Dynamic Source Routing* (DSR) [43], in order to contrast the operations of the traffic aware routing against that of the traditional routing.

2.3.1 Ad hoc On-demand Distance Vector (AODV)

AODV [66, 68] is an RFC and one of the well-known routing protocols for MANETs. AODV is a reactive protocol that establishes routes in an on-demand basis, and therefore it reduces the number of required flooding operations for route discovery. The operation of AODV protocol consists of two processes: *route discovery* and *route maintenance*

When a source node needs to send data, but does not already have a valid route to the destination, it initiates a route discovery process in order to locate the destination. The source broadcasts a *route request* packet to its neighbours which in turn forward the route request to their neighbours, and so on, until either the destination or an intermediate node with a fresh route to the destination is located. Each node that forwards a route request creates a *reverse route* back to the source node itself. Once the route request reaches the destination or an intermediate node with a fresh route, the destination/intermediate node responds by unicasting a *route reply* packet to the source node. Each node that participates in forwarding the route reply back to the source creates a *forward route* to the destination. Nodes along the path from source to destination are not required to have the knowledge about which nodes are forming the path. Each node stores the next hop rather than the entire path.

Route maintenance is triggered when a node detects that a route is no longer valid, it removes the routing entry and sends a *route error* packet to the neighbours that are actively using the route; informing them that this route is no longer valid. For this purpose, AODV uses an active neighbour list to keep track of the neighbours that are using a particular route. The nodes that receive this packet repeat this procedure. The packet is eventually received by the affected sources that can choose to either stop sending data or request a new route by sending out a new route request.

AODV uses destination *sequence numbers* to ensure that routes are loop-free and contain the most recent route information. Each node maintains its own sequence number. The source node includes in the route request the most recent sequence number which has been obtained for the destination. Intermediate nodes can reply to the route request only if they have a route to the destination whose corresponding destination sequence number is greater than or equal to that contained in the route request.

2.3.2 Ad hoc On-demand Multipath Distance Vector Routing (AOMDV)

AOMDV [55] has been developed for MANETs as a reactive multipath protocol establishes routes in an on-demand basis. The protocol is an extension of AODV, and it has been shown in [55] that it can guarantee loop-freedom and disjointness of the generated paths. AOMDV computes disjoint routes without the use of *source routing* [70]. The operation of AOMDV protocol consists of two processes: *route discovery* and *route maintenance*

Similar to AODV, when a source node wants to send data, but does not already have a valid route to the destination node, it initiates a route discovery process in order to locate the destination. The source broadcasts a *route request* packet to its neighbours which in turn forward the route request to their neighbours, and so on, until either the destination or an intermediate node with a fresh route to the destination is located. Unlike AODV, the duplicate copies that arrive later are not discarded. Some of these copies can be used to form alternate reverse paths. Thus all duplicate copies are examined in AOMDV for potential alternate paths. However reverse paths are formed only using those copies that meet loop-freedom and link disjointness criteria [55].

When an intermediate node receives a route request copy, it checks whether there are one or more valid forward paths to the destination that have not been used in any previous route replies for this route request. If so, the node generates a *route reply* and sends it back to the source along the reverse path. Otherwise, the node re-broadcasts the route request if it has not been previously forwarded. When the destination receives route request copies, it forms reverse paths in the same way as intermediate nodes. However, the destination generates a route reply in response to every route request copy that arrives via a loop-free path.

A route maintenance process is initiated when the last used path to a given destination breaks. AOMDV uses *route error* packet to declare a route failure. The node that detects the link failure generates a route error packet sending it back to the source. The route error packet removes the route along its way to the source. When the source receives the route error packet it removes the path from its table and decide whether to send a new route request or to stop sending data to the destination.

2.3.3 Dynamic Source Routing (DSR)

DSR [43] is a reactive on-demand routing protocol which has been designed for MANETs. DSR uses source routing [70] where each data packet to be routed is carrying in its header the complete ordered list of nodes through which it must use to reach the destination. The operation of DSR protocol consists of two mechanisms: *route discovery* and *route maintenance.*

Route discovery is initiated when a source node wishes to obtain a route to a destination. The operation starts by flooding a *route request* packet in the network starting from the source node. Each node that rebroadcast the route request packet adds its address to source route in the route request packet and then forwards it. The route request is propagated in the network until it is received by the destination node or by an intermediate node that knows a route to the destination. Each node maintains a *route cache* of source routes that it has learned or overheard. The destination responds on receiving the route request packet by sending a *route reply* packet back to the source node, which contains the route traversed by the received route request packet. Also an intermediate node that knows a route to the destination can respond back to the source.

Route maintenance is responsible for detecting changes in the network topology that affects the used routes. When a route is broken, the source node is notified by a *route error* packet. Nodes that receive the route error packet remove that source route from its route cache. The source can then attempt to use any other route to that destination already in its cache or can initiate a new route discovery to find a new route.

2.4 Traffic aware techniques

This section describes traffic aware techniques namely *degree of nodal activity* suggested in the routing protocol "Load-Balanced Ad hoc Routing" (LBAR) [37], *traffic density* suggested in "Load Aware Routing in Ad hoc" (LARA) [75] and Hotspot Mitigation Protocol (HMP) [48]. Both the degree of nodal activity and traffic density belong to the category of end-to-end techniques, where as the Hotspot Mitigation Protocol belongs to the category of on-the-spot techniques (please refer to Section 1.2.2 in Chapter 1 for definition of these two classes of traffic aware techniques).

2.4.1 Degree of nodal activity

The "degree of nodal activity" has been suggested as a technique (or metric) for selecting the route with the least traffic load in "Load-Balanced Ad hoc Routing" (LBAR) [37] which is a DSR-like routing protocol. LBAR is a reactive routing protocol that attempts to find a path, which possesses the least traffic load based on a given cost function. The cost function is calculated using two components: *nodal activity* and *traffic interference*. Nodal activity of a node is defined as the number of active paths passing through that node. An active path is an established path from a source to a destination. Traffic interference is defined as the sum of nodal activity for the node's immediate neighbours. The cost of a route is defined as the sum of nodes' nodal activity plus the activity of their neighbouring nodes. The path with the minimum cost is considered as having minimum traffic and is selected between the source and destination nodes. For the subsequent chapters we refer to the "*degree of nodal activity*" technique as "*nodal activity*" for short.

The traffic interference TI_i of node i is computed using equation 2.1, where A_j^i is the nodal activity of the node j; an immediate neighbour of node i.

$$TI_i = \sum_{\forall j} A_j^i \qquad (2.1)$$

The cost of a route r is calculated using equation 2.2, where A_i is the nodal activity of node i and TI_i is its traffic interference, where i is a node on the route r; excluding the source and destination nodes.

$$C_r = \sum_{i \in r} (A_i + TI_i) \qquad (2.2)$$

Route discovery:

In LBAR, the route discovery process is initiated whenever a source node needs to communicate with another node. The source node broadcasts a *route request* (or setup if we adopt the terminology of [37]) packet to its neighbours. The packet carries the cost seen from the source to the current node. A node that receives a route request packet forwards it to its neighbours after updating the cost based on its nodal activity value and traffic interference value. In order to prevent looping when route request packets are routed, the route request packet contains a list of all node IDs used in establishing the path from source node to the current intermediate node. The destination node collects arriving route request packets within a route-select waiting period, which is a predefined timer for selecting the best-cost path. After the waiting period expires the destinations sends a *route reply* packet to the source node along the selected path. When the source node receives a route reply, it recognises that a path has been established to the destination and then starts data transmission.

Route maintenance:

Route maintenance is triggered whenever a node on the active path moves out of the communication range, the case of which an alternate path must be found. If the source node moves away from the active path, the source has to reinitiate the route discovery procedure to establish a new route to the destination. When either the destination node or some intermediate node moves outside the active path, route maintenance is initiated to re-establish the broken path. In case the next hop becomes unreachable, the node upstream of the broken hop propagates a *route error* packet to the destination node. The destination then picks up an alternative path and then sends a route reply packet to the initiator of the route error packet. If the destination has no alternative path, it propagates a route error packet to the source, which will initiate a new route discovery if need be.

2.4.2 Traffic density

The "*traffic density*" has been proposed as a metric for selecting the route with the minimum traffic load in LARA [75] which is a variation of the DSR protocol. LARA uses *traffic density* to represent the degree of contention at the MAC layer. This metric is used for selecting a route with the minimum traffic load during the route discovery phase. The LARA protocol requires that each node maintain a record of the latest *traffic queue* estimations at each of its neighbours in a table called the neighbourhood table. A *traffic queue* is defined as the average value of the interface queue length measured over a period of time. *Traffic density* of a node is defined as the sum of *traffic queue* of that node plus the *traffic queues* of all its neighbours. The path with the minimum cost is selected to be the route between the source and destination nodes.

The traffic queue q_i of node i is computed using a number of samples, say N, where a sample is taken every sample-period, as written in equation 2.3, where $q_i(\mathrm{k})$ is kth sample of the queue.

$$q_i = \frac{\sum_{k=1}^{N} q_i(k)}{N} \tag{2.3}$$

The traffic density TD_i at node i is computed as written in equation 2.4, where q_j is the queue length of node's i neighbouring node j.

$$TD_i = q_i + \sum_{\forall j, j\, adj\, i} q_j \tag{2.4}$$

The cost of the route r is calculated by adding the traffic density of the nodes that are forming the route as in equation 2.5, excluding traffic density at the source and destination nodes

$$C_r = \sum_{i \in r} TD_i \tag{2.5}$$

Route discovery:

In LARA, the route discovery process is initiated whenever node needs to send out data. In the route discovery process, the source broadcasts a *route request* packet that contains a sequence number, a source ID and a destination ID. A node that receives the route request, broadcasts the request further, after appending its own traffic density to the packet. This process continues until the route request packet reaches the destination. After receiving the first route request, the destination waits for a fixed time-interval for more route request packets to arrive. When the timer expires, the destination node selects the best route from among the candidate routes and sends a *route reply* to the source. When the source node receives the route reply, it can start data transmission. If it does not receive any route reply within a route discovery period, it can restart the route discovery procedure afresh.

Route maintenance:

Route maintenance is triggered whenever a node on the active path moves out of the communication range. In such a case, an alternate path must be found. If a link failure occurs during a data transmission session, the source is informed of the failure via a *route error* packet. On receiving a route error packet, the source initiates a new route request and queues all subsequent packets for that destination until a new route is found.

2.4.3. Hotspot Mitigation Protocol (HMP)

This research focuses mainly on end-to-end traffic techniques. Nonetheless, in order to contrast the operations of the end-to-end techniques with those of their on-the-spot counterparts, this section briefly describes the Hotspot Mitigation Protocol (HMP) [48], which belongs to the category of on-the-spot techniques.

In HMP, nodes monitor local conditions including buffer occupancy, packet loss, and MAC delay conditions. Hotspots are any nodes which are characterized by excessive contention, congestion, and resource exhaustion in the network. The main goal of HMP is to redirect new routes away from hotspots. On the emergence of a hotspot local actions are taken such as suppressing a new route to ensure that routed traffic does not compound congestion problems.

HMP uses a number of parameters to identify hotspots. These parameters are associated with MAC-delays, packet loss, and buffer occupancy. A node is declared as a hotspot if a combination of MAC-delays, packet loss, buffer occupancy, and remaining energy reserves exceed certain predefined system thresholds. The MAC-delay parameter is used by the protocol to detect MAC-delay violations. The MAC-delay is defined as the measured time for the successful transmission of a data packet at the MAC layer. If the measured MAC-delay exceeds the MAC-delay threshold then this is considered as a MAC-delay violation. If the number of these violations exceeds a predefined value then the protocol declares the node as a hotspot.

A hotspot detection interval starts when the first MAC-delay violation is observed and lasts until the node either declares itself a hotspot or a data packet is successfully delivered without a MAC-delay violation. Also a link failure or route error resets all associated counters or parameters used by HMP to detect hotspots.

The buffer occupancy parameter is also used to identify hotspots. If the buffer occupancy exceeds a predefined threshold and there is a MAC-delay violation the node declares itself a hotspot. The threshold is set to a buffer level that is less than the buffer overflow mark. This approach is used because buffer occupancy information alone is insufficient to declare a hotspot unless the buffer overflow mark is exceeded. HMP also uses packet loss as a parameter to identify hotspots. If packet loss violates a predefined threshold the node declares itself a hotspot.

2.5 An algorithmic framework for routing with end-to-end traffic aware techniques

This section describes the routing framework that is used in the implementation of end-to-end traffic aware techniques examined by this research work. The description will be useful for the following chapters.

Like in existing reactive routing protocols [43, 68] whenever a route to a destination node is required, a *route discovery* process is initiated in order to locate the destination. To do so, a *route request* packet is broadcasted to the immediate neighbours which, in turn, forward the route request to their neighbours. Each node that receives a copy of the route request adds the cost parameters information to the route request packet and then forwards it to its neighbours. The forwarding process continues until the route request packet reach the desired destination. Duplicate copies of the route request that satisfy the loop-freedom criterion used in AOMDV [55] are also forwarded. Figure 2.1(a) illustrates an example of route request journey in the

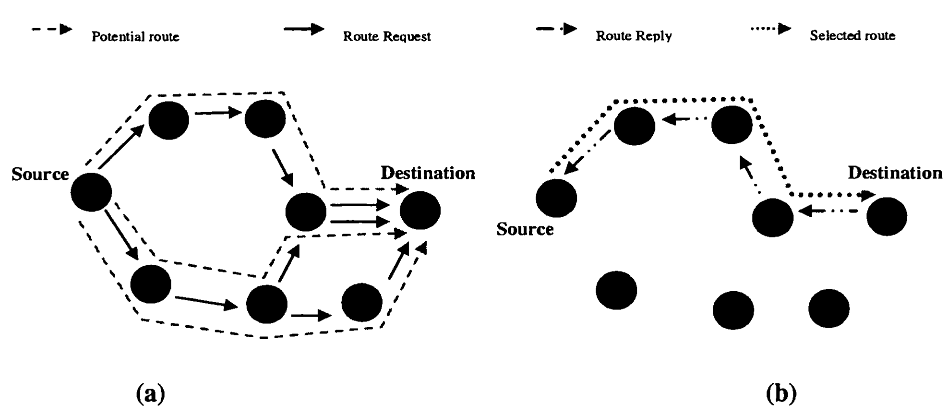


Figure 2.1: Illustration of route discovery (a) route request journey and potential routes (b) route reply path and the selected route.

network and potential routes. The figure shows three potential routes between source and destination.

On the reception of the first copy of route request, the destination waits for a predetermined *selection time* waiting for other copies of the route request that will form alternate routes in order to select one of them based on the selection criteria set by the traffic aware technique. The cost of a route is calculated using the information specified by the traffic aware technique in question (e.g., the number of packets at the interface queue in the case of traffic density) that is carried in the route request packet. When the timer expires the destination selects one of the routes according to the traffic aware technique selection method. Once a route with the minimum cost is selected the destination sends a *route reply* back to the source node. The source node then uses this route to send data to that destination. Figure 2.1(b) shows an example of route reply journey in the network, and accordingly the selected route which will be in the reverse direction. Figure 2.2 illustrates the algorithmic framework of the route discovery process in end-to-end traffic aware routing.

For nodal activity and traffic density, each network node periodically exchanges its traffic conditions with the immediate neighbouring nodes at fixed intervals, typically every 1 second [37, 75]. However for our proposed technique, as will be discussed in later chapters, these exchanges are optional. When a node detects a link failure in a given route, the route is deleted and a *route error* packet is sent to inform the source and the other nodes which forward the packet that this route is no longer valid; this can help to avoid the problem of stale routes. Nodes that receive the route error packet delete the affected routes. Figure 2.3 illustrates the algorithm framework of the *route maintenance* process in end-to-end traffic aware routing. The source upon receiving the route error packet can initiate a new route discovery to find a new route to replace the broken one by broadcasting a new route request. A new route discovery is initiated rather than using any stored routes instead of the broken route in order to be aware of the current mobility condition and its effect on the established routes.

```
End-to-end traffic-aware routing algortihm framework (Route discovery) {
// executed by every node in the nerwork
  If a route to a destination is required
      Broadcast a route request packet;
  When a route request copy is received by a node other than destination {
      If first route request copy {
              Add cost parameters information to the one in the route request packet
              Forward the route request to the neighbours;
              }
      Else If duplicate route request copy&& meets with the loop-free and link-disjointness criteria in AOMDV [55] {
              Add cost parameters information to the one in the route request packet
              Forward the route request to the neighbours;
              }
      Else discrard route request;
  }
  At the destination:
      When a route request is received by the destination
              Wait for other potentional request copies within selction interval from the receiption of the first copy;
              // every route request copy forms an alternative route
              // the selection interval is set by the traffic aware technique
      If selction interval is finsihed {
              Select on of the routes according to the selection criteria set by the teachnique in use;
              Send a route reply to the source node which initiated the route request;
      }
} End (Route discovery)
```

Figure 2.2: Outline of the algorithmic framework of the route discovery process in end-to-end traffic-aware routing.

```
End-to-end traffic-aware routing algortihm framework (Route maintenace) {
// executed by every node in the nerwork
    If link breakage is detected {
        Delete the affected routes;
        Notify affected node(s) by sending route error packet
        // each node holds a prucursors list contains the nodes that use a specific route
    }
    If route error packet is received {
        Delete the affected routes;
        Notify affected node(s) by sending route error packet
    }
} End (Route maintenance)
```

Figure 2.3: Outline of the algorithm framework of the route maintenance process in end-to-end traffic-aware routing.

2.6 The NS-2 simulator

This section introduces briefly the well-known network simulator *ns-2* [1]. Ns-2 is a discrete event network simulator. It is organized according to the OSI model [38] and it supports wired and wireless networks. Ns-2 has been heavily used in the field of MANETs [8, 10, 14, 23, 25, 44, 48, 59, 69, 92]. This is due to the fact that it is open source and include detailed simulation of the important operations of such networks [10, 38]. The development efforts of the simulator have been supported by DARPA and NSF [2, 38]. It worth noting that the simulator has been extensively validated in [10, 42] and verified in [10, 50]. The following paragraphs provide some details about the physical and link layers.

The ns-2 simulator includes a radio propagation model that supports propagation delay, capture effects, and carrier sense [72]. The radio model uses characteristics similar to the commercial Lucent's WaveLAN with nominal bit rate of 2Mb/s and a nominal range of 250 meters with omni-directional antenna. The propagation model has been devised by experts in modulation and combines both a free space propagation model and a two-ray ground reflection model [72].

Ns-2 implements the standard IEEE 802.11 Distributed Coordination Function (DCF) MAC protocol [39]. In this standard the transmission of each unicast packet is preceded by a Request-to-Send(RTS)/Clear-to-Send(CTS) control packets exchange between neighbouring nodes to reduce the probability of collisions due to hidden terminals problem [39]. The RTS and CTS are short control packets containing the length of the data packet that will follow. The exchange of RTS/CTS reserves the wireless channel for transmission of the unicast packet. Once the source node receives the CTS, it transmits its data packet. In this way, neighbouring nodes hearing these exchanges do not transmit during data transmission. Each correctly received unicast packet should be followed by an Acknowledgment (ACK) to the sender; otherwise the sender retransmits the packet a limited number of times (e.g. 7 times) until this ACK is received [39]. Broadcast packets, on the other hand, are not preceded by an RTS/CTS exchange nor acknowledged by their recipients, but they are sent only when the medium is clear.

2.7 Justification of the method of study

In this work, extensive simulations are conducted to explore performance-related issues of traffic aware routing in MANETs. This section briefly discusses the choice of simulation as the proper method of study for the purpose of this dissertation, justifies the adoption of ns-2 as the preferred simulator, and further provides information on the techniques used to reduce the opportunity of simulation errors.

After some consideration, simulation was chosen as the method of study in this dissertation. Particularly, when this research work was undertaken, analytical models with respect to multi-hop MANETs were considerably coarse in nature [6] which made them unsuitable to aid the study of routing protocols with a reasonable degree of accuracy; it should be noted, however, that understanding of multi-hop wireless communications has improved in recent times [79]. In addition, since the range of this study on traffic aware routing in MANETs involves numerous mobiles nodes, even a moderate deployment of nodes as an experimental test-bed could involve substantial and too expensive cost. As such, simulation was chosen as it provides a reasonable trade-off between the accuracy of observation involved in a test bed implementation and the insight and completeness of understanding provided by analytical modelling.

In order to conduct simulations the popular ns-2 simulator has been used extensively in this work. Ns-2 has been chosen primarily because it is a proven simulation tool utilised in many previous MANET studies [8, 10, 14, 15, 23, 25, 44, 48, 69] as well as in other network studies [9]. While developing modifications to the simulator, special care was taken in order to ensure that the algorithms implemented would function as designed and that the simulator would not exhibit unwanted side-effects; this was accomplished through meticulous use of the validation suite provided with ns-2 as well as careful piecemeal testing of implemented features.

2.8 Assumptions

In the following chapters, extensive simulation results will be presented to evaluate the performance of our traffic aware techniques in MANETs. The subsequent assumptions have been used during this research and have also been extensively used in other similar existing studies, e.g. [10, 15, 23, 31, 37, 59, 69, 75].

- Mobiles nodes have sufficient power supply to function throughout the simulation time. At no time does a mobile node run out of power or malfunction because of lack of power.
- The number of nodes in a given topology remains constant throughout the simulation time.
- All nodes are identical and are equipped with IEEE 802.11 transceivers.
- Nodes can move at any time without notice.
- All nodes are willing to participate in forwarding packets to other nodes in the network.
- The routing protocol operates at the network layer. As in all previous studies on traffic aware routing [31, 48, 75], there is exchange of information (e.g. number of packets at the interface queue) with the MAC layer for the operation of the traffic aware routing protocols.
- Dealing with security attacks (e.g. malicious attacks or denial of services [57]) is not part of the traffic aware routing protocols. Such attacks are assumed to be handled with by a suitable security protocol [77, 93].

2.9 Summary

A lot of research efforts have been devoted to developing efficient routing algorithms [20, 22, 34, 43, 46, 58, 62, 63, 67, 68, 71, 81] for MANETs. Although most of the suggested protocols have shown a degree of success in dealing with node mobility, they have not fully tackled the critical issue of the limited resources in MANETs when making selecting new routes. However, as the resources in MANET have great impact on the operation of the network and its operational lifetime, there have recently emerged noticeable interests in designing routing solutions which consider resources availability in a typical MANET. Such solutions include routing that incorporate traffic aware techniques [28, 37, 48, 64, 75], that aim to distribute traffic load over the network nodes in order to improve the performance of the routing algorithm and to minimise as much as possible the usage of node's important resources, such as channel bandwidth and battery power.

This chapter has provided a background on traditional routing algorithms, such as AODV, AOMDV, and DSR. Furthermore, it has provided a description of two representative end-to-end traffic aware techniques, namely degree of nodal activity proposed in [37] and traffic density proposed in [75]. It has also briefly described the hotspot mitigation protocol, a representative of the on-the-spot techniques.

This chapter has also described the routing framework that is used in the implementation of end-to-end traffic aware techniques considered in this research work. It has also briefly described the ns-2 simulator that is used to conduct the performance evaluation of traffic aware techniques and briefly discussed the choice of simulation as a tool of study in this work. Finally, it has outlined the list of assumptions which apply throughout the dissertation.

In the subsequent chapters, we will conduct a performance comparison between the end-to-end traffic aware techniques, and then suggest a new end-to-end traffic technique that can overcome the limitations of the existing approaches. Finally, we will conduct a performance comparison between the two main classes of traffic aware techniques; end-to-end and on-the-spot.

Chapter 3

Performance analysis of end-to-end traffic aware techniques

3.1 Introduction

Traffic aware routing has been receiving considerable attention recently [28, 37, 64, 75, 84, 96] from the research community as it could play a major role in extending network lifetime (or operation), which is described as the time taken for the network to get partition as a result of nodes run out of power and hence the network is not fully functional, as it targets the reservation of nodes' resources which could lead to performance improvement in terms of important performance metrics including *throughput* and *end-to-end delay*. The research efforts on traffic aware routing have focused on various types of routing protocols; notably reactive [37, 75], proactive [76] and hybrid [64]. The present study focuses on reactive mechanisms as a starting point in the performance analysis of traffic aware routing protocols.

As has been mentioned in the previous chapters, traffic aware techniques can be classified into two categories: *end-to-end* and *on-the-spot*. The performance merits of the existing end-to-end traffic aware techniques have been analysed and compared against traditional routing algorithms [43, 68]. For instance, traffic aware solutions such as *traffic density* and *nodal activity* proposed in [75] and [37], respectively, which are representative of end-to-end techniques, have shown performance improvement over well-known traditional routing algorithms like AODV [68] and DSR [43]. In the study of Saigal *et al* [75], a routing protocol with traffic density has been compared against DSR [43] and DLAR [31]. Their study has shown that traffic density achieves improvement in throughput as well as end-to-end delay over both DSR and DLAR. Moreover, in the study of Hassanein and Zhou [37], where nodal activity has been proposed, a routing protocol with nodal activity has outperformed both AODV and DSR in terms of both end-to-end delay and packet delivery ratio. There has also been a performance comparison among the existing on-the-spot techniques in the work of [48]. However, to the best of our knowledge, there has not been a similar study that evaluates and compares the relative performance merits of end-to-end techniques.

Motivated by the above, the main goal of this chapter is to conduct an extensive performance analysis by means of ns-2 simulations of the important techniques suggested for end-to-end traffic aware routing, notably, traffic density and nodal activity; the operations of these techniques have been discussed in details in Chapter 2 (please see Section 2.4). Our simulation study is the first comparison study to be reported in the literature and will help to highlight the relative performance merits of the proposed techniques under various network operating conditions. The techniques are assessed using three widely used performance metrics: throughput, end-to-end delay and routing overhead.

The rest of this chapter is organised as follows. Section 3.2 illustrates the simulation model, the system parameters and performance metrics. Section 3.3.1 analyses the effects of network size in terms of number of nodes and the size of network area on the performance of traffic aware techniques while Section 3.3.2 analyses the effects of packet length. Finally, section 3.4 concludes the chapter.

3.2 Simulation model

As in the previous studies of [8, 10, 14, 15, 23, 25, 44, 48, 69], the simulation model consists of the following main components: simulation area, simulation time, number of nodes, mobility model, maximum node speed, number of traffic flows, and traffic rate. The model is represented by two scenario files, which are the topology scenario and traffic scenario. The topology scenario corresponds to how nodes are distributed over the simulation area and their movement during simulation time. The traffic scenario file contains information such as: type of data, number of flows, and traffic rate. In all scenarios nodes are assumed to be equipped with the wireless standard IEEE 802.11 [39]. The physical radio characteristics of each node; such as the antenna gain, transmit power, and receiver sensitivity, are chosen to be similar to the commercial Lucent's WaveLAN [45], with nominal bit rate of 2Mb/s and a nominal range of 250 meters with omni-directional antenna. The propagation model has been devised by experts in modulation and combines both a free space propagation model and a two-ray ground reflection model [72].

In order to properly conduct an extensive performance analysis of traffic aware techniques we have devised a simulation model. The model is intended to avoid the effect of source and destination pairs being within one hop of each other as this will not assist in the assessment of traffic aware techniques since it will automatically select the direct communication link between the source and destination and hence the selection criterion of the traffic aware technique is not assessed. In this model we have selected a group of nodes forming the source and destination pairs are stationary while the rest of the nodes in the network are mobile. Traffic flows do not share the same source nodes, each flow has a source and destination pairs that are different from the one used in other flows. When the source node starts off sending data traffic to the destination, traffic will continue flowing until the end of simulation.

Keeping certain network nodes stationary helps to create scenarios that clearly reveal the performance merits of the different competing traffic aware techniques. In such scenarios, of course, there are no fixed paths between source and destination pairs as intermediate nodes that form the paths are mobile. Nodes in ad hoc network may run out of power or switch themselves off to save energy. However in the simulated scenarios each node has enough battery to stay alive during the simulated time. This is to allow us to study the behaviour of the techniques under the same scenarios in order to allow direct and fair comparisons between the techniques away from the effect of losing nodes, though it would be interesting to discover as a next step of this research. Figure 3.1 illustrates how sources and destinations are placed in our simulation scenarios.

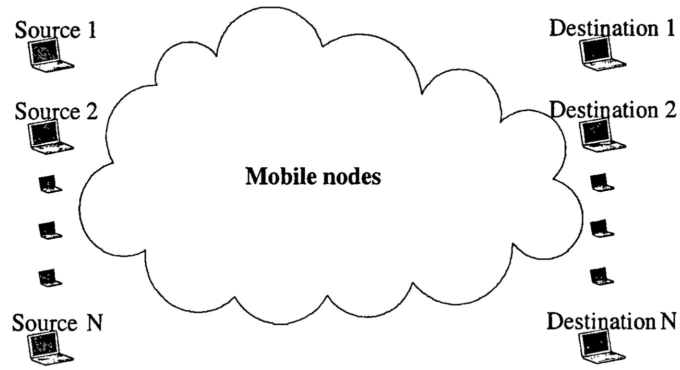


Figure 3.1: Illustration of how sources and destinations are placed in the simulation scenarios.

As in most previous studies [10, 31, 37, 59, 69, 75], we assume that all nodes are equal (e.g. physical radio characteristics of network interface and battery power) and that all transmission errors are recovered. Also, we assume that all nodes within the network are willing to participate in forwarding packets to other nodes in the network. Dealing with security threats such as malicious attacks or denial of services [57] is important for the operation of the network. However we assume that dealing with security attacks is not part of the routing technique used and that such attacks are recovered with the help of a security protocol [77, 93]. We also assume that nodes can move at any time without notice as this is a normal behaviour in MANETs. The routing protocol operates at the network layer, and we assume that there is interaction with the MAC layer to exchange any relevant information about the interface queue (e.g. number of packets waiting at the interface queue) similar to previous studies [31, 48, 75].

We have implemented the traffic aware techniques, traffic density and nodal activity, under the AODV-like routing algorithm, referred to as AOMDV in [55]. AOMDV is a multi-path algorithm that supports loop-free multiple paths. The ns-2 source code for this algorithm is already available [54] and it is easy to modify this source code to simulate the traffic density and degree of nodal activity metrics rather than writing the simulator model from scratch. The framework algorithm used to implement the techniques is described in Chapter 2 (please see Section 2.5).

3.2.1 Simulation parameters

The performance analysis is based on the simulation of 75 and 100 wireless nodes forming a MANET over a flat space of size 900m×1000m, and 1200m×1000m, respectively. The reason for choosing these numbers combination is to avoid the occurrences of frequent partitions in the network. Each simulation runs for a period of 900 seconds. These setting could represent MANET scenarios in real life; e.g. University campus section or Festival location, although the number of nodes could be larger than the one presented in these scenarios and the operational time could be longer but this is to keep the simulation running time manageable. Flows with Constant Bit Rate (CBR) data have been used as it was important to challenge the protocols with identical loads and environmental conditions in order to enable direct and fair comparison between the techniques. The traffic rate is varied between 2, 4 and 8 packets/s representing low, medium and high traffic loads relative to each other, respectively. Results from simulations presented in this chapter have shown that at traffic rate of 8 packets/s significant amount of packets are dropped due to congestion, less than 25% of sent packets have been successfully delivered. For details about the network capacity in MANET the interested might refer to [69]. The number of CBR flows is 5 flows. The packet size has been set to either 512 or 128 bytes, for each traffic scenario. Mobile nodes move according to the widely used random waypoint model [53]. In this mobility model each node remains stationary for a pause time period. When the pause time expires, the node selects a random destination in the simulation space and moves towards it. When the node reaches its destination, it pauses again for the same pause time. This behaviour is repeated throughout the simulation time.

In all the simulated scenarios the pause time has been set to 0 seconds to allow all time mobility. The node maximum speeds varied between 1, 2, 5, 7, 10, 15 and 20m/s; from human slow walk speed to vehicle speeds. For each speed we have made runs for 30 randomly generated topologies. Simulation parameters are illustrated in Table 3.1. It is worth noting that most of the values for these parameters have been adopted in the literature [10, 14, 15, 23, 25, 44, 48, 69]. Furthermore, such values have been selected because they make the computing resources and times to the run most of the simulation scenarios manageable.

3.3 Performance Analysis

This section analyses the performance of *traffic density* and *nodal activity* techniques. The performance analysis has been conducted using the simulation model and parameters described in Section 3.2

The performance of the two techniques is measured by three important performance metrics: *throughput*, *end-to-end delay* and *routing overhead*. The throughput is the amount of data received (measured in bits per second) at the final destination over the simulated time averaged over the number of flows. This measure provides an indication of the efficiency of the technique as it shows the amount of data that the protocol is able to deliver to destinations. The end-to-end delay is the average time interval between the generation of a packet in a source node and the successful delivery of the packet at the destination node. This delay accounts for all possible delays that can occur in the source and all intermediate nodes. The routing overhead is the number of routing (control) packets generated during the simulated time in order to establish and maintain paths and to exchange traffic information among network nodes as dictated by the

TABLE 3.1: THE SYSTEM PARAMETERS USED IN THE SIMULATION EXPERIMENTS FOR TRAFFIC DENSITY AND NODAL ACTIVITY

Parameter	Values
Number of nodes	75, 100
MAC layer	IEEE 802.11
Transmission range	250m
Simulation area	900m×1000m,1200m×1000m
Simulation time	900s
Mobility model	Random waypoint model
Maximum speed	1, 2, 5, 7, 10, 15 and 20m/s.
Pause time	0s
Traffic type	CBR
Packet size	128, 512 bytes
Packet rate	2, 4, 8 packets/s
Number of flows	5
Number of runs per data point	30

operation of a given traffic aware metrics. For routing packets sent over multiple hops, each hop counts as one transmission. The routing overhead measures the scalability of the routing protocol and its efficiency in terms of consuming a node's battery power. The high routing overhead could affect the performance in terms of data throughput and end-to-end delay [59]. The high overhead is a result of factors like the unsuccessful delivery of route requests and the unsuccessful delivery of route replies. The interested for an in-depth discussion might refer to [59] for further details.

The statistics have been collected using 95% confidence intervals. In most cases the error bars have been found to be quite small; the error bars have not been included in all the figures for the sake of clarity and tidiness.

3.3.1 Performance impact of network size and traffic rate

This section conducts a performance analysis of the behaviour of the nodal activity and traffic density techniques for different network sizes of 75 and 100 nodes forming a MANET over a flat space of size 900m×1000m, and 1200m×1000m, respectively. We also study the impact of speed on the performance of the two techniques. The reason for choosing these numbers combination is to avoid the occurrences of frequent partitions in the network. The simulated network areas share the same width, but have different length. Traffic has been varied between 2, 4, and 8 packets/s, where 512-byte packets have been used. In the figures that are presented below, the x-axis describes the variations in node's maximum mobility speed, while the y-axis represents results of the performance metric of interest.

Throughput:

Figure 3.2 shows the throughput for the two traffic aware techniques when 75 nodes are used over a 900m×1000m topology area. Figure 3.2 (a) depicts the behaviour of the two techniques for a traffic rate of 2 packets/s. As can be seen in the figure, the achieved throughput by the two techniques has been affected by the variation of mobility speed. For a very low mobility at speed of 1m/s, nodal activity and traffic density exhibit close performance. This is potentially because the mobility activates the operation of the techniques by selecting new routes; less changeable topologies would result in keeping the same selected route and ignoring the existence of other choices. Increasing the mobility speed makes the difference between the two techniques more noticeable. Although the difference is not large, traffic density achieves higher throughput than that of nodal activity. For nodal activity the increase in speed to 20 m/s has resulted in a decrease in the achieved throughput of around 12% compared to that achieved at 1 m/s, and for traffic density the decrease is around 10%. The decrease in throughput is due to the fact that the increase in mobility causes frequent topology changes, resulting in more frequent broken routes as the mobility speed increases.

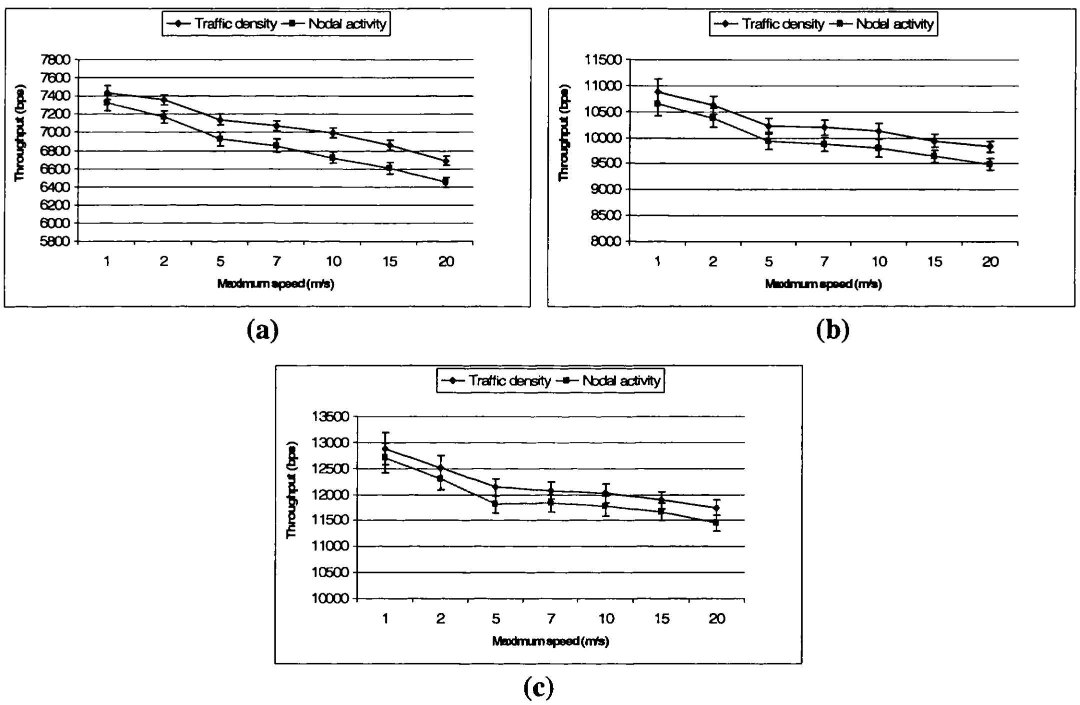


Figure 3.2: Throughput of the traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 3.2(b) shows the achieved throughput when traffic rate is 4 packets/s. Traffic density shows better performance over nodal activity, where the difference in throughput varies between 2-3.5%. The decrease in throughput due to speed increase has reached the limit of 1160 bps for nodal activity corresponding to 10.9% of its maximum achieved throughput, which is

10650 bps. In contrast, for traffic density, the decrease in throughput reaches 1050 bps that is equivalent to 9.7% of its maximum achieved throughput, which is 10890 bps.

Figure 3.2(c) depicts the achieved throughput under traffic rate of 8 packets/s. Similar to the results in Figure 3.2(b), the results reveal that traffic density outperforms nodal activity with a difference up to 300 bps (which is approximately 2.5% of its maximum achieved throughput). The decrease in throughput due to speed variation reaches the limit of 1250 bps for nodal activity corresponding to around 10% of its maximum achieved throughput, whereas for traffic density it reached 1140 bps that is equivalent to about 9% of its maximum achieved throughput.

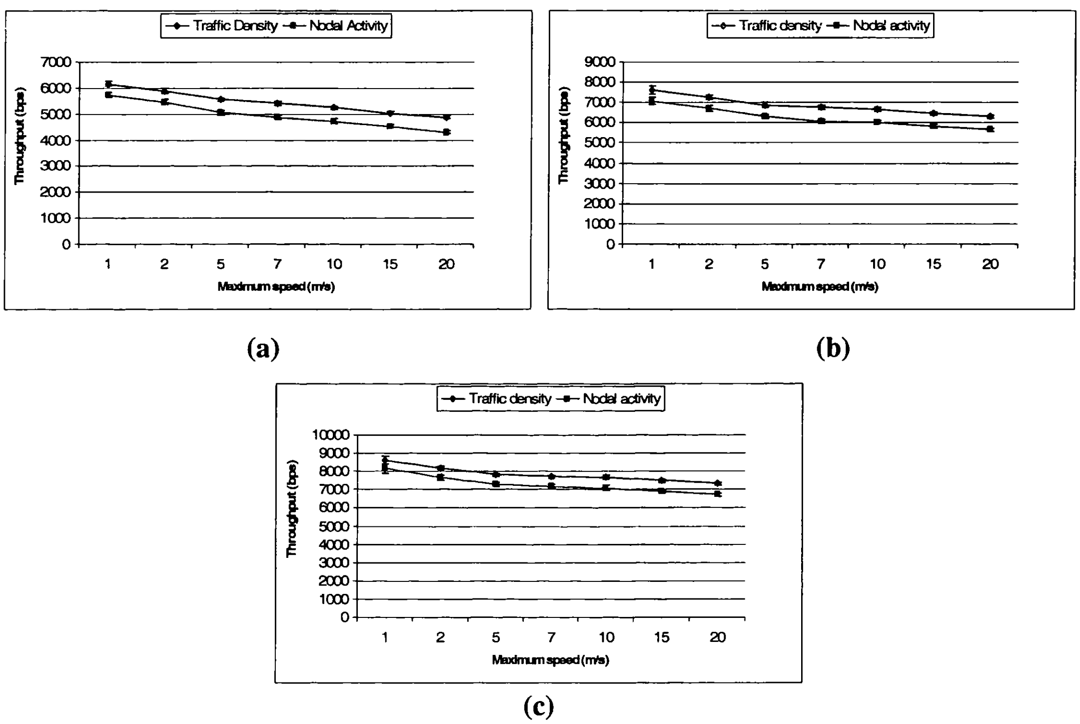


Figure 3.3: Throughput of the traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

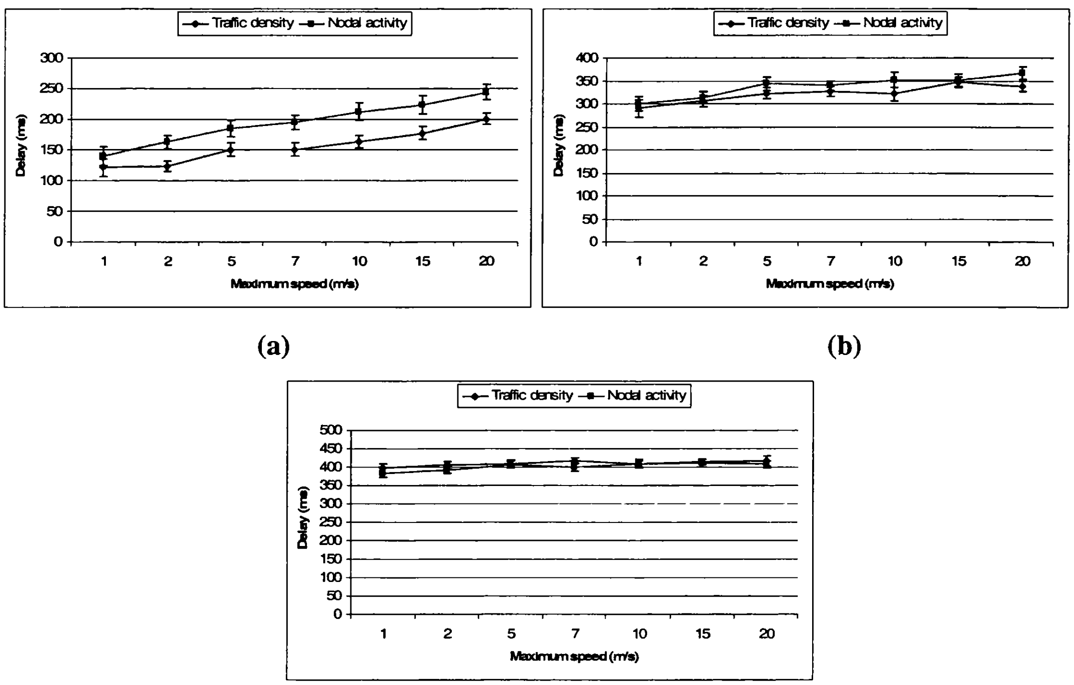
Figure 3.3 shows the throughput for the two traffic aware techniques when 100 nodes are used over a 1200m×1000m topology area. Figure 3.3(a) demonstrates the behaviour of the two techniques for traffic with a rate of 2 packets/s. We can observe from the figure that the throughput decreases as the mobility increases for both techniques. This is similar to what has been observed in the previous figure for a smaller network size. For traffic density, the decrease is up to 21% of its maximum achieved throughput (which is 6160 bps), while for nodal activity the decrease is 25% of its maximum achieved throughput (which is 5740 bps). Nonetheless, traffic density shows better throughput compared to nodal activity under various mobility speeds with a difference up to 570 bps (which corresponds to 9.3% of its maximum achieved throughput).

Figure 3.3(b) shows the performance behaviour when the traffic rate is increased to 4 packets/s. The figure shows that traffic density outperforms nodal activity with up to 650 bps (which corresponds to 8% of its maximum achieved throughput). Although the throughput decreases while nodes' speed increases the throughput difference between the techniques is almost the same. The decrease in throughput due to the increase in speed is up to 17% for traffic density, and for nodal activity the loss is 20%. Similar results are depicted in Figure 3.3(c) when higher traffic rate of 8 packets/s is used.

End-to-end delay:

Figure 3.4 shows the end-to-end delay for traffic density and nodal activity techniques under relatively light, moderate and high traffic with rates of 2, 4 and 8 packets/s, respectively, when 75 nodes are used over a 900m×1000m topology area. In Figure 3.4(a), the difference in performance between the two techniques is shown clearly in terms of end-to-end delay. Traffic density outperforms nodal activity over different speeds. Traffic density achieves on average about 40 ms less than that of nodal activity. The two techniques exhibit close performance at very low mobility of a speed of 1 m/s. The increase in speed affects both methods, but it is slightly more noticeable for nodal activity. Increasing the speed from 1m/s to 20m/s results in an increase in the end-to-end delay by 100 ms for nodal activity and 80 ms for traffic density. This is because the number of broken routes increases as the mobility speed increases which require a route discovery initiation to establish a new route. Since the route discovery latency is included in the end-to-end delay then the end-to-end delay increases.

Figure 3.4(b) shows the end-to-end delay for the two techniques when the traffic rate is increased to 4 packets/s. Traffic density and nodal activity exhibit close performance and the difference between the two is negligible. The speed variation has not shown significant effect on the delay except that it increases slightly as mobility speed increases. This is because more traffic has been injected into the network leading to an increase in the end-to-end delay as the network becomes congested (more than 35% of the packets has been dropped) which reduces the impact of higher speeds. A similar behaviour is shown in Figure 3.4(c) when higher traffic rate of 8 packets/s is used.



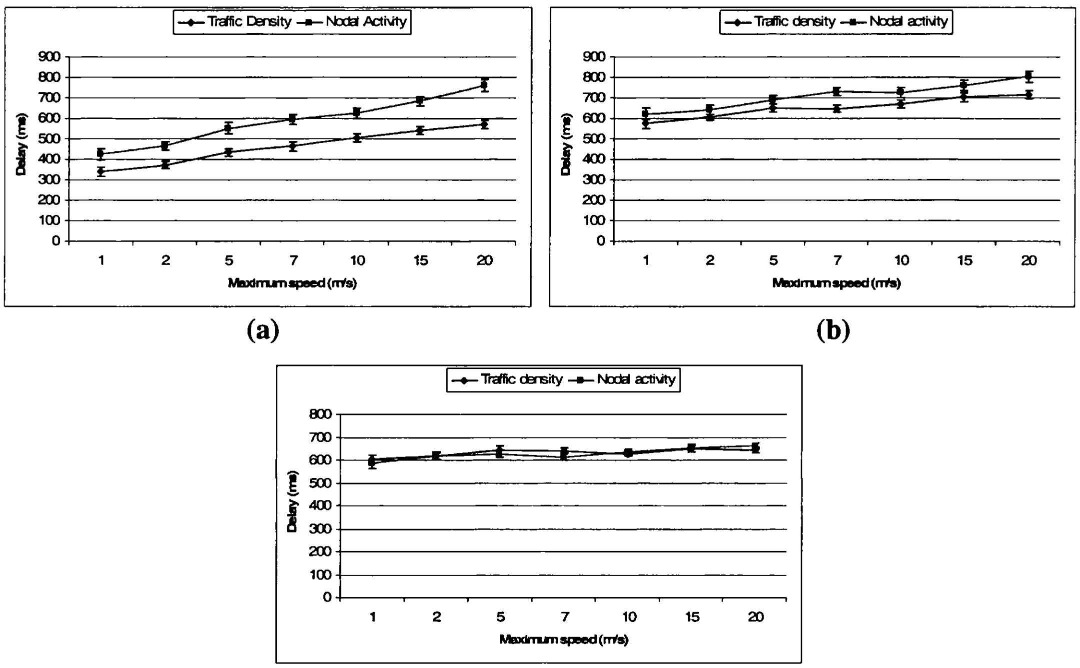
(a) (b)

(c)

Figure 3.4: End-to-end delay of the traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 3.5 shows the behaviour of the two techniques when the traffic with rate is 2, 4 and 8 packets/s, respectively, and 100 nodes are used over a 1200m×1000m topology area. Figure 3.5(a) shows that traffic density clearly outperforms nodal activity with a difference up to 190ms. The mobility speed has a great impact on both techniques. However the effect is more noticeable for nodal activity. For traffic density, the increase in the end-to-end delay is about

230 ms (about 67% of the minimum end-to-end delay), while for nodal activity the increase is about 340ms (about 80% of its minimum achieved end-to-end delay). On the other hand, the impact of increasing the network size can be seen clearly. The minimum end-to-end delay has significantly increased compared to that in Figure 3.4(a). For instance the minimum end-to-end delay for traffic density has increased from 120ms, in Figure 3.4(a), to 340ms, in Figure 3.5(a). This behaviour is because longer paths have been used.



(a) (b) (c)

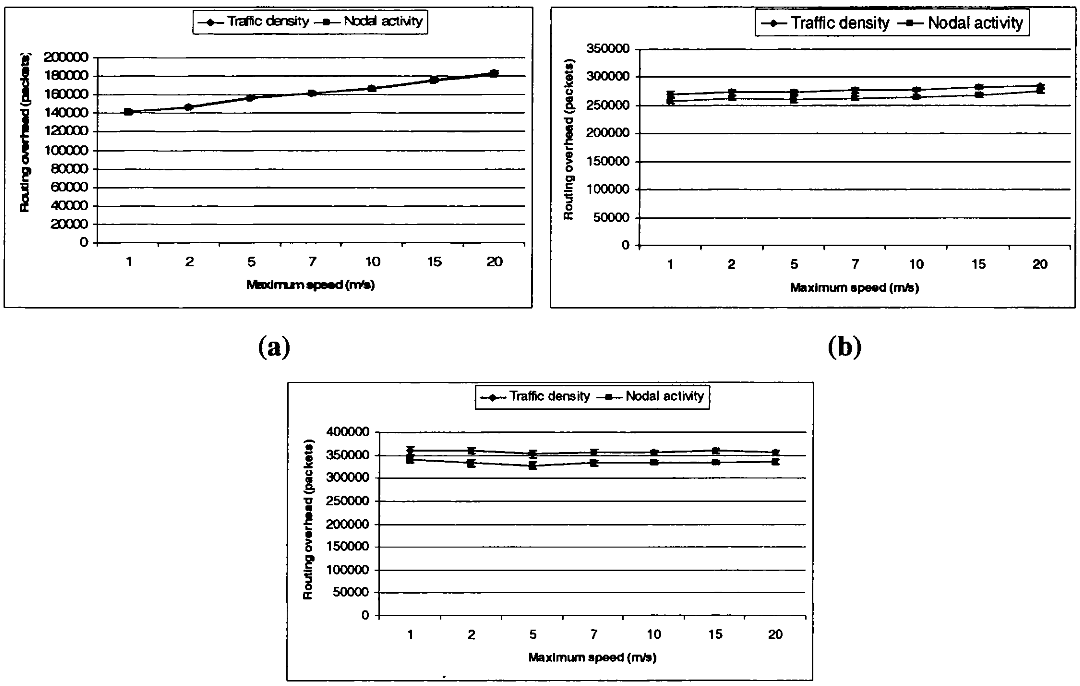
Figure 3.5: End-to-end delay of the traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Similarly in Figure 3.5(b) reveal that traffic density has a lower end-to-end delay, with a difference ranging between 40-85ms compared to nodal activity. But the minimum end-to-end delay in this figure is higher than that of figure 3.5(a) by 230ms. Moreover, the impact of speed is less than that in Figure 3.5(a). For traffic density, the increase in the end-to-end delay due to the increase in speed is about 140 ms (about 25% of the minimum end-to-end delay), while for nodal activity the increase is about 180ms (about 29% of its minimum achieved end-to-end delay). On the other hand, when increasing the traffic rate to 8 packets/s, the two techniques exhibit similar performance as it is shown in Figure 3.5(c) and the impact of speed is small. This is potentially because significant amount of packets are dropped. The calculation of end-to-end delay does not consider dropped packets, only less than 25% of sent packets have been successfully delivered. This also justifies why the maximum end-to-end delay in this figure is less than that in Figures 3.5(a) and 3.5(b). Nevertheless we should bear in mind that even if the two techniques have shown similar delay, traffic density has the upper hand in terms of throughput under the same scenarios.

Routing overhead:

Figure 3.6 show the generated routing overhead by the two techniques when the traffic rate is 2, 4 and 8 packets/s, respectively, and when 75 nodes are used over a 900m×1000m topology area. The figure 3.6(a) reveals that when the traffic rate is 2 packets/s the routing overhead increases noticeably as the mobility speed increases. However, both techniques exhibit similar performance behaviour in terms of this metric. The overhead increases by around 29% (that is around 40,000 routing packets) at the speed of 20 m/s compared to that at 1 m/s for both nodal activity and traffic density. This is because the occurrence of broken routes becomes more frequent as the mobility speed increases, which triggers route maintenance and route discovery sessions more often. This, in turn, causes an increase in the routing overhead. The reason why nodal activity and traffic density may generate similar overheads is because they share the same method of route discovery and both of them exchange neighbourhood information (e.g. number of active flows).

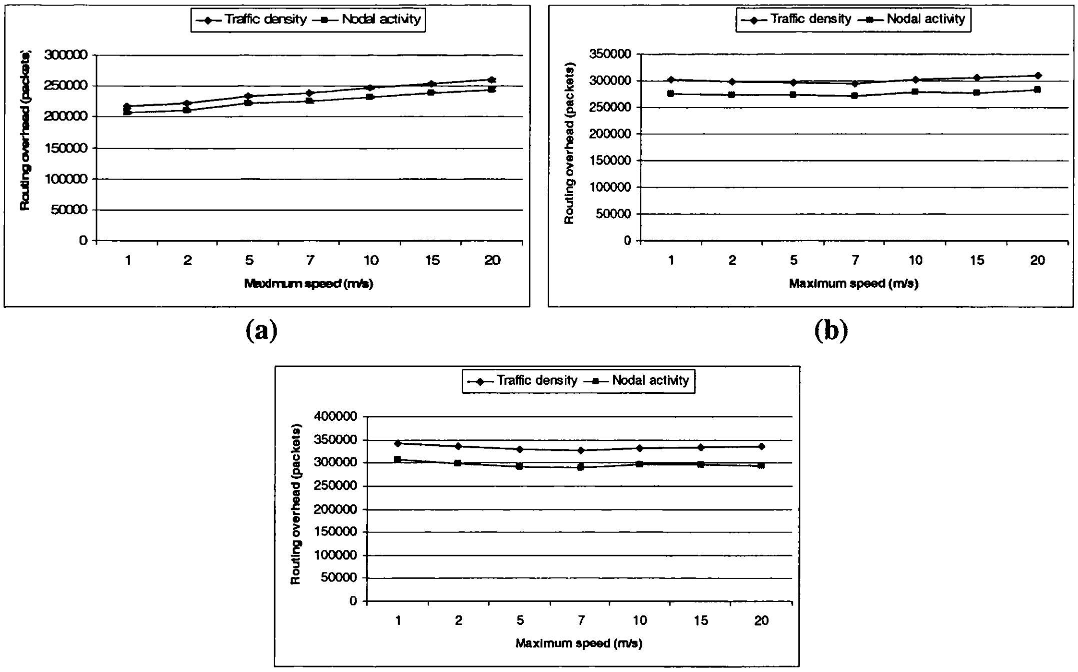
Figure 3.6(b) demonstrates the generated overhead from the two techniques for the relatively higher traffic rate of 4 packets/s. The figure shows that nodal activity generates 4.4% less overhead than traffic density. This is because nodal activity generates less route requests than traffic density. The maximum generated overhead is much higher than that in Figure 3.6(a). However the variation in mobility speed shows a difference on the generated overhead by the technique itself. The overhead increases up to 5.4% (18,000 routing packets) at speed of 20 m/s compared to that at 1 m/s for nodal activity and for traffic density up to 6.9% (15,000 routing packets). In Figure 3.6(c) where the traffic rate is 8 packets/s, nodal activity generates 6.5% less overhead compared to that generated by traffic density. However there is less impact on the generated overhead by the two techniques due to the variation in the mobility speed. This is because the impact of increasing the traffic reduces the impact of the speed.



(a) (b) (c)

Figure 3.6: Routing overhead of the traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 3.7 shows the routing overhead in traffic density and nodal activity for the traffic rates of 2, 4 and 8 packets/s, respectively, and when 100 nodes are used over a 1200m×1000m topology area. Figure 3.7(a) shows that the nodal activity generates less overhead, and as a result provides better performance. Figure 3.7(b) depicts traffic density with a higher overhead and nodal activity is the one with better performance, with a difference about 9% of overhead. Similarly in Figure 3.7(c) at traffic rate 8 packets/s, nodal activity outperforms traffic density with a difference around 11%.



(a) (b)

(c)

Figure 3.7: Routing overhead of the traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

3.3.2 Performance impact of data packet length

In this section, we demonstrate the effect of varying the packet length on the performance of traffic aware techniques. The parameters used in the previous Section 3.2.1 for the network size of 75 nodes with simulation area 900mx1000m are recalled here. The same mobility scenarios are used again and the same traffic scenarios are used except the change regarding the packet size that is set to 128 bytes.

Throughput:

Figure 3.8 depicts the performance of the two techniques in terms of throughput for the traffic rates of 2, 4 and 8 packets/s, respectively. Figure 3.8(a) shows when the traffic rate is 2 packets/s there is a performance advantage for traffic density over nodal activity, especially at higher mobility speeds. However, the difference between the two techniques is only 3% of the maximum achieved throughput. It is worth noting that the throughput when using a 128-byte packet should be lower than that when 512-byte packet since the same traffic rate is used but the packet size is smaller. Similar to Figure 3.2(a) where a 512-bytes packet is used, for a very low mobility at speed of 1m/s, nodal activity and traffic density exhibit close performance. The decrease in throughput is more apparent at higher speeds. Increasing the speed to 20 m/s results in a decrease in throughput up to 10.5% for nodal activity and 8% for traffic density compared to that achieved at 1 m/s.

Figure 3.8(b) shows the behaviour of the two techniques under medium traffic rate of 4 packets/s. Traffic density performs better than nodal activity. Although the throughput decreases while nodes' speed increases the difference in throughput between the techniques remains almost the same. The average difference between the two techniques is about 210bps, which is 5% of the maximum achieved throughput. While under the same scenario in Figure 3.2(b) but using 512-byte packet, the average difference is 3% of the maximum achieved throughput. Increasing the speed to 20 m/s results in a decrease in throughput of around 11.5% for nodal activity compared to that achieved at 1 m/s, and for traffic density the decrease is around 9.5%.

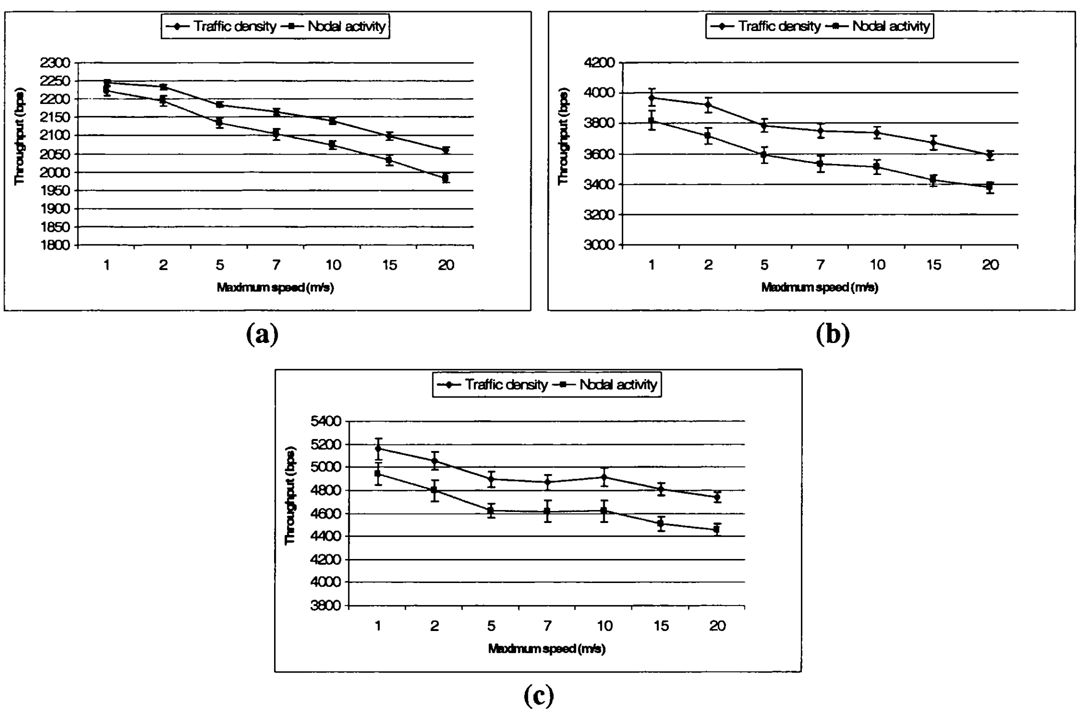


Figure 3.8: Throughput of the traffic density and nodal activity techniques for a network of 75 nodes over topology area 900mx1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

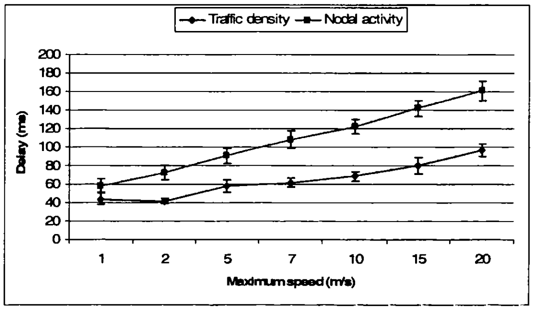
Figure 3.8(c) which depicts the achieved throughput under traffic rate of 8 packets/s shows a performance advantage in the favour of traffic density. The maximum difference between the two techniques is 270 bps, which is 5% of the maximum achieved throughput. While under the same scenario but using 512-byte packet, in Figure 3.2(c), the maximum difference is 2.5% of the maximum achieved throughput. The increase in speed to 20 m/s make throughput decreases by 10% for nodal activity compared to that achieved at 1 m/s, and for traffic density the decrease is 8%.

End-to-end delay:

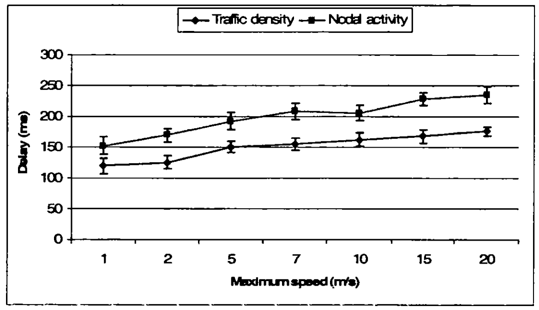
Figure 3.9 presents results for the performance behaviour of the two techniques in terms of the end-to-end delay for a traffic rate of 2, 4 and 8 packets/s, respectively. In Figure 3.9(a) the performance difference between the two techniques is shown clearly. Traffic density outperforms nodal activity over different speeds. Traffic density achieves a better end-to-end delay of 65 ms less than that achieved by nodal activity at high speeds, which is 40% of the maximum end-to-end delay (160ms). The two techniques exhibit close performance at very low mobility of a speed of 1 m/s. The increase in speed affects both techniques. However it is slightly more apparent for nodal activity. Increasing the speed from 1 m/s to 20 m/s results in an increase in the end-to-end delay by 100 ms for nodal activity, and 55 ms for traffic density.

Figure 3.9(b) depicts the end-to-end delay of the two techniques for a traffic rate of 4 packets/s. Traffic density has better performance over the simulated speeds with a performance difference ranges between 30-60 ms, while the maximum end-to-end delay is about 240ms. The increase in speed affects both methods, but it is slightly more noticeable for nodal activity. Increasing the speed from 1 m/s to 20 m/s results in an increase in the end-to-end delay by 85 ms for nodal activity, which is lower than that shown in Figure 3.9(a). For traffic density, the increase is 55 ms, which is comparable to that shown in Figure 3.9(a).

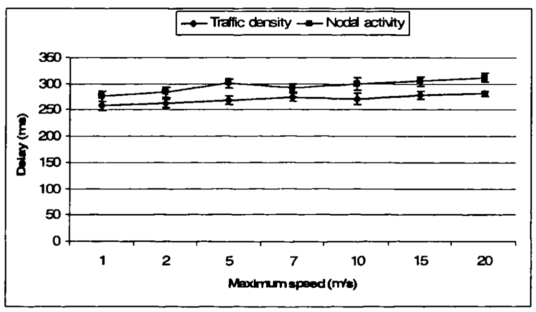
Figure 3.9(c) depicts the end-to-end delay of the two techniques for a traffic rate of 8 packets/s. The two techniques exhibit close performance various mobility speeds. Nonetheless, traffic density exhibits better performance over the simulated speeds with a performance difference ranges between 20-30 ms. The increase in speed affects both methods, but it is slightly more for nodal activity. Increasing the speed from 1 m/s to 20 m/s results in an increase in the end-to-end delay by 35 ms for nodal activity, and by 20 ms for traffic density.



(a)



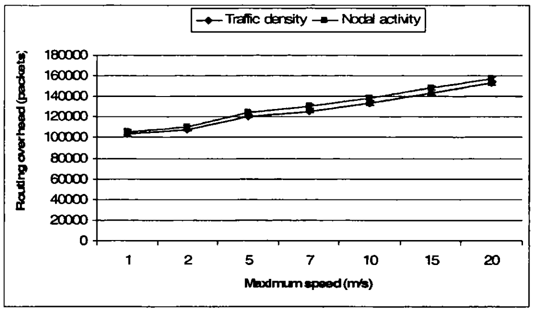
(b)



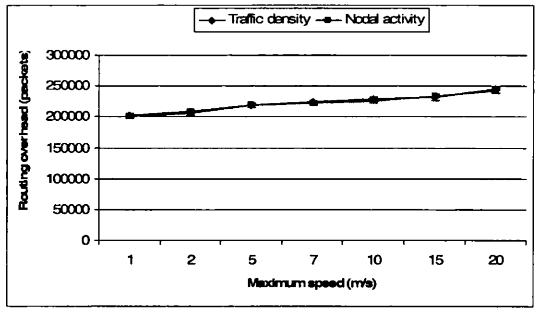
(c)

Figure 3.9: End-to-end delay of the traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

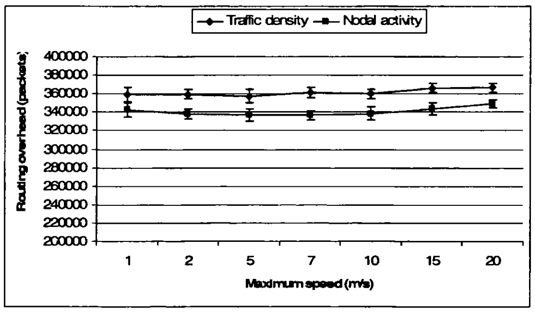
On the other hand, the impact of packet size can be seen clearly. When comparing the end-to-end delay in figures 3.4 and 3.9 it is noticeable that the minimum and maximum end-to-end delays significantly increase when using larger packet size. This is because the network can handle small size packets more efficiently and faster. Also it is clear that the minimum and maximum end-to-end delays in Figure 3.9(a) are less than that in Figure 3.4(a) where 512-byte packets are used. A similar behaviour is depicted when comparing Figure 3.4(b) with Figure 3.9(b) and Figure 3.4(c) with Figure 3.9(c), but with one addition that is both the minimum and the maximum end-to-end delays in Figures 3.9(b and c) are less than the minimum end-to-end delay in Figures 3.4(b and c).



(a)



(b)



(c)

Figure 3.10: Routing overhead of the traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Routing overhead:

Figure 3.10 shows the performance of the two techniques in terms of routing overhead for traffic rates of 2, 4 and 8 packets/s, respectively. In Figure 3.10(a), both techniques generate similar routing overhead. However the routing overhead increases noticeably while the mobility speed increases. The overhead increases by 48% (about 50,000 routing packets) at speed of 20 m/s compared to that at 1 m/s for both nodal activity and traffic density.

Both techniques use a similar method for discovering a route (see Section 2.5). However they are different in the way of selecting the route. If the selected route is stable this would require less routing packets to maintain the route and hence less overhead. The reason why nodal activity and traffic density generate comparable overheads is because they share the same method of route discovery and both of them exchange neighbourhood information (e.g. number of active flows). The difference in overhead would happen if the traffic aware technique selects a less stable route (e.g. more congested).

Figure 3.10(b) demonstrates the generated overhead from the two techniques for the traffic rate 4 packets/s. In this scenario, nodal activity generates comparable overhead to that generated by traffic density. The variation of mobility speed does not show any significant difference on the generated overhead by the two techniques. Although the maximum generated overhead is much higher than that in Figure 3.10(a), the overhead increase due to mobility is roughly similar. The overhead increases by 20.6% (41,000 routing packets) at the speed of 20 m/s compared to that at 1 m/s for nodal activity and by 21.8% (46,000 routing packets) for traffic density. In Figure 3.10(c), nodal activity generates less overhead compared to that generated by traffic density when the traffic rate is 8 packets/s. However, the overhead generated by the two techniques is not affected significantly by the variation in the mobility speed. Nonetheless, nodal activity generates on average 5.7% (20,000 routing packets) less than that of traffic density over different speeds.

3.4 Conclusions

This chapter has conducted the first performance comparison of two existing end-to-end traffic aware routing techniques, namely *traffic density* and *degree of nodal activity*, to assess their performance behaviour under similar working environments. The first part of the comparative analysis has been carried out through studying the effects of varying the network size in terms of the topology dimensions and the number of nodes deployed over that topology. The second part was intended to demonstrate the behaviour of the two techniques when varying data packet length.

Throughout the simulation experiments the two techniques tended to show close performance in terms of throughput and delay for small sized networks. This could be due to the fact that the two techniques might be using paths with comparable load conditions, but these paths are not necessarily the same. However with increasing the network size from 75 to 100 nodes, the two techniques demonstrate performance differences. Traffic density shows a better end-to-end delay up to about 32% lower than that for nodal activity. Furthermore, in terms of throughput, traffic density outperforms nodal activity with a difference up to 9%. However, for routing overhead the nodal activity technique achieves better performance with a difference up to 12%. This is because nodal activity generates less route requests than traffic density.

Traffic density achieves higher throughput than in nodal activity when the packet length is decreased from 512 to 128 bytes, with a relative difference around 5%. Further, the difference in end-to-end delay between the two techniques becomes more noticeably in favour of the traffic density technique with a difference of approximately 40% less than in nodal activity. However, the difference becomes less noticeable with respect to the routing overhead as the packet length decreases.

To sum up, simulation results have shown that the traffic density technique outperforms its nodal activity counterpart in both throughput and end-to-end delay in most of the simulated scenarios. But, the latter technique exhibits comparable routing overhead in certain considered cases. The subsequent chapter will describe a new end-to-end traffic aware technique that overcomes the limitations of both the existing traffic density and nodal activity techniques. The performance results presented in the next chapter will show that the new traffic aware approach exhibits superior performance characteristics over the existing traffic density and nodal activity techniques in most studied scenarios.

Chapter 4

Load density: A new traffic aware technique for MANETs

4.1 Introduction

The main idea behind traffic aware routing is to avoid the condition that could occur in traditional routing [43, 68] where few nodes have to carry excessive loads compared to the rest of the network [28, 37]. The excessive load which includes both the past and current traffic load would result in high battery consumption at these nodes and regions in the network being more congested than others. In order to distribute the load over the network nodes, a distribution method is required. Achieving a good load distribution in MANETs, where nodes in the network have comparable loads, is a challenging task. The issue is complicated not only because a full knowledge of the network status is required in order to make good decisions, but also because it has to deal with load distribution in the presence of nodes mobility. The existing techniques for distributing the load in MANETs use either the end-to-end approach or the on-the-spot approach as classified earlier in Chapter 1. Distributing the load over nodes using the end-to-end approach, which is the focus of this chapter, is achieved by using a route selection criterion when routes are being established.

Existing end-to-end traffic aware techniques, e.g. [37, 75], suffer from a number of limitations in their estimation of the load. In order to make a good judgment about the load, it may not be *sufficient* to only capture the number of active paths as is in the case of the nodal activity technique [37] or estimate the number of packets at the MAC layer interface queue as in the traffic density technique [75]. The number of active paths could be useful if nodes keep the correct estimation of number of active paths. However mobility might affect the accuracy of number of active paths at the intermediate nodes of a given path leading to inaccurate cost of that path, unless these changes were detected on time. Furthermore, the number of active flows does not capture the history of experienced traffic at intermediate nodes and could represent only the current traffic load. The number of packets at the interface queue is useful for capturing contention at the MAC layer. However, this is might not be sufficient to represent the load experienced by the node over time which has consumed part of its battery and reduced its operational lifetime as it could represent only the current transiting traffic. One of the most limited resources of a mobile node is its power, and since most of the power consumption of a node is due to sending and receiving packets [87], reducing the amount of packets transferred by a node helps to reduce its power consumption, leading to an increase in its lifetime. A more detailed discussion on the limitations of these techniques will follow in Section 4.2.

Motivated by the above, the goal in this chapter is to develop and evaluate a new traffic aware technique, based on the end-to-end approach, referred to here as *load density*, which can overcome the limitations of the existing techniques. To achieve this, the new proposed technique takes more accurate and continuous information (i.e. load history) about the traffic experienced by the nodes compared to the existing techniques. This information is computed by using the history of lengths of packets passed over the nodes and the number of packets waiting at the nodes' interface queue. The use of packets lengths in the calculations rather than just using the number of packets gives more precise estimation of the amount of traffic forwarded by the node, and it covers any variation in packets' sizes, while the number of packets at the interface queue provides an indication of the degree of contention over the wireless communication medium. One of the main features of the suggested *load density* technique is that the amount of exchanges of load information among neighbouring nodes is lower compared to the other similar methods since it depends only on the exchanged information via the route request, which results in reduced communication overhead, as will be shown later in the subsequent sections.

This chapter also conducts an extensive evaluation of our proposed load density technique by means of ns-2 simulations, and compares its performance against existing techniques. The obtained results from the previous chapter, Chapter 3, on the comparison between two typical end-to-end techniques, namely *nodal activity* and *traffic density*, have shown a superiority of traffic density over nodal activity in terms of throughput and end-to-end delay. However, both techniques have shown close performance in some of the simulated scenarios, e.g. when a network of 75 nodes over topology area 900m×1000m and a 512-byte packet are used. Moreover, nodal activity has shown better performance in terms of routing overhead. Therefore to conduct a fair and realistic comparison, we compare the performance of our load density against that of both traffic density and nodal activity techniques. Also, in our comparison we evaluate our proposed technique under the same assumptions and the same scenarios that have been used in Chapter 3. In addition, we extend the performance comparison by examining the effect of the background traffic.

The rest of this chapter is organised as follows. Section 4.2 discusses the limitations of existing traffic aware techniques. Section 4.3 describes the proposed traffic aware technique and presents its algorithm. Section 4.4.1 analyses the effects of the network size in terms of number of nodes and the size of network area on the performance of the traffic aware techniques. Section 4.4.2 analyses the effects of packet length. Section 4.4.3 analyses the effects of number of flows. Finally, section 4.5 concludes the chapter.

4.2 Limitations of existing traffic aware techniques

The existing end-to-end traffic aware techniques attempt to distribute the load by using a metric (or cost function) to select the route with a minimum load. Such techniques are represented by nodal activity [37] and traffic density [75]. However these techniques suffer from a number of limitations in their judgment of the load.

The number of active paths, which is used in the nodal activity technique [37], could be useful if nodes do not move frequently as the number of active paths at the intermediate nodes of a given path may be inaccurate. The calculated number of active paths could be more than the actual number of currently used paths such as when an intermediate node move outside the path, leading into inaccurate judgment of the path's load. The effect of this might increase when the activity information obtained from the neighbours of the intermediate nodes forming the path is not accurate. Another limitation, that is worth mentioning, could come from the inherent characteristics of traffic flows (e.g. similar traffic rates, similar packet lengths) when the flows have less identical characteristics, the case in which a direct comparison between the number of active paths of two paths (or two nodes) might not be accurate.

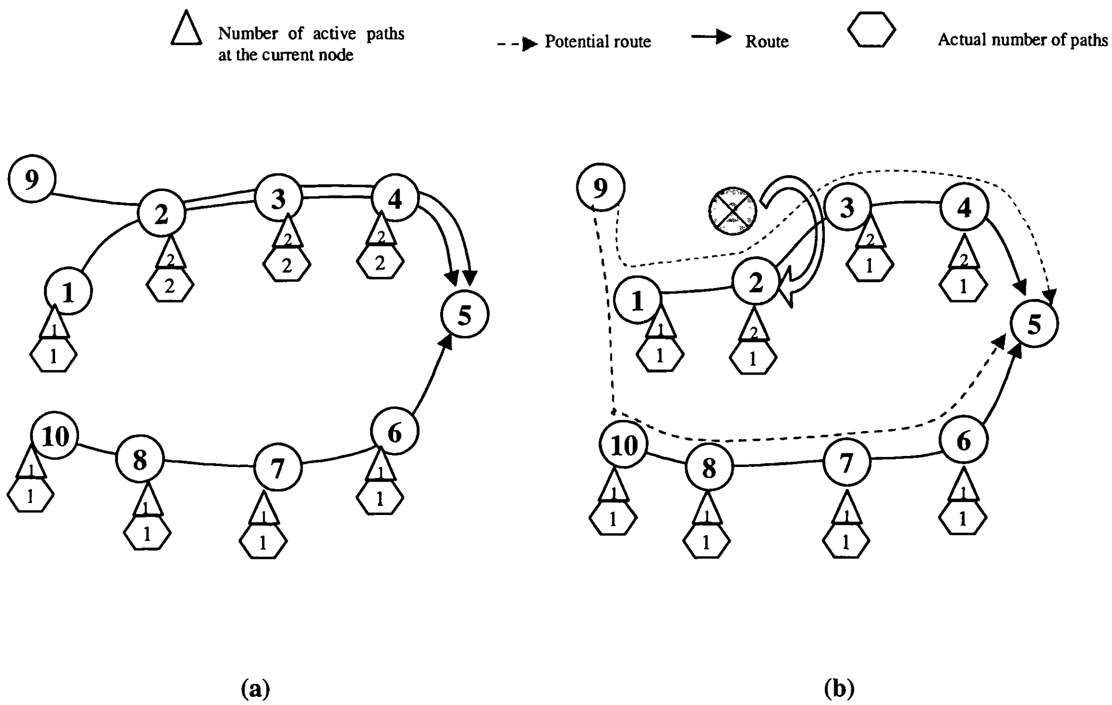


Figure 4.1: An example to show the potential inaccuracy in the calculated number of active paths in nodal activity.

Figure 4.1 provides an illustrative example that describes how the calculated number of active paths in nodal activity can be different from actually existing number of paths. The example consists of a set of nodes and three flows as shown in Figure 4.1(a). Associated with each node the number of active paths (the path is considered inactive if it has not been used for a period of time, e.g. 3 seconds) and the number of actual paths (the paths that are actually going through that node). In Figure 4.1(b), node 2 moves away from the range of node 9, this will result in a route break between node 9 and node 5, and hence the route between the two nodes which passes through nodes 2, 3, and 4 is actually no longer active. However nodes 2, 3, and 4 still consider that the number of active paths as 2, while it is actually 1.When node 9 detects that the link with node 2 is broken, it will try to establish a new route to node 5. Two potential routes can be established: 9-1-2-3-4-5 and 9-1-10-8-7-6-5 as it is shown in Figure 4.1(b). The calculated cost of second route is lower than the calculated cost for the first, so the second route is selected. However the real cost for the first one is lower. Furthermore the second route is longer in terms of number of hops which might result in an increased end-to-end delay.

The number of packets at the interface queue, which is used in the traffic density technique [75], could be useful for capturing contention at the MAC layer. However, this is not sufficient to represent the load experienced by the node over time. One of the most limited resources of a mobile node is its power, and since most of the power consumption of a node is due to sending and receiving packets [19], reducing the amount of packets transferred (forwarded) by a node helps to reduce power consumption, leading to an increase in its lifetime [87]. Contention information could be used to distribute the current workload over nodes but it might not be useful for distributing the load over long time period as the level of contention often varies over time.Further, the number of packets at the interface might not be accurate when the packets that are being forwarded over nodes have different lengths. So, a metric is required that can deal with most of the cases that could arise in the network; that is, a metrics that can capture the load as well as contention along the path.

4.3 The proposed load density technique

The proposed technique, named *load density*, uses two main components; the *load history* information represented by the total traffic passed over a node, and the *contention* information represented by and the number of packets waiting at the nodes' interface queue in order to take the possible contention in the network into consideration in the calculations. The new load density technique is embedded into the reactive AODV-like routing protocol as we have done in the analysis of the existing techniques nodal activity and traffic density, in Chapter 3. The framework used to implement traffic aware techniques including load density has been described earlier in Section 2.5. The sequel will describe in more detail the operation of load density.

Route discovery:

The route discovery process starts whenever a node wants to communicate with another node for which it does not have a known route. The source node broadcasts a *route request* packet to its neighbours. Every node receives the route request packet will forward it to its neighbours after updating the cost information carried in the route request packet, by adding the values of its load-history and contention information (see Section 4.3.1) to those carried in the packet. The cost information carried in the route request packet, which includes the load history and the contention information, represents the cost seen from the source to the current node. The process of forwarding the route request continues until the packet is received by the destination node. The destination collects the arriving route request packets within a route *selection period*; activated upon receiving the first route request packet, for selecting the best-cost route. Once the selection period is expired the destination selects the route with the best cost based on three parameters: traffic load history, contention information and route length (in hops) and then sends a *route reply* packet to the source node. When the source node receives the route reply packet, the path is then established and communication can be started.

The route selection algorithm searches for routes that have contention below the predetermined *contention threshold*; in order to give preference to routes with low contention in order to keep good performance, and then among these routes it selects the route with minimum load history. If there are other routes with *acceptable-load-difference* from the route with minimum load history, then the route with lowest number of hops is selected. Otherwise, it selects the route with minimum load history. The route selection process is illustrated in more details in Figure 4.2. The current settings of *selection-period*, *contention-threshold*, and *acceptable-load-difference* presented in the route selection algorithm have been set as illustrated in Table 4.1.The values of contention-threshold and acceptable-load-difference have been set to accommodate the selection of the route with minimum traffic load history in most of the cases. The value of contention-threshold is set to the size of the interface queue (which is 50 packets). The selection-period has been set similar to nodal activity and traffic density.

TABLE 4.1: ROUTE SELECTION ALGORITHM THRESHOLD VALUES.

Parameter	Value
selection-period	50 ms
contention-threshold	50 packets
acceptable-load-difference	1024 Bytes

Load density traffic-aware routing algortihm *(Route selection) {*
// For selecting the route three parameters are used:
// traffic load history, contention information and route length in hops.

Collect all route requests packets sent from source S and received within the selection-period
// The selection-period is started when the first route request packet is received.
// Each route request packet corresponds to a route from source to destination
Find the set of routes R that has contention value ≤ contention-threshold
From the set R find the route r with minimum traffic-load-history
Compare the routes' traffic-load history with r's traffic-load history If the difference < acceptable-load-difference
then select the route with the lowest number of hops and send reply to the source
else select r as route and send reply to the source
If all routes available have contention values > contention-threshold
then select the route with minimum traffic-load history and send reply to the source
} End (Route selection)

Figure 4.2: Route selection algorithm in load density technique.

Route maintenance:

The route maintenance is triggered when there is a change in the topology that affects the validity of an active route is detected. If either the source node, an intermediate node, or the destination node on an active route moves out of the communication range, an alternative route must be found. Once a node detects that the next hop is unreachable, it propagates a *route error* packet to inform the source node and the other nodes which forward the packet that this route is no longer valid. Nodes that receive the route error packet delete the affected routes. The source upon receiving the route error packet can initiate a new route discovery to find a new route instead of the broken one by broadcasting a new route request. A new route discovery process is initiated rather than using any stored routes instead in order to deal with mobility and its effect on the established routes.

4.3.1. Computation of the route cost

The cost function in load density has two main components: the data traffic load (in bytes) forwarded by nodes and the number of packets present at the interface queue. Each node keeps information about the amount of traffic passed over it in addition to the interface queue history represented by the averaged number of packets occupying the queue over a period of time. The route cost is calculated by gathering traffic load history and contention information for the nodes along the route. The contention information for a node represents the number of packets at the interface queue.

The load history of a node i is calculated by accumulating the data packets lengths passed over that node as expressed in equation 4.1, where p is a data packet passing over node i.

$$load\_history(i) = \sum_{p \in i} length(p) \qquad (4.1)$$

The traffic load of a route say r is calculated using equation 4.2 below, where i is a node on the route r, excluding the source and destination nodes.

$$traffic\_load(r) = \sum_{i \in r} load\_history(i) \qquad (4.2)$$

The average number of packets at the interface queue N_Q is computed using a number of samples, say n, where a sample is taken every sample-period (10 ms), as written in equation 4.3

$$N_Q(i) = \frac{\sum_{k}^{n} q_k^i}{n} \qquad (4.3)$$

The contention of the route r is calculated by adding the contention of the nodes that are forming the route as in equation 4.4, excluding the contention at the source and destination nodes

$$route\_contention(r) = \sum_{i \in r} N_Q(i) \tag{4.4}$$

The contention information is based on the local contention information known at the node rather than relying on the information exchanged with the neighbours. Each node updates its local contention information every 500ms i.e. after taking 50 samples of the interface queue.

Example:

To illustrate the process of calculating the cost of routes and how a route is selected among the available routes, let us consider a simple scenario that is depicted in Figure 4.3. The network consists of 9 nodes with a presumable load history and contention information associated with each node as it is shown in the figure. Nodes 1 and 6 represent the source (S) and destination (D) pairs. Initially the source node; node 1, initiates a route request packet that is broadcasted to its neighbours, which are node 2 and 9. Nodes 2 and 9 add their load history and contention information to the route request packet, and then forward it their neighbours. Every node that receives the route request packet updates the packet by adding its load history and contention information to that in the route request packet and then forwards the packet to its neighbours. The route request packet maintains a hop count; each time it is forwarded the hop counter is incremented by 1. The process of updating and forwarding the packet continues until the packet

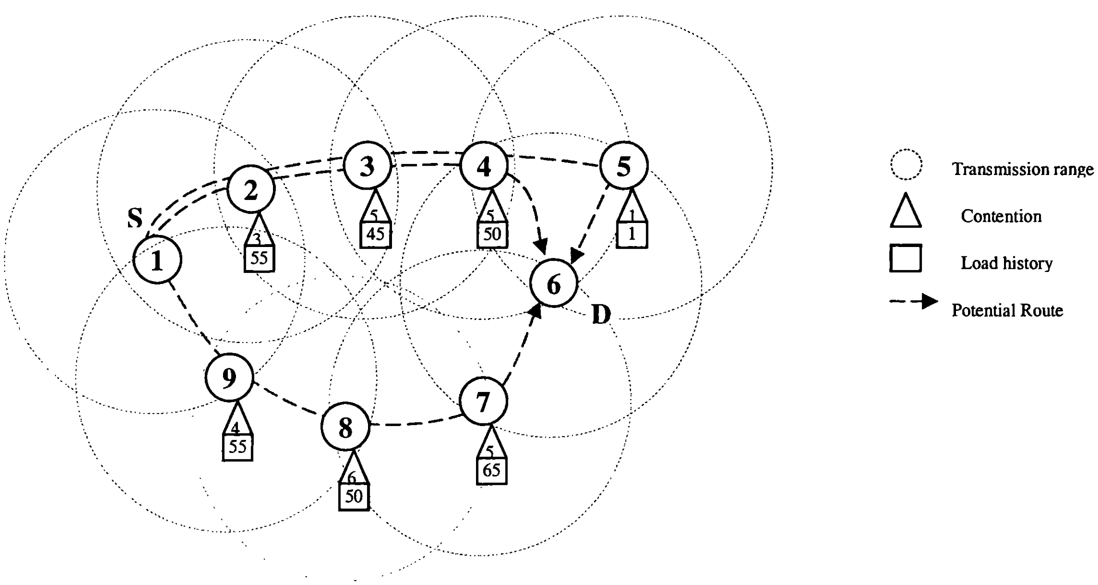


Figure 4.3: Illustration example of route selection in load density technique.

is delivered by the destination. From this scenario three routes from source to destination are available: 1-2-3-4-6, 1-2-3-4-5-6, and 1-9-8-7-6. Let us name these routes as r1, r2 and r3 respectively.

The cost parameters (load, contention and hop count) associated with the three routes are as follows: r1(150, 13, 4), r2(151 , 14, 5), r3(170, 15, 4). According to the route selection algorithm in Figure 4.2, we first look for routes with a contention value less than the contention threshold to give preference to routes with low contention. All three routes satisfy this condition. The next step is to find the route with minimum traffic load, which is r1 in our example. The final step is to find out if there are other routes with comparable traffic load but with lower hop count (i.e. shorter route). Route r2 has a comparable traffic load but since it has higher hop count, r1: 1-2-3-4-6 is selected to be the route between node 1 and node 6.

4.4 Performance Analysis

This section analyses the performance of the suggested end-to-end load density technique, and compares it against the existing *traffic density* and *nodal activity*. The performance analysis has been conducted using the same simulation model and parameters as outlined in Chapter 3 (see Section 3.2), and for the sake of clarity we include the list of parameters here (see Table 4.2).

The performance of the three techniques is measured by four performance metrics: *throughput*, *end-to-end delay*, *routing overhead* and *normalised load covariance* . The throughput is the amount of data received (measured in bits per second) at the final destination over the simulated time averaged over number of flows. This measure provides an indication of the efficiency of the technique as it shows the amount of data that the network is able to deliver to destinations. The end-to-end delay is the average time interval between the generation of a packet at a source node and its successful delivery at the destination node. This delay accounts for all possible delays that can occur in the source and all intermediate nodes. The routing overhead is the number of routing (control) packets generated during the simulated time in order to establish and maintain paths and to exchange traffic information among the network nodes as dictated by the operation of a given traffic aware metric.

The normalised load covariance is calculated by dividing the load covariance of the data forwarded by the number of packets successfully delivered at the destinations. Load covariance is defined according to equation 4.5, where node's load l_i is the sum of data packet lengths passed over node i. The mean μ is the total sum of the nodes' load over the number of nodes M participating in forwarding data packets. The rational behind this metric is to measure the ability of distributing the load while successfully delivering data packets. The lower this metric the better performance is since the load covariance measures how far are nodes' load from the mean.

$$\text{Load covariance} = \frac{\sum_{i}^{M}(l_i - \mu)^2}{M * \mu} \tag{4.5}$$

TABLE 4.2: THE SYSTEM PARAMETERS USED IN THE SIMULATION EXPERIMENTS

Parameter	Values
Number of nodes	75, 100
MAC layer	IEEE 802.11
Transmission range	250m
Simulation area	900m×1000m,1200m×1000m
Simulation time	900s
Mobility model	Random waypoint model
Maximum speed	1, 2, 5, 7, 10, 15 and 20m/s.
Pause time	0s
Traffic type	CBR
Packet size	128, 512 bytes
Packet rate	2, 4, 8 packets/s
Number of flows	5
Number of runs per data point	30

4.4.1 Performance impact of network size

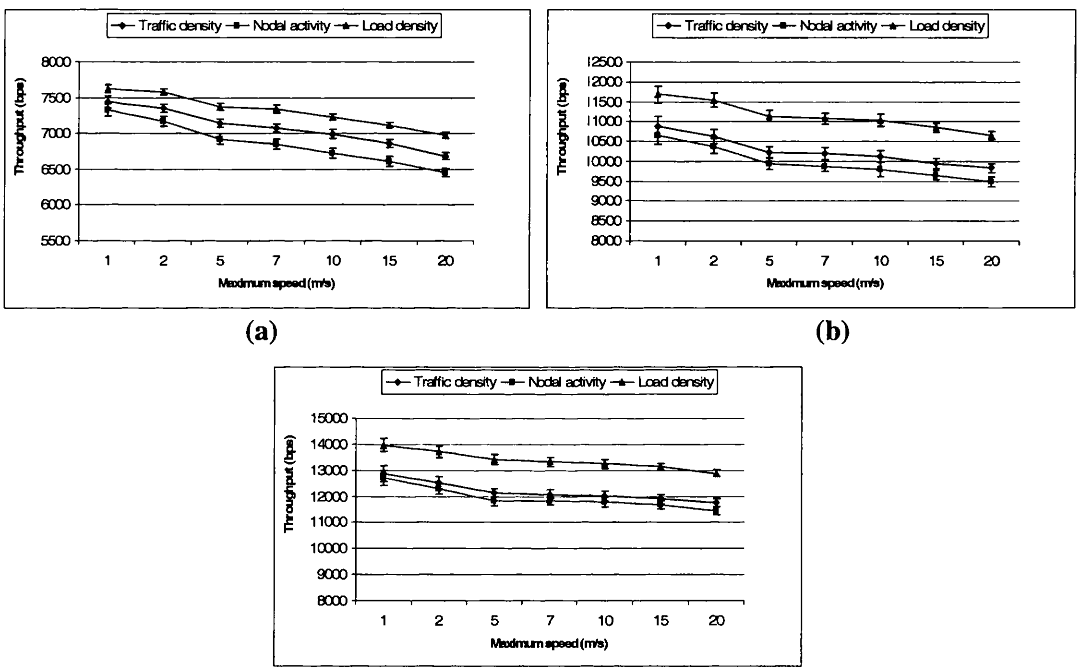
The analysis is conducted using two network sizes of 75 nodes forming a MANET over a simulation area of 900m×1000m and 100 nodes forming a MANET over a simulation area of 1200m×1000m. The reason for choosing these network settings is to avoid the occurrences of frequent partitions in the network. The simulated network areas share the same width of 1000 m, but have different lengths of 900m and 1200m. The node maximum speed has been set to 1, 2, 5, 7, 10, 15 and 20 m/s; representing human slow walk speed to vehicle speeds. Traffic rate has been varied between 2, 4, and 8 packets/s, where 512-byte packets have been used. In the figures that are presented below, the *x*-axis describes the variations in node's maximum mobility speed, while the *y*-axis represents results of the performance metric of interest. The results have been conducted from steady state simulations. A 95% confidence interval is calculated for each point. In all of the cases the error bars are quite small. These error bars are not shown on all the figures in order to increase their clarity.

Throughput:

Figure 4.4 shows the throughput for the three traffic aware techniques when 75 nodes are used over a 900m×1000m topology area. Figure 4.4 (a) depicts the behaviour of the techniques for a traffic rate of 2 packets/s. As it is shown in the figure, the throughput achieved by the three techniques has been affected by the variation of mobility speed. For a very low mobility at maximum speed of 1m/s, load density, nodal activity and traffic density exhibit comparable performance compared to higher mobility speeds. This is because the mobility activates the operation of the techniques by selecting new routes; less changeable topologies would result in keeping the same selected route and ignoring the existence of other choices. However increasing the mobility speed makes the difference between load density and the other techniques more noticeable. Although the difference is not large, load density achieves higher throughput than that of nodal activity and traffic density. Load density achieves on average over the varied speeds about 3.5% higher throughput than the throughput of traffic density, and about 6.5% higher than that of nodal activity.

Figure 4.4(b) shows the achieved throughput when traffic rate is increased to 4 packets/s. The figure reveals that load density outperforms both traffic density and nodal activity at all simulated speeds, with a difference of 7.6% (880bps) from the throughput achieved by traffic density and a difference of 11.7% (1170 bps) from the throughput achieved by nodal activity.

The figure also demonstrates the impact of changing the mobility maximum speed from 1 m/s to 20 m/s on the achieved throughput. For instance, for load density the decrease in throughput due to mobility speed increase reaches the limit of 1020 bps; that is about 8.7% of its maximum achieved throughput which is 11680 bps. In contrast, for traffic density, the throughput decrease reaches 1050 bps that is equivalent to 9.7% of its maximum achieved throughput, which is 10890 bps, and for nodal activity the decrease in throughput corresponds to 10.9% of its maximum achieved throughput, which is 10650 bps. The decrease in throughput is due to the increase in mobility which causes frequent topology changes. This results in more broken routes as the speed increases.

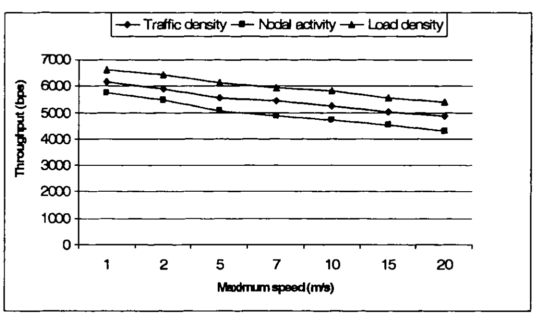


(a) (b) (c)

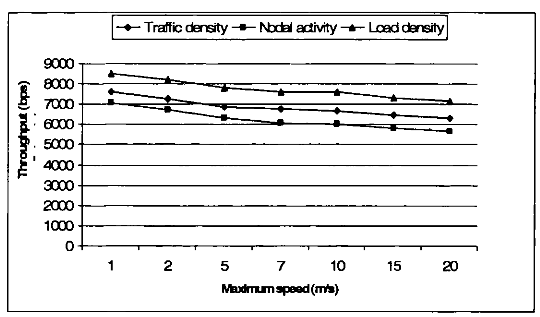
Figure 4.4: Throughput of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 4.4(c) depicts the achieved throughput when the traffic rate is 8 packets/s. Similar to the results in Figure 4.4(b), load density outperforms the other two techniques, with a difference of 1200bps higher than the throughput achieved by traffic density, and a difference up to 1450 bps from the throughput achieved by nodal activity, over various mobility speeds. These differences correspond respectively to 9% and 10.8% of the average throughput achieved by load density (which is 13390bps).

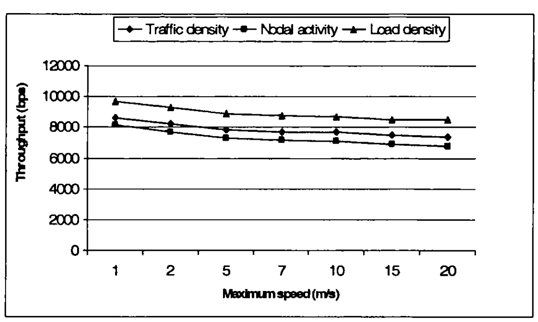
Figure 4.5 shows the throughput for the three traffic aware techniques when the network size is scaled up to 100 nodes distributed over a 1200m×1000m topology area. Figure 4.5(a) demonstrates the behaviour of the three techniques for a traffic rate of 2 packets/s. We can observe from the figure that the throughput decreases as the mobility increases for the three techniques. This agrees with what have been observed in Figure 4.4 when the area of 900m×1000m has been considered. For load density the decrease in throughput is 18.5% of its maximum achieved throughput which is 6640 bps. For traffic density and nodal activity, the decrease compared to their maximum achieved throughput which is 6160 bps and 5740 bps, respectively, is up to 21% and 25%, correspondingly. Nonetheless, load density outperforms both traffic density and nodal activity under various mobility speeds. Load density achieves on average over the varied speeds about 9% of its average throughput higher than the throughput of traffic density, and about 17.5% higher than that of nodal activity. Also it worth mentioning that



(a)



(b)



(c)

Figure 4.5: Throughput of the load density, traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

the throughput achieved for smaller network area 900mx1000m is higher than that achieved in this figure. This is because there is more potential for a route to break because nodes are given more space to move in freely.

Figure 4.5(b) shows the performance behaviour when the traffic rate is increased to 4 packets/s. The figure gives a clearer picture of the performance of load density compared to Figure 4.5(a) as the difference in the achieved throughput between load density and the other techniques is higher. Load density outperforms traffic density on average by 13%, while compared to nodal activity it is about 24%. The figure also demonstrates some steadiness in the throughput difference between the techniques while varying the speed. Although the throughput decreases as node speed increases, the throughput difference between the techniques is approximately the same. Similar results are depicted in Figure 4.5(c) when higher traffic rate of 8 packets/s is used.

End-to-end delay:

Figure 4.6 shows the end-to-end delay for load density, traffic density and nodal activity techniques under different traffic rates that are relatively, light, moderate and high traffic with rates of 2, 4 and 8 packets/s, respectively, when 75 nodes are used over a 900m×1000m topology area. Figure 4.6(a) illustrates the behaviour of the three techniques when the traffic rate is 2 packets/s. Although the techniques exhibit close performance in terms of end-to-end delay at very low mobility of a maximum speed of 1 m/s, the difference in performance between the techniques is clearly more noticeable at higher speeds. The figure shows that while traffic density outperforms nodal activity especially for higher speeds, load density outperforms the other two techniques. The end-to-end delay in load density is about 24 ms lower than in traffic density and 65 ms lower than in nodal activity, which represents a 15% and 32% reduction, respectively. Furthermore, the impact of increasing the speed affects the three methods, but it is slightly more noticeable for nodal activity and traffic density. Increasing the speed from 1m/s to 20m/s results in an increase of 100 ms in the end-to-end delay for nodal activity, 80 ms for traffic density, and 50 ms for load density. This is because the number of broken routes increases as the mobility speed increases which require a route discovery initiation to establish a new route. Since the route discovery latency is included in the end-to-end delay then the end-to-end delay increases.

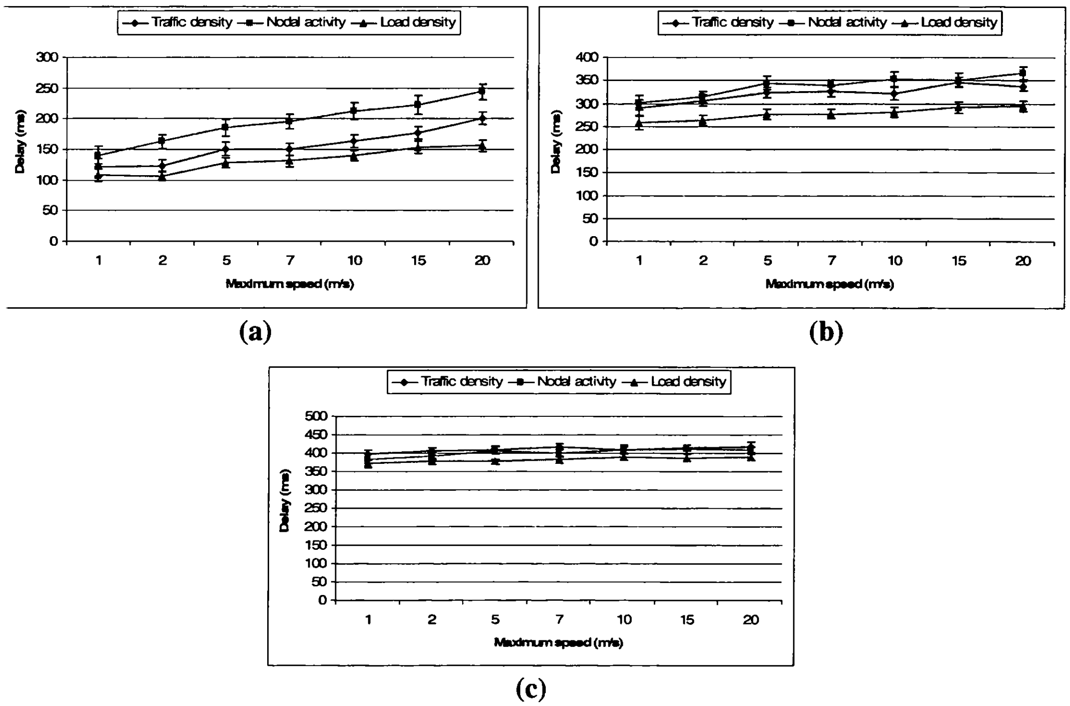


Figure 4.6: End-to-end delay of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 4.6(b) shows performance results for the three techniques when the traffic rate is increased to 4 packets/s. While traffic density and nodal activity exhibit close performance and the difference between the two is negligible, load density outperforms the two techniques by 14% and 18%, respectively, from traffic density and nodal activity. The speed variation has not shown significant effect on the end-to-end delay except that it increases slightly as mobility speed increases. This is because increasing the traffic rate causes congestion in the network (more than 30% of the packets has been dropped) which reduces the impact of higher speeds, leading to an increase in the end-to-end delay compared to Figure 4.6(a). A similar behaviour is shown in Figure 4.6(c) when higher traffic rate of 8 packets/s is used.

Figure 4.7 shows the behaviour of load density compared to traffic density and nodal activity when the traffic rate is 2, 4 and 8 packets/s, respectively, and 100 nodes are used over a

1200m×1000m topology area. It is noticeable when comparing the performance in Figures 4.7(a, b, c) against that in Figures 4.6(a, b, c) where smaller network size of 75 nodes over a 900m×1000m topology area is used that the minimum end-to-end delay is higher than the maximum end-to-end delay of the smaller network size although the same traffic scenarios are used. This observed behaviour is due by two reasons. Firstly, nodes in the larger topology have more freedom in mobility as they have more space to move in, this would result in more broken routes. Secondly, the existence of longer paths would increase the delay especially when there is a broken links where there is a need to discover a new path.

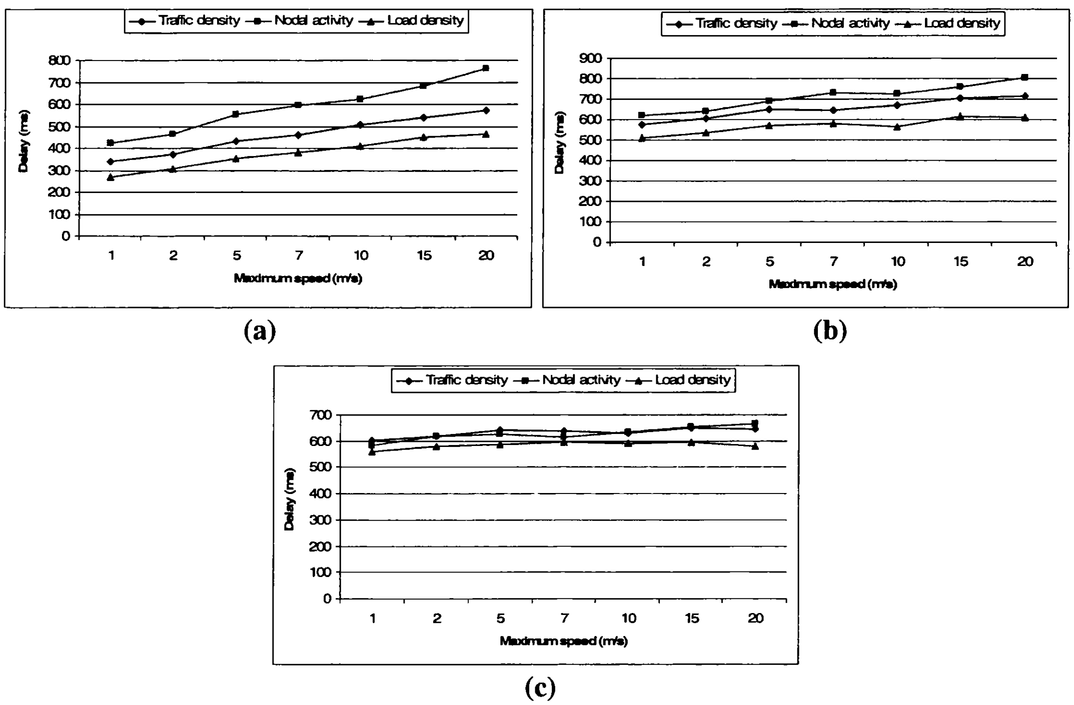


Figure 4.7: End-to-end delay of the load density, traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 4.7(a) shows clearly that load density outperforms both traffic density and nodal activity, with a difference of 85ms on average from traffic density and from nodal activity 210ms. The impact of increasing the mobility speed has a great effect on all techniques. However the effect is more noticeable on nodal activity. For load density, the increase in the end-to-end delay is around 190 ms while it is around 230 ms and 340ms for traffic density and nodal activity, respectively.

Similarly in Figure 4.7(b) when the traffic rate is increased to 4 packets/s, load density has better performance than the other two techniques by exhibiting lower end-to-end delay values for the simulated mobility speeds. While traffic density has a lower end-to-end delay compared to nodal activity, with a difference ranging between 40-85ms, load density has a lower end-to-end delay with a difference ranging between 70-100ms compared to traffic density and hence better performance than both. On the other hand, the impact of speed on the three techniques when the traffic rate is increased is less than that in Figure 4.7(a) when lower traffic rate is used. For load density, the increase in the end-to-end delay due to the increase in speed is around 100 ms (about 20% of its minimum end-to-end delay), while for traffic density, the increase in the end-to-end delay due to the increase in speed is 140 ms (about 25% of its minimum end-to-end delay), and for nodal activity the increase is 180ms (about 29% of its minimum achieved end-to-end delay).

In contrast, when increasing the traffic rate to 8 packets/s, the three techniques exhibit close performance and sometime similar performance in the case of nodal activity and traffic density as it is shown in Figure 4.7(c). The results also reveal that in this case the impact of speed is low. This is potentially because significant amount of packets are dropped. The calculation of the end-to-end delay does not consider dropped packets. Only 28% and less of sent packets have been successfully delivered. This also justifies why the maximum end-to-end delay in this figure is less than that in Figures 4.7(a) and 4.7(b). Nevertheless we should bear in mind that even if the load density has shown close end-to-end delay to the other techniques in some of the simulated speeds, it has achieved lower end-to-end delay especially at higher speeds of 15 m/s and 20 m/s, and also it has the upper hand in terms of throughput under the same scenarios.

Routing overhead:

Figure 4.8 shows the generated routing overhead by load density, traffic density and nodal activity techniques when the traffic rate is 2, 4 and 8 packets/s, respectively, and when 75 nodes are used over a 900m×1000m topology area. Figure 4.8(a) reveals clear advantage for load density against traffic density and nodal activity. This is because load density does not rely in its operation on exchanging neighbourhood information as it is in traffic density and nodal activity (e.g. number of active flows) since such exchange is optional and thus reduces the routing overhead. For instance at the maximum speed of 1 m/s load density generates only half of the routing overhead generated by either traffic density or nodal activity. The figure also reveals that when the traffic rate is 2 packets/s, the routing overhead increases noticeably as the mobility speed increases. The overhead increases by around 40,000 routing packets at the speed of 20 m/s compared to that at 1 m/s for the three methods. This is because the occurrence of broken routes becomes more frequent as the mobility speed increases, which triggers route maintenance and route discovery sessions more often. This, in turn, causes an increase in the routing overhead.

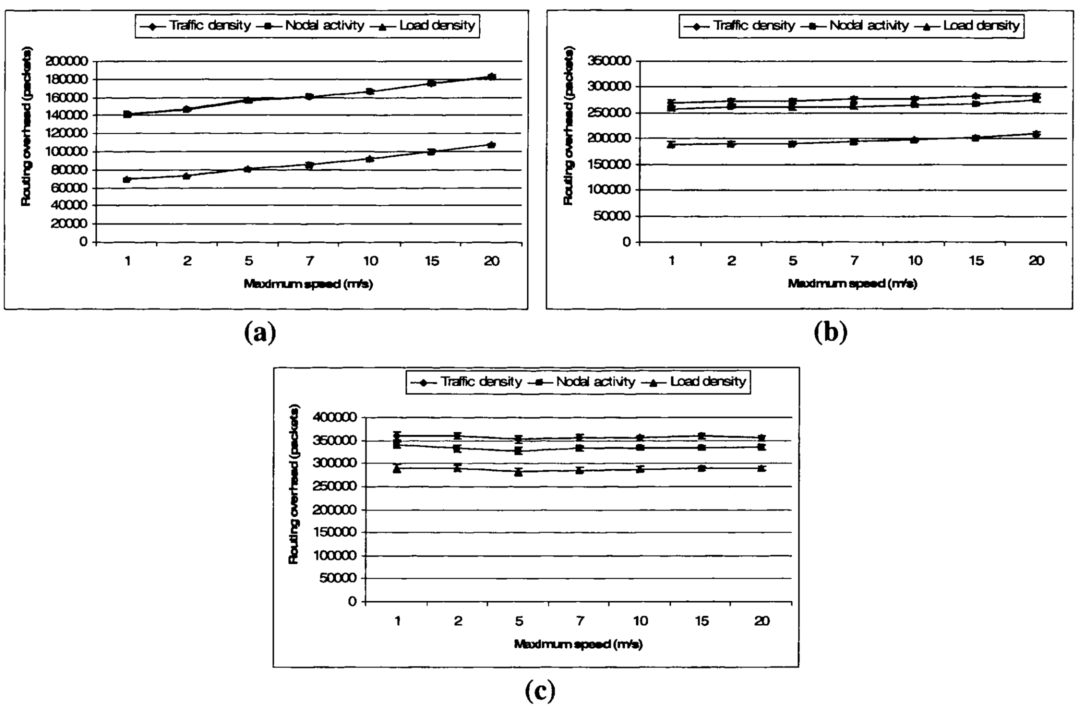
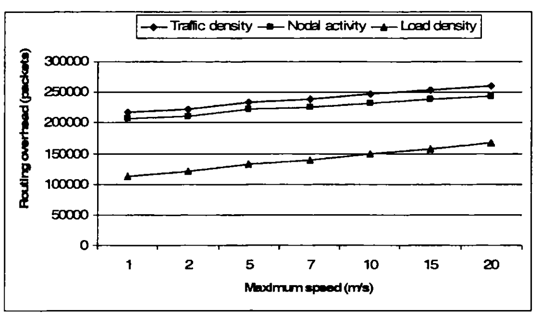


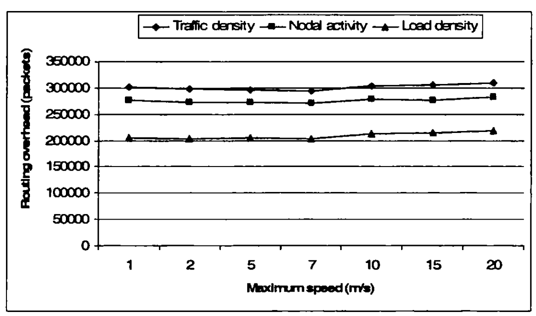
Figure 4.8: Routing overhead of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 4.8(b) demonstrates that the generated overhead from the three techniques for the relatively higher traffic rate of 4 packets/s. The results show that load density still outperforms nodal activity and traffic density by 68,000 and 80,000 routing packets, respectively. Nodal activity generates less overhead than traffic density. On the other hand, the impact of increasing the mobility speed on the generated overhead reaches the limit of 21,000 routing packets at the speed of 20 m/s compared to that at 1 m/s for load density. Similar results in Figure 4.8(c) are depicted when the traffic rate is increased to 8 packets/s, where load density outperforms both nodal activity and traffic density by 46,000 and 69,000 routing packets, respectively.

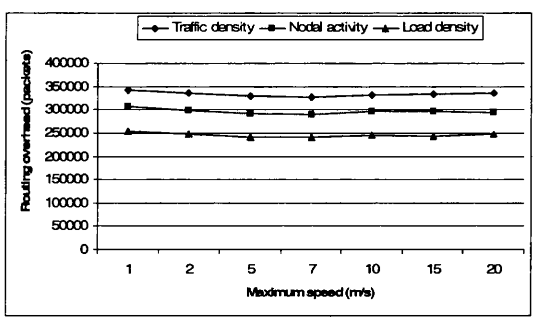
Figure 4.9 shows the routing overhead generated by load density, traffic density and nodal activity for the traffic rates of 2, 4 and 8 packets/s, respectively, and when 100 nodes are used over a 1200m×1000m topology area. Figure 4.9(a) shows that load density clearly outperforms the other techniques as it generates less overhead with at least a difference of 80,000 routing packets. Figure 4.9(b) depicts traffic density with a highest overhead followed by nodal activity and then by load density. A clear advantage for load density is also depicted in the figure as it generates about 67,000 routing packets less than that for nodal activity and about 92,000 routing packets less than that for traffic density. Similarly in Figure 4.9(c) at traffic rate 8 packets/s, load density outperforms both traffic density and nodal activity.



(a)



(b)



(c)

Figure 4.9: Routing overhead of the load density, traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Normalised load covariance:

Figure 4.10 demonstrates the normalised load covariance for load density, nodal activity and traffic density for traffic rates of 2, 4 and 8 packets/s, and when 75 nodes are used over a 900m×1000m topology area. The figure reveals a clear performance advantage in favour of load density over the other two techniques, where as nodal activity and traffic density show comparable performance with a negligible difference in favour of traffic density. For instance in Figure 4.10(b) at traffic rate of 4 packets/s load density outperforms nodal activity by 10%, and outperforms traffic density by 8%. The load density technique achieves better normalised load covariance because it considers in its decision of selecting routes the history of the load transiting through the nodes.

Figure 4.11 demonstrates the normalised load covariance for the three techniques when the traffic is 2, 4 and 8 packets/s, and when 100 nodes are used over a 1200m×1000m topology

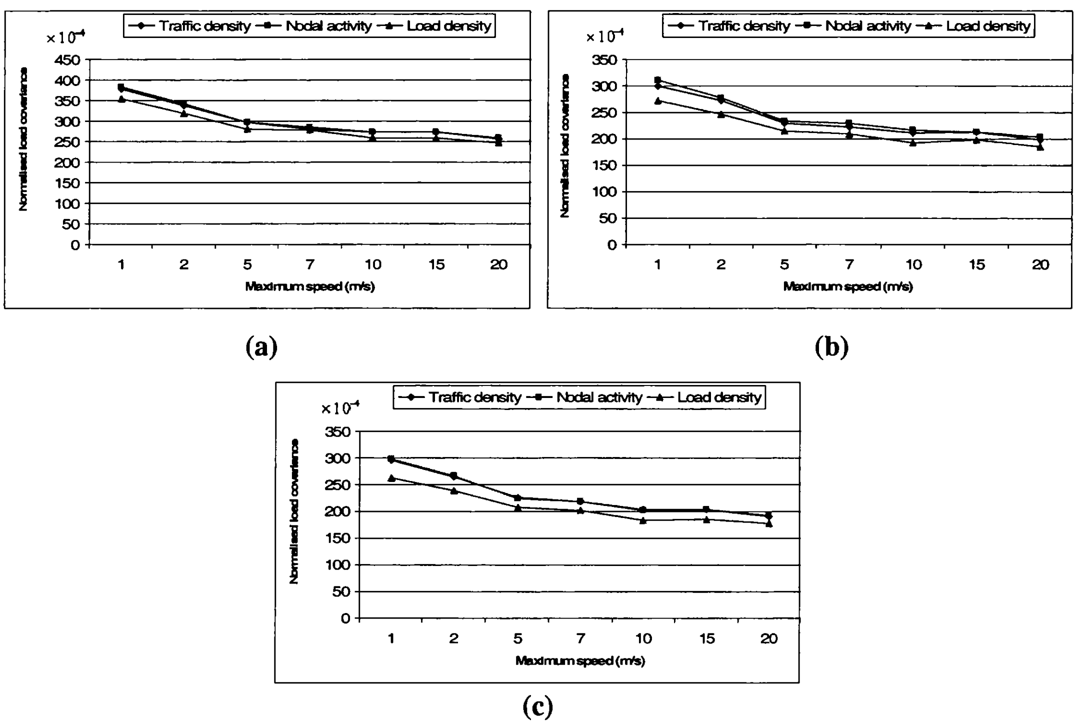
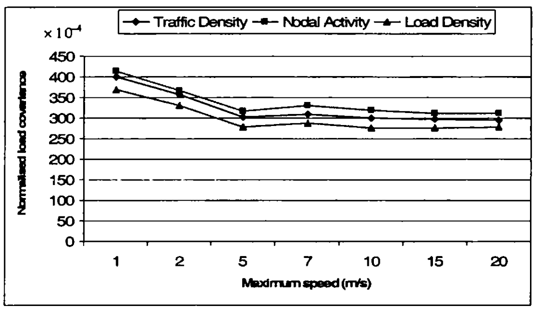
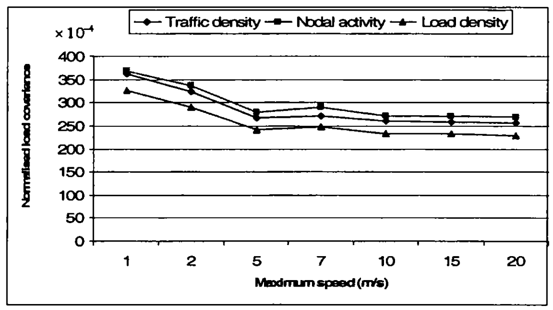


Figure 4.10: Normalised load covariance of of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

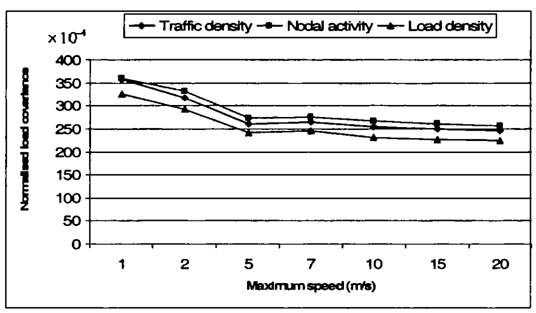
area. The performance difference between nodal activity and traffic density is clearer compared to the smaller size network in Figure 4.10. The figure reveals that traffic density outperforms nodal activity with an average difference around 4% at traffic rate of 4 packets/s. The figure also reveals that load density still outperforms the other techniques, for instance at traffic rate of 4 packets/s the difference is around 14% from nodal activity and from traffic density the difference is around 10%.



(a)



(b)



(c)

Figure 4.11: Normalised load covariance of the load density, traffic density and nodal activity techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

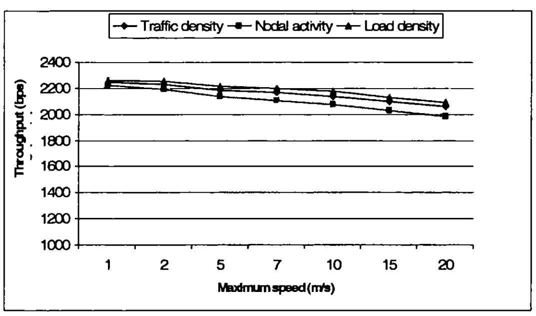
4.4.2 Performance impact of packet length

This section demonstrates the effect of varying the packet length on the performance of the traffic aware techniques. The parameters used in the previous Section 4.4.1 for the network size of 75 nodes with simulation area 900m×1000m are recalled here. The same mobility scenarios are used again and the same traffic scenarios are used except the change regarding the packet size that is set to 128 bytes.

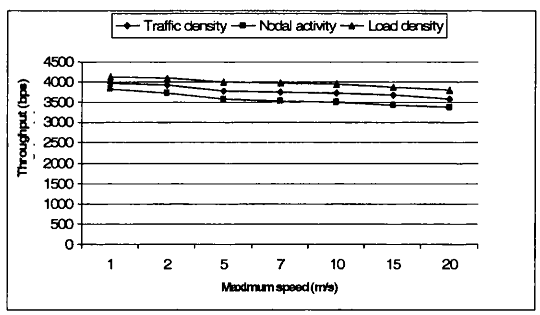
Throughput:

Figure 4.12 depicts the performance of the three techniques in terms of throughput for the traffic rates of 2, 4 and 8 packets/s, respectively. Figure 4.12(a) shows that when the traffic rate is 2 packets/s there is a performance advantage for load density over traffic density and nodal activity especially at higher mobility speeds. However, the difference between the three techniques is small. Load density has on average 1.2% of its maximum achieved throughput (2260 bps) higher than traffic density and an average of 3.6% higher than nodal activity. Similar to the results in Figure 4.4(a), where packet length is 512-byte and mobility speed is 1m/s, load density, nodal activity, and traffic density exhibit comparable performance.

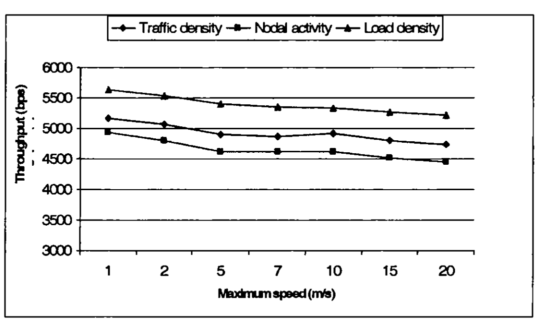
Figure 4.12(b) shows the behaviour of the two techniques under the relatively medium traffic rate of 4 packets/s. The difference in performance between the three techniques is clearer compared to Figure 4.12(a). The figure shows that load density outperforms traffic density and nodal activity over different mobility speeds, with a difference of 200 bps better than traffic density and difference of 400bps from nodal activity, which respectively corresponds to about



(a)



(b)



(c)

Figure 4.12: Throughput of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

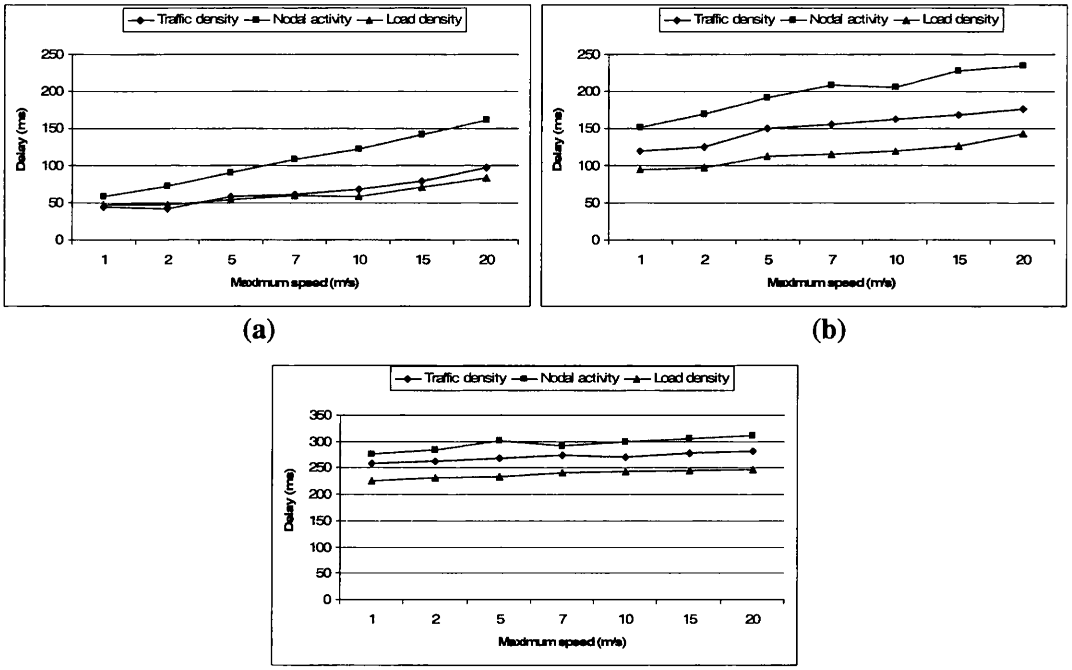
5% and 10% of the maximum achieved throughput by load density (that is 4140 bps). While under the same scenario in Figure 4.4(b) but using 512-byte packet, the difference is about 7% from traffic density and from nodal activity is about 10% of the maximum throughput achieved by load density (that is 11680 bps). Although the throughput decreases while nodes' speed increases the difference in throughput between the techniques remains almost the same.

Figure 4.12(c) which depicts the throughput when traffic rate is 8 packets/s shows load density outperforms both traffic density and nodal activity. The difference between load density and traffic density is about 9% and between load density and nodal activity is about 14% of the maximum throughput achieved (5630 bps), where as under the same scenario but using 512-byte packet, in Figure 4.4(c), the difference is about 8% from traffic density and 9% from nodal activity of the maximum achieved throughput which is 13970 bps.

End-to-end delay:

Figure 4.13 presents the performance behaviour of load density, traffic density and nodal activity techniques in terms of the end-to-end delay for a traffic rate of 2, 4 and 8 packets/s, respectively. Figure 4.13(a) demonstrates the end-to-end delay when traffic rate is 2 packets/s. As the figure illustrates load density shows close performance with traffic density in general. However, both load density and traffic density outperforms nodal activity over different speeds. The three techniques exhibit close performance at very low mobility of maximum speed of 1 m/s. The increase in speed affects the three techniques. However it is slightly more apparent for nodal activity. On the other hand when comparing the performance of the three techniques in this figure to that in Figure 4.6(a) where 512-byte packet is used, it is noticeable that when 128-byte packet is used the minimum end-to-end delay is less than 50% of the minimum end-to-end when 512-byte packet is used.

Figure 4.13(b) depicts the end-to-end delay of the three techniques when the traffic rate is increased to 4 packets/s. The figure illustrates that the load density outperforms both traffic density and nodal activity with an end-to-end delay ranges between 95-145 ms over the



(a) (b)

(c)

Figure 4.13: End-to-end delay of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

simulated speeds from low mobility with a maximum speed of 1 m/s to a high mobility with a maximum speed of 20 m/s. The figure also demonstrates that the maximum end-to-end delay of load density is less that the minimum end-to-end delay of nodal activity, and less than the end-to-end delay of traffic density at the maximum speed of 5 m/s.

Figure 4.13(c) depicts the end-to-end delay of the three techniques for a traffic rate of 8 packets/s. The increase in speed affects all methods, but the impact of increasing the speed from 1 m/s to 20 m/s has shown small increase in the end-to-end delay compared to the maximum end-to-end-delay of each technique. Nonetheless, load density exhibits better performance over the simulated speeds.

To summarise, when comparing the end-to-end delay in Figures 4.6 and 4.13 it is noticeable that the minimum and maximum end-to-end delays significantly increase when using larger packet size. It is clear that the minimum and maximum end-to-end delays in Figure 4.13(a) are less than that in Figure 4.6(a) where 512-byte packets are used. A similar behaviour is depicted when comparing Figure 4.6(b) with Figure 4.13(b) and Figure 4.6(c) with Figure 4.13(c), but with one addition that is both the minimum and the maximum end-to-end delays in Figures 4.10(b and c) are less than the minimum end-to-end delay values in Figures 4.6(b and c).

Routing overhead:

Figure 4.14 shows the performance of the three techniques in terms of routing overhead for traffic rates of 2, 4 and 8 packets/s. Figure 4.14(a) demonstrates the generated overhead for traffic rate of 2 packets/s. The figure reveals that load density outperforms both nodal activity and traffic over the simulated speeds. Load density generates on average about half of the routing overhead generated by either nodal activity or traffic density.

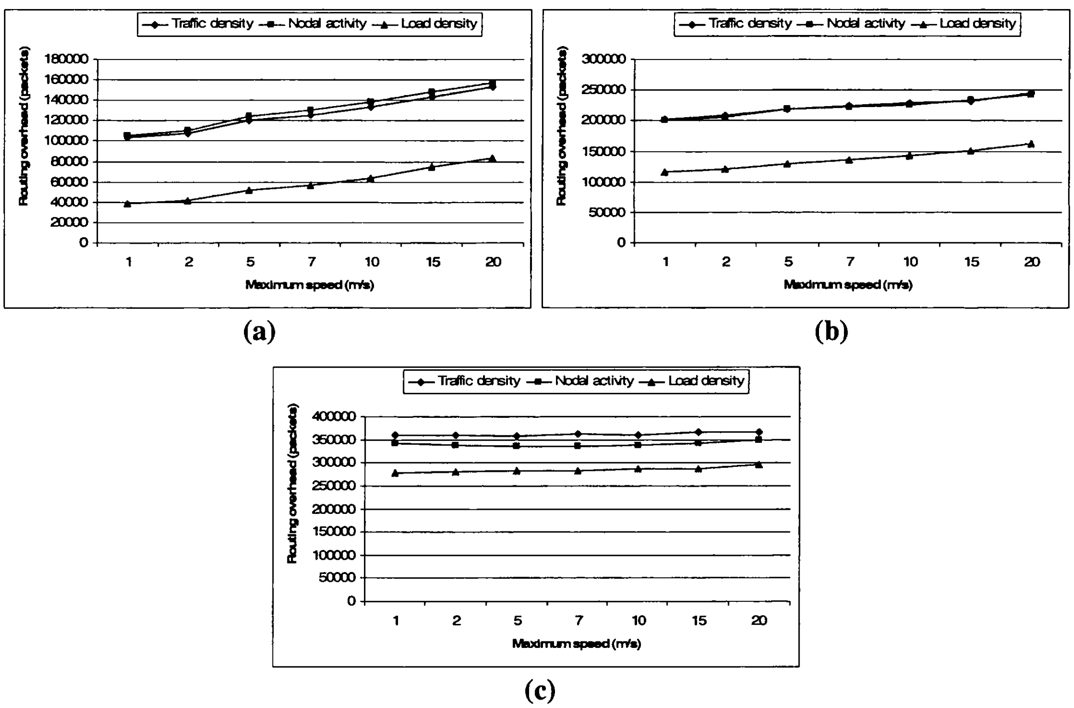
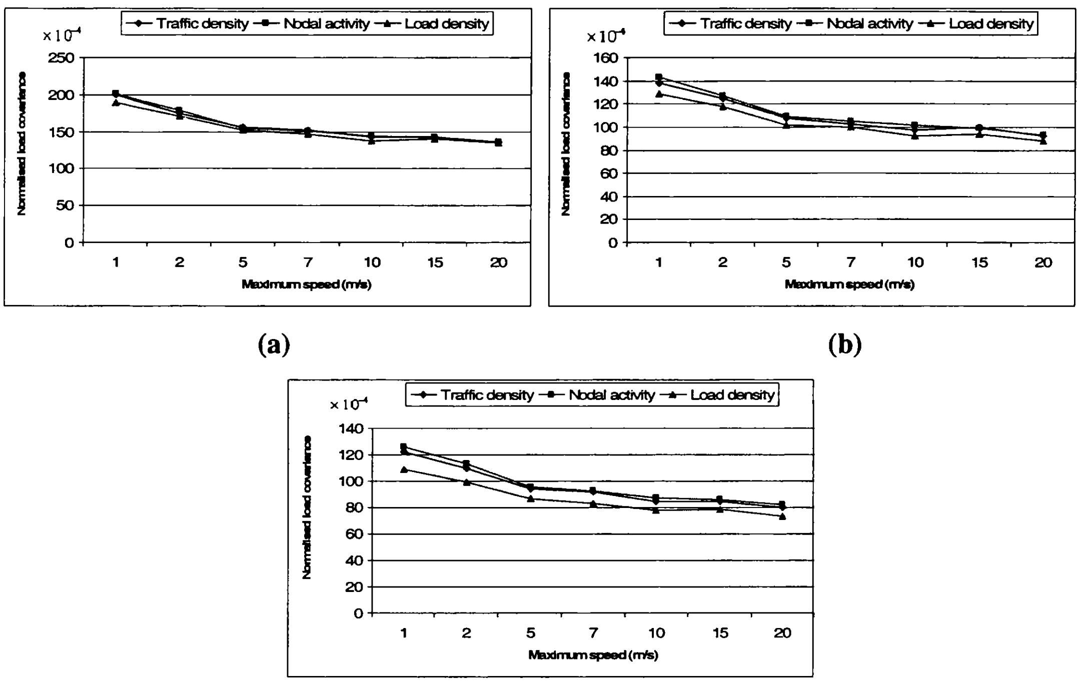


Figure 4.14: Routing overhead of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 4.14(b) demonstrates the generated overhead by the three techniques when the traffic rate increased to 4 packets/s. The figure shows that load density outperforms the other techniques with about 80,000 routing packets less than either nodal activity or traffic density over various speeds. Similar results are also depicted in Figure 4.14(c) when the traffic rate is increased to 8 packets/s. However nodal activity generates less overhead compared to that generated by traffic density. On the other hand, the generated overhead by the three techniques is not affected significantly by the variation in the mobility speed. This is because the routing overhead has already reached a high level in an effort by the routing algorithm to handle the heavy traffic that creates congested condition (more than 40% of the packets has been dropped). This helps to reduce the impact of speed compared to Figures 4.14(a and b).



(a) (b)

(c)

Figure 4.15: Normalised load covariance of the load density, traffic density and nodal activity techniques for a network of 75 nodes over topology area 900m×1000m with a 128-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Normalised load covariance:

Figure 4.15 demonstrates the normalised load covariance for load density, nodal activity and traffic density for traffic rates of 2, 4 and 8 packets/s. The three techniques show comparable performance at traffic rate of 2 packets/s as it is demonstrated in Figure 4.15(a). However when the traffic rate is increased to 4 packets/s, in Figure 4.15(b), the performance difference becomes clearer. The figure reveals that load density outperforms both nodal activity and traffic density by 7% and 5%, respectively, where as nodal activity and traffic density show comparable performance. Similarly, when the traffic rate is increased to 8 packets/s in Figure 4.15(c) load density outperforms nodal activity by 11% and outperforms traffic density by 9%.

4.4.3 Performance impact of increasing the number of flows

The analysis is conducted using a network size of 75 nodes forming a MANET over a simulation area of 900m×1000m. We have taken only one of the network sizes used in previous section to keep the time spent on running the experiments manageable. The node maximum speed has been set to 1, 2, 5, 7, 10, 15, and 20 m/s. Traffic rate has been varied between 2, 4, and 8 packets/s, where 512-byte packets have been used. The number of flows has been set to 10, 20, and 30 flows. For each of the settings we monitor the performance of 5 flows that are formed according to the network model in Chapter 3 (see Section 3.2) where the source of traffic and its destination are stationary, while the rest flows are considered as background traffic. For background traffic flows, the sources and destinations are not stationary and are free to move in any direction and speed. Table 4.3 illustrates the simulation parameters used in the experiments conducted in this section. In the figures that are presented below, the x-axis

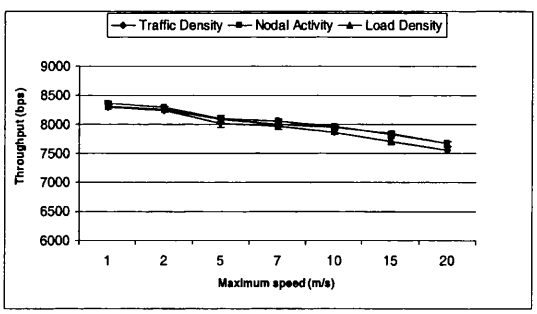
TABLE 4.3: THE SYSTEM PARAMETERS USED IN THE SIMULATION EXPERIMENTS FOR BACKGROUND TRAFFIC

Parameter	Values
Number of nodes	75
MAC layer	IEEE 802.11
Transmission range	250m
Simulation area	900m×1000m
Simulation time	900s
Mobility model	Random waypoint model
Maximum speed	1, 2, 5, 7, 10, 15 and 20m/s.
Pause time	0s
Traffic type	CBR
Packet size	512 bytes
Packet rate	2, 4, 8 packets/s
Number of flows	10, 20, 30 flows
Number of runs per data point	30

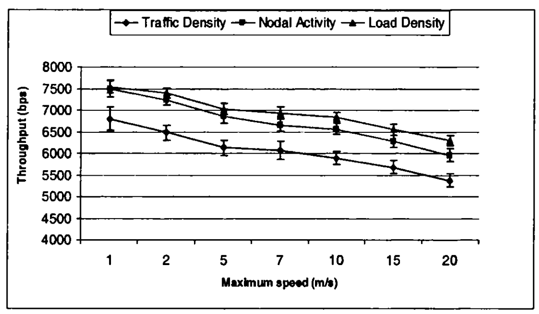
describes the variations in node's maximum mobility speed, while the y-axis represents results of the performance metric of interest. The results have been conducted from steady state simulations. A 95% confidence interval is calculated for each point.

Throughput:

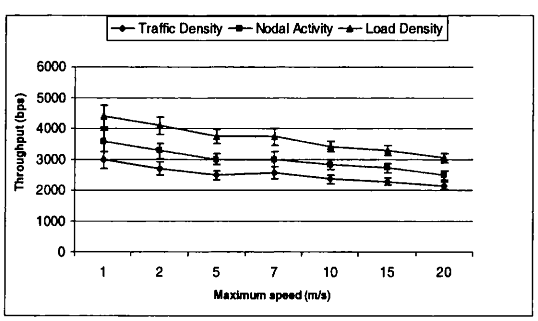
Figure 4.16 demonstrates the achieved throughput for load density, nodal activity and traffic density when traffic rate is 2 packets/s, and when 10, 20 and 30 flows are used. In Figure 4.16(a), when 10 flows are used, the techniques show close performance as most of the injected traffic has been delivered. However, when the number of flows is increased to 20 flows, in Figure 4.16(b), load density and nodal activity show comparable performance with a small difference of 220 bps on average, that is 3%, to the favour of load density. In contrast, load density outperforms traffic density with a difference around 880bps, on average over the simulated speeds, that is approximately 13% of its average throughput (6950bps), while nodal activity outperforms traffic density with a difference of 660bps that is about 10% of its average



(a)



(b)



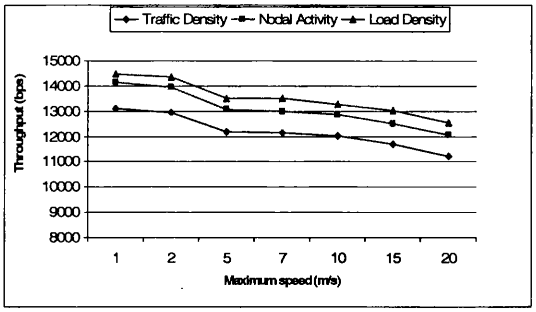
(c)

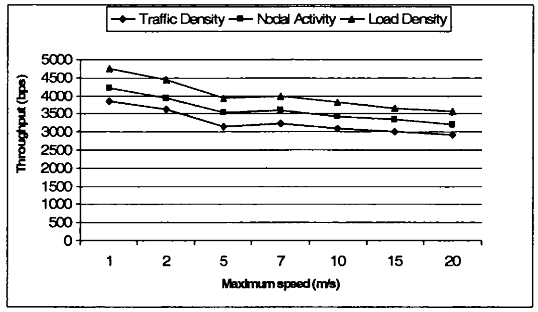
Figure 4.16: Throughput of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

throughput (which is 6720bps). Similar behaviour is depicted in Figure 4.16(c) when 30 flows are used, but the difference in favour of load density has increased to 1170bps from traffic density, and from nodal activity the difference has increased to 690bps, which is respectively about 32% and 19% of its average throughput (3670bps).

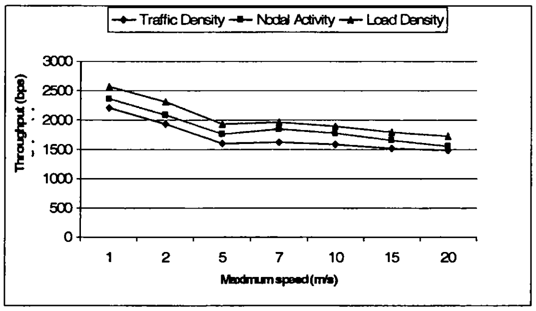
The impact of changing the maximum speed from 1m/s to 20m/s has affected all techniques. For instance in Figure 4.16(a), for load density the decrease in throughput due to increasing mobility speed reaches the limit of 640bps; that is about 7.7% of its maximum achieved throughput (which is 8300bps). In contrast, for traffic density the decrease in throughput is around 730bps that is about 9% of its maximum throughput, and for nodal activity the decrease is 700bps that is about 8.5% of its maximum throughput. The decrease in throughput is due to the frequent topology changes caused by mobility which results in more broken routes as speed increases.

Figure 4.17 shows the achieved throughput by the three techniques when the traffic rate is increased to 4 packets/s. Figure 4.17(a) shows that nodal activity outperforms traffic density by 890bps on average, which is approximately by 7%. However, it also shows that load density outperforms traffic density by 1340bps, and outperforms nodal activity by 450bps, which correspond respectively to about 10% and 3% of its average throughput (13530 bps). Similar behaviour is depicted in Figure 4.17(b) when increasing the number of flows to 20, but the performance difference is smaller than that in Figure 4.17(a), this because of the sharp drop in the number of delivered packets which has affected all techniques, only less than 28% of the packet has been delivered. The drop in the delivery ratio is because increasing the number of flows to 20 would mean that 40 nodes (which is more than half of the network size); 20 sources and 20 destinations, are trying to establish a communication with each other. Nonetheless, load density outperforms traffic density by 760bps and outperforms nodal activity by 410bps, which correspond respectively to around 18% and 10% of its average throughput (4020 bps). Similar behaviour is depicted in Figure 4.17(c) when the number of flows is increased to 30.



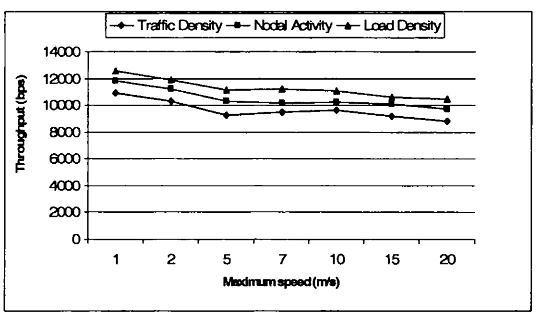


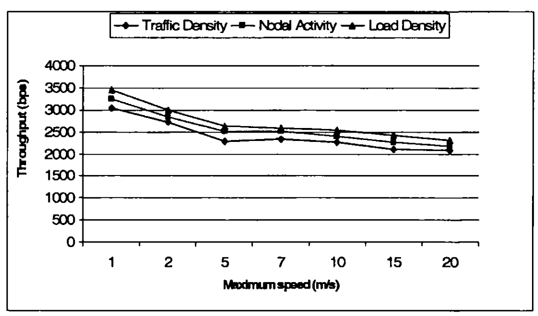
(a) (b)



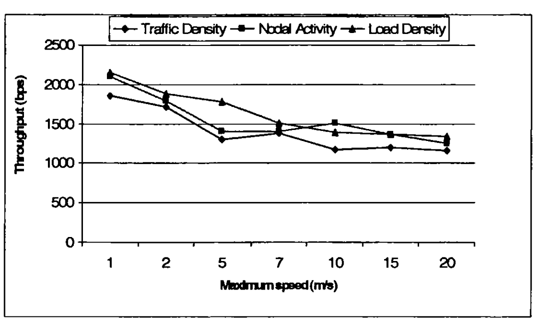
(c)

Figure 4.17: Throughput of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows





(a) (b)



(c)

Figure 4.18: Throughput of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 4.18 demonstrates the achieved throughput when the traffic rate is increased to 8 packets/s. The increase in traffic rate has shown great impact on the techniques as the throughput drops compared to lower traffic rates, when 2 packets/s and 4packets/s in Figures 4.16 and 4.17, since this has increased the congestion in the network. The impact gets more sever as we increase the number of flows as it is shown in Figures 4.18(b, c). For instance, with 30 flows, only 6% and less of the packets have been delivered to their destinations.

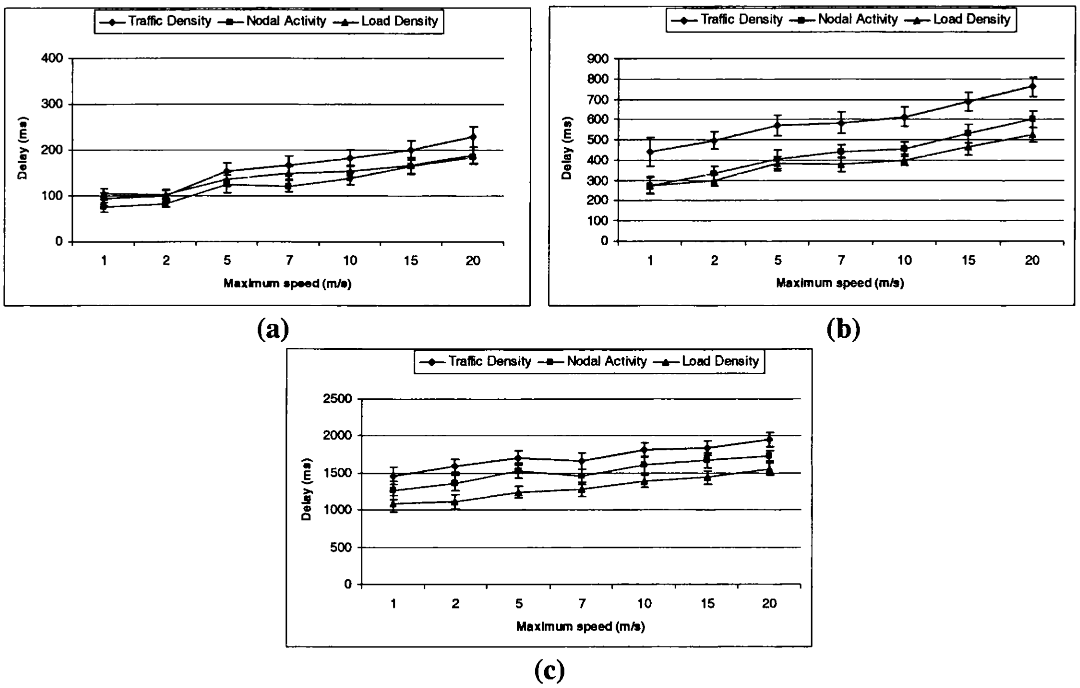
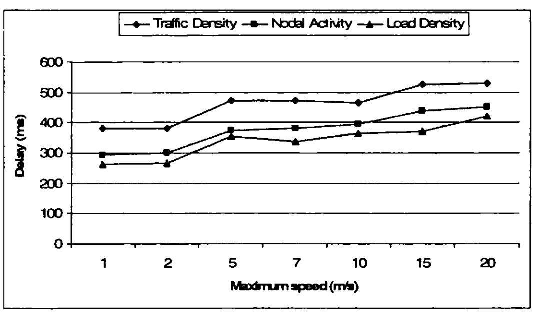


Figure 4.19: End-to-end delay of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

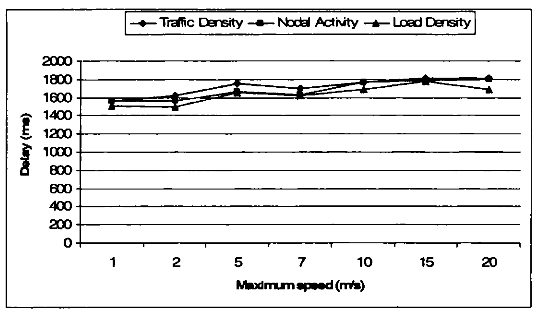
End-to-end delay

Figure 4.19 shows the end-to-end delay for load density, nodal activity and traffic density when traffic rate is set to 2 packets/s, for 10, 20 and 30 flows. In Figure 4.19(a) when 10 flows are used the techniques show comparable results. However, when increasing the number of flows to 20 in Figure 4.19(b), load density technique outperforms traffic density and nodal activity with a difference reaches the limit of 200ms from traffic density and from nodal activity the difference reaches 45ms; on average over the simulated speeds. In contrast, in Figure 4.19(c), as we inject more traffic to the network, when 30 flows are used, the network becomes congested, only around half of the injected traffic and less has been delivered to destination, and as the figure shows, a sharp increase in the end-to-end delay compared to that in Figure 4.19(b). Nonetheless, the figure shows load density outperforms traffic density and nodal activity with an average difference of 420ms and 220ms, respectively

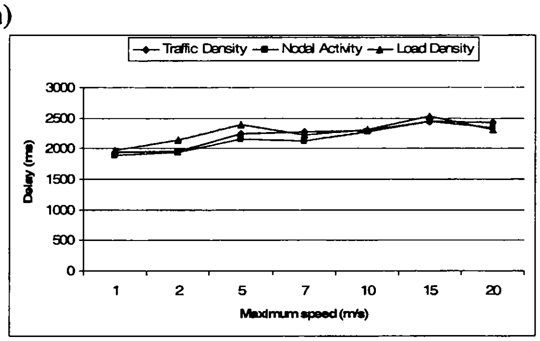
Figure 4.20 shows the performance for the three techniques, when the traffic rate is increased to 4 packets/s. Figure 4.20(a) demonstrates that the load density outperforms traffic density and nodal activity with an average difference of 120ms and 40ms, respectively. In contrast, nodal activity outperforms traffic density with on average 80ms. Figure 4.20(b) shows the end-to-end delay when the number of flows is increased to 20. As the figure demonstrates there is sharp increase in the delay for all techniques, this because of the amount of traffic injected into the



(a)



(b)



(c)

Figure 4.20: End-to-end delay of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

network, which has affected both delay and throughput for all techniques, where less than 28% of the packets has been delivered to the destination. Nonetheless, the techniques show comparable performance. Similarly in Figure 4.20(c) as we increase the number of flows to 30, the techniques show comparable performance.

Figure 4.21 demonstrates the end-to-end delay by the techniques when the traffic rate is increased to 8 packets/s. Figure 4.21(a) demonstrates that when 10 flows are used, load density and nodal activity show comparable performance while both of them show slightly better performance than traffic density. However, as the number of flows is increased to 20 and 30, in Figures 4.21(b, c), traffic density and nodal activity show comparable performance. In contrast traffic density and nodal activity outperform load density, with a difference reached the limit of 130ms in Figure 4.21(c). Nonetheless we should bear in mind that load density has achieved higher throughput, and hence dropped less packets compared to traffic density and nodal activity, which could justify the higher end-to-end delay.

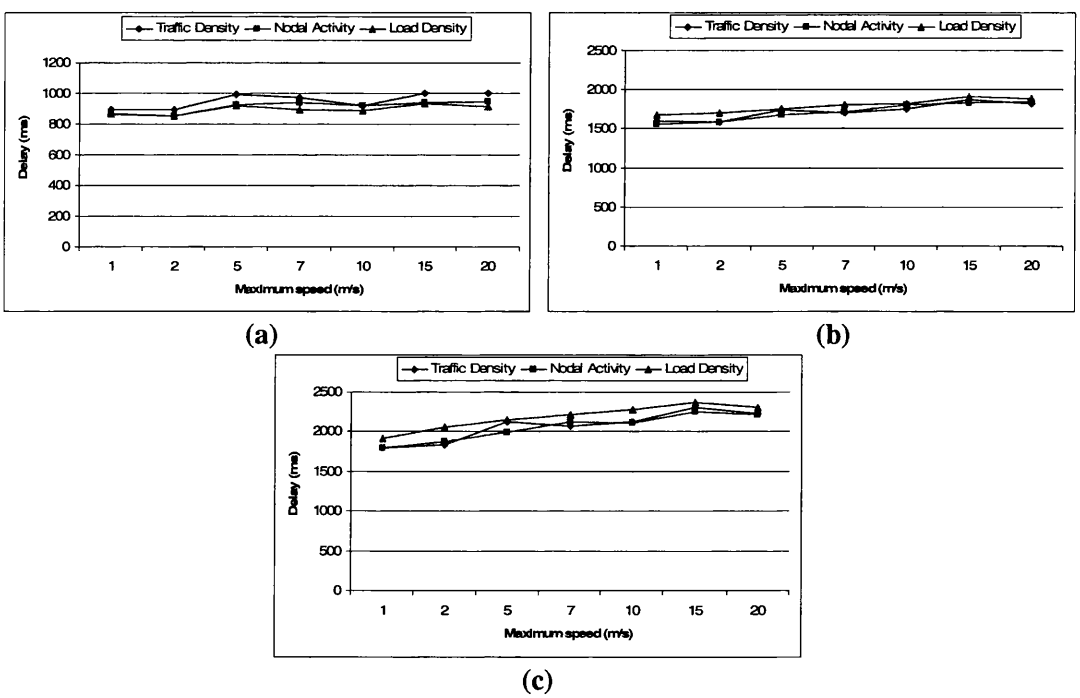
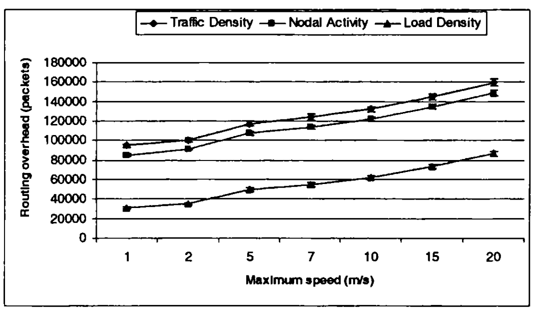


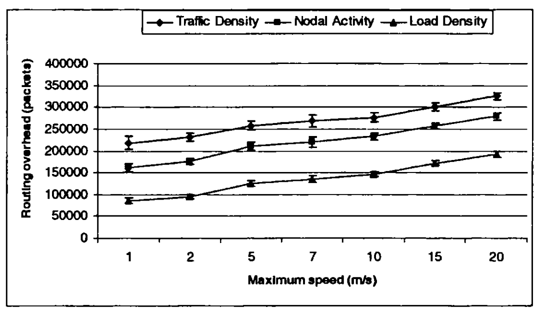
Figure 4.21: End-to-end delay of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Routing overhead:

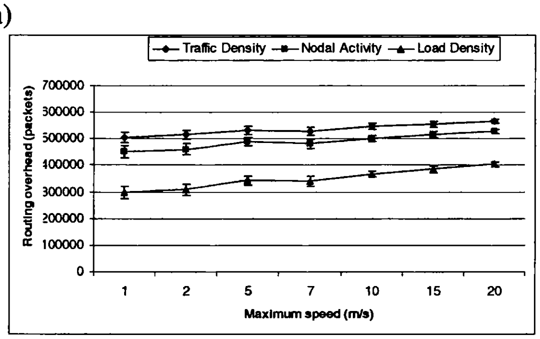
Figure 4.22 shows the generated routing overhead by load density, traffic density and nodal activity when the traffic rate is 2packets/s, and when number of flows is 10, 20 and 30 flows. Figure 4.22(a) shows comparable routing overhead generated by traffic density and nodal activity with an average of around 125,000 and 115,000 routing packets, respectively, while the load density technique generates on average about 58,000 routing packets. In Figure 4.22(b) when the number of flows is increased to 20 flows. Load density generates on average about 136,000 routing packets, traffic density generates on average around 268,000 routing packets, and nodal activity generates on average around 220,000 routing packets. Similar behaviour is depicted when increasing the number of flows to 30 flows, in Figure 4.22(c). Load density generates on average about 350,000 routing packets, while traffic density generates on average around 535,000 routing packets, and nodal activity generates around 490,000 routing packets.



(a)



(b)



(c)

Figure 4.22: Routing overhead of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 4.23 shows the routing overhead when the traffic rate is increased to 4packets/s, and when number of flows is set to 10, 20 and 30 flows. In Figure 4.23(a) when 10 flows are used, load density generates on average about 114,000 routing packets, traffic density generates on average about 216,000 routing packets, and nodal activity generates around 192,000 routing packets, and hence load density technique outperforms traffic density and nodal activity with a difference of about 100,000 and 78,000 routing packets, respectively. Similar results are depicted in Figures 4.20(b, c) under 20 and 30 flows. In contrast increasing the mobility has shown little impact on the performance of the techniques as we increase the number of flows, the impact of increasing the traffic has overcome the impact of the speed.

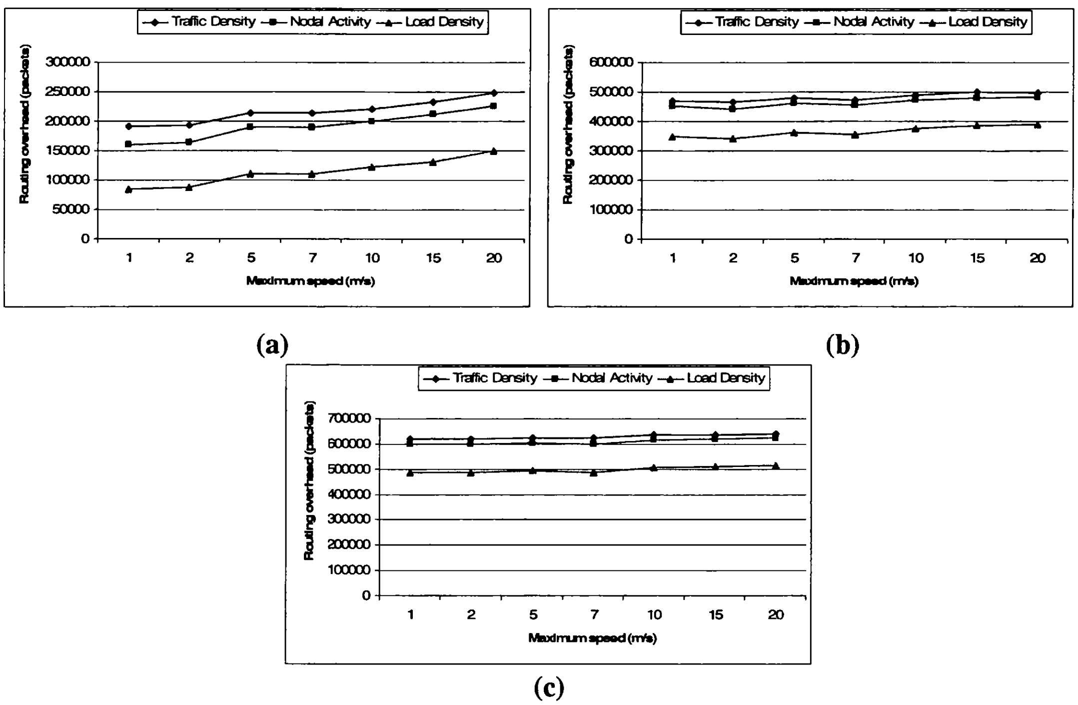
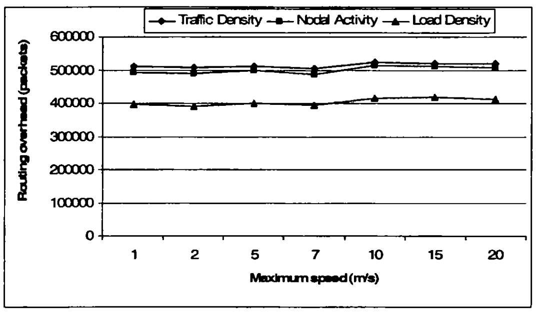
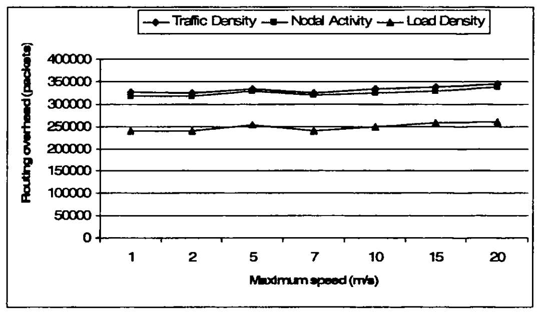


Figure 4.23: Routing overhead of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

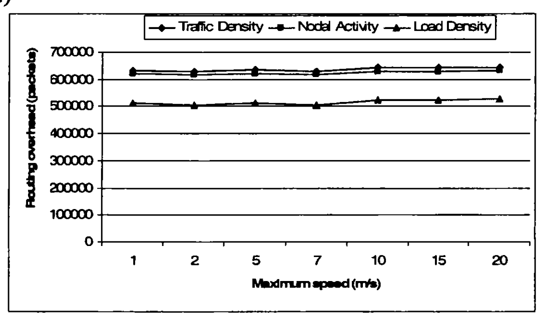
Figure 4.24 shows the routing overhead when the traffic rate is increased to 8 packets/s, and when number of flows is 10, 20 and 30 flows. In Figure 4.20(a) when 10 flows are used, load density generates on average about 83,000 routing packets less than that generated by traffic density, and about 77,000 routing packets less than that generated by nodal activity. In contrast in Figures 4.8(b, c) under 20 and 30 flows, respectively, load density outperforms traffic density, on average, by around 110,000 and 120,000 routing packets, respectively, and it outperforms nodal activity by around 94,000 routing packets under 20 flows and by 107,000 routing packets under 30 flows.



(a)



(b)



(c)

Figure 4.24: Routing overhead of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Normalised load covariance:

Figure 4.25 demonstrates the normalised load covariance for load density, nodal activity and traffic density when traffic rate is 2 packets/s, and when 10, 20 and 30 flows are used. In Figure 4.25(a) when 10 flows are used, load density outperforms both nodal activity and traffic density by 11% and 6.5%, respectively. Similarly, when the number of flows is increased to 20, in Figure 4.25(b) load density outperforms nodal activity by 7% and outperforms traffic density by 10%. In contrast when the number of flows is increased to 30 flows, in Figure 4.25(c) the difference between the techniques becomes apparent as load density outperforms nodal activity by 16% and outperforms traffic density by 29%. Load density achieves a better normalised load covariance because it considers the history of the load transiting through the nodes in its decision of selecting routes.

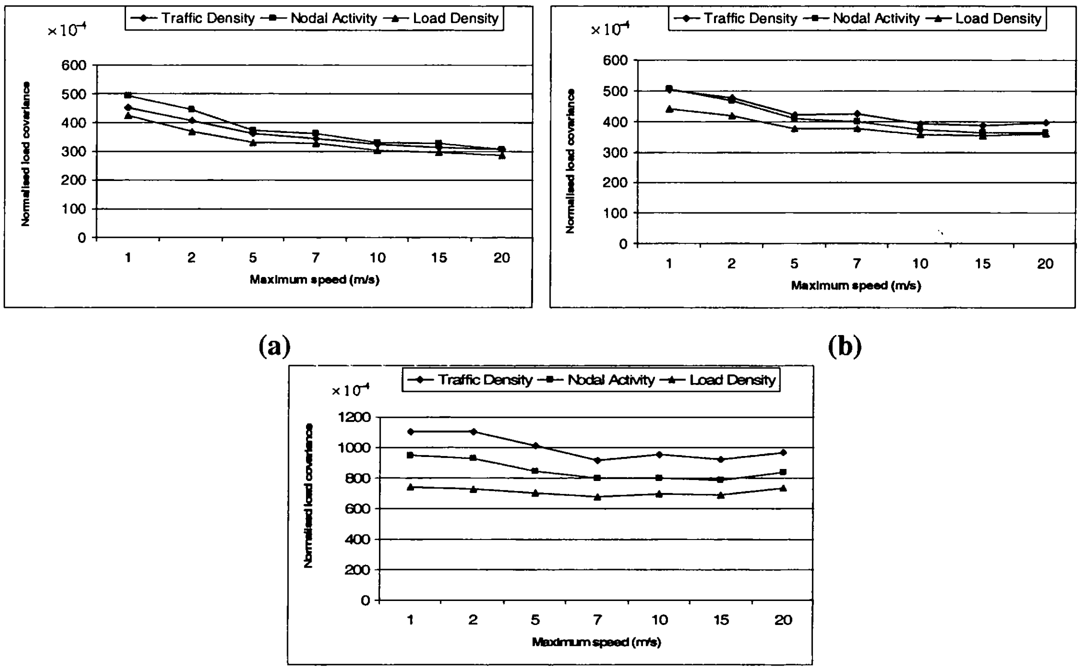


Figure 4.25: Normalised load covariance of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 4.26 depicts the performance of the three techniques when traffic rate is increased to 4 packets/s. The figure reveals that load density outperforms both nodal activity and traffic density regardless of the used number of flows. Figure 4.26(a) shows that load density outperforms both nodal activity and traffic density by 10% and 11%, respectively when 10 flows are used, where as nodal activity and traffic density show comparable performance. Where as when the number of flows is increased to 20, in Figure 4.26(b) load density outperforms nodal activity by 9% and outperforms traffic density by 17.5%. Similarly in Figure 4.26(c) when the number of flows is increased to 30 flows, load density outperforms nodal activity by 8% and outperforms traffic density by 20%.

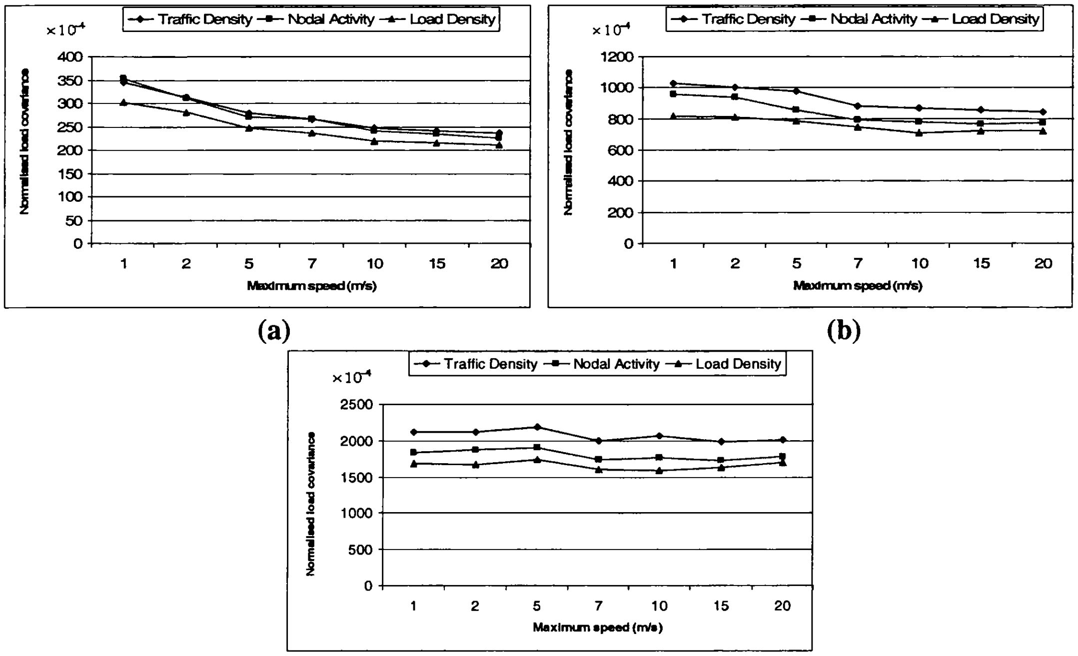


Figure 4.26: Normalised load covariance of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 4.27 shows the normalised load covariance for the techniques when traffic rate is increased to 8 packets/s. The figure reveals that load density still outperforms nodal activity and traffic density; it also reveals that nodal activity outperforms traffic density regardless of the number of flows with a difference that reaches around 12%. This is because when the network is congested due to heavy traffic, the MAC interface queue is at its highest levels of occupation. But most of the packets are likely to be dropped mostly because the MAC layer cannot deliver the packets. So the technique may not select the current best routes as the interface queue in this situation might not reflect the actual on going traffic (or traffic successfully transmitted). To illustrate in more detail, in Figure 4.27(a) when 10 flows are used, load density outperforms both nodal activity and traffic density by 12% and 15.5%, respectively. In contrast when the number of flows is increased to 20, in Figure 4.27(b) load density outperforms nodal activity by 5% and outperforms traffic density by 13.5%. Where as when the number of flows is increased to 30 flows, in Figure 4.27(c) load density outperforms nodal activity by 7% and outperforms traffic density by 18.5%.

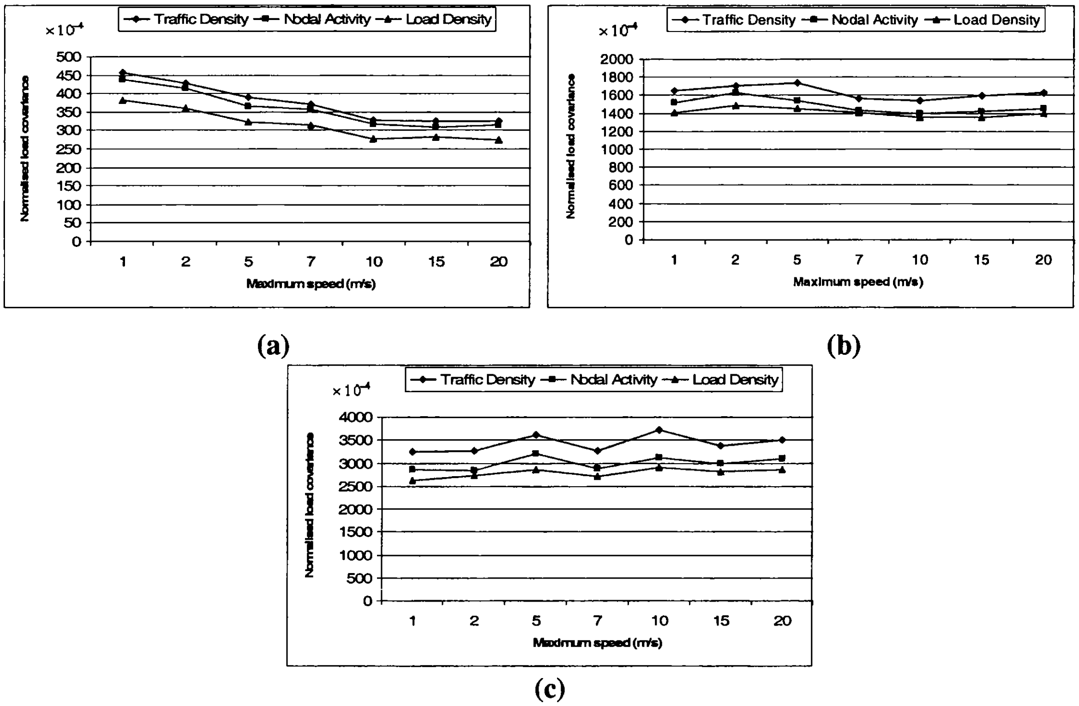


Figure 4.27: Normalised load covariance of the load density, nodal activity and traffic density techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

4.5 Conclusions

This chapter has proposed a new end-to-end traffic aware technique, namely *load density*. The technique utilises three parameters for path selection decision which are: the node's load, contention and route length. We have conducted an extensive evaluation of the proposed method and compared it against two existing end-to-end traffic aware techniques, namely *traffic density* and *nodal activity*, in order to assess the performance behaviour of our load density technique under different working environments. The simulation environments have incorporated various traffic rates and mobility speeds. The first part of the assessment has been carried out through studying the effects of varying the network size in terms of the topology dimensions and the number of nodes deployed over that topology. The second part of the assessment was intended to demonstrate the behaviour of the techniques when varying the length of data packets. The final part of the analysis was intended to study the behaviour of the techniques when varying the number of flows.

In the first part of the assessment which considers the effect of network size, load density outperforms both nodal activity and traffic density in terms of throughput, regardless of the network size. However the performance difference between load density and the other techniques increases when increasing the network size. For instance, at a traffic rate of 2 packets/s the difference between load density and traffic density rises from an average of 3% to become 9% when the network size increases from 75 to 100 nodes, while the difference when comparing load density and nodal activity rises from an average of 6% to become 20%. This is potentially because there are more route breakages in larger networks which affect the throughput achieved by all techniques, but apparently load density uses less congested routes compared to the other techniques. A similar behaviour is depicted for end-to-end delay. The difference between load density and traffic density on average is up to 15% for a network of 75 nodes, while for larger network of 100 nodes it goes up to 18%, and compared to nodal activity it is 32% and 36%, respectively. Moreover, in terms of routing overhead, load density generates about half of the overhead generated by the traffic density or nodal activity. Furthermore, in terms of normalised load covariance, load density has demonstrated a performance advantage of 10% and 14% over traffic density and nodal activity respectively.

In the second part of the assessment, decreasing the packet length from 512 to 128 bytes has shown that load density still achieves a higher throughput than the other two techniques with a relative difference up to 9% and 15% higher than in traffic density and nodal activity, respectively. In terms of end-to-end delay, load density has demonstrated a performance advantage with a relative difference as 23% and 42% lower end-to-end delay, respectively from traffic density and nodal activity. Moreover, in terms of routing overhead, there was a clear advantage in favour of load density regardless of packet length. Furthermore, in terms of normalised load covariance, load density has demonstrated a performance advantage over the other techniques.

In the final part of the assessment which considers the effect of varying the number of flows, load density outperforms nodal activity and traffic density. However a performance merit for nodal activity is revealed compared to traffic density. For instance when a number of 30 flows is used and at traffic rate of 2 packets/s load density outperforms nodal activity and traffic density in terms of throughput by 23% and 47% respectively. In terms of end-to-end delay, load density outperforms nodal activity with a difference a round 14%, and outperforms traffic density by around 24%. However the three techniques have shown comparable results such as when the network is extremely congested, for instance when 20 and 30 flows is used and traffic rate is 4 packets/s. Moreover in terms of routing overhead, load density has shown a clear performance advantage by generating around 28% less routing overhead from nodal activity and from traffic density it is around 34%. Furthermore, in terms of normalised load covariance, load density outperforms the other techniques with a difference around 16% from nodal activity and from traffic density it is around 29%

To summarise, the new load density technique has demonstrated a performance advantage over the other end-to-end techniques; traffic density and nodal activity. This comes to an agreement with the motivation behind the development of our new method and shows the performance improvement over the existing techniques. The subsequent chapter will describe a new on-the-spot traffic aware technique derived from our end-to-end technique. It will then conduct the first performance comparison between the two main approaches to traffic aware routing: end-to-end and on-the-spot.

Chapter 5

Performance Comparison of end-to-end and on-the-spot traffic aware techniques

5.1 Introduction

The major focus in previous chapters of this dissertation has been on the end-to-end category of traffic aware techniques. This chapter deals with the second category of traffic aware techniques, notably on-the-spot [28, 48]. These techniques use information that capture the current status of a small region in the network, and do not require the collection of information about the cost of the path between the end nodes; the source and destination. A given node actively monitors load conditions. This could detect overload conditions in its surrounding part of the network and needs to act upon it. However, an additional overhead can be generated from attempting to create a new route away from the overloaded nodes.

End-to-end techniques require gathering information about the load at intermediate nodes in the process of forming a path between a given pair of nodes to obtain a good estimation of load in the network in an attempt to avoid overloaded nodes. This could lead a better judgment for distributing the traffic load. However, these techniques avoid interactively distributing the load, and tend to wait until the route maintenance phase is initiated. Depending on the frequency of route maintenance, the gathered information might become out of date as it may not represent the current network load. On-the-spot techniques have been suggested as a solution to distribute the load among nodes on the argument that the end-to-end approach might be hard to achieve good performance levels [28, 48]. Despite the fact that there has been a number of research studies on traffic aware routing [28, 37, 48, 64, 75], there has not been so far an attempt to compare the relative performance merits of end-to-end and their on-the-spot counterparts.

This chapter conducts an extensive comparative study by means of ns-2 simulations of the two main categories of traffic aware routing, notably end-to-end and on-the-spot. To the best of our knowledge, this is the first research that conducts such a performance comparison. To this end, we have adapted our new traffic aware technique presented in Chapter 4 to suggest a new "on-the-spot" traffic aware technique, which will be referred to here as *on-the-spot load density*. Although there exist several on-the-spot techniques in the literature [28, 48, 95], we have opted for our *on-the-spot load density* so that the competing techniques have comparable implementation requirements, and as a consequence, the performance comparison is made fair and realistic.

The rest of this chapter is organised as follows. Section 5.2 describes the proposed on-the-spot traffic aware technique. Section 5.3.1 analyses the effects of the network size in terms of number of nodes and the size of network area on the performance of the traffic aware techniques while Section 5.3.2 analyses the effects of number of flows. Finally, section 5.4 concludes the chapter.

5.2 Description of on-the-spot load density technique

The proposed *on-the-spot load density* technique is determined using two components; the *load history* represented by the total traffic passed over a node, and the *contention* represented by the number of packets waiting at the node's interface queue and the interface queue of the nodes in its neighbourhood (on the spot) in order to take into consideration the degree of contention around the node. In our end-to-end version of load density the nodes do not use explicit exchanges of contention information in the neighbourhood as is the case in the on-the-spot. However, they use implicit exchange of information on contention at the nodes forming the route to the destination, embedded in the route request. So, in order to mirror this we have used explicit exchange. The new on-the-spot technique is embedded into the AODV-like routing protocol which has been used for the end-to-end load density. The operation of the technique consists mainly of three processes: *load conditions monitoring*, *route discovery* and *route maintenance*. The sequel will describe in more detail the operation of on-the-spot load density.

Load conditions monitoring:

The overload condition has two components: the node's history of forwarded data traffic (in bytes) and the contention in the node's local area (i.e.1-hop distance). The load history of a node i is calculated by accumulating the data packets lengths (in bytes) passed over that node as expressed in equation 5.1, where p is a data packet passing over node i.

$$load\_history(i) = \sum_{p \in i} length(p) \qquad (5.1)$$

Contention is represented by the average number of packets at the node's interface queue and the average number of packets at the neighbouring nodes' interface queue. The average number of packets at the interface queue N_Q is computed using a number of samples, say n, where a sample is taken every sample-period (10 ms), as written in equation 5.2. Each node updates $N_Q(i)$ every 1 second i.e. after taking 100 samples of the interface queue.

$$N_Q(i) = \frac{\sum_{k}^{n} q_k^i}{n} \qquad (5.2)$$

The contention at node i is estimated using equation 5.3, given below, where $N_Q(j)$ is the average number of packets at the interface queue at node j. Each node exchanges the contention information with its neighbours every exchange interval using "hello" packets, typically every 1 second. The hello packet is broadcasted only for one hop i.e. only to the immediate neighbours. Neighbours that receive the hello packet update their neighbourhood contention information.

$$Contention(i) = N_Q(i) + \sum_{\forall j \in nb(i)} N_Q(j) \qquad (5.3)$$

Nodes update their overload condition periodically (every 1 second). A node sets overload condition in any of the following conditions: If contention exceeds the predetermined *contention threshold* or if the average interface queue N_Q reaches more than half its full capacity, or if load history exceeds the *load history threshold.* When an overload condition is declared at a node, no action is taken at the time, until new traffic arrives at the node in the hope that the overload condition has eased. When a data packet arrives at an intermediate node with a declared overload condition, the packet is forwarded as normal. However, a route error is sent to the source to initiate a new route discovery. Figure 5.1 illustrates this process in more detail. In contrast, for control packets, the case is different. When a route request arrives at a node, it is updated as explained in the route discovery algorithm in Figure 5.2, whereas any other control packets are just forwarded.

```
On-the-spot traffic-aware routing algortihm (Overload monitoring) {
// executed by every node in the nerwork
// Overload conditions reperesnt both contention and load history
// overload_condition is initialy false
    Update overload conditions every time interval { // e.g. every 1 second
            If contention > contention-threshold
              or node's interface queue > contention-threshold / 2
             or load_history> load-threshold{
               overload_condition = true;
            } else {
               overload_condition = false;
            }
    }
    When forwarding a data packet check overload conditions{
        If overload_condition {
            Forward packet as normal
            Send a route error packet to the source node to initiate a new route request
        } else {
            Forward packet as normal
        }
} End (Overload monitoring)
```

Figure 5.1: Outline of the load monitoring algorithm in on-the-spot technique.

Route discovery:

The route discovery process starts whenever a node needs to communicate with another node for which it does not have a known route. The source node broadcasts a *route request* packet to its neighbours. Every node receives the route request packet forwards it to its neighbours after updating the overload flags in the route request packet. The overload flags include *load history flag* and *contention flag*. These flags are set according to the state of the forwarding node. The load history flag is set if *load-history* exceeds a certain *load threshold.* Similarly the contention flag is set if *contention* exceeds a certain *contention threshold* or if average interface queue is more than half full. The rational behind the overload flags is to allow the nodes to use the only available route between source and destination even if it was overloaded. The process of forwarding the route request packet continues until it is received by the destination node.

The destination collects the arriving route request packets within a route selection period; activated upon receiving the first route request packet, for selecting the overload-free route. Once the selection period is completed the destination selects the route according to the following preference: the first route that has no overload conditions (i.e. no flags are set), if not; the first route that has unset load history flag, if not available it selects first the route that has unset contention flag, otherwise it takes the first route regardless of its status, and then sends a *route reply* packet to the source node. When the source node receives the route reply packet, the path is then established and communication can commence. The route discovery and selection process is illustrated in details in Figure 5.2.

The current settings of the *selection period, load history threshold* and *contention threshold* used in the route discovery algorithm have been set as illustrated in Table 5.1. The value of the contention threshold is set to the size of the interface queue (which is 50 packets). For the load history threshold we have used three settings, which are 1.5MB, 3MB and unlimited. The 3MB and 1.5MB representing respectively the half and quarter of the data delivered in the network, when 4 packets/s for 5 flows and a network with 75 nodes is used. The unlimited setting will ignore the load history condition and decision is based only on the contention. The selection period has been set to the same value used for the other techniques discussed in this research.

TABLE 5.1: PARAMETERS OF THE ON-THE-SPOT TECHNIQUE.

Parameter	Value
Selection period	50 ms
Contention threshold	50 packets
Load history threshold	1.5 MB, 3 MB, and unlimited (contention only)

On-the-spot traffic-aware routing algortihm *(Route discovery) {*
// executed by every node in the nerwork
If a route to a destination is required
Broadcast a route request packet;
When a route request copy is received by a node other than destination {
Update the overload condition flags as follows {
If contention > contention-threshold or node's interface queue > contention-threshold / 2
set the contention-flag in the route request packet;
If load_history> load-threshold
set the load-flag in the route request packet
}
If first route request copy
Forward the route request to the neighbours;
Else If duplicate route request copy && meets with the loop-free and link-disjointness criteria in AOMDV[55]
Forward the route request to the neighbours;
Else discrard route request;
}

At the destination:
When a route request is received by the destination
Wait for other potentional request copies within selction interval from the receiption of the first copy ;
// every route request copy forms an alternative route
// the selection interval is set by the traffic aware technique
//The rational of this that in some cases the available routes go through nodes with overload conditions,
//and in order to avoid the problem of having no route we use a route with overload conditions.
Select the route in the following order {
1. *Select a roout with no overload conditions, ie. The one that has unset load and contetnion flags*
2. *Select a route that has unset load-flag*
3. *Select a route that has unset contention-flag*
4. *Otherwise, select any route, e.g. the first route.*

}
Send a route reply to the source node which initiated the route request;

} End (Route discovery)

Figure 5.2: Outline of the route discovery algorithm in on-the-spot technique.

Route maintenance:

The route maintenance is triggered when there is a change in the topology that affects the validity of an active route. If either the source, an intermediate, or the destination on an active route moves out of the communication range, an alternative route must be found. Once a node detects that the next hop is unreachable, it propagates a *route error* packet that this route is no longer valid to inform the source node and the other nodes which might be affected. The nodes that receive the route error packet delete the affected routes. The source upon receiving the route error packet can initiate a new route discovery to find a new route instead of the broken one by initiating a new route request. A new route discovery process is initiated rather than using any stored routes instead in order to deal with mobility and its effect on the established routes. The route maintenance is illustrated in Figure 5.3.

```
End-to-end traffic-aware routing algortihm framework (Route maintenace) {
// executed by every node in the nerwork
    If link breakage is detected {
        Delete the affected routes;
        Notify affected node(s) by sending route error packet
        // each node holds a prucursors list contains the nodes using a specific route
    }
    If route error packet is received {
        Delete the affected routes;
        Notify affected node(s) by sending route error packet
    }
} End (Route maintenance)
```

Figure 5.3: Route maintenance algorithm in on-the-spot technique.

Example:

To illustrate the process of finding a route using on-the-spot technique, let us consider a simple scenario that is depicted in Figure 5.4. The network consists of 10 nodes with presumable load history and contention information associated with each node, and these values are presented in the following order: node's interface queue, sum of neighbours' interface queue, and the load history of the node, as it is shown in the figure. Nodes 1 and 6 represent the source (S) and destination (D) pairs. Initially the source node; node 1, initiates a route request packet that is broadcasted to its neighbours, which are node 2 and 9. Nodes 2 and 9 check if their load history and contention information exceed the predetermined thresholds, if it does then the overload flags in the route request packet is updated accordingly, eventually in this case it does not, and then the route request is forwarded to their neighbours. Every node that receives the route request packet updates the overload flags (the contention and load flags) in the packet if applicable, and then forwards the packet to its neighbours. For instance, node 10 receives two copies of the route request, one from node 3 and another one from node 8. For node 10 the contention (10+45=55 packets) exceeds the contention threshold (50 packets). In this case the contention flag of the two route request copies is set; to indicate that this route is going through a node that is experiencing a contention overload condition, and then it is forwarded. The process of updating and forwarding the packet continues until the packet is delivered to the destination.

From this scenario 5 routes from source to destination are available: 1-2-3-10-6, 1-9-8-7-6, 1-9-8-10-6, 1-2-3-4-6 and1-2-3-4-5-6. Let us name these routes as r1, r2, r3, r4 and r5 respectively. According to the route selection algorithm in Figure 5.2, we try to find a route that has overload flags which are unset. Assume that route request copies arrived in the above order (i.e. r1, r2, r3, r4, r5). In this case route r1 is dropped from the selection, and the route r2 is selected to be the route between node 1 and node 6 since it meets with the criteria as it has unset overload flags.

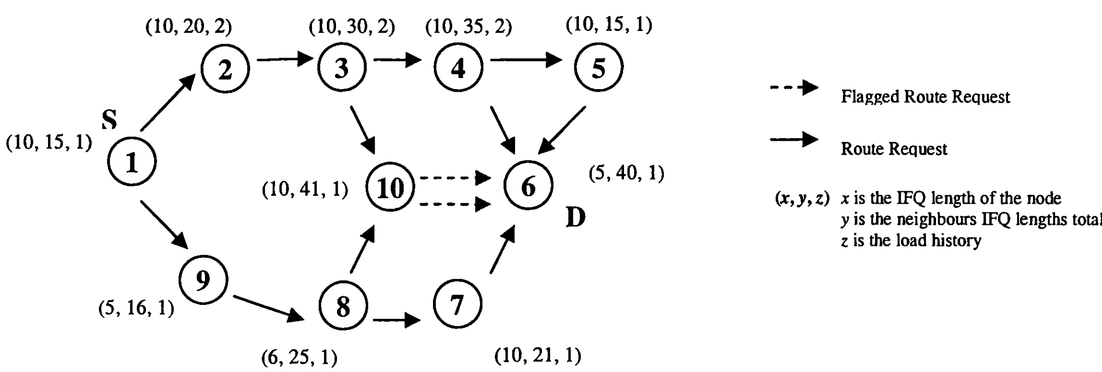


Figure 5.4: Example to illustrate the propagation of route request in on-the-spot technique.

5.3 Performance Analysis

This section analyses the performance of the suggested on-the-spot load density technique, and compares it against its end-to-end load density counterpart. The performance analysis is conducted using the same simulation model and parameters as outlined in Chapter 3 (see Section 3.2), and for the sake of clarity we include the list of parameters here (see Table 5.2). For the rest of the chapter we will use the term *end-to-end* interchangeably with *end-to-end load density*, and the term *on-the-spot* interchangeably with *on-the-spot load density*, for short.

TABLE 5.2: THE SYSTEM PARAMETERS USED IN THE SIMULATION EXPERIMENTS

Parameter	Values
Number of nodes	75, 100
MAC layer	IEEE 802.11
Transmission range	250m
Simulation area	900m×1000m,1200m×1000m
Simulation time	900s
Mobility model	Random waypoint model
Maximum speed	1, 2, 5, 7, 10, 15 and 20m/s.
Pause time	0s
Traffic type	CBR
Packet size	512 bytes
Packet rate	2, 4, 8 packets/s
Number of flows	5
Number of runs per data point	30

The performance of the two techniques is measured by four performance metrics: *throughput*, *end-to-end delay, normalised load covariance* and *routing overhead.* The throughput is the amount of data received (measured in bits per second) at the final destination over the simulated time averaged over number of flows. This measure provides an indication of the efficiency of the technique as it shows the amount of data that the network is able to deliver to destinations. The end-to-end delay is the average time interval between the generation of a packet at a source node and its successful delivery at the destination node. This delay accounts for all possible delays that can occur in the source and all intermediate nodes. The normalised load covariance is calculated by dividing the load covariance of the data forwarded by the number of packets successfully delivered at the destinations. Load covariance is defined according to equation 5.4, where node's load l_i is the sum of data packet lengths passed over node i. The mean μ is the total sum of the nodes' load over the number of nodes M participating in forwarding data packets. The rational behind this metric is to measure the ability of distributing the load while successfully delivering data packets. The lower this metric the better performance since the load covariance measures how far are nodes load from the mean. The routing overhead is the number of routing (control) packets generated during the simulated time in order to establish and maintain paths and to exchange traffic information among the network nodes as dictated by the operation of a given traffic aware metric.

$$\text{Load covariance} = \frac{\sum_{i}^{M}(l_i - \mu)^2}{M * \mu} \tag{5.4}$$

5.3.1 Performance impact of network size

The analysis is conducted using two network sizes of 75 nodes forming a MANET over a simulation area of 900m×1000m and 100 nodes forming a MANET over a simulation area of 1200m×1000m. The reason for choosing these numbers combination is to avoid the occurrences of frequent partitions in the network. The simulated network areas share the same width of 1000 m, but have different lengths of 900m and 1200m. The node maximum speed has been set to 1, 2, 5, 7, 10, 15 and 20 m/s; representing human slow walk speed to vehicle speeds. Traffic rate has been varied between 2, 4, and 8 packets/s, where 512-byte packets have been used. In the figures that are presented below, the *x*-axis describes the variations in node's maximum mobility speed, while the *y*-axis represents results of the performance metric of interest. The results have been conducted from steady state simulations. A 95% confidence interval is calculated for each point. In most of the cases the error bars are quite small. These error bars are not shown on all the figures in order to increase their clarity.

Throughput:

Figure 5.5 shows the throughput for the end-to-end load density and two variations of the on-the-spot technique that are different in the load history threshold settings, referred to as on-the-spot 1 (with load history threshold of 3MB) and on-the-spot 2 (with load history threshold of 1.5MB), when 75 nodes are used over a 900m×1000m topology area. Figure 5.5 (a) depicts the behaviour of the techniques for a traffic rate of 2 packets/s. Although at this traffic rate there are few overload conditions detected for both on-the-spot 1 and on-the-spot 2, the figure shows that the end-to-end load density achieves a slightly higher throughput than that of on-the-spot schemes, with a difference of 210bps on average over the varied speeds, which is about 3% of the average throughput of end-to-end load density (7320bps). The reason for this behaviour is that even with low contention or load history values, end-to-end load density would still select one of the available routes.

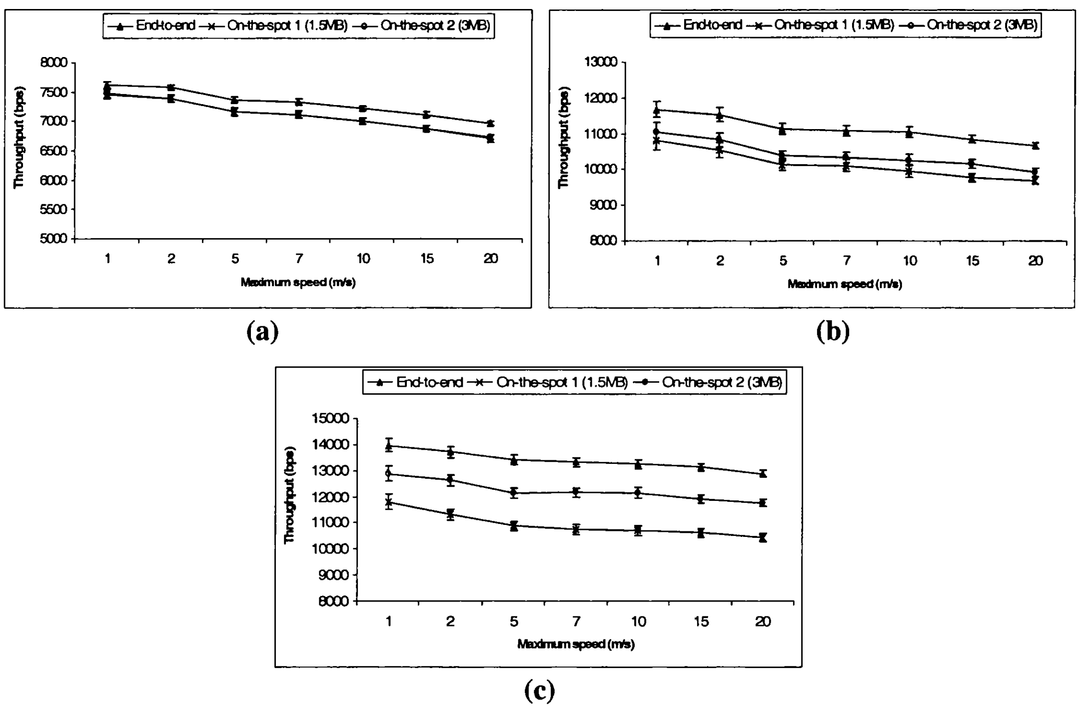


Figure 5.5: Throughput of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 5.5(b) shows the throughput when traffic rate is increased to 4 packets/s. As the figure demonstrates, increasing the traffic rate reveals the performance differences between on-the-spot schemes as the two will react differently depending on the load history threshold. The figure also shows that end-to-end load density outperforms on-the-spot schemes at all simulated speeds, with an average difference of about 1000bps from on-the-spot 1 and from on-the-spot 2 it is about 720bps, which is respectively about 9% and 6.5% of the throughput average of end-to-end load density (1140bps). On the other hand, the figure demonstrates the impact of changing the maximum mobility speed from 1m/s to 20m/s on the throughput. For instance, for the end-to-end technique the decrease in throughput due to increasing the mobility speed reaches the limit of 1020 bps; that is about 9% of its maximum throughput represented by the highest throughput achieved by the technique over all speeds, which is 11680 bps. In contrast, for on-the-spot 1, the throughput decrease reaches 1100 bps, that is approximately 10% of its maximum throughput which is 10800 bps, and for on-the-spot 2, the throughput decrease reaches 1120 bps that is about 10% of its maximum achieved throughput which is 11050 bps. This is because the increase in mobility causes frequent topology changes resulting in more broken routes.

Figure 5.5(c) depicts the throughput when the traffic rate is 8 packets/s. The figure demonstrates that the end-to-end method outperforms its counterparts of on-the-spot schemes with an average difference of 2460bps from on-the-spot 1, and from on-the-spot 2 the average difference is 1150bps, which is respectively about 18% and 8% of the end-to-end technique average throughput (which is 13390bps). It also demonstrates that on-the-spot 2 outperforms on-the-spot 1. This is due to the setting of the load history threshold as the on-the-spot 1 technique seems to overreact. Finding the best load threshold is a challenging issue and requires a further investigation as it requires a good knowledge of traffic injected into the network, and this might be hard to obtain.

Figure 5.6 shows the throughput for the traffic aware techniques, end-to-end and on-the-spot, when the network size is scaled up to 100 nodes distributed over a 1200m×1000m topology area. Figure 5.6(a) demonstrates the behaviour of the techniques for a traffic rate of 2 packets/s. We can observe from the figure that the throughput decreases as the mobility increases for the two techniques. This agrees with what have been observed in Figure 5.5 when the area of

900m×1000m has been considered. For the end-to-end technique the decrease in throughput is 18.5% of its maximum achieved throughput which is 6640 bps. For the on-the-spot schemes the decrease compared to their maximum achieved throughput which is 6620 bps, is 20%. Nonetheless, the end-to-end load density outperforms its counterparts under various mobility speeds. The end-to-end technique achieves on average over the varied speeds about 7.5% of its average throughput (5990bps) higher than the throughput of on-the-spot techniques. Also it worth mentioning that the throughput achieved for smaller network area 900m×1000m is higher than that achieved in this figure. This is because there is more potential for a route to break with larger areas because nodes are given more space to move in. Figure 5.7 shows the average number of route changes in the two network sizes over different mobility speeds, considering the 30 topologies we have used for each speed.

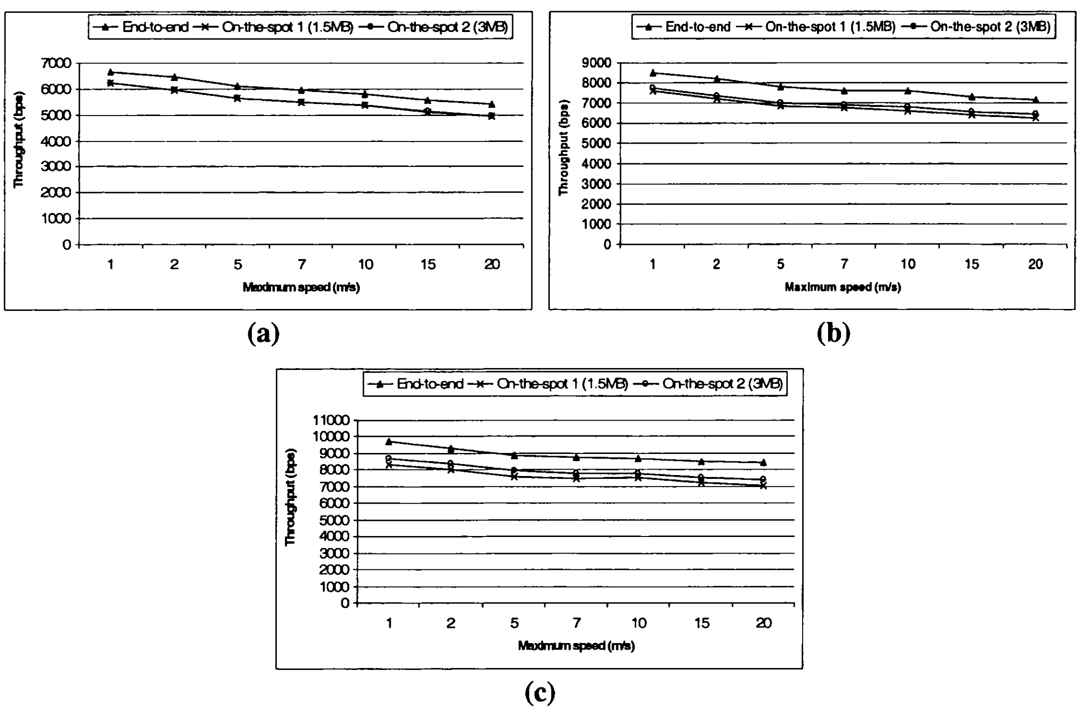


Figure 5.6: Throughput of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 5.6(b) shows the performance behaviour when the traffic rate is increased to 4 packets/s. As the figure demonstrates, the end-to-end load density outperforms the two schemes of on-the-spot technique with a difference of 11.9% and 9.7% of its average throughput (which is 7740bps) from on-the-spot 1 and on-the-spot 2, respectively. Similar results are depicted in Figure 5.6(c) when higher traffic rate of 8 packets/s is used. The figure shows that the end-to-end technique achieves on average 14.5% of its average throughput (which is 8900bps) higher than the throughput of on-the-spot 1, and 11% higher than the throughput of on-the-spot 2.

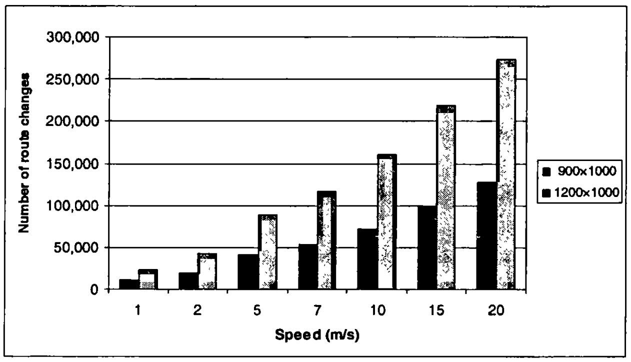
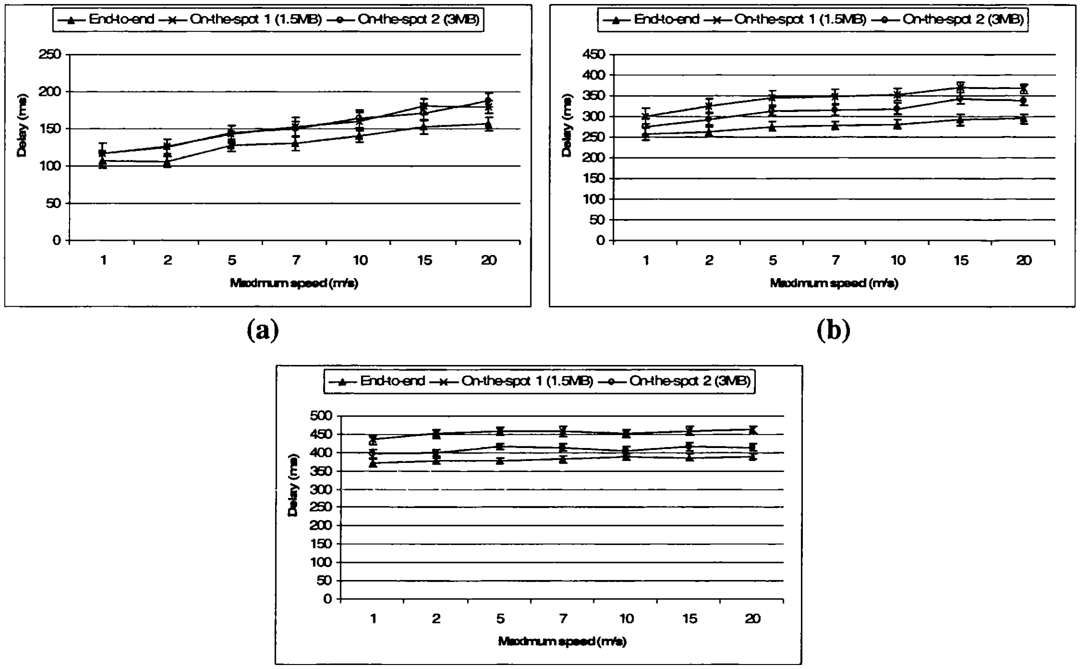


Figure 5.7: Number of route changes in the used networks: 75 nodes over 900m×1000m and 100 nodes over 1200m×1000m.

End-to-end delay:

Figure 5.8 shows the end-to-end delay for end-to-end load density and on-the-spot load density techniques under different traffic rates that are relatively, light, moderate and high traffic with rates of 2, 4 and 8 packets/s, respectively, when 75 nodes are used over a 900m×1000m topology area. Figure 5.8(a) demonstrates the behaviour of the two techniques when the traffic rate is 2 packets/s. As the figure shows, the two on-the-spot schemes exhibit comparable performance while the end-to-end technique exhibit slightly better performance from both of them. Although the difference between the end-to-end and on-the-spot schemes is only 20ms on average, this corresponds to 15% of the average end-to-end delay of the end-to-end technique (which is 130ms). On the other hand, the impact of increasing the speed affects all methods. Increasing the speed from 1m/s to 20m/s results in an increase of 70 ms in the end-to-end delay for on-the-spot 1, 60 ms for on-the-spot 2, and 50 ms for end-to-end technique. This is because the number of broken routes increases as the speed increases, which require a route discovery initiation to establish a new route. Since the route discovery latency is included in the end-to-end delay then the end-to-end delay increases.



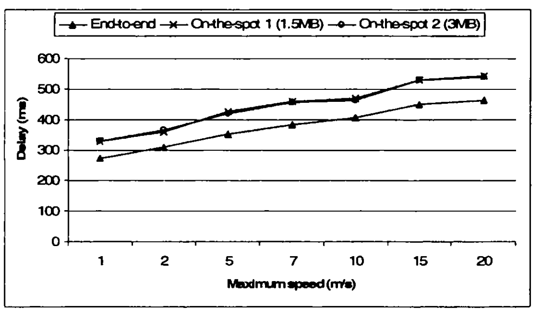
(a) (b)

(c)

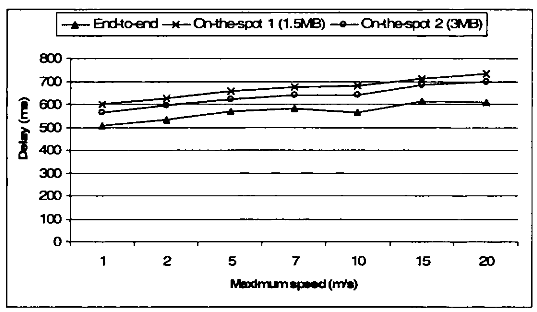
Figure 5.8: End-to-end delay of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 5.8(b) shows the performance for the techniques when the traffic rate is increased to 4 packets/s. The figure demonstrates a clearer picture about the performance comparison between the techniques compared to Figure 5.8(a) as difference between the techniques is getting wider. The end-to-end load density outperforms the schemes of on-the-spot technique with a difference of about 65ms from on-the-spot 1, and from on-the-spot 2 the difference is about 35ms. The speed variation, on the other hand, has not shown significant effect on the end-to-end delay except that it increases slightly as mobility speed increases. This is because increasing the traffic rate causes congestion in the network (more than 30% of the packets has been dropped) which reduces the impact of speed, leading to an increase in the end-to-end delay compared to Figure 5.8(a). A similar behaviour is shown in Figure 5.8(c) when higher traffic rate of 8 packets/s is used.

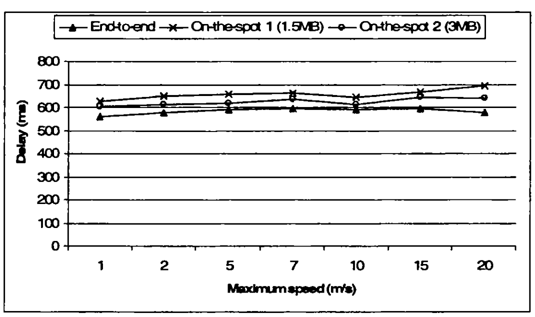
Figure 5.9 demonstrates the behaviour of end-to-end load density compared to on-the-spot load density when the traffic rate is 2, 4 and 8 packets/s, respectively, and 100 nodes are used over a 1200m×1000m topology area. It is noticeable when comparing the performance in Figures 5.9(a, b, c) against that in Figures 5.8(a, b, c) where smaller network size of 75 nodes over a 900m×1000m topology area is used that the minimum end-to-end delay has greatly increased compared to the end-to-end delay of the smaller network size although the same traffic scenarios are used. This observed behaviour is due by two reasons. Firstly, nodes in the larger topology have more freedom in mobility as they have more space to move in, this would result in more broken routes as it has been illustrated in Figure 5.7. Secondly, the existence of longer paths would increase the delay especially when there is a link break as this requires discovering a new route. Figure 5.10 demonstrates the average minimum path length in hop count for the used traffic flows in the two used network sizes.



(a)



(b)



(c)

Figure 5.9: End-to-end delay of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 5.9(a) shows that the two schemes of on-the-spot technique exhibit similar performance; it shows also that the end-to-end load density outperforms on-the-spot schemes, with an average difference about 70ms. The impact of increasing the mobility speed has shown great effect on all techniques. For end-to-end load density, the increase in the end-to-end delay is around 190 ms while for on-the-spot schemes it is around 210ms.

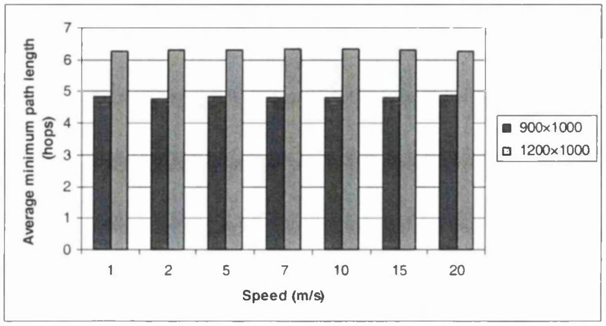


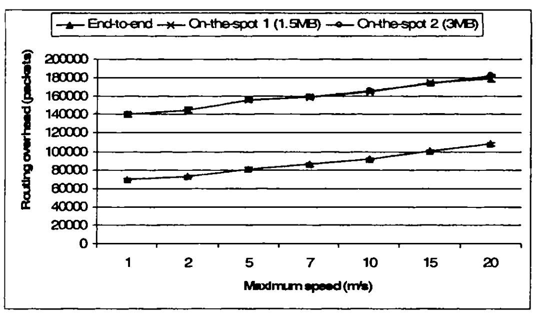
Figure 5.10: Average minimum path length in hop count in the used networks: 75 nodes over 900m×1000m and 100 nodes over 1200m×1000m.

Similarly in Figure 5.9(b) when the traffic rate is increased to 4 packets/s, end-to-end technique has better performance than the two schemes of on-the-spot technique by exhibiting lower end-to-end delay values for the simulated mobility speeds. While on-the-spot 2 has on average 35ms lower end-to-end delay compared to on-the-spot 1, end-to-end load density has on average 65ms lower end-to-end delay compared to on-the-spot 2 and hence better performance than both. On the other hand, the impact of speed on the three techniques when the traffic rate is increased is less than that in Figure 5.9(a) when lower traffic rate is used. For end-to-end technique, the increase in the end-to-end delay due to the increase in speed from 1m/s to 20m/s is around 100ms, while for on-the-spot schemes, the increase in the end-to-end delay due to the increase in speed is 135ms.

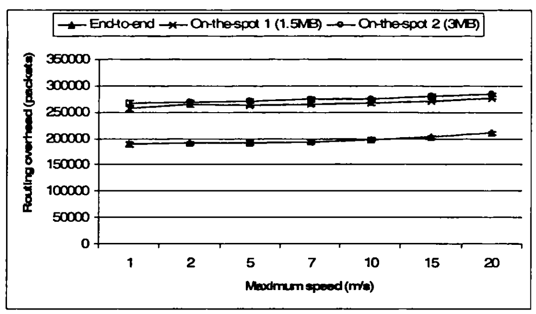
In contrast, when increasing the traffic rate to 8 packets/s, the techniques exhibit close performance as it is shown in Figure 5.9(c). However, the end-to-end technique still outperforms on-the-spot schemes with on average 75ms from on-the-spot 1 and 40ms from on-the-spot 2. The results also reveal that in this case the impact of speed is low. This is potentially because significant amount of packets are dropped. The calculation of the end-to-end delay does not consider dropped packets. Only 28% and less of sent packets have been successfully delivered. This also justifies why the maximum end-to-end delay in this figure is less than that in Figures 5.9(a) and 5.9(b).

Routing overhead:

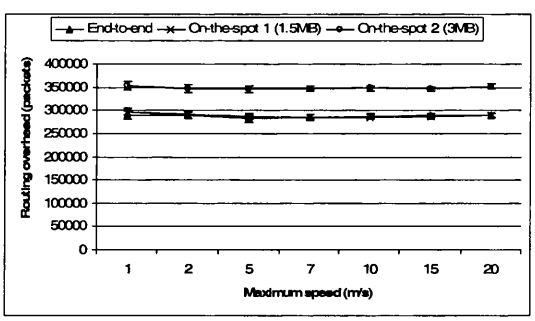
Figure 5.11 shows the generated routing overhead by two methods; end-to-end and on-the-spot when the traffic rate is 2, 4 and 8 packets/s, and when 75 nodes are used over a 900m×1000m topology area. Figure 5.11(a) shows comparable routing overhead generated by on-the-spot schemes with an average of 160,000 routing packets, while the end-to-end technique generate on average about 87,000 routing packets. The figure also reveals that when the traffic rate is 2 packets/s, the routing overhead increases noticeably as the mobility speed increases. The overhead increases by around 40,000 routing packets at the speed of 20 m/s compared to that at 1 m/s for all techniques. This is because the increase in mobility causes frequent topology changes resulting in broken routes which become more frequent as speed increases.



(a)



(b)



(c)

Figure 5.11: Routing overhead of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 5.11(b) demonstrates that the generated overhead from the techniques for the relatively higher traffic rate of 4 packets/s. The results show an increase of more than 100,000 routing packets in the generated overhead by all techniques compared to that in Figure 5.11(a) due to the increase in traffic rate. On the other hand, the impact of increasing the mobility speed from 1m/s to 20m/s on the generated overhead reaches the limit of 21,000 routing packets for end-to-end load density, 20,000 routing packets for on-the-spot 1 and for on-the-spot 2 it is 18,000 routing packets. Figure 5.11(c) shows the generated overhead when the traffic rate is increased to 8 packets/s. As the figure demonstrates, on-the-spot 1 and end-to-end load density exhibit close performance, while on-the-spot 2 generates on average about 60,000 routing packet extra compared to that of on-the-spot 1 and end-to-end technique. However for on-the-spot 1 this is not considered an advantage bearing in mind its throughput and end-to-end compared to that of on-the-spot 2.

Figure 5.12 shows the routing overhead under for the traffic rates of 2, 4 and 8 packets/s, and when 100 nodes are used over a 1200m×1000m topology area. Figure 5.12(a) shows that end-to-end load density outperforms on-the-spot schemes with a difference around 97,000 routing

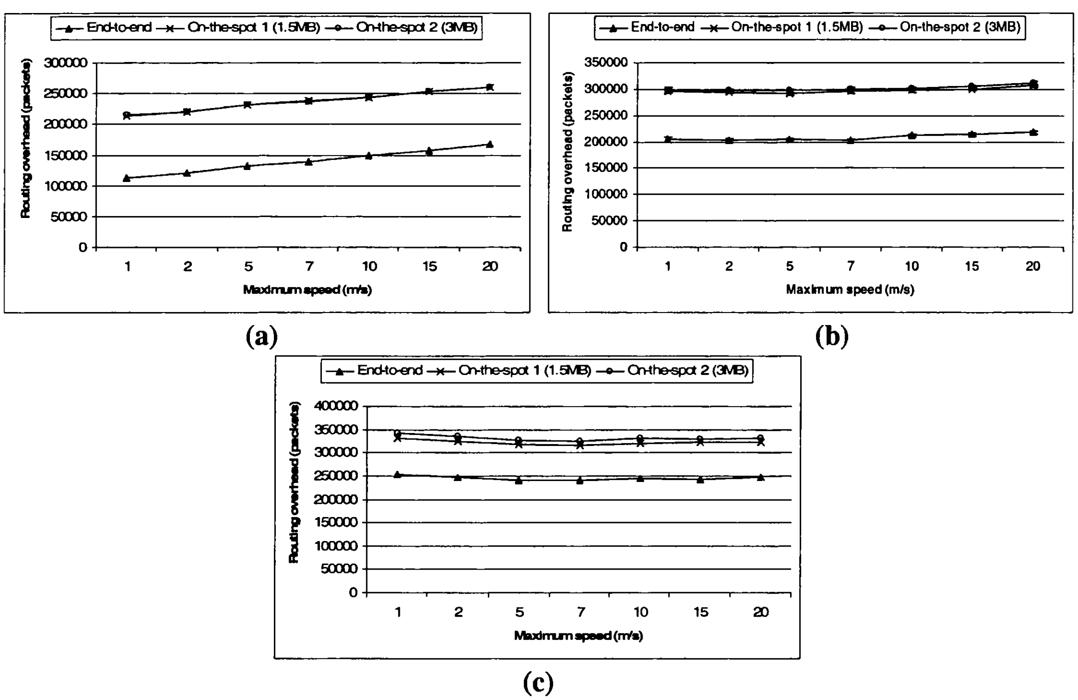


Figure 5.12: Routing overhead of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

packets, on average. Similarly in Figures 5.12(b, c) at traffic rates 4 packets/s and 8 packets/s, respectively, end-to-end load density outperforms both on-the-spot 1 and on-the-spot 2.

Normalised load covariance:

Figure 5.13 demonstrates the normalised load covariance for end-to-end load density and on-the-spot load density schemes for traffic rates of 2, 4 and 8 packets/s, and when 75 nodes are used over a 900m×1000m topology area. The figure reveals that the end-to-end load density generally outperforms on-the-spot schemes with a difference reaches the limit of 18% from on-the-spot 1 and around 9.5% from on-the-spot 2, at traffic rate of 8 packets/s. In contrast the two on-the-spot schemes show similar performance at traffic rates of 2 packets/s. However when the traffic rate is increased to 4 and 8 packets/s, on-the-spot 2 outperforms on-the-spot 1 with a difference reaches around 4% and 9.5%, respectively. This is related to the setting of the load history threshold (1.5 MB and 3MB), which has been set representing, respectively, the quarter and half of the data delivered in the network, at traffic rate 4 packets/s.

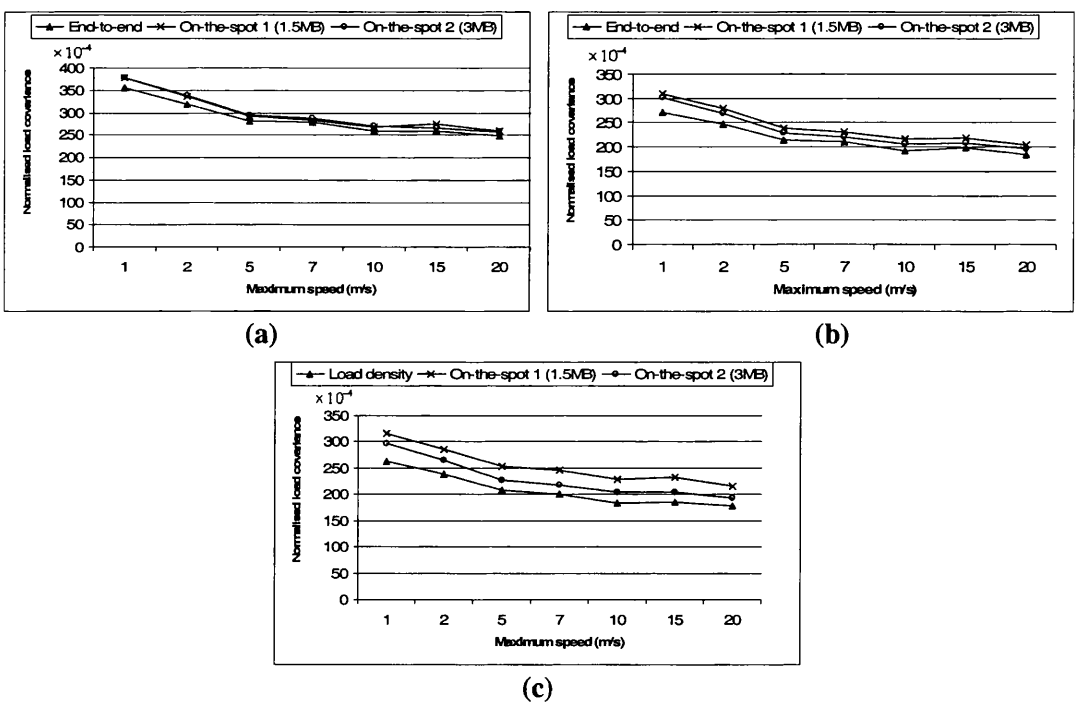


Figure 5.13: Normalised load covariance of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

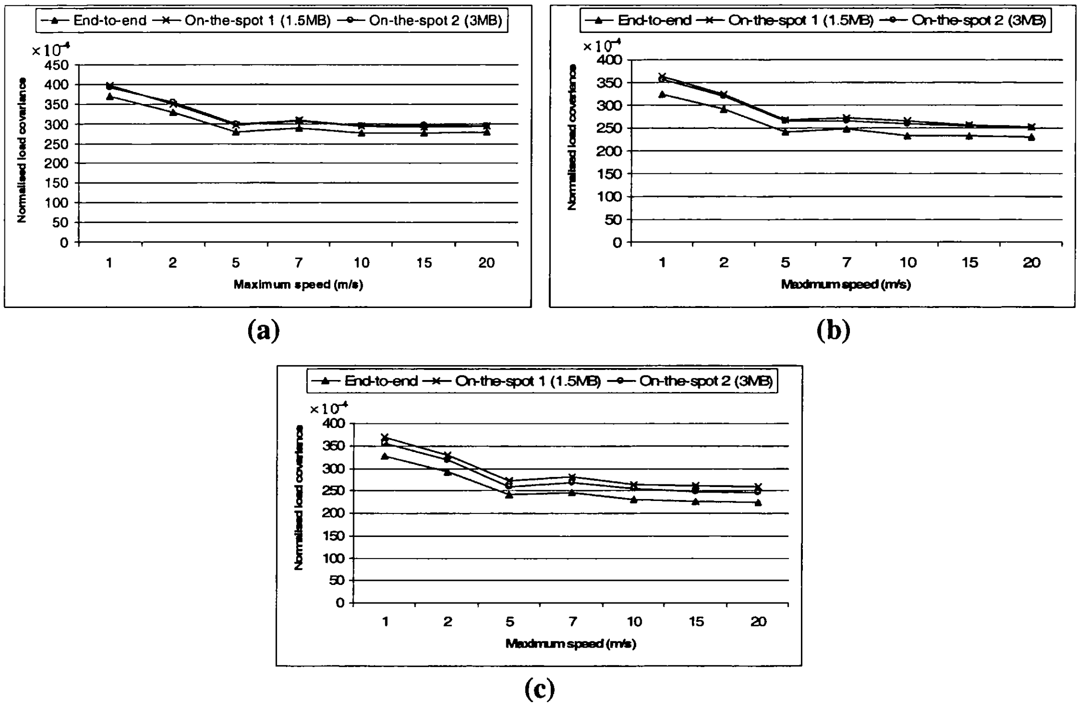


Figure 5.14: Normalised load covariance of the end-to-end, on-the-spot 1 and on-the-spot 2 techniques for a network of 100 nodes over topology area 1200m×1000m with a 512-byte packet. (a) traffic rate =2 packets/s, (b) traffic rate =4 packets/s, (c) traffic rate =8 packets/s

Figure 5.14 shows the normalised load covariance for the three techniques for traffic rates of 2, 4 and 8 packets/s, and when 100 nodes are used over a 1200m×1000m topology area. The figure reveals that on-the-spot schemes generally show comparable performance. Whereas end-to-end load density show a performance advantage compared to on-the-spot 1 and on-the-spot 2 with a difference that reaches the limit of 12% and 9%, respectively.

5.3.2 Performance impact of increasing the number of flows

The analysis is conducted using a network size of 75 nodes forming a MANET over a simulation area of 900m×1000m. We have considered only one of the network sizes used in the previous section in order to keep the time spent on running the experiments manageable. The node speed has been set to 1, 2, 5, 7, 10, 15, and 20 m/s. Traffic rate has been varied between 2, 4, and 8 packets/s, where 512-byte packets have been used. The number of flows has been set to 10, 20, and 30 flows. For each of the settings we monitor the performance of 5 flows that are

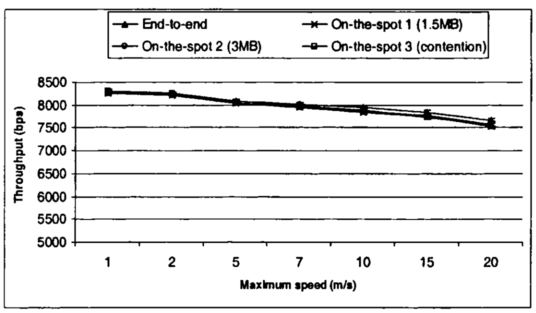
TABLE 5.3: THE SYSTEM PARAMETERS USED IN THE SIMULATION EXPERIMENTS FOR BACKGROUND TRAFFIC

Parameter	Values
Number of nodes	75
MAC layer	IEEE 802.11
Transmission range	250m
Simulation area	900m×1000m
Simulation time	900s
Mobility model	Random waypoint model
Maximum speed	1, 2, 5, 7, 10, 15 and 20m/s.
Pause time	0s
Traffic type	CBR
Packet size	512 bytes
Packet rate	2, 4, 8 packets/s
Number of flows	10, 20, 30 flows
Number of runs per data point	30

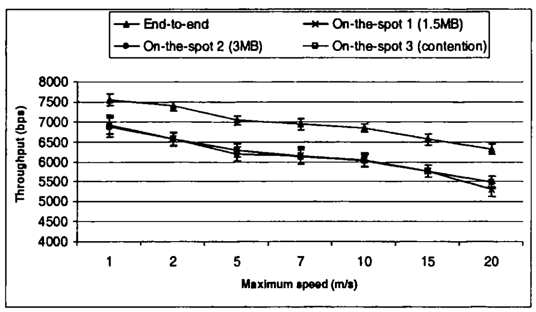
formed according to the network model in Chapter 3 (see Section 3.2) where the source of traffic and its destination are stationary, while the rest flows are considered as background traffic. For background traffic flows, the sources and destinations are not stationary and are free to move in any direction and speed. Table 5.3 illustrates the simulation parameters used in the experiments conducted in this section. In the figures that are presented below, the x-axis describes the variations in node's maximum mobility speed, while the y-axis represents results of the performance metric of interest. The results have been conducted from steady state simulations. A 95% confidence interval is calculated for each point. The comparison between end-to-end load density and on-the-spot load density includes the three load history threshold settings for on-the-spot illustrated in table 5.1.

Throughput:

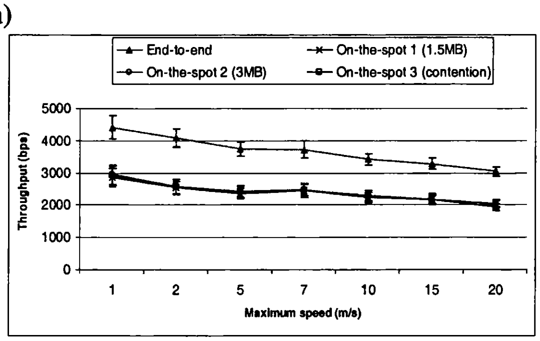
Figure 5.15 demonstrates the achieved throughput for end-to-end load density and the schemes of on-the-spot load density when traffic rate is 2 packets/s, and when 10, 20 and 30 flows are used. The techniques show comparable performance when 10 flows are used in Figure 5.15(a). However end-to-end load density shows better performance in terms of normalised load covariance as illustrated in Figure 5.24(a). Figure 5.15(b) shows the throughput for the techniques when the number of flows is increased to 20 flows. As the figure illustrates end-to-end load density outperforms the three schemes of on-the-spot technique, with a difference around 800bps from on-the-spot 2 that is 11.5% of its average throughput (which is 6950bps), on average over the simulated speeds. The figure also reveals negligible differences between on-the-spot schemes. Similar behaviour is depicted in Figure 5.15(c) when 30 flows are used, but the difference has increased to 1260bps from on-the-spot 2 that is approximately 34% of its average throughput (3670bps). The impact of changing the mobility maximum speed from 1m/s to 20m/s has affected all techniques. For instance in Figure 5.15(a), the throughput for end-to-end load density decreases as the mobility speed increases and the decrease reaches the limit of 640bps; that is about 7.7% of its maximum achieved throughput (which is 8300bps). In contrast, for on-the-spot schemes the decrease in throughput is around 730bps that is about 8.8% of their maximum throughput.



(a)

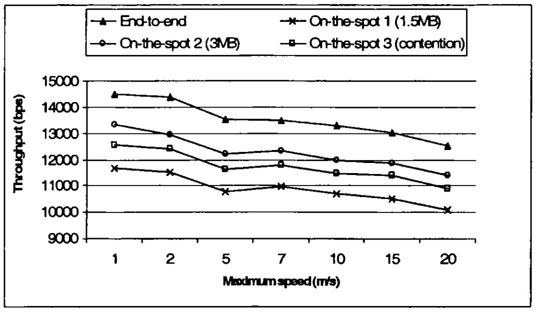


(b)

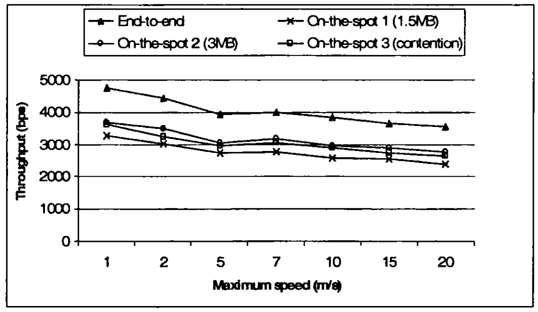


(c)

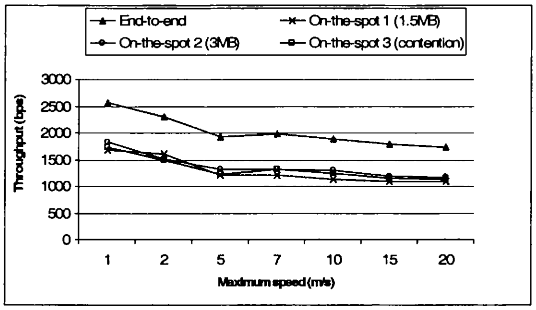
Figure 5.15: Throughput of the end-to-end, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows



(a)



(b)



(c)

Figure 5.16: Throughput of the end-to-end, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 5.16 shows the achieved throughput by end-to-end load density and on-the-spot schemes when the traffic rate is increased to 4 packets/s, and when 10, 20 and 30 flows are used. Figure 5.16(a) show that on-the-spot 2 (with load history threshold of 3MB) outperforms on-the-spot 1 (with load history threshold of 1.5MB) with on average 1400bps, and outperforms on-the-spot 3 (unlimited threshold; i.e. contention only) with on average 560bps. However, it also shows that end-to-end load density outperforms on-the-spot 2 by 1230bps on average, and hence outperforms the other schemes. Furthermore, the figure reveals the importance of load history component as on-the-spot 2 outperforms on-the-spot 3. In contrast when the number of flows is increased to 20 flows in Figure 5.16(b) on-the-spot 2 still outperforms the other schemes, but the difference is smaller than that in Figure 5.16(a), this because of the sharp drop in the number of delivered packets. The drop has affected both on-the-spot load density schemes as well as end-to-end load density where less than 28% of the packet has been delivered. The drop in the delivery ratio is because increasing the number of flows to 20 would mean that 40 nodes (which is more than half of the network size); 20 sources and 20 destinations, are trying to establish a communication with each other. Nonetheless, end-to-end load density outperforms on the-spot 2 by 880bps, on average, and hence outperforms the other schemes. Similar behaviour is depicted in Figure 5.16(c) when the number of flows is increased to 30.

Figure 5.17 demonstrates the achieved throughput by end-to-end load density and on-the-spot schemes under 10, 20 and 30 flows, when the traffic rate is increased to 8 packets/s. The increase in traffic rate has shown great impact on the techniques as the throughput drops compared to lower traffic rates, when 2 packets/s and 4packets/s in Figures 5.15 and 5.16, since this has increased the congestion in the network. The impact gets more sever as we increase the number of flows as it is shown in Figures 5.17(b, c). For instance, with 30 flows, only 6% and less of the packets have been delivered to their destinations. Nonetheless, the end-to-end load density has shown outperformance compared to on-the-spot schemes.

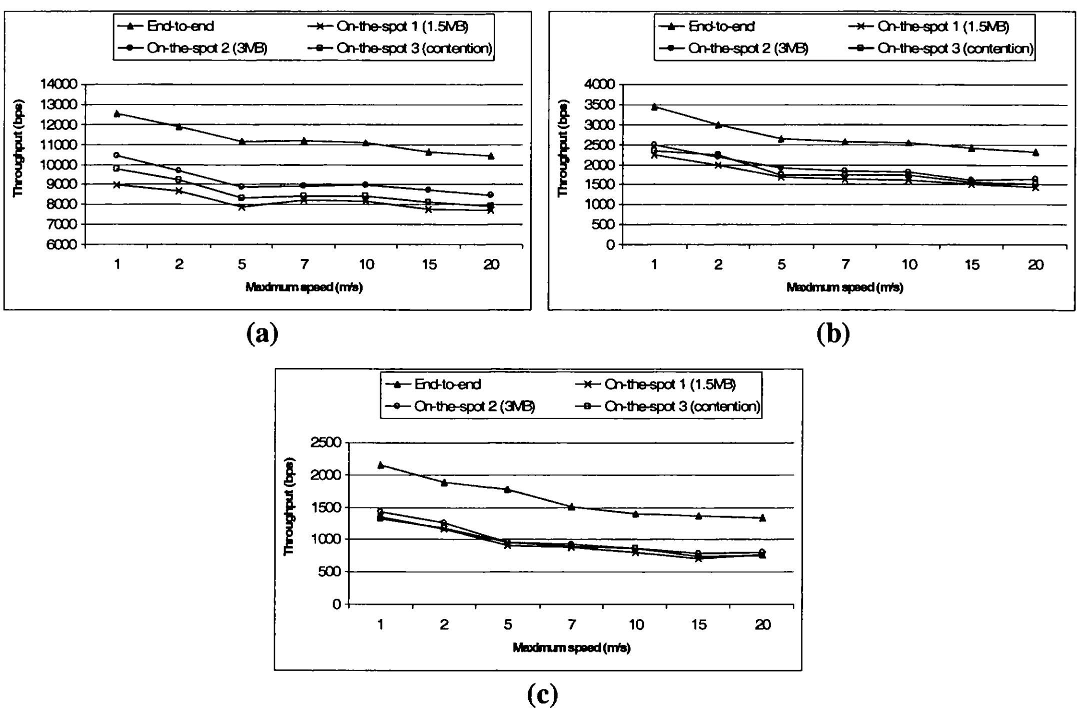


Figure 5.17: Throughput of the end-to-end, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

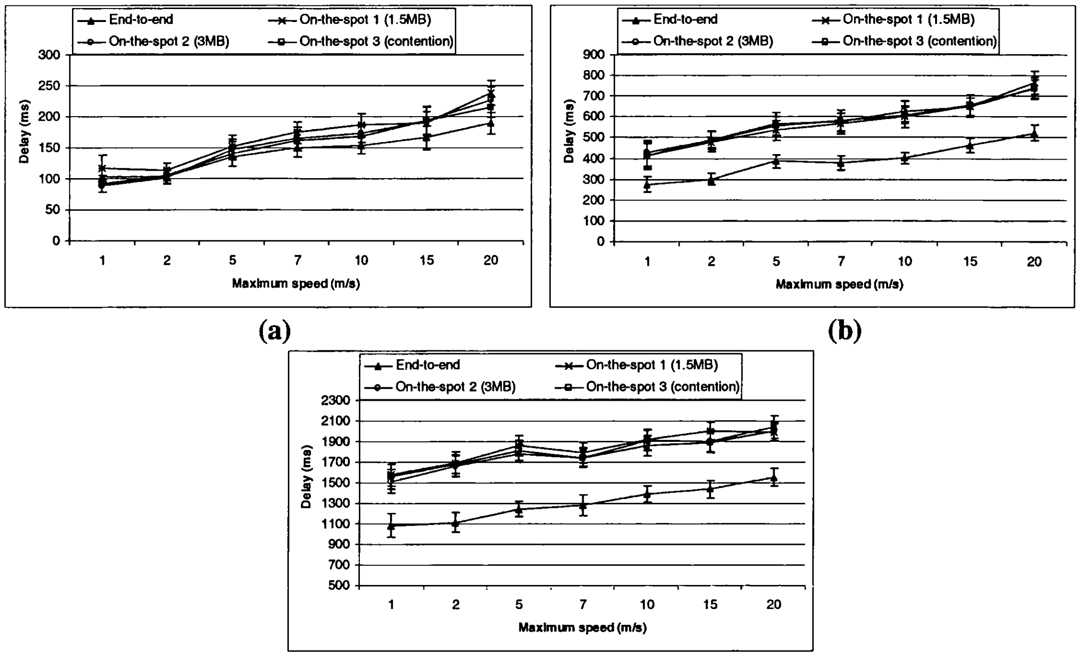


Figure 5.18: End-to-end delay of the end-to-end load density, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

End-to-end delay

Figure 5.18 shows the end-to-end delay for end-to-end load density and on-the-spot load density schemes when traffic rate is set to 2 packets/s, for 10, 20 and 30 flows. Figure 5.18(a) demonstrates the performance when 10 flows are used. As the figure reveals, the techniques show comparable results. However, when increasing the number of flows to 20 in Figure 5.18(b), the techniques show different performance. The end-to-end technique outperforms the schemes of on-the-spot with a difference reaches the limit of 190ms; on average over the simulated speeds. Figure 5.18(c) shows the performance of the techniques when 30 flows are used. As we inject more traffic to the network, the network becomes congested, around half of the traffic and less has been delivered to destination, and as the figure shows, a sharp increase in the end-to-end delay compared to that in Figure 5.18(b). Nonetheless, the figure shows end-to-end load density outperforms on-the-spot 2 and on-the-spot 3 with an average difference of 490ms, and outperforms on-the-spot 1 with an average difference of 530ms.

Figure 5.19 shows the achieved end-to-end delay by end-to-end load density and on-the-spot schemes under 10, 20 and 30 flows, when the traffic rate is increased to 4 packets/s. Figure 5.19(a) demonstrates that the end-to-end load density outperforms the three schemes of on-the-spot load density with a difference of 110ms, 220ms, and 390ms, respectively from on-the-spot 2, on-the-spot 3, and on-the-spot 1. In contrast, on-the-spot 2 (with load history threshold of 3MB) outperforms on-the-spot 1 (with load history threshold of 1.5MB) with on average 275ms, and outperforms on-the-spot 3 (with unlimited threshold; i.e. contention only) on average with 100ms. However when the number of flows is increased to 20 in Figure 5.19(b) all techniques demonstrate a sharp increase in the end-to-end delay. This is because of the amount of traffic injected into the network, which has affected both delay and throughput for all techniques, where less than 28% of the packets has been delivered to the destination.

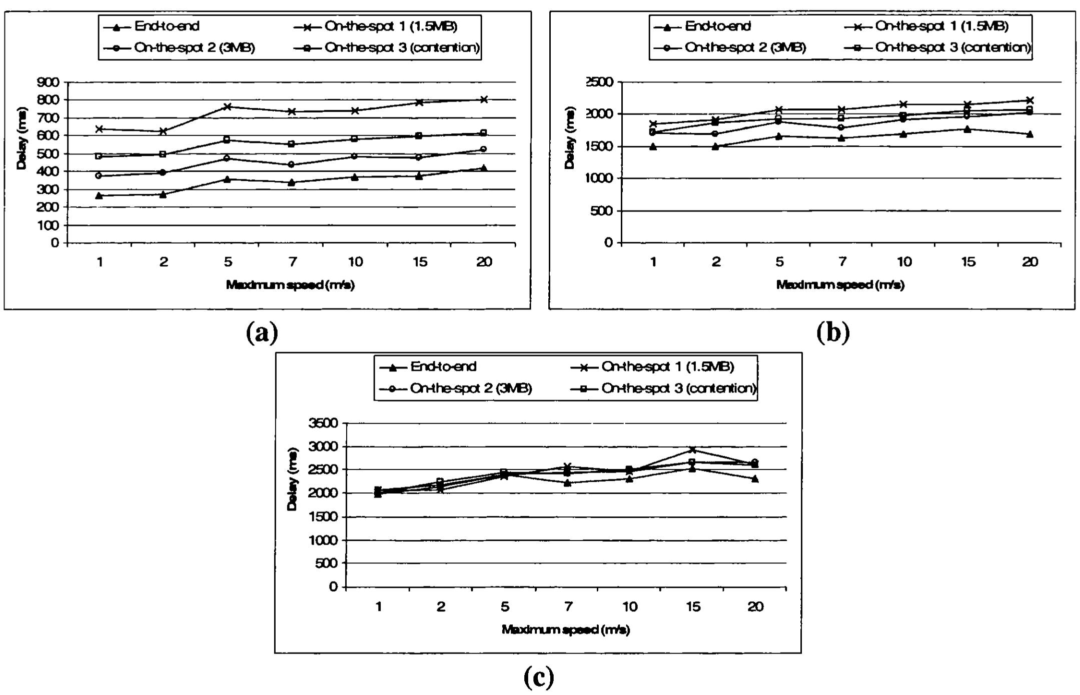
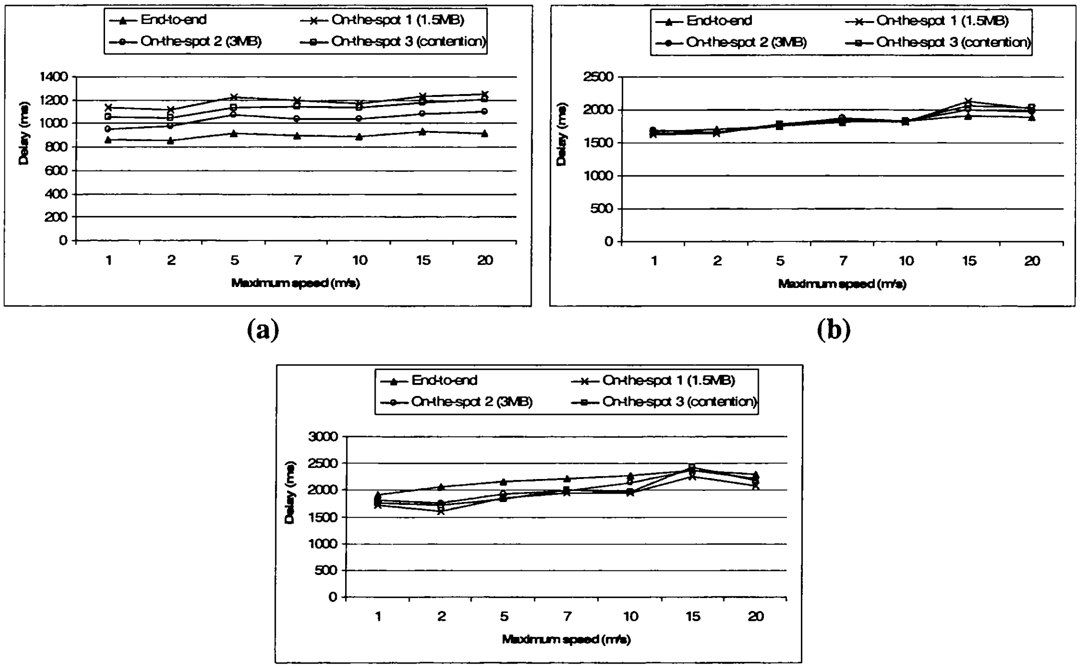


Figure 5.19: End-to-end delay of the end-to-end load density, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Nonetheless, the end-to-end technique still outperforms its on-the-spot counterpart. In contrast in Figure 5.19(c) as we increase the number of flows to 30, the techniques show comparable performance.



(c)

Figure 5.20: End-to-end delay of the end-to-end load density, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 5.20 demonstrates the end-to-end delay by techniques under 10, 20 and 30 flows, and when the traffic rate is increased to 8 packets/s. Figure 5.20(a) demonstrates that while the end-to-end load density outperforms the on-the-spot schemes, on-the-spot 2 (with load history threshold 3MB) outperforms the other schemes. However, as the number of flows is increased to 20, in Figure 5.20(b), end-to-end load density and on-the-spot schemes show comparable performance. In contrast when the number is increased to 30, in Figure 5.20(c), on-the-spot load density schemes outperform the end-to-end load density, with a difference reached the limit of 270ms. However, the figure also shows that on-the-spot 1 and on-the-spot 3 outperform on-the-spot 2, with an average difference of 115ms and 50ms, respectively. Nonetheless we should

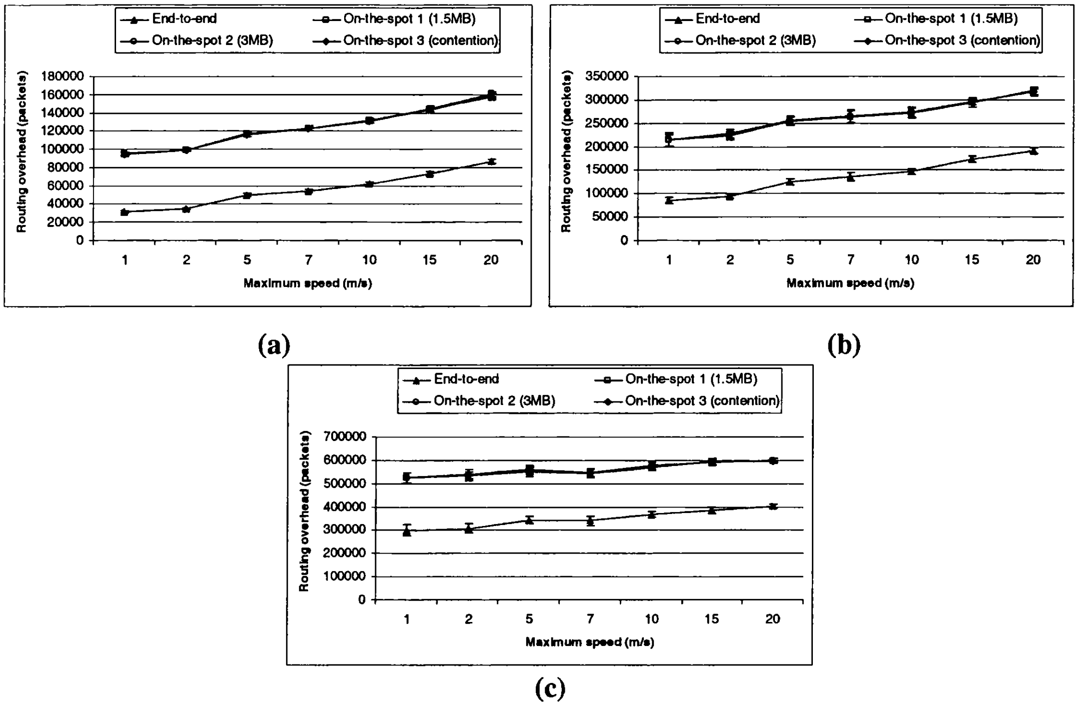


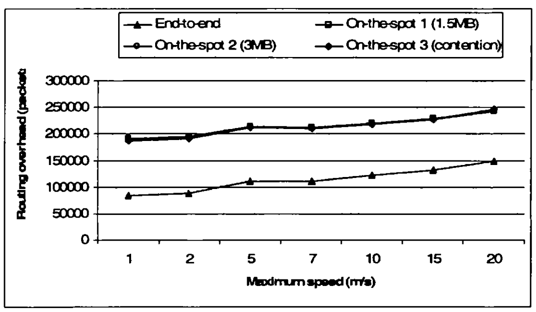
Figure 5.21: Routing overhead of the end-to-end load density, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

bear in mind that end-to-end load density has achieved higher throughput, and hence dropped less packets compared to on-the-spot schemes, which could justify the higher end-to-end delay.

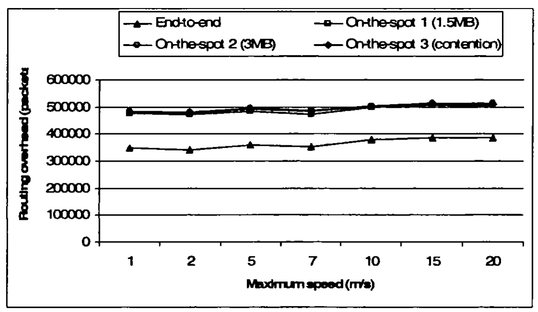
Routing overhead:

Figure 5.21 shows the generated routing overhead by the end-to-end load density and on-the-spot schemes when the traffic rate is 2packets/s, and when number of flows is 10, 20 and 30 flows. Figure 5.21(a) shows comparable routing overhead generated by on-the-spot schemes with an average of around 125,000 routing packets, while the end-to-end technique generate on average about 58,000 routing packets. The figure also reveals that routing overhead increases noticeably as the mobility speed increases. The overhead increases by around 56,000 routing packets for end-to-end, and for on-the-spot schemes by around 64,000 routing packets, at the speed of 20 m/s compared to that at 1 m/s for all techniques.

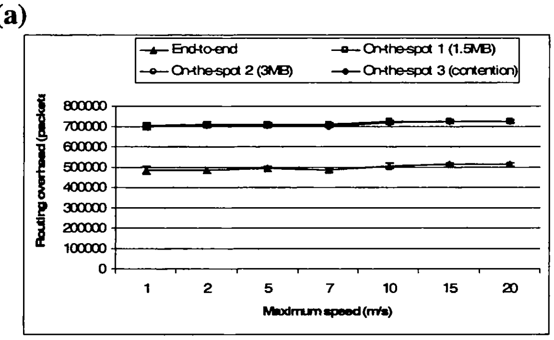
Figure 5.21(b) demonstrates the generated overhead by the techniques when the number of flows is increased to 20 flows. The figure shows an increase in the generated overhead compared to that in Figure 5.21(a), on average, by more than the double. The end-to-end load density generates on average about 136,000 routing packets, while on-the-spot schemes generate on average around 265,000 routing packets. On the other hand, the impact of increasing the mobility speed from 1m/s to 20m/s on the generated overhead reaches the limit of around 105,000 routing packets for all techniques. Similar behaviour is depicted when increasing the number of flows to 30 flows, in Figure 5.21(c). End-to-end load density generates on average about 350,000 routing packets, while on-the-spot load density schemes generate on average around 560,000 routing packets.



(a)



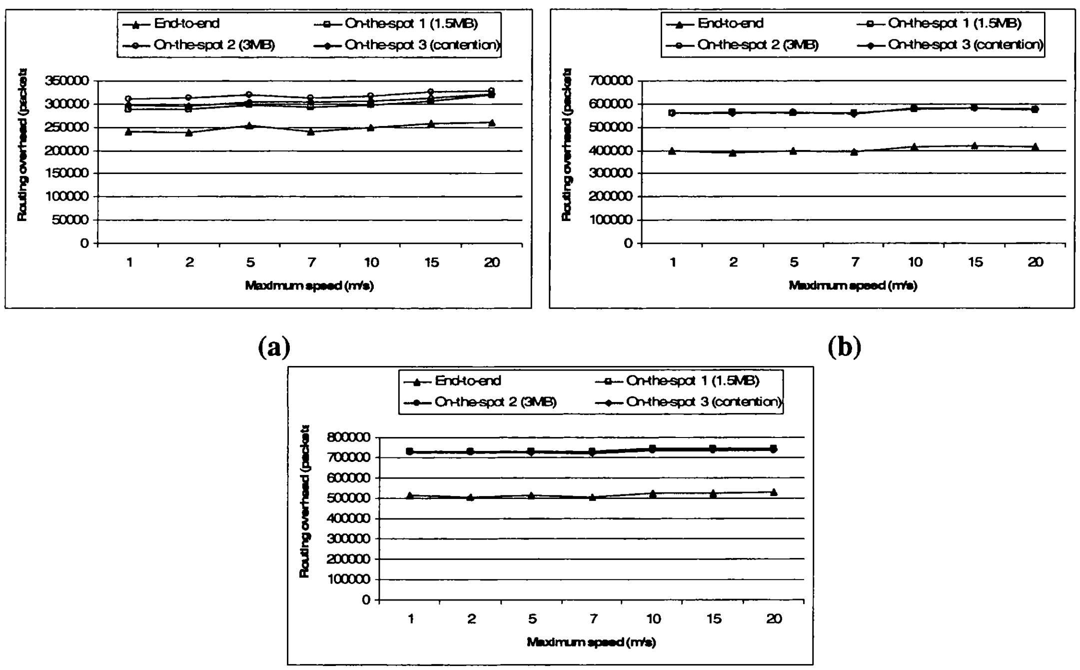
(b)



(c)

Figure 5.22: Routing overhead of the end-to-end load density, on-the-spot 1, on-the-spot 2 and on-the-spot 3 techniques for a network of 75 nodes over topology area 900mx1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 5.22 shows the routing overhead when the traffic rate is increased to 4packets/s, and when number of flows is set to 10, 20 and 30 flows. In Figure 5.22(a) when 10 flows are used, end-to-end load generates on average about 114,000 routing packets, and on-the-pot schemes generates around 215,000 routing packets, and hence the end-to-end technique outperforms the on-the-pot technique with a difference of about 100,000 routing packets. Similarly in Figures 5.22(b, c) under 20 and 30 flows, respectively, end-to-end load density outperforms on-the-spot load density, on average, by around 130,000 and 214,000 routing packets, respectively. In contrast increasing the mobility has shown little impact on the performance of the techniques as we increase the number of flows, the impact of increasing the traffic has overcome the impact of the speed.



(c)

Figure 5.23: Routing overhead of the end-to-end load density, on-the-spot 1 (at 1.5 MB threshold), on-the-spot 2 (at 3 MB threshold) and on-the-spot 3 (contention only) techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

Figure 5.23 shows the routing overhead when the traffic rate is increased to 8 packets/s, and when number of flows is 10, 20 and 30 flows. In Figure 5.22(a) when 10 flows are used, end-to-end load generates on average about 50,000-69,000 routing packets less than that generated by on-the-pot schemes. In contrast in Figures 5.23(b, c) under 20 and 30 flows, respectively, the end-to-end load density outperforms on-the-spot load density, on average, by around 164,000 and 215,000 routing packets, respectively.

Normalised load covariance:

Figure 5.24 shows the normalised load covariance for the end-to-end load density and on-the-spot schemes when the traffic rate is 2packets/s. In Figures 5.24(a, b, c) when number of flows is set to 10, 20 and 30 flows, respectively, the three on-the-spot schemes demonstrate comparable performance, where as the end-to-end load density outperforms the three on-the-spot schemes on average by around 7%, 10% and 29% respectively for 10, 20, and 30 flows.

In contrast in Figure 5.25, when increasing the traffic rate to 4 packets/s, the performance difference between on-the-spot schemes takes shape. This is due to the setting of the load history threshold (1.5MB, 3MB and unlimited). In Figure 5.25(a) when number of flows is set to 10, on-the-spot 2 and on-the-spot 3 show comparable performance, while the two schemes outperform on-the-spot 1 by around 9.5% and 7%, respectively. The end-to-end technique still

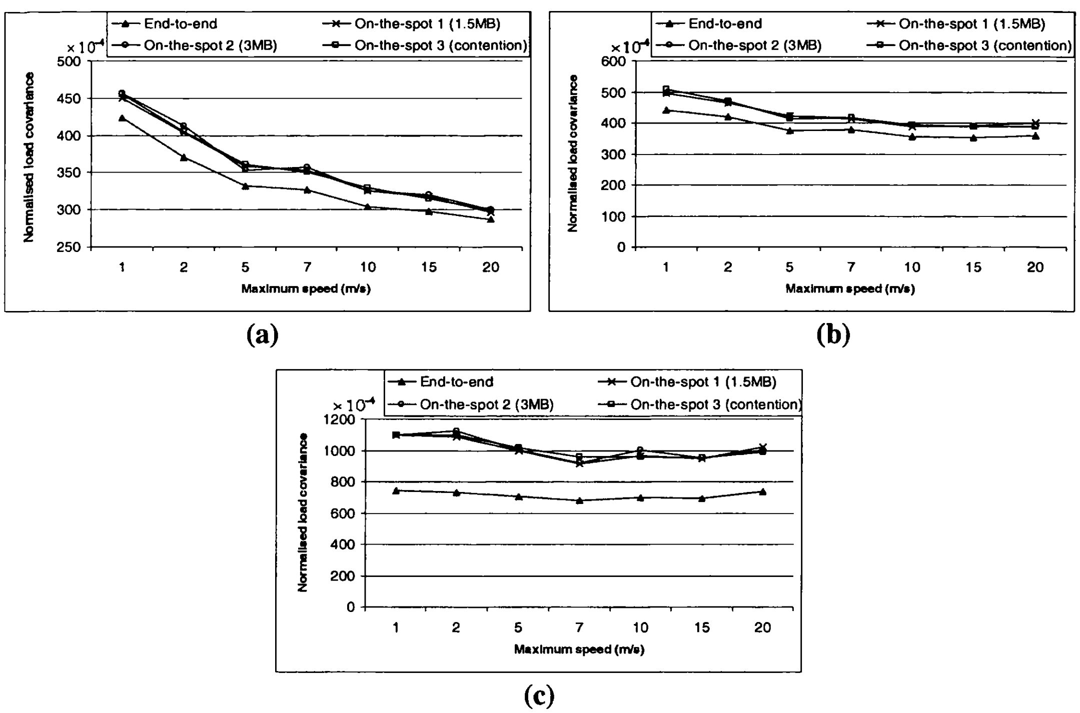


Figure 5.24: Normalised load covariance of the end-to-end load density, on-the-spot 1, on-the-spot 2 and on-the-spot 3 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 2 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows

outperforms the three on-the-spot schemes with difference around 18%, 10%, and 12%, respectively from on-the-spot 1, on-the-spot 2 and on-the-spot 3. Similar results are revealed in Figure 5.25(b) when the number of flows is increased to 20. However when the number of flows is increased to 30 in Figure 5.25(c), on-the-spot schemes show comparable performance, but the end-to-end load density still outperforms the three schemes by around 27%.

Figure 5.26 shows the normalised load covariance for the techniques when the traffic rate is 8 packets/s. The figure reveals that the end-to-end load density outperforms the three schemes of on-the-spot regardless of the number of flows, by about 18%-24% from on-the-spot 2, 21%-25% from on-the-spot 3, and 24%-29% from on-the-spot 1. The figure also reveals that when 10 and 20 flows are used on-the-spot schemes outperform each other, but with a small difference, in the following order: on-the-spot 2, on-the-spot 3, and on-the-spot 1.

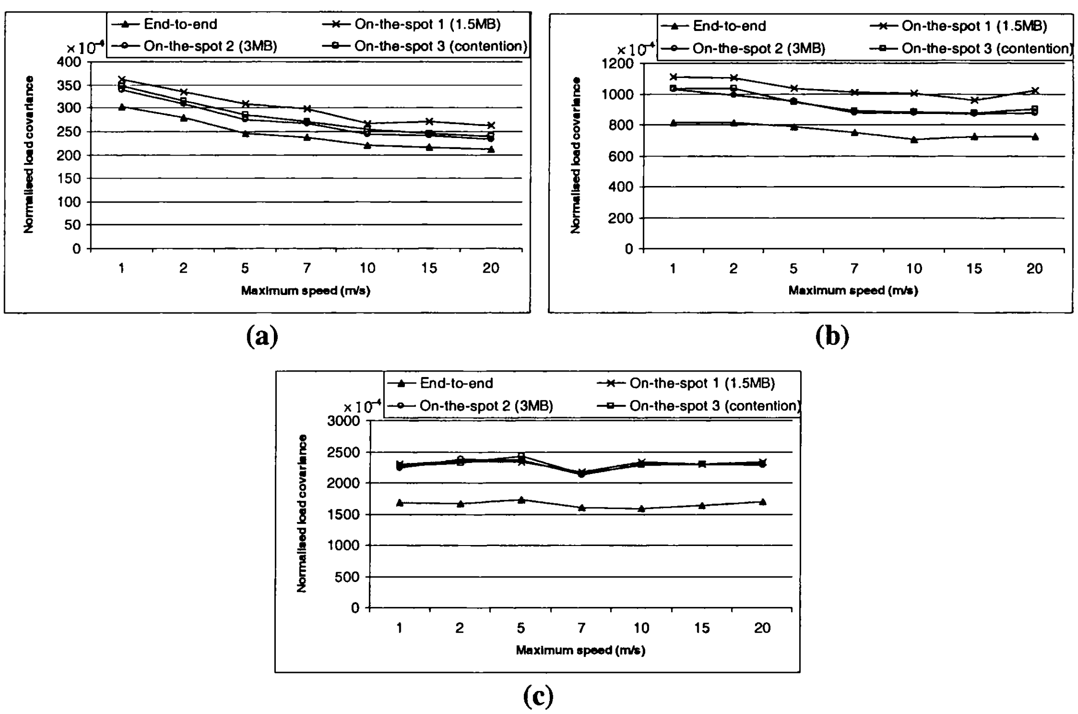
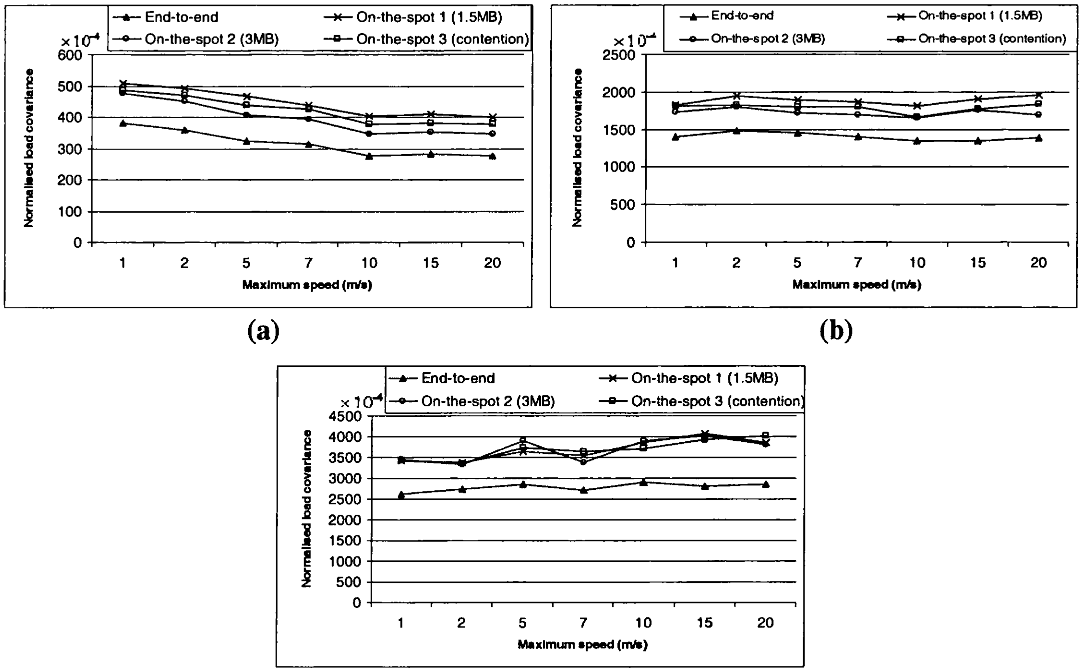


Figure 5.25: Normalised load covariance of the end-to-end load density, on-the-spot 1, on-the-spot 2 and on-the-spot 3 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 4 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows



(a) (b) (c)

Figure 5.26: Normalised load covariance of the end-to-end load density, on-the-spot 1, on-the-spot 2 and on-the-spot 3 techniques for a network of 75 nodes over topology area 900m×1000m with a 512-byte packet and traffic rate of 8 packets/s. (a) traffic flows =10 flows, (b) traffic flows =20 flows, (c) traffic flows =30 flows.

5.4 Conclusions

This chapter has presented an on-the-spot traffic aware technique named here as *on-the-spot load density*, adapted from our end-to-end traffic aware technique, named *load density*. The technique utilises two components in it decision which are: *load history* and *contention*. Some challenges have emerged during the adaptation process represented mainly by setting the load history threshold, so we have used three different settings. We have conducted an extensive evaluation of on-the-spot load density traffic aware technique for the three settings of load threshold and compared it against the end-to-end load density, in order to assess the performance behaviour the two approaches under different working environments. The simulation environments have incorporated various traffic rates and mobility speeds. The first part of the assessment has been carried out through studying the effect of varying the network size in terms of the topology dimensions and the number of nodes deployed over that topology. The second part of the assessment was intended to demonstrate the behaviour of the techniques when varying the number of flows.

In the first part of the assessment which considers the effect of network size, end-to-end load density outperforms the schemes of on-the-spot in terms of throughput, regardless of the network size. In contrast on-the-spot 2 (with a load threshold of 3MB) have shown comparable performance with on-the-spot 1 (with a load threshold of 1.5MB) when the traffic rate is relatively low at 2 packets/s. However at higher rates on-the-spot 2 outperforms on-the-spot 1. For instance, at traffic rate of 8 packets/s with a network size of 75 nodes, on-the-spot 2 achieves 10% of its average throughput higher than on-the-spot 1. Nonetheless, the assessment has shown that throughput decreases as the network size increases form 75 to 100, while the performance difference between end-to-end load density and on-the-spot 2 increases. For instance, the difference between end-to-end technique and on-the-spot 2 rises from an average of 3% to become 8% when the network size increases from 75 to 100 nodes. This is because there are more route breakages in larger networks which have affected the throughput achieved by all techniques, but apparently the end-to-end method uses less congested routes compared to on-the-spot. In terms of end to end delay, the end-to-end load density outperforms on-the-spot 2 on average by 13% for a network of 75 nodes, while for larger network of 100 nodes it goes up to 15%. Moreover, in terms of routing overhead, end-to-end load density generates up to 45% less overhead than that generated by on-the-spot 2. Furthermore, in terms of the normalised load covariance, the end-to-end method has demonstrated a performance advantage over on-the-spot 2 with a difference that reaches 9%.

In the second part of the assessment, varying the number of flows has shown that the end-to-end load density still achieves higher throughput than the on-the-spot schemes with a difference that could be up to 52%. The on-the-spot schemes have shown comparable performance in general. However in the cases where there is a difference, on-the-spot 2 has shown the upper hand over on-the-spot 1 and on-the-spot 3 (with unlimited threshold; i.e. contention only), with a difference approximately 12% and 5%, respectively. In terms of end-to-end delay, it was apparent that increasing the number of flows from 10, 20 to 30 flows increase the end-to-end delay for all techniques, when traffic rates are 2packets/s and 4 packets/s. However it was less noticeable at traffic rate of 8 packets/s for 20 and 30 flows as the network is severely congested. The end-to-end delay in the end-to-end load density could be 27% lower than in the on-the-spot 2. Moreover, in terms of routing overhead, there was a clear advantage in favour of end-to-end load density regardless of number of flows or traffic rates. Furthermore, in terms of the normalised load covariance, load density outperforms on-the-spot 2 by 29%.

To summarise, the end-to-end load density technique has demonstrated a performance advantage over the on-the-spot load density technique in terms of throughput, end-to-end delay, routing overhead and normalised load covariance. The on-the-spot presented in this chapter is one version of on-the-spot load density, further improvements might be possible. Despite the fact that end-to-end load density has shown clear outperfomance, it might be of our interest to further investigate other versions. The subsequent chapter will conclude this thesis and describe potential future research directions.

Chapter 6

Conclusions and future directions

6.1 Introduction

In recent years, MANETs have attracted great interest in the research community [19, 29, 65, 78, 83, 89]. This has been motivated by the advances in wireless devices and the potential applications that could be realised using these networks [65, 83] ranging from common voice conferencing to complicated high-risk military operations. Much research effort [3, 7, 24, 32, 33, 37, 48, 57, 65, 69, 74, 77, 80, 82, 86] has been devoted to tackling a number of challenging issues in order to ensure that this range of applications can run successfully on MANET systems.

One of major challenges in MANETs is the design of an efficient routing protocol that can accommodate their dynamic nature and frequent topology changes; the topology can change unpredictably, so the routing protocol should be able to adapt automatically. To deal with such issues a number of routing algorithms have been proposed [20, 22, 34, 43, 46, 58, 62, 63, 67, 68, 71, 81]. Most of these algorithms have considered finding the shortest path when building a route from source to destination. However, this can lead to some nodes being more overloaded than the others [28]. MANET resources, such as node power and channel bandwidth are often at a premium [5, 12, 16, 19, 52], and therefore it is important to optimise their use as much as possible. Consequently, a traffic aware technique to distribute the load is very desirable in order to make good utilisation of nodes' resources.

The main goal of traffic aware routing is to avoid a situation that could occur in traditional routing, [43, 68] where a small number of nodes have to carry excessive traffic loads compared to the rest. In a MANET, this excessive load would result in high battery consumption at these nodes also in addition to the usual problem of causing some network regions to be much more congested than others. Achieving a good load distribution in MANETs, where nodes in the network have comparable loads, is a challenging task. The issue is complicated not only because a full knowledge of the network status is required in order to make optimal decisions, but also because it has to deal with node mobility. In order to distribute the load as far as possible over the available nodes, a number of traffic-aware techniques [28, 37, 48, 64, 75] have recently been suggested. These can be classified into two categories: *end-to-end* and *on-the-spot*, based on the way they establish and maintain routes between source and destination.

The importance of traffic-aware routing stems from the role that it plays in extending the network's lifetime (or operational time) as it targets the reservation of node's power and better utilisation of channel bandwidth which could lead to performance improvement in terms of important performance metrics including throughput and end-to-end delay.

6.2 Summary of the results

The major focus of this research has been the study, analysis, and development of traffic-aware techniques for reactive routing protocols in MANETs. Summarised below are the major contributions made in this research work.

- Existing end-to-end traffic aware techniques, e.g. [37, 75], estimate load at a given node using either the length of the MAC layer interface queue in the case of traffic density [75] or the number of active traffic flows in degree of nodal activity [37]. The performance merits of these approaches have already been analysed and compared against the traditional routing algorithms, such as AODV and DSR [37, 75]. However, there been no similar study that evaluates and compares the relative performance merits of the suggested end-to-end techniques. Motivated by these observations, in the first part of this dissertation, we have conducted an extensive performance analysis of the two major end-to-end traffic aware techniques, traffic density and nodal activity. Our study is the first comparative analysis to be reported in the literature which highlights the relative performance merits of these approaches under a variety of network operating conditions.

- In order to conduct a suitable performance analysis of traffic aware techniques, we have designed and built a versatile simulation model. Extensive simulation experiments under a variety of traffic conditions and mobility scenarios have been carried out to evaluate the performance of the nodal activity and traffic density methods. The effect of varying the network size in terms of the topology dimensions and the number of nodes deployed over that topology have been investigated, e.g. 75 nodes and 100 nodes over a simulation area of 900m×1000m and 1200m×1000m, respectively, and the mobility speed has been varied from 1m/s to 20m/s. Also, different packet lengths of 128 and 512 bytes have been used to examine the behaviour of the techniques. In such scenarios, our comparative study has revealed that in most of the considered cases traffic density technique outperforms its nodal activity counterpart in both throughput and end-to-end delay. For instance, the throughput in traffic density could be approximately 9% higher than in nodal activity when the network size increases to 100 nodes. Similarly, for the same scenarios, the end-to-end delay in traffic density could be 21% lower than in nodal activity. But, the latter technique exhibits lower routing overhead with a difference up to 12%. This is because nodal activity generates fewer route requests than traffic density. Furthermore, in terms of normalised load covariance, the two techniques exhibit comparable performance, but with a small difference up to around 4% in favour of traffic density.

- In contrast, when the background traffic increases significantly, nodal activity technique outperforms traffic density. This is because when there are a relatively high number of flows in the network there is more chance that the nodes will update their routing table faster since two or more flows could be using paths that have nodes in common. Hence when a route breaks for one flow the route error generated will update other existing stale routes and, consequently, the number of active flows will be more accurate. For instance, the throughput in nodal activity could be 19% higher than in traffic density. The end-to-end delay in nodal activity could be 27% lower than in traffic density. Further, the routing overhead in nodal activity could be 19% lower than in traffic density. The normalised load covariance in nodal activity could be 15% lower than in traffic density. Our findings reveal that it may not be always useful to use feedback from the MAC layer, as in traffic density, to help in the decision of distributing traffic among the network nodes.

- The existing end-to-end traffic aware techniques represented by nodal activity [37] and traffic density [75] suffer from a number of limitations in their judgment of the load, and this is related to estimation of the current traffic load as well as the past experienced traffic load. In order to make a good estimation of the load, it may not be *sufficient* to only capture the number of active paths as in the nodal activity technique or estimate the number of packets at the interface queue over a period of time as in the traffic density technique. The number of active paths could also be useful if the nodes do not move frequently since mobility can affect this value at the intermediate nodes of a given path, leading to inaccurate estimation of the load on that path. The number of packets at the interface queue is useful for capturing contention at the MAC layer. However, this is not sufficient to capture the load experienced by a node over a period of time. In view of these limitations, we have developed a new traffic aware technique, referred to as *load density*, that can overcome the limitations of the existing methods by using a combination of traffic load history and traffic contention, which would reflect the amount of data passed over each node and capture the current degree of contention around the node's neighbourhood.

- Numerous simulation experiments have been carried out in order to compare the performance of our proposed technique against that of the existing traffic density and nodal activity. The simulation environments have incorporated various traffic conditions and different network scenarios, using the same parameter values as above. The simulations have shown that the new load density technique has a performance advantage over traffic density and nodal activity in terms of throughput, end-to-end delay, routing overhead and normalised load covariance. This is in line with the motivation behind the development of our new method and shows that the performance improvement over the existing techniques is present in the various network settings and traffic conditions considered. Our technique has shown to strike a balance between keeping the performance at good levels while distributing the load as much as possible and keeping a low routing overhead. For instance, the throughput in load density could be 23% and 47% higher than in nodal activity and traffic density, respectively, while the normalised load covariance in load density could be up to 16% and 29% less than in nodal activity and traffic density, respectively. The end-to-end delay in load density is 14% and 24% lower than in nodal activity and traffic density, respectively. Further, the routing overhead in load density could be 28% and 34% lower than in nodal activity and in traffic density, respectively.

- Finally, we have investigated the relative performance merits of the end-to-end and local approaches of traffic aware routing in MANETs to determine the situations where each approach could be suitably applied. To ensure a realistic and fair comparison, we have adapted our suggested end-to-end load density technique in order to develop a new "on-the-spot" traffic aware technique, referred to here as *on-the-spot load density.* As with the end-to-end version, on-the-spot load density uses two components to estimate load: *load history* and *contention.* However, the decision of whether to be part of a path is made "on the spot" by the intermediate node as the path is being established, rather by the end nodes; the destination in the case of end-to-end load density, unless that path is the only path available to the destination in which case the node should participate in the path in order to keep good performance levels. Three variations of the on-the-spot load density that use three different settings for the load history threshold have been considered.

- Simulation results show that in most of the scenarios considered the end-to-end load density technique exhibits a performance advantage over the on-the-spot load density technique in terms of throughput, end-to-end delay, routing overhead and normalised load covariance. For instance, the throughput in the end-to-end load density could be 52% higher than in the on-the-spot load density, while the normalised load covariance could be 29% less than in the on-the-spot technique. The end-to-end delay in the end-to-end load density could be 27% lower than in the on-the-spot technique. Further, the routing overhead in load density could be 38% lower than in on-the-spot technique. This is because the on-the-spot technique relies on local decisions taken by intermediate nodes which might not be always good especially if all the potential paths to a destination pass through nodes with declared overload condition in which case an optimal selection of a path could not be feasible as the destination node cannot know which is the best path because all the paths have been marked with an overload condition.

6.3 Directions for the future work

There are several interesting issues and open problems that require further investigation. A selection of these is listed below.

- In this research, we have considered the *random waypoint model* [53] to simulate node mobility. This model has been used widely in the literature [10, 15, 31, 37, 44, 47, 48, 69]. However, there are other mobility models [8, 13, 40, 60] that have been proposed to simulate mobility in MANETs. Although these models are not as widely used as the random waypoint model, it would be interesting to examine the behaviour of traffic aware techniques under different mobility models, such as the random trip model [8] and obstacle mobility model [40].

- The proposed load density technique presented in Chapter 4 has three parameters for selecting routes. This results in three variations in the decision as to selecting routes. In this research, we have only examined one variation that would, in most cases, select the route with the lowest load, with the preference of selecting routes that have contention value below a certain contention threshold. However, other variations are possible and require further investigation. One possible variation would select a route with the lowest contention value among those routes that are within an acceptable load difference from the minimum load threshold.

- Synthetic traffic workload has been used in this work. It is important to study the behaviour of traffic aware techniques using application traces collected from real experiments. The obstacle in applying application traces stems from the lack of such traces. However as more real world MANET experiments are conducted, such traces should become available in the future.

- The performance of traffic aware techniques has been assessed with CBR traffic for direct and fair comparison. A natural extension of this work would analyse the behaviour of these techniques under other traffic types such as TCP and VBR.

- This dissertation has presented extensive analysis of traffic aware techniques based on reactive routing as a starting point. However there are other traffic aware techniques proposed for proactive [76] and hybrid routing [64]. Studying these techniques would be a useful direction for further research work in the future.

- In this research, the performance of traffic aware techniques has been analysed assuming in a homogeneous network in a pure ad hoc mode. It would be interesting to investigate the behaviour of the traffic aware techniques in wireless mesh networks [11, 91] and in heterogeneous networks where the MANET is connected to a wired network [44]. Such investigation may require adapting the existing traffic aware techniques to suit the new environments.

- Most existing studies, including ours, have used simulation in order to evaluate the performance of the algorithms and protocols suggested for MANETs. However, it might not be possible to examine large-scale scenarios (e.g. a MANET with a large number of nodes) using the simulation approach due to time and complexity constraints. Therefore, developing analytical models for MANETs would be desirable as they allow to investigate the performance behaviour of these systems under scenarios (e.g. large networks operating under heavy traffic), which might not be possible to consider by simulations.

- Simulation is a valuable tool for studying MANET systems. However simulation always requires certain assumptions and simplification to be made (e.g. on radio properties and nodes mobility) in order to keep the simulation model's complexity at a manageable level. As a result, however, the model may not capture all the factors that might affect system performance. So far, there has been a little work on the deployment and performance measurement of real world MANETs . In real world experiments the whole system is tested in real world settings. Such experiments can validate the conclusions reached by simulations and also the findings would be particularly valuable for realistic calibration of future simulation models. Provided that sufficient computing resources are available it would be useful and complementary to this to conduct such experiments.

References

[1] "The Network Simulator ns-2," http://www.isi.edu/nsnam/ns, 2006.

[2] "The ns manual," http://www.isi.edu/nsnam/ns/ns-documentation.html, 2006.

[3] M. Abolhasan, T. Wysocki, and E. Dutkiewicz, "A review of routing protocols for mobile ad hoc networks," *Ad Hoc Networks*, vol. 2, no. 1, 2004, pp. 1-22.

[4] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, 2005, pp. 325-349.

[5] R. S. Al-Qassas, A. Al-Ayyoub, and M. Ould-Khaoua, "Bandwidth-efficient routing protocol for mobile ad hoc networks," *IEE Proceedings-Software*, vol. 150, no. 4, 2003, pp. 230-234.

[6] F. Alizadeh-Shabdiz and S. Subramaniam, "Analytical models for single-hop and multi-hop ad hoc networks," *Mobile Networks and Applications*, vol. 11, no. 1, 2006, pp. 75-90.

[7] G. Allard, P. Jacquet, and B. Mans, "Routing in extremely mobile networks," in *Challenges in Ad Hoc Networking*, vol. 197, K. A. Agha, I. G. Lessous, and G. Pujolle, Eds., 2006, pp. 129-138.

[8] J.-Y. L. Boudec and M. Vojnovic, "The random trip model: stability, stationary regime, and perfect simulation," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, 2006, pp. 1153-1166.

[9] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, X. Ya, and Y. Haobo, "Advances in network simulation," *IEEE Computer*, vol. 33, no. 5, 2000, pp. 59-67.

[10] J. Broch, D. A. Maltz, D. B. Johnson, Y. C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," *Proceedings of the Fourth Annual International Conference on Mobile Computing and Networking (MobiCom'98)*, Dallas, TX, 1998, pp. 85-97.

[11] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: commodity multihop ad hoc networks," *IEEE Communications Magazine* vol. 43, no. 3, 2005, pp. 123-131.

[12] S. Cai, L. Wang, and X. Yang, "An ad hoc multicast protocol based on passive data acknowledgement," *Computer Communications*, vol. 27, no. 18, 2004, pp. 1812-1824.

[13] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communication & Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, vol. 2, no. 5, 2002, pp.

483-502.

[14] T. Camp, J. Boleng, B. Williams, L. Wilcox, and W. Navidi, "Performance comparison of two location based routing protocols for ad hoc networks," *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, 2002, pp. 1678-1687.

[15] G. Chakrabarti and S. Kulkarni, "Load balancing and resource reservation in mobile ad hoc networks," *Ad Hoc Networks*, vol. 4, no. 2, 2006, pp. 186-203.

[16] S. Chakrabarti and A. Mishra, "QoS issues in ad hoc wireless networks," *IEEE Communications Magazine*, vol. 39, no. 2, 2001, pp. 142-148.

[17] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000)*, 2000, pp. 22-31.

[18] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, "Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks " *Proceedings of International Conference on Mobile Computing and Networking (MobiCom'2001)*, 2001, pp. 85-96.

[19] W. Chen, N. Jain, and S. Singh, "ANMP: Ad hoc Network Management Protocol," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, 1999, pp. 1506-1531.

[20] C.-C. Chiang, "Routing in clustered multihop, mobile wireless networks with fading channel," *Proceedings of IEEE Singapore International Conference on Networks (SICON'97)*, 1997, pp. 197-211.

[21] I. Chlamtac, M. Conti, and J. J. N. Liu, "Mobile ad hoc networking: Imperatives and challenges," *Ad Hoc Networks*, vol. 1, no. 1, 2003, pp. 13-64.

[22] T. Clausen and P. Jacquet, "Optimized link state routing protocol," IETF MANET, Internet Draft, Jul. 2003.

[23] R. de Oliveira and T. Braun, "A dynamic adaptive acknowledgment strategy for TCP over multihop wireless networks," *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005)*, 2005, pp. 1863-1874.

[24] S. Doshi, S. Bhandare, and T. X. Brown, "An on-demand minimum energy routing protocol for a wireless ad hoc network," *ACM Mobile Computing and Communications Review*, vol. 6, no. 3, 2002, pp. 50-66.

[25] H. Elaarag, "Improving TCP performance over mobile networks," *ACM Computing Surveys*, vol. 34, no. 3, 2002, pp. 357-374.

[26] M. Elias M. Campista, I. M. Moraes, P. Miguel Esposito, A. Amodei Jr, D. De O. Cunha, L. Henrique M.K. Costa, and O. Carlos M.B. Duarte, "The ad hoc return channel: a low-cost solution for Brazilian interactive digital TV," *IEEE Communications Magazine*, vol. 45, no. 1, 2007, pp. 136-143.

[27] L. R. Ford, Jr., and D. R. Fulkerson, *Flows in Networks*. Princeton University Press: Priceton N.J, 1962.

[28] J. Gao and L. Zhang, "Load balanced short path routing in wireless networks," *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, 2004, pp. 1098-1107.

[29] L. M. Gavrilovska and V. M. Atanasovski, "Ad hoc networking towards 4G: Challenges and QoS solutions," *7th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services*, 2005, pp. 69-80.

[30] J. Geier, *Wireless LANs*, 2nd ed. Sams Publishing, 2001.

[31] M. Gerla and S.-J. Lee, "Dynamic load-aware routing in ad hoc networks," *Proceedings of the IEEE International Conference on Communications (ICC)*, Helsinki, Finland, 2001, pp. 3206-3210.

[32] C. Gui and P. Mohapatra, "Overlay multicast for MANETs using dynamic virtual mesh," *Wireless Networks*, vol. 13, no. 1, 2007, pp. 77-91.

[33] M. Guizani, "Security and trust in mobile ad hoc networks," *Proceedings - 4th Annual Communication Networks and Services Research Conference, CNSR 2006*, 2006, pp. 4-5.

[34] Z. Haas and M. Pearlman, "The performance of query control schemes for the zone routing protocol," *IEEE/ACM Transactions on Networking*, vol. 9, no. 4, 2001, pp. 427-438.

[35] M. Haenggi and D. Puccinelli, "Routing in ad hoc networks: a case for long hops," *IEEE Communications Magazine*, vol. 43, no. 10, 2005, pp. 93-101.

[36] Y. Hao, J. Shu, M. Xiaoqiao, and L. Songwu, "SCAN: self-organized network-layer security in mobile ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 2, 2006, pp. 261-273.

[37] H. Hassanein and A. Zhou, "Load-aware destination-controlled routing for MANETs," *Computer Communications*, vol. 26, no. 14, 2003, pp. 1551-1559.

[38] L. Hogie, P. Bouvry, and F. Guinand, "An overview of MANETs simulation," *Electronic Notes in Theoretical Computer Science*, vol. 150, no. 1, 2006, pp. 81-101.

[39] IEEE Standards Association, "IEEE standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," IEEE, 1999.

[40] A. Jardosh, E. M. Belding-Royer, K. C. Almeroth, and S. Suri, "Towards realistic mobility models for mobile ad hoc networks," *Proceedings of the 9th Annual international Conference on Mobile Computing and Networking*, San Diego, CA, USA, 2003, pp. 217-229.

[41] M. Joa-Ng and I.-T. Lu, "Peer-to-peer Zone-based Two-level Link State Routing for Mobile Ad Hoc Networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, Aug 1999, pp. 1415-1425.

[42] D. B. Johnson, "Validation of wireless and mobile network models and simulation," *Proceedings of the DARPA/NIST Workshop on Validation of Large-Scale Network Models and Simulation*, Fairfax, VA, 1999, pp.

[43] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks,"

in *Mobile Computing, Chapter 5*, T. Imielinski and H. Korth, Eds.: Kluwer Publishing Company, 1996, pp. 153-181.

[44] S. Joo-Han, V. W. S. Wong, and V. C. M. Leung, "Efficient on-demand routing for mobile ad hoc wireless access networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 7, 2004, pp. 1374-1383.

[45] A. Kamerman and L. Monteban, "WaveLAN-II: A high-performance wireless LAN for the unlicensed band," *Bell Labs technical journal*, vol. 2, no. 3, 1997, pp. 118-133

[46] Y. Ko and N. H. Vaidya, "Location-Aided Routing (LAR) in mobile ad hoc networks," *Wireless Networks*, vol. 6, no. 4, 2000, pp. 307-321.

[47] S. Kurkowski, T. Camp, and M. Colagrosso, "MANET simulation studies: the incredibles," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 9, no. 4, 2005, pp. 50-61.

[48] S.-B. Lee, J. Cho, and A. T. Campbell, "A hotspot mitigation protocol for ad hoc networks," *Ad Hoc Networks*, vol. 1, no. 1, 2003, pp. 87-106.

[49] B. M. Leiner, D. L. Nielson, and F. A. Tobagi, "Issues in packet radio network design," *Proceedings of the IEEE*, vol. 75, no. 1, 1987, pp. 6-20.

[50] J. Li, C. Blake, D. S. J. D. Couto, H. I. Lee, and R. Morris, "Capacity of ad hoc wireless networks," *Proceedings of the 7th annual international conference on Mobile computing and networking* Rome, Italy, 2001, pp. 61-69.

[51] Q. Li, J. Aslam, and D. Rus, "Online power-aware routing in wireless ad-hoc networks," *Proceedings of International Conference on Mobile Computing and Networking (MobiCom'2001)*, 2001, pp. 97–107.

[52] S. Lim, W.-C. Lee, G. Cao, and C. R. Das, "A novel caching scheme for improving Internet-based mobile ad hoc networks performance," *Ad Hoc Networks*, vol. 4, no. 2, 2006, pp. 225-239.

[53] G. Lin, G. Noubir, and R. Rajmohan, "Mobility models for ad hoc network simulation," *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004)*, 2004, pp. 454-463.

[54] M. K. Marina, "The ns-2 code for AOMDV, http://www.cs.sunysb.edu/~mahesh/aomdv/," 2004.

[55] M. K. Marina and S. R. Das, "On-demand multipath distance vector routing in ad hoc networks," *International Conference on Network Protocols (ICNP)*, 2001, pp. 14-23.

[56] J. Moy, "OSPF Version 2, RFC 2178 (draft standard)," Internet Engineering Task Force July 1997.

[57] C. S. R. Murthy and B. S. Manoj, *Ad Hoc Wireless Networks: Architectures and Protocols* Prentice Hall PTR: Upper Saddle River, N.J., 2004.

[58] S. Murthy and J. J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks," *ACM Mobile Networks and Application , Special Issue on Routing in Mobile Communication Networks*, vol. 1, no. 2, 1996, pp. 183-197.

[59] V. Naumov and T. Gross, "Scalability of routing methods in ad hoc networks," *Performance Evaluation*, vol. 62, no. 1-4, 2005, pp. 193-209.

[60] J. M. Ng and Z. Yan, "Reference region group mobility model for ad hoc networks," *Second IFIP International Conference on Wireless and Optical Communications Networks (WOCN 2005)*, 2005, pp. 290-294.

[61] N. Nikaein, W. Shiyi, and C. Bonnet, "HARP: Hybrid Ad hoc Routing Protocol," *International Symposium on Telecommunications*, 2001, pp.

[62] R. Ogier, F. Templin, and M. Lewis, "Topology Dissemination Based on Reverse-Path Forwarding (TBRPF)," IETF MANET, Internet Draft, Oct. 2003.

[63] V. D. Park and M. S. Corson, "A Highly Adaptive Distributed Routing Algorithm for Mobile Wireless Networks," *Proceedings of the 16th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 1997)*, Kobe, Japan, 1997, pp. 1405-1413.

[64] M. R. Pearlman, J. H. Zygmunt, S. Peter, and S. T. Siamak, "On the impact of alternate path routing for load balancing in mobile ad hoc networks," *Proceedings of the 1st ACM international symposium on Mobile ad hoc networking & computing (MobiHoc)*, Boston, Massachusetts, 2000, pp. 3-10.

[65] C. E. Perkins, *Ad hoc Networking*. Addison-Wesley: Boston, 2001.

[66] C. E. Perkins, E. M. Belding-Royer, and S. R. Das, "Ad hoc On-Demand Distance Vector (AODV) Routing, RFC 3561," Internet Engineering Task Force July 2003.

[67] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance Vector Routing (DSDV) for Mobile Computers," *Proceedings ACM SIGCOMM'94*, London, 1994, pp. 234-244.

[68] C. E. Perkins and E. M. Royer, "Ad-hoc On-demand Distance Vector Routing," *Proceedings 2nd IEEE Workshop on Mobile Computing Systems and Applications*, New Orleans, LA, 1999, pp. 90-100.

[69] C. E. Perkins, E. M. Royer, S. R. Das, and M. K. Marina, "Performance comparison of two on-demand routing protocols for ad hoc networks," *IEEE Personal Communications, [see also IEEE Wireless Communications]*, vol. 8, no. 1, 2001, pp. 16-28.

[70] R. Perlman, *Interconnections: Bridges and Routers*. Addison-Wesley: Reading, Massachusetts, 1992.

[71] V. Ramasubramanian, Z. J. Haas, and E. G. Sirer, "SHARP: A Hybrid Adaptive Routing Protocol for Mobile Ad Hoc Networks," Annapolis, Maryland, USA, 2003, pp. 303 - 314.

[72] T. S. Rappaport, *Wireless communications, principles and practice*. Prentice Hall, 1996.

[73] E. M. Royer and C.-K. Toh, "A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks," *IEEE Personal Communications* vol. 6, no. 2, 1999, pp. 46-55.

[74] P. M. Ruiz, F. J. Ros, and A. Gomez-Skarmeta, "Internet connectivity for mobile ad hoc

networks: solutions and challenges," *IEEE Communications Magazine*, vol. 43, no. 10, 2005, pp. 118-125.

[75] V. Saigal, A. K. Nayak, S. K. Pradhan, and R. Mall, "Load balanced routing in mobile ad hoc networks," *Computer Communications*, vol. 27, no. 3, 2004, pp. 295-305.

[76] B. Sangman, "A hierarchical routing method for load-balancing,' in *High Performance Computing - HiPC 2003*, vol. 2913/2003: Springer Berlin / Heidelberg, 2003, pp. 142-151.

[77] K. Sanzgiri, B. Dahill, B. N. Levine, C. Shields, and E. M. Belding-Royer, "A Secure Routing Protocol for Ad Hoc Networks," *Proceedings of the 10th IEEE International Conference on Network Protocols*, 2002, pp. 78-89.

[78] J. Saurabh and D. P. Agrawal, "Wireless community networks," *IEEE Computer*, vol. 36, no. 8, 2003, pp. 90-92.

[79] R. Sollacher, M. Greiner, and I. Glauche, "Impact of interference on the wireless ad-hoc networks capacity and topology," *Wireless Networks*, vol. 12, no. 1, 2006, pp. 53-61.

[80] X. Tao, T. Kunz, and D. Falconer, "Throughput maximizing routing in a MANET: Protocol and analysis," *Proceedings International Conference on Wireless Networks, Communications and Mobile Computing*, 2005, pp. 175-180.

[81] C.-K. Toh, "Associativity based routing for ad hoc mobile networks," *Wireless Personal Communications Journal, Special Issue on Mobile Networking and Computing Systems*, vol. 4, no. 2, 1997, pp. 103-139.

[82] C.-K. Toh, "Maximum battery life routing to support ubiquitous mobile computing in wireless ad hoc networks," *IEEE Communications*, vol. 39, no. 6, 2001, pp. 138-147.

[83] C.-K. Toh, *Ad Hoc Mobile Wireless Networks: Protocols and Systems*. Prentice Hall: Upper Saddle River, NJ, December 2001.

[84] D. A. Tran and H. Raghavendra, "Congestion Adaptive Routing in Mobile Ad Hoc Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 11, 2006, pp. 1294-1305.

[85] A. C. Viana, M. D. de Amorim, S. Fdida, and J. F. de Rezende, "Self-organization in spontaneous networks: the approach of DHT-based routing protocols," *Ad Hoc Networks*, vol. 3, no. 5, 2005, pp. 589-606.

[86] X. Wei-Qiang and W. Tie-Jun, "TCP Issues in Mobile Ad Hoc Networks: Challenges and Solutions," *Journal of Computer Science and Technology*, vol. 21, no. 1, 2006, pp. 72-81.

[87] Y. Wei, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, 2002, pp. 1567-1576.

[88] K. Woo, C. Yu, D. Lee, H. Y. Youn, and B. Lee, "Non-blocking, localized routing algorithm for balanced energy consumption in mobile ad hoc networks," *Proceedings of Int Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS 2001)*, 2001, pp. 117–124.

[89] T.-Y. Wu, C.-Y. Huang, and H.-C. Chao, "A survey of mobile IP in cellular and mobile ad-hoc network environments," *Ad Hoc Networks*, vol. 3, no. 3, 2005, pp. 351-370.

[90] A. Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad hoc routing," *Proceedings of the 7th annual international conference on Mobile computing and networking*, Rome, Italy, 2001, pp. 70-84.

[91] W. Xudong, C. Sunghyun, and J. P. Hubaux, "Wireless mesh networking: theories, protocols, and systems," *IEEE Wireless Communications, [see also IEEE Personal Communications]*, vol. 13, no. 2, 2006, pp. 8-9.

[92] Z. Ye, S. V. Krishnamurthy, and S. K. Tripathi, "A framework for reliable routing in mobile ad hoc networks," *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003)*, 2003, pp. 270-280.

[93] S. Yi, P. Naldurg, and R. Kravets, "Security-aware ad hoc routing for wireless networks," *Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*, Long Beach, CA, USA, 2001, pp. 299-302.

[94] C. Yu, B. Lee, and H. Y. Youn, "Energy efficient routing protocols for mobile ad hoc networks," *Wireless Communications and Mobile Computing*, vol. 3, no. 8, 2003, pp. 959-973.

[95] Y. Yuan, H. Chen, and M. Jia, "An adaptive load-balancing approach for ad hoc networks," *Proceedings International Conference on Wireless Communications, Networking and Mobile Computing*, 2005, pp. 743-746.

[96] L. Zhang, Z. Zhao, Y. Shu, L. Wang, and O. W. W. Yang, "Load balancing of multipath source routing in ad hoc networks," *IEEE International Conference on Communications(ICC)*, 2002, pp. 3197-3201.

Publications during the course of this research

- R. S. Al-Qassas, M. Ould-Khaoua and L.M. Mackenzie, Performance Evaluation of a New End-to-End Traffic-Aware Routing in MANETs, Proc. ICPADS'06, vol. 2, IEEE Computer Society Press, 2006, pp. 49-54.

- R. S. Al-Qassas, L.M. Mackenzie and M. Ould-Khaoua, Performance Study of End-to-End Traffic-Aware Routing, Proc. 39th Annual Simulation Symposium, IEEE Computer Society Press, SCS, Huntsville, AL, April 2 - 6, 2006, pp. 26-33.

- R. S. Al-Qassas, M. Ould-Khaoua and L.M. Mackenzie, A New End-to-End Traffic-Aware Routing for MANETs, Proc. 21st UK Performance Engineering Workshop, University of Newcastle, July, 2005, pp.81-90, ISSN 1368-1060.

- R. S. Al-Qassas, M. Ould-Khaoua and L.M. Mackenzie, Performance Study of Traffic-Aware Routing in Mobile Ad hoc Networks, Proc. 6th Annual Network Symposium, Liverpool John Moores University, EPSRC, UK, June 2005, pp. 279-284.

- R. S. Al-Qassas, M. Ould-Khaoua and L.M. Mackenzie, Simulation study of traffic-aware routing in mobile ad hoc networks, Technical Report; TR-2005-193, Department of Computing Science, University of Glasgow, April 2005.

- R. S. Al-Qassas and M. Ould-Khaoua, Proactive routing in mobile ad hoc networks, Proc. 20th UK Performance Engineering Workshop, University of Bradford, July 7-8, 2004, pp. 330-339, ISBN: 0-95410000-3-4.

- R. S. Al-Qassas, A. Al-Ayyoub, and M. Ould-Khaoua, Bandwidth-efficient routing protocol for mobile ad hoc networks, IEE Proc.-Software, vol. 150, no. 4, 2003, pp. 230-234.