2007

Resource-efficient strategies for mobile ad-hoc networking

Li, Zhuoqun

http://hdl.handle.net/10026.1/325

http://dx.doi.org/10.24382/1367
University of Plymouth

RESOURCE-EFFICIENT STRATEGIES FOR MOBILE AD-HOC NETWORKING

Z. Li

Ph.D. November 2007

RESOURCE-EFFICIENT STRATEGIES FOR MOBILE AD-HOC NETWORKING

by

ZHUOQUN LI

A thesis submitted to the University of Plymouth
in partial fulfillment for the degree of

DOCTOR OF PHILOSOPHY

School of Computing, Communications and Electronics
Faculty of Technology

Sponsored by the BIOPATTERN EU Network of Excellence (EU Contract 508803)

November 2007

Resource-Efficient Strategies for Mobile Ad-hoc Networking

Zhuoqun Li

Abstract

The ubiquity and widespread availability of wireless mobile devices with ever increasing inter-connectivity (e.g. by means of Bluetooth, WiFi or UWB) have led to new and emerging next generation mobile communication paradigms, such as the Mobile Ad-hoc NETworks (MANETs). MANETs are differentiated from traditional mobile systems by their unique properties, e.g. unpredictable nodal location, unstable topology and multi-hop packet relay. The success of on-going research in communications involving MANETs has encouraged their applications in areas with stringent performance requirements such as the e-healthcare, e.g. to connect them with existing systems to deliver e-healthcare services anytime anywhere. However, given that the capacity of mobile devices is restricted by their resource constraints (e.g. computing power, energy supply and bandwidth), a fundamental challenge in MANETs is how to realize the crucial performance/Quality of Service (QoS) expectations of communications in a network of high dynamism without overusing the limited resources.

A variety of networking technologies (e.g. routing, mobility estimation and connectivity prediction) have been developed to overcome the topological instability and unpredictability and to enable communications in MANETs with satisfactory performance or QoS. However, these technologies often feature a high consumption of power and/or bandwidth, which makes them unsuitable for resource constrained handheld or embedded mobile devices. In particular, existing strategies of routing and mobility characterization are shown to achieve fairly good performance but at the expense of excessive traffic overhead or energy consumption. For instance, existing hybrid routing protocols in dense MANETs are based in two-dimensional organizations that produce heavy proactive traffic. In sparse MANETs, existing packet delivery strategy often replicates too many copies of a packet for a QoS target. In addition, existing

tools for measuring nodal mobility are based on either the GPS or GPS-free positioning systems, which incur intensive communications/computations that are costly for battery-powered terminals. There is a need to develop economical networking strategies (in terms of resource utilization) in delivering the desired performance/soft QoS targets.

The main goal of this project is to develop new networking strategies (in particular, for routing and mobility characterization) that are efficient in terms of resource consumptions while being effective in realizing performance expectations for communication services (e.g. in the scenario of e-healthcare emergency) with critical QoS requirements in resource-constrained MANETs.

The main contributions of the thesis are threefold:

(1) In order to tackle the inefficient bandwidth utilization of hybrid service/routing discovery in dense MANETs, a novel "track-based" scheme is developed. The scheme deploys a one-dimensional track-like structure for hybrid routing and service discovery. In comparison with existing hybrid routing/service discovery protocols that are based on two-dimensional structures, the track-based scheme is more efficient in terms of traffic overhead (e.g. about 60% less in low mobility scenarios as shown in Fig. 3.4). Due to the way "provocative tracks" are established, the scheme has also the capability to adapt to the network traffic and mobility for a better performance.

(2) To minimize the resource utilization of packet delivery in sparse MANETs where wireless links are intermittently connected, a store-and-forward based scheme, "adaptive multi-copy routing", was developed for packet delivery in sparse mobile ad-hoc networks. Instead of relying on the source to control the delivery overhead as in the conventional multi-copy protocols, the scheme allows each intermediate node to independently decide whether to forward a packet according to the soft QoS target and local network conditions. Therefore, the scheme can adapt to varying networking situations that cannot be anticipated in conventional source-defined strategies and deliver packets for a specific QoS targets using minimum traffic overhead.

(3) The important issue of mobility measurement that imposes heavy communication/computation burdens on a mobile is addressed with a set of resource-efficient "GPS-free" solutions, which provide mobility characterization with minimal resource utilization for ranging and signalling by making use of the information of the time-varying ranges between neighbouring mobile nodes (or groups of mobile nodes). The range-based solutions for mobility characterization consist of a new mobility metric for network-wide performance measurement, two velocity estimators for approximating the inter-node relative speeds, and a new scheme for characterizing the nodal mobility. The new metric and its variants are capable of capturing the mobility of a network as well as predicting the performance. The velocity estimators are used to measure the speed and orientation of a mobile relative to its neighbours, given the presence of a *departing node*. Based on the velocity estimators, the new scheme for mobility characterization is capable of characterizing the mobility of a node that are associated with topological stability, i.e. the node's speeds, orientations relative to its neighbouring nodes and its past epoch time.

Author's Declaration

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award.

This study was funded in part by the BIOPATTERN EU Network of Excellence (EU Contract 508803).

Publications:

1. Z. Li, L. Sun and E. C. Ifeachor, Method and Apparatus for Determining the Speed and Orientation of Networked Mobile Stations, UK Patent filed on June 2007

2. Z. Li, L. Sun and E. C. Ifeachor, Energy-Efficient Packet Relaying in Sparse Mobile Sensor Networks, in Handbook of Wireless Mesh & Sensor Networking, George Aggelou (Editor), accepted for publication by McGraw-Hill.

3. Z. Li, L.Sun and E. C. Ifeachor, GPS-free Mobility Metrics for Mobile Ad Hoc Networks, IET Proceedings on Communications, vol.1, issue 5, pp. 970-976, October 2007.

4. Z. Li, L. Sun and E. C. Ifeachor, Range-based Relative Velocity Estimators for Networked Mobile Devices, Submitted to IEEE Transactions on Vehicular Technology (revised, under the second round review).

5. Z. Li, L. Sun and E. C. Ifeachor, Range-based Mobility Estimations in MANETs with Application to Link Availability Prediction, in Proceedings of IEEE International Conference on Communications (IEEE ICC 2007), Glasgow, UK, June 2007.

6. Z. Li, L.Sun and E. C. Ifeachor, Adaptive Multi-Copy Routing for Intermittently Con-

nected Mobile Ad Hoc Networks, in Proceedings of IEEE GlobeCom Technical Conference (IEEE GlobeCom 2006), San Francisco, USA, November 2006.

7. Z. Li, L.Sun and E. C. Ifeachor, Track-based Hybrid Service Discovery in Mobile Ad-hoc Networks, in Proceedings of the 16th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications (IEEE PIMRC 2005), Berlin, Germany, September 2005.

8. Z. Li, L.Sun and E. C. Ifeachor, Challenges of Mobile Ad-hoc Grids and their Applications in e-Healthcare, in Proceedings of Second International Conference on Computational Intelligence in Medicine and Healthcare (CIMED' 2005), Lisbon, Portugal, July 2005.

9. Z. Li, L. Sun, Z. Qiao and E. C. Ifeachor, Perceived Speech Quality Driven Retransmission Mechanism for Wireless VoIP, in Proceedings of IEE Fourth International Conference on 3G Mobile Communication Technologies, London, UK, June 2003, pp. 395 – 399.

Signed .....

Date 27/11/2007

Word Count *50440*

Acknowledgments

This thesis is the account of three years of devoted work in the field of mobile ad-hoc networking at the University of Plymouth, UK, which would not have been possible without the support and guidance of many.

First, I would like to express my sincere gratitude to my supervisor and director of studies, Professor Emmanuel C. Ifeachor, for the great freedom that he offered me to pursue my research interests, for his continuous encouragement throughout this project, and for his extensive knowledge, patient instructions, critical comments that have helped me to become an independent researcher. I would also like to thank my second supervisor Dr. Lingfen Sun, for her support and guidance during my M.Res and Ph.D studies, for the tremendous amount of time and efforts she has spent in discussing my work with me and in improving the quality of my papers and this thesis.

I would like to thank Mr Zizhi Qiao, Dr. Cindy S.F. Goh, Dr Brahim Hamadicharef and other wonderful colleagues in Signal Processing and Multimedia Communication (SPMC) Research Group, University of Plymouth. It is my great pleasure to have been working with them in the group. I would also like to thank Dr. Xingang Wang for the valuable discussions with him on many research topics. I would like to acknowledge Mr. Ola E. Dallokken and Mr. Muhammad T. Rahman. I have benefited from collaborating with them for their M.sc projects.

Special thanks go to my dear wife, Shaofen, who has given up her job and left her families just for sharing happiness and facing difficulties together with me.

Finally, this thesis is dedicated to my parents and my sister, for their endless love, and their encouraging support for the every important decision that I have made. I hope they will never regret and be proud of me and themselves.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMR	Adaptive Multi-copy Routing
AODV	Ad-hoc On-demand Distance Vector Routing
BER	Bit Error Rate
CBR	Constant Bit Rate
CDF	Cumulated Distribution Function
CPU	Central Processing Unit
CSM	City Section Mobility
DHT	Distributed Hash Table
DSR	Dynamic Source Routing
DTN	Delay Tolerant Network
ECG	Electrocardiogram
FER	Frame Error Rate
GPS	Global Positioning System
GPSR	Greedy Perimeter Stateless Routing
IID	Identical and Independent Distributed
LAR	Location-Aided Routing

LOS	Line-Of-Sight
LPA	Link Prediction Algorithm
MAC	Medium Access Control
MANET	Mobile Ad-hoc Network
MOS	Mean Opinion Score
NBE	Normalized Bias of the speed Estimates
NLOS	Non-Line-Of-Sight
OLSR	Optimized Link-State Routing protocol
PDA	Personal Digital Assistant
QoS	Quality of Service
R2WP	Restricted Random Waypoint mobility
RD	Random Direction mobility
RPGM	Reference Point Group Mobility
RREP	Route Reply
RREQ	Route Request
RSSI	Received Signal Strength Indication
RVE	Range-based relative Velocity Estimator
RWP	Random Waypoint Mobility
SNR	Signal-to-Noise Ratio

TCP	Transmission Control Protocol
TDoA	Time-Difference-of-Arrival
ToA	Time-of-Arrival
UDDI	Universal Description Discovery and Integration
UPNP	Universal Plug And Play
UWB	Ultra-WideBand
WLAN	Wireless Local Area Network
ZRP	Zone Routing Protocol

Chapter 1

Introduction

This Chapter is organized as follows. The motivations of the project are presented in Section 1.1. Section 1.2 gives the research questions. The aims and objectives of the project are outlined in Section 1.3. The main contributions of the project are summarized in Section 1.4. A brief overview and the organization of the thesis are given in Section 1.5.

1.1 Motivations

The ubiquity and widespread availability of wireless mobile devices with ever increasing inter-connectivity (e.g. by means of Bluetooth [1], WiFi [2], WiMax [3] or UWB [4]) has led to new and emerging next generation mobile communication systems such as the mobile ad-hoc networks (MANETs), which are formed by wireless links between moving nodes. Providing satisfactory performance/Quality of Service (QoS) for communications in this type of networks is very challenging due to their unique and special attributes (e.g. unpredictable nodal location, unstable topology and multi-hop packet relay). On-going research on mobile ad-hoc networking has tackled some of the problems and encouraged their applications in areas with challenging performance requirements such as e-healthcare. For example, to connect ad-hoc networks with the Internet to deliver e-healthcare services anytime anywhere. However, given the fact that the capacity of mobile devices is significantly restricted by the resources available (e.g. computing power, storage capability, communication bandwidth and energy supply), a

major challenge that remains in such scenarios is how to provide the best possible and sustainable soft QoS (i.e. to gurantee QoS if networking conditions permit [5]) with a network of resource constrained mobile devices.

The performance of communications in mobile ad-hoc networks have been improved by a family of technologies, which address various aspects of networking such as medium access control, routing, mobility modeling/characterization, topology control, link duration prediction and localization, etc. In accordance with the node density of the network, these technologies have to deploy distinct approaches for dense cases as well as sparse cases.

Routing and mobility characterization are two of the most important aspects of mobile ad-hoc networking. In MANETs, most of the building blocks of the communication system need to be aware of the characteristics of mobile movements in order to improve their effectiveness and adaptivity. And without distributed routing protocols to cope with the frequent topological changes, communications in MANETs would not have been possible. Recent advances in these two areas have brought many solutions that make communications in ad-hoc networks a reality. For example, protocols such as AODV [6], EDSR [7] and ZRP [8] were proposed to route packets in cases of high node density and Epidemic [9], Spray&Wait [10] were developed for networks of very low node density. Mobility models (e.g. Random Waypoint, Random Walk and City Section, etc.) [11] and estimators (e.g. Doppler-frequency-based) have also been developed to capture/predict the characteristics of the movements of mobiles, based on which the stability of a wireless link, a routing path or the performance of communications in a MANET can be evaluated.

However, there still exists a fundamental problem that restricts the application of mobile ad-hoc networks to services with critical QoS requirements. In particular, the building blocks of existing technologies require the consumption of substantial extra system resources (e.g. the consumption of battery energy for wireless transmission/receptions [12] or CPU power for positioning algorithms [13]) for delivering the desired performance and QoS. As a result, they are not sustainable for the ad-hoc networks of handheld devices with significant resource

constraints. For instance, existing hybrid service/routing discovery protocols are based on two-dimensional structures (e.g. zone-based), which shows good performance in discovering a route to a specific host/service but with an expense of significantly increased traffic overhead. In intermittently connected sparse MANETs, the source of a packet tend to increase the *replication factor* of a packet for a delivery delay target when the networking conditions are getting worse. The source-defined delivery strategy is not optimal as it cannot anticipate the networking situations experienced by the relay nodes, which often leads to a waste of bandwidth and energy because of unnecessary packet transmissions. Uneconomical networking strategies can be also found in the area of mobility characterization. In order to determine the nodal mobility, and hence to predict the quality of communications in a network, usually localization systems such as the GPS [14] or GPS-free positioning are deployed to trace nodal locations over time. These methods are able to estimate nodal mobility with satisfying accuracy. But, their operations involve intensive computing/communication activities and, as a result, are too expensive (in terms of energy consumptions) for battery-powered terminals.

In order to provide sustainable QoS-sensitive communications services in the ad-hoc network of resource-constrained mobile devices, it is critically necessary to design new networking technologies (e.g. routing and mobility characterization) that can effectively achieve the performance expectations of the services with efficient utilizations of system resources for their activities.

1.2 Research Questions

According to the expected number of neighbours that are connected to a node at the same time, Mobile Ad-Hoc Networks (MANETs) can be classified as dense MANETs, in which nodes maintain one or more links most of the time, and sparse MANETs, in which nodes have less than one connections on average (e.g. an ad-hoc network of buses [15]). This thesis seeks to address the different research questions and problems that arise in dense and sparse

MANETs:

- In a dense MANET, how can routes to a destination or a specific service be discovered in a way that minimizes traffic overhead without compromising a soft QoS target (e.g. packet delivery latency and success rate)?

  Answer to this question leads to a new hybrid service/routing discovery scheme. The novelty of the scheme is the track-based one-dimensional proactive areas for hybrid routing. The track-based structure of the proposed scheme produces much less proactive traffic than traditional hybrid routing protocols that are based on two-dimensional structures. It has also the capability to adapt the density of proactive tracks to traffic patterns for better efficiency. This work will be discussed in Chapter 3.

- In sparse MANETs that can cover multiple geographical areas of diverse communication conditions, how can packets be delivered efficiently and adaptively to the local networking situations?

  Answer to this question leads to a new *store-and-forward* based packet delivery scheme for sparse mobile ad-hoc networks, Adaptive Multi-copy Routing (AMR). The novelty of the AMR scheme is that, instead of relying on the source, in the scheme intermediate nodes independently constrain packet relaying according to the local network conditions and the soft QoS targets (e.g. the expected end-to-end delay and packet loss rate). By exploiting the diversity of networking conditions of sparse networks, the efficiency (in terms of traffic overhead and energy consumptions) of the AMR scheme is increased significantly. This work is discussed in Chapter 4

- For sparse and dense MANETs, how can the mobility of a single node and the whole network be reliably measured, predicted and characterized in a simple and energy-efficient way?

  Answer to this leads to a set of "range-based" solutions including a new mobility metric

for performance measurement, two estimators for approximations of inter-node relative speeds in sparse/dense networks and a new scheme for mobility characterizations. Previous proposals on mobility estimations are either not energy-efficient as complex localization systems (e.g. GPS) are required or not reliable as precise knowledge of the signal propagation is involved. The range-based GPS-free mobility metric and its variants are described in Chapter 5. The novelty of the proposed solutions is that they use only ranging information between nodes for more efficient and channel-independent mobility estimations. The velocity estimators and a new method for estimating the node density and determining their applicability are presented in Chapter 6. The scheme of range-based mobility characterization and its application in predicting link availability are discussed in Chapter 7

1.3 Project Aim and Objectives

The main aim of this project is to develop new efficient and effective networking strategies (particularly routing and mobility characterization) that can deliver the desired performance/QoS with minimized resource utilizations for performance-critical communications in ad-hoc networks of resource-constrained mobile devices.

Specific objectives of the research are:

- To provide an understanding of the performance of existing ad-hoc networking technologies (e.g. routing and mobility characterization) in terms of QoS benchmarks (e.g. packet delivery ratio and end-to-end latency) as well as their resource utilizations (e.g. bandwidth overhead and energy consumption).

- For ad-hoc networks of both sparse and dense topology, to develop novel resource efficient technologies to discover routes for packet transmission/access to services with expected soft QoS and to characterize the mobility of a mobile node/a whole ad-hoc network.

- To Investigate the applications of these networking technologies for QoS-adaptive communications (e.g. distributed adaptation of packet relay in sparse networks) and QoS predictions (e.g. to predict the loss rate of packet delivery and the availability of wireless links) in MANETs.

1.4 Contributions of Thesis

The project has mainly focused on the investigation of the performance of existing networking technologies of mobile ad-hoc networks in terms of QoS provisioning and resource consumptions, the development of new efficient networking strategies (specifically routing and mobility characterization) that can achieve performance/QoS expectations, as well as their applications in distributed QoS-adaptive communications and performance prediction. Specifically, the contributions of the project are the following:

1. In order to tackle the problem of heavy traffic overhead of hybrid service/routing discovery in dense MANETs, a new "track-based" hybrid service/routing discovery scheme is presented for dense mobile ad-hoc networks. The novelty of the scheme is that it adopts an one-dimensional track-like structure for hybrid routing and service discovery. In comparison with existing schemes that are based on two-dimensional structures, the one-dimensional structure of the scheme significantly reduces the consumption of the system resources (e.g. bandwidth and energy). Due to the way "provocative" tracks are established, the scheme is also able to adapt its behavior (e.g. to be more reactive or proactive) to the network traffic and mobility for a better QoS provisioning. The publication associated with this contribution is [16].

2. To minimize the resource utilization of packet delivery in sparse MANETs where wireless links exhibit intermittent availability, a store-and-forward based scheme, "adaptive multi-copy routing", is presented. Instead of relying on the source to control packet de-

livery as in conventional multi-copy protocols, the proposed scheme shifts the control of multi-hop relaying from the source to intermediate nodes. This method allows the scheme to adapt to varying networking situations that cannot be anticipated in conventional source-defined strategies and to deliver packets for specific QoS (e.g. end-to-end delay and packet loss) targets with minimum traffic overhead. The associated publications are [17] and [18].

3. The important issue of mobility measurement that requires either external facilities or intensive communications/computations is addressed with a new resource efficient "GPS-free" mobility metric that can be used for performance measurement in MANETs. The main novelty of the proposed mobility metric is that it is able to capture the inter-node relative motions in a 2D plane and hence the network average of relative speed, in real-time, without using any costly localization systems. Variants of this metric are presented for prediction of QoS measures (e.g. inter-group inter-meeting time and packet delivery rate) for networks with group and random mobility. A calibration method is also proposed to improve the accuracy of the metrics in noisy environments. The associated publication is [19].

4. In order to achieve resource efficient mobility measurement for a mobile node, two range-based velocity estimators are proposed to estimate the inter-node relative velocity in sparse and dense MANETs. The novelty of the proposed estimators is that the only information required for producing velocity estimates is the time-varying inter-node distances. In contrast to the existing methods, the proposed velocity estimators have some major benefits, such as being robust in noisy conditions, independent of the characteristics of the wireless channel/signal, and energy-efficient as costly positioning systems are not needed. An efficient algorithm for predicting the node density of the network is also given to allow mobiles to determine the applicability of the two estimators. The publication associated with this contribution is [20].

5. A new scheme that is extended from the range-based velocity estimators is presented as a new solution of resource efficient mobility characterization. The scheme is able to characterize the mobility of a node that are associated with the availability of a link, i.e. a mobile node's speeds, orientations relative to its neighbouring nodes and its epoch time in the recent past. As an extension of the range-based velocity estimators, the scheme has similar advantages such as being reliable for channels of unknown properties, accurate in noisy environment, and efficient as it is localization-free. The associated publication is [21].

1.5 Outline of Thesis

The thesis is organized as follows:

Chapter 2 gives a brief background information about MANETs, fundamental ad-hoc networking technologies and the challenges of QoS provisioning. Routing and service discovery mechanisms for dense and sparse networks are introduced in in Section 2.3. Previous studies on modeling and characterization of mobile mobility are presented in Section 2.4. Section 2.5 introduces the concept of QoS, some QoS measures that are commonly used to benchmark networking technologies, and several design considerations of QoS policies for MANETs.

The study of efficient routing technologies for dense and sparse mobile ad-hoc networks are presented in Chapter 3 and Chapter 4, respectively.

Chapter 3 addresses the issues related to hybrid routing and service discovery in dense ad-hoc networks. The problems of existing technologies related to this topic are reviewed in Section 3.2. Section 3.3 presents a new *track-based* scheme for hybrid service/routing discovery. Settings of the experiments to compare the performance of the proposed scheme with existing ones and the analysis of the results are given in Section 3.4.

Chapter 4 presents the existing problems and solutions on packet delivery in sparse mobile ad-hoc networks. A review of conventional routing protocols for networks with sparse node

distributions is given in Section 4.2. A new adaptive multi-copy packet delivery scheme is proposed in Section 4.3. Section 4.4 presents the experiments that are designed to validate the performance of the proposed scheme and analyzes the experimental results.

The research on efficient mobility characterization in sparsely or densely populated networks are presented in Chapter 5, Chapter 6 and Chapter 7.

Chapter 5 discusses the measurement of network performance using resource efficient mobility metrics. An introduction of traditional mobility metrics is given in Section 5.2. Section 5.3 presents a new GPS-free mobility metric and its variants. Several experiments for performance evaluation are presented in Section 5.4. In this Section, a calibration method is discussed for improving the accuracy of the mobility metrics in noisy environments.

Chapter 6 addresses the issues of velocity prediction for networked mobile devices. Section 6.2 introduces the state of the art in the area of mobile velocity measurement/prediction. Two new estimators are presented in Section 6.3 for estimating inter-nod relative velocity in sparse networks and dense networks, respectively. Their performance is validated in Section 6.4. In Section 6.5, a new method is proposed for a mobile node to estimate network density in order to examine the applicability of either of the two estimators.

Chapter 7 introduces a new scheme for efficient characterization of the nodal mobility that is associated with topological stability. The scheme is extended from the range-based velocity estimators introduced in Chapter 6. The details of the scheme is given in Section 7.3. The performance of the scheme is presented in Section 7.4. Section 7.5 discusses an application of the proposed scheme for link availability prediction.

Chapter 8 reviews the achievements of the project, suggests future work and concludes the thesis.

Chapter 2

A Review of Mobile Ad-hoc Networking and Quality of Service Provisioning

2.1 Introduction

The work presented in this dissertation is centred on an emerging type of network architecture, mobile ad-hoc networks (MANETs). This Chapter introduces the concept, characteristics and benefits of MANETs as well as the challenges of QoS provisioning in this kind of networks. A brief introduction of the background information on MANETs is given in Section 2.2. Two fundamental aspects of mobile ad-hoc networking, routing and mobility characterization, are described in Section 2.3 and Section 2.4, respectively. Section 2.5 gives details of commonly used QoS metrics and some design considerations of QoS provisioning in MANETs. Section 2.6 summarizes this Chapter.

2.2 Mobile Ad-hoc Networks

The recent advances in wireless communication technologies and the fast growing availability of mobile devices have engendered a revolutionary communication paradigm, mobile ad-hoc networks. Differentiated from traditional wireless/mobile networks such as WLAN and Cellular networks, a mobile ad-hoc network is temporary network that can be constructed from a group of mobile devices anytime, anywhere. Peer to peer communications in an ad-hoc net-

work are relayed neighbouring nodes in a hop-by-hop manner. Therefore, there is no need for static infrastructures or centralized network backbones in multi-hop ad-hoc networks. Given that their popularity increases day-by-day*, the mobile ad-hoc network working group was charted in 1997 in order to discuss and develop solutions in this area.

Fig. 2.1 shows an example mobile ad-hoc network, which consists of several wireless-enabled PDAs and Laptops. Communications between one node and another in the network could cover more than one hop due to the limited transmission range of wireless interfaces (e.g. 160m for a ORiNOCO card operating outdoors at a data rate of 11Mb/s [23]). Intermediate nodes that happened to be in the communication paths act as routers and forward packets towards the hosts. Generally, the fundamental attributes that distinguish mobile ad-hoc networks from conventional wireless/mobile networks and lead to new research problems can be summarized as follows:

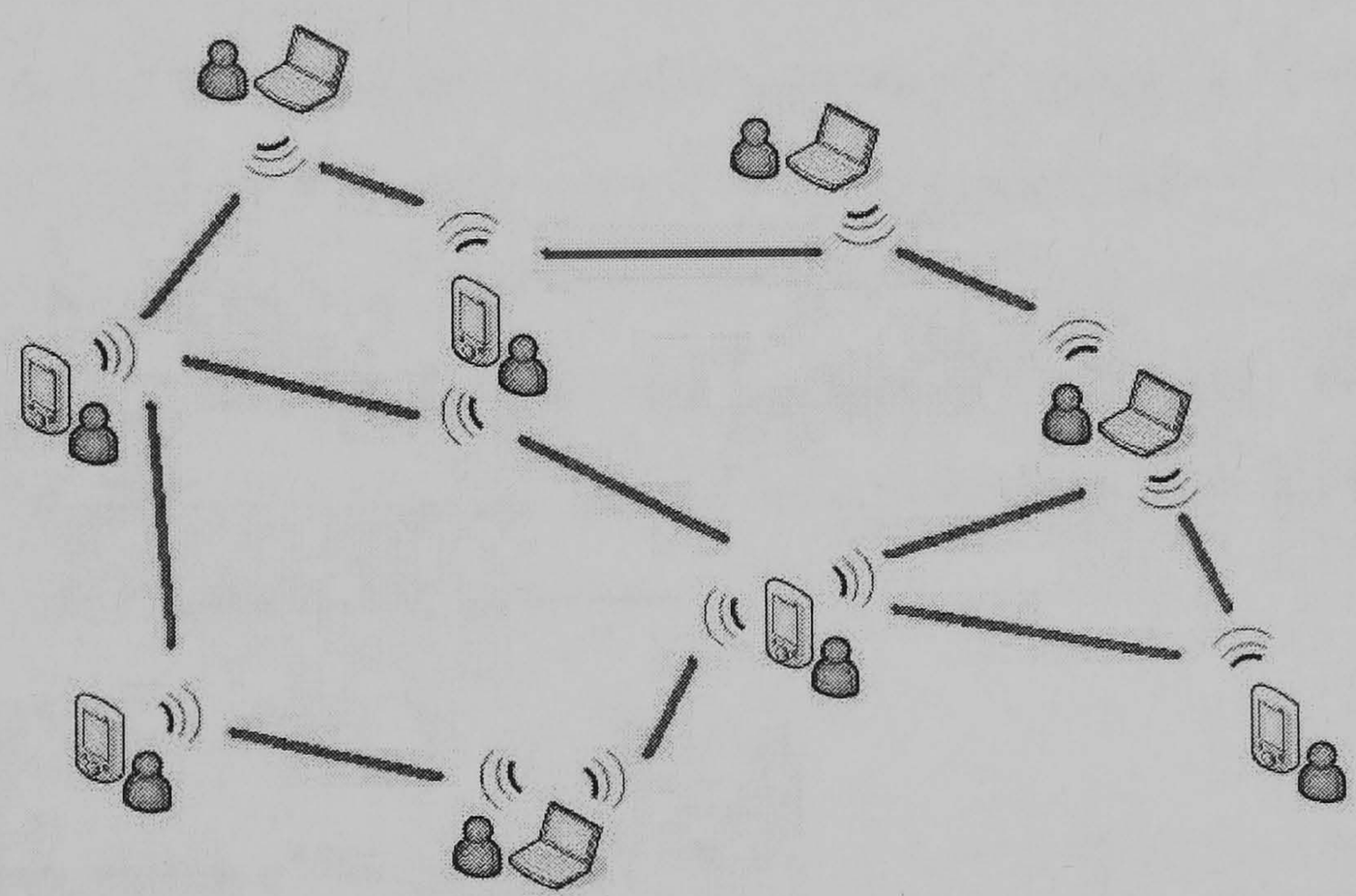


Figure 2.1: An example of mobile ad-hoc networks

- *Multi-hop delivery*:

*MANET networking capacities have been recently added to the radio links of CISCO routers [22]

In mobile ad-hoc networks, the direct transmission range of a node is limited by the transmitting signal strength. When packets are sent out from a source node to its destination outside the direct transmission range, one or more intermediate nodes are required to relay the packets in a hop-by-hop manner. This is in contrast to traditional WLAN and cellular networks where the wireless communications are only "one-hop". As the multi-hop path of packet delivery are connected by wireless links between mobile nodes that could change their positions at anytime, routing mechanisms, one of the most challenging tasks in mobile ad-hoc networking, are required to find a right route to the destination and to maintain the established path.

- *Distributed Management and Control*

  Without a fixed or centralized infrastructure, in ad-hoc networks every nodes is responsible for dynamically discovering and establishing connections to neighbouring nodes that they can directly communicate with. Mobile nodes have also to regularly disseminate their local topology to each other in order to update their global views of the network. Ad-hoc networks are thus highly self-governing, self-organizing and robust to single-point failures. However, to make sure every node has a fresh and correct view of the network, an open research question is how to determine the frequency of disseminating topological information so as to let mobile nodes aware of the most recent topological changes with minimized traffic overhead.

- *Limited Computing/Communication Resources*:

  The computing and communication capabilities as well as the power supply of mobile devices has been being improved since they were invented. However, the ever-increasing demand for mobile access to more complex services/programes has been always challenging to the computing power of existing devices. The wireless channels over which mobile nodes communicate can not easily provide as much bandwidth as that of a wired connection. High bandwidth wireless technologies (e.g. UWB) have developed but their

uses are limited in short distance (e.g. within 10m). The battery lifetime of mobile devices has been dramatically increased in the last decade and more improvement on the mobile energy supplies can be expected in the future. However, as long as the common expectation for mobiles with smaller hardware size and longer recharging intervals exists, energy supply will be always a bottle-neck for mobile computing. Therefore, a research challenge that will remain in the foreseeable future is how to make the networking technologies more resource efficient, i.e. to be more lightweight, bandwidth economical and energy saving.

- *Unpredictable and Unreliable topology*:

  In wireless/mobile networks, the topology is normally unreliable due to the adverse conditions of wireless channels (e.g. fading, shadowing, interference, collisions, etc). In MANETs, the situation is made worse by mobile nodes that constantly moving and even joining and leaving the network at will. Therefore, the connectivity of links in MANETs may vary with time leading to the unpredictability of the network topology. In order to get packets delivered, mobile nodes have to regularly update the links directing to destination nodes and to determine obsolete neighbours that have moved away. Therefore, it is very important for a node to have knowledge of the moving patterns of nodes in its vicinity in order to estimate the lifetime of its wireless links and existing paths. Mobility characterization, i.e. modelling the movement patterns of mobiles and predicting their speeds and orientations, is a fundamental technology that can support effective predictions of topological dynamics (e.g. the availability of links and end-to-end paths) and even QoS indicators at application level. An opening research topic is how to predict the mobility pattern of neighbouring nodes and even that of the whole network efficiently and distributely at a single mobile node. Network partitioning is another unique challenge in mobile ad-hoc networks. The reasons could be either a very low node density or the group mobility behaviour [24]. Group mobility is referred to the phenomena in which

mobile nodes with similar mobility characters form mobility groups that may exhibit distinct movement patterns. Network partitioning would intermittently break a connected network topology into several isolated partitions. In such situations, traditional routing protocols would fail. A new generation of routing schemes is being developed to make use of the intermittent links to achieve the eventual delivery of packets.

The unique properties of MANETs suit the special demand of many applications for a mobile communication. For example, MANETs can be easily, quickly deployed in areas of various size with low installation costs at anytime anywhere, have the capability self-optimizing for structural changes, and possess the ability of self-healing upon unexpected communication failures. Typical applications of MANETs can be found in areas such as wildfire fighting and e-healthcare emergency where quickly setting up a network is critical, in disaster management and military exercises where the environment is so hostile that constructing a fixed infrastructure is very difficult. The principle of ad-hoc networking can be also applied to network of wireless sensors scattered in a wild area or embedded in a building/vehicle for distributed data collections. A more commercial application of ad-hoc networking is wireless mesh networking [25], which is still multi-hop-based but comprised of both mobile and fixed nodes. Wireless mesh networks have received a lot of interests as an economical way to extend the current Internet to remote regions. More successful applications of MANETs would emerge if the aforementioned technical challenges (e.g. routing and mobility characterization) were addressed by innovative and effective solutions.

2.3 Fundamental Mobile Ad-hoc Networking: Service and Routing Discovery

In the IP networking domain, routing is referred to the specific mechanism that responsible for finding a suitable path along which packets can be delivered from the source to the desti-

nation. In pervasive computing, communications between nodes are not bound by specific IP addresses but the special services held by the hosts. A service can be a software or hardware entity such as a computational function, a segment of data file, communication channels or storage spaces, etc. Therefore, technologies for service discovery has been developed to address this need. Previously, service discovery is mostly accomplished in the application layer by a range of service lookup methods/protocols such as the Distributed Hash Tables (DHTs) based Chord [26] and LANE [27], and other distributed on demand service discovery schemes such as UPnP [28] and SLP [29]. As these protocols rely on underlying routing protocols to handle network dynamics and to discover paths to service hosts/brokers for service lookups, a more efficient way of locating services is to perform service discovery in the network layer (e.g. to integrate with routing protocols) [7] [30] [31].

A common structural view of the integration of service discovery mechanisms with routing protocols is demonstrated in Fig. 2.2. Fig. 2.2a shows the protocol stack which integrate service discovery with an ad-hoc routing protocol (e.g. AODV [6]). In this architecture, the conventional routing module is integrated into the service discovery module, which accepts a description of services demanded (e.g. keywords and IP addresses) from applications, packetizes them into service request packets before pushing them to the link layer. Upon receiving service replies, the path to the discovered hosts is updated in routing module whereas the availability of the service providers is reported to the applications. If one of the applications is a service provider, it should also respond to requests and send back service reply packets to the requesting nodes. Fig. 2.2b shows the format of service request/reply packets, which are extensions of conventional routing request/reply messages. For instance, the destination IP address is incorporated into the service description, which can be keywords, port numbers or IP addresses representing the hosts whose path need to be determined. The service discovery process would reduce to a normal route discovery if the service description only contains an IP address.

In MANETs, the average number of neighbouring nodes in a node's signal coverage, often referred to as node degree or node density, is a very important factor that affects the quality of

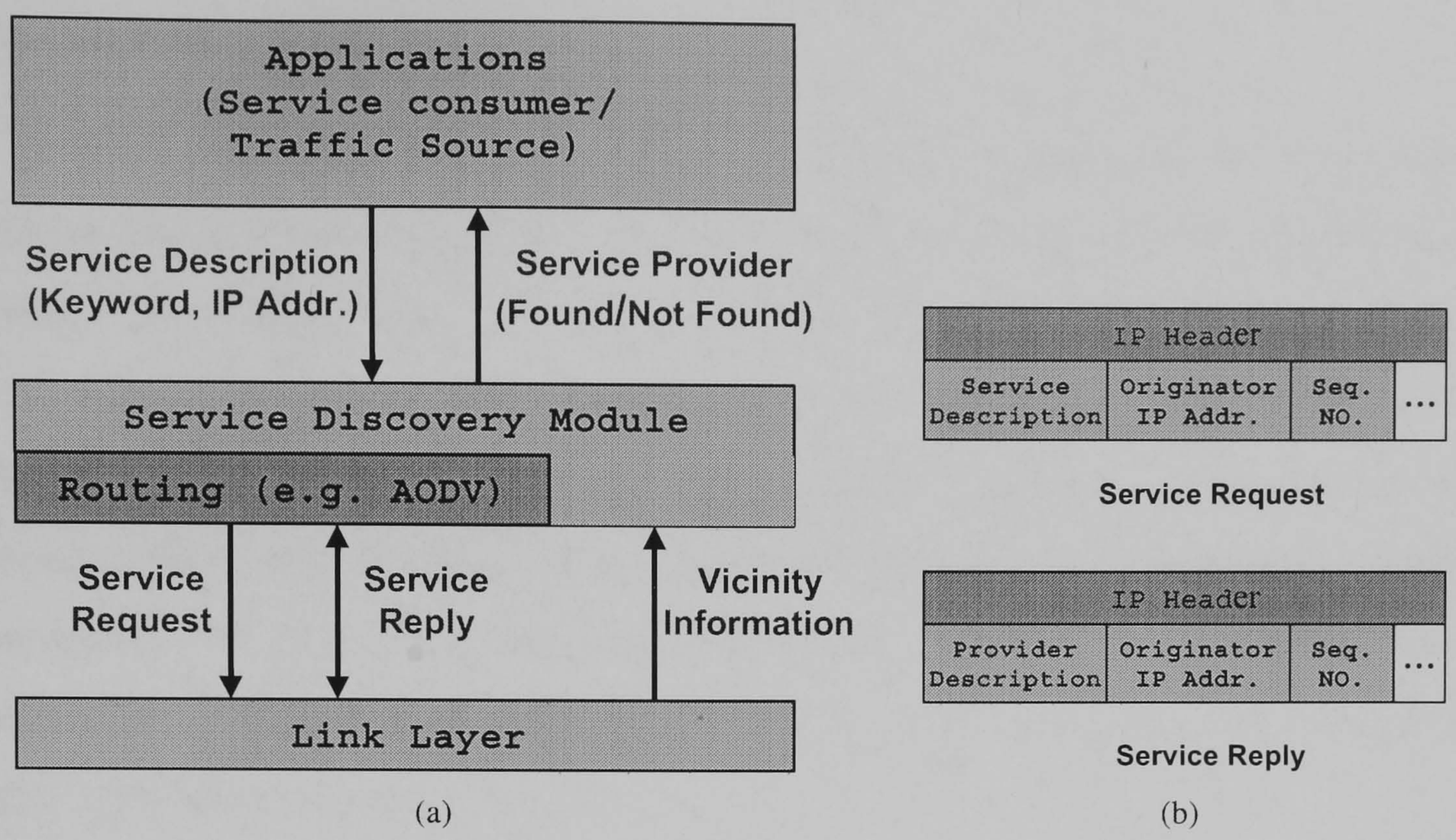


Figure 2.2: Network level service discovery (a) the protocol stack with message flows, (b) the format of service request/reply packets.

communications. For instance, there are more source-to-destination paths in denser networks, which has also a higher chance of transmission collisions. For routing mechanisms, distinct strategies are required to deal with the situations when at least an end-to-end path exists and when no end-to-end route can be found, most of the time.

2.3.1 Service/Routing Discovery in Dense Mobile Ad-hoc Networks

For routing discovery in ad-hoc networks of high density, numerous solutions have been proposed in recent research, some of which have been further extended to support service discovery. The existing service/routing discovery protocols of interests are those working in the network layer, which can be broadly classified as proactive [32], reactive [33] [6] [34], virtual backbone/hierachical [30] [35] [36], reactive/proactive hybrid [37] [8] [38] [39], and position-based routing [40] [41] [42].

Proactive routing discovery

proactive routing protocols, also referred to as table-driven protocols, carry out routing discovery/maintainnce in a way that is commonly used in the wired networks. Routing mechanisms of this category require each node to maintain a routing table that contains information on the route to all of the possible destinations in the network, such as destination IP, the next hop and number of hops, etc. To let mobiles aware of link changes and to maintain the consistency of the routing table, disseminating of route update messages throughout the network is necessary. The table-driven routing protocols have the benefit of instant access to routing information at the expense of extra proactive traffic. Example proactive routing protocols are DSDV [32], WRP [32] and OLSR [43]

Reactive service/routing discovery

In pure reactive service discovery approaches, e.g. DSR [33], AODV [6], GOSSIP [34], nodes advertise service/routing requests on demand and periodical exchange of link/service state information is not required. Only when a node needs to know the path to nodes holding the resources that it wants, it floods service requests throughout the network. Response messages are sent back by nodes which hold the very resources or IP address being requested or cached information about routes to the right destination. Efforts to maintain proactive areas are saved in such an on-demand way of service/routing discovery. Fig. 2.3 demonstrates a typical process of discovering a path to a destination in the AODV protocol. As shown in Fig. 2.3a, source S need to send a message to destination D but it does not have a valid route to D. It initiates a routing discovery process to find out the path to D. It broadcasts a route request (RREQ) packet to its neighbours, which then relay the request to their own neighbours. The process would continue until either the destination D or other intermediate nodes with a valid route to D receive the RREQ message. Having received a copy of the RREQ message, the destination/intermediate node responds by unicasting a route reply (RREP) packet back to the source S using the reverse

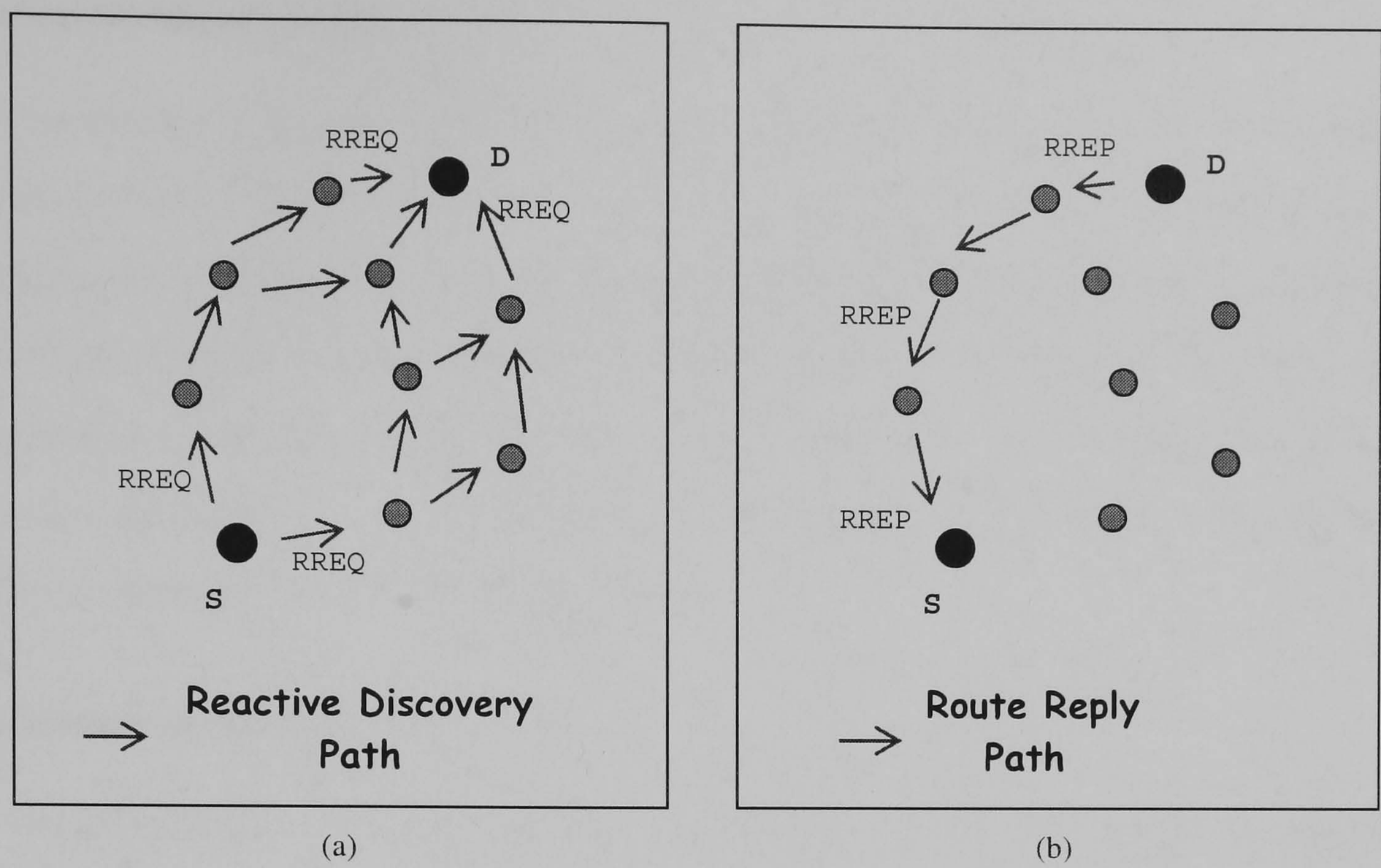


Figure 2.3: AODV routing (a) reactive routing discovery, (b) route reply propagation

of the path that the RREQ message has traveled through (see Fig. 2.3b). While relaying RREQ (RREP) messages forward (backward), intermediate nodes update their routing tables and record the next hop of the route to the source and the destination for future use.

Hybrid service/routing discovery

Hybrid service/routing discovery protocols such as CARD [37], ZRP [8] and its variants [38] [39] try to exploit the advantages of both proactive and reactive strategies. These protocols limit periodic exchanges of route updates to a node's neighbours in a two-dimensional area, i.e. zones or vicinities of nodes several hops away. Paths to a specific node (or a service in a node) inside the proactive area are updated regularly so that they are ready on demand. For nodes (or nodes hosting specific resources) not reachable in a node's own zone or vicinity, these protocols would initiate reactive route/service queries that are either unicasted to remote 'contacts' [37] or bordercasted to neighbouring proactive areas/zones [8] [38] [39].

Virtual backbone routing

Approaches in the category of virtual backbone [30] [35] [36] try to maintain a virtual hierarchical architecture in the ad-hoc network. The virtual hierarchy provides a way for the implementation of service/routing discovery approaches developed for wired networks for MANETs. However, maintaining a virtual backbone in MANETs is costly as extra maintenance traffic overhead is required to keep the backbone/hierarchy stable. Due to its dependence on nodes from which the backbone or high level hierarchy are built up, this category of protocols are prone to single point failures in their backbones.

Position-based routing

Position-based routing protocols such as LAR [41] and GPSR [42] attempts to reduce the traffic overheads of reactive algorithms by exploiting location information of mobile nodes. Position-based protocols need nodes aware of their locations using positioning facilities such as the GPS. A location lookup system is also needed to provide the source with the location of the destination, which is often embedded into the header of a packet. Therefore, intermediate nodes can forward packets only to a neighbouring node that is closer to the destination. In comparison to reactive protocols, the advantage of position-based routing is the significant reduction in control packet overhead without sacrificing the success rate of packet delivery. However, the position-based routing has some major drawbacks that limit its application. For example, the design of a reliable location lookup system is very challenging in MANETs. It is also not economical to equip every mobile with the functionality to determine its location in realtime.

2.3.2 Routing in Sparse Mobile Ad-hoc Networks

In MANETs of low node density, the distribution of mobile nodes can be so sparse that no direct source-to-destination path can be found most of the time. Traditional routing protocols

that only forward packets in connected paths are not suitable for sparse MANETs due to the very weak connectivity. To make communications viable in sparse MANETs a special routing solution, *store-and-forward* based routing, has been proposed for packet delivery in mobile networks of sparse topology by making use of the intermittently connections between mobile nodes. Fig. 2.4 illustrates the principle of *store-and-forward* based routing in a sparse MANET of 3 nodes. As shown in Fig. 2.4, in the network of node s, r, and d, there is no connected path between node s and d at the time (e.g. 10am) when s has some packets destined to d. Node s would hold these packets temporally and forward them to any passing-by nodes (e.g. to r at 10:30am). Given that the nodes do not stop moving in the area of limited size, either node s or r may (e.g. at time 11am) meet d after a period so that the packets could be offloaded at their destination. The distribution of the length of the time period is determined by a number of factors such as the speed of mobiles, the size of the network area and the node density. There is a possibility that the time period is infinite and a packet will not be able to reach its destination by the time that it is dropped from the buffers.

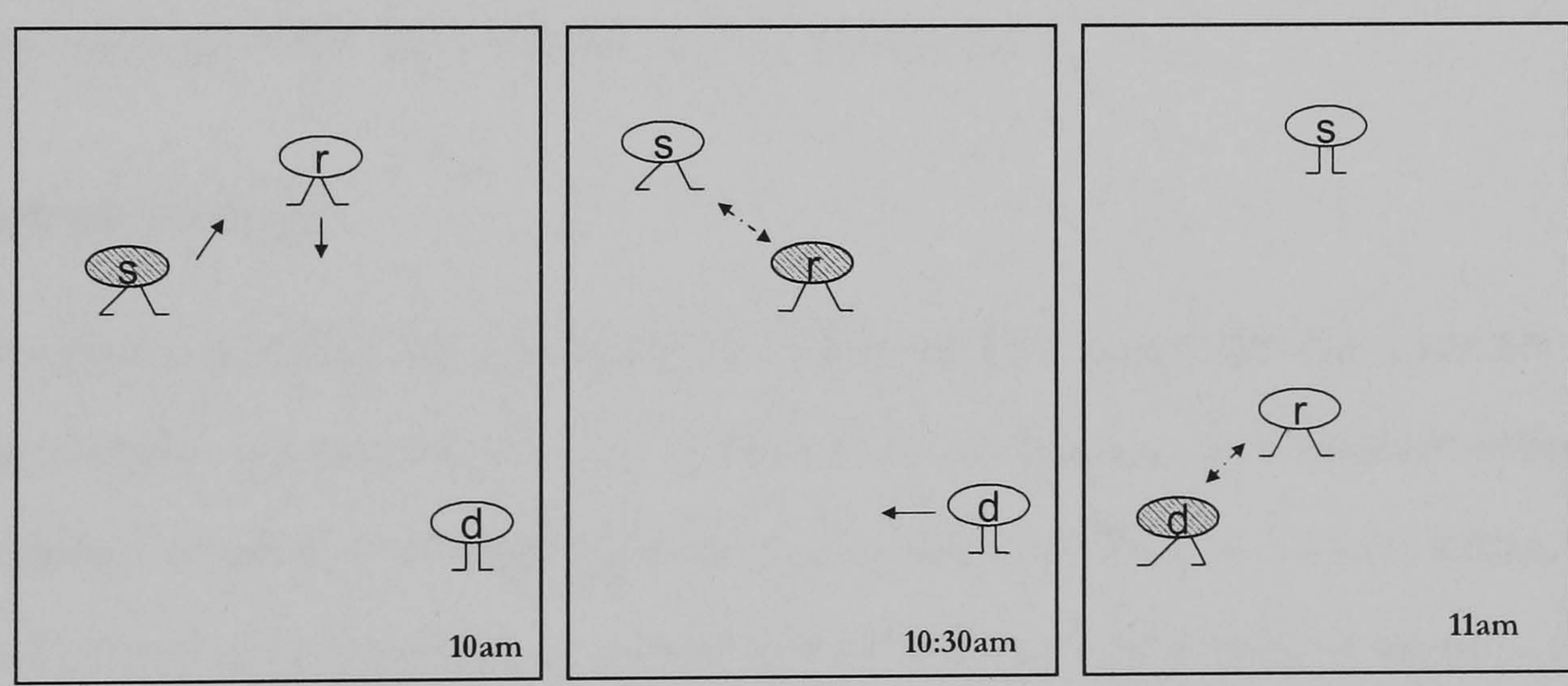


Figure 2.4: Illustration of the store-and-forward based packet delivery

Based on the concept of *store-and-forward*, numerous routing schemes have been developed in recent studies for networks that share the same properties of being sparse and mobile (e.g. Intermittently Connected Mobile Networks or Delay Tolerant Networks). These proposals can be broadly classified as deterministic methods [44] [45] and probabilistic methods. A

more detailed survey of routing in sparse MANETs/Delay Tolerant Networks (DTNs)can be found in [46].

Deterministic methods

In the deterministic approaches, an end-to-end route is determined before messages are transmitted by the source. The determination of the delivery path is depend on the available knowledge on the characteristics of the network topology and the trajectories of mobiles at a specific time (or time interval). An example deterministic routing for sparse MANETs is the *space-time* routing framework proposed in [44]. The *space-time* routing forward packets in pre-specified schedules and paths that are designed according to the *a priori* knowledge of nodal trajectories. A similar routing framework proposed in [45] incorporates several strategies that can be used in accordance with the amount of time-varying topological information (or *knowledge oracles* [45]) available. For example, if the information on aggregate statistics of the inter-node contacts is known, a routing table can be constructed using the Dijkstra algorithm with time-invariant edge costs and the average waiting time.

Probabilistic methods

Probabilistic methods are proposed for more common cases that the network evolution is not predictable nor known *a priori*. In this class of methods, intermediate nodes decided independently whether to forward packets at realtime, according to various policies that can be unconditional forwarding [9] or conditional forwarding based on QoS targets [10], utility threshold or mobility pattern of the contact [47], etc. Epidemic routing [9] is the simplest probabilistic method that allows every relay node to forward packets to their contacts unconditionally. Epidemic routing can quickly disseminate packet throughout a network and no *oracles* of the network are needed. However, unconditional forwarding often leads to high bandwidth wastage. More bandwidth economical probabilistic methods such as only forward packets if pre-defined conditions are met. For example, PROPHET [48] only forward packets to contacts

having higher chance of meeting the destination. *Mobyspace* based routing only replicate packets to neighbours having similar mobility pattern to the destination. Multi-copy routing [10] allows packets to be relayed as long as the total number of packet replicates has not reach a preset limit, the *replication factor*. Complementary techniques such as erasure-coding [49] [50] can be also applied for the probabilistic methods to reduce the cases of long delivery delays without increasing traffic overhead.

2.4 Fundamental Mobile Ad-hoc Networking: Mobility Characterization

In ad-hoc networks of mobile nodes, the main factor that impairs the quality of communications is the frequent link breakages and topological reconfigurations caused by relative movements between mobile nodes. Therefore, it is critically important to understand the various moving patterns of mobiles and to analyze their impacts on the performance. It is also imperative to predict the mobility of a node as well as that of a whole network and to make ad-hoc networking technologies adaptive to mobility .

2.4.1 Mobility Modeling and Abstraction

Mathematical synthetic models have been used to realistically reproduce typical movements of mobile nodes in specific scenarios or environments. A well designed mobility model should be able to capture as much as possible the behavior of mobiles in reality while not being too complex and unanalyzable. Mobility models are widely used with network simulators to generate movement patterns for mobile nodes in order to examine the performance of networking protocols. Most of previous research in mobility modeling is carried out in two directions. One direction is to design new models for better abstractions of various real world scenarios. The other is to analyze the statistical properties of the mobility models, and to study their impli-

cations in designing networking technologies and their effects on the performance of communications. An survey on mobility models that were mostly proposed before year *2000* can be found in [11]. Another survey that covers more recent research advances in mobility modeling is given in [51]. Most of existing mobility models can be roughly classified into three categories according to their degree of randomness [51], namely, statistical models, constrained topology based models and trace based mobility models.

Statistical Models

Mobility models in which the motion of mobiles are total random are classified as statistical models. Total randomness of motion means that mobiles can move to anywhere in the network area with randomly chosen speeds and directions. Example statistical models are the Random Walk (or Random Direction) model [52] and the Random Waypoint model [11].

Random Waypoint (RWP) model is the most popular and studied mobility model. In this model, a mobile randomly chooses a destination called waypoint and moves towards it in a straight line with a constant velocity that is selected randomly between a speed range of [min speed, max speed]. After it reaches the waypoint, it pauses for some time before repeating the same procedure. Random Direction (RD) model is very similar to the RWP model. In this model, a mobile randomly chooses a direction from [0, 2π] and moves forward in a straight path. After some random time or the mobile reached the boundary of the network area, new speed and orientation are set for the next epoch. The properties of these two models have been well studies. For instance, C. Bettstetter in [53] studied the node distributions of RWP model (non-uniform distribution) and the the RD model (uniform distribution), as well as the resulting network connectivity. The Random Waypoint model is also shown to have the *speed decay* phenomena. The related discussion can be found in [54].

Constrained Topology Based Models

Constrained topology based mobility models have relatively limited motion randomness due to the predefined constraints reflecting the situations of real-world scenarios. Some typical constrained topology based mobility models are the Reference Point Group Mobility (RPGM) model [55], City Section Mobility (CSM) model [56], Freeway Mobility model [57] and the Obstacle Mobility model [58].

The RPGM model were introduced as an abstraction of the movements of a group of mobiles that have similar/associated mobility as well as a mobile's motion within a group. In this model, each group has a logical centre. The motion of a group is defined by its logical centre, which moves according to the RWP model. The movement of an individual node is specified by the combination of the group motion and a random motion vector. The CSM model (also referred as the Manhattan Grid model) tries to reproduce the movements of mobiles in a Manhattan-like city that consists of building blocks and streets. In this model, mobiles can only move in streets. The maximum speed limits can be configured in accordance to the type of cities/roads of interests. Similar to the CSM model, the Freeway Mobility models nodal movements in the highways. Naturally, mobiles in this model has a much higher speed limit and much less turn-around freedom than those in the CSM model. The Obstacle Mobility model considers several facts in human movements in mobility modelling. For example, people move towards specific destinations rather than randomly chosen points. Obstacles such as buildings could block peoples movements. And usually, people intend to move along (often the shortest) pathways instead of random trajectories. A similar model that has more advanced features (e.g. generating motion constraints from a map) is introduced in [59].

Trace based Mobility Models

Movements of mobiles in trace based mobility models are deterministic as they are defined based on existing trace data. The benefit of trace based mobility models is that these models

can faithfully represent the real movements of mobile nodes and help to disclose valuable proprietary information of actual mobility patterns. Attempts to build trace based mobility models have been carried out in [60] [61], in which the daily movements of campus WLAN users are traced and translated to mathematical models. Mobility models produced by these efforts tend to be tightly coupled with the specific environments and mobile hosts with which the traces are collected. Therefore, the trace based mobility models can not be applied to generic scenarios where the movement pattern of mobiles does share the same features that only exist in the origin of the traces.

2.4.2 Mobility Metrics and Measurement

The distinct behavior of mobiles in different scenarios lead to different mobility patterns that can be captured by various mobility models. However, mobility models alone can not give enough information on how the various mobility patterns impact the network topology and the performance of routing protocols in diverse ways. Mobility metrics are introduced to quantitatively and qualitatively describe the influence of mobility patterns on network performance. As most of the existing mobility metrics are derived on the basis of the current and past motions of mobiles. It is important to devise efficient mobility measuring methods for situations when knowledge of velocities/orientations of mobile hosts is not available.

Mobility Metrics

Common mobility metrics are derived either based on the mobility of mobiles in physical terms (e.g. speed or relative speed), or based on the physical observations and mathematical modeling of the topological connectivity (e.g. link duration, path duration and remoteness [62])

.

- *Average relative speed*

  The *average relative speed* of all pairs of nodes in the network is proposed in [63] as a

mobility metric for performance evaluation. In comparing to the absolution speed, the relative speed between mobile nodes in a network is more appropriate as it is more related to the topological stability.

- *Average link duration or link change rate*

  Link duration is the time interval during which two nodes are within each others radio coverage. A link changes its state from up (down) to down (up) if a node move out from (in to) the other's transmission range. *Link change rate* is the total number of link state changes (up/down) in unit time. *Link duration* is shown in [64] to have a higher correlation with the network performance (e.g. packet loss ratio and end-to-end delay) than *link change rate*. However, as it would normally take a long time to observe link durations, the *average link duration* is more suitable as a long term indicator of the network performance.

- *Average path duration*

  Average path duration [65] is the average time interval during which all links in a source-to-destination path exist. The average path duration is proportional to the nodal transmission range while being inversely proportional to the average path length (e.g. hop count) and inter-node relative speed.

- *Remoteness*

  Remoteness is proposed in [62] as a basis for deriving a standard mobility measure. The measure is derived from average relative speed. Based on the observation that relative movements between nodes that makes most impact on the link status are those near the transmission range of each other. A *remoteness* function is defined as a monotonically increasing function of the distance between two nodes within the transmission range. But the function has a peak value with a distance at the communication range so that the remoteness function is most sensitive to nodal movements around the rim of the

radio coverage. The proposed standard mobility measure is defined as the average of the derivative of *remoteness* over all node pairs. The benefits of such a measure as shown in [62] is that it is flexible for customizations and consistent with the link change rate in a wide range of scenarios.

Mobility Measurement and Estimation

Detection and measurement of the movements of mobiles (e.g. speed and orientation) are critical in deriving some of the mobility metrics (e.g. average relative speed and *remoteness*). The measured mobile movements can be collected for creating the mobility profile of mobiles within a time frame. It also important to predict the intentional motion of mobiles for estimating link durations [66] and for improving handover processes [67].

If the mobiles are configured with GPS receivers, their movements can be easily deduced from their position changes over time. For scenarios in which the GPS receivers are not available or GPS signals are not reliable, mechanisms exploiting the properties of radio signal emitted from moving devices (e.g. the *Doppler* effect [67]) can be used for speed estimation. The *Doppler* effect is the change in frequency and wavelength of a radio wave that is perceived by an observer moving relative to the source of the waves. The maximum *Doppler* frequency, which is the ratio of the mobile speed to the wavelength of the carrier [67], is often used to deduce the relative speed between two mobiles (or one mobile and one fixed station). In reality, it is difficult to determine the maximum *Doppler* frequency especially when the propagation model of the wireless channel is unknown.

As the relative movements between a pair of nodes are always accompanied with variations of distances between them, the time-varying inter-node ranging information can be also exploited for velocity estimation. In [66], a speed estimation method that does not rely on the *Doppler* effect is proposed. The method utilizes 3 distance samples of a node pair in a short time to build a nonlinear equation system, from which estimates of relative velocity between the node pair can be derived. A critical problem of this method is that it is vulnerable to mea-

surement noises, which may be caused by multi-path, Non-Line-Of-Sight (NLOS) and other environmental factors that interfere with signal propagation, in the ranging information. In Chapter 6, a new ranging-based velocity estimator that is more robust in noisy environments is introduced as a part of the Ph.D research.

2.5 Quality of Service Provisioning in Mobile Ad-hoc Networks

Quality of Service (QoS) is defined as a set of service requirements that a network must meet while transporting a stream of packets from the source to the destination [68]. QoS provisioning techniques allow the network to guarantee a set of performance criteria representing the service requirements, e.g. throughput, bit error rate at the wireless link layer, packet delivery delay, packet loss ratio, jitter at the IP network level, and user perceived voice/video/data quality at the application level. The requirements of QoS provisioning can be classified as soft QoS and hard QoS. Soft QoS [5] means that there may exist transient time periods when the required QoS is not guaranteed due to poor networking conditions. But the required QoS is ensured as soon as the network conditions return to normal. A guranteed QoS 100% of the time is referred to as a hard QoS

QoS provisioning in mobile ad hoc networks faces many challenges that are introduced by the unique properties of infrastructureless, decentralized, unpredictable topology and limited resources. To deal with these issues, QoS provisioning mechanisms need to be distributed and self-adaptive to topological reconfigurations. Typical tasks of QoS provisioning for MANETs are: efficient and effective routing, medium access control, mobility management and estimation, bandwidth management, QoS guaranteed delivery, and prediction and measurement of QoS metrics.

2.5.1 Quality of Service metrics

QoS metrics are commonly used by applications to specify their QoS requirements and by networking technologies as measurable indicators of network performance. Numerous QoS metrics have been deployed by various networking technologies in performance evaluation. According to the relevant networking layer in which they are frequently seen being used, a rough classification of the QoS metrics is presented hereafter [69]:

Physical Layer Metrics

- Bit Error Rate (BER), defined as the ratio of the number of incorrectly received bits to the total transmitted bits in a transmission. BER is often used as a measure of the signal quality of one/multi hop(s) of wireless channel(s).
- Signal-to-Noise Ratio (SNR), defined as the ratio of a signal power to the noise power that is corrupting the signal. SNR can only measured at Physical layer but is often used by upper layer techniques to evaluate the link quality.

Link and MAC Layer Metrics

- Link change rate, in addition to be a mobility metric as introduced in the previous section, is also used as a QoS measure of the reliability of wireless links. A link that changes its status frequently is regarded as of low quality and unreliable.
- MAC delay, referred to as the time taken to transmit a packet over a wireless channel with contention-based MAC, including the total propagation time in the channel and the time to acknowledge the data [69]. MAC delay is a good indicator of the situation of link level congestions.
- Frame Error Rate (FER), defined as ratio of the number of corrupted frames received to the total transmitted MAC frames in an unit time. Its a reflection of the performance of

physical layer measures in the link layer.

Network Layer Metrics

- Packet loss rate, defined as the ratio of the total number of packets lost in their journey to the the destination to the total number of packet sent out by the source. The network level packet loss rate is one of the most important QoS metrics, which has a direct impact on the perceived service quality at the application layer and is often a product of multiple low layer factors, i.e. BER, FER and MAC level congestions.

- End-to-End delay, referred to as the time taken for a packet to travel from the source to the destination. End-to-end delay is an important measure for real-time applications, in which late arrival packets are regarded as lost.

- Delay Jitter or Variance, defined as a variation in the end-to-end delay of received packets at the destination. At the source, packets are sent in a continuous stream with the packets being spaced evenly apart. Due to network congestion, improper queuing, or configuration errors, the spacing between adjacent packets arriving at the destination can vary instead of remaining constant. Jitter has a significant impact on the perceived quality of time sensitive applications (e.g. real-time audio/video streaming). In some applications such as VoIP, jitter buffers are deployed to discard packets that arrive too late and to constrain the delay jitter at an acceptable level.

Transport Layer Metrics

- Throughput, defined as the amount of data that is delivered to from one peer of a communication link/path to to another in a network. Throughput is usually measured in bit per second (bps). As a transport level metric, throughput is not only related to the low level technologies (e.g. bandwidth of the wireless link) but also depends on the transport protocol being used. For example, TCP tends to exhibit low throughput than UDP in a

same congested network due to its congestion control mechanism [70].

- Energy consumption per packet, which measures the amount of energy required for delivering a packet including those consumed in transmitting multi-copies of a packet or in end-to-end retransmissions. The common unit of this metric is *joule* per bit or *joule* per packet.

Application Layer Metrics

Application layer QoS metrics are varied among different type of applications. For multimedia applications such as Voice (and/or Video) over IP, the most popular quality benchmark is the Mean Opinion Score (MOS) [71], which provides a quantitative indication of the perceived quality of received media after decoding/transmission. In [72], a new and generic metric called dataMOS is introduced to quantify the user perception of the performance of data-network based applications For instance, in web surfing, dataMOS can be used as a numerical measure of the overall quality, browsing speed, and webpage design as perceived by the users of a website.

2.5.2 Design Considerations and Trade-offs of QoS guarantees in Mobile Ad-hoc Networking

To overcome the unique problems of MANETs while trying to preserve QoS guarantees, several issues and trade-offs need to be considered in designing communication protocols and technologies.

Energy efficiency

Energy-efficiency has been a major concern in ad-hoc networking due to the fact that most of mobile devices are powered by battery. Energy consumptions in communications can be reduced by limiting the traffic overhead of per packet delivery [18], topology/signal strength

control [73], or cross-layer power management systems [74]. However, for packets with QoS requirements, providing QoS guarantees inevitably requires extra energy consumption. A common design strategy for this is to minimize the activities of QoS mechanisms to the lowest level that just can achieve the promised QoS targets.

Bandwidth efficiency

Bandwidth is another type of precious resource in ad-hoc networks of mobile devices. Despite the fact that wireless connections are getting broader bandwidth, the reality is that super-high communication bandwidth (e.g. UWB) are normally only achievable in a short and probably line-of-sight (LOS) distance. Environmental factors such as interference also make designed bandwidth lower than expected and leave data communications an even narrower spectrum. Traffic overhead reduction has been an important design consideration of networking technologies. General strategies include replacing *Broadcast* with *Unicast*, minimizing the proactive overhead (e.g. neighboure discovery and route updates) and restricting the excess traffic of QoS schemes for given QoS objectives. A QoS-adaptive packet delivery scheme is introduced in this dissertation in Chapter 4.

External resource utilization

In order to enhance the performance of communications in MANETs, external devices are often deployed to assist networking protocols. Typical utilization of external resource in mobile ad-hoc networking is the use of GPS receiver to obtain nodal position to support techniques such as the location-assisted packet delivery [42]. However, the use of external resources is not economical in terms of the cost of the mobile devices and energy efficiency. On going research has been trying resolve this problem with novel techniques that can provide the same functionality as the external devices. For instance, GPS-free localization is proposed in [75] to replace the costly GPS receivers. Methods of GPS-free mobility estimation and characterization are also presented in this dissertation in Chapter 5,6 and 7.

Distributed and self-adaptive operation

The traditional centralized and hierarchical QoS frameworks cannot work well in MANETs due to its decentralized nature. It is expected that QoS designs in an ad-hoc network are distributed to individual nodes, which can carry out optimal operations based on their localized knowledge of the network in order to achieve the global QoS objectives. A single mobile node in a QoS framework should also have the capability to survey the network dynamics (preferably based on localized knowledge/methods)and be self-adaptive to changes in the communication conditions.

QoS targets vs. Capacity

It has been shown that various QoS targets are achievable by carrying out extra operations and dedicating more resources of the system. Nevertheless, there is always a trade-off between the QoS requirements of a certain class of traffic and other aspects of the system performance. For example, although sending more redundant packet copies can decrease packet delivery delay in MANETs, the number of replicates of a packet should be carefully controlled so that it does not damage the system's capacity [76]. Therefore, analysis of the impact of networking operations required for a specific QoS target on the capacity and the performance of the whole ad-hoc network system is necessary in devising QoS strategies.

Individual QoS targets vs. public QoS requirements

In addition to the conflict between QoS targets of a traffic class and the system capacity, there are also inconsistencies between QoS objectives of a packet stream and those of other packet streams as well as the QoS promises of a network as a whole. Normally, traffic classes are assigned with different priorities so that QoS mechanisms can treat them accordingly. However, providing QoS to high priority traffic classes at a node may consume too much resources (e.g. energy) and affect the provision of QoS for traffic flows arriving later, even those that

are assigned with a higher priority. An optimal QoS scheme is one that can achieve a balance between preserving individual QoS guarantees for on-going traffic and improving the sustainability of the public QoS policies.

2.6 Summary

In this Chapter, a review is given on the unique properties of mobile ad-hoc networks (MANETs) as well as the two area of networking technologies of interests, i.e. routing and mobility characterization. An introduction on the basic concepts of Quality of Service (QoS), QoS provisioning and common QoS metrics used in performance evaluation is presented as well. As a basis for design considerations of new efficient and effective networking technologies, this Chapter also gives a discussion on the general challenges and trade-offs of providing QoS guarantees in MANETs. The issues and research problems regarding QoS provisioning in MANETs that are introduced in this chapter have directed the PhD studies that are presented in Chapters 3 to 7.

Chapter 3

Track-based Hybrid Service/Routing Discovery

3.1 Introduction

Service discovery, which allows devices to transparently and seamlessly locate available software/hardware entities throughout the heterogeneous communication networks, is a critical component for on-demand communications and collaborations in pervasive computing environments. In the basic paradigm of service discovery, service consumers advertise requests containing attributes representing the services they need. A few service providers who want to share their resources listen on a specified interface for service requests and reply to those matching the services they hold. Routing can be seen as a special case of service discovery, in which a node with a specific IP address is the entity to be discovered.

Service discovery is a challenging task in mobile ad-hoc networks due to lack of stable infrastructures as well as adverse conditions of the wireless channel. For example, conventional directory based service discovery standards such as Jini [77] and UDDI [78] can not be implemented straightforwardly as they require the support of a stable infrastructure.

This Chapter present a novel *Track*-based hybrid service (routing) discovery scheme for MANETs. The key novelty of the proposed scheme is that it performs service discovery based on the one dimensional proactive structure (e.g. tracks), instead of the existing two dimensional

proactive structures (e.g. zones). This design can lead to a significant reduction of proactive traffic due to its light-weight organization of the hybrid structure. Another novelty is that the density of proactive tracks is adaptive to the patterns of network traffic for less proactive traffic.

The remainder of this Chapter is organized as follows. Section 3.2 introduces the related work on existing service/routing discovery protocols for MANETs. Section 3.3 presents an overview of the protocol architecture of the *Track*-based scheme and details of its service discovery mechanism. The simulation-based performance evaluation and results are give in Section 3.4. Section 3.5 summarizes this Chapter.

3.2 Related Work

Related research activities are in the areas of resource discovery/routing for mobile ad-hoc networks. Most of them can be broadly classified as: virtual backbone [30] [35], purely reactive [7] [6] and hybrid reactive/proactive [37] [8] [38] [39].

Approaches in the category of virtual backbone [30] [35] try to maintain a virtual hierarchical architecture in the ad-hoc network to support directory based service discovery. However, the virtual hierarchy requires constant management overhead and can be susceptible to single node failures.

Purely reactive service discovery approaches, e.g. MPP [7], successfully waive periodical updates by advertising service requests on-demand. When a node needs to know about the location of specific resources/services, it floods service requests throughout the ad-hoc network . Response messages will be sent back by nodes which provide the very resources being requested or cached information about routes to the right destination. Although the purely reactive schemes do not require constant exchanging of control packets, flooding the whole network for every query is inefficient, especially when the service provider can be easily located at a node's neighbourhood.

Hybrid approaches such as CARD [37], ZRP [8] and its variants [38] [39] try to benefit

from the advantages of both proactive and reactive strategies. These protocols limit periodic exchanges of route updates to a node's neighbourhood, i.e. zones or vicinities of nodes several hops away. Resources inside the proactive area will be ready on demand. For resources not reachable in a node's own zone or vicinity, these protocols would dynamically initiate service discoveries in a reactive way. They also try to reduce excess traffic flooding by unicasting queries to remote 'contacts' [37] or bordercasting to neighbouring zones [8] [38] [39]. However, recursively unicasting queries to remote areas are likely to incur long discovery delays, whereas bordercasting is vulnerable to query failures due to out-of-date topological knowledge of the complicated proactive zones. A common drawback of these approaches is that substantial bandwidth overhead needs to be spent on maintaining the proactive area even in the absence of service queries or data traffic.

These limitations of current approaches motivated the development of a more efficient and flexible hybrid service discovery protocol, which should be able to minimize proactive overhead as well as querying overhead with the light-weight proactive structure and self-adaptability to different traffic patterns.

3.3 The Architecture of the *Track*-based Scheme

The main contribution of this study is a novel *Track*-based hybrid service discovery scheme. Unlike conventional Zone-based reactive/proactive hybrid schemes [37] [8] [38] [39], the proposed scheme limits its proactive areas only to the one-dimensional tracks connecting service consumers and providers. By doing so, it requires much less proactive overhead than the two dimensional structure such as zones. The proposed approach is also flexible in maintaining proactive areas. On one hand the tracks are built on existing service consumer to service provider associations. The busier the network, the more proactive the system behaviour becomes. On the other hand, the efforts in taking care of the proactive areas can be merged with those required for maintaining communications between service consumers and providers in

the post-discovery stage. The *Track*-based scheme is designed to be integrated with existing routing/networking protocols to achieve better efficiency than application layer protocols (e.g. Lanes [27]).

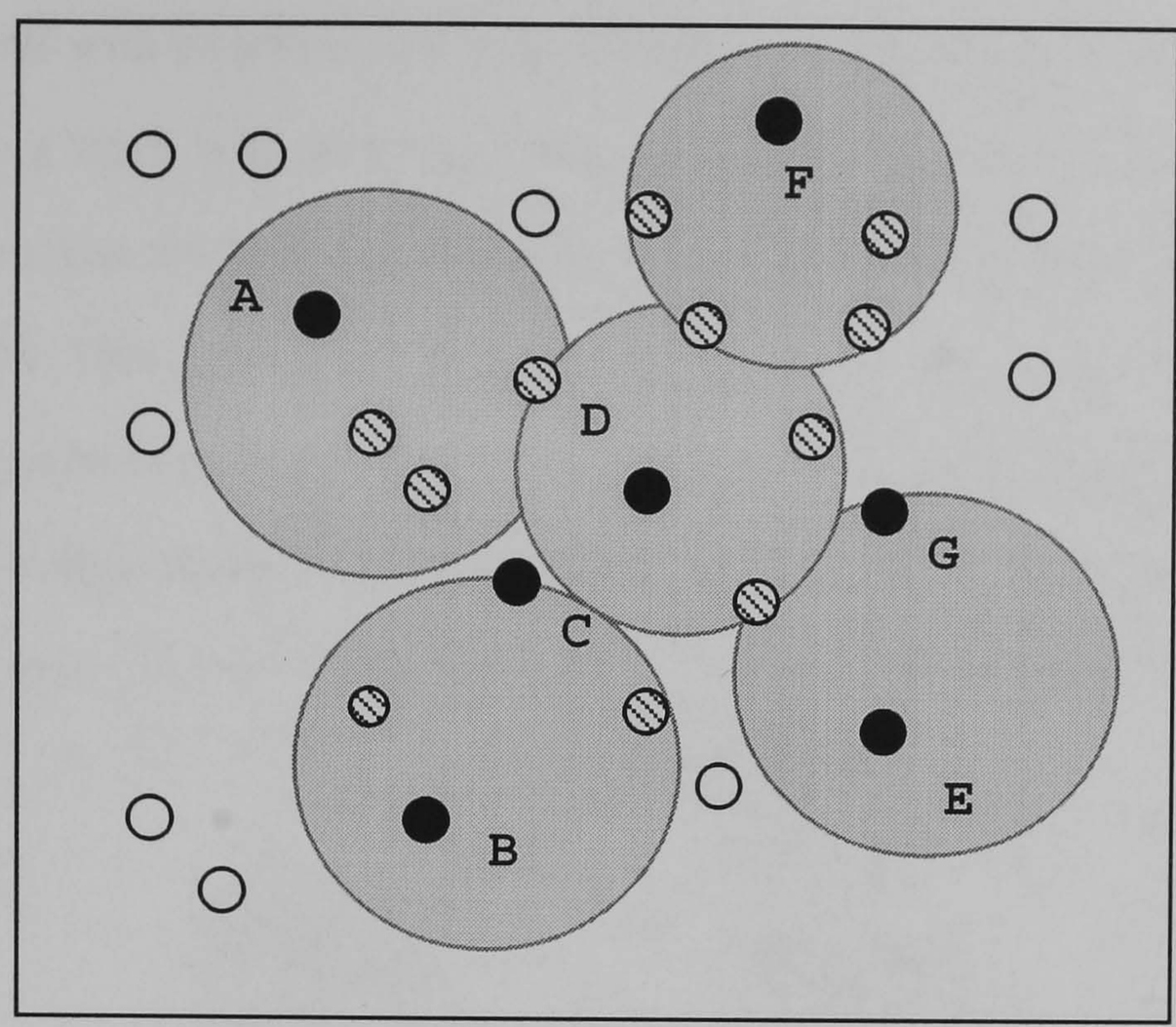
3.3.1 Protocol description

The *Track*-based scheme is comprised of a proactive component and a reactive component for service discovery. It differs from Zone-based hybrid approaches [8] [38] [39] in that it maintains one dimensional track-like structures as proactive areas instead of two dimensional zones. The difference is illustrated in Fig. 3.1, where grey areas represent proactive zones or tracks. Dark, shadow and white nodes correspond to service consumers (e.g. B, C, E and G) or providers (e.g. A, D and F), nodes involved in proactive areas, and the remaining individual nodes, respectively. As illustrated in Fig. 3.1a, proactive areas in the Zone-based service discovery are circular zones consisting of nodes several hops away. Maintaining such a two dimensional topology requires a lot of proactive overhead, especially when node density is high. On the contrary, as depicted in Fig. 3.1b the one dimensional structure of proactive areas in the *Track*-based scheme would require minimal maintenance overhead, as each node in a track only needs to keep trace of its predecessor and its successor in the same track.

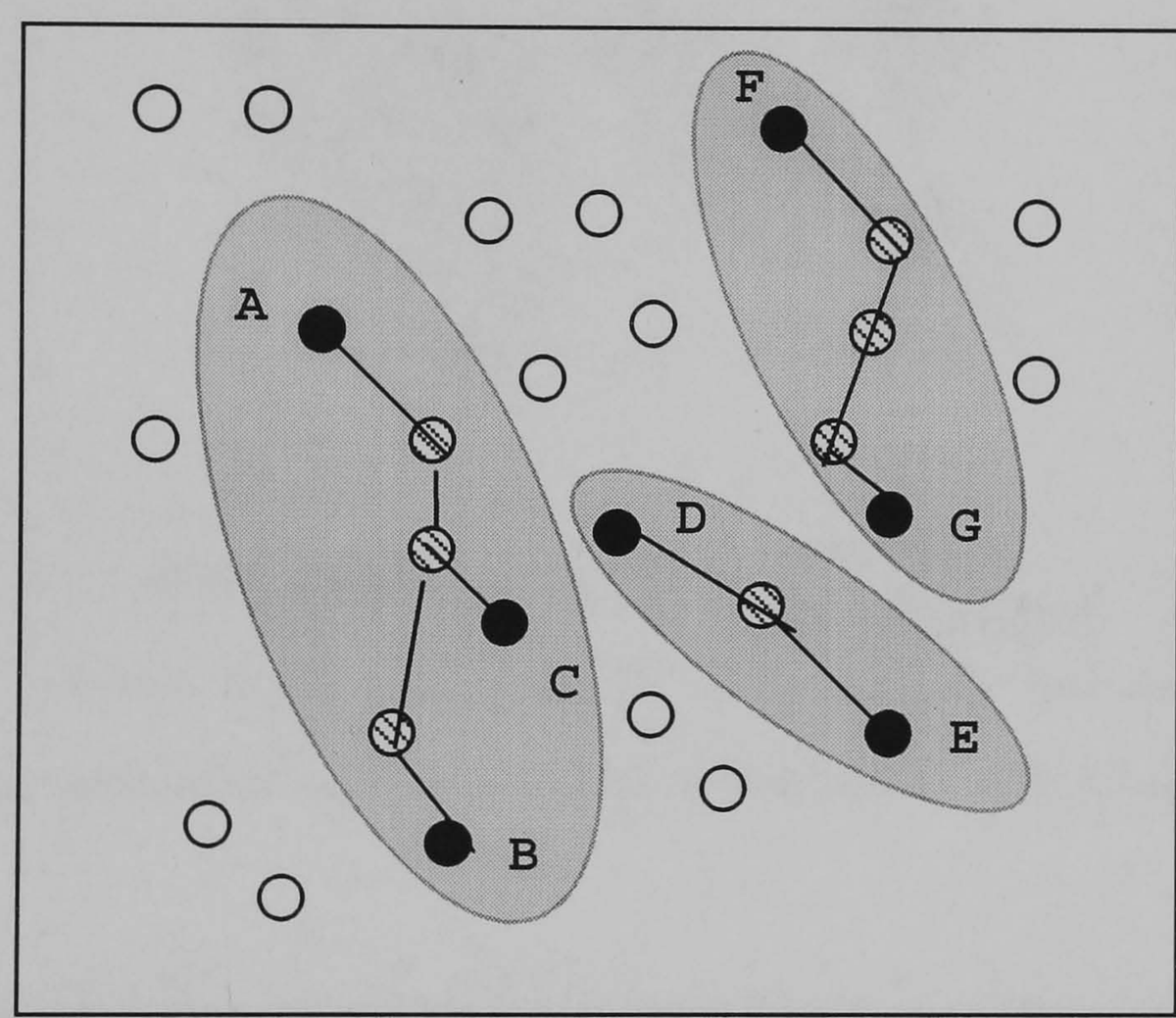
The *Track*-based scheme is also flexible in adapting the number of tracks when the traffic pattern changes. The Zone-based strategies requires every node to maintain its proactive zone to support the inter-zone bordercasting mechanism [8] [79], whereas in the *Track*-based scheme, proactive tracks are built from the associations between service consumer and service provider. Thus, the density of tracks grows or decreases with the number of these associations.

3.3.2 Proactive service discovery component

The proactive component of the proposed scheme is responsible for establishing and updating topological information of tracks through periodical validation messages. It also proactively



(a) the Zone-based service discovery



(b) the *Track*-based service discovery

Figure 3.1: Instances of proactive areas

maintains the states of available resources in a track.

Unlike conventional approaches in which reconstruction of proactive areas is carried out to cope with topological changes regardless the demands of their users, in the *Track*-based scheme the proactive areas are created and dismissed dynamically. When a service consumer success-

fully associates itself with its service provider a track is created between them at the same time. Nodes involved in a track will advertise their services to the whole track. As illustrated in Fig. 3.2, proactive areas are built on consumer-to-provider associations (e.g. $B \rightarrow A$, $C \rightarrow A$, $E \rightarrow D$ and $G \rightarrow F$). The lifetime of a track can be the same as the consumer-to-provider association but will be terminated after a certain idle period. This approach enables the *Track*-based scheme to eliminate unnecessary proactive efforts for relay nodes never used.

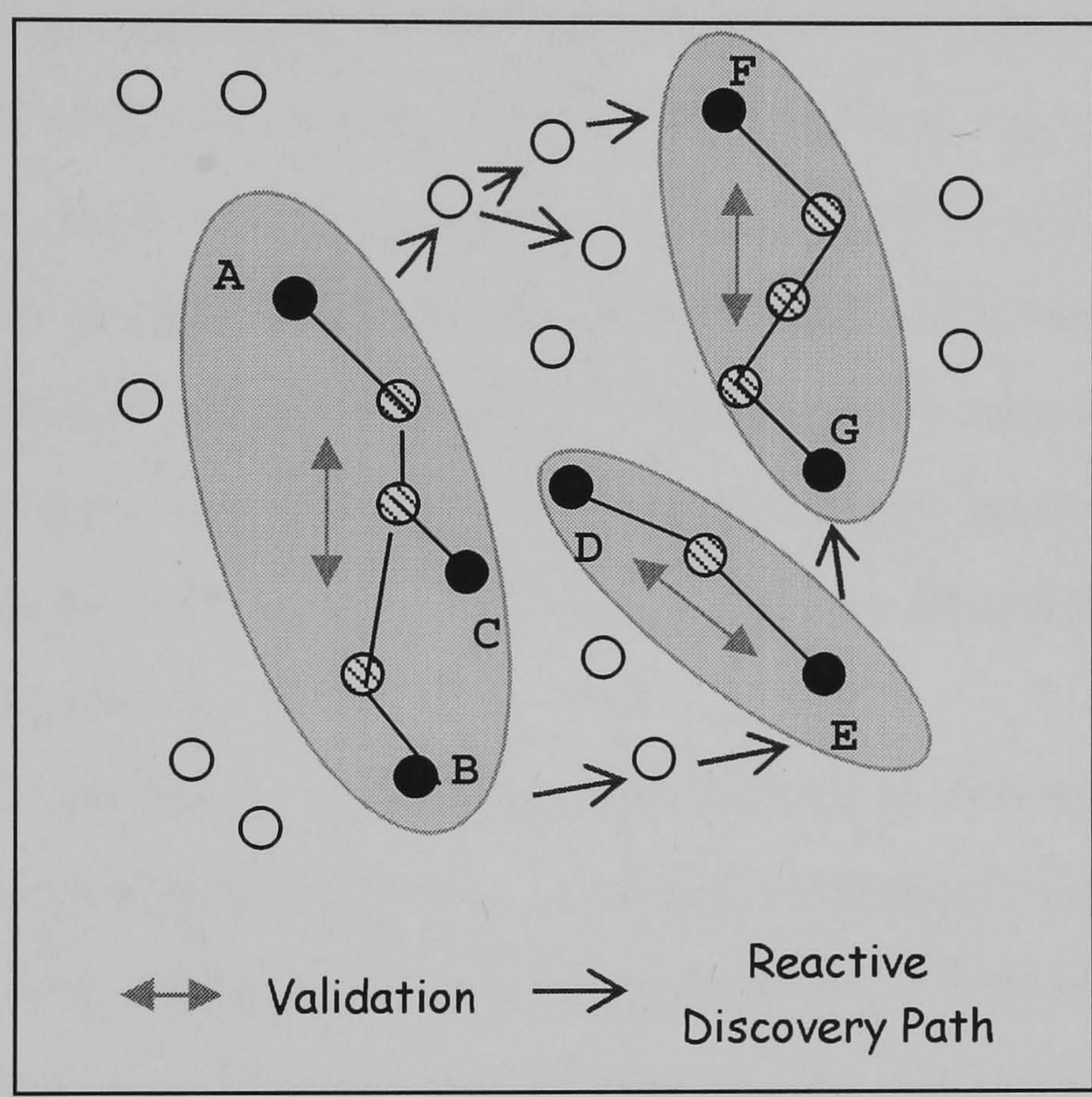


Figure 3.2: General view of the service discovery path between service consumer B to service provider F

The maintenance mechanism of a track is similar to that of a 'contact' as described in [37]. Periodical validation messages are dispatched from a service consumer to keep tracing its service provider and to detect broken connections in the track structure. Each relay node in a track forwards the validation message to its successor in the same track and replies to its predecessor with a validation acknowledgement. If the validation is not acknowledged for a certain time, the node assumes the connection to its successor is broken and invokes the local recovery procedure to repair the broken link. If the connection still can not be recovered after this procedure,

a path-broken message is sent back to the source of validation message to inform the service consumer. Note that a node may join multiple tracks at the same time, especially in networks of high node density. The maximum number of tracks that a node can participate in a time may be limited so as to constrain the traffic overhead for maintaining proactive tracks.

The local recovery mechanism is designed to help tracks survive frequent link failures. However, the recovering overhead for a broken track should be minimized so as to maximize the benefits of keeping such a track. Generally, the missing successor of a node could just move away when the link broken is detected. In addition, this successor may still be accessible by other one-hop away neighbours. It is also possible that some one-hop neighbours have obtained alternative paths towards the destination. Links to these neighbours can be used to replace the broken one. Based on these considerations, a two-hop broadcast method is used to recover the broken connections of a track. This method enables an efficient local (two-hop away) discovery of alternative paths to the end of a track without initiating global path discoveries from the service consumer.

The bandwidth overhead (i.e. transmission and reception of control packets) required to maintain a track can be estimated as $2\sigma(L-1)+2\mu(\Delta+\Delta^2)$ packets (emitted and received) per track, where σ is the validation rate for an L hops long track, μ is the link breakage rate and Δ the average degree of a node. It is a sum of validation and local recovery overhead where $2\sigma(L-1)$ represents the former and $2\mu(\Delta+\Delta^2)$ the latter.

3.3.3 Reactive service discovery component

The reactive component in the *Track*-based scheme is employed to disseminate service requests that can not be matched locally. Fig. 3.2 illustrates the reactive discovery path between consumer B and service provider F. When node B needs a specific type of resource or service, it first checks the availability of such service in its own service table which also comprises of service information of track $B \rightarrow A$. If the local service table cannot provide enough access information of the required service, the reactive component in node B creates a service request

and emits it to neighbouring nodes.

Since tracks are created on demand and not every node should be involved in a track, there would be mixed tracks and individual nodes (grey areas and white nodes in Fig. 3.2) coexisting in the network. The primary challenge in this unique design of the *Track*-based hybrid scheme is how to efficiently disseminate requests among the mixed tracks and individual nodes in the ad-hoc network.

In the proposed scheme, the proactive areas are made to be entities to participate in the procedure of reactive discovery, in that a track acts as a node in disseminating service requests. We define the *effective degree* Δ_E for efficient broadcasting among tracks and individual nodes. The value of Δ_E is decided by the reactive component according to the locality information of each track. The target is to achieve the maximum efficiency over the trade-off between querying overhead and querying delays/success rate. A track may choose to rebroadcast request at ends of the track and/or nodes with most neighbours, while limiting the receivers in its rebroadcasting under Δ_E. Fig. 3.2 shows the difference in the *effective degree* Δ_E between tracks $B \rightarrow A$ (Δ_E=2) and $E \rightarrow D$ (Δ_E=1). By adapting the value of Δ_E in real-time, the reactive component gets a fine grain control of reactive querying packets generated by each track.

Upon receiving a service request, a reply is sent back if the receiving node knows the location of the services/resources requested. Otherwise, individual nodes rebroadcast the request to their neighbours. For nodes belonging to tracks (e.g. node E and G in Fig. 3.2), the request is checked in the table of services in this track and is made aware through the whole track before the request is relayed to the neighbouring areas. The same request received by any node of the track next time is considered redundant and will be discarded. As shown in Fig. 3.2, although track $G \rightarrow F$ received the same requests from different nodes, only the first copy of the received requests will be processed.

As the density of proactive areas in the *Track*-based scheme is growing with the traffic, in an idle traffic pattern few proactive areas would be built up by this scheme. In this case the

performance of the *Track*-based scheme is similar to that of a purely reactive approach, i.e. the Ad-hoc On-demand Distance Vector Routing (AODV) [6]. Querying traffic (transmissions and receptions) at this stage is expected to be $2M_n + f_{rep}$ packets per query, where M_n is the number of links in an ad-hoc network with size n and f_{rep} represents the number of service reply packets. Service reply packets are unicasted from nodes knowing the location of the requested service to the source of a request. The value of f_{rep} can be roughly estimated according to the distance between the replying nodes and the requesting source [80]. With t tracks existing in the n nodes network, the reactive overhead of this approach can be estimated as $2(M'_n + \sum_{i=0}^{t} \Delta_{Ei}) + f_{rep}$, where M'_n is the number of links excluding those connected to the t tracks and Δ_{Ei} is the effective degree of track i. As M_n is growing faster than the network size, reduce M_n to M'_n contributes to the reduction of reactive packets overhead per request.

3.4 Simulation and Performance Evaluation

3.4.1 Simulation Settings

This Section presents the parameters and settings in simulating the behaviours of the *Track*-based scheme as well as the protocols used for performance comparison, namely, flooding and bordercasting. We implemented these service discovery mechanisms in the ns-2 [81] network simulator.

The link layer and physical interface were set to approximately reflect the characteristics of the Lucent WaveLAN™ card with 250m nominal propagation range. The distributed coordination function (DCF) of IEEE 802.11b was utilized as the MAC protocol with a maximum data rate of 2 MBit/s. The conditions of wireless links are assumed to be perfect in the simulations.

The simulations are carried out with several topology settings. In order to maintain a constant node density (an average of 8 for Δ), mobile nodes are uniformly distributed in an area growing from 700m$\times$700m to 1800m$\times$1800m, while the number of nodes ranges from 20 to

120 to reflect realistic scenarios. The mobility model in the environment was simulated using a Random-WayPoint (RWP) mobility model. In this model, each node randomly chooses a destination point in the area and moves towards it at a randomly velocity up to the chosen maximum velocity. When a node arrives at its destination, it then waits for a specified pause time before continuing the same pattern of motion. In our simulations, maximum velocities ranged between 0 m/s and 20 m/s (or 72km/hour), while the pause time was kept at 0 seconds to build a picture of constant moving mobile nodes.

In the simulation, the maximum velocity is fixed at 5m/s (or 18km/hour) while the simulation topologies are being varied. While changing the maximum velocity, the number of nodes was fixed to be 40 and the simulation area 1000m$\times$1000m.

Among the service discovery protocols simulated, *flooding* and the reactive component of the *Track*-based scheme are extensions to AODV, while *bordercasting* is derived from ZRP with 2-hop zone radius, and is configured with query detection and early termination as described in [79]. To simulate service discovery mechanisms, other parameters that require considerations are the number of service consumer to provider pairs and the service query rate. Three pairs of service providers for three categories of services are specified in the simulations resulting in a total of six providers. The number of service consumers and query rate in traffic patterns is also varied and noted as T(C, Q), i.e. each of C service consumers initiates Q queries every second for each kind of service in turn. If not stated, the default traffic pattern is chosen as T (12, 1), which represent a medium traffic load in the network, for varying network size or mobility. While we varying the traffic patterns, the network size is fixed at 40 and maximum speed fixed at 5m/s.

In the simulation results illustrated, each data point corresponds to a mean of 100 repeated measurements with different motion patterns and topologies. Error bars are also plotted to indicate the 95% confidence interval of each data point.

Table 3.1: The Control Packets Overhead of Different Traffic Patterns (Tx+Rx packets/query)

Traffic pattern	T(8,0.5)	T(12,1)	T(14,1.5)
Reactive overhead	429.44	335.88	140.59
Proactive overhead	97.28	126.22	322.52

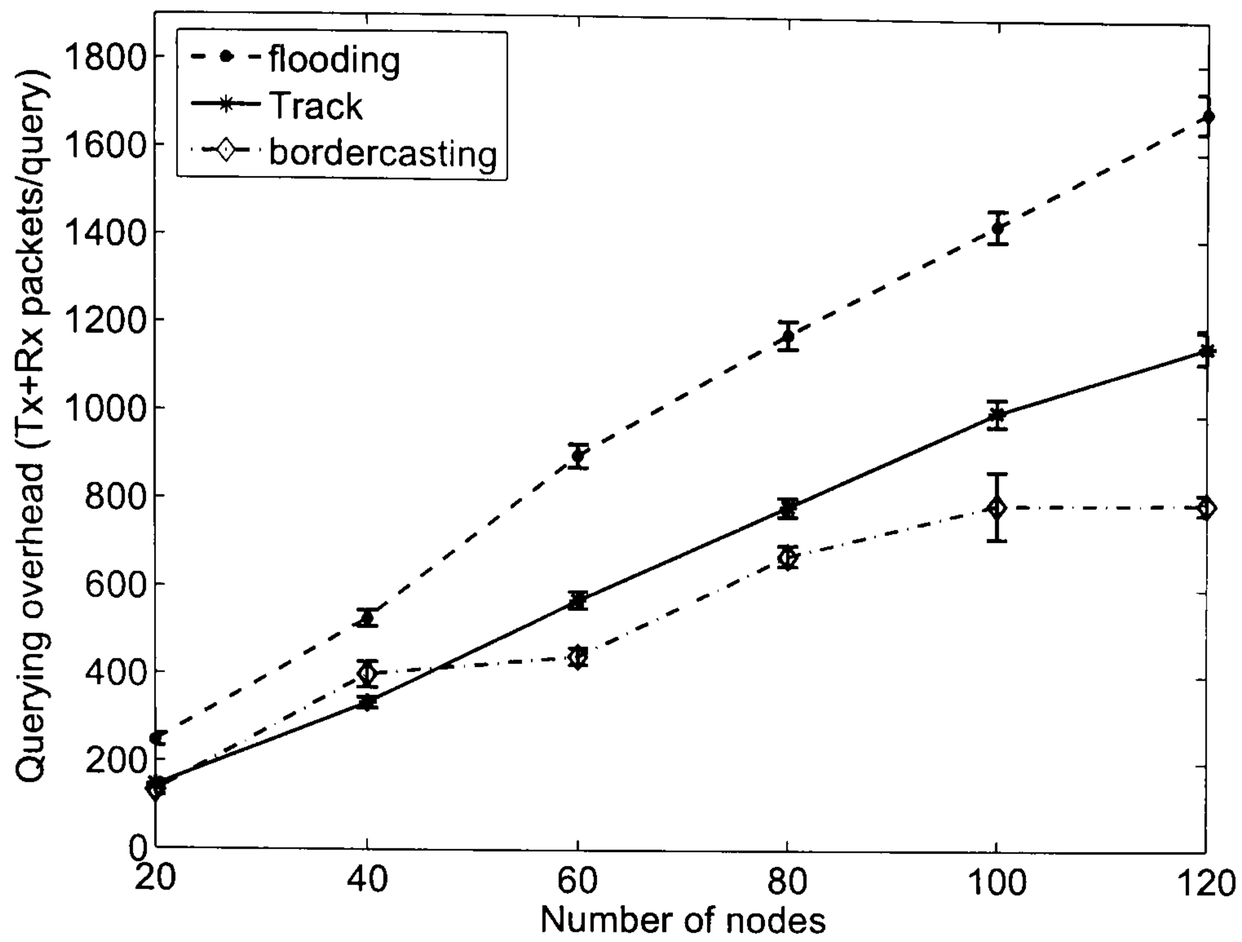
3.4.2 Quality of Service Measures

In order to compare the proposed scheme with *flooding* and *bordercasting*, the system's performance is measured with several QoS metrics, i.e. control packets overhead, querying delay and first attempt success rate. Control packets overhead consists of querying overhead and proactive overhead. Querying overhead gives a quantitative view on traffic produced for a service query, while proactive overhead represents those required to maintain proactive areas of an ad-hoc network averaged over a certain period of time. HELLO beacons are not included as the quantity of these beacons is independent from the discovery schemes used. The querying delay equals the time required between when a node sends out a service request and when successfully receive a positive response. The first attempt success ratio is used to reflect the robustness of each protocol by the ratio of successful queries in one try over the total number of generated queries. This metric is more straightforward than those used in literature [37] [39] where the success rate is obtained in simulations with up to 3 retries for failure queries.

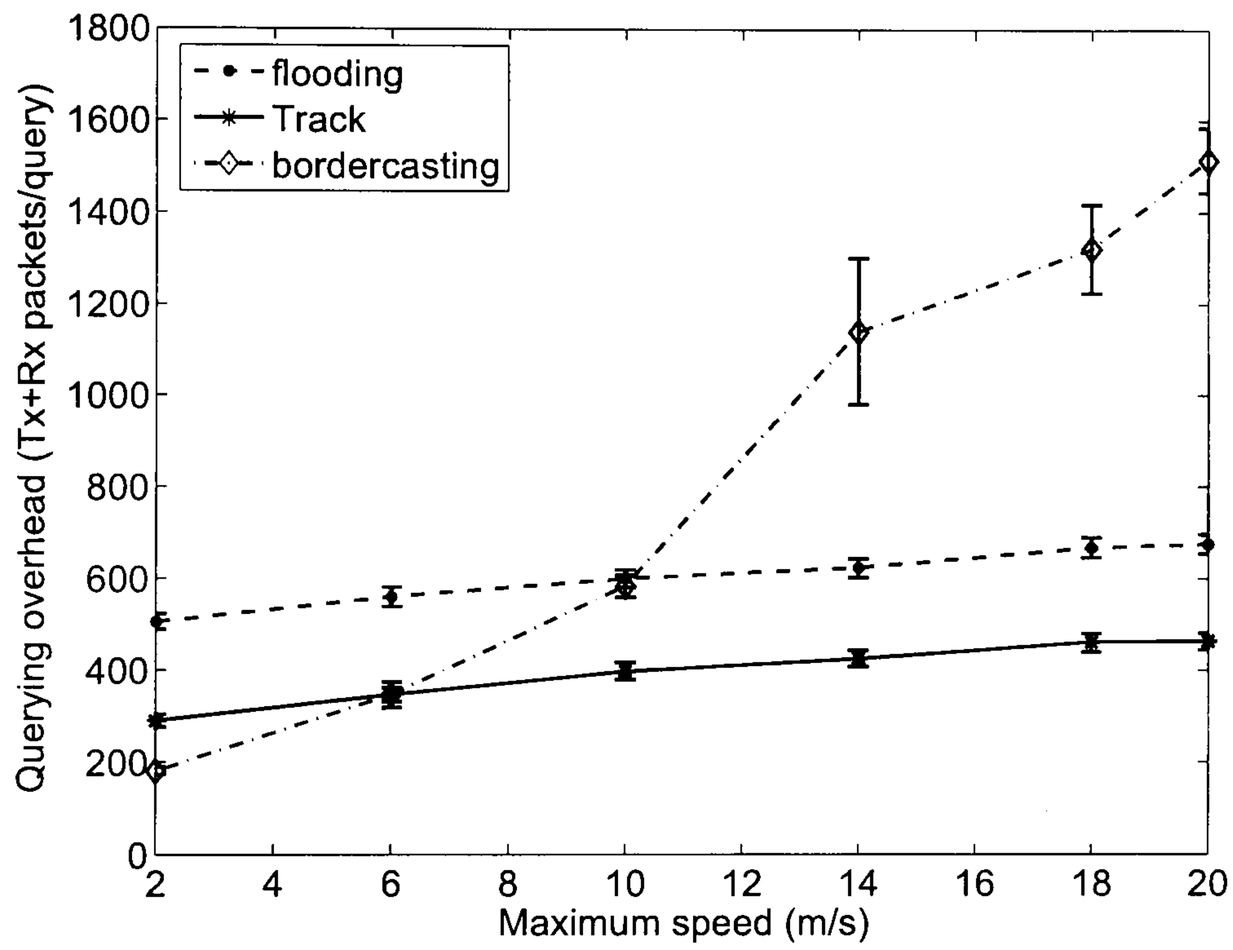
3.4.3 Control packets overhead

The behaviour of the *Track*-based scheme) is firstly listed in Table 3.1 for various traffic patterns (e.g. from 8 service consumers, 0.5 queries per second to 14 service consumers with 1.5 queries per second). We can see from Table 3.1 that the reactive traffic decreases when the network becomes busier. This is the natural cause of increased track density. Higher track density makes the behaviour of the *Track*-based scheme become more proactive, as also indicated by the curve of proactive traffic.

Fig. 3.3 presents a comparison of the querying overhead between three protocols against



(a) Querying overhead vs. Network Size



(b) Querying overhead vs. Mobility

Figure 3.3: Querying overhead

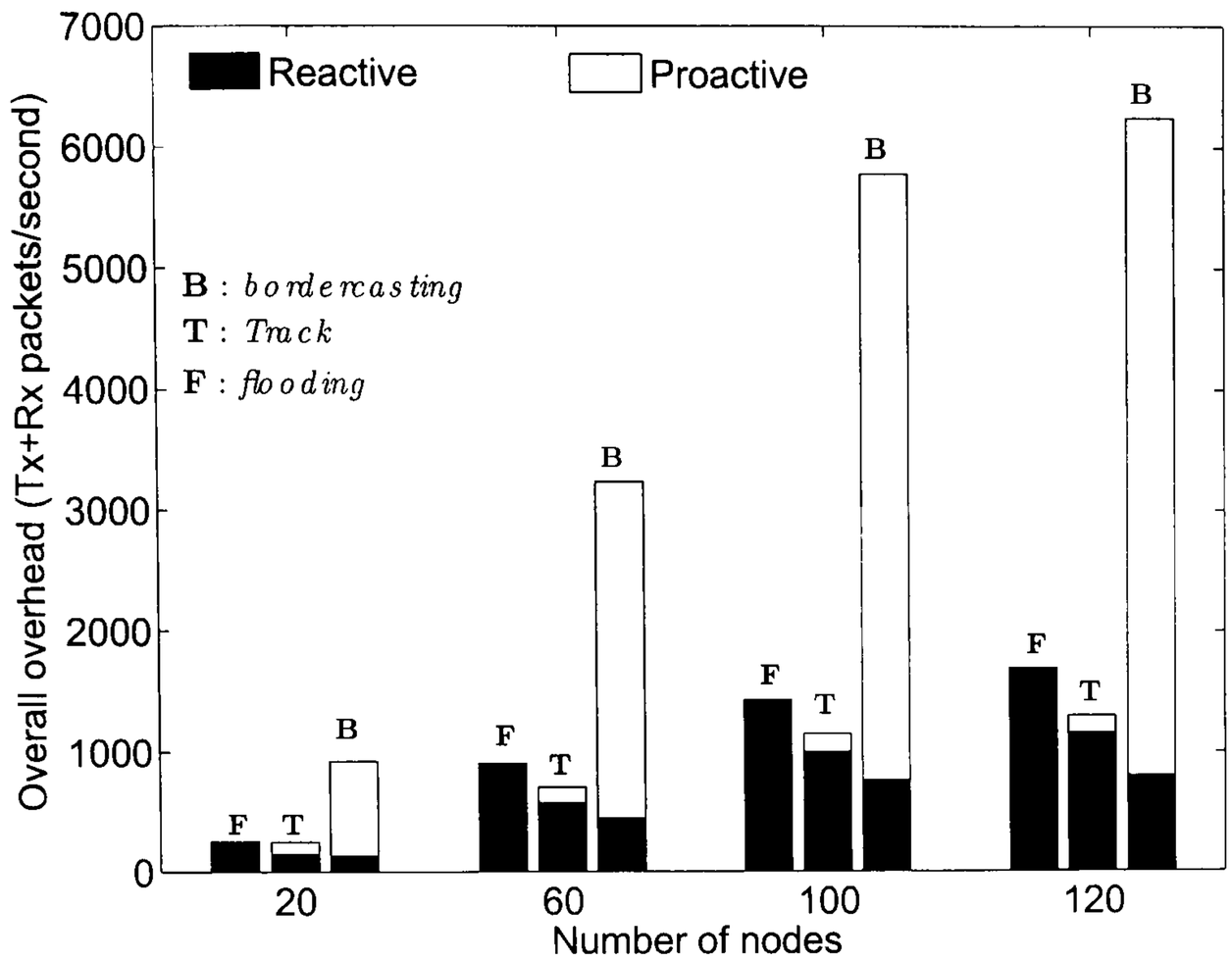
varying network size and mobility. The overhead measured includes transmission as well as reception of control packets in each operation. Therefore the querying overhead of *flooding* is about $2M_n$ packets per query. From Fig. 3.3, we can see that *Track* achieved at least 30% less traffic than *flooding* in transmitting requests and replies in all the scenarios simulated. Also from Fig. 3.3a, it seems that *bordercasting* outperforms the rest with its efficient inter-zone

bordercasting techniques. However, Fig. 3.3b shows that the performance of *bordercasting* degrades dramatically when the maximum speed is increased. This result is due to that *bordercasting* delivers requests according to its bordercast tree topology that may be inadequate in frequent link failures or collisions caused by high mobility.

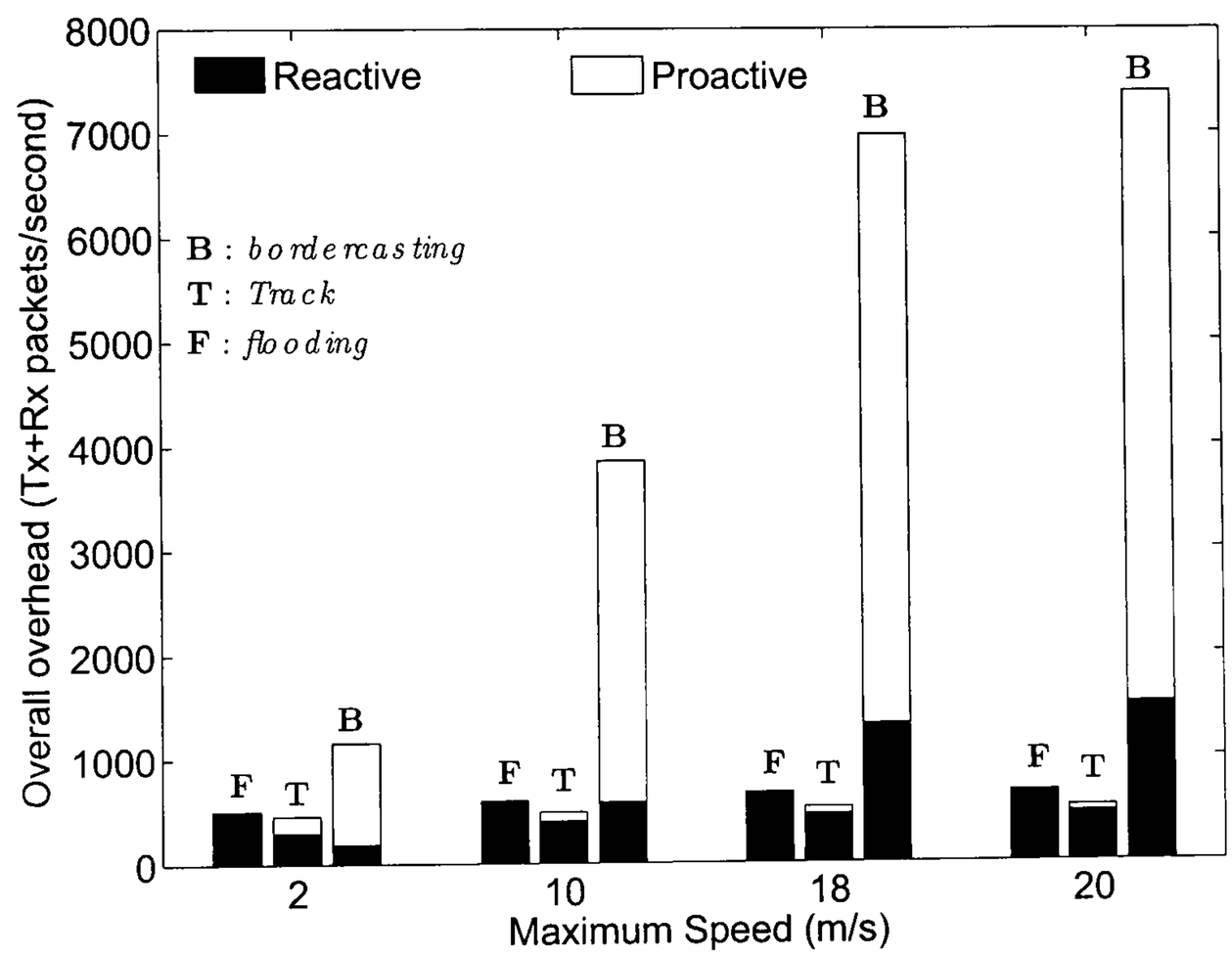
Fig. 3.4 gives the overall control packets overhead comparison with indications of the contributions of both reactive and proactive overhead. The overall overhead measured are quantified as packets per second when the reactive traffic is generated from 1 query per second. As shown in Fig. 3.4, proactive traffic constitutes of 80-90% of the overall overhead of *bordercasting* in all scenarios. As the two dimensional proactive zone of *bordercasting* requires frequent exchanging of control packets (e.g. link state packets) to sense topology reconfigurations. Such maintenance expenses are very sensitive to the size or mobility of the ad-hoc network. On the contrary, reactive traffic constitutes of 60-90% of the overall overhead of *Track*. For *flooding*, all of the overall overhead is produced only in reactive queries. However, with such a low query rate, the proactive traffic of the *Track*-based scheme does not lead to worse performance than *flooding*. In fact, Fig. 3.4a and Fig. 3.4b show that the overall traffic for *Track* is about 25% less than that of *flooding* in large network size or high mobility scenarios. There are several reasons why the *Track*-based scheme can achieve such a significant reduction in proactive traffic. Firstly, the one dimensional structure of tracks requires very light traffic for maintenance or local recovery. Secondly, the number of tracks are adapted to fit the traffic pattern of the mobile ad-hoc network, whereas in *bordercasting* the proactive zones are independent of the pattern of traffic.

3.4.4 Querying delay

The average querying delays of successful responded queries are given in Fig. 3.5a against increasing network size and Fig. 3.5b against varying maximum speed. Generally, hybrid approaches perform better than *flooding*, as services located in a proactive area are ready to use for all the members of this area. *bordercasting* achieved the best position due to its robust inter-



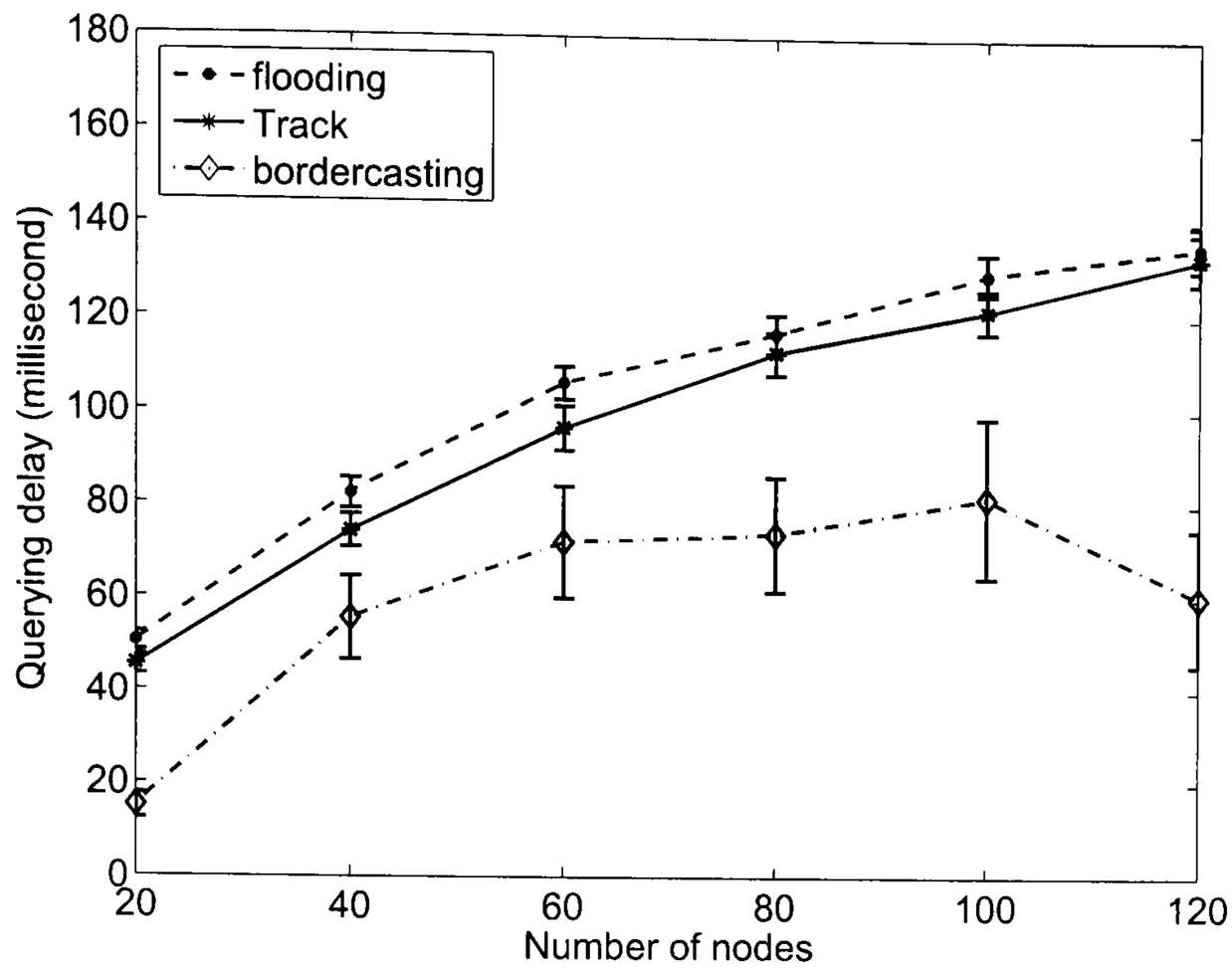
(a) Overall overhead vs. Network Size



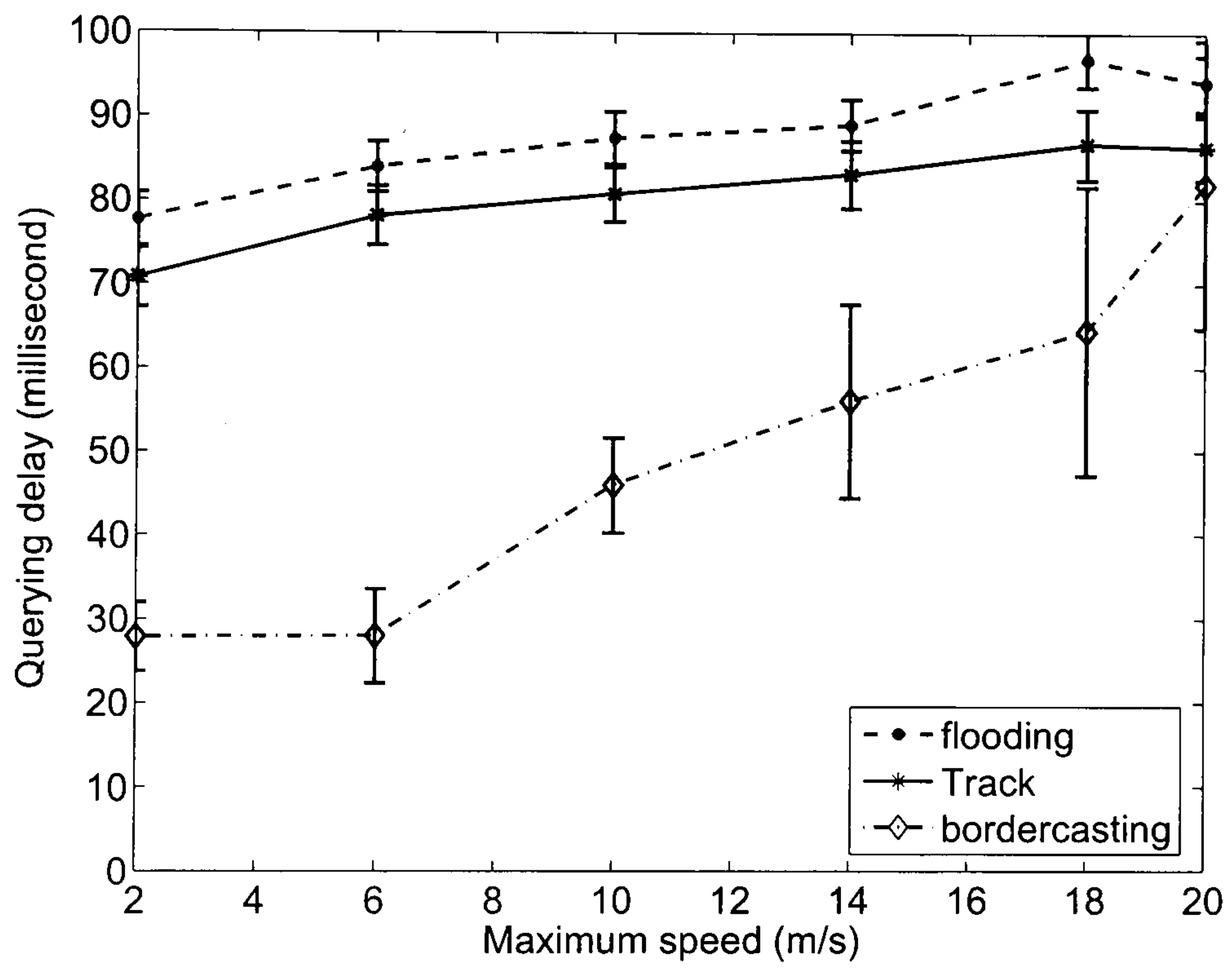
(b) Overall overhead vs. Mobility

Figure 3.4: Overall control packets overhead

zone bordercast mechanism which disseminates requests among proactive zones throughout the network. The querying delay of each approach tends to increase with network size or mobility, in which service requests would have to travel for longer distances and experience more link breakages. Fig. 3.5a shows one exception produced by *bordercasting* at network size 120.



(a) Querying delay vs. Network Size



(b) Querying delay vs. Mobility

Figure 3.5: Querying delay

The reason behind this exception is because in large networks queries are more vulnerable to collisions and more difficult to be disseminated over long distance. Hence the point plotted at network size 120 for *bordercasting* represents mostly short range queries. This problem will be given more explanations in the following subsections. The *Track*-based scheme is worse than

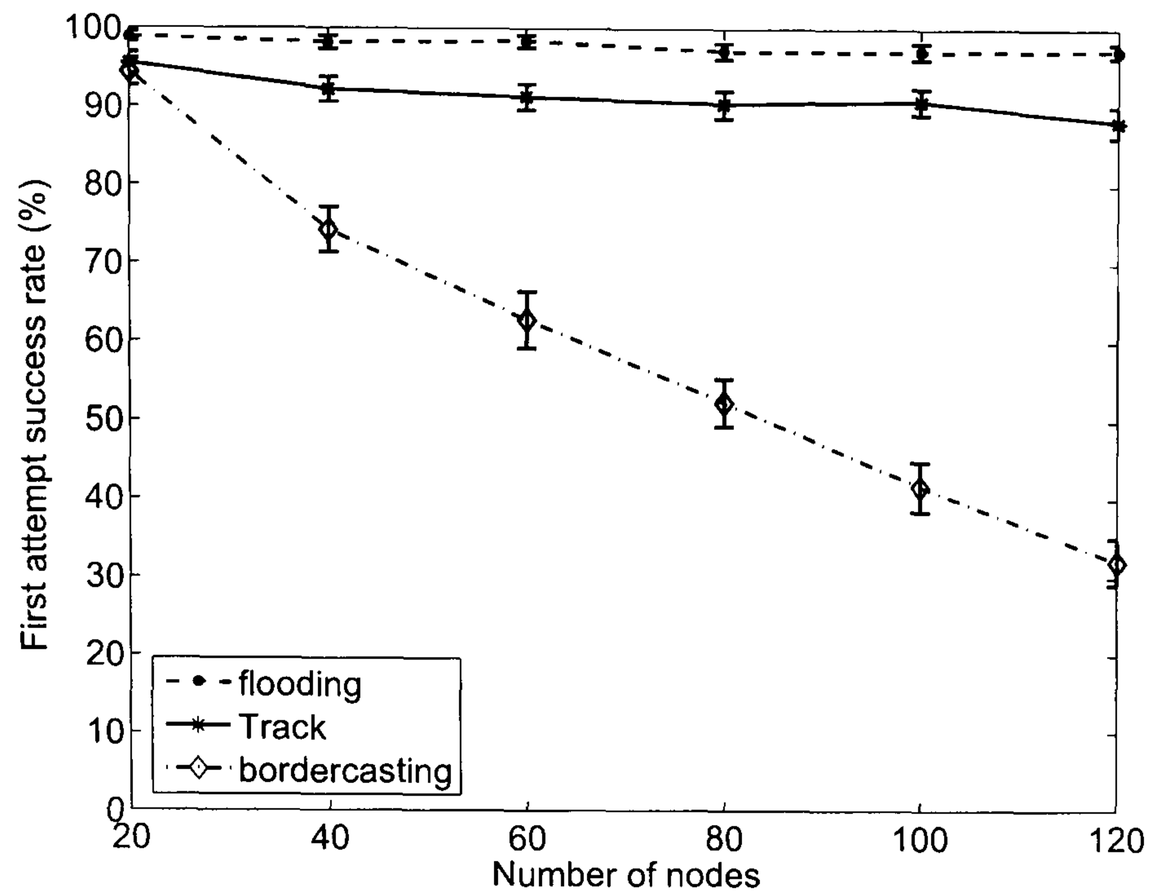
bordercasting in terms of querying delay as it has to disseminate requests among the tracks and individual nodes mixed network. However, it is an efficient trade-off between the amount of overhead reduced by the *Track*-based scheme and the extra time required for a query.

3.4.5 First attempt success rate

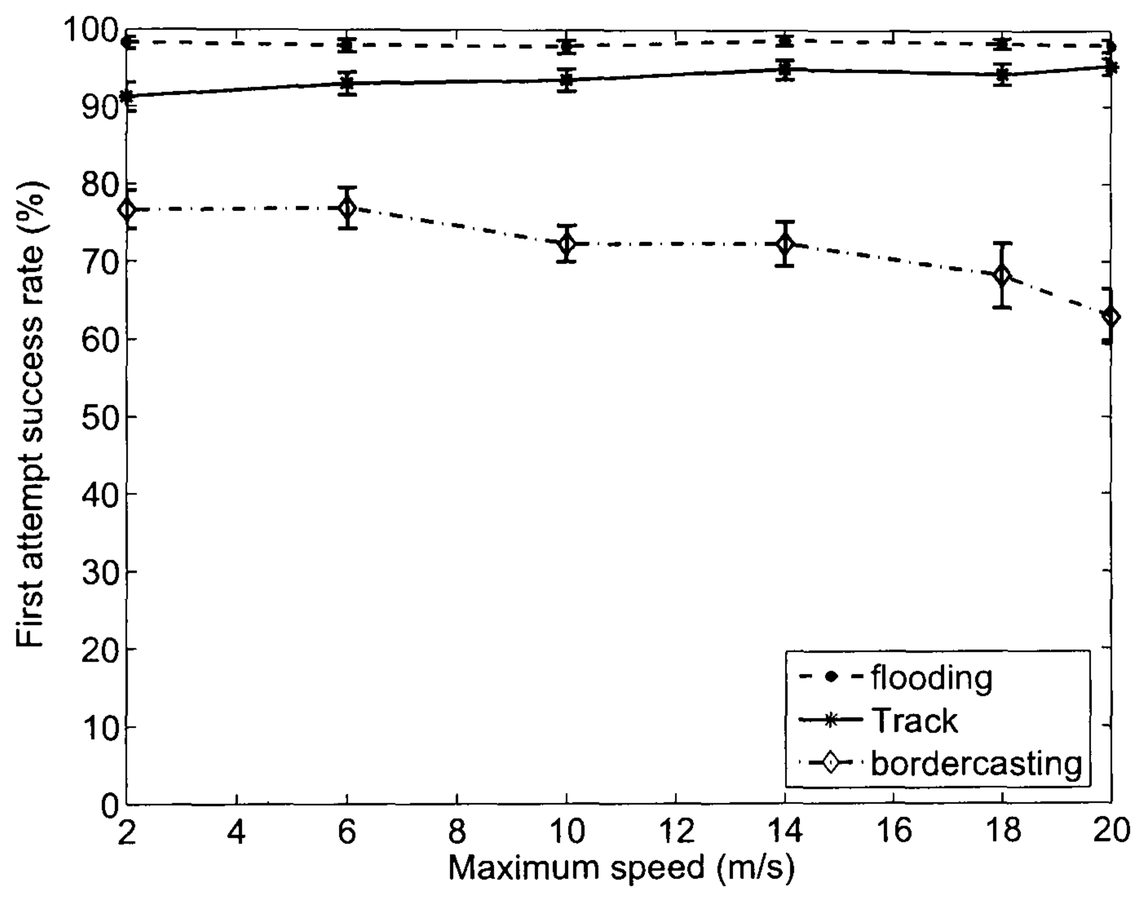
The success rate of attempting a service query in one try is illustrated in Fig. 3.6a and Fig. 3.6b against growing network size and mobility, respectively.

Despite *bordercasting* is powerful in delivery of requests among proactive zones and is efficient in reducing querying traffic, the inter-zone bordercasting mechanism - the key component of ZRP and its variants - heavily relies on the up-to-date topology information of the proactive zone to accurately disseminate requests [39]. Because link state packets are more likely to be lost in collisions or frequent link breakages, it is more difficult to precisely maintain the current version of topology as the network size or mobility increase. Therefore, the first attempt success rate of *bordercasting* is dropping significantly with the increasing of network size or mobility, as can be seen from Fig. 3.6a and Fig. 3.6b. Although collision-free solutions [82] may help to release this problem, we can still observe the weakness of relying on proactively maintained topology to unicast service requests.

Also plotted in Fig. 3.6, the curves of *Track* and *flooding* do not show much sensitivity on changing network size or mobility. *Track* keeps the first attempt success rate over 90% in most cases. This value is generally decreasing with the network size. However, from Fig. 3.6b it is approaching that of *flooding* when the nodal mobility is increased. This is because more tracks are likely to be broken in higher mobility scenarios resulting in the more reactive behaviour of the *Track*-based scheme. From Fig. 3.6, we can see that the *Track*-based scheme does not trade much of the querying success rate off for the reduced querying or proactive overhead.



(a) First attempt success rate vs. Network Size



(b) First attempt success rate vs. Mobility

Figure 3.6: First attempt success rate

3.5 Summary

This Chapter presents the *Track*-based scheme for service (and/or routing) discovery in mobile ad-hoc networks. The *Track*-based scheme is a proactive and reactive hybrid service discovery approach. It tries to minimize the proactive traffic with its one dimensional track-like structure of proactive areas. It also adapts the density of proactive tracks to different

traffic patterns for better efficiency. The proposed scheme has been implemented in the ns-2 network simulator. Its performance is compared with purely reactive (e.g. *flooding*) and Zone-based hybrid (e.g. *bordercasting*) protocols against several QoS measures (e.g. control packets overhead, querying delay and first attempt success rate) in a wide range of network scale and mobility scenarios. Preliminary results show that, the *Track*-based approach requires minimal overhead to maintain its light-weight hybrid framework and performs more efficiently in balancing the trade-offs between control packets overhead and other performance measures among the approaches compared.

Chapter 4

Adaptive Multi-copy Routing for Sparse Mobile Ad hoc Networks

4.1 Introduction

In MANETs where nodes are sparsely distributed, fully connected source-to-destination paths do not exist most of the time. As traditional routing protocols can only delivery packets over connected end-to-end paths, they would fail in such networks. This Chapter focuses on energy-efficient packet delivery for sparse MANETs. Recently, *store-and-forward* based packet delivery has been proposed to address this. However, existing *store-and-forward* strategies (e.g. multi-copy relaying) do not adapt to the variations in the network conditions due to their reliance on the source to determine the relaying process. In this Chapter, a new *store-and-forward* based scheme, Adaptive Multi-Copy Routing (AMR), is presented for packet delivery in MANETs of sparse node distribution. In this scheme, instead of relying on the source, intermediate nodes independently constrain packet relaying according to the current network conditions in order to limit the delivery delay according to a given QoS target (e.g. the targeting end-to-end delay and packet arrival rate*) making it more energy-efficient. The proposed scheme are implemented in the ns-2 simulator and its performance is validated in simulations that cover a wide range of scenarios.

*Due to the application level constraints or limited buffer space, late arrival packets may be discarded at their destinations. Packet arrival rate is referred to the proportion of packets that arrive at their destination with an end-to-end latency that is equal to or less than the delay constraint/target.

This Chapter is organized as follows. Section4.2 inroduces related work on packet delivery in sparse or intermittently connected MANETs. In section4.3, the principle and the mechanisms of the proposed scheme of Adaptive Multi-copy Routing (AMR) is presented in detail. The simulation based performance evaluation of the AMR scheme is given in Section4.4. Section4.5 summarizes this Chapter.

4.2 Related Work

To perform communications in an ad-hoc network of sparsely scattered nodes, the most common solution is the *store-and-forward* routing. In the *store-and-forward* routing paradigm, a node that has packets to send or relay would store these messages in its buffers if it is not possible to forward the packets. Packets in the buffers are then relayed from a node to another over the intermittently available links accompanied with their movements. Thus, a packet could be eventually delivered if, over time, the nodes carrying a copy of the packet continue to move around and one of them reached its destination.

A number of routing strategies, based on the concept of *store-and-forward*, have recently been introduced for sparse MANETs or those of similar properties (e.g. intermittently connected or delay tolerant networks). Epidemic routing [9] carries out packet relaying in a way that is reminiscent the concept of flooding. In epidemic routing, a node that has packets to send or relay would store these messages in its buffers if it is not possible to forward the packets. Whenever the node meets another node, the two nodes exchange copies of packets that they do not have in common. Thus, as nodes continue to move around, the packets they are carrying would reach their destination over time. In practice, epidemic routing suffers from high usage of network bandwidth. A simple way to reduce the excessive traffic produced by epidemic routing is to only replicate packets to nodes that have higher delivery probability [48] or to those with similar mobility pattern to the destination [83]. A more aggressive way to reduce the traffic overhead is to use single-copy routing [84], i.e. each relay node forwards at most

one copy per packet. Although these relaying strategies can substantially reduce the delivery overhead, they increase delivery delay and decrease the success rate of delivery. To address the delay-overhead dilemma, multi-copy relaying strategies are proposed in [10] and [85] to limit the delivery overhead with a *replication factor* $\mathcal{R}$, which is defined as the exact number of copies to be produced for each packet for an expected delivery delay. The $\mathcal{R}$ packet copies can be distributed through a process that can be represented as a balanced binary-tree [10] or the locally-optimal tree [85]. The EBEC scheme [50] is a more complex version of multi-copy routing. EBEC utilize erasure coding to generate $\mathcal{R} \times K$ message blocks per packet for enhanced redundancy. These message blocks are selectively distributed among nodes according to their probability of reaching the destination for less delivery delay.

A common problem in conventional multi-copy routing protocols [10] [50] [85] is that they rely heavily on the source to control the relaying process. A source determines the relaying process by setting the *replication factor* before it initiates packet relaying. However, it is normally difficult and very complicated for the source to estimate a suitable $\mathcal{R}$ based on its limited knowledge about the network. If local network conditions (e.g. network mobility or delivery probability) change during the relaying process, the predefined *replication factor* would no longer be appropriate leading to a degraded performance.

In this Chapter, an efficient *store-and-forward* based scheme, Adaptive Multi-Copy Routing (AMR) is presented, for packet delivery in sparse MANETs. In this scheme, instead of using a source-defined *factor*, each of the intermediate relay nodes independently decide whether to further replicate a packet based on the local network conditions and the target of end-to-end delay and packet arrival rate. Thus, the scheme does not rely on the source for determining the *replication factor* and is more efficient in coping with unpredictable variations in the local network conditions than source-defined relaying strategies. Therefore it is more suitable for highly dynamic sparse MANETs.

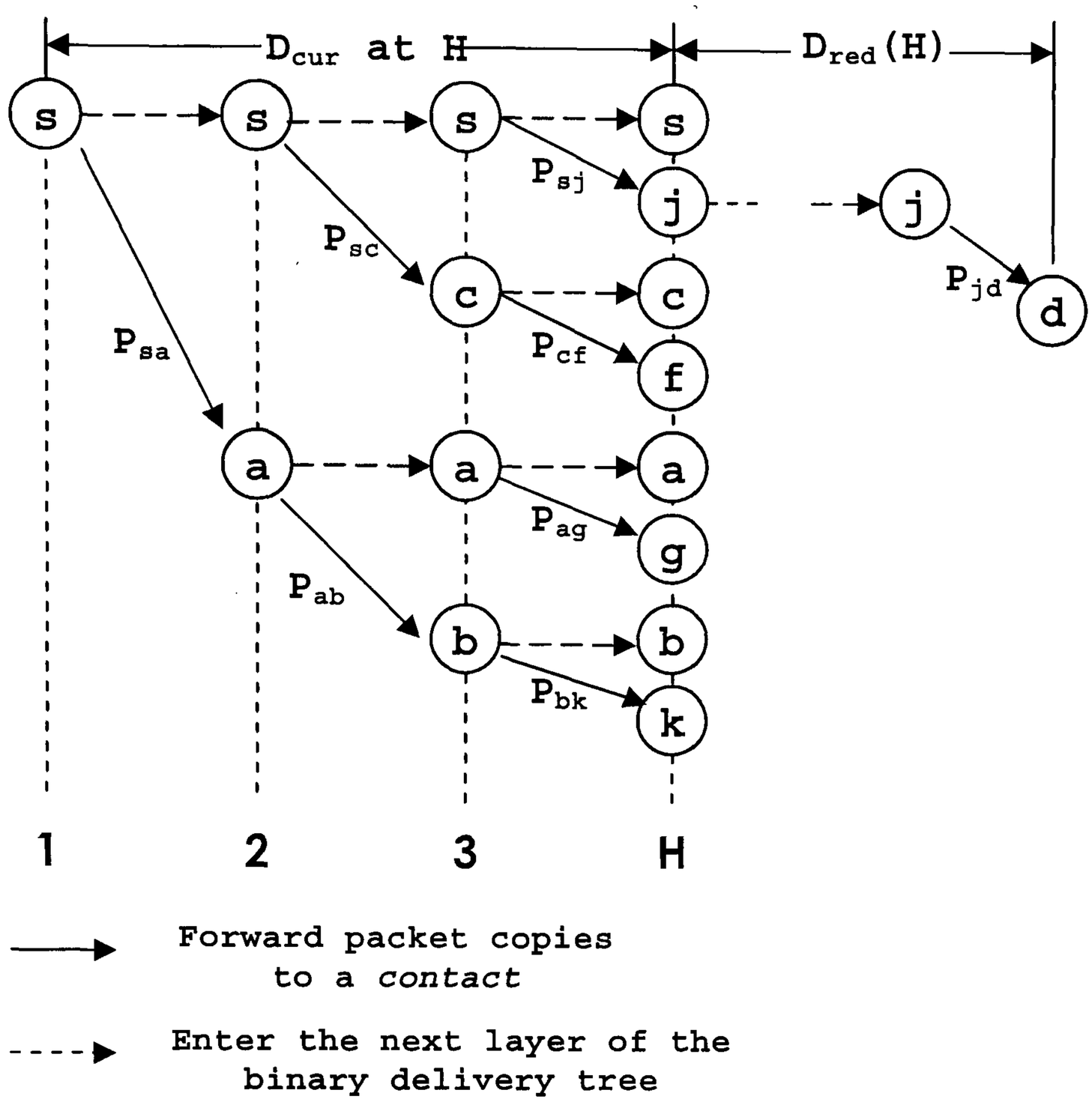


Figure 4.1: The Binary Relaying Tree

4.3 Multi-copy Packet Relaying with Distributed Adaptations

In this section, the proposed Adaptive Multi-copy Routing (AMR) scheme is introduced. The relaying process in AMR is partially inspired by the binary spray&wait routing [10]. As shown in Fig. 4.1, the relaying process of a packet in AMR starts from the source (node s) and finishes at the destination (node d). When s meets† node a, a copy of the packet is forwarded to a. The packet is successfully handed over to a with a probability $\mathcal{P}_{sa}$, which is defined as the one-hop delivery probability and is determined by the packet error rate of the wireless link,

†by a meets b we mean $d_{ab} < r$, where d_{ab} is the distance between a and b and r is the radio propagation radius of a mobile node. a (b) is called a ***contact*** of b (a) when a meets b.

link duration and the position of the packet in the buffer, etc. Having passed a, s keeps moving on and giving a copy of the packet to each node it meets on the way. Nodes successfully receiving a copy of the packet (including the source) would further replicate the copy to those that they meet on their way until either they think that they have spread enough packet copies or they have delivered the packet to the destination. Such a replication process of a packet can be represented as a binary tree with the source of the packet as its root. The distribution of delivery delay and overhead of a packet is determined by the depth of the binary delivery tree and the networking parameters (e.g. the mobility of mobiles, the size of the network ground and the node density). There is a possibility that the delivery delay become infinite and a packet will not be able to reach its destination by the time when it is dropped off the buffers.

Unlike the conventional multi-copy protocols in which the source chooses a suitable depth of the delivery tree for a delay target by specifying the *replication factor* $\mathcal{R}$. The AMR scheme lets the relay nodes independently decide the depth of the relaying process for each packet they are carrying. As a packet has spent some time (referred as $\mathcal{D}_{cur}$, see Fig. 4.1) to reach a relay node, the purpose of limiting the depth of the delivery tree at a relay node is to limit the residual delivery delay (referred as $\mathcal{D}_{red}$, see Fig. 4.1), instead of the end-to-end delivery delay, for a delay (and/or packet arrival rate) target. Such a distributed adaptation of the relaying process has two major benefits: firstly, the relay nodes have fresher knowledge of the networking conditions than the source hence their estimations of the delivery delay is more accurate. Secondly, estimating the residual delivery delay of a packet at the relay nodes is simpler than to estimate the end-to-end delivery delay at the source, i.e. $\mathcal{D}_{red}$ is independent of the number of nodes in the network. The details of the algorithms and mechanisms of the AMR scheme is given in the following subsections.

4.3.1 Distributed Estimation of the Residual Delivery Delay

Distributed estimation of the residual delivery delay $\mathcal{D}_{red}$ of a packet is the key mechanism of the AMR scheme. $\mathcal{D}_{red}$ is a function of both $\mathcal{R}$ and the mean intermeeting times[‡] $\bar{\mathcal{T}}_{int}$. Recent study [10] has suggested that $\frac{\bar{\mathcal{T}}_{int}}{\mathcal{R}}$ is a good approximation of the residual delivery delay when all nodes in the network perform Independent and Identically Distributed (IID) random walks. However, the estimation of the mean intermeeting times $\bar{\mathcal{T}}_{int}$ in [10] does not reflect the impact of network mobility. Groenevelt in [86] has proposed mathematic models to estimate the mean intermeeting times between nodes moving at a specific speed for both Random Waypoint model (RWP) and Random Direction Model (RD) at their steady states. Suppose that all nodes in the network are constant moving in the RWP model, the event of one node encounters another is a *Poisson* process and the intermeeting times is can be modeled by a exponential distribution. The Cumulated Distribution Function (CDF) of the intermeeting times can be expressed as:

$$\mathcal{F}(x, \bar{\lambda}) = \begin{cases} 1 - e^{-\bar{\lambda}x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4.1}$$

where $\bar{\lambda}$ is the expected arrival rate of *contacts*. The expected intermeeting time (denoted as $\bar{\mathcal{T}}_{int}$) can be estimated with the following equation [86]:

$$\bar{\mathcal{T}}_{int} = \frac{1}{\bar{\lambda}} = \frac{\mathcal{L}\mathcal{W}}{2\omega r \bar{v}} \tag{4.2}$$

where $\omega \approx 1.3683$ is a specific constant for the RWP model, r is the radio radius, $\bar{v}$ is the mean relative speed between nodes of the network, and $\mathcal{L}$ and $\mathcal{W}$ is the side length and width of the geographical network area, respectively.

If a node does not have the knowledge of one or more parameters listed in Eq. 4.2, it is still

[‡]Defined as the duration between the time when a node meet the other node and the next time when they meet again

possible to estimate the mean of intermeeting times $\bar{\mathcal{T}}_{int}$ using the mean of the intermeeting times samples that it has collected. As the size of the network area, the radio radius and the model of the nodal movements are normally fixed, it is reasonable to assume that only the mean relative speed $\bar{v}$, which varies with time or the type of mobile devices, is unknown to a node. It is easier for a node to sample relative speeds from any *contacts* than to sample intermeeting times from those that it have met before. A single node can estimate the time-varying $\bar{v}$ by:

$$\bar{v} \approx \frac{1}{|\mathcal{V}^{\tau}|} \sum_{\hat{v}_i \in \mathcal{V}^{\tau}} \hat{v}_i \tag{4.3}$$

where $\hat{v}_i$ is the relative speed sample that a node measured from a passing *contacts* i and $\mathcal{V}^{\tau}$ is the set of speed samples that a node collected during past τ seconds. τ should be designed to enable a node to collect enough samples (e.g. over 30 [87]) for an accurate estimation of $\bar{v}$ and to make the estimate emphasize on recent changes in the network mobility.

If the relaying process of a packet can be characterized by the binary delivery tree as depicted in Fig. 4.1, a node can estimate the number of copies that has been replicated for the packet as $2^{\mathcal{H}-1}$ provided that it knows the current depth $\mathcal{H}$ of the delivery tree. Supposing the one-hop delivery probability is 1, a single relay node could estimate the residual delivery delay $\mathcal{D}_{red}$ when the depth of the binary delivery tree is $\mathcal{H}$ by:

$$\mathcal{D}_{red}(\mathcal{H}) = \frac{\bar{\mathcal{T}}_{int}}{2^{\mathcal{H}-1}} = \frac{1}{2^{\mathcal{H}-1}\bar{\lambda}} \tag{4.4}$$

Due to the fact that a thinned *Poisson* process is still a *Poisson* process [86], the distribution of residual delivery delays in the system is again exponential with $2^{\mathcal{H}-1}\bar{\lambda}$ being the *rate parameter*. Therefor, the CDF of the residual delivery delay conditioned on the depth of the delivery tree being $\mathcal{H}$ can estimated by:

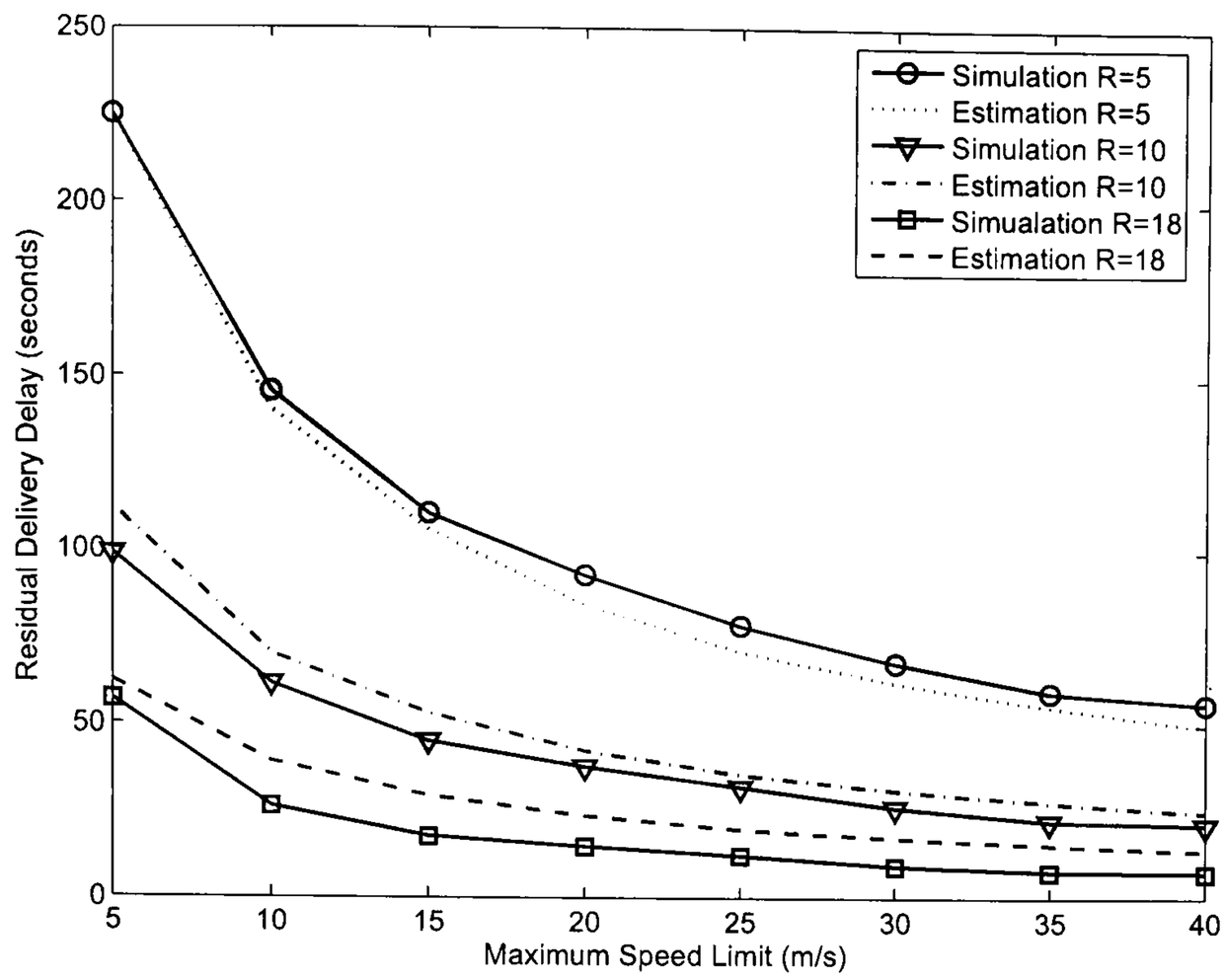


Figure 4.2: The residual delivery delay of a packet when there are R of its replicates in the network

$$\mathcal{F}_{red}(x, \bar{\lambda}|\mathcal{H}) = \begin{cases} 1 - e^{-2^{\mathcal{H}-1}\bar{\lambda}x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4.5}$$

Fig. 4.2 shows residual delivery delays of packets estimated by relay nodes using Eq. 4.4 as well as those obtained from simulations (please refer to Section4.4 for experimental settings) after 5, 10 and 18 copies of a packet have been distributed in a network of perfect one-hop delivery probability. Fig. 4.2 confirms that, given the knowledge of both the value of $\mathcal{R}$ and the networking conditions, a relay node does have the ability to accurately estimate the residual delivery delay of a packet.

4.3.2 Estimating the Delivery Probability

The estimation of $\mathcal{D}_{red}$ in Eq. 4.4 is based on the assumption that a node could deliver a packet to a *contact* with a probability of 1. However, the probability of successful one-hop

delivery in reality is susceptible to a variety of factors (e.g. Signal-to-Noise Ratio (SNR), contention, link duration or the size of MAC buffer). As packet losses due to low SNR is normally addressed in the link layer and contention rarely happens in the sparse topology, in the AMR scheme, only the delivery probability with limited link duration is considered, i.e. the probability that the transmission of a packet can be finished before the link breaks.

As illustrated in Fig. 4.3, node j is passing by its *contact* d with a relative speed $\hat{v}_{jd}$. Let c denote the closet distance between j's trajectory and node d and $\mathcal{T}_{lk}$ be the link duration. We have $\mathcal{T}_{lk} = \frac{2\sqrt{r^2-c^2}}{\hat{v}_{jd}}$. Let φ_i be the time required for transmitting the i_{th} packet to d and φ_{i-} be the amount of time that packet i has to wait until its transmission. We have $\varphi_{i-} = \sum_{k=0}^{i-1} \varphi_k + \varphi_{ctrl}$ $(i > 0)$ and $\varphi_{i-} = \varphi_{ctrl}$ $(i = 0)$, where φ_{ctrl} is the constant time for setting up a connection between the two nodes. Therefore, for a given φ_i and a relative speed estimate $\bar{v}$ the delivery probability of packet i (denoted as $\mathcal{P}_i$) is the probability that c is small enough to make $\mathcal{T}_{lk} > \varphi_i + \varphi_{i-}$, i.e. $\mathcal{P}_i(\varphi_i) = P(c \leq \sqrt{r^2 - (\frac{(\varphi_i+\varphi_{i-})\bar{v}}{2})^2})$. As c is uniformly distributed over [0, r) when the nodal movements reaches the steady state (assuming a RWP mobility model) [88], node j can estimate $\mathcal{P}_i$ for the i_{th} packet being forwarded to d by:

$$\begin{aligned} \mathcal{P}_i(\varphi_i) &= \int_0^{\sqrt{r^2-(\frac{(\varphi_i+\varphi_{i-})\bar{v}}{2})^2}} \frac{1}{r}\, dx \\ &= \begin{cases} \sqrt{1-(\frac{(\varphi_i+\varphi_{i-})\bar{v}}{2r})^2} & \varphi_i \leq \frac{2r}{\bar{v}} - \varphi_{i-} \\ 0 & \varphi_i > \frac{2r}{\bar{v}} - \varphi_{i-} \end{cases} \end{aligned} \tag{4.6}$$

Eq. 4.6 is based on the assumption that the mean relative speed $\bar{v}$ is known. For estimation of the distribution of link duration with unknown relative speed except of its distribution, the readers are referred to [89] or [90]. Knowing the time for transmitting a packet i to a contact and the current depth $\mathcal{H}$ of the delivery tree, a relay can now estimate the residual delivery time

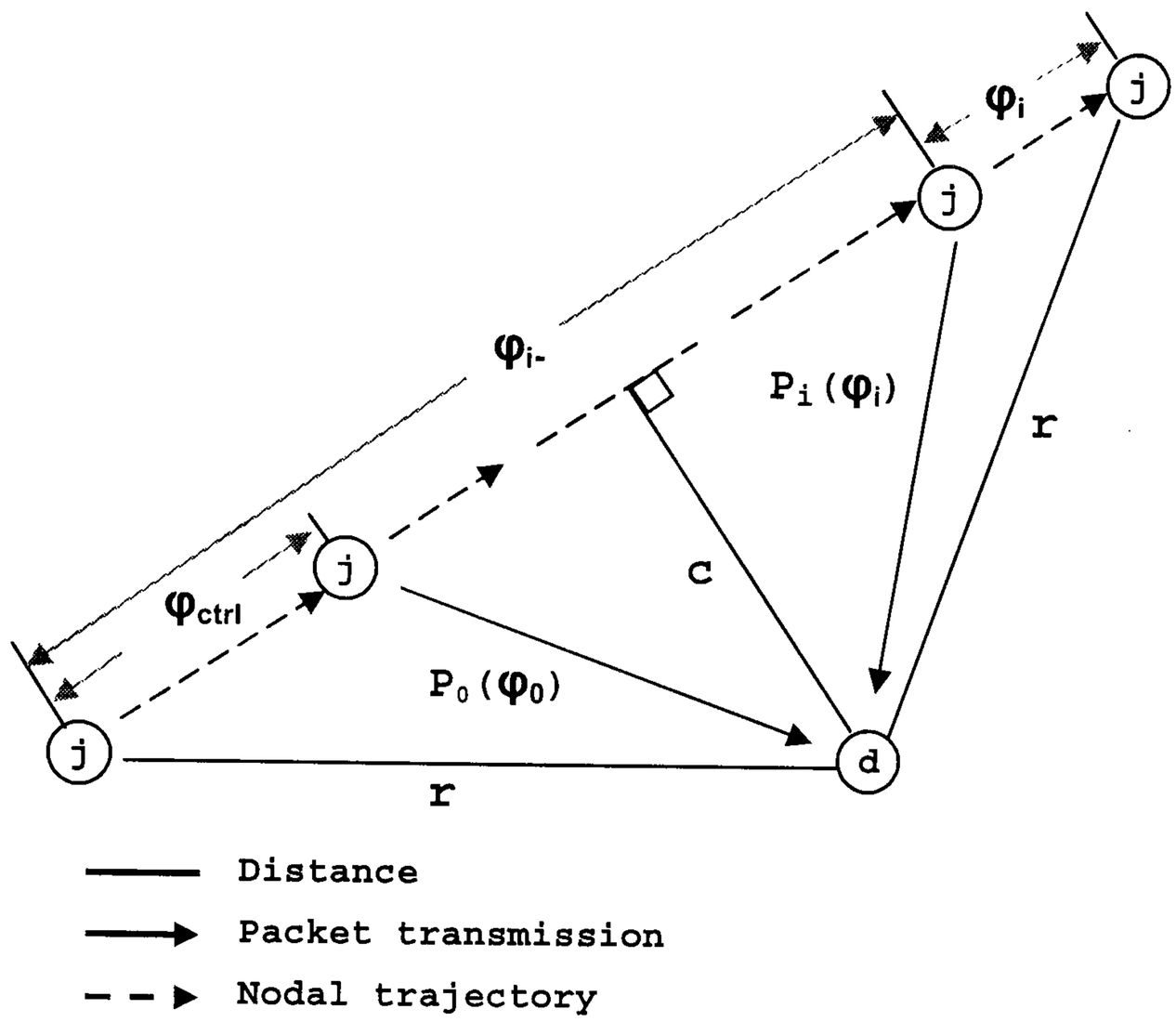


Figure 4.3: Delivery probability with limited link duration

as:

$$\mathcal{D}_{red}(\mathcal{H}, \varphi_i) = \frac{\bar{\mathcal{T}}_{int}}{2^{\mathcal{H}-1}\mathcal{P}_i(\varphi_i)} = \frac{1}{2^{\mathcal{H}-1}\mathcal{P}_i(\varphi_i)\bar{\lambda}} \tag{4.7}$$

and the CDF of the residual delay conditioned at $\mathcal{H}$ and $\mathcal{P}_i$ can be expressed mathematically as:

$$\mathcal{F}_{red}(x, \bar{\lambda}|\mathcal{H}, \mathcal{P}_i) = \begin{cases} 1 - e^{-2^{\mathcal{H}-1}\mathcal{P}_i\bar{\lambda}x} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4.8}$$

4.3.3 Distributed Adaptation of the Relaying Process for a QoS Target

To allow relay nodes to independently adapt the relaying process, three new fields are added into the packet header, namely, the current depth of the binary delivery tree $\mathcal{H}$, the target of end-to-end delivery delay $\mathcal{D}_{tg}$ and packet arrival rate $\mathcal{A}_{tg}$. $\mathcal{D}_{tg}$ and $\mathcal{A}_{tg}$ is specified at the source in accordance with the application layer requirements. $\mathcal{H}$ is increased for one at each relay

node before being replicated to the next relay. The pseudo code for the implementation of the adaptation mechanisms at each relay node is listed in Fig. 4.4. In the implementation, a relay can estimate the delivery probability of a packet to a *contact* using its knowledge of the packet size, the number of packets being forwarded to the *contact* and the bandwidth. From Fig. 4.4 we can see that, based on the information embedded in a packet's header, a single node can make its own decision on whether to forward this packet to a *contact* or to wait until it meet the destination. On average, relay nodes in the same layer of the binary relaying tree of a packet would receive the packet around the same time (e.g. in a $\mathcal{N}$ nodes network the mean time that a packet spent for travelling from the $\mathcal{H}_{th}$ to the $\mathcal{H}+1_{th}$ layer of the binary delivery tree is about $\frac{\bar{T}_{int}}{\mathcal{N}-2^{\mathcal{H}-1}}$). Assuming both the buffer size and the traffic load are constant, on average the communication traffic between the relay nodes and their *contacts* is constant hence they will draw the same conclusion on the delivery probability for the same packet. Therefore, the $\mathcal{H}_{th}$ layer relay nodes would make the same decision on whether to forward a packet. The forwarding process would continue until $\mathcal{H}$ is sufficient to guarantee the proportion of packets arrive at their destinations in time equal to or higher than the expected packet arrival rate $\mathcal{A}_{tg}$.

4.3.4 Extra Traffic Due to the Imprecise Adaptation

As the number of copies the AMR scheme produced for a packet tend to be powers of 2 with a specific delay target (e.g. if only 10 copies are needed for a delay target, the final number of copies produced by the AMR scheme would be $16 = 2^4 > 10$), we say such a binary tree based adaptation is imprecise. The extra copies $\Delta\mathcal{R}$ distributed for a packet would obviously bring more traffic than that are enough to achieve a certain delay target. However, the ratio of the extra copies over the network size $\frac{\Delta\mathcal{R}}{\mathcal{N}}$ is shown to become constant when the network size increases. Suppose that the number of nodes in the network grows to a very large size $\mathcal{N}$ ($2^{\mathcal{H}+1} > \mathcal{N} \geq 2^{\mathcal{H}}$), the biggest possible number of extra copies $\Delta\mathcal{R}_{max}$ produced by the imprecise adaptation is equal to $2^{\mathcal{H}} - 2^{\mathcal{H}-1} - 1$. We have $\frac{\Delta\mathcal{R}}{\mathcal{N}} \leq \frac{\Delta\mathcal{R}_{max}}{\mathcal{N}}$. Therefore, the extra

```
/* input: Packet_Header pkt, Contact_ID cID */
Function: AMR_Relay ( pkt, cID) ;

  if ( cID == pkt->dest ) {
          Relay_a_Packet ( pkt, cID ) ;
           return ;
  }

  if ( pkt->H  >= DEPTH_LIMIT ) return ;

  Lambda = Expected_Contact_Arrvial_Rate ();

  D_cur = CURRENT_TIME – pkt ->time_stamp;

  P_delivery = P( pkt->size, Mean_Relative_Speed ) ;

  CDF_red =  F_red (pkt->D_tg - D_cur, Lambda, pkt->H - 1, P_delivery )

  if (CDF_red >= pkt->A_tg) {
          pkt->H = DEPTH_LIMIT ;
          return ;
  }

  else {
          pkt->H ++ ;
          copy_packet ( pkt, copy_of_pkt) ;
          Relay_a_Packet ( copy_of_pkt, cID ) ;
          return ;
  }
```

Figure 4.4: Pseudo-code for the distributed adaptation of the relaying process of a packet

copies per node produced for a packet due to the imprecise adaptation is limited by:

$$\frac{\Delta\mathcal{R}_{max}}{\mathcal{N}} = \frac{2^{\mathcal{H}} - 2^{\mathcal{H}-1} - 1}{2^{\mathcal{H}} + \varepsilon} \tag{4.9}$$

where $\varepsilon = \mathcal{N} - 2^{\mathcal{H}}$ and $\varepsilon \ll \mathcal{N}$. From Eq. 4.9, a conclusion can be drew that that $\frac{\Delta\mathcal{R}}{\mathcal{N}}$ is bounded by 0.5 as the size of the network is growing to infinity.

4.4 Simulation Results and Evaluation

This section presents the details of the simulation model and four sets of simulation based experiments that are designed to evaluate the performance of the AMR scheme (in terms of traffic overhead, energy-efficiency and adaptability) under various QoS targets.

4.4.1 Simulation Model

The performance of the AMR scheme is evaluated using the *ns-2* network simulator. The simulated MANET consists of $\mathcal{N} = 40$ nodes uniformly distributed in a square area with the size of $\mathcal{L} \times \mathcal{W}$. Each node in the network utilizes IEEE 802.11b as the MAC protocol and is capable of transmitting data at the rate of 2 MBit/s. In our simulations, the minimum value for $\mathcal{L}$ or $\mathcal{W}$ is 600m. The transmit power (denoted as P_t) of the wireless interface in a node is fixed at 0.075W for a nodal radio radius of 50m. These settings create sparse network topologies with a maximum node density (given by $\frac{\pi r^2 \mathcal{N}}{\mathcal{L}\mathcal{W}}$) of 0.87 $node/m^2$. All the mobile nodes in the network move around according to the Random Waypoint Model (RWP). The pause time of the RWP model is kept at 0 to produce continuous movements. The movement speed of a mobile node is uniformly distributed over $[v_{min}, v_{max}]$, where v_{min} is fixed at 2.5m/s (or 9km/hour) and v_{max} is varied from 5m/s (or 18km/hour) to 40m/s (or 144km/hour) to create simulation scenarios with different mobility. If not stated, the default value of v_{max} is 20m/s (or 72km/hour). Every node in the network is acting as the packet source, relay and receiver at the same time. However, no data traffic is produced in the 1100 seconds after the simulation starts to allow the mobility model to reach its steady state. After that, each node generates a packet destined to each of the rest of the network every 0.5 second for a duration of 500s. The simulation then continues for another 900 seconds to allow most of the packets to reach their destination. Each simulation is repeated 30 times with different random seeds to provide smoother results.

4.4.2 Performance Evaluation

Experiment 1

The first experiment is designed to test the performance of the AMR scheme regarding its adaptability to varying network mobility with an ideal delivery probability of 1. The performance of the Epidemic routing is also evaluated severing as a benchmark protocol in the scenario of tight delay budget. In this experiment, the side length and width of the network

ground is configured at the minimum value 600m. The QoS target of this experiment is to achieve 125s of mean delivery delay (or about 63.3% of packets should arrive at their destinations in less than 125s)[§]. The results of this experiment is demonstrated in Fig. 4.5. We can see from Fig. 4.5a that the relay nodes in the AMR scheme keeps adapting to the increasing network mobility for the specified 125s delay target. Although it does not achieve the delay target when the network mobility is low (e.g. the maximum speed limit v_{max} is below 10m/s), the AMR scheme has achieved a delivery delay as quick as that achieved by the fastest possible scheme, the epidemic routing[¶]. When the network mobility climbs up, we can see that the AMR scheme quickly adjusts the depth of the delivery process at each relay node resulting in delivery delays that are just below the delay target. It seems that both epidemic routing and AMR can achieve the delivery target and the former is preferable to the latter as it is simpler. However, as shown in Fig. 4.5b, by adapting the depth of delivery process to the varying network mobility, in general the AMR scheme minimizes the traffic overhead to achieve the given delay target.

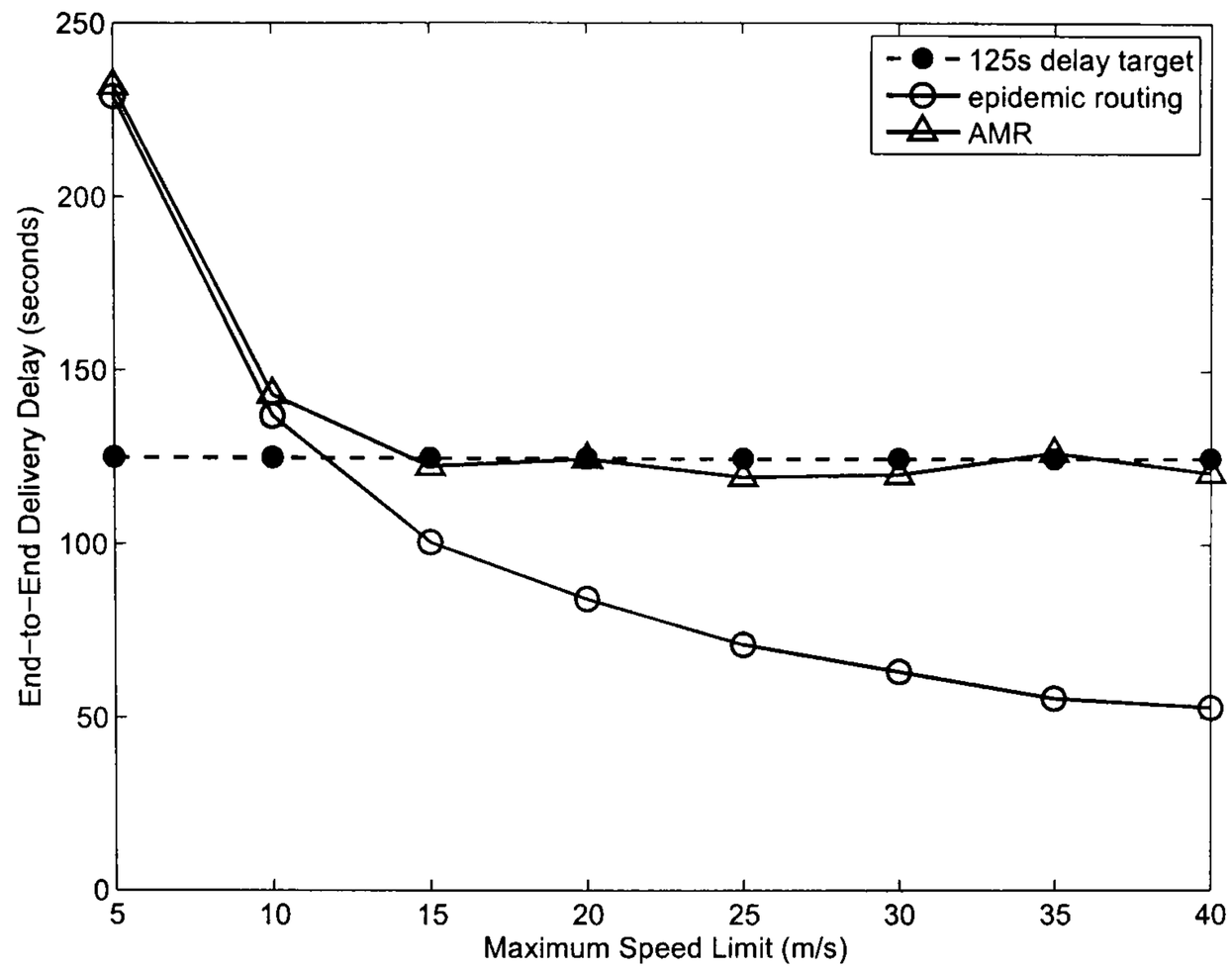
Experiment 2

The second experiment is designed to demonstrate the adaptability of the AMR scheme to a variety of delay targets and delivery probabilities. The size of the network ground is configured as $600m \times 600m$ and the speed limit v_{max} is fixed at 20m/s. The results obtained from scenarios of varying QoS targets (e.g. expected mean end-toe-end delay) with a constant delivery probability of 1 is plotted in Fig. 4.6. The curve of the AMR scheme in Fig. 4.6 sticks to that of the delay targets spanning from 90 seconds to 250 seconds, which confirms that the excellent adaptability of the AMR scheme to the wide range of QoS targets.

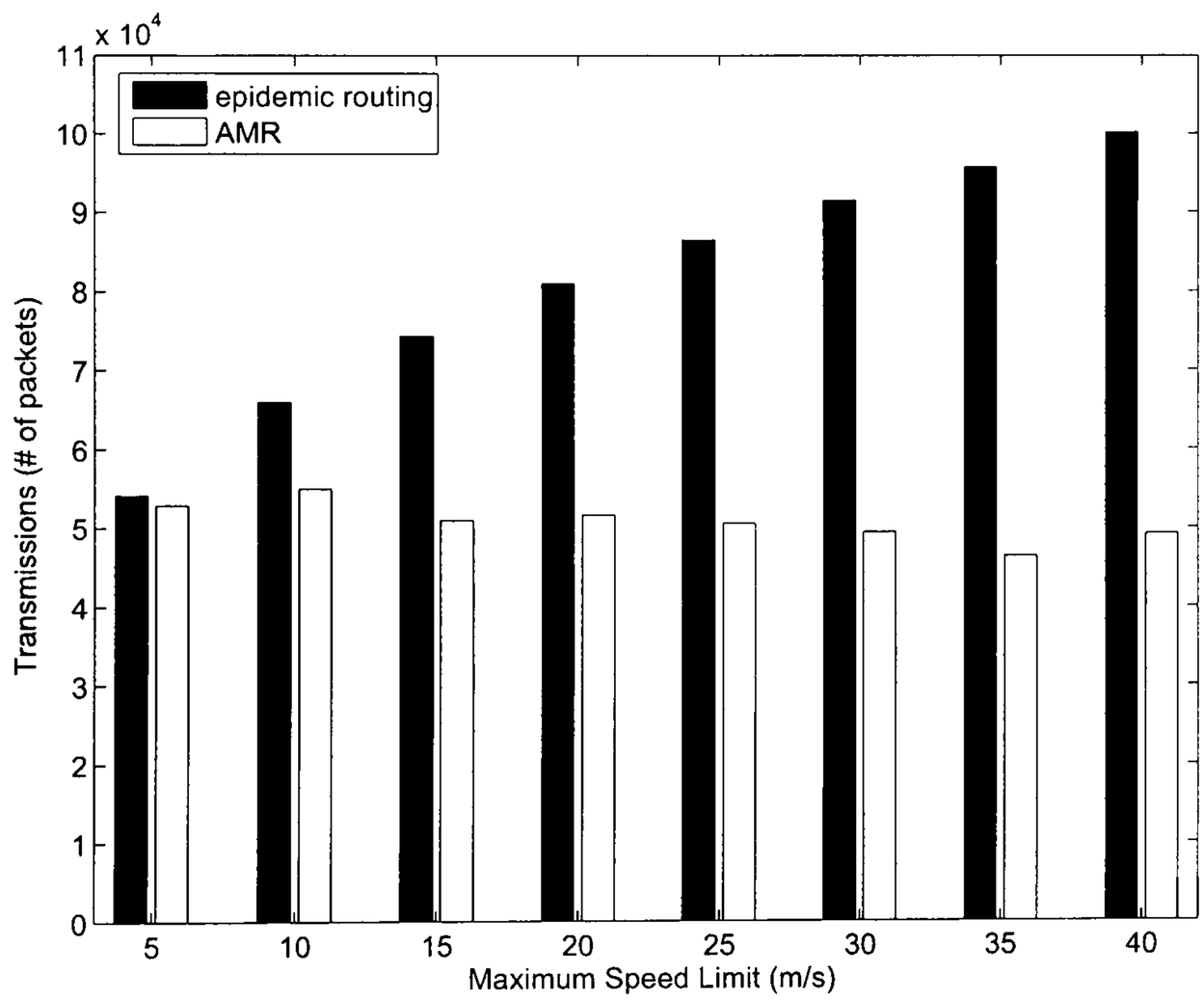
Fig. 4.7 demonstrates the performance of the AMR scheme in varying delivery probabilities given a fixed delay target of 250 seconds. A multi-copy scheme with a fixed *replication*

[§]For exponential distributions, the CDF of the mean is about 0.6326

[¶]By limiting the traffic generation rate, we can achieve a so low contention probability for epidemic routing that makes it become the fastest solution.



(a)



(b)

Figure 4.5: Performance comparison with delay target 125 seconds (a) Delivery delay (b) Total packet transmissions

factor of two is chosen as the benckmark protocol for the relaxed delay target. The 2-copy multi-copy scheme is shown in Fig. 4.7 where the delivery probability is 1 to have a similar performance with the AMR scheme for the given delay target and network mobility. However,

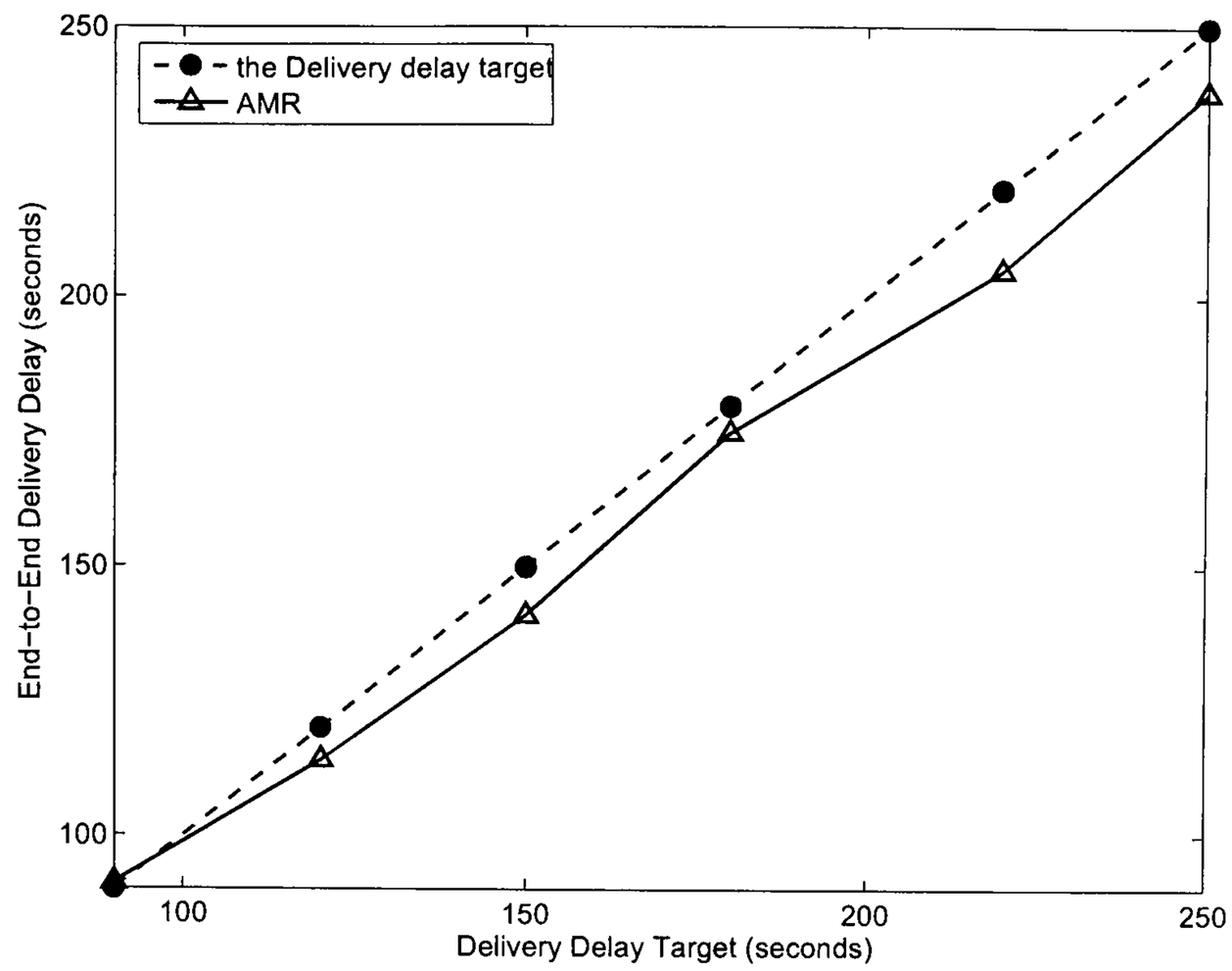


Figure 4.6: the Delay targets Vs. the delivery delays achieved by the AMR scheme

when the delivery probability decreases the fixed 2-copy multi-copy scheme misses the delay target due to its lack of adaptivity. On the contrary, Fig. 4.7 shows that the varying delivery probabilities do not stop the AMR scheme from reaching the mean delay target. There are bigger gaps between the delivery delay of the AMR scheme and the delay target when the delivery probability is around 0.7. These gaps are resulted from the imprecise adaptation as explained in Section4.3.

Experiment 3

The third experiment is designed to evaluate the performance of the AMR scheme (in terms of mean delay and total energy consumptions of packet transmissions) against two multi-copy strategies of different $\mathcal{R}$ factor (e.g. 2 and 20). The side length and width of the network ground is set to be 1200m and 600m, respectively. The resulted node density is 0.43 $node/m^2$. The QoS target is to achieve 250s of mean delivery delay. For this experiment, the size of a packet $\mathcal{S}$ is fixed at 1050 bytes (including MAC and IP headers). Without taking the energy dissipation in retransmissions into account, transmitting a packet over one-hop consumes $P_t \frac{\mathcal{S}}{BW}$ *joules* of

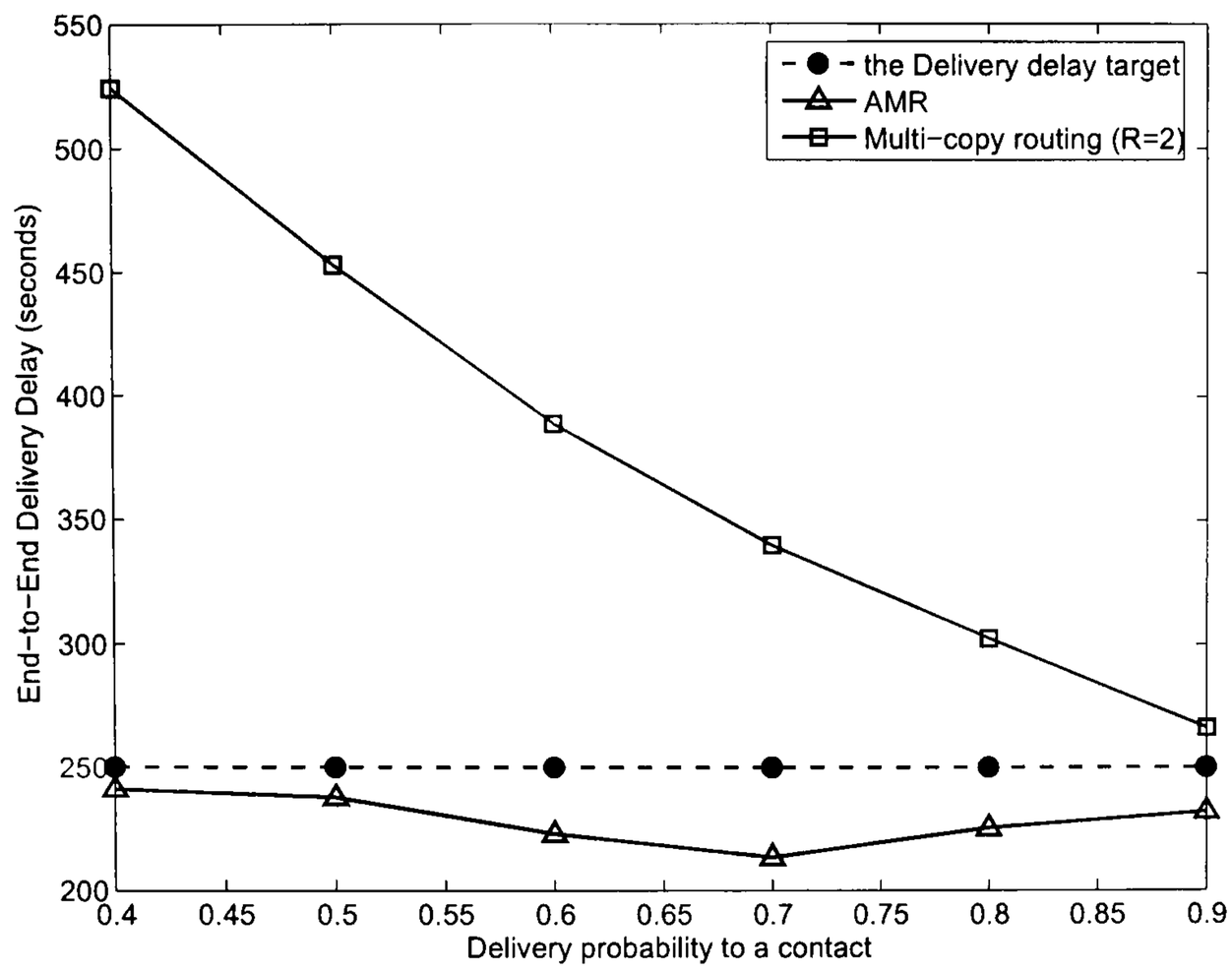
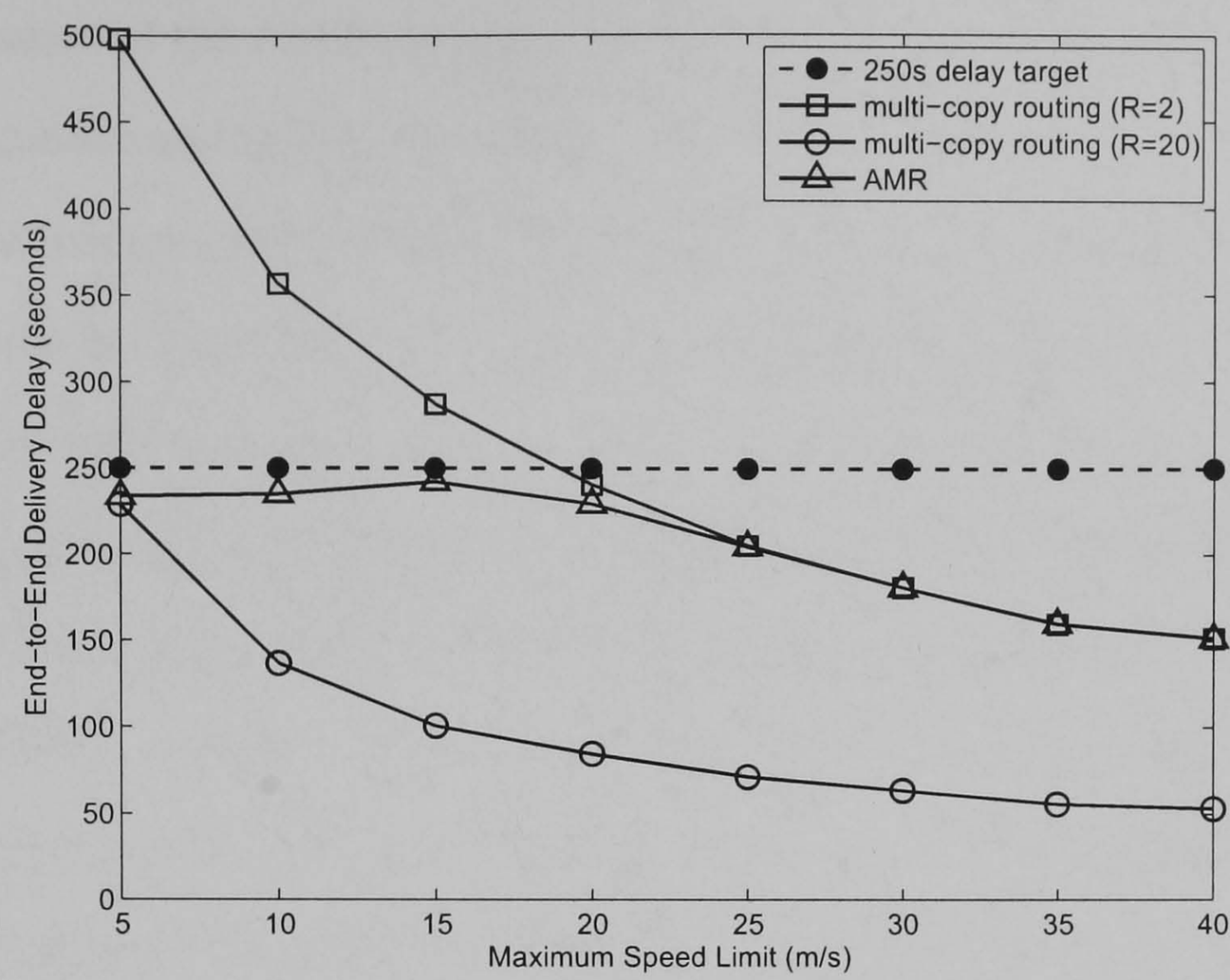


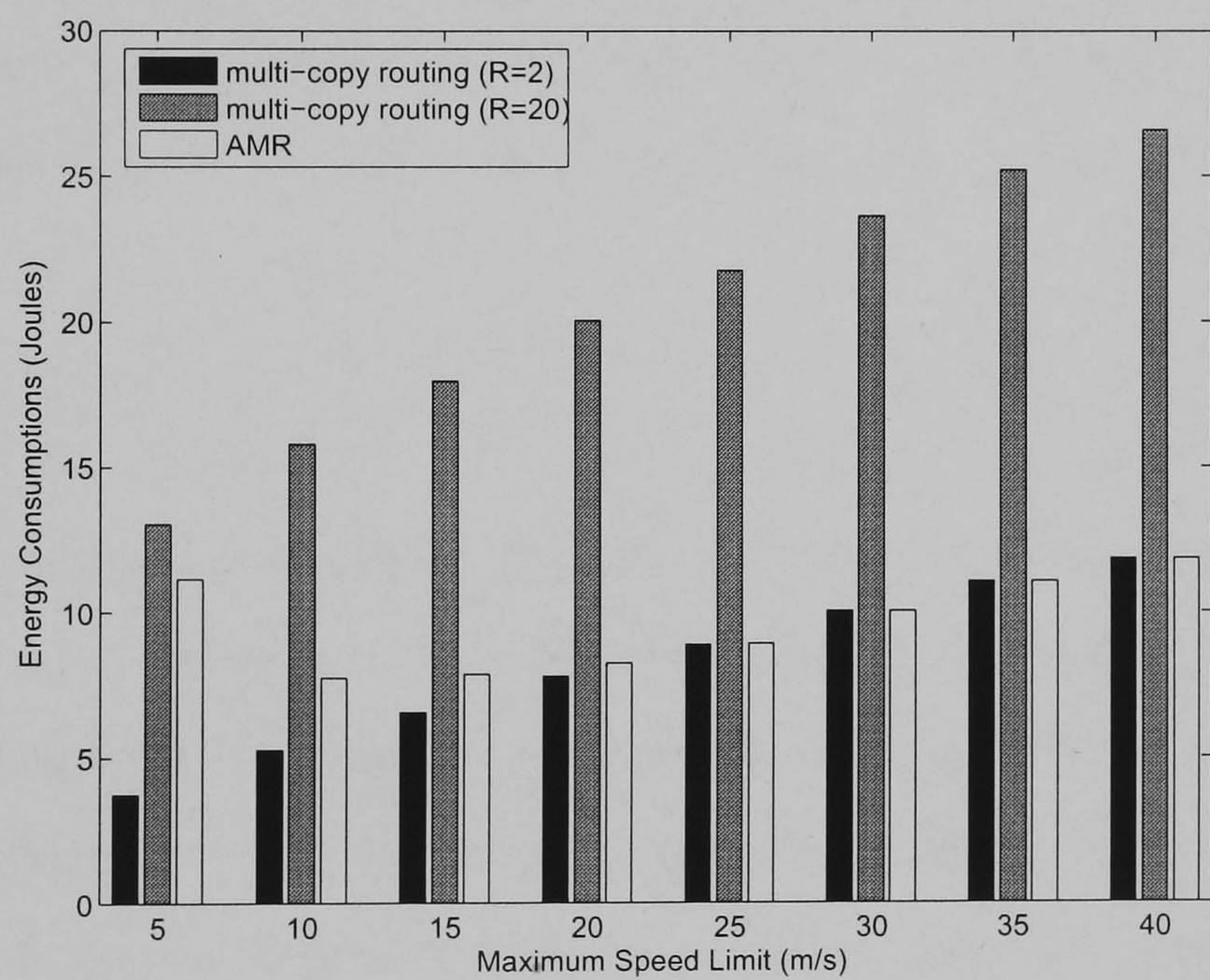
Figure 4.7: the Delivery probability Vs. the end-to-end delivery delay

energy, where P_t denotes the transmit power of wireless radios and BW for the data rate.

The performance of the AMR scheme in comparison to the 2-copy and the 20-copy multi-copy schemes with varying speed limits and constant 100% delivery probability is shown in Fig. 4.8. The end-to-end latency of packet delivery of the three schemes is given in Fig. 4.8a. The energy consumed by these schemes for packet transmissions is given in Fig. 4.8b. We can see from Fig. 4.8a that, with the AMR scheme, relay nodes keeps adapting packet delivery to the varying network mobility. When the network mobility is low (e.g. $v_{max} \leq 20m/s$) the AMR scheme is able to set a right depth $\mathcal{H}$ for the delivery process resulting in end-to-end delays just below the specified 250s budget. When the network mobility climbs up, the AMR scheme further constrained the delivery process making $\mathcal{H}$ to be 2. We can observe that in these cases the performance of the AMR scheme is similar to that of the multi-copy scheme with $\mathcal{R} = 2$. Although from Fig. 4.8a we can see that the multi-copy strategy with $\mathcal{R} = 20$ achieved the delivery target in all the scenarios. The energy consumptions of the scheme shown in Fig. 4.8b are up to 3 times of that of the AMR scheme in scenarios of high mobility. Fig. 4.8b also demonstrates that the 2-copy multi-copy scheme used less energy than the AMR scheme



(a)



(b)

Figure 4.8: Performance comparison with delay target 250 seconds (a) Delivery delay (b) Total energy consumptions for packet transmissions

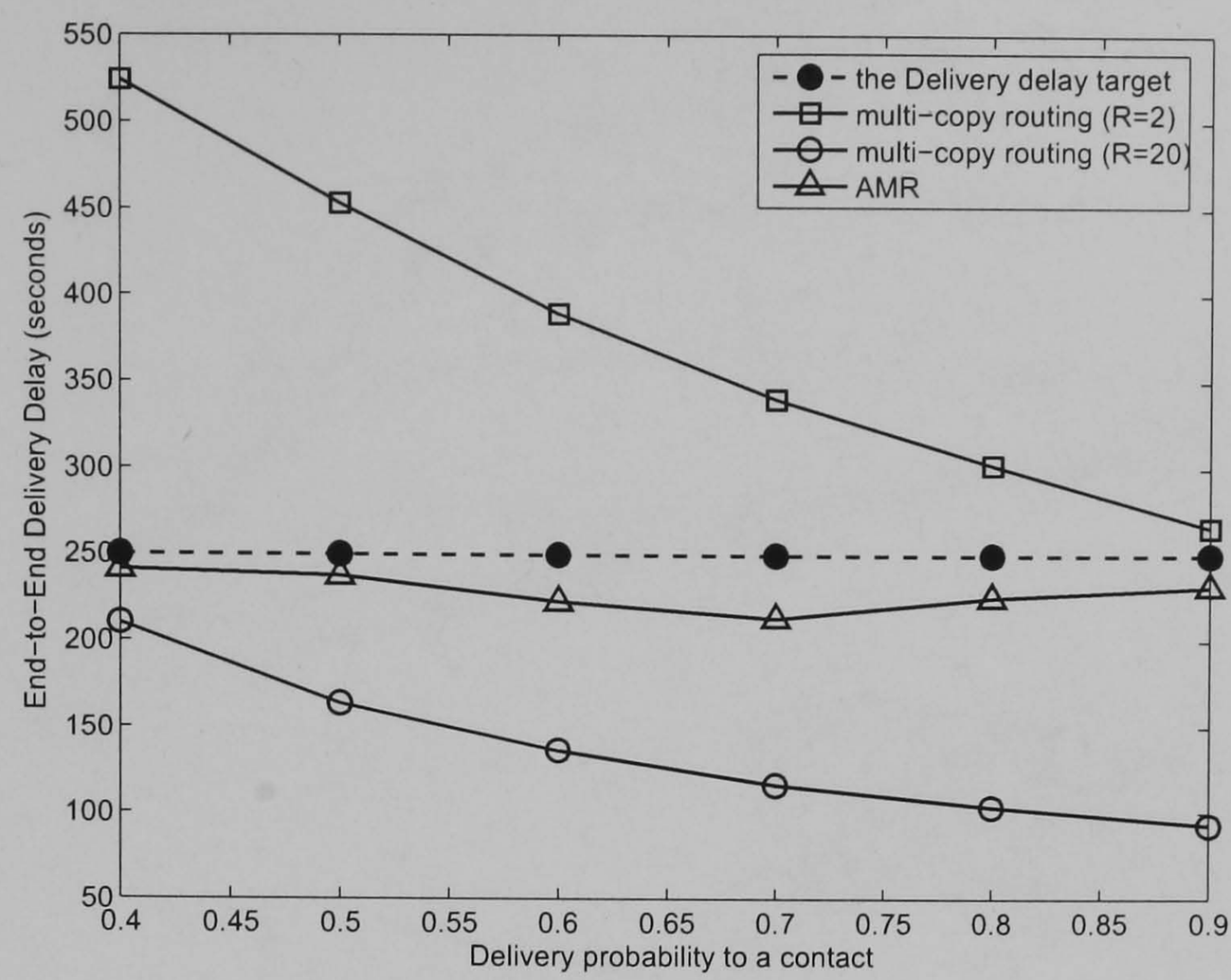
in scenarios of $v_{max} \leq 20m/s$. But the extra energy is actually utilized by the AMR scheme in the low mobility situations to achieve close-to-target delivery delays for an overall optimal balance between the constraints of delay and energy consumptions.

The performance of the AMR scheme in comparison to the 2-copy and the 20-copy multi-copy schemes against varying delivery probabilities is shown in Fig. 4.9a for end-to-end latency and Fig. 4.9b for energy consumptions. The speed limit v_{max} is fixed at 20m/s. The 2-copy multi-copy scheme has been shown in Fig. 4.9a to exhibit similar performance with the AMR scheme when the delivery probability is 1 and v_{max} is 20m/s. However Fig. 4.9a and Fig. 4.9b shows that, due to lack of adaptability, the scheme missed the 250s delay target for all the less-than-one delivery probabilities despite it consumes the least amount of energy for packet transmissions. On the contrary, we can observe from Fig. 4.9 that the adaptability of the AMR scheme enables it to spend some extra energies according to the present delivery probability so that the end-to-end delay target can be achieved. The delivery delay of the multi-copy scheme can be enhanced by increasing $\mathcal{R}$, i.e. with $\mathcal{R} = 20$ in Fig. 4.9a the scheme can achieve the delay target regardless of the delivery probabilities. However, the downside of increasing $\mathcal{R}$ to more than needed is, again, the waste of energy in unnecessary transmissions, as demonstrated in Fig. 4.9b.

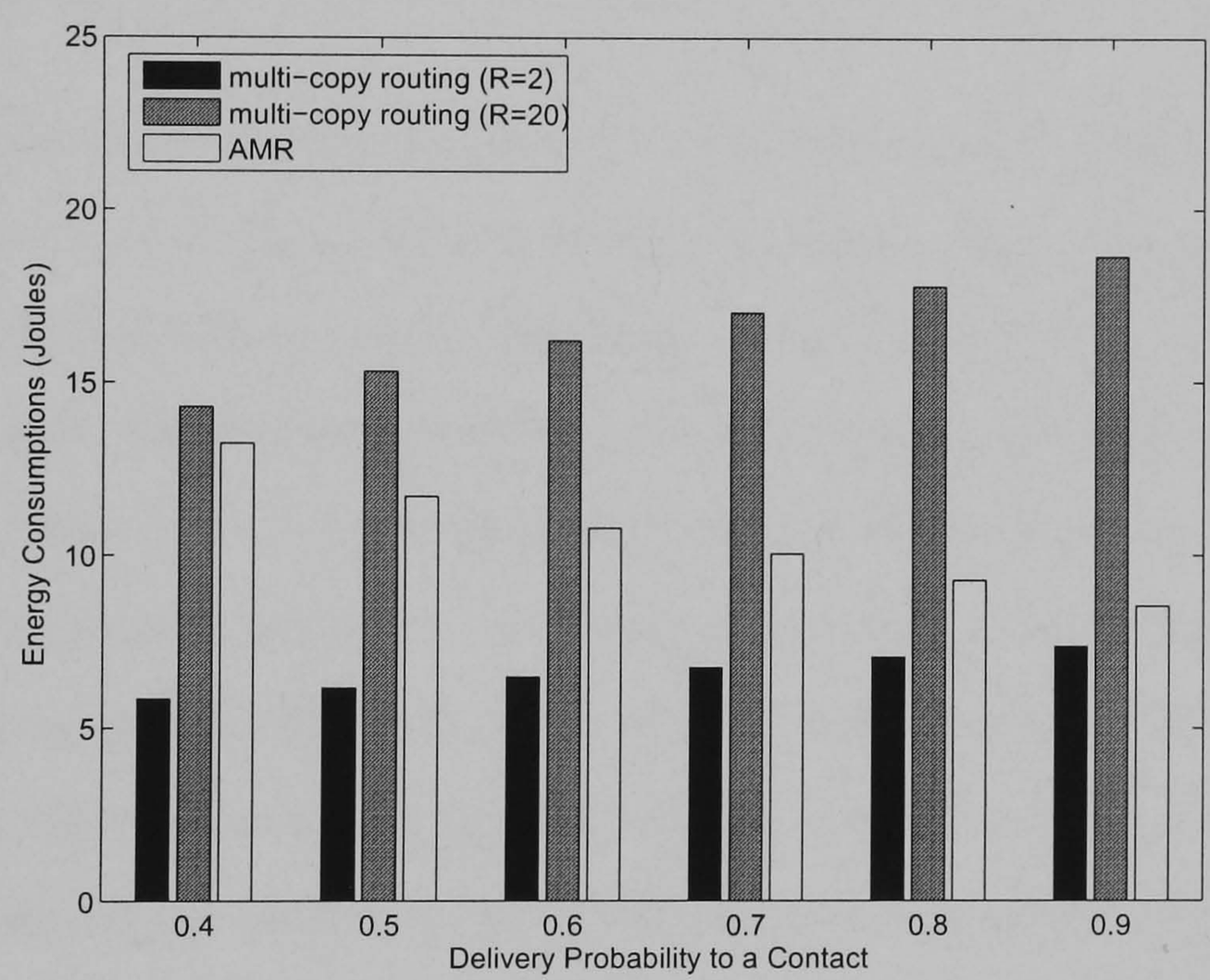
Experiment 4

The fourth experiment is devised to verify that the AMR scheme can achieve more specific QoS objectives, i.e. a certain percentage of packets (referred to as packet arrival rate) should arrive at their destination within a delivery delay target. For this experiment, the size of the network ground is set to be $1800m \times 1800m$, resulting in a node density of 0.09 $node/m^2$.

The AMR scheme is first given QoS objectives of 500s delay with arrival rate of 0.4, 0.6 and 0.8 packet/s for packet delivery in scenarios of various mobility levels. The resulting CDFs of packet delivery delays achieved by the AMR scheme as well as those that are achieved by the epidemic and the 2-copy multi-copy schemes are plotted in Fig. 4.10. As shown in Fig. 4.10, the AMR scheme has achieved most of the QoS objectives. In scenarios of low mobility level (e.g. $v_{max} \leq 10m/s$, relay nodes in the AMR scheme increases the number of packet replicates for the harder objectives (e.g. *confidence* of 0.6 and 0.8) and flood packets in a



(a)



(b)

Figure 4.9: Performance comparison against one-hop delivery probabilities (a) Delivery delay (b) Total energy consumptions for packet transmissions

way that is similar to those in the epidemic scheme, in order to achieve faster delivery. When the network mobility is increased, intermediate nodes resort to a smaller *replication factor* for energy-efficiency. In high mobility scenarios (e.g. $v_{max} \geq 25m/s$), the AMR scheme reduces

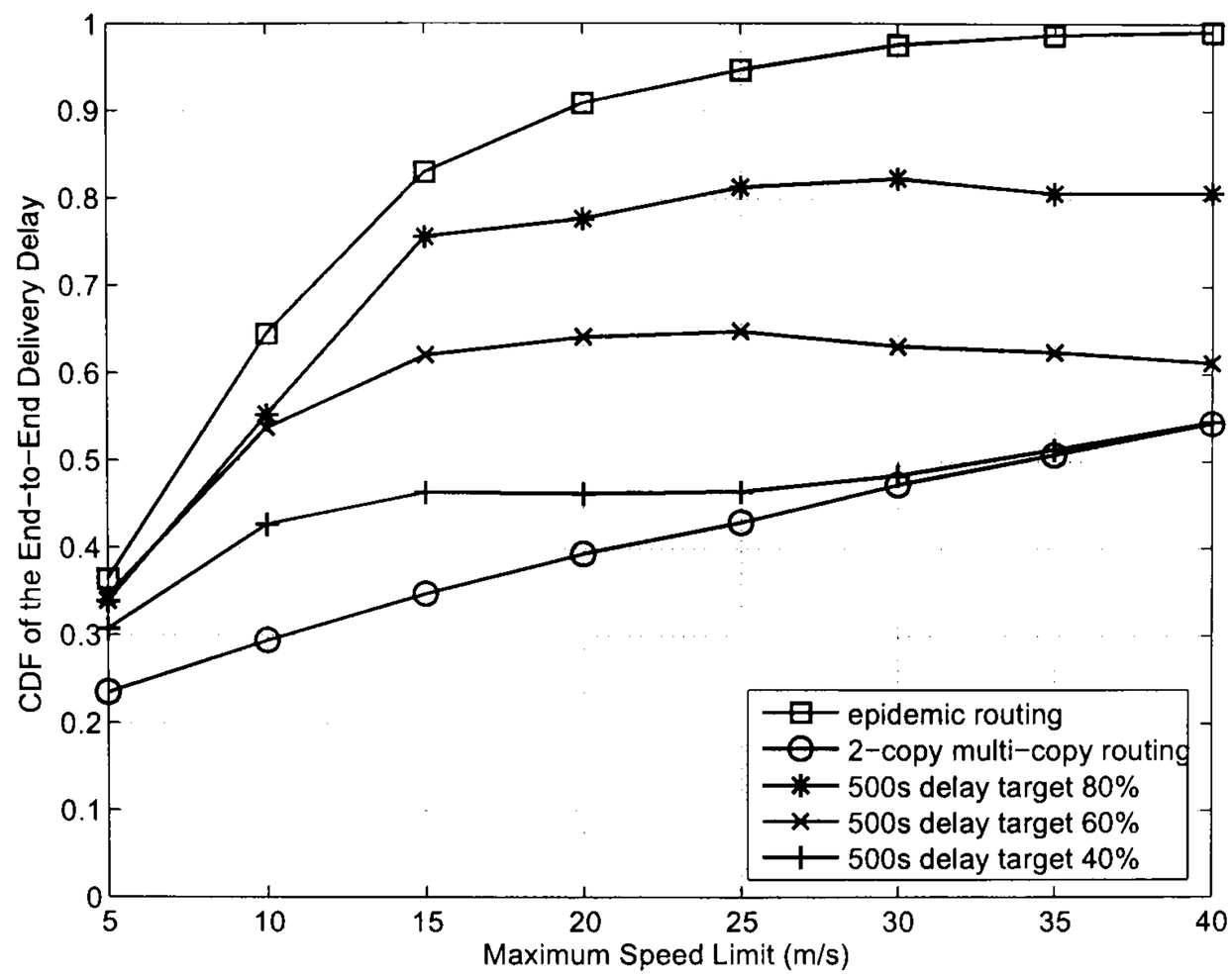


Figure 4.10: the Delay targets Vs. the delivery delays achieved by the AMR scheme

average number of packet replicates to only 2 for the relatively easy objective of 500s delay with 0.4 *confidence*, resulting in a similar performance to the 2-copy multi-copy scheme.

Another set of QoS objectives (e.g. 900s delay with packet arrival rate of 0.4, 0.6 and 0.8) are also given to the AMR scheme to deliver packet in scenarios of varying one-hop delivery probability (e.g. from 0.3 to 1) but fixed mobility ($v_{max} = 20m/s$). The results are compared with those of the epidemic and 2-copy multi-copy scheme and plotted in Fig. 4.11. Similar to the case of varying mobility, Fig. 4.11 shows that the AMR scheme produces as many packet replicates as the epidemic scheme to deal with the low delivery probabilities (e.g. 0.5 or less) and to achieve the tougher objectives (e.g. packet arrival rate of 0.8 for 900s delay). In good networking conditions (e.g. delivery probability 0.6 or higher), the AMR scheme automatically adapts its behavior for less traffic overhead. For easy objectives (e.g. 0.4 *confidence* for 900s delay), the AMR scheme constrains packet relaying in a way that is similar to the 2-copy multi-copy scheme.

As each node is assumed to be equipped with a big enough relay buffer and allocated enough time for the multi-hop packet relay, all the protocols investigated achieved a delivery ratio of

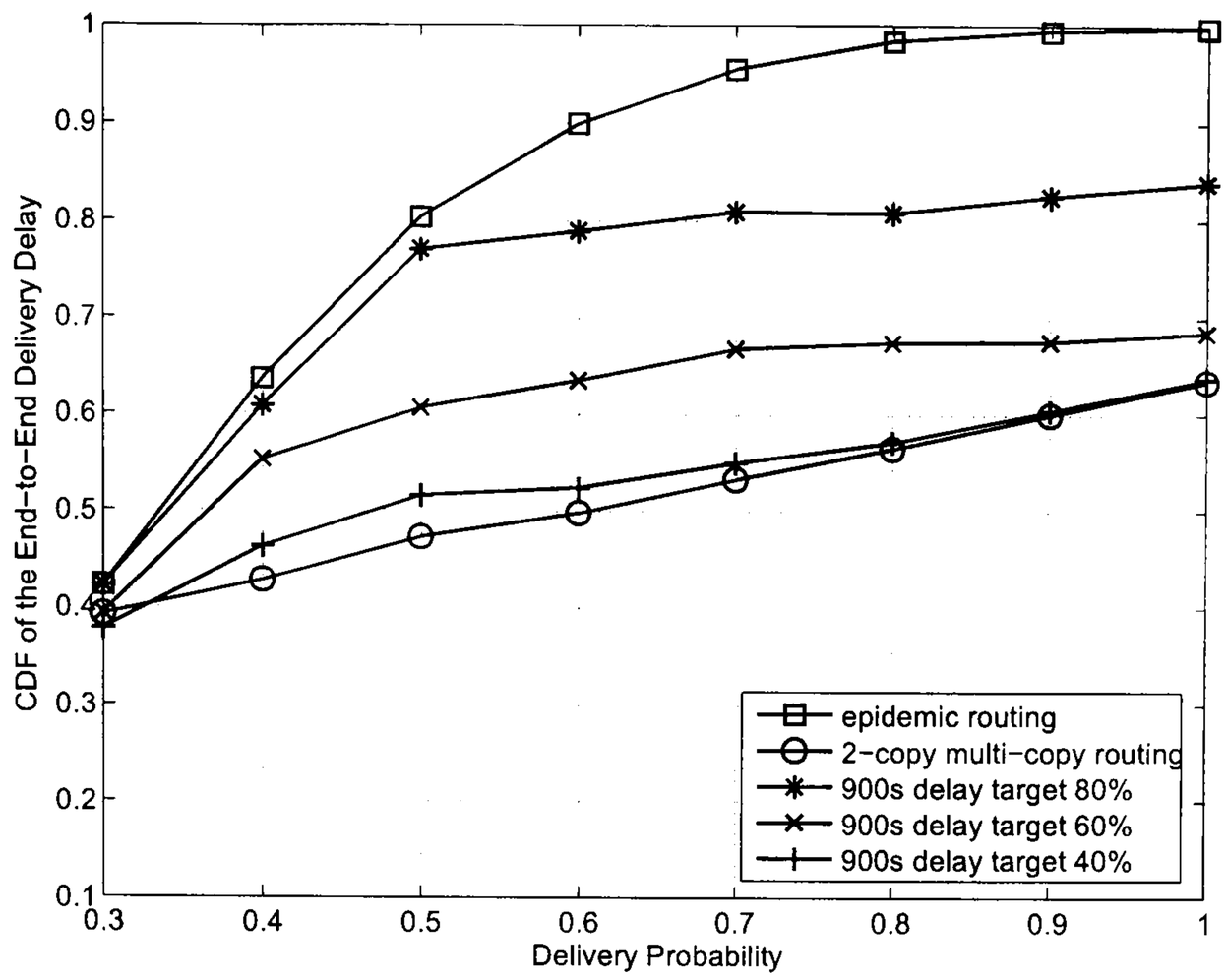


Figure 4.11: the Delivery probability Vs. the end-to-end delivery delay

over 98% in all of the four experiments.

4.5 Conclusion

Data delivery in sparse mobile ad-hoc networks can not be accomplished with traditional routing technologies that are designed for fully connected paths due to the intermittent availability of links. The problem is addressed by the recently proposed *Store-and-forward* based routing (e.g. multi-copy relaying). However, existing *Store-and-forward* based routing protocols, such as the multi-copy relaying strategies, lack the adaptability to frequent variations in network conditions as they rely on the source to control the relaying process. This Chapter introduces a *store-and-forward* based scheme, the Adaptive Multi-Copy Routing (AMR), which addresses this issue and provides an energy-efficient way for packet delivery in sparse MANETs. In the AMR scheme, instead of using a source-defined *replication factor*, relay nodes independently decide whether to forward a packet based on the given budget of delivery delay and current network conditions. Simulation results confirm that the AMR scheme is able

to exploit the varying network mobility and/or delivery probability that cannot be anticipated in the source-defined strategies and to deliver packets within specific QoS targets with minimum traffic overhead and energy consumptions.

Chapter 5

Range-based Mobility Metrics

5.1 Introduction

Communications in mobile ad-hoc networks rely on distributed routing protocols which can cope with frequent changes in network topologies caused by nodal mobility. The performance of routing protocols in MANETs depends on their ability to adapt to the dynamics of the network and to work out the best strategy for packet delivery. Thus, it is important to have quantitative mobility metrics that can reflect the performance of communications for various levels of network mobility and topological stability.

Normally, the measurement of mobility metrics involves complex positioning systems to determine the time-varying locations of mobiles. Tracking the locations of mobile nodes is costly in terms of energy consumption or may not work properly in some scenarios (e.g. in doors when a GPS receiver [14] is used).

In this Chapter, a new mobility metric, the *intra-vicinity dependency* is presented, for performance measurements. The main novelty is that it can fully capture the relative motions between a node and its vicinity (e.g. its one-hop neighbours) in a 2D plane, in real-time, using distance estimates and simple triangulation. It can be used as a substitute for the average relative speed when information about the true speed of mobile nodes is not available. Variants of this metric are also proposed for predicting the performance of networks that follow

group and random mobility models (e.g. inter-group inter-meeting times* and packet delivery rate†). To make the proposed mobility metrics more robust in noisy environments, a calibration method is also proposed for improving accuracy. The performance of the metrics is validated by four sets of experiments covering various mobility models (e.g. the Random WayPoint model (RWP), Restricted Random Waypoint (R2WP), the City Section Mobility model (CSM), and the Reference Point Group Mobility (RPGM) model) [11] [91]. The accuracy of the proposed localization-free metrics in approximating the average relative speed between mobile nodes/-groups are compared in two experiment sets with several existing methods [62] [63]. Another two sets of experiments are also carried out to test the performance of the proposed metrics when they are used to predict the inter-group inter-meeting times for networks moving in the RPGM model [11] and when they are deployed to estimate the packet delivery rate for networks following the RWP model [11].

This Chapter is organized as follows. Related work on mobility metrics is introduced in Section 5.2. Section 5.3 present the definitions and the associated algorithms of the proposed mobility metric and its variants. Section 5.4 gives technical details of the four simulation-based experiments. Summaries of this Chapter can be found in Section 5.5.

5.2 Related Work

Common measures of nodal mobility and topological dynamics [57] [92] are derived from the relative velocities between pairs of mobile nodes. To obtain the relative velocity between a pair of nodes, knowledge of their locations or motion vectors (speed, direction, etc.) in real-time is necessary. Such knowledge can be inferred from changing positions of mobile nodes with the help of a localization system e.g. GALILEO [93] GPS [14] or GPS-free positioning systems [75] [94] [95]. However, the GPS signal is normally too weak to be of any use in-

*Defined as the duration between the time when a group of nodes encounters another group and the next time when they meet again

†Defined as the total number of packets received divided by the total number of packets transmitted

doors. The GPS receiver may also be impractical for some mobile devices due to size or power constraints. In addition, GPS-free positioning systems [75] [94] [95] are too complex and computationally expensive just for measuring mobility. Some mobility metrics that do not require localization systems have been recently proposed. For example, in [64] the link duration is used as a mobility measure for adaptive protocols. However, a problem of this approach is that measurement of link duration cannot always be carried out in real-time as the links could exist for several minutes in low speed scenarios.

The mobility metric proposed in [63] does not need the use of localization systems either. It is based on the time derivative of the distance (or *distance change rate*) between pairs of neighbouring nodes. Kwak et al. [62] extended this work to create a mobility measure that puts more weight on nodal movements that are closer to the communication range. These mobility metrics have been shown to have strong correlation with the *link change rate* of a network. However, the approach of measuring the inter-node mobility (the *distance change rate*) is not accurate enough as it can only reflect the relative motions in the direction of line of sight. Considering the fact that the relative movements between nodes in the 2D space involve their relative motions in the horizontal and vertical scales, the *distance change rate* actually underestimates the network mobility. This problem limits the applications of these metrics to only mapping functions (e.g. from the metrics to the link change rates) that do not require the exact value of the relative motion.

5.3 The Range-based GPS-free Mobility Metrics

This section presents the proposed mobility metric, the *intra-vicinity dependency*, and its variants (i.e. the *inter-group mobility* and the *intra-vicinity mobility*), as well as the associated measurement methods and algorithms.

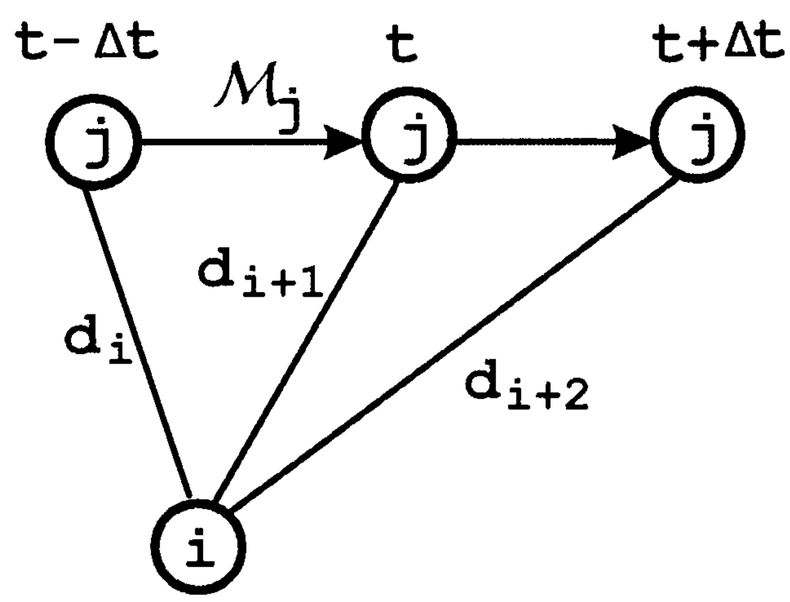


Figure 5.1: The relative movements and distances between two mobile nodes i and j

5.3.1 The Intra-Vicinity Dependency

As mobile nodes move around, the relative motions between a node and its one-hop neighbours cause their relative positions, and hence the network topology, to change in time. Thus, the impact of network mobility on the communication performance can be inferred from the average rate of changes in the inter-node relative positions (referred to as *change rate* in the rest of the Chapter). However, measuring the *change rates* of relative positions between nodes cannot be accomplished by simply taking the time derivative of the inter-node distances [62] [63]. As shown in Fig. 5.1, minimal deviations in the distance samples (e.g. d_i, d_{i+1} and d_{i+2}) measured at every Δt do not match the obvious relative movements between node i and j, i.e. $|d_{i+1} - d_i| \ll \mathcal{M}_j$.

To fully capture the relative motions between a node and its one-hop neighbours, a mobility metric, *intra-vicinity dependency* is proposed. The measurement of this metric is based on knowledge of a mobile node's position relative to the other pairs of nodes in its neighbourhood. It is assumed that each node periodically measures the distances to its one-hop neighbours by means of received signal strength (RSSI) [96], time-of-arrival (ToA) [96] or time-difference-of-arrival (TDoA) [97] measurements. Every node also reports its distance estimates to all of its one-hop neighbours at a frequency of $1/\Delta t$. Thus, every node knows not only the distances to its one-hop neighbours but also those between neighbouring nodes within the vicinity. Let $\mathcal{R}_i$ be the set of one-hop neighbours of node i and $\mathcal{U}_i$ the set of pairs of i's one-hop neighbours that

are also one-hop away from each other. We have $\mathcal{U}_i = \{(j,k)|j \in \mathcal{R}_i \wedge k \in \mathcal{R}_i \wedge j \in \mathcal{R}_k \wedge k \in \mathcal{R}_j\}$, where ($j$, k) can be any combinations of node couples within node i's vicinity that satisfy this condition. For each pair of nodes $(j,k) \in \mathcal{U}_i$, as shown in Fig. 5.2, the distance l from node i to the midpoint of the line between nodes j and k as well as the included angle θ are used to interpret the relative position between node i and node pair (j, k). Given the estimates of distances between these nodes (e.g. d_{ij}, d_{jk} and d_{ik}), the values of l and θ can be obtained by simple triangulation as follows:

$$l = \sqrt{d_{ij}^2 + (\frac{d_{jk}}{2})^2 - d_{ij}d_{jk}cos\varphi_{ik}} \tag{5.1}$$

$$\theta = cos^{-1}\frac{l^2 + (\frac{d_{jk}}{2})^2 - d_{ik}^2}{ld_{jk}} \tag{5.2}$$

where φ_{ik} is the included angle of d_{ij} and d_{jk} given by:

$$\varphi_{ik} = cos^{-1}\frac{d_{ij}^2 + d_{jk}^2 - d_{ik}^2}{2d_{ij}d_{jk}} \tag{5.3}$$

As node i moves, its position relative to node pair (j, k) will change in time. For example, as shown in Fig. 5.2, its position moves from (l, θ) at time $t - \Delta t$ to ($l^{'}$, $\theta^{'}$) at time t. Let $b = |lcos\theta|$, $a = lsin\theta$, $b^{'} = |l^{'}cos\theta^{'}|$ and $a^{'} = l^{'}sin\theta^{'}$ be node i's position relative to node pair (j, k) in the horizontal and vertical scales at time $t - \Delta t$ and t, respectively. We have node i's movement relative to node pair (j, k) over Δt as follows:

$$\mathcal{M}_{jk}^{i}(t - \Delta t, t) = \sqrt{(b^{'} - b)^2 + (a^{'} - a)^2} \tag{5.4}$$

The *intra-vicinity dependency*, $\mathcal{MV}_i$, is defined as a measure of the average relative movements between a mobile node i and the pairs of nodes in its one-hop vicinity. $\mathcal{MV}_i$ can be

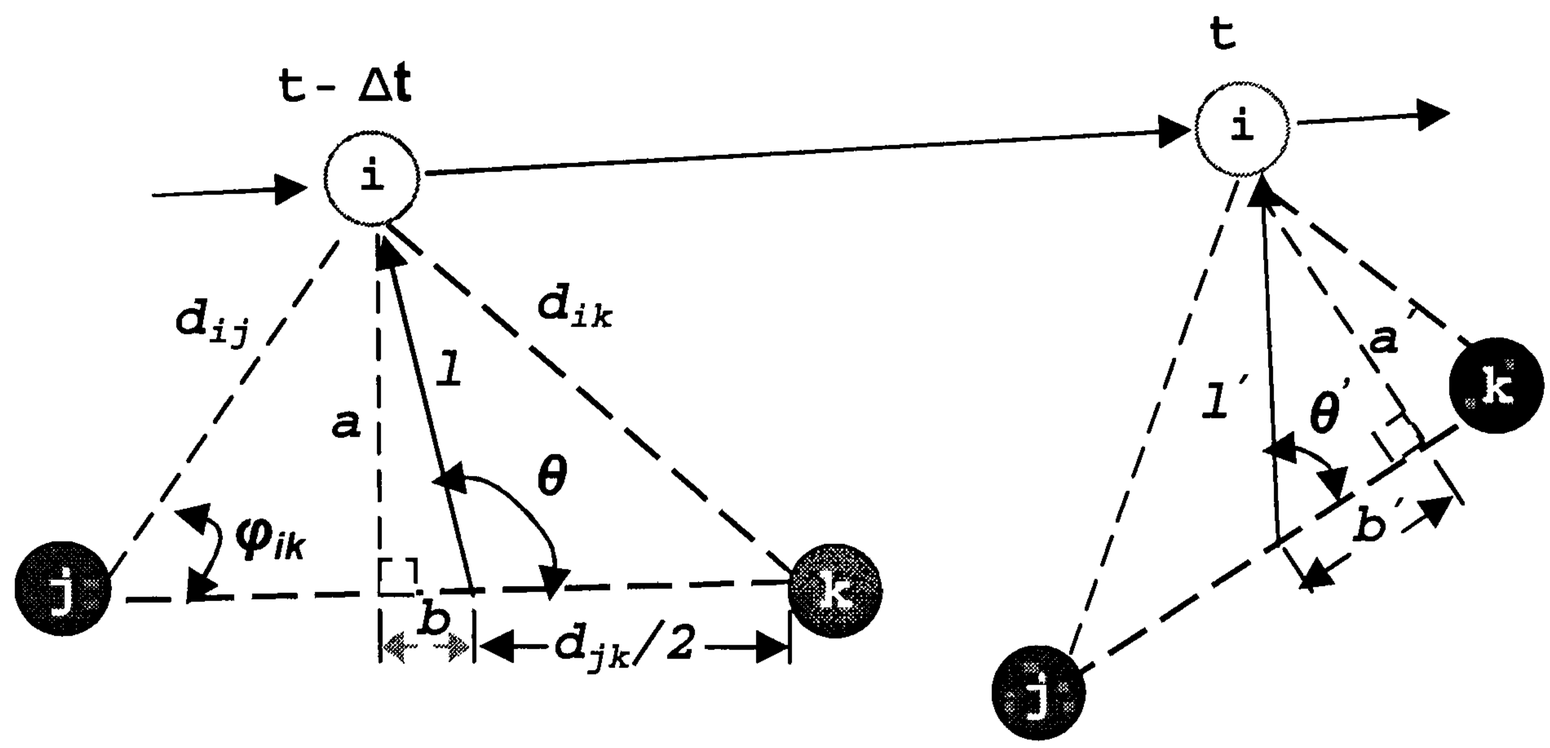


Figure 5.2: The measurement of the relative movement between node i and its neighbouring node pair (j, k) over time slot Δt

measured at every time slot, Δt, by:

$$\mathcal{MV}_i(t-\Delta t, t) = \frac{1}{|\mathcal{U}_i|} \sum_{j,k \in \mathcal{U}_i} \frac{\mathcal{M}^i_{jk}(t-\Delta t, t)}{\Delta t} \tag{5.5}$$

where $|\mathcal{U}_i|$is the cardinality of the set $\mathcal{U}_i$.

5.3.2 The Inter-Group Mobility

When the movements of mobile nodes follow a group mobility model (e.g. RPGM), the connections between nodes in the same mobility group is relatively stable and topological changes are mainly due to the relative movements between mobility groups. Therefore, it is necessary to define a mobility metric for the inter-group relative motions. The *inter-group mobility* $\mathcal{MG}_i$ is defined as a metric of the average relative moving speed between node i's mobility group and the neighbouring groups. In Fig. 5.2, if node i belongs to one mobility group and node pair (j, k) belongs to another, a good approximation of the relative movement between these two groups can be obtained using Eq. 5.4. Let $\mathcal{U}_i^G$ be the set of node pairs that are one-hop away from node i but not belonging to i's mobility group, we have $\mathcal{U}_i^G \subseteq \mathcal{U}_i$. $\mathcal{MG}_i$

can be given at every Δt as:

$$\mathcal{MG}_i(t-\Delta t,t) = \frac{1}{|\mathcal{U}_i^G|}\sum_{j,k\in\mathcal{U}_i^G}\frac{\mathcal{M}_{jk}^i(t-\Delta t,t)}{\Delta t} \tag{5.6}$$

where $|\mathcal{U}_i^G|$is the cardinality of the set $\mathcal{U}_i^G$, which can be obtained by deploying the Sequential Clustering (SC) algorithm [98] to classify node i's movements relative to its vicinity, i.e. $\{\mathcal{M}_{jk}^i(t-\Delta t,t), j,k \in \mathcal{U}_i\}$.

5.3.3 The Intra-Vicinity Mobility

The *intra-vicinity mobility* is proposed to capture the impact of nodal mobility on the network connectivity. The *intra-vicinity dependency* introduced at the beginning of this section is a reflection of the amount of inter-node relative movements for networks that follow random mobility models (e.g. RWP). However, the network connectivity is not determined by the *intra-vicinity dependency* alone. The impact of the inter-node relative movements on the network connectivity also depends on the radio radius, r, i.e. the longer the radio radius the less the impact of the relative mobility on the network connectivity and vice versa. The *intra-vicinity mobility*, $\mathcal{MR}_i$, is defined as a metric of the average relative movements between a mobile node i and the node pairs in its one-hop vicinity within the radio range r. The value of $\mathcal{MR}_i$ over time Δt can be simply derived from Eq. 5.5 as follows:

$$\mathcal{MR}_i(t-\Delta t,t) = \frac{\mathcal{MV}_i(t-\Delta t,t)}{r} \tag{5.7}$$

5.4 Simulation and Evaluation

In this section, the ns-2 simulator is deployed to simulate various network scenarios covering a wide range of nodal mobility. The performance of the proposed mobility metrics is evaluated in these scenarios. The experimental results collected from the simulations are pre-

sented.

5.4.1 The Simulation Model

The simulated mobile ad-hoc network consists of $\mathcal{N} = 40$ nodes moving around in a square area with the size of $\mathcal{L} \times \mathcal{L}$. Using both of the mobility modelling tools described in [91] and [99], the simulations cover a variety of scenarios and mobility models (e.g. the Random Waypoint model (RWP), the Restricted Random Waypoint (R2WP), the City Section Mobility model (CSM) and the Reference Point Group Mobility (RPGM) model).

The Restricted RWP mobility [91] was originally introduced to model the movement of mobiles within/between towns distributed in a large area. In this study, R2WP is deployed to model the nodal mobility of an indoor scenario (e.g. a shopping mall with three large shops/-rooms), in which mobiles have higher probability to perform restricted random movements in one of the three rooms and less chance to move between rooms. The CSM model (or the Manhattan Grid model in [99]) try to model nodal mobility in the street network of a metropolitan city. Using this model, mobiles move in predefined streets that are separated by 12 building blocks and have a possibility of 0.6 to perform 90-degree turns at intersections. In the RPGM model, the 40 nodes are divided into four mobility groups (10 nodes per group). The motion of a group is defined by its logical centre, which moves according to the RWP model. The movement of an individual node is specified by the combination of the group motion and a random motion vector $\overrightarrow{\mathbf{RM}}$ [11]. In this experiment, the length and the direction of $\overrightarrow{\mathbf{RM}}$ have uniform distributions over a chosen range of [0, 120m] and [0, 2π), respectively.

For these mobility models, the pause probability of each node (group) is kept at 0 to produce continuous nodal (group) movements. The movement speed of a mobile node is uniformly distributed over the interval $[v_{min}, v_{max}]$, where v_{min} is fixed at 2.5m/s and v_{max} is varied from 5m/s (18km/hour) to 40m/s (144km/hour) to create different mobility scenarios.

The inter-node distances and mobility metrics are measured at intervals of $\Delta t = 1$ second for a duration of 300 seconds starting 1250 seconds after initialization. In the figures generated

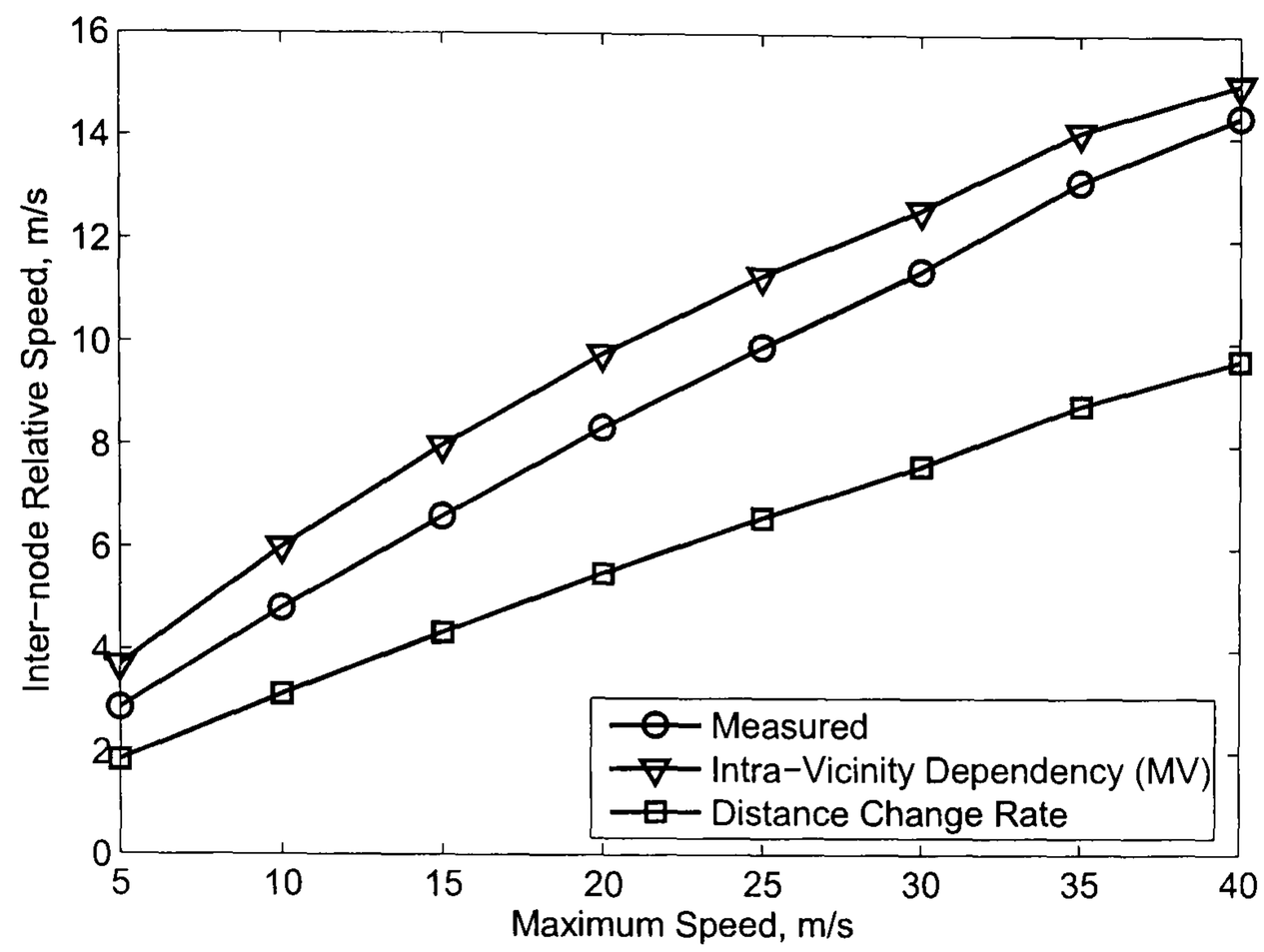


Figure 5.3: Comparisons between the Intra-Vicinity Dependency and the Distance Change Rate as a substitute for the actual average inter-node relative speed for the RWP model

from the simulations, each data point corresponds to a mean of the results generated in 30 repeated experiments performed with different random seeds.

5.4.2 The Experiments and Results

Four sets of experiments are carried out to evaluate the performance of the proposed mobility metrics and to demonstrate their applications in various scenarios.

Experiment 1:

The first experiment is to evaluate the performance of the *intra-vicinity dependency* and the *inter-group mobility* as a substitute for the average inter-node relative speed for networks following the RWP, R2WP and CSM models and the inter-group relative speed for those following the RPGM model, respectively. In this experiment, r is fixed at 150m while $\mathcal{L}$ is set at 800m.

Fig. 5.3 gives a comparison between $\mathcal{MV}$ (the network and time mean of $\mathcal{MV}_i$) and the

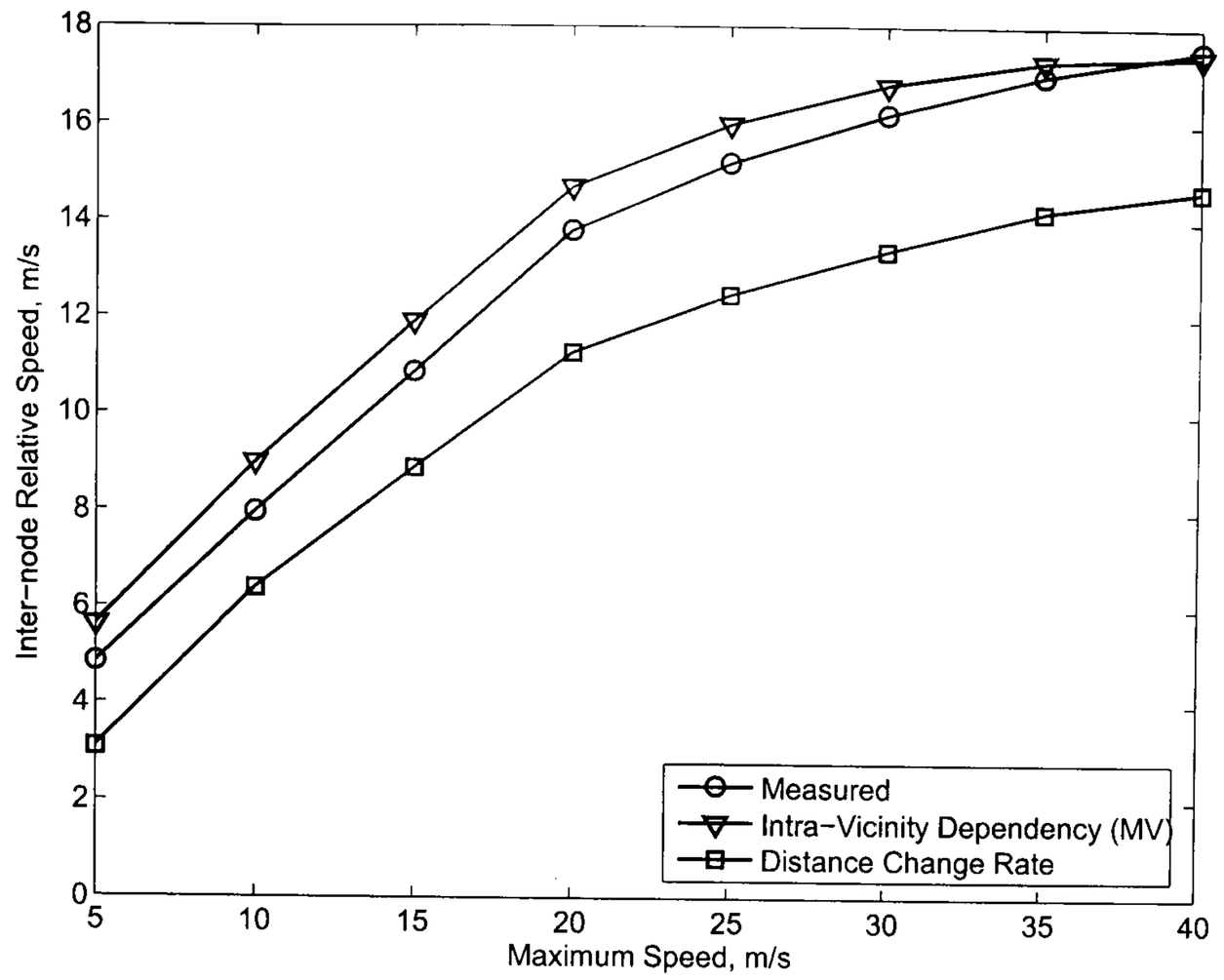


Figure 5.4: Comparisons between the Intra-Vicinity Dependency and the Distance Change Rate as a substitute for the actual average inter-node relative speed for the R2WP model

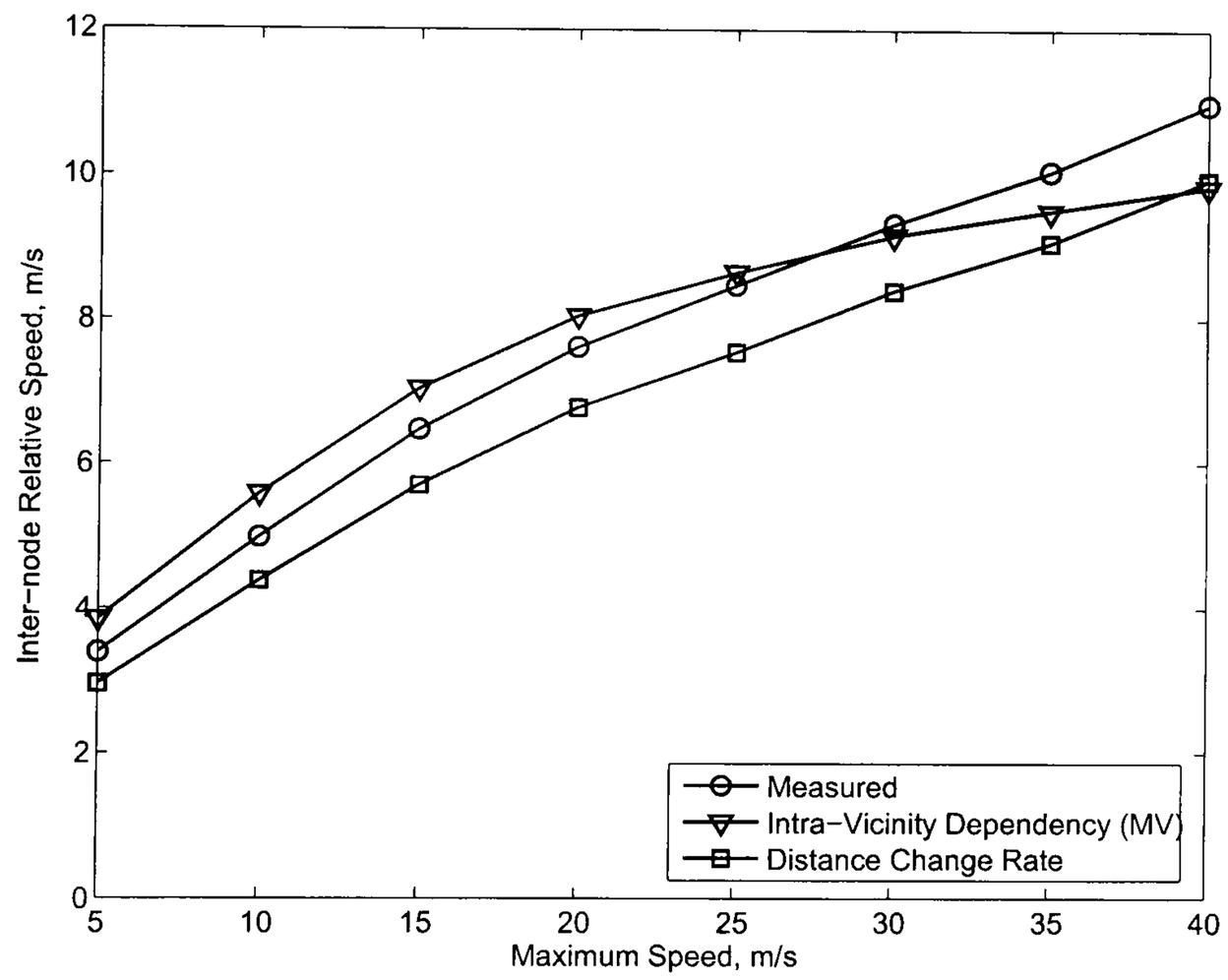


Figure 5.5: Comparisons between the Intra-Vicinity Dependency and the Distance Change Rate as a substitute for the actual average inter-node relative speed for the CSM model

mean inter-node *distance change rate*‡ [63] [62] regarding their similarity to the average inter-

‡Given the time varying distance estimate between node i and j, $d_{ij}(t)$, the distance change rate between i and j is given by: $\frac{d}{dt}d_{ij}(t)$

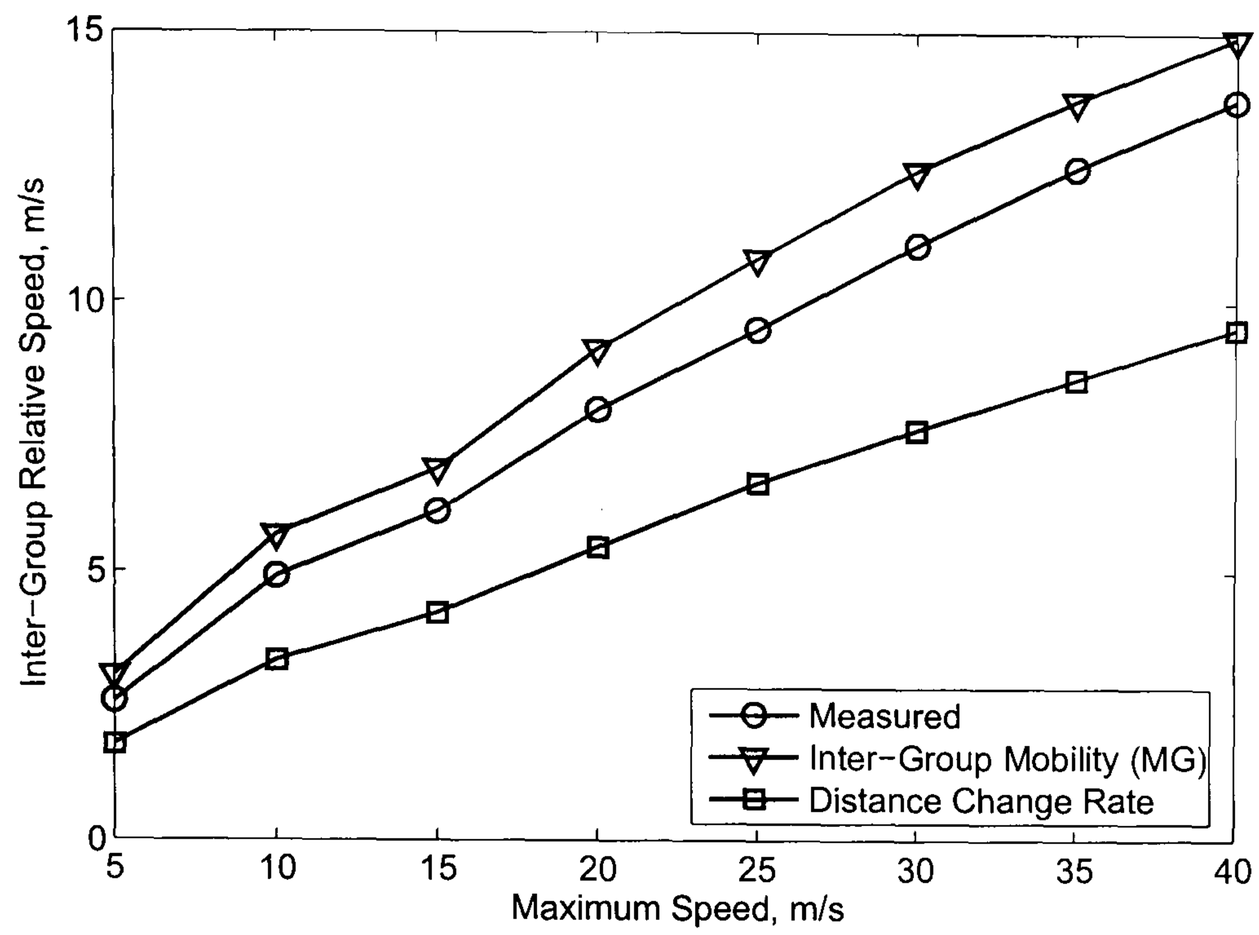


Figure 5.6: Comparisons between the Inter-Group Mobility and the Distance Change Rate as a substitute for the actual average inter-group relative speed for the RPGM model

node relative speed calculated from the location information[§] of mobile nodes moving in the RWP model. As shown in Fig. 5.3, with an average deviation of less than 1.5m/s the mean *intra-vicinity dependency*, $\mathcal{MV}$, is a good estimate of the average inter-node relative speed over the wide range of maximum speed limits. On the contrary, as it only captures the relative motion in the direction of line of sight, the *distance change rate* underestimates the average inter-group relative speed especially with high speed limits.

Fig. 5.4 presents the results of the same comparison between $\mathcal{MV}$ and the mean *distance change rate* for the R2WP model. As shown in Fig. 5.4, the performance of $\mathcal{MV}$ is still promising. The average deviation between $\mathcal{MV}$ and the measured mean relative speed is only 0.7m/s. Compared to Fig. 5.3, the gap between the *distance change rate* and the measured relative speed appears to be smaller. This may be due to the fact that nodal mobility in the R2WP model are less-random, as their orientations are often restricted by the shape of the rooms.

[§]Let $(\mathcal{V}_x^i, \mathcal{V}_y^i)$ and $(\mathcal{V}_x^j, \mathcal{V}_y^j)$ be node i and j's motion vector, respectively. The relative speed between i and j can be calculated by: $\sqrt{(\mathcal{V}_x^i - \mathcal{V}_x^j)^2 + (\mathcal{V}_y^i - \mathcal{V}_y^j)^2}$

In Fig. 5.5, a similar comparison between $\mathcal{MV}$ and the mean inter-node *distance change rate* has been also made for networks moving in the CSM model. Fig. 5.5 shows that, in the CSM model, $\mathcal{MV}$ generally has a good match with the average inter-node relative speed, except that at the speed limit of 35m/s or higher $\mathcal{MV}$ is as inaccurate as the *distance change rate*. This is due to the negative effects of significant number of 90-degree turns from cars arriving at intersections. Fig. 5.5 also shows that the *distance change rate* has a better accuracy in approximating the average inter-node relative speed in the CSM model than in the RWP and the R2WP model (see Fig. 5.3 and Fig. 5.4). The reason is because in the CSM model most of the relative movements between mobiles are one-dimensional in the network of streets hence they can be well captured by the *distance change rate*.

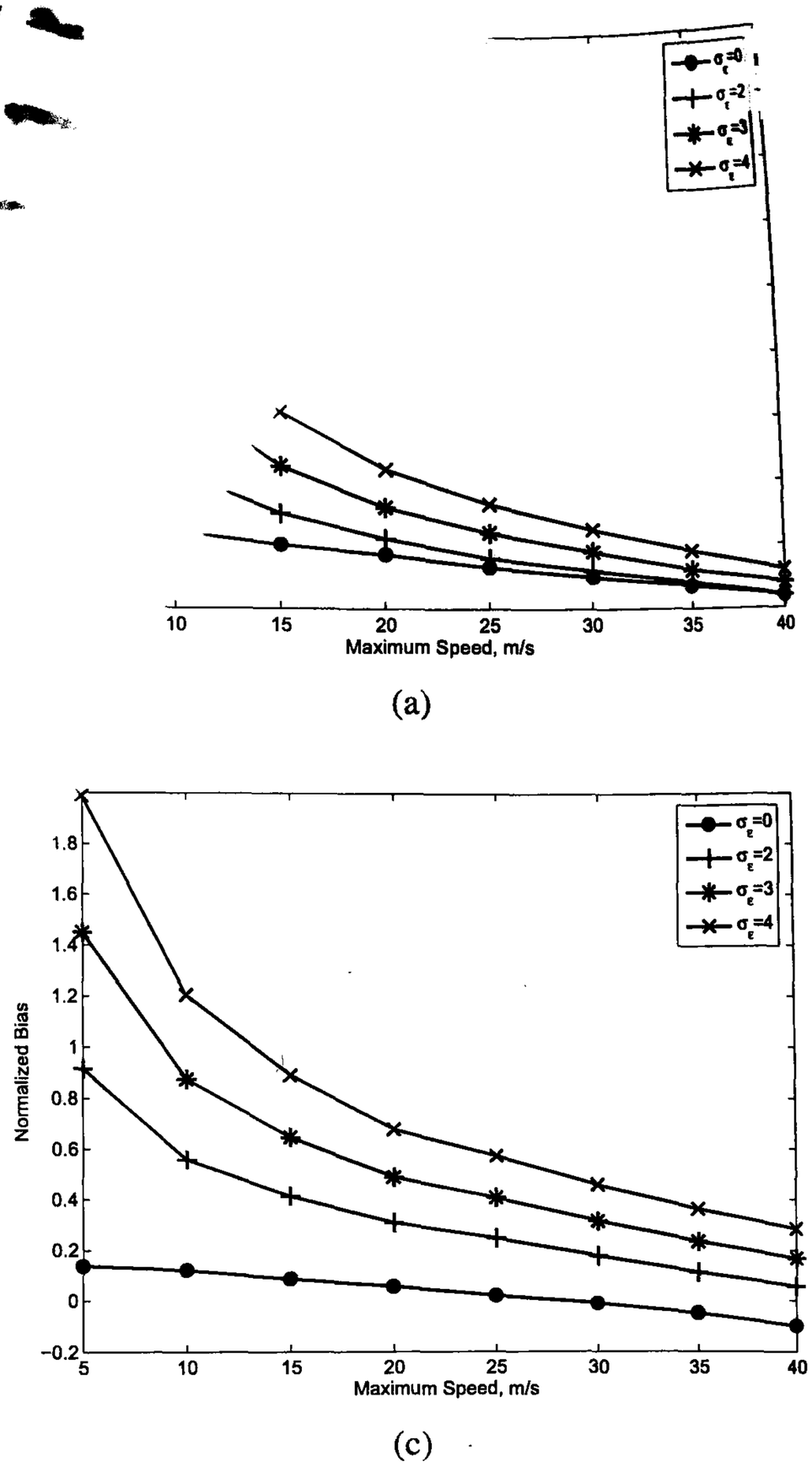
Fig. 5.6 compares $\mathcal{MG}$ (the network and time average of $\mathcal{MG}_i$) and the mean *distance change rate* against the actual inter-group relative speed obtained from the nodal location information. As shown in Fig. 5.6, $\mathcal{MG}$ is a more accurate approximation of the inter-group relative speed than *distance change rate* regardless of the speed limits (the average deviation is less than 1m/s). The *distance change rate* is shown to have underestimated the inter-group relative speed due to the aforementioned reason.

Experiment 2:

In the second experiment, the performance of the *intra-vicinity dependency* and the *inter-group mobility* with noisy range information is examined. Without taking the effect of Non-Line-of-Sight (NLOS) into account, a distance estimate (e.g. measured using TDoA) from a noisy environment at time t (denoted as $\hat{d}_t$) can be simply modelled as:

$$\hat{d}_t = d_t + \epsilon_t \tag{5.8}$$

where d_t is the true distance and ϵ_t is the measurement noise that can be modelled as a zero-mean Gaussian random variable with variance σ_ϵ^2. In the experiment, σ_ϵ ranges between 0m



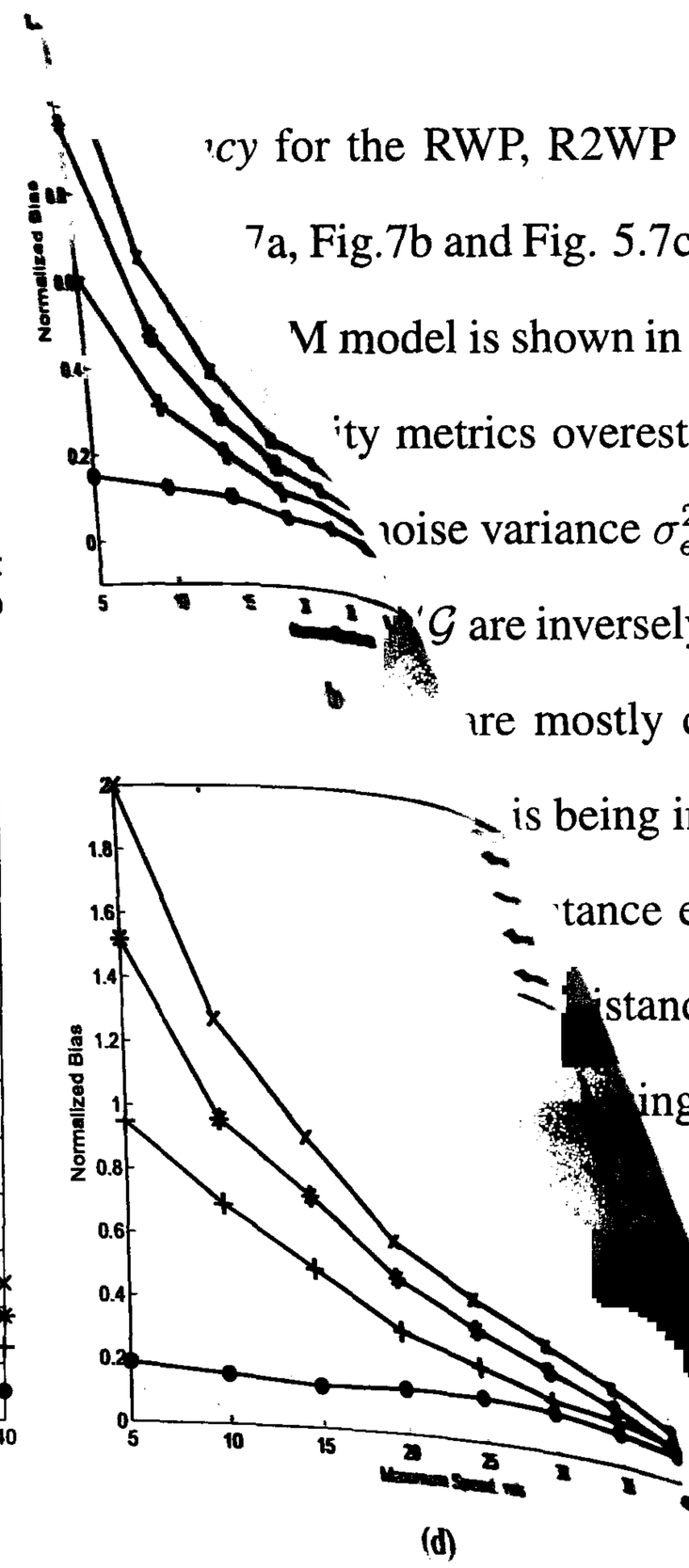


Figure 5.7: the Normalized Bias of the mobility metrics. (a) Intra-Vicinity Dependency, RWP model, (b) Intra-Vicinity Dependency, R2WP model, (c) Intra-Vicinity Dependency, CSM model, (d) Inter-Group Mobility, RPGM model

eed. The normalized bias of $\mathcal{MG}$ can be

cy for the RWP, R2WP and the CSM

7a, Fig.7b and Fig. 5.7c, respectively.

M model is shown in Fig. 5.7d. We

ty metrics overestimate the real

oise variance σ_ϵ^2 regardless of

$\mathcal{G}$ are inversely proportional

re mostly caused by the

is being increased. Fig.

tance errors (e.g. to

stance errors for a

ing $\mathcal{MV}$ or $\mathcal{MG}$

ainst various

stimated as

ig. 5.8 to

ce has a

to 4m, which is achievable with off-the-shelf products [100]. I evaluated the normalized bias σ_ϵ between $\mathcal{MV}$, the mean *intra-vicinity dependency*, and the true average inter-node relative speed, as well as that between $\mathcal{MG}$, the mean *inter-group mobility*, and the actual average relative speed between neighbouring mobility groups. The normalized bias of $\mathcal{MV}$, denoted as $\sigma_{\mathcal{MV}}$, is expressed as:

$$\sigma_{\mathcal{MV}} = \frac{\mathcal{MV}}{\bar{v}} - 1 \tag{5.9}$$

where $\bar{v}$ is the true average inter-node relative speed. The normalized bias of $\mathcal{MG}$ can be expressed in a similar way as in Eq. 5.9.

The normalized bias of the *intra-vicinity dependency* for the RWP, R2WP and the CSM model against increasing speed limits are shown in Fig. 5.7a, Fig.7b and Fig. 5.7c, respectively. The normalized bias of the *inter-group mobility* for the RPGM model is shown in Fig. 5.7d. We can observe from Fig. 5.7 that in noisy conditions the mobility metrics overestimate the real network mobility with a normalized bias that increases with the noise variance σ_ϵ^2 regardless of the mobility models. However, the normalized bias for $\mathcal{MV}$ and $\mathcal{MG}$ are inversely proportional to the nodal speed. This is due to the fact that estimation errors are mostly caused by the environmental noise and their values keep stable while the nodal speed is being increased. Fig. 5.7 indicates that a *calibration factor* can be derived from estimated distance errors (e.g. to deploy a calibrated Ray Tracing software to analyze and model the ToA distance errors for a specific environment [101]) and the average inter-node relative speed (e.g. using $\mathcal{MV}$ or $\mathcal{MG}$ as a substitute).

Fig. 5.8 is a plot of the estimation errors (i.e. $|\mathcal{MV}-\bar{v}|$) for the R2WP model against various combinations of inter-node relative speeds (measured by $\mathcal{MV}$) and noise levels (estimated as σ_ϵ). We have also produced a 3rd order polynomial surface fitting for the data in Fig. 5.8 to predict estimation errors based on nodal speed and noise variance. The fitted surface has a goodness (R-square) value of 0.994. The fitting function, $\mathcal{F}^c(\mathcal{MV}, \sigma_\epsilon)$, is given as:

$$\begin{aligned}\mathcal{F}^c(\mathcal{MV}, \sigma_\epsilon) = {} & 1.5132 + 1.2799\mathcal{MV} - 0.3563\sigma_\epsilon + 0.6004\mathcal{MV}^2 + 0.0496\sigma_\epsilon^2 - 0.1893\mathcal{MV}\sigma_\epsilon \\ & -0.0127\mathcal{MV}^3 - 0.0018\sigma_\epsilon^3 + 0.0055\mathcal{MV}\sigma_\epsilon^2 - 0.0184\mathcal{MV}^2\sigma_\epsilon \qquad (\mathcal{MV}>0,\ \sigma_\epsilon>0)\end{aligned} \tag{5.10}$$

The function $\mathcal{F}^c(\mathcal{MV}, \sigma_\epsilon)$ can be used to produce a calibration factor to reduce the estimation errors caused by inaccurate distance information. As a result, a calibrated *intra-vicinity*

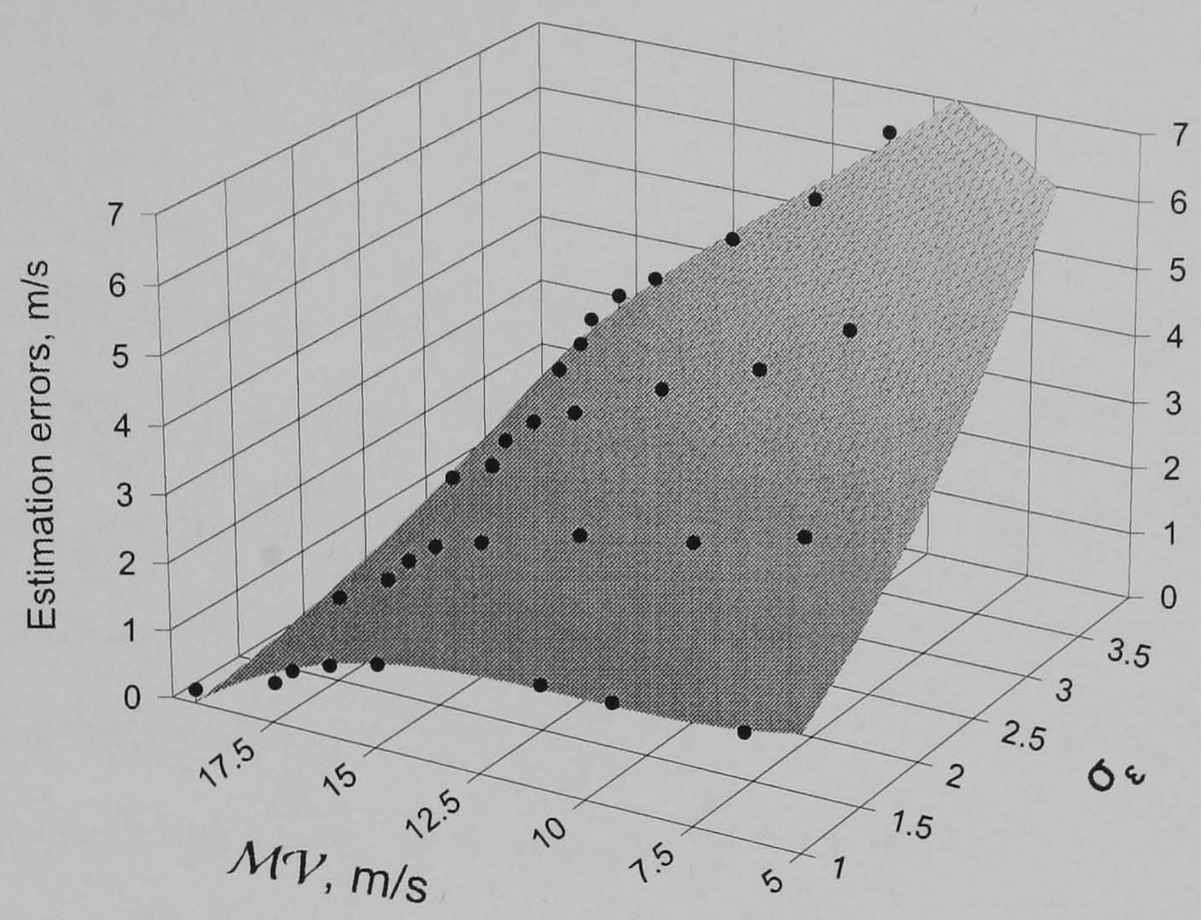


Figure 5.8: Estimation errors vs. the Intra-Vicinity Dependency and noise variance, R2WP model

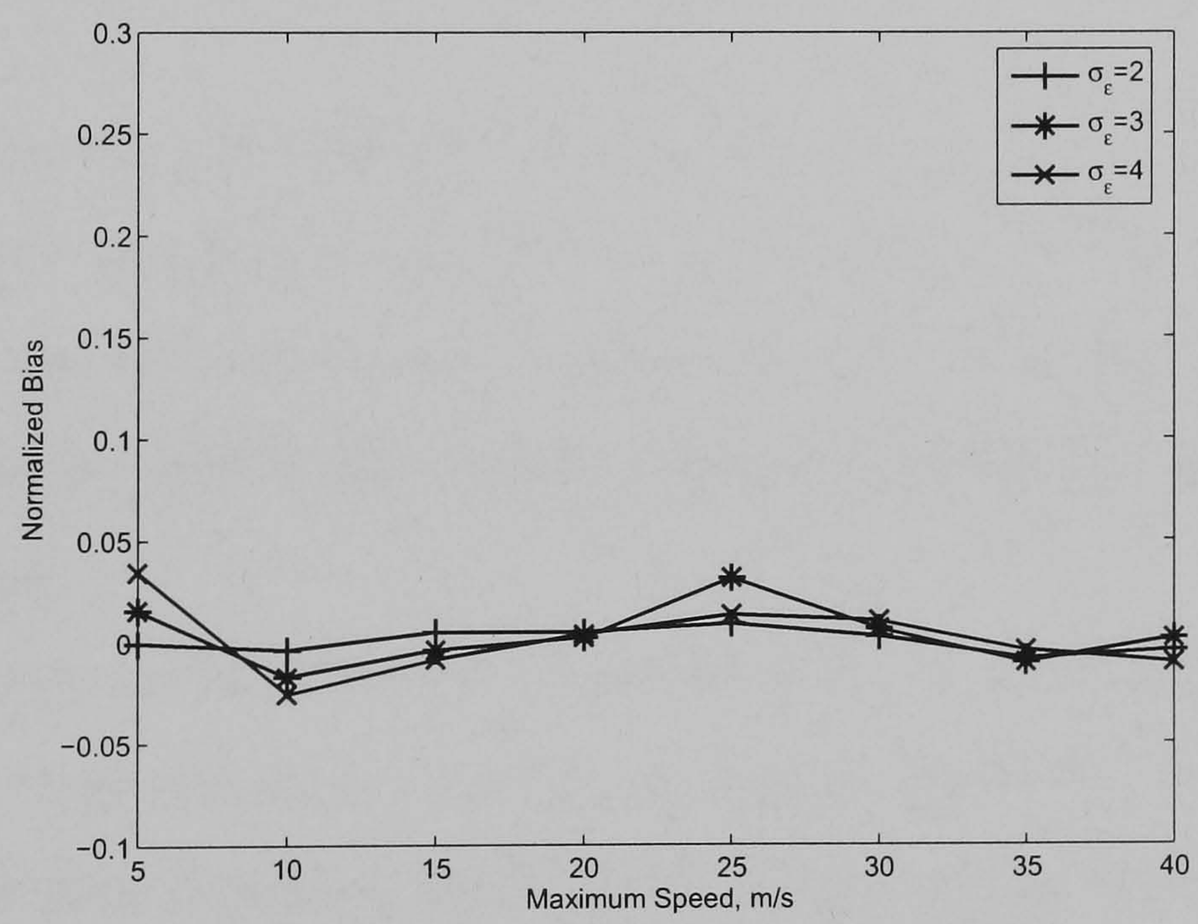


Figure 5.9: the Normalized Bias of the calibrated Inter-Vicinity Dependency, R2WP model

dependency, $\mathcal{MV} - \mathcal{F}^c(\mathcal{MV}, \sigma_\epsilon)$, can be used for noisy environments. Fig. 5.9 shows the near-perfect normalized bias for the calibrated $\mathcal{MV}$, which proves that the method is effective

for the R2WP model. Calibration factors for other mobility models can be derived with the same function but using different parameters, as the mobility models have different levels of sensitivity to noisy distance estimates due to their distinct nodal movement patterns.

Experiment3:

The third experiment is to demonstrate the application of the *inter-group mobility* for estimating the inter-group inter-meeting times for networks following the RPGM model. The simulation settings of this experiment are similar to those of the first experiment, except that the simulation period after initialization is extended to 1500 seconds for collecting the actual inter-group inter-meeting times. In the RPGM model, the logical centre of each mobility group is actually moving according to the RWP model. Therefore, the mean inter-meeting times between mobility groups $\mathcal{T}^G$ can be predicted using Eq. 5.11 [86, p. 93].

$$\mathcal{T}^G = \frac{\mathcal{L}^2}{2\omega r^G \bar{\mathcal{V}}} \tag{5.11}$$

where $\omega \approx 1.3683$ is a specific constant for the RWP model, $\bar{\mathcal{V}}$ is the mean inter-group relative speed and r^G is the combined radio range of a mobility group. We have $r^G \geq r$ and $r^G \rightarrow r$ when $r \gg |\overrightarrow{RM}|$. The value of r is used to approximate r^G for simplicity. Assuming the speed information of mobile nodes is not available, I tried to substitute $\bar{\mathcal{V}}$ with $\mathcal{MG}$ or the mean *distance change rate*.

Fig. 5.10 compares the performance of estimating the mean inter-group inter-meeting times using $\mathcal{MG}$ (the average *inter-group mobility*) and the average *distance change rate* as substitutes for the true average inter-group relative speed. As demonstrated in Fig. 5.10, using $\mathcal{MG}$ for $\bar{\mathcal{V}}$ in Eq. 5.11 the predicted group inter-meeting times is very close to those that are measured using simulation. The use of the *distance change rate* approach leads to an overestimation of $\mathcal{MG}$ due to the fact that the *distance change rate* is an underestimation of the average inter-group relative speed $\bar{\mathcal{V}}$. When the network mobility is low, such an overestimation can exceed

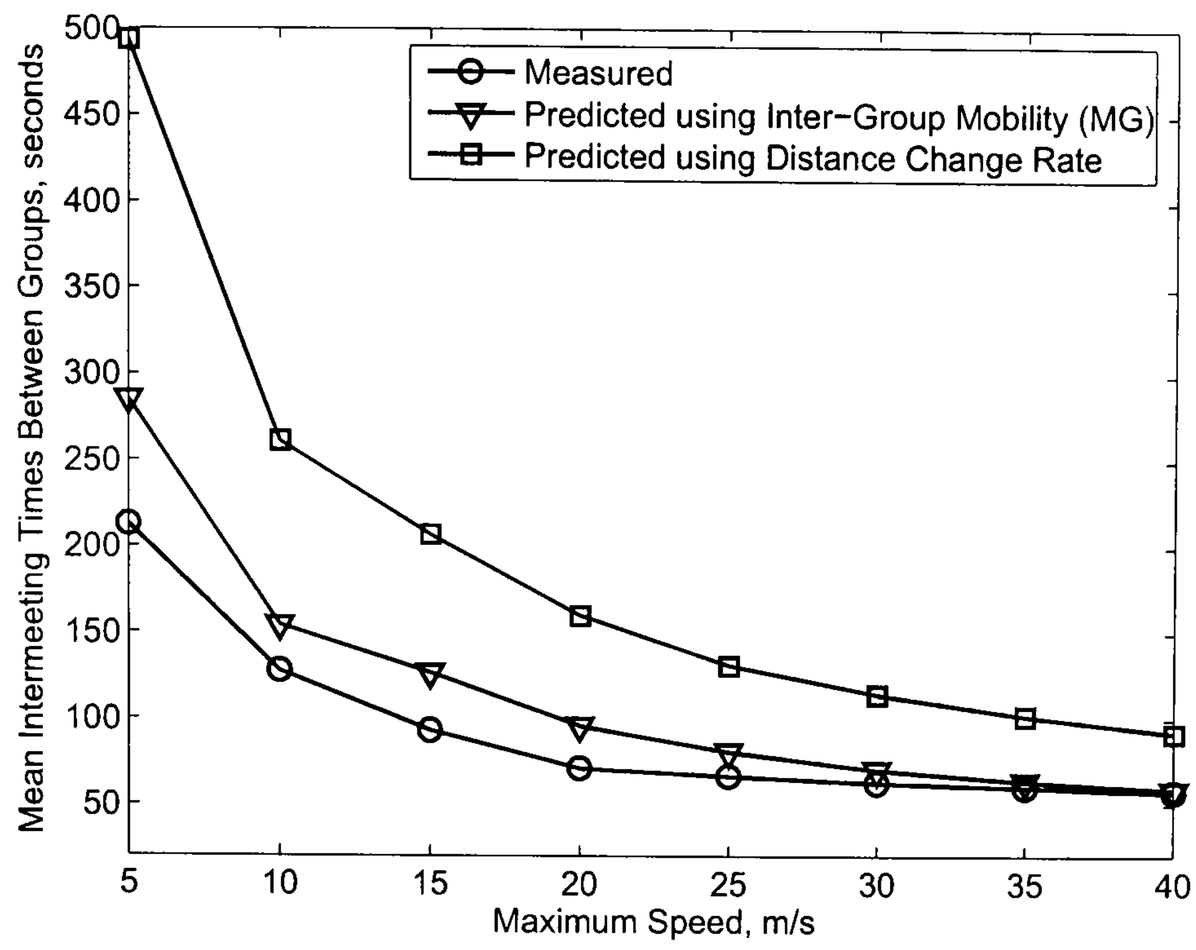


Figure 5.10: Intermeeting Times Between Groups vs. Maximum Speed Limit

200 seconds.

Experiment 4:

The fourth experiment is to investigate the network performance when the *intra-vicinity mobility* is used as a metric for the packet delivery rate. In this experiment, three UDP data source/sink pairs are randomly chosen from a network following the RWP mobility model to generate a total of 50 packets per second of Constant Bit Rate (CBR) traffic in the network. More source/sink pairs are also tested. But in these cases more packets are lost due to congestion, which makes it difficult to distinguish the impact of mobility from that of congestion. The radio radius r is varied from 100m to 250m (with a step of 50m) and the side length of the simulation area $\mathcal{L}$ from 400m to 1000m (with a step of 200m) to create 4 sets of network scenarios with a constant node density.

Fig. 5.11 gives a plot of the CBR data packets delivery rate as the mean *intra-vicinity mobility* $\mathcal{MR}$ increases. From Fig. 5.11 we see a strong relationship between the packet delivery rate and $\mathcal{MR}$, for which a 3rd order polynomial curve fitting function is derived (Eq.

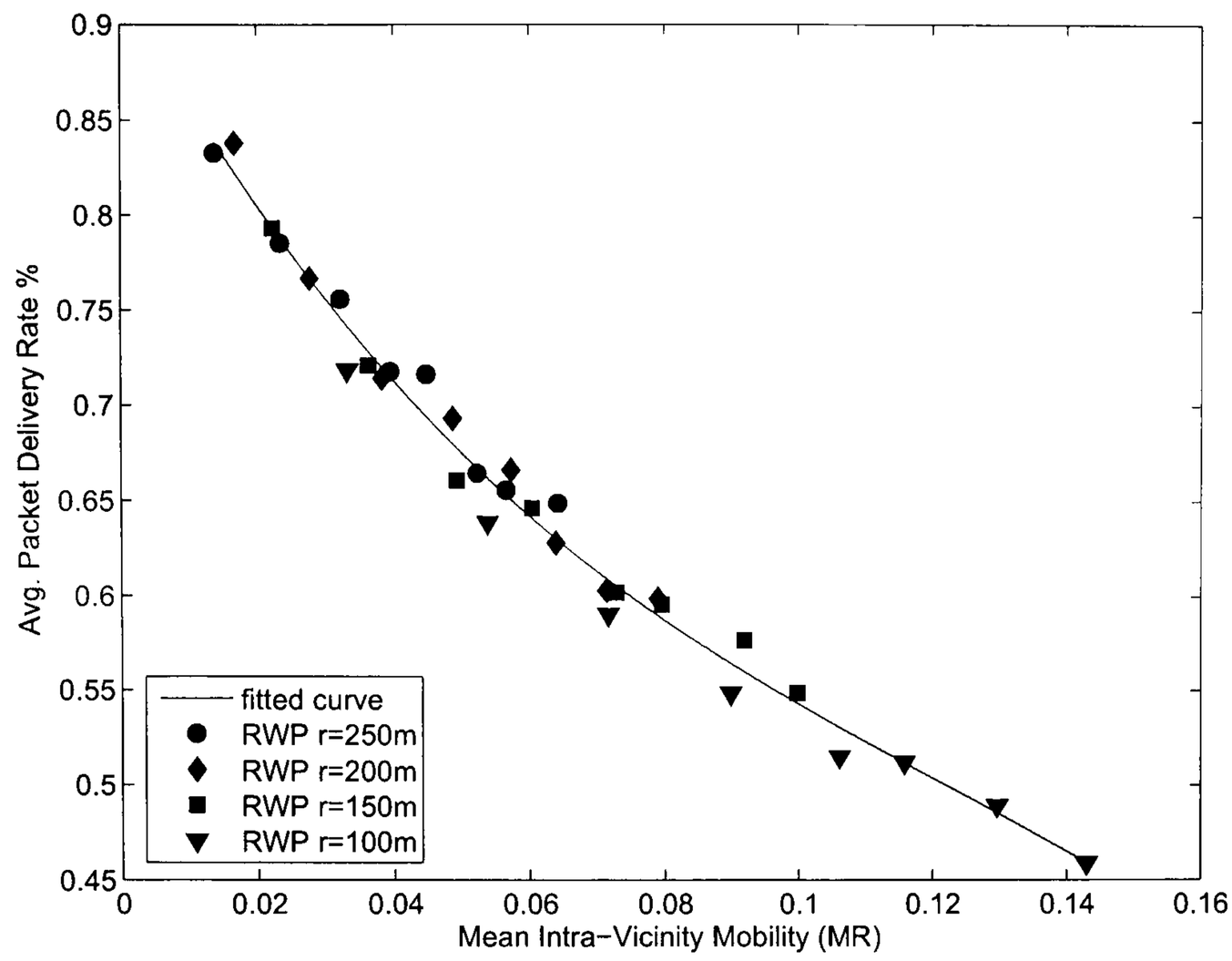


Figure 5.11: Average Packet Delivery Rate vs. Mean *Intra-Vicinity Mobility*

5.12). The goodness of the fit is 0.985. The function $\mathcal{F}^r(\mathcal{MR})$ can be used by adaptive protocols to estimate the performance of packet delivery given the measured network mobility.

$$\mathcal{F}^r(\mathcal{MR}) = -108.1\mathcal{MR}^3 + 39.14\mathcal{MR}^2 - 6.627\mathcal{MR} + 0.9224 \qquad (\mathcal{MR}>0) \tag{5.12}$$

5.5 Summary

In this Chapter, a new range-based GPS-free mobility metric, the *intra-vicinity dependency*, is presented for measuring the performance of mobile ad-hoc networks. This metric can fully capture the relative motions between a mobile node and its vicinity in a 2D plane, in real-time, using simple triangulation. Variants of this metric are also proposed for predicting the performance of networks that follow group and random mobility models (e.g. inter-group intermeeting times and packet delivery rate). To deal with estimation errors introduced by noisy distance estimates, a calibration method is also proposed for improving the accuracy of the

mobility metrics. The performance of the metrics is validated using ns-2 based simulations with several mobility models (i.e. RWP, R2WP, CSM, and RPGM). Experimental results show that, without the help of any location/speed information of mobile nodes, the proposed metrics enable a more accurate approximation of the average inter-node/inter-group relative speed than the existing method. It is also shown in the experiments that the proposed metrics yield excellent performance when they are used to predict the inter-group intermeeting times for networks that follow the RPGM model and to estimate the packet delivery rate for those that follow the RWP model. The GPS-free metrics are ideal candidates for mobility measurement in energy-constrained mobile sensor/ad-hoc networks. The reliability and accuracy of these metrics make them useful for ad-hoc routing protocols to adapt packet delivery for topological dynamics.

Chapter 6

Range-based Relative Velocity Estimations

6.1 Introduction

The relative velocity between mobile devices is one of the key impairment factors that affect the quality of communications in MANETs. Velocity estimations of mobile devices has been widely used in traditional cellular mobile networks in handover processes, cell-layer-assignment and dynamic channel assignment [102]. Accurate and reliable estimation of the relative velocity between mobile nodes is also important for the emerging wireless mobile ad hoc networks. The performance of communications in mobile ad hoc networks suffers from frequent topological changes caused by the relative movements between mobile nodes. Knowledge of the inter-node relative velocity would help to predict topological dynamics (e.g. link availability prediction [103]) and to optimize routing protocols (e.g. mobility metrics [57]) for improving the quality of communication.

Conventionally, velocity estimates are obtained from localization systems such as GPS. If the characteristics of signal propagation can be identified, it is also possible to obtain the velocity estimates based on the statistics of the received signal. A recently proposed method [66] makes use of the time-varying inter-node range information for velocity estimations provided that the range estimates are noise-free.

Chapter 5 introduced a GPS-free mobility metric that can be used as a substitute for the network-wide average of inter-node relative speed in MANETs. The metric is useful in es-

timating the network average of inter-node relative speed and in quantifying the impact of mobility on the performance of a network. But it cannot be used to estimate the relative speed between a pair of nodes. This Chapter presents a new Range-based method for relative Velocity Estimations (RVE). Two Range-based relative Velocity Estimators (i.e. RVEs and RVEd) are derived for both Sparse and Dense ad-hoc networks. In addition to being less dependent on the characteristics of the wireless channel, the proposed method is more tolerant of the multi-path or Non-Line-of-Sight (NLOS) errors contained in range measurements. Thus, it is more robust than the existing method in noisy communication environments. An efficient analytical algorithm for predicting the expected node degree in networks with uniform or non-uniform node distributions is also given in order to provide mobile nodes the capability of examining the applicability of the two estimators.

In two simulation-based experiments designed for networks of sparse and dense node density, the proposed method produced velocity estimates that exhibit excellent match with the actual values regardless of the nodal speed limits or the distribution of noises. Compared to the existing method, the proposed method is shown to achieve an improvement factor of about 33 (in terms of normalized bias) with Gaussian multi-path noises of 4m standard deviation, or 20 with Uniform NLOS noises of 32m maximum.

The organization of this Chapter is as follows. Related works on velocity estimations for mobile systems are introduced in Secioton 6.2. In Section 6.3, two Range-based Relative Velocity Estimators (i.e. *RVEs* and *RVEd*) are presented for sparse and dense MANETs, respectively. Section 6.4 gives details of the simulation settings, the simulation results and the numerical analysis. A new approach for predicting the expected node degree for networks of uniform/non-uniform node distribution is presented in Section 6.5. This Chapter is summarized in Section 6.6.

6.2 Related Work

Hitherto, estimation of relative velocity between mobile devices was either based on a localization system, such as GALILEO [93] the Global Positioning System (GPS) [14] or GPS-free positioning system [75] [94] [95], or based on analyzing the characteristics of received signal (e.g. time-frequency characteristics [67] or instantaneous frequency of the received signal [102]).

The GPS signal is normally too weak to be of any use indoors. When used outdoors, GPS signals are often reflected or blocked by buildings. GPS-free positioning systems [75] [104] [94] may be used as alternatives to GPS to trace the movement of a mobile node. However, such positioning systems require the presence of stable anchor nodes, which may not exist in ad-hoc networks where every node is moving constantly.

Various techniques have been proposed for velocity estimations based on characteristics of the wireless channel and/or received signal. For example, Doppler frequency is utilized in [67] [105] [106] [107] for mobile velocity estimation. In [102] Azemi et al. propose to improve the performance of such methods by using the first moment of the instantaneous frequency of the received signal in a Rayleigh fading channel. The algorithm proposed in [108] uses the normalized autocorrelation values of the received signal to provide coarse estimations of the mobile velocity. However, a common problem in these velocity estimators is that they rely on precise knowledge of the characteristics/statistics of specific wireless channels or received signal. This limits their applicability to only scenarios where the pattern of signal propagation is constant.

The relative movements between nodes cause the inter-node distance to change over time. In a recent study, time-varying range information between mobile nodes, instead of the characteristics of the wireless channel/signal, was used for velocity estimation in order to predict link availability [66]. In [66], a Link Prediction Algorithm (LPA) is presented, which estimates relative moving speed/direction between nodes by utilizing inter-node range estimates sampled

within a short period to build a nonlinear system of equations. The velocity estimates are then obtained by solving this system of equations. However, the method is based on the assumption that the range information is measured in noise-free environment [66]. This approach is not feasible in noisy communication scenarios which are the norm.

In this Chapter, a new method of Range-based relative Velocity Estimation (RVE) is proposed, which is less dependent on signal characteristics and is more robust in noisy communication environments. Two velocity estimators, namely *RVEs* and *RVEd*, are derived for sparse and dense ad-hoc networks, respectively. Using triangulation, the only information needed by the proposed method is the inter-node distance measured in real time, i.e. provided by techniques such as Time-of-Arrival (ToA) [100]. Further, the proposed method is more tolerant of the multi-path or NLOS errors contained in range measurements hence is more robust than the existing method for noisy communication environments.

6.3 The Range-Based Relative Velocity Estimations

In this section, the two range-based relative velocity estimators (i.e. *RVEs* and *RVEd*) are presented for networks with sparse and dense node density, respectively. The proposed method requires the information of a mobile node's distance to the nodes in its neighbourhood for estimation of their relative velocity. It is assumed that each node periodically measures the distances to its one-hop neighbours by means of either Received Signal Strength Indication (RSSI), Time-of-Arrival (ToA) or Time-Difference-of-Arrival (TDoA) measurements. A distance estimate produced by any of these methods at time i, $\hat{d}_i$ can be modeled as:

$$\hat{d}_i = d_i + \epsilon_{M,i} + \zeta_{N,i}\epsilon_{N,i} \tag{6.1}$$

where d_i is the true value of the distance and $\epsilon_{M,i}$ denotes the measurement errors caused by multi-path signal. $\epsilon_{N,i}$ represents the measurement noises in Non-Line-of-Sight (NLOS)

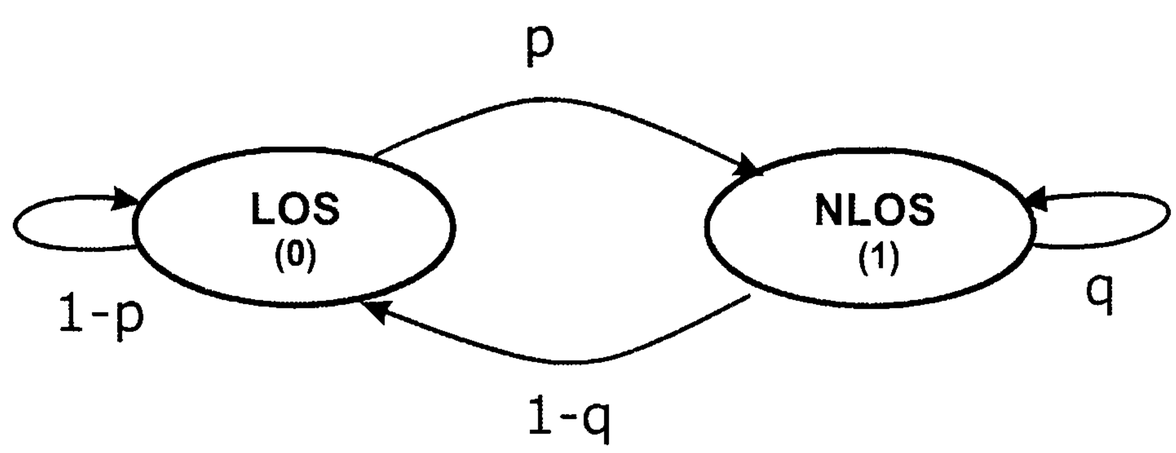


Figure 6.1: The change of channel status modeled as a two-state Markov process

conditions. Both $\epsilon_{M,i}$ and $\epsilon_{N,i}$ can be modeled as Gaussian random variables (zero-mean for the former and positive mean for the later) in a specific environment or by a uniform distribution (with a positive lower range for $\epsilon_{N,i}$) if the signals are measured from various environments [109] [96] [110]. $\zeta_{N,i}$ is a discrete random variable with two possible values, i.e. *0* and *1*. It is deployed to indicate the time-correlated status of the signal propagation conditions (e.g. LOS or NLOS). The alternation between the two status can be modeled as a two-state Markov process (or a Gilbert-Elliot model) [111]. As shown in Fig. 6.1, the channel conditions are alternating between two states, LOS ($\zeta_{N,i} = 0$) and NLOS ($\zeta_{N,i} = 1$). p is the probability that the signal propagation environment will become NLOS given that the current state is LOS. q is the probability that the channel status will be kept as NLOS given that the current state is NLOS. At steady state, the probability of the channel being in the NLOS state ($Pr(\zeta_N) = 1$) can calculated as $\frac{p}{p+1-q}$ and that in the LOS state ($Pr(\zeta_N) = 0$) is given by $1 - \frac{p}{p+1-q}$.

In addition to measuring distances to neighbouring nodes, every node also piggybacks its collection of distance estimates to the periodically broadcasted *HELLO beacons* so that the information is disseminated to all of its one-hop neighbours. Thus, every node knows not only the distances to its one-hop neighbours but also those between neighbouring nodes within the one-hop vicinity. The distance measurements and exchanges among neighbouring nodes are carried out at a frequency of $1/\Delta t$. The time interval Δt should be finely tuned to achieve an optimal trade-off between the resolution requirement of the velocity estimation and the energy consumptions of ranging algorithms as well as the bandwidth usage in broadcasting distance information.

6.3.1 Estimating Relative Velocity in Sparse Networks

The *RVEs* approach is specifically designed for mobile networks with sparse node density, i.e. networks in which the maximum node degree* is 1. When a node is passing through a neighbouring node's radio coverage, its distance to the neighbouring node is first decreasing before it reaches the closest point and then increasing until it exceeds the node radio radius. In this Chapter, *departing neighbours* is used to denote a node's neighbours whose distance is increasing and *arriving neighbours* is for others. If this node does not change its path when it is passing by this neighbour, its trajectory is a straight line and is perpendicular to the closet distance between these two nodes. As shown in Fig. 6.2, d_e is node i's true distance to node j at the time when j reached i's radio coverage. $\hat{d}_e$ is its estimate. $\hat{d}_{p-1}$, $\hat{d}_p$ and $\hat{d}_{p+1}$ are the distance estimates sampled at consecutive time slots (e.g. $t_p - \Delta t$, t_p and $t_p + \Delta t$). $\hat{d}_p$ can be used to approximate the closest distance between node j's trajectory and node i if it satisfies $\{\hat{d}_{p-1} > \hat{d}_p \wedge \hat{d}_{p+1} > \hat{d}_p\}^\dagger$. Let $\mathcal{M}_j(t)$ be the movement of j relative to i during time t with the relative velocity v_{ij}, i.e. $\mathcal{M}_j(\Delta t) = v_{ij}\Delta t$. According to the theorem of Pythagoras, we can easily obtain the estimate of v_{ij}, $\hat{v}_{ij}$, by:

$$\begin{aligned} \hat{v}_{ij} = \frac{\mathcal{M}_j(\Delta t)}{\Delta t} &= \frac{\sqrt{\hat{d}_{p-1}^2 - \hat{d}_p^2}}{\Delta t} \\ &= \sqrt{\frac{d_{p-1}^2 - d_p^2}{\Delta t^2} + \frac{\varepsilon(\hat{d}_{p-1}, \hat{d}_p)}{\Delta t^2}} \end{aligned} \tag{6.2}$$

*Defined as the number of one-hop neighbours (or links) of a node.

†Without loss of generality, it is assumed that this relationship is not affected by measurement noise as the noise levels in adjacent time slots are similar.

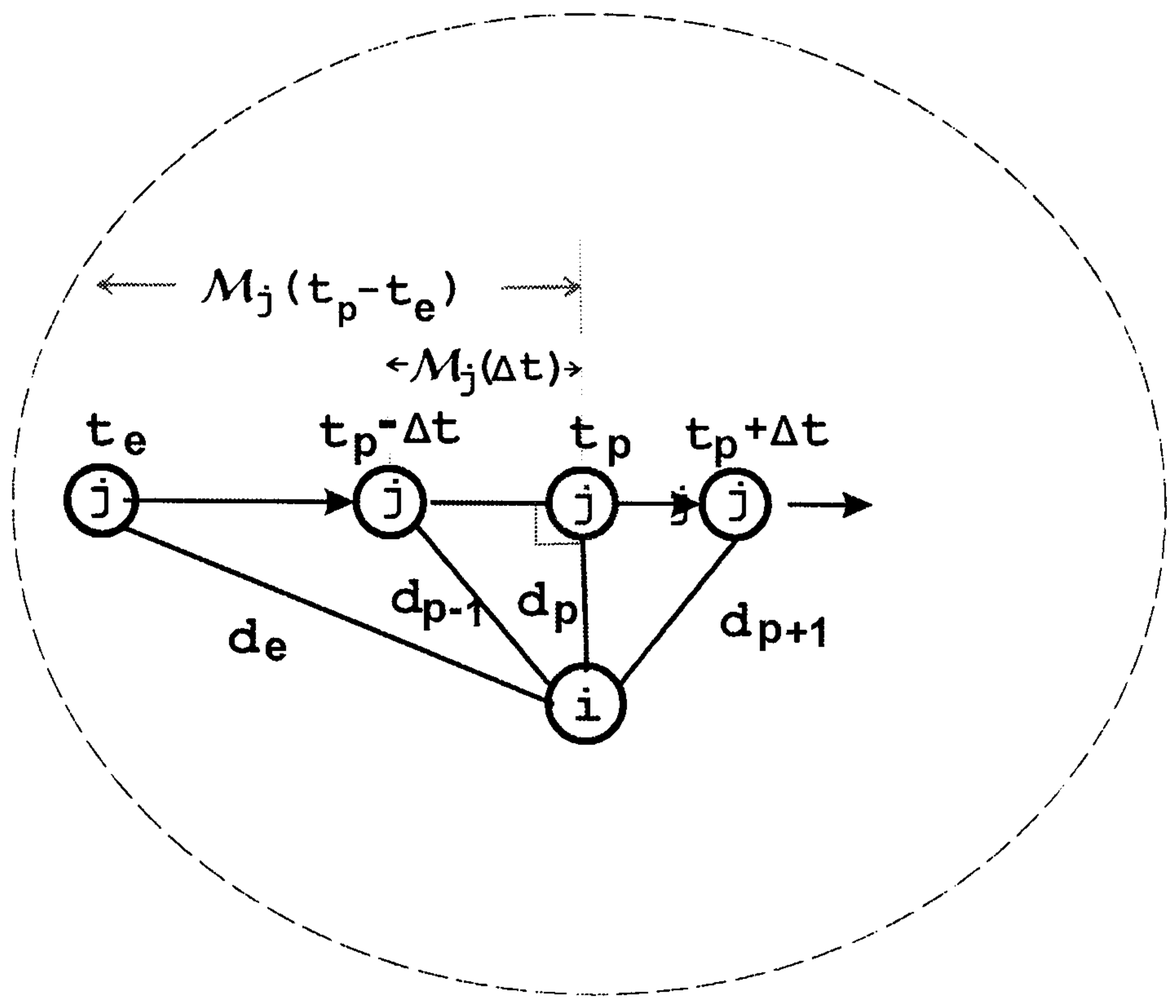


Figure 6.2: Estimation of relative velocity for sparse networks - the *RVEs* approach

where $\varepsilon(\hat{d}_{p-1}, \hat{d}_p)$ is the error component resulted from two noisy range estimates, $\hat{d}_{p-1}$ and $\hat{d}_p$. Its average is denoted as $\bar{\varepsilon}$. $\hat{v}_{ij}$ can be also determined from a longer time scale by:

$$\begin{aligned} \hat{v}_{ij} = \frac{\mathcal{M}_j(t_p - t_e)}{t_p - t_e} &= \frac{\sqrt{\hat{d}_e^2 - \hat{d}_p^2}}{t_p - t_e} \\ &= \sqrt{\frac{d_e^2 - d_p^2}{(t_p - t_e)^2} + \frac{\varepsilon(\hat{d}_e, \hat{d}_p)}{(t_p - t_e)^2}} \end{aligned} \tag{6.3}$$

The two variants of *RVEs* given in Eq. 6.2 and Eq. 6.3 are referred to as $RVEs(\Delta t)$ and $RVEs(t_p - t_e)$, respectively. When the range information is perfect (e.g. $\varepsilon = 0$), $RVEs(\Delta t)$ is preferable to $RVEs(t_p - t_e)$. As the later is vulnerable to the changes in either i's or j's trajectory during time $t_p - t_e$. However, in the general cases when the range information is noisy, $RVEs(t_p - t_e)$ is more accurate than $RVEs(\Delta t)$ as on average $\frac{\bar{\varepsilon}}{t_p - t_e} \leq \frac{\bar{\varepsilon}}{\Delta t}$. The performance of the two *RVEs* approaches will be compared in Section IV.

If not other specified, in the rest of the Chapter the *RVEs* approach is only referred to $RVEs(t_p - t_e)$. With the help of *RVEs* a *departing neighbour* is a *velocity-known* node.

6.3.2 Estimating Relative Velocity in Dense Networks

RVEs is simple and efficient. But it can only provide a neighbouring node's velocity estimate after it becomes a *departing node* and may not suit applications that require velocity estimates of neighbouring nodes as soon as they enter the radio coverage. In fact, when the density of the network increased (e.g. when the node degree is at least 2), a node can quickly estimate the relative velocity of its *arriving nodes* by making use of mobile nodes with known relative velocity (e.g. *departing nodes*) in its one-hop vicinity.

Assuming that among node i's one-hop neighbours there is at least one node (e.g. j) whose velocity has been estimated, i.e. by the *RVEs* approach. The other variant of range-based velocity estimations, *RVEd*, is proposed to exploit nodes with known velocity to estimate the relative velocity of other nodes that are covered by both i's and j's radio range (e.g. k). Fig. 6.3 shows node i's positions and distances relative to its neighbours j and k observed at time $t - \Delta t$ and t. The changes in the distances among these nodes also cause their included angles to vary over time. As these nodes periodically exchange with each other the distance information to their one-hop neighbours, node i is able to calculate these time-varying included angles at every time slot (e.g. Δt). Let $\mathcal{K}_0$ be node i's distance to j measured at time $t - \Delta t$ and $\mathcal{J}_0$ be that to k measured at the same time. According to the *Cosine Rule*, the included angle of $\mathcal{K}_0$ and $\mathcal{J}_0$, φ_i, is given by:

$$\varphi_i = cos^{-1}\frac{\mathcal{K}_0^2 + \mathcal{J}_0^2 - \mathcal{I}_0^2}{2\mathcal{K}_0\mathcal{J}_0} \tag{6.4}$$

Similarly, given the distance estimates $\mathcal{K}_1$ and $\mathcal{J}_1$ to node j and k at time t, respectively, the included angle $\varphi_i^{'}$ can be calculated by:

$$\varphi_i^{'} = cos^{-1}\frac{\mathcal{K}_1^2 + \mathcal{J}_1^2 - \mathcal{I}_1^2}{2\mathcal{K}_1\mathcal{J}_1} \tag{6.5}$$

The change of the angle φ_i to $\varphi_i^{'}$ from time $t - \Delta t$ to t is the result of both node j's and k's

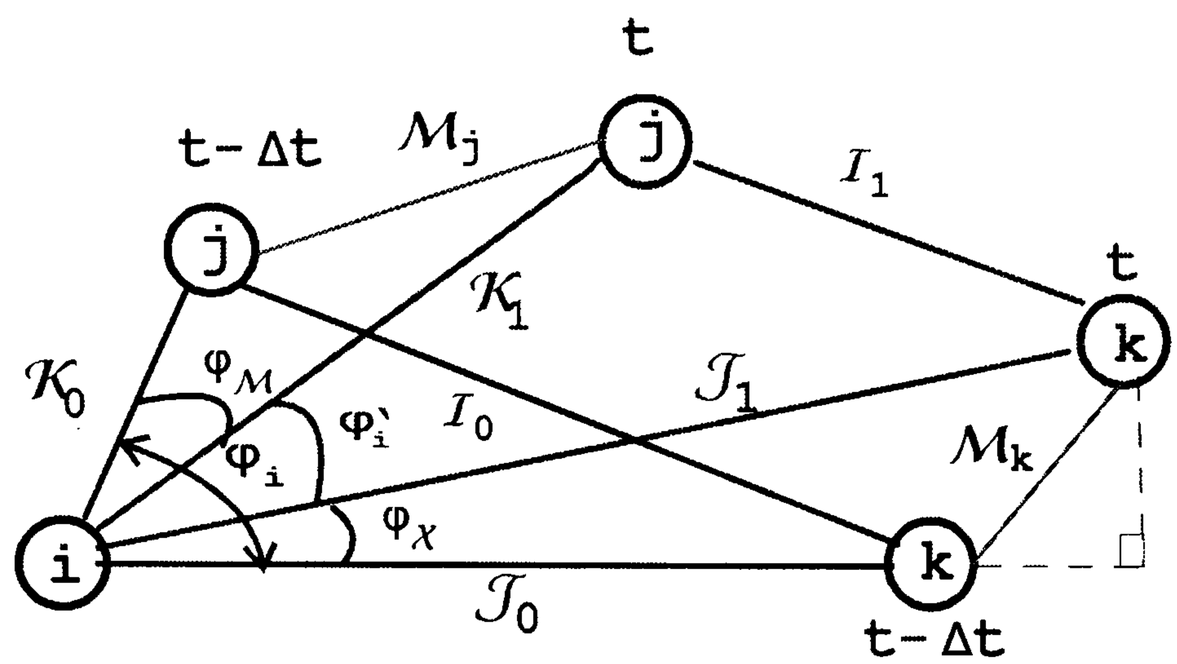


Figure 6.3: Changes in the ranges and the orientations between node i, j and k caused by their relative movements during time slot Δt

movements relative to i during Δt. Assuming that during Δt j and k does not cross each other[‡], there could be 4 possible relative movement patterns among node i, j and k in the time slot (see Fig. 6.4). Let $\varphi_{\mathcal{M}}$ be the included angle of $\mathcal{K}_0$ and $\mathcal{K}_1$, $\varphi_{\mathcal{X}}$ be the included angle of $\mathcal{J}_0$ and $\mathcal{J}_1$. The various relative movement patterns demonstrated in Fig. 6.4 result in four different relationships between φ_i, $\varphi_i^{'}$, $\varphi_{\mathcal{M}}$ and $\varphi_{\mathcal{X}}$, which can be summarized in Eq. 6.6 as:

$$\varphi_{\mathcal{X}} = \begin{cases} |\varphi_i - \varphi_i^{'} - \varphi_{\mathcal{M}}| & \varphi_i \geq \varphi_i^{''} \\ |\varphi_i - \varphi_i^{'} + \varphi_{\mathcal{M}}| & \varphi_i < \varphi_i^{''} \end{cases} \tag{6.6}$$

where $\varphi_i^{''}$ is the included angle of $\mathcal{J}_0$ and $\mathcal{K}_1$. From Eq. 6.6, we could obtain the value of $\varphi_{\mathcal{X}}$ for deriving node k's velocity relative to i during Δt if $\varphi_{\mathcal{M}}$ and $\varphi_i^{''}$ is known. Since the relative speed of node j is known, we can calculate $\varphi_{\mathcal{M}}$ ($\varphi_{\mathcal{M}} \geq 0$) by:

$$\varphi_{\mathcal{M}} = cos^{-1} \frac{\mathcal{K}_0^2 + \mathcal{K}_1^2 - \mathcal{M}_j^2}{2\mathcal{K}_0\mathcal{K}_1} \tag{6.7}$$

where $\mathcal{M}_j$ ($\mathcal{M}_j = \hat{v}_{ij}\Delta t$) is the distance that node j have moved during Δt relative to node i.

[‡]The cases when node j and k cross each other are not discussed, as they rarely happen if the time slot Δt is short (e.g. 1 second). Even if the two nodes do cross each other on some occasions, they can be detected and treated as exceptions.

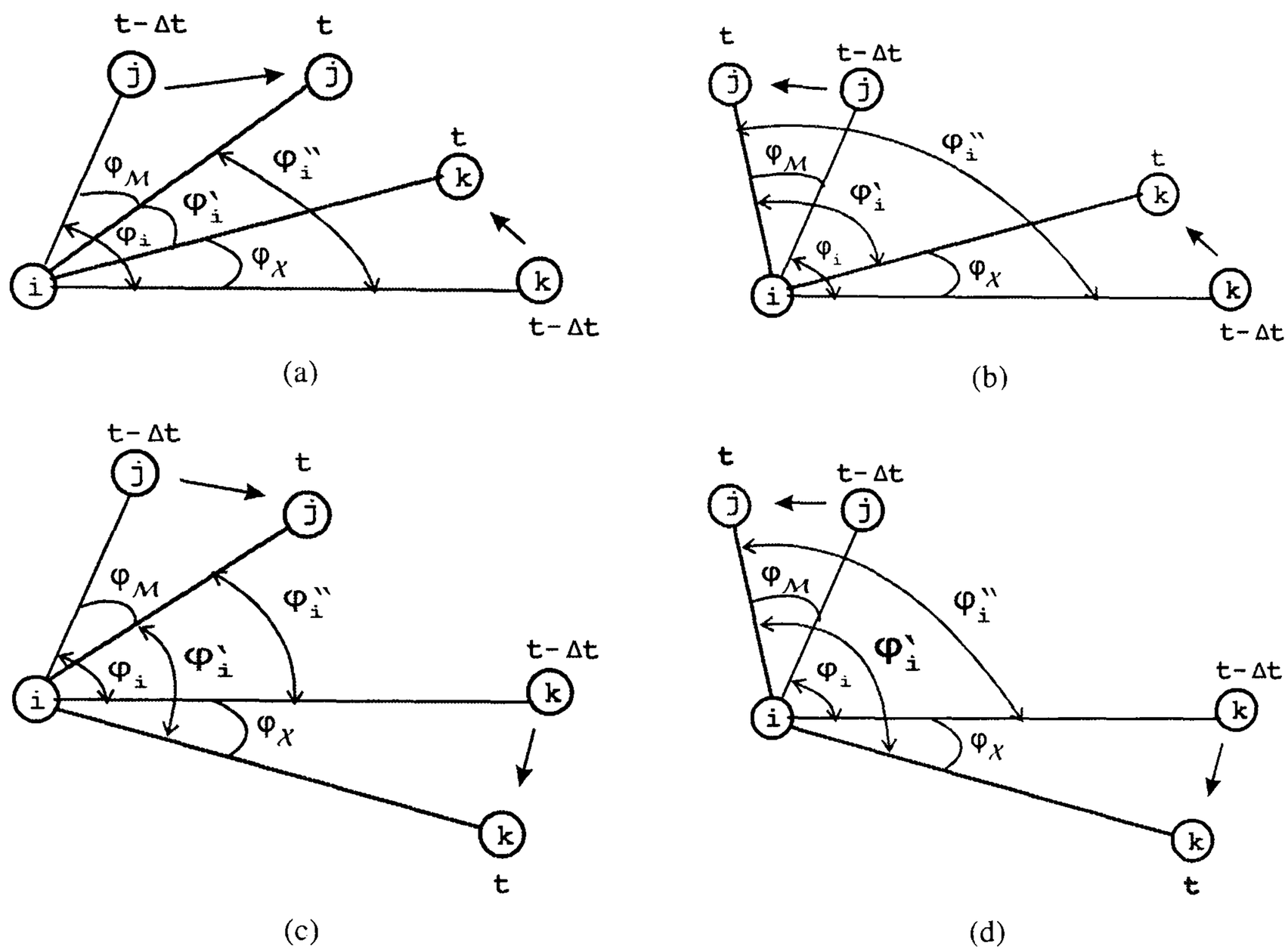


Figure 6.4: Node j and k's possible movement patterns relative to i during time slot Δt results in various relationships between φ_i, $\varphi_i^{'}$, $\varphi_{\mathcal{M}}$ and $\varphi_{\mathcal{X}}$. (a) $\varphi_i = \varphi_i^{'} + \varphi_{\mathcal{M}} + \varphi_{\mathcal{X}}$ ($\varphi_i \geq \varphi_i^{''}$), (b) $\varphi_i = \varphi_i^{'} - \varphi_{\mathcal{M}} + \varphi_{\mathcal{X}}$ ($\varphi_i < \varphi_i^{''}$), (c) $\varphi_i = \varphi_i^{'} + \varphi_{\mathcal{M}} - \varphi_{\mathcal{X}}$ ($\varphi_i \geq \varphi_i^{''}$), (d) $\varphi_i = \varphi_i^{'} - \varphi_{\mathcal{M}} - \varphi_{\mathcal{X}}$ ($\varphi_i < \varphi_i^{''}$).

In noisy environments, the included angles (e.g. φ_i, $\varphi_i^{'}$ and $\varphi_{\mathcal{M}}$) are calculated from distance estimates with measurement errors. As a result, the estimate of $\varphi_{\mathcal{X}}$ (denoted as $\hat{\varphi}_{\mathcal{X}}$) would also contain an error component, which is mathematically expressed as:

$$\begin{aligned}
\hat{\varphi}_{\mathcal{X}} &= \begin{cases} |(\varphi_i + \varepsilon_i) - (\varphi_i^{'} + \varepsilon_i^{'}) - (\varphi_{\mathcal{M}} + \varepsilon_M)| & \varphi_i \geq \varphi_i^{''} \\ |(\varphi_i + \varepsilon_i) - (\varphi_i^{'} + \varepsilon_i^{'}) + (\varphi_{\mathcal{M}} + \varepsilon_M)| & \varphi_i < \varphi_i^{''} \end{cases} \\
&= \begin{cases} |\varphi_i - \varphi_i^{'} - \varphi_{\mathcal{M}}| + |\varepsilon_i - \varepsilon_i^{'} - \varepsilon_M| & \varphi_i \geq \varphi_i^{''} \\ |\varphi_i - \varphi_i^{'} + \varphi_{\mathcal{M}}| + |\varepsilon_i - \varepsilon_i^{'} + \varepsilon_M| & \varphi_i < \varphi_i^{''} \end{cases}
\end{aligned} \tag{6.8}$$

where ε_i, $\varepsilon_i^{'}$ and ε_M are the error components of φ_i, $\varphi_i^{'}$ and $\varphi_{\mathcal{M}}$, respectively. Eq. 6.8 indicates

that the accuracy of $\hat{\varphi}_{\mathcal{X}}$ is largely determined by ε_M if the distribution of ε_i is similar to that of $\varepsilon_i^{'}$.

It is very difficult to determine $\varphi_i^{''}$ as its opposite side connecting k at time $t - \Delta t$ and j at time t is not measurable. Without knowing $\varphi_i^{''}$, we will have two estimates of $\varphi_{\mathcal{X}}$ from Eq. 6.6 with one of them being the true estimate. The true estimate can be identified from two sets of values of $\hat{\varphi}_{\mathcal{X}}$. The second set can be produced by using *RVEd* with another velocity-known node or with node j at Δt time later[§]. Let $\varphi_{\mathcal{X}}^A(\varphi_0^a, \varphi_1^a)$ and $\varphi_{\mathcal{X}}^B(\varphi_0^b, \varphi_1^b)$ be the two sets of values produced by the *RVEd* approach with two velocity-known nodes or two time slots. The true estimate should be the one possessing same (or similar) values in both sets. A mathematical expression of this is:

$$\hat{\varphi}_{\mathcal{X}} = \varphi_{\mathcal{X}}^A \bigcap \varphi_{\mathcal{X}}^B \tag{6.9}$$

Therefore, we can calculate k's movement relative to node i during time slot Δt, $\mathcal{M}_k$ by:

$$\mathcal{M}_k = \sqrt{(\mathcal{J}_1 cos\varphi_{\mathcal{X}} - \mathcal{J}_0)^2 + (\mathcal{J}_1 sin\varphi_{\mathcal{X}})^2} \tag{6.10}$$

and the estimated relative velocity $\hat{v}_{ik}$ by:

$$\hat{v}_{ik} = \frac{\mathcal{M}_k}{\Delta t} \tag{6.11}$$

6.4 Performance Evaluation

6.4.1 Simulation Settings

To evaluate the performance of the proposed method of velocity estimations (i.e. RVEs and RVEd), the ns-2 simulator is deployed to simulate sparse and dense topologies in two experiments, which covers different levels of mobility and environmental noises. The simulated mobile ad hoc network consists of $\mathcal{N} = 40$ nodes moving within in a square area with the size

[§]Supposing that node k does not change its velocity during the past $2\Delta t$ time.

of $\mathcal{L} \times \mathcal{L}$. $\mathcal{L}$ is set as 1400m to make networks of sparse topology and 800m for those of dense topology. For both type of networks, the radio radius of a node is fixed at 150m. Using the mobility modelling tool, Bonnmotion [99], the simulations cover a wide range of nodal mobility characterized by the Random Waypoint Model (RWP) [11]. For the RWP model, the pause time is kept at 0 to produce continuous movement. The movement speed of a mobile node is uniformly distributed over $[v_{min}, v_{max}]$, where v_{min} is fixed at 3.5m/s and v_{max} is limited in the range between 5m/s and 40m/s.

A mobile node measures the distance to its one-hop neighbours and exchanges this information with all of its neighbouring nodes at intervals of $\Delta t = 1$ second. The initial status of signal propagation is LOS ($\zeta_N = 0$). The estimates of inter-node distance have multipath errors that can be modeled by a zero-mean Gaussian distribution with variance σ_M^2 or a uniform distribution over the range of $[-\sigma_M, \sigma_M]$. The measurement errors caused by NLOS conditions are modeled by a positive-mean Gaussian distribution with mean μ_N and variance σ_N^2, or an uniform distribution over the range of $(0, \mu_N]$. Based on the empirical data provided in [100] [109], the maximum value of σ_M and μ_N is set to be 4m and 32m, respectively. σ_N is fixed at 3m in this thesis.

Having collected all the distance estimates within its one-hop vicinity, each node carries out the range-based relative velocity estimation for each of its passing-by neighbours for a duration of 500 seconds starting after 1250 seconds of warm-up period to avoid speed decay [54]. In all figures generated from the simulation, each data point corresponds to a mean of 30 repeated experiments with different random seeds.

6.4.2 Experiment 1: Sparse Network case

Experiment 1 is designed to test the performance of the proposed RVEs method and that of an existing velocity estimation method (e.g. the Link Prediction Algorithm *LPA* [66]) in networks where mobile nodes are sparsely distributed.

Given perfect range information, the estimates of inter-node relative velocity produced by

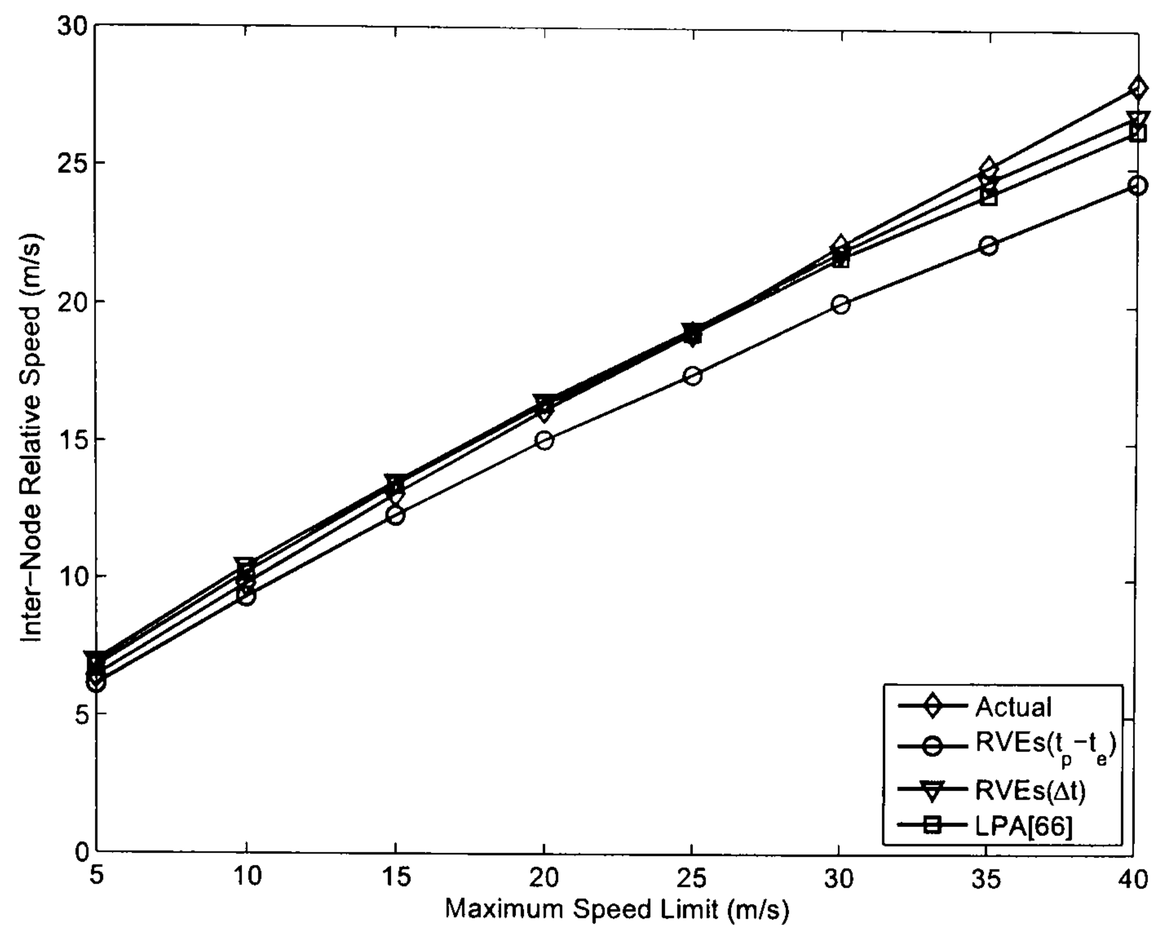


Figure 6.5: Sparse Network Case: Inter-node Relative Speed Vs. Maximum Speed Limit (noise free)

the *LPA*, $RVEs(t_p - t_e)$ and $RVEs(\Delta t)$ approaches are compared to the true values in Fig. 6.5. We can see from Fig. 6.5 that the estimates provided by these methods are very close to the true values over a wide range of nodal mobility scenarios. The accuracy of $RVEs(t_p - t_e)$ is slightly lower than other methods as it assumes that the trajectory of one node within the neighbouring node's radio coverage is a straight line. When the range information become noisy, the advantage of $RVEs(t_p - t_e)$ becomes obvious. Fig. 6.6 shows the performance of these three methods with noisy range estimates with Gaussian multipath errors ($\sigma_M = 2m$). From Fig. 6.6 we can observe that the *LPA* method is significantly affected by the noises and its accuracy drops to far below the two *RVEs* approaches. The accuracy of the $RVEs(t_p - t_e)$ method is now better than $RVEs(\Delta t)$. As discussed in Section II, this is due to that $RVEs(t_p - t_e)$ chooses to estimate the relative velocity in a longer period which helps to further reduce the impact of the erroneous distance information.

The performance of the *LPA* and *RVEs* (i.e. $RVEs(t_p - t_e)$) methods are further examined in LOS conditions for zero-mean Gaussian noises of variance σ_M and Uniform errors of range $[-\sigma_M, \sigma_M]$ in Fig. 6.7(a) and Fig. 6.7(b), respectively. In Fig. 6.7, the normalized bias between

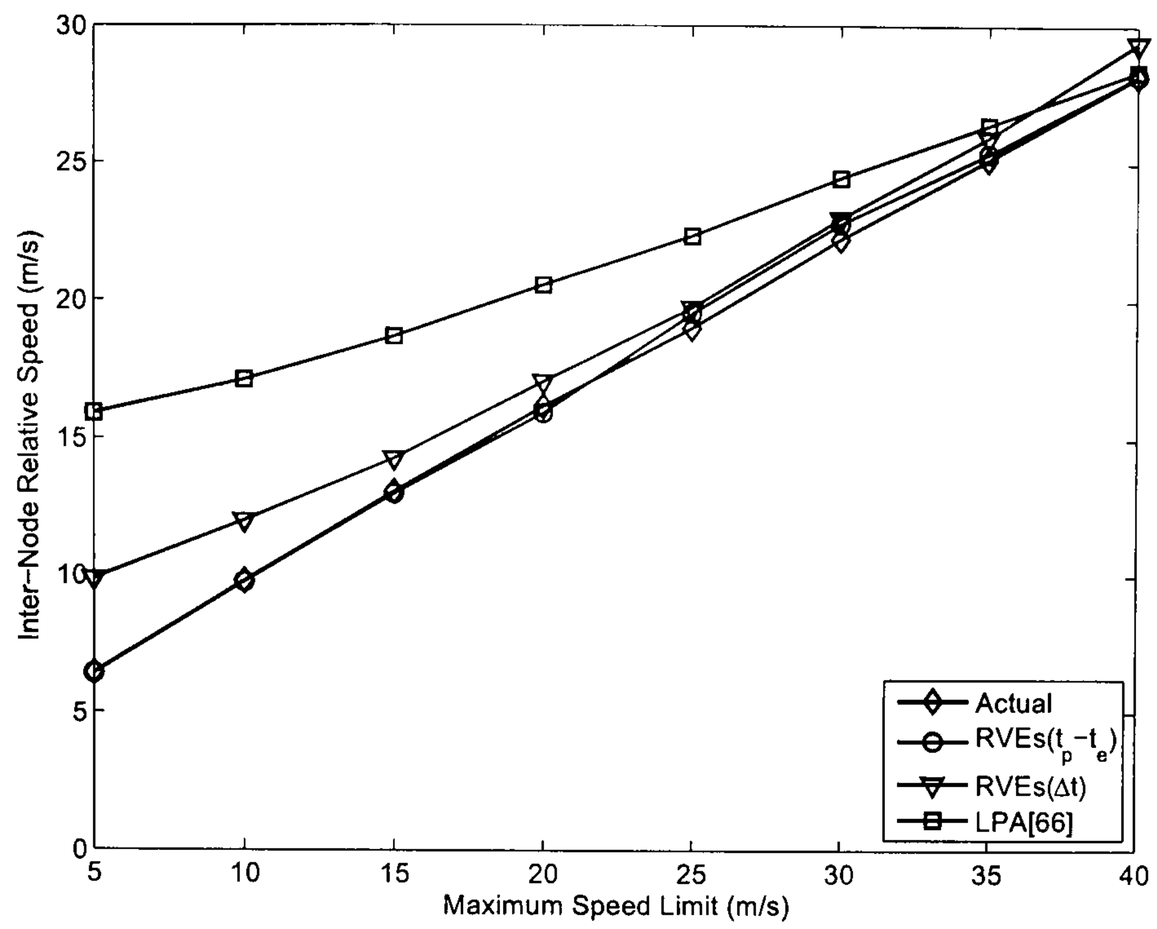


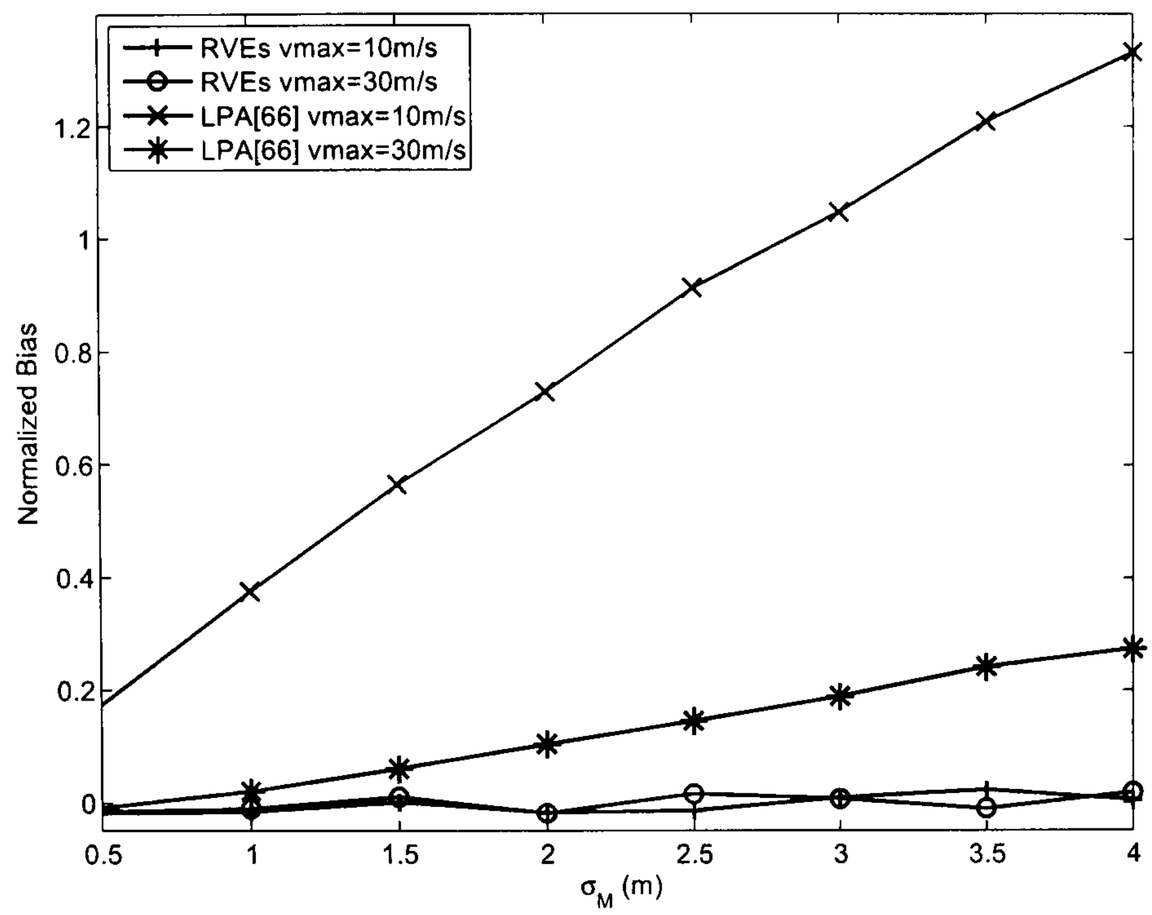
Figure 6.6: Sparse Network Case: Inter-node Relative Speed Vs. Maximum Speed Limit, LOS state (p=0, q=0), Gaussian multipath error model ($\sigma_\epsilon = 2m$).

their estimates of relative velocity and the true values are obtained with distance information measured from a range of noise levels. The Normalized Bias of the speed Estimates (NBE) is given by:

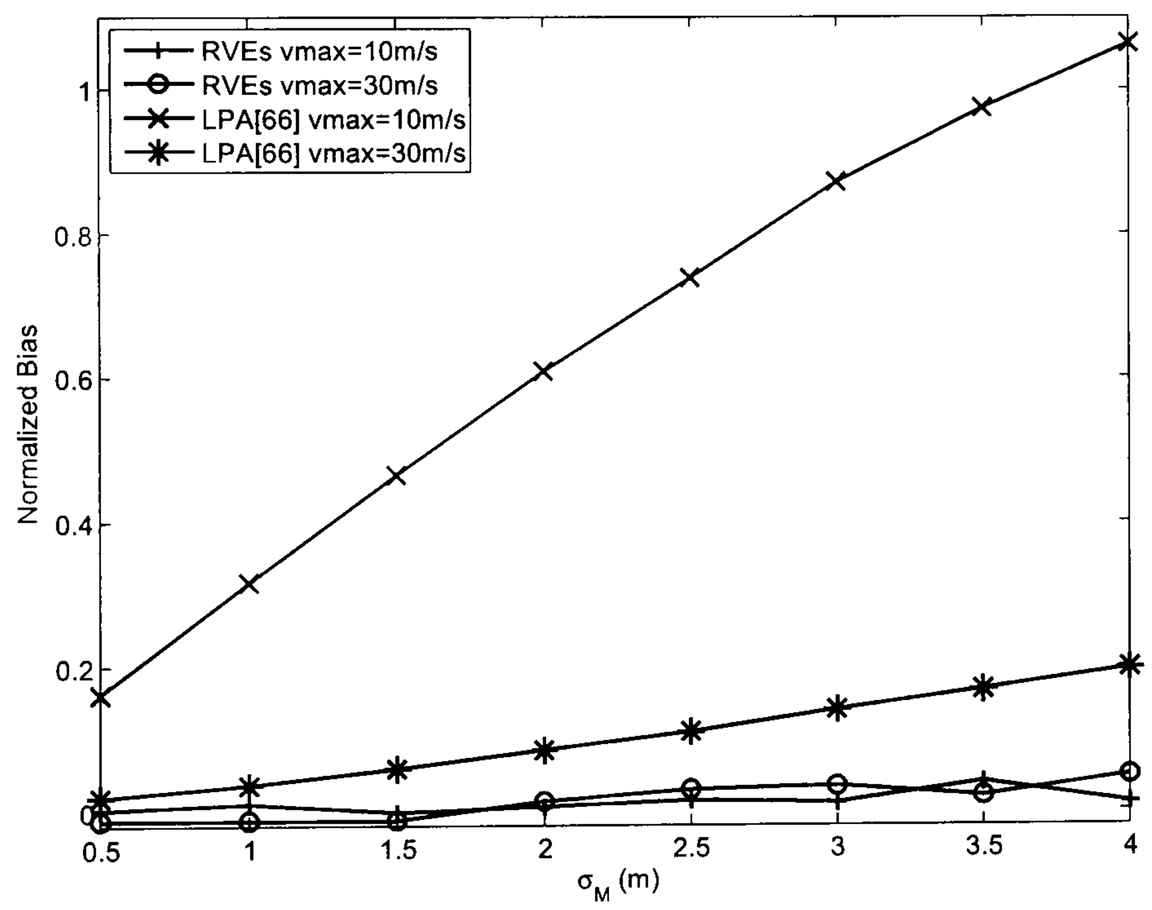
$$NBE = \frac{E[\hat{v}]}{\bar{v}} - 1 \tag{6.12}$$

where $\bar{v}$ is the mean of the true values of relative velocities measured from simulation and $E[\hat{v}]$ is the mean of the velocity estimates. NBE indicates how close the velocity estimates are to the true values. As demonstrated in both Fig. 6.7(a) and Fig. 6.7(b), the *RVEs* approach once again demonstrates its robustness against various level of noises (σ_M is in the range between 0.5m to 4m) for both low ($v_{max} = 10m/s$) and high nodal mobility levels ($v_{max} = 30m/s$). The NBE of the *RVEs* method from the true speed value is kept within the range of [-0.02, 0.04] in the scenarios regardless of the noise level or the maximum speed limit. On the other hand, Fig. 6.7 shows that the NBE of the estimates provided by *LPA* is increasing with σ_M with either of the two noise distributions. When σ_M is 4m and v_{max} is 10m/s, the NBE of *LPA* reaches a value of about 1.3.

Similar results can be observed in Fig. 6.8, in which the normalized bias of *RVEs* and *LPA*



(a)



(b)

Figure 6.7: Sparse Network Case: Normalized bias of the range-based velocity estimator Vs. Noise Variances, LOS state (p=0, q=0), (a) Gaussian multipath error model, (b) Uniform multipath error model.

in NLOS conditions is plotted. The NLOS noises are modeled by an Uniform distribution over the range of (0, μ_N]. The value of μ_N is increased from 4m to 32m at a step of 4m. In these NLOS settings, as shown in Fig. 6.8, the NBE of *RVEs* is limited in the range of [0.02, 0.17] whereas that of *LPA* can be as big as 2.9.

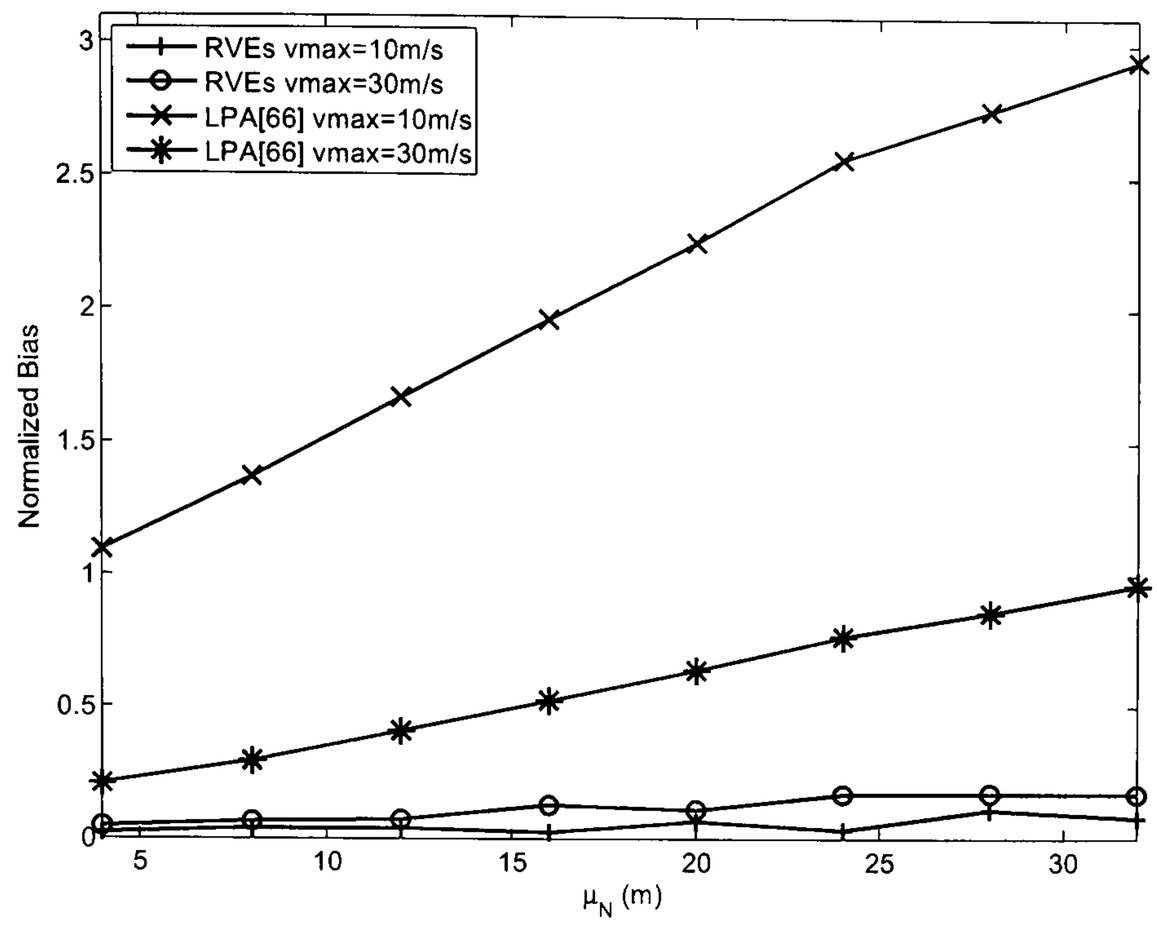


Figure 6.8: Sparse Network Case: Normalized bias of the range-based velocity estimator Vs. Noise Variances, NLOS state (p=1, q=1), Gaussian multipath error model ($\sigma_\epsilon = 2m$), Uniform NLOS error model.

From Fig. 6.6-6.8, it seems that the *LPA* method has better performance in scenarios of higher network mobility. In fact, this is only because the true velocity has higher values in scenarios of faster mobility, which makes the gaps between the true values and the estimates provided by *LPA* appear smaller.

6.4.3 Experiment 2: Dense Network case

Experiment 2 is designed to test the performance of the proposed RVEd method and that of the *LPA* algorithm in networks where mobile nodes are densely distributed.

Fig. 6.9 shows the true velocity values, the corresponding estimates produced by the *RVEd* and the *LPA* method for v_{max} between 5m/s and 40m/s. We can see from Fig. 6.9 that, similar to the sparse network case, these methods yield excellent results in the noisy-free situation and the advantage of *RVEd* over *LPA* is not obvious. Given Gaussian multipath errors with 2m of variance in LOS conditions, Fig. 6.10 shows that the *RVEd* method distinguishes itself from the *LPA* method with accurate estimates while *LPA* shows significantly decreased accuracy. The

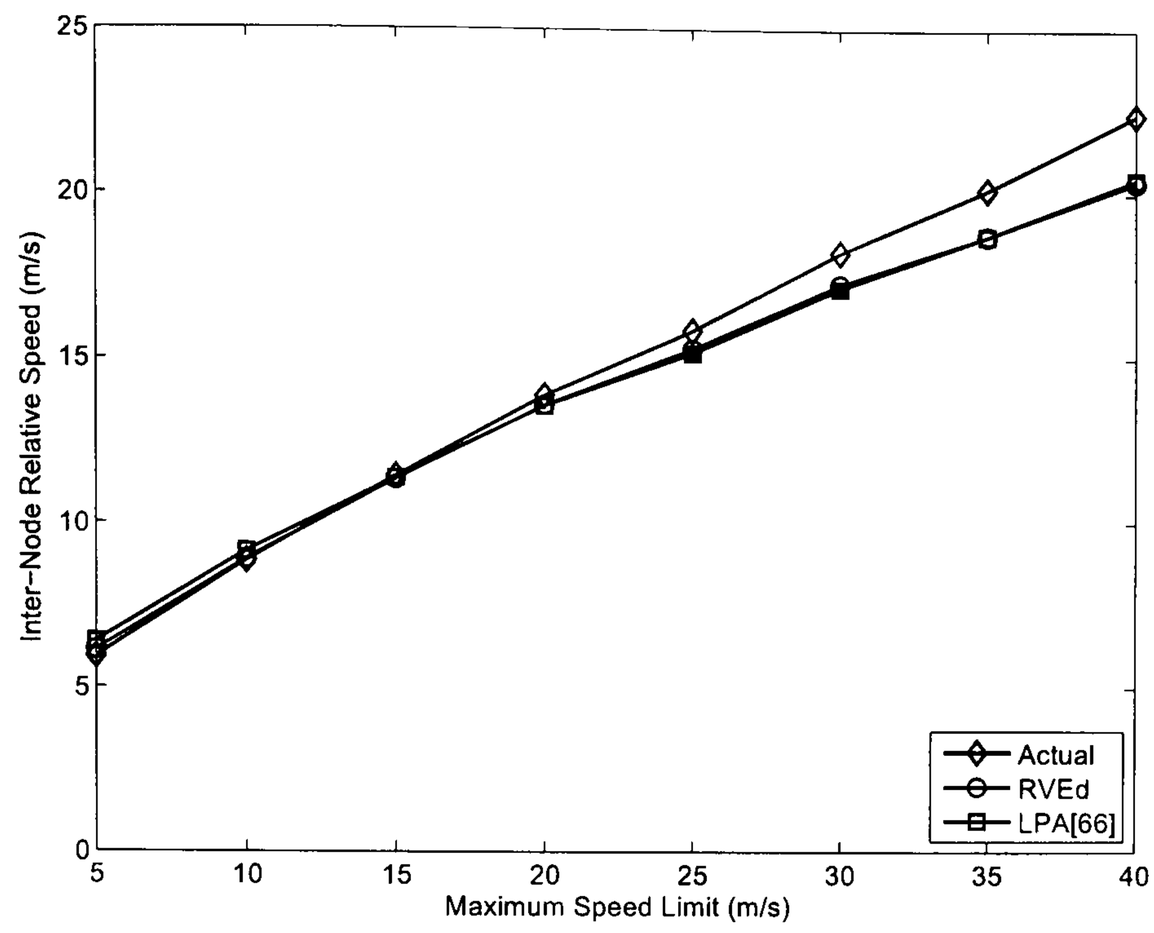


Figure 6.9: Dense Network Case: Inter-node Relative Speed Vs. Maximum Speed Limit (noise free)

NBE of the estimates of the *LPA* and *RVEd* methods are shown in Fig. 6.11(a) for zero-mean Gaussian noises with variance σ_M and Fig. 6.11(b) for Uniform errors distributed between $(-\sigma_M, \sigma_M)$. The value of σ_M is increased from 0.5m to 4m at a step of 0.5m. As shown in Fig. 6.11, regardless of the error distributions, the NBE of the *RVEd* approach is within the range of [-0.06, 0.06] at either of the two speed limits, while that of *LPA* can be up to 2 (an improvement factor of up to 33).

Fig. 6.12 shows the performance of the *RVEd* and *LPA* methods with NLOS errors that are uniformly distributed over (0m, 32m] for the dense network case. The NBE of *RVEd* is shown to fall into the range of [-0.01, 0.2] regardless of the true nodal speed. The corresponding NBE range of the *LPA* method is between 0.5 and 4.0, which indicates an significant improvement factor of up to 20. The results given in Fig. 6.11-6.12 are consistent with those of Fig. 6.7-6.8 for the sparse network case in that the *LPA* method is sensitive to both the level of noise variances and nodal mobility while the proposed *RVEd* method remains stable and exceptionally accurate regardless of the type of noises and the level of nodal mobility.

The parameters (e.g. p and q) of the 2-state Markov model of channel conditions (e.g. LOS

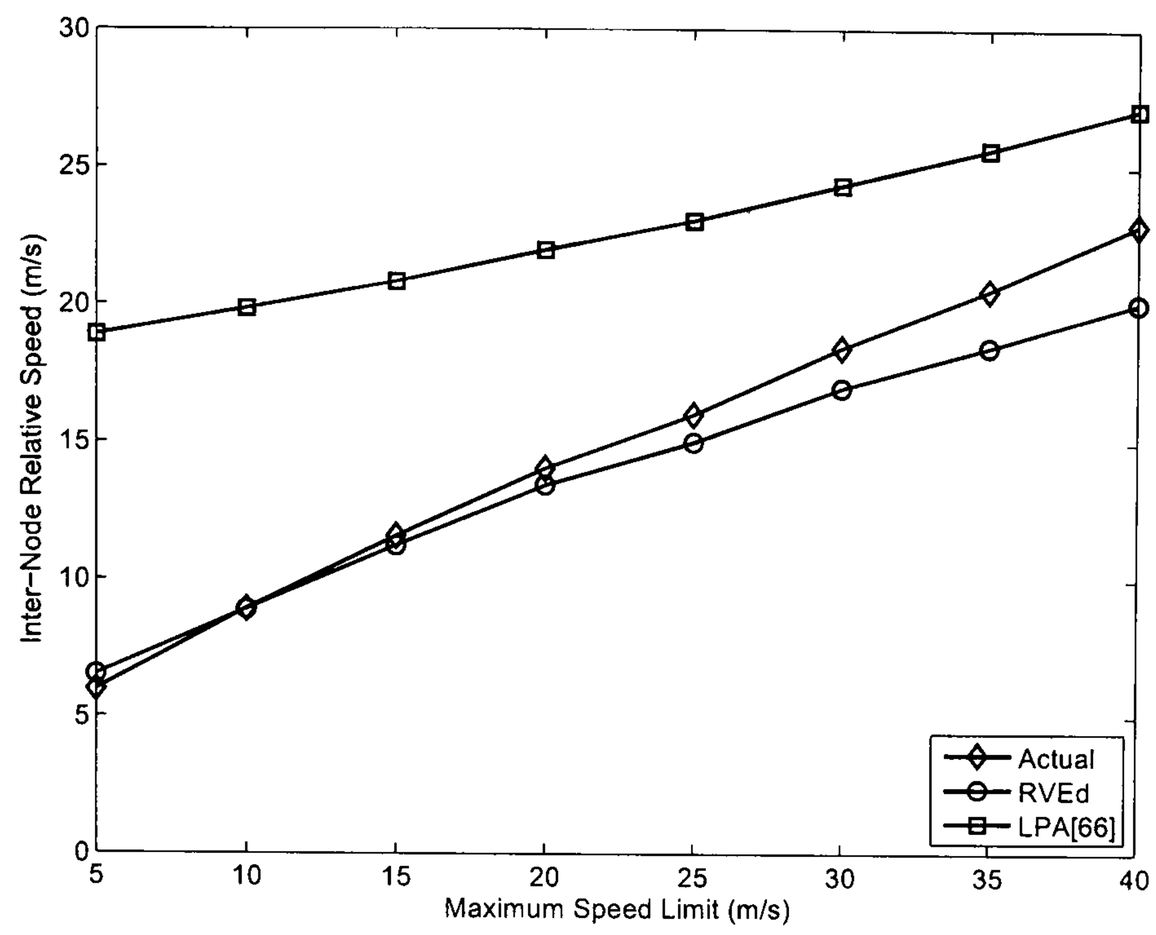
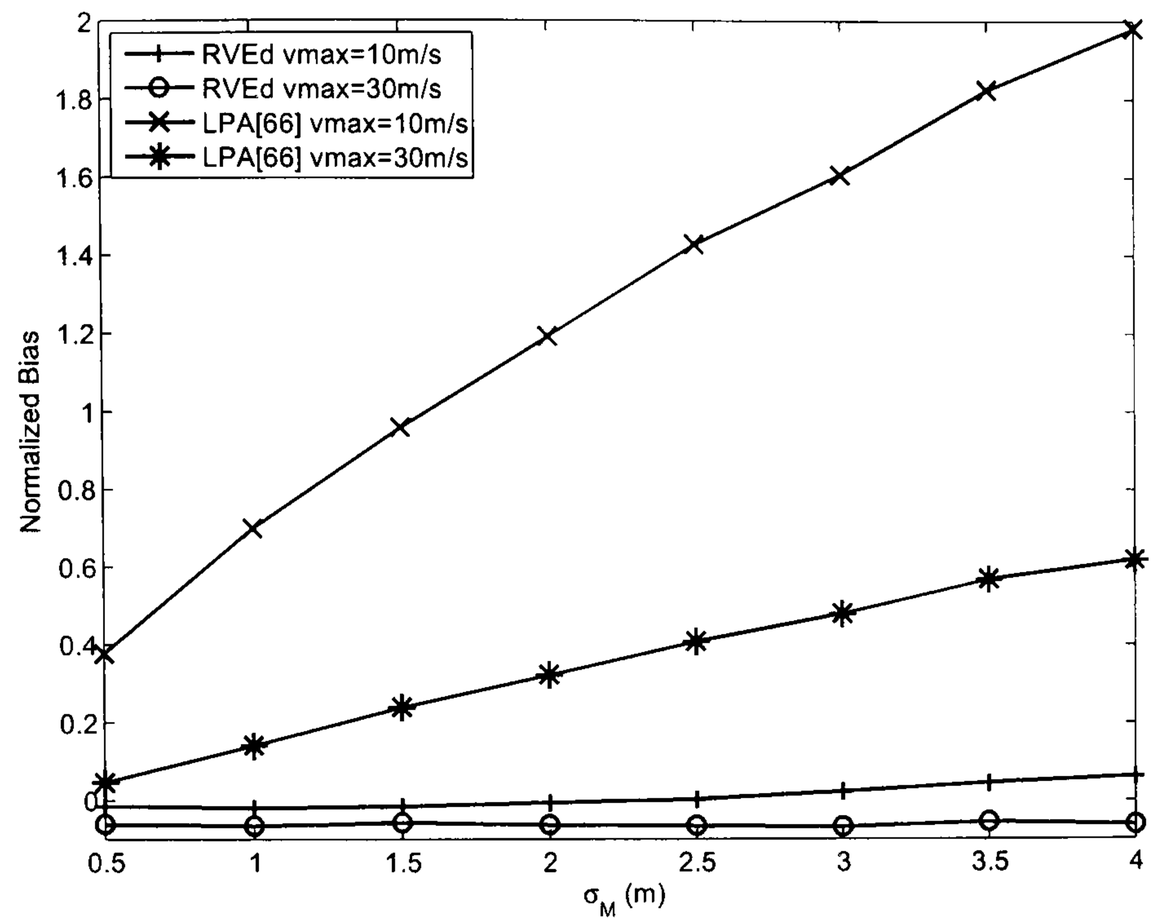
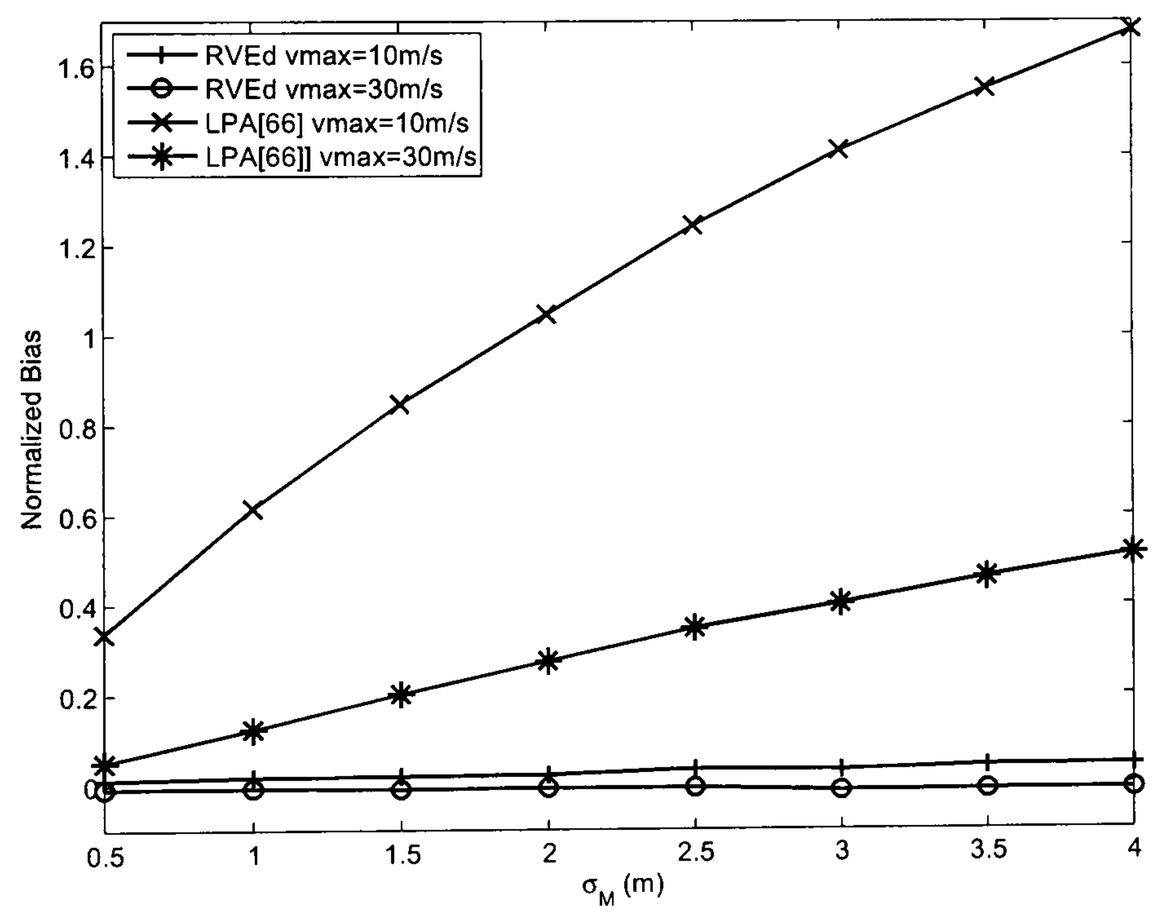


Figure 6.10: Dense Network Case: Inter-node Relative Speed Vs. Maximum Speed Limit, LOS state (p=0, q=0), Gaussian multipath error model ($\sigma_\epsilon = 2m$)

and NLOS) are varied and the results are plotted in Fig. 6.13. The probability of jumping into NLOS conditions from LOS communications or that of remaining in NLOS depend on the geographical features of the environment (e.g. the density of buildings, the arrangement of offices and pedestrian crossings, etc). The performance of the *RVEd* method in frequent NLOS scenarios (p=0.6) and that in infrequent NLOS cases (p=0.3) are evaluated with the probability of remaining in NLOS status, q, being varied from 0.2 to 0.9. The distribution of NLOS errors in these tests is considered as Gaussian with a mean of 15m and a variance of 3m. As can be observed in Fig. 6.13, the parameter p in fact does not have much impact of the NBE of the *RVEd* method. However, an increase in the parameter q (or the time-correlation of NLOS conditions) does help to improve the accuracy of *RVEd*. This is because that, as the *RVEd* method is more susceptible to the deviations in the distance measurement noises of the two time slots involved than the absolute values of the measurement errors, it would achieve a better accuracy with a higher q, under which consecutive time slots are more likely to experience the same environmental conditions and the measurement errors intend to have similar mean values. Note that *RVEd* appears to be more accurate with higher nodal speed (e.g. $v_{max} = 30m/s$).



(a)



(b)

Figure 6.11: Dense Network Case: Normalized bias of the range-based velocity estimator Vs. Noise Variances, LOS state (p=0, q=0), (a) Gaussian error model, (b) Uniform error model.

This is again because of the smaller gaps between the true values and the estimates in higher mobility.

The effects of different ranging intervals Δt (i.e. 1s, 3s and 5s) on the performance of the *RVEd* estimator is shown in Fig. 6.14 for LOS scenarios. In the scenarios the measurement

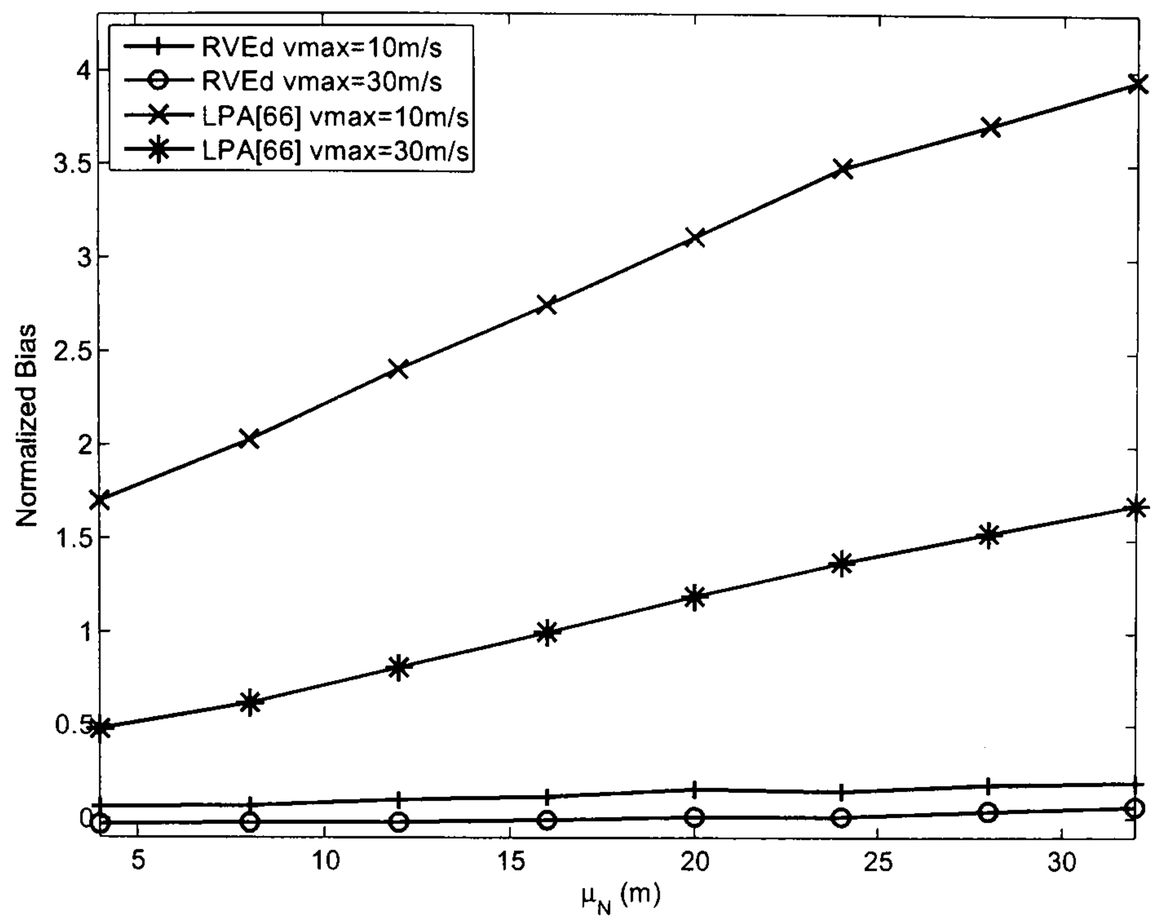


Figure 6.12: Dense Network Case: Normalized bias of the range-based velocity estimator Vs. Noise Variances, NLOS state (p=1, q=1), Gaussian multipath error model ($\sigma_\epsilon = 2m$), Uniform NLOS error model.

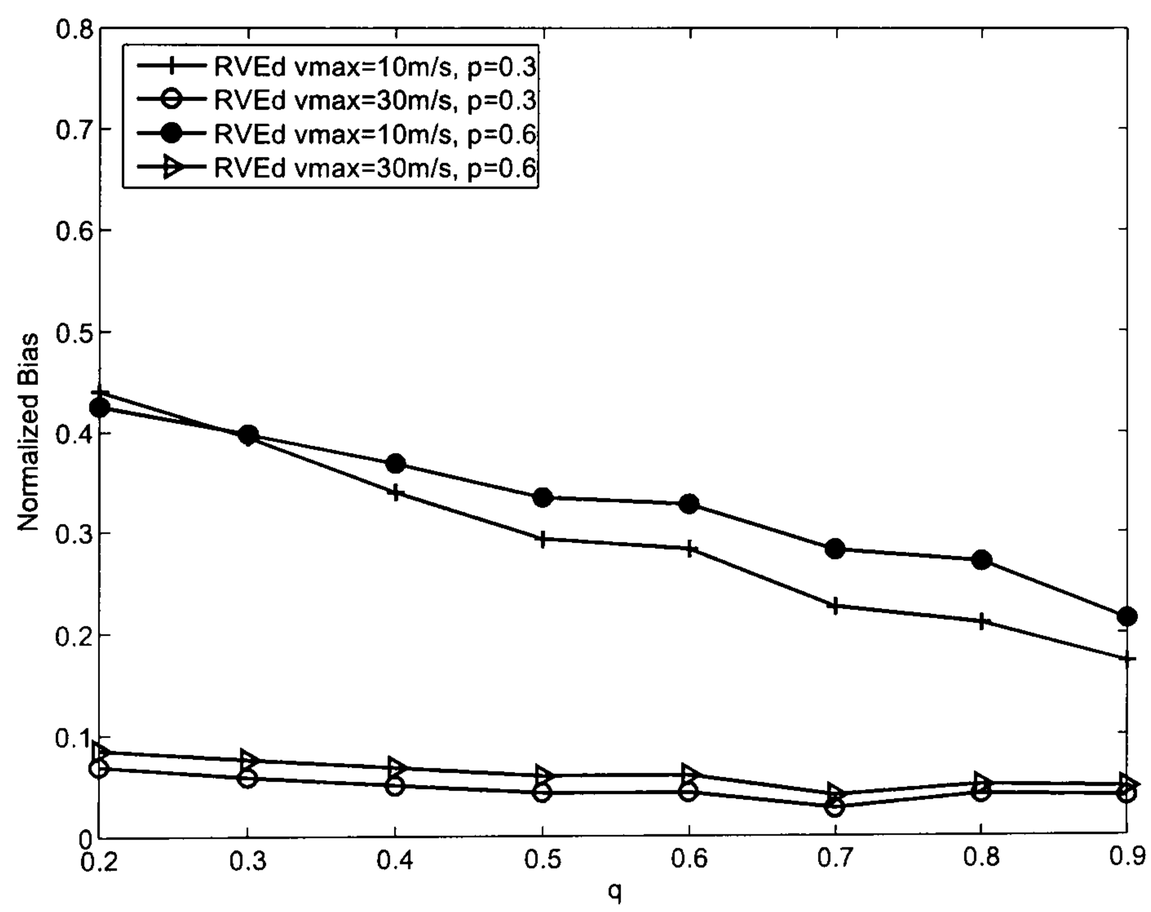


Figure 6.13: Dense Network Case: Normalized bias of the range-based velocity estimator Vs. the probability of turning into NLOS state (q), Gaussian multipath error model ($\sigma_\epsilon = 2m$), Gaussian NLOS error model ($\sigma_N = 3m$).

errors are modeled as zero-mean Gaussian with $4m^2$ of variance and the nodal mobility is constrained by speed limits ranging from 5m/s to 40m/s. As shown in Fig. 6.14, when mobile nodes move at a low speed (e.g. $v_{max} < 10m/s$), varying the value of Δt does not have

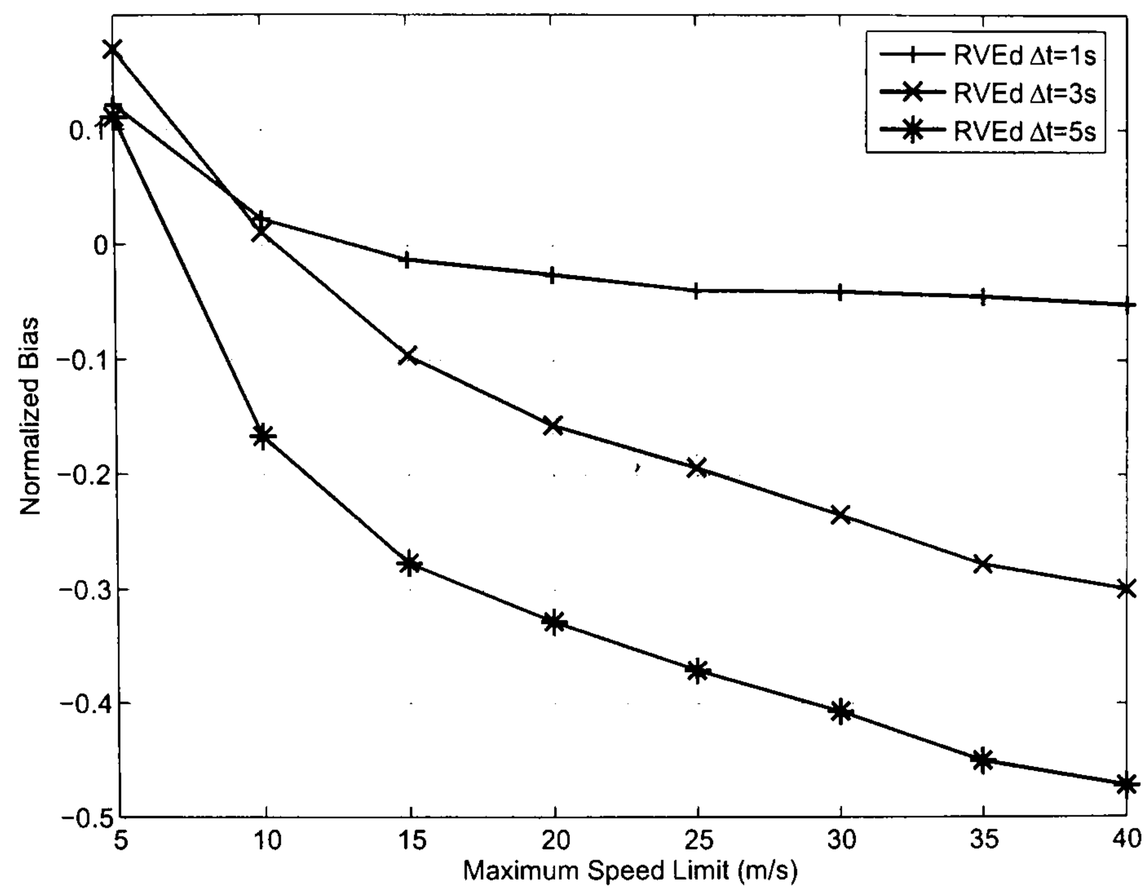


Figure 6.14: Dense Network Case: the effect of the ranging interval Δt on the accuracy of the range-based velocity estimator, Gaussian multipath error model ($\sigma_\epsilon = 2m$).

obvious affects on the performance of *RVEd*. However, when the nodes are in fast moving ($v_{max} \geq 15m/s$), an increase in Δt has a significant impact on the accuracy of the *RVEd* estimator. For instance, when v_{max} is at 40m/s, a 4-second increase in Δt (e.g. from 1s to 5s) causes the normalized bias to rise from -0.05 to -0.47. This is because in high mobility scenarios mobile nodes are able to change their speed/orientation in a very short time slot, which can not be captured by the *RVE* estimators if the ranging interval Δt is relatively too long.

6.5 Estimation of the Expected Node Degree

As one of the two proposed methods (i.e. *RVEd*) works best in relatively dense networks (i.e. with node degree of 2 or over), it is important to verify the expected node degree of a given network to determine the applicability of either of the two methods as well as the achievable resolution of velocity estimates. Fig. 6.15 shows a schematic block diagram of the communications between an application, the velocity estimators and a node degree estimator. As shown in Fig. 6.15, the applicability of either the *RVEs* or *RVEd* method can be determined

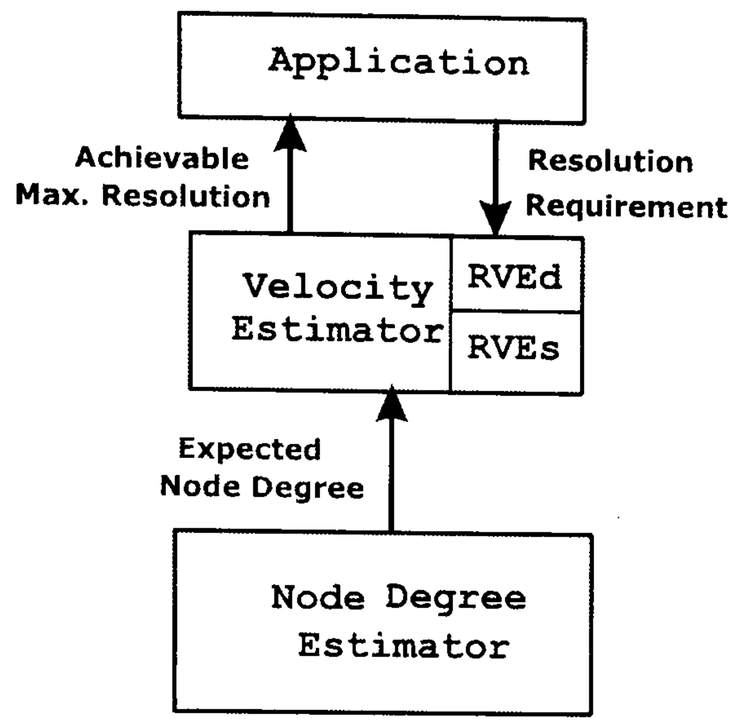


Figure 6.15: Communications between an application, the velocity estimators and a node degree estimator

based on the prediction made by a node degree estimator. Based on the mean link duration and the type of *RVE* method being used, the information on the achievable maximum resolution of velocity estimates can be also assessed and provided to the application. Given this information, the application could propose a realistic resolution requirement for to the velocity estimator.

Recent studies (e.g. [112] and [113]) have proposed statistical methods to estimate the expected node degree of a network given the number of nodes, the radio radius and the size of the network area. However, the proposed methods in these studies can give an accurate prediction of the expected node degree only for networks in which the stationary distribution of nodal positions is uniform. A method of node degree prediction for networks with non-uniform node distributions can be found in [53]. This method is very complex and its simplified approximation given in [53] only accounts for a certian type of networks, e.g. the radio range normalized against the radius of the circular network ground is not bigger than 0.3. In this section, a new approach for predicting the expected node degree is presented. The approach departs from the current methods in that it has a simple and unified form for both networks with uniform and non-uniform node distributions.

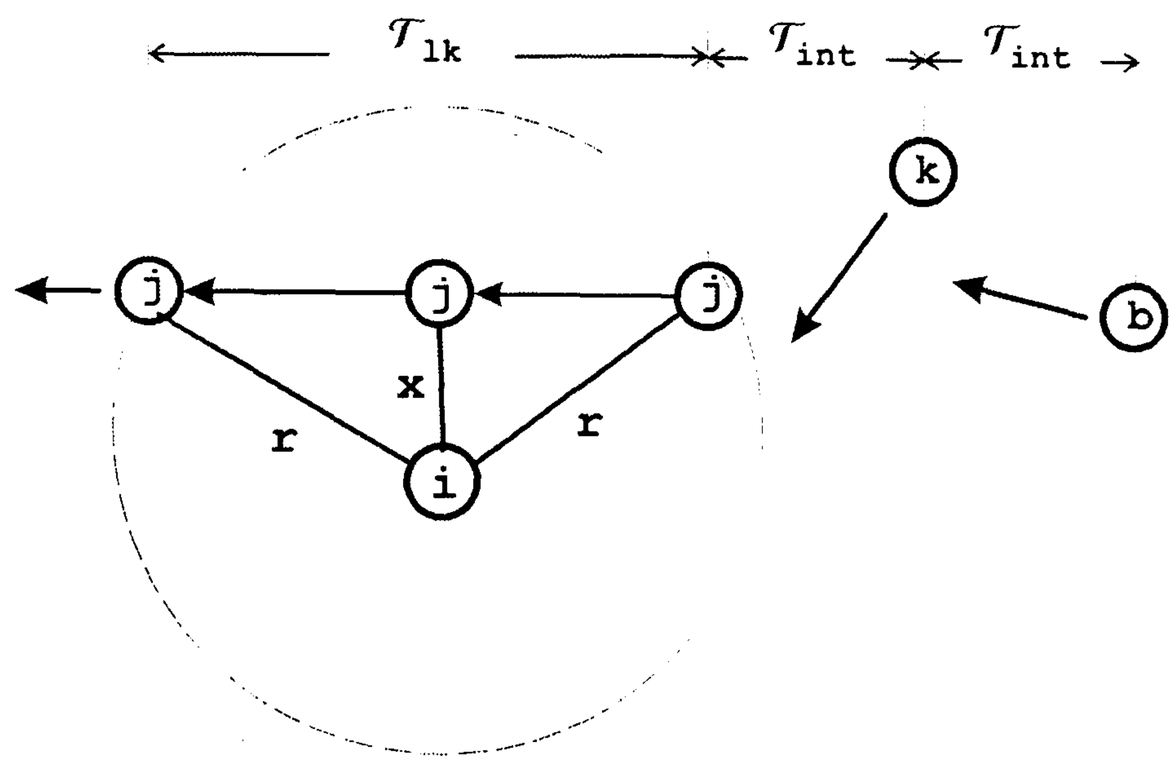


Figure 6.16: link duration and intermeeting times

6.5.1 The Proposed Approach for Estimating the Expected Node Degree

Fig. 6.16 shows a mobile ad hoc network where the movements of nodes are Identical and Independent Distributed (IID) processes and can be characterized as a certain mobility model, which may cause the stationary distribution of the nodal positions to be uniform (e.g. the Random Direction (RD) model) or non-uniform (e.g. the Random WayPoint (RWP) model) [53]. In fact, in such a network with an arbitrary mobility model a single mobile node and its one-hop neighbours can be seen as a simple queueing system, in which the arriving of neighbours at the node's radio coverage can be modeled as a poisson process with expected intervals $E[\mathcal{T}_{int}]$ or *arrival rate* $\bar{\lambda} = \frac{1}{E[\mathcal{T}_{int}]}$. In this queueing system, the average time that a node's neighbours spent in the "queue" is the expected link duration and the average number of neighbours stay in the "queue" is the expected node degree.

Let $\mathcal{T}_{lk}$ be the link duration, its expected value, $E[\mathcal{T}_{lk}]$ can be estimated as:

$$E[\mathcal{T}_{lk}] = \frac{2\sqrt{r^2 - E[x]^2}}{\bar{v}} = \frac{\sqrt{3}r}{\bar{v}} \tag{6.13}$$

where $\bar{v}$ is the expected relative moving speed between two mobile nodes. x is the shortest distance a node can get to its neighbours, which is a random variable uniformly distributed over $[0, r]$. We have its expected value $E[x] = \frac{r}{2}$. Eq. 6.13 is a rough approximation as

it ignores the fact that mobile nodes may change their trajectory and speed during the link duration. However, as the experimental results will show, the link duration estimated by Eq. 6.13 is accurate enough to serve our purpose, i.e. to check whether the network is dense enough for using *RVEd*.

Groenevelt in [86] has proposed analytical methods to predict the expected arriving rate $\bar{\lambda}_s$ of a specific neighbour in both the RD model and the RWP model. We have the expected arriving intervals of any of the $\mathcal{N}-1$ neighbours $\bar{\lambda} = (\mathcal{N}-1)\bar{\lambda}_s$ Therefore, assuming the mobile nodes are continuously moving, the expected arrival rate of an arbitrary neighbour can be derived from [86] for the RD model as:

$$\bar{\lambda} \approx \frac{2r\bar{v}(\mathcal{N}-1)}{\mathcal{L}^2} \tag{6.14}$$

and for the RWP model as:

$$\bar{\lambda} \approx \frac{2\omega r\bar{v}(\mathcal{N}-1)}{\mathcal{L}^2} \tag{6.15}$$

where $\omega \approx 1.3683$ is a constant specific to the RWP model.

Based on the principle of Little's law [114], a new algorithm has been derived in the study to predict the expected node degree of a network following the RD or the RWP mobility model, $E[n]$ by:

$$\begin{aligned} E[n] &= E[\mathcal{T}_{lk}]\bar{\lambda} \\ &= \begin{cases} \frac{3.464r^2(\mathcal{N}-1)}{\mathcal{L}^2} & RD\ model \\ \frac{3.464\omega r^2(\mathcal{N}-1)}{\mathcal{L}^2} & RWP\ model \end{cases} \end{aligned} \tag{6.16}$$

6.5.2 Experiment 3: Node Degree Estimation

A third experiment is designed to verify the proposed approach for node degree prediction with extensive simulations reproducing networks of various node density. In this experiment,

the nodal mobility is regulated by either the RWP or the RD model ¶. I first try to create a set of network scenarios of decreasing node density by increasing the side length of the network area $\mathcal{L}$ from 600m to 1400m while fixing the number of nodes $\mathcal{N}$ at 40 and the radio radius r at 150m. For all of these scenarios, both the RD and the RWP mobility model are implemented with a maximum speed v_{max} of 20m/s. A warm up period of 1250 seconds is set in order to minimize the effect of speed decay in the RWP model and to keep a similar average speed in these two mobility models. The resulting analytical estimates as well as the simulation results are plotted in Fig. 6.17.

Fig. 6.17 compares the expected node degrees measured from simulation and the analytical results approximated by $E[n]$ (Eq. 6.16) for networks with either RWP or RD mobility in the scenarios of various side length. From Fig. 6.17, we can see excellent agreements between the actual data and the analytical results given by $E[n]$. The deviations between the actual and the analytical results are mostly less than 0.5. Note that, in order to diminish the border effect [115, appendix A], the actual average node degree are measured from the nodes that are located with the centering area of $500m \times 500m$ when the side length is 600m.

Another set of network scenarios has been also created with varying node density by changing r from 75m to 175m while setting $\mathcal{L}$ as 800m and $\mathcal{N}$ as 40. The mobility models involved in these scenarios are still RD and RWP. The analytical results for this set of network scenarios and the corresponding true values from simulations are given in Fig. 6.18. From Fig. 6.18, we can observe that $E[n]$ yields excellent performance in predicting the expected node degree in the scenarios of varying radio radius. Larger gaps between the analytical and the actual results of about 0.75 are found at radio radius of less than 100m, which may resulted from the inaccuracy of both $E[\mathcal{T}_{lk}]$ and $\bar{\lambda}$.

Another set of network scenarios is produced to test the impact of network mobility on the expected node degree. In these scenarios the node density is fixed ($\mathcal{L} = 800m$, $\mathcal{N} = 40$ and $r = 150m$) while the maximum speed limits v_{max} is varied from 5m/s to 40m/s. Fig. 6.19(a)

¶The RD movements in the simulations are generated by a mobility modeling tool described in [91]

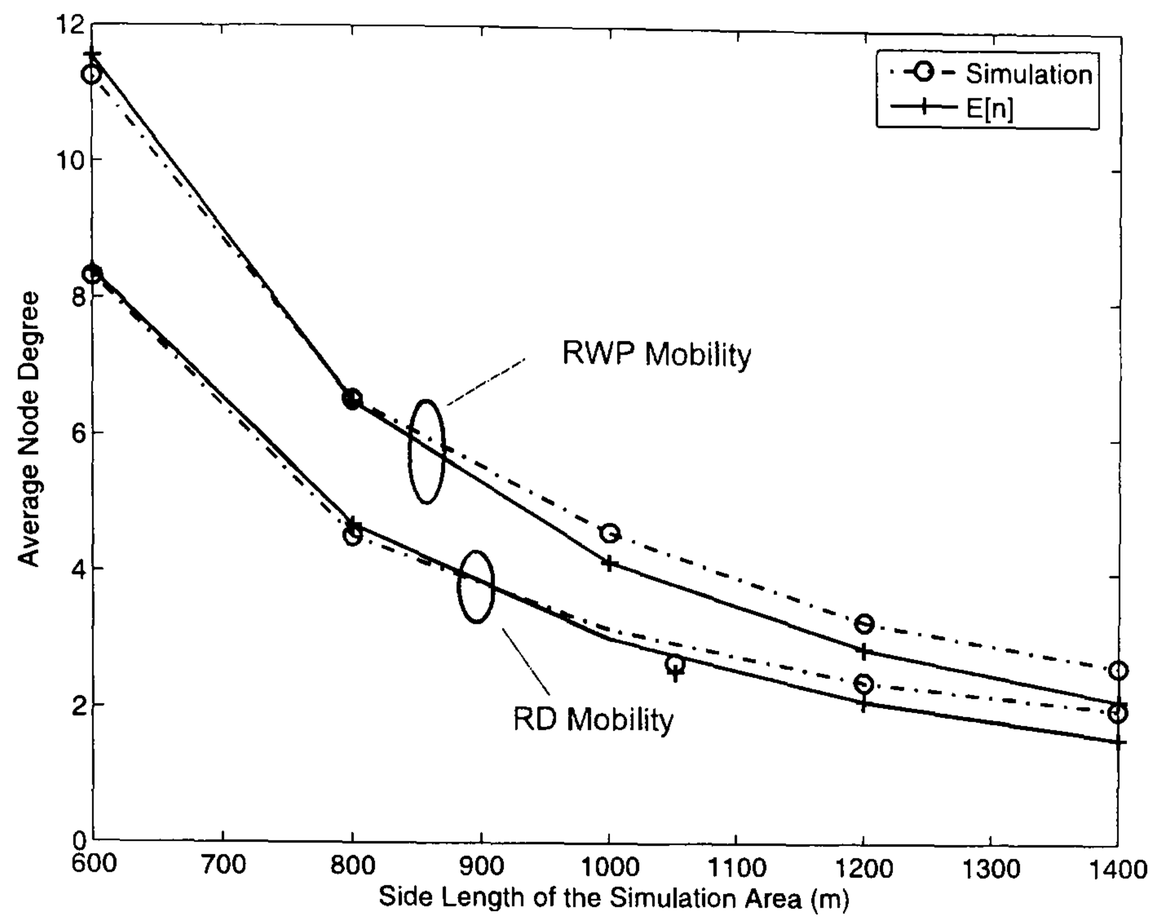


Figure 6.17: Average Node Degree Vs. Side Length of Simulation Area

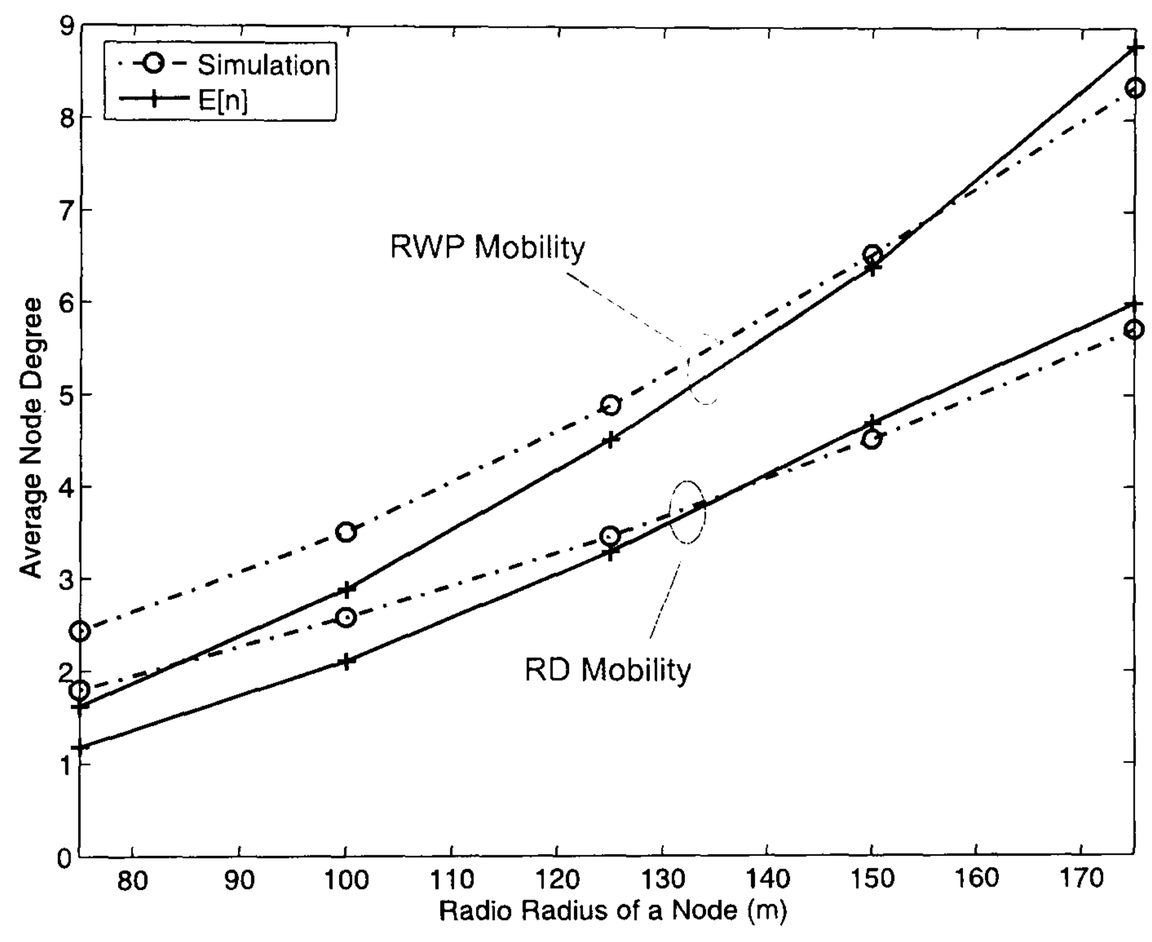


Figure 6.18: Average Node Degree Vs. Radio Radius

gives the actual expected node degree of these scenarios and the analytical results given by $E[n]$ as well as an approach given in [53] for the RWP model (denoted as $E[D_{rwp}]$). The actual results and those given by $E[n]$ as well as an approach from [53] for the RD model (denoted as $E[D_{rd}]$) are plotted in Fig. 6.19(b). We can see from Fig. 6.19 that, in general, the expected node degree is well approximated by both $E[n]$ and the approaches provided by [53] in the

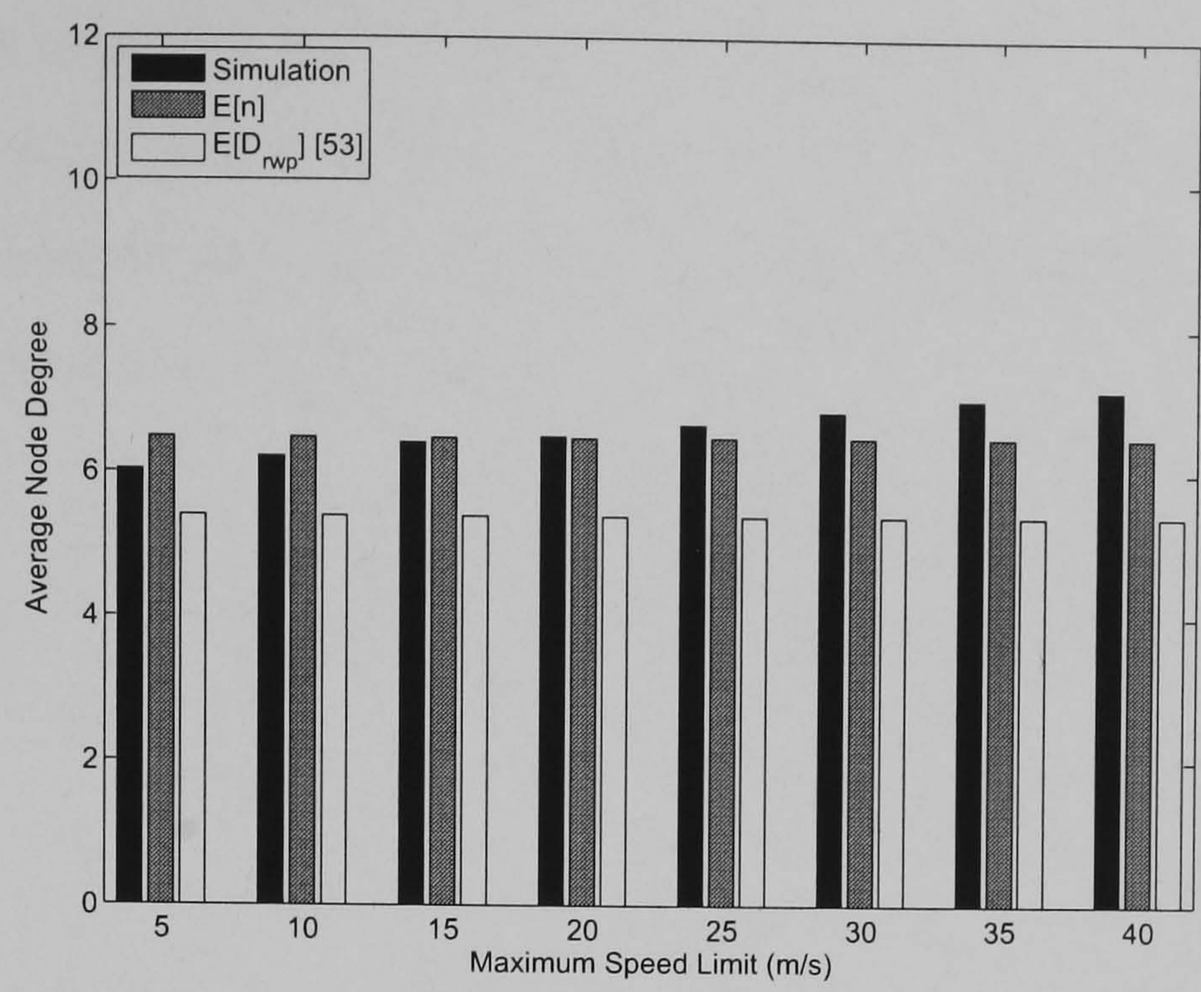
scenarios of varying network mobility and the former is slightly more accurate. From Fig. 6.19, we can also observe that the increase in the nodal movement speeds actually increases the expected node degree. The affect of nodal speed on the expected node degree is not addressed in either Eq. 6.16 or the approaches proposed in [53]. But, this can be explained by the proposed queueing theory based model, in which a node's neighbours are more likely to change their trajectories at higher mobility because of the shorter epoch time. Hence high speed nodes are likely to spent more time in the "queue" (e.g. link duration) resulting in an increased node degree.

The accuracy of the proposed method could be improved if a more realistic model of link duration is used. Nevertheless, the proposed method is precise enough to serve our purpose, i.e. to examine the applicability of the relative velocity estimator *RVEd* in a specific network scenario.

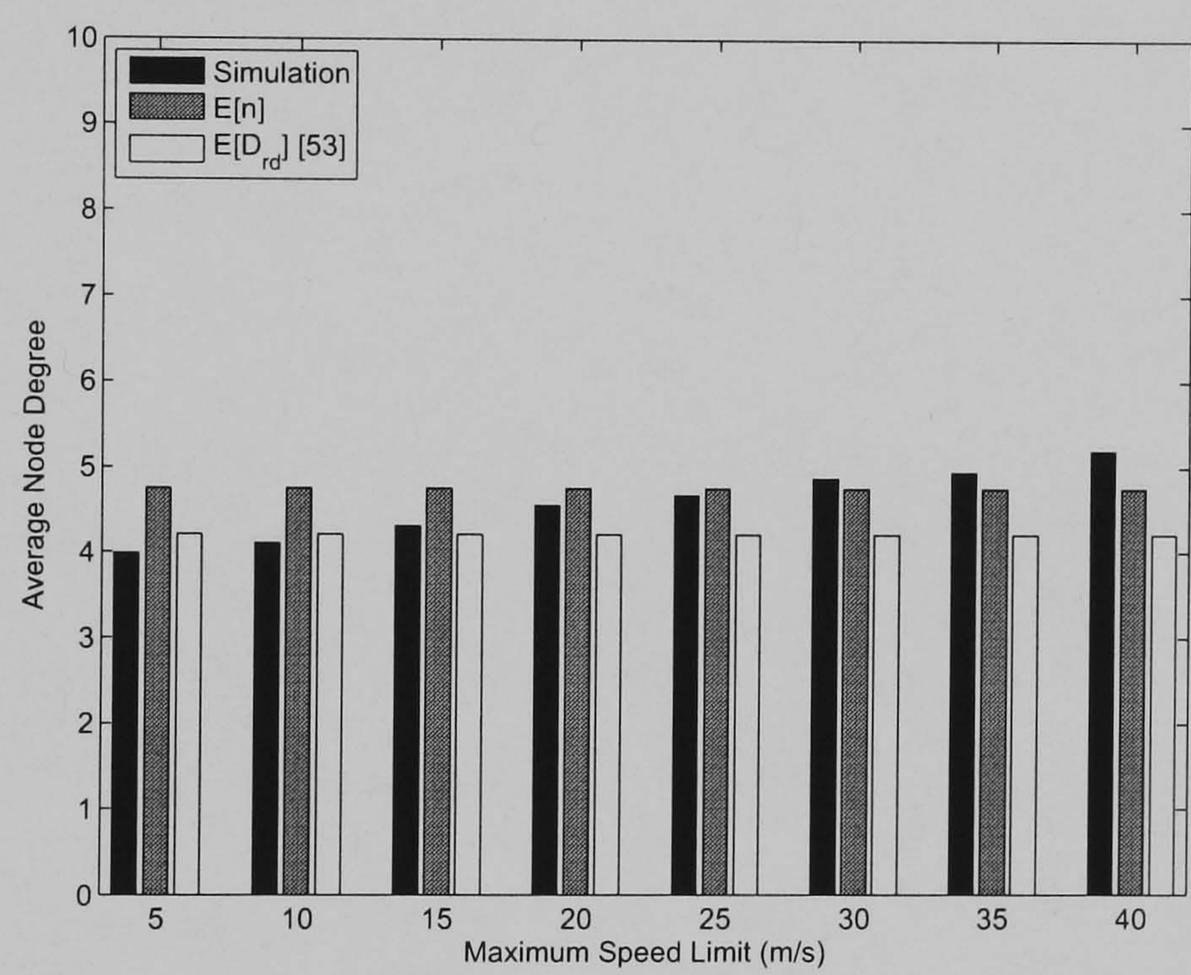
6.6 Summary

Knowledge of the inter-node relative velocity is important in mobile ad-hoc networking, as it is closely correlated with the performance of communications. Conventional ways of measuring the relative velocity involves complex localization systems such as the costly GPS or statistical approaches that rely on precise knowledge of the characteristics of wireless channels or signal propagation.

This Chapter presents a new method of Range-based relative Velocity Estimation that is less dependent on signal or channel characteristics and is more robust in noisy communication scenarios. Two velocity estimators (i.e. *RVEs* and *RVEd*), are derived for measuring relative velocity in sparse and dense ad-hoc networks, respectively. Using triangulation, the only information needed by the proposed methods is the time-varying ranges between mobile devices. Further, the proposed method is more tolerant of the multi-path or NLOS errors contained in range measurements than than existing methods. Thus, the new method is more robust in noisy



(a)



(b)

Figure 6.19: Average Node Degree Vs. Maximum Speed Limits. (a) Random Way Point Model (b) Random Direction Model

communication environments. A simple analytical algorithm for predicting the expected node degree in networks with uniform/non-uniform node distributions is also provided for examining the applicability of either of the two proposed methods. Simulation results show that the proposed methods work well in estimating the inter-node relative velocity for both sparse

and dense network cases regardless of the nodal movement speeds or the distribution of environmental noises. It is also shown that, compared to the existing method, multi-path or time-correlated NLOS noises in the range estimates have only negligible effects on the performance of the proposed velocity estimation methods.

Chapter 7

Range-based Mobility Characterization

7.1 Introduction

The quality of communications in mobile ad-hoc networks suffers from frequent topological changes caused by the relative movements between mobile nodes. Accurate and reliable characterization of nodal mobility (e.g. how often they move away from each other) in realtime is necessary for predicting the stability of the network topology for different configurations/-mobility models. For example, knowledge of the relative velocity between mobiles nodes and the average epoch time* is useful in predicting link availability [117] and provisioning of adaptability for routing schemes [118]. In this Chapter, a triplet $\langle V, \Theta, \Psi \rangle$ is firstly defined to characterize the nodal mobility. Referred to as a *mobility triplet*, it consists of mobility parameters that are directly correlated with the stability of a network (e.g. relative speeds and orientations between mobile nodes and their average epoch time).

Based on the range-based velocity estimators introduced in Chapter 6, a new ranging based scheme is presented to measure the stability-related nodal mobility in realtime. In contrast to existing methods that are based on either localization systems or features of a wireless channel, the proposed scheme estimates the parameters of the *mobility triplet* $\langle V, \Theta, \Psi \rangle$ only from the information of time-varying inter-node distances while retaining its reliability and accuracy

*As defined in [116], an epoch is a segment of a mobile node's path, during which the node travels in a constant direction at a constant speed. The time a node spent on one epoch is defined as the epoch time.

even in noisy environments. The performance of the proposed scheme is validated by computer simulations in networks of different properties. The scheme is also applied to a tool for link availability estimation [117] as a less complex way of obtaining mobility parameters. Preliminary results show that, with the proposed scheme, the performance of the tool for networks of different mobility models (e.g. Random WayPoint and Random Walk [11]) is comparable to when the nodal mobility is assumed to be known *a priori*.

The organization of this Chapter is as follows. The related work on mobility estimations for mobile networks are introduced in Section7.2. In Section 7.3, a new scheme of estimating the parameters of the *mobility triplet* is presented. The simulation settings for performance evaluation of the scheme and the obtained results are discussed in Section7.4. Section 7.5 gives a example application of the proposed scheme, in which the scheme is applied to link availability predictions. The preliminary results collected for this application are also given in this section. This Chapter is summarized in Section 7.6.

7.2 Related Work

Existing work related to the topic of interest in this Chapter is mobility characterization and link lifetime prediction. Real-time estimation/tracking of nodal mobility have been well studied in recent literatures [119] [120] [121] in order to support seamless connectivity and QoS provisioning. The existing proposals for mobility estimation/tracking either require the present of a centralized infrastructure [119] or assume the mobiles to be aware of the *environmental context* [121] or to be configured with a positioning system [120] (e.g. a GPS receiver [14] or a GPS-free positioning system [75]). For an ad-hoc network of battery-powered mobile devices, none of these proposals are optimal. Because ad-hoc networks are decentralized and the operations for retrieving a mobile's *environmental context* or geographical position are costly in terms of energy consumption.

Link lifetime prediction is another important topic in QoS provisioning. The average link

lifetime is an important QoS metric that reflect the reliability and quality of communications in a MANET. Predicted link expiry time can be also used for packet scheduling between neighbouring nodes. In [117], a prediction-based link availability estimation is introduced. It has been shown that, given the knowledge of nodal mobility (e.g. velocity and epoch time), this method can accurately predict the availability of a link. But, how the information of nodal mobility can be efficiently obtained has not been discussed further in the literature.

In Chapter 6, the range-based relative Velocity Estimations (RVE) have been introduced for MANETs. Using barely the information of the inter-node distance measured in realtime by techniques such as Time-of-Arrival (ToA) [100], the RVE estimators have the benefit of being less dependent on signal characteristics and more robust in noisy scenarios. In this Chapter, a triplet $\langle V, \Theta, \Psi \rangle$ is defined to represent the mobility parameters that are directly correlated with the topological stability, i.e. link availability [116] [117]. A new scheme of mobility estimations is also developed, which allows mobile nodes to derive the triplet $\langle V, \Theta, \Psi \rangle$ in realtime, using only the inter-node ranging information. The scheme is applied to the prediction-based link availability estimation as a more lightweight way of obtaining mobility parameters.

7.3 The Scheme of Range-based Mobility Estimations

In this section, the proposed scheme of mobility estimations based on the RVE estimations (e.g. *RVEd*) presented in Chapter 6 is presented. As introduced in Chapter 6, in an ad-hoc network of $\mathcal{N}$ mobile nodes the RVE estimators utilize the time-varying inter-node distances/orientations to approximate the relative speed between a pair of nodes. The information of inter-node ranging can be measured on a periodical basis by means of RSSI, ToA or TDoA. Every node also reports its collection of distance estimates to all of its one-hop neighbours. Thus, a node in the network knows not only the distances to its one-hop neighbours but also those between its neighbouring nodes within the radio coverage. The measurements and exchanges of distance information are carried out at a frequency of $1/\Delta t$, which can be finely

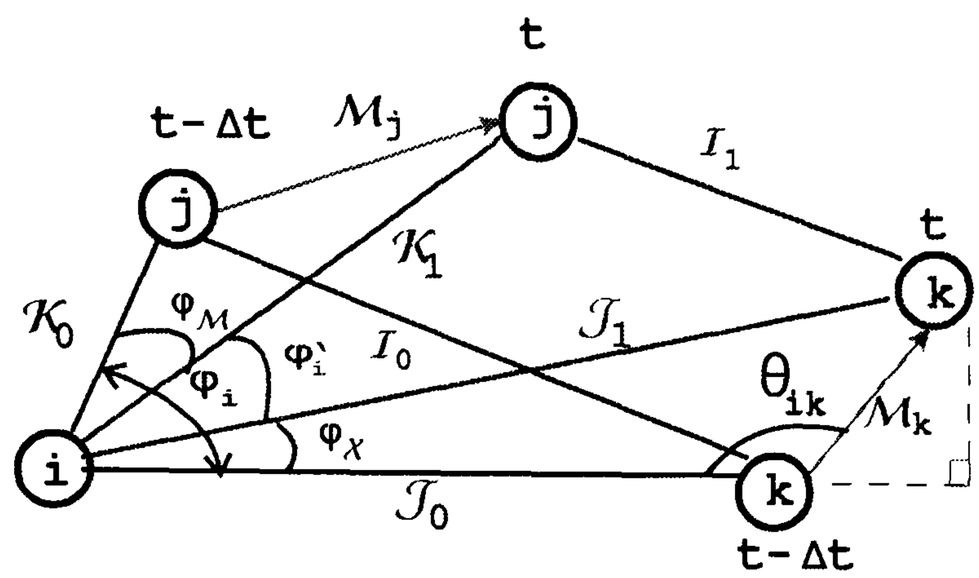


Figure 7.1: the changes in the ranges and the orientations between node i, j and k caused by their relative movements during a short period of time

tuned according to the resolution requirement of the mobility estimation.

Let $\mathcal{R}_i$ be the set of node i's one-hop neighbours that are also one-hop away to each other. In the proposed scheme for mobility estimation, a triplet of mobility parameters $\langle V_i, \Theta_i, \Psi_i \rangle$ is defined for a node i, where $V_i = \{\hat{v}_{ik} | k \in \mathcal{R}_i\}$ and $\Theta_i = \{\hat{\theta}_{ik} | k \in \mathcal{R}_i\}$ are the sets of estimated i's speeds and orientations relative to its one-hop neighbours, respectively. $\Psi_i = \{\psi_n | n = 1, 2, ..., C\}$ is the collection of C past epoch times of node i.

7.3.1 Estimating Inter-node Relative Speeds and Orientations

As shown in Fig. 7.1, given the presence of at least two neighbours within the radio coverage ($(j, k) \in \mathcal{R}_i$) including a *departing node*, i.e. node j, a node i can derive $\hat{v}_{ik}$ and $\hat{\theta}_{ik}$ in realtime using the *RVEd* method (the details of this method is given in Chapter 6).

Random relative movements among the three nodes in a short period of time can be categorized into four different patterns (see Fig. 6.3, Chapter 6). From the four patterns of relative position shift among the three nodes, two distinct mathematical relationships among the inter-node relative orientations (e.g. φ_i, $\varphi_i^{'}$, $\varphi_{\mathcal{M}}$ and $\varphi_{\mathcal{X}}$) can be derived, with which two possible values of $\varphi_{\mathcal{X}}$ can be obtained. In order to identify the true value of $\varphi_{\mathcal{X}}$, the algorithm has to be recalculated with another *departing node* or with the same *departing node* (e.g. j) but different ranging information measured at Δt time later.

Having determined the value of $\varphi_{\mathcal{X}}$, k's speed relative to node i during a time slot Δt, $\hat{v}_{ik}$,

can be estimated by:

$$\hat{v}_{ik} = \frac{\mathcal{M}_k}{\Delta t} = \frac{\sqrt{\mathcal{J}_1^2 + \mathcal{J}_0^2 - 2\mathcal{J}_1\mathcal{J}_0 cos\varphi_{\mathcal{X}}}}{\Delta t} \tag{7.1}$$

and the relative orientation $\hat{\theta}_{ik}$ by:

$$\hat{\theta}_{ik} = cos^{-1}\frac{\mathcal{J}_0^2 + \mathcal{M}_k^2 - \mathcal{J}_1^2}{2\mathcal{J}_0\mathcal{M}_k} \tag{7.2}$$

7.3.2 Estimating Epoch Time

The epoch time ψ_n of node i ($\psi_n \in \Psi_i$) can be determined by recording the time when the node starts a new epoch (or finishes the last epoch). Without the help of any localization systems or any speed meters, a node can not detect the changes in its own moving speed or orientation. Nevertheless, provided the presence of 2 or more neighbours most of the time, a node can still detect the start/end point of an epoch be exploiting the fact that the changes in its moving direction/speed affect its relative speeds/orientations to all of its neighbours. The pseudo code for the implementation of an algorithm for estimating epoch time in a node is listed in Fig. 7.2.

In the implementation, node i initializes a counter with value *0* at the beginning of updating sets V_i and Θ_i. The value of the counter is increased by *1* if a change in the speed or orientation to a neighbour in set $\mathcal{R}_i$ is detected. After this round of relative speed and orientation updates, current time is recorded as a starting point of a new epoch if the value of the counter equals $|\mathcal{R}_i|$, i.e. the number of node i's one-hop neighbours whose velocities have been estimated using the *RVEd* method. Therefore, epoch time is obtained from intervals between the starting points of consecutive epochs. Elapsed time of the current epoch is also recorded. The use of this information for link availability estimation is given in Section 7.5.

```
Velocity_Change_Count = 0 ;

for n in Neighbor_Set R

    Curr_v_n = Estimate_Relative_Speed ( n ) ;
    Curr_θ_n = Estimate_Moving_Orientation ( n ) ;

    if ( Prev_v_n != Curr_v_n || Prev_θ_n != Curr_θ_n )
            Velocity_Change_Count++;
            Prev_v_n = Curr_v_n ;
            Prev_θ_n = Curr_θ_n ;
    end if

end for

if ( Velocity_Change_Count == Number_of_Neighbors_in_R )

     Epoch_Times [i++] = Current_Time – Last_Epoch_Time;
     Last_Epoch_Time = Current_Time ;

else Curr_Epoch_Elapsed = Current_Time – Last_Epoch_Time;

end if
```

Figure 7.2: Pseudo Code for Estimations of Epoch Times

7.4 Performance Evaluation of the Scheme of Mobility Estimations

To validate the performance of the proposed scheme of mobility estimations, the ns-2 network simulator is deployed to simulate various network scenarios of different speed limits and mobility models in two experiments.

7.4.1 Simulation Model

The simulated mobile ad-hoc network consists of $\mathcal{N} = 40$ nodes moving around in a square area with the size of $\mathcal{L} \times \mathcal{L}$. The radio radius r of a node is fixed at 250m. Using both of the mobility modelling tools described in [99] and [91], the simulations cover a wide range of nodal mobility characterized by the Random Waypoint Model [11] and an enhanced version of the Random Walk model [91]. For both of the mobility models, the pause time is kept at *0* to produce continuous movement. The speed of a mobile node is uniformly distributed over

$[v_{min}, v_{max}]$, where v_{min} is fixed at 3.5m/s and v_{max} ranges between 5m/s and 40m/s.

Each mobile node measures its distance to the one-hop neighbours and exchanges this information with all of its neighbouring nodes at intervals of $\Delta t = 1$ second. Without considering the effects of NLOS conditions, the estimates of inter-node distance have measurement noise that are modeled by a zero-mean Gaussian distribution with variance σ_ϵ^2. The default value of σ_ϵ is 2m. The mobility estimations are carried out in a duration of 900 seconds starting after 1050 seconds of warm-up period.

In the figures generated from the simulation results, each data point corresponds to a mean of 30 repeated experiments with different random seeds.

7.4.2 Experiment 1

The experiment is to compare the epoch time estimated by the proposed scheme with the actual values measured with the integrated localization module in the simulators. Both the Random WayPoint and the Random Walk mobility models are involved in this experiment. In these two models, mobile nodes are bounced back when they reach the boundary of the network area.

For the scenarios of Random WayPoint mobility, the side length of the network area is varied from 400m to 800m for an increasing length of average epochs. The average duration of actual epochs in the network and the estimates provided by the proposed scheme are given in Fig. 7.3. As shown in Fig. 7.3, the actual values of average epoch time agree with their estimates regardless of the speed limits or the side lengths. In the scenarios with 400m and 600m of side lengths the estimates seems to be slightly overestimated. This may due to that in smaller areas mobile nodes are bounced back more often, which decreases the accuracy of the proposed method.

In the Random Walk model, the epochs of a mobile's movement is exponentially distributed with a mean of ψ. Fig. 7.4 gives the actual and estimated epoch time in scenarios of Random Walk mobility with various speed limits and mean epoch time (e.g. 20s, 40s and 60s). We

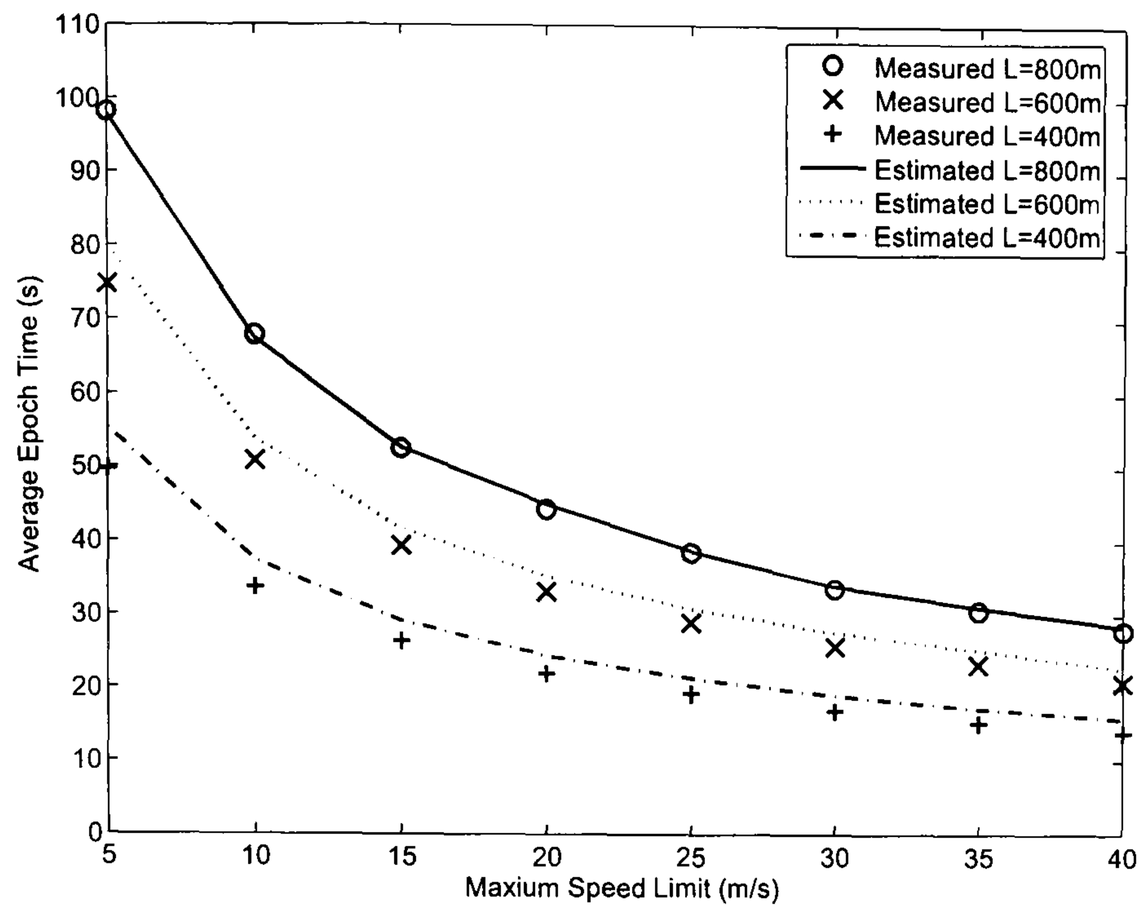


Figure 7.3: Average Epoch Time Vs. Maximum Speed Limit, Random WayPoint Mobility

can see from Fig. 7.4 that the estimates produced by the proposed scheme match the actual values again with the Random Walk mobility. It should be noted that both the measured and the estimated epoch time are shorter than the specified mean ψ. This is because a mobile often starts a new epoch even when the current one is not finished due to that it has reached the network boundary and is being reflected back. The frequency of such "reflections" is increasing with the nodal speed, which is confirmed by Fig. 7.4.

7.5 Link Duration and Availability Estimations with the *Mobility Triplet*

In this section, the proposed mobility estimation scheme is applied to a tool of link availability prediction presented in [117]. The purpose is to demonstrate that the proposed scheme can be used to provide networking tools/protocols with mobility information (e.g. relative speeds and epoch distributions) that are normally obtained from complex localization/movement tracking systems or assumed to be known *a priori*.

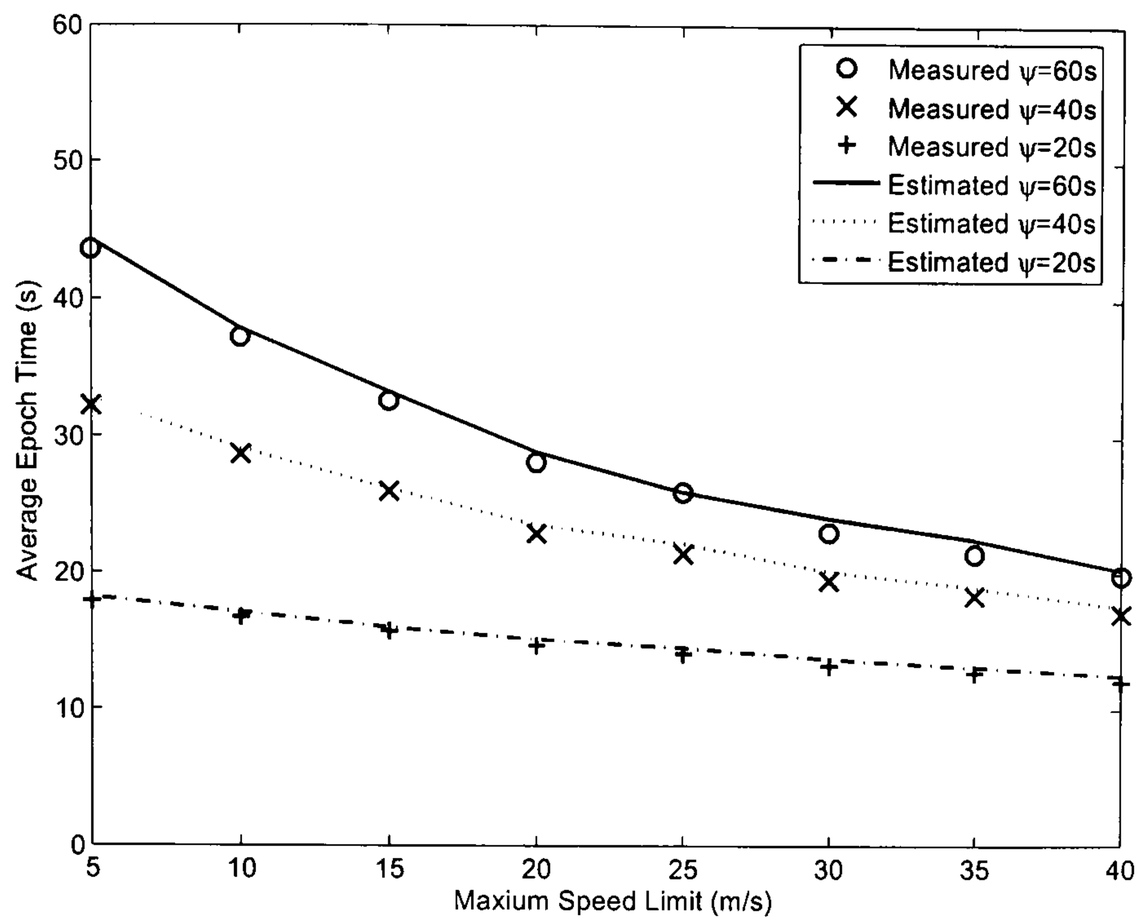


Figure 7.4: Average Epoch Time Vs. Maximum Speed Limit, Random Walk Mobility

7.5.1 the Prediction-Based Link Availability Estimation

Jiang et al. [117] proposed the prediction-based link availability estimation, which involves two consecutive stages: firstly, the estimation of the projected link life time (denoted as T_p), which is the continuous period from the time when the estimation is made until the link is broken, assuming that the relative speed and orientation between the two nodes are constant during this period. Secondly, as the real link life may not really last to T_p if either of the two nodes changed their epochs (see Fig. 7.5), the prediction of the probability (denoted as $L(T_p)$) that the real link life will be T_p or over.

For estimating the projected link life time T_p between two nodes (e.g. node *i* and *k* in Fig. 7.5), the authors in [117] recommended the use of velocity information obtained from GPS or a measurement-based prediction using 3 samples of the inter-node distance. For the prediction of $L(T_p)$ between *i* and *k*, one can use the following expression proposed in [117]:

$$L(T_p) = L_1(T_p) + L_2(T_p, \phi) \tag{7.3}$$

where $L_1(T_p)$ represents the case that neither of the two nodes changed their epochs within

time T_p. Let $\mathcal{A}_i(x) = \mathcal{P}\{\psi \leq x\}$ denote CDF of node i's epoch time. For a link connecting node i and k, $L_1(T_p)$ is given by:

$$L_1(T_p) = \mathcal{A}_i(T_p)\mathcal{A}_k(T_p) \tag{7.4}$$

$L_2(T_p, \phi)$ represents the case that either of the two nodes change their velocities at time ϕ within T_p ($\phi < T_p$). $L_2(T_p, \phi)$ can be calculated as [117]:

$$L_2(T_p, \phi) = \frac{\phi + p(T_p - \phi)\mathcal{A}_i(T_p - \phi)\mathcal{A}_k(T_p - \phi)}{T_p} + \varepsilon \tag{7.5}$$

where p=0.5 is the simplified probability that the two nodes will approach each other after an epoch change occurs. ε is an estimate of the probability that there are more than 1 epoch changes since ϕ (within T_p) and the contributes of these epoch changes to the link availability is positive. The value of ε is suggested to be measured in realtime due to the difficulty of its exact calculation [117].

The equations listed above indicate that the information of the predicted link lifetime T_p and the distribution of mobile nodes' epoch times is essential for this method of link availability prediction. In the rest of the Chapter, the proposed mobility triplet $\langle V, \Theta, \Psi \rangle$ is used for the calculation of Eq. 7.3-7.5 as a less complicated way of estimating link lifetime and distribution of nodal epochs that was not considered in [117].

7.5.2 Estimating Link Durations

With the mobility parameters $\langle V, \Theta, \Psi \rangle$ estimated by the proposed scheme, the projected link lifetime T_p can be approximated within 2 time slots, instead of 3 as suggested in [66] and [117]. This shortens the amount of time that a mobile needs to wait until it can have an estimation of T_p, especially when the inter-node distances are sampled over long time intervals.

As shown in Fig. 7.5, d_0 and d_1 are the distance estimates between i and k sampled at time

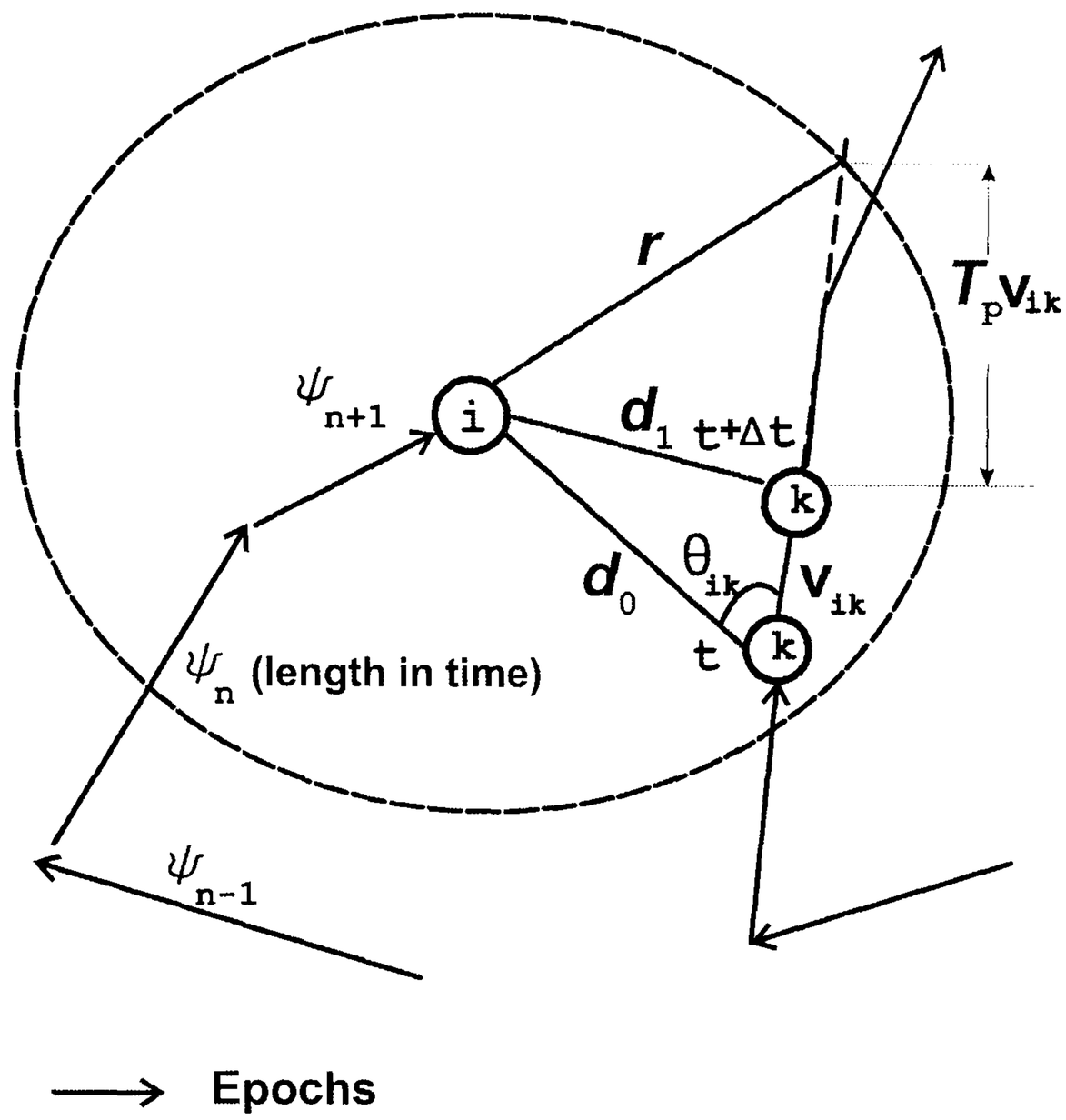


Figure 7.5: Epochs and link duration resulted from node i and k's movements

t and $t + \Delta t$, respectively. v_{ik} ($v_{ik} \in V_i$) is the relative speed and θ_{ik} ($\theta_{ik} \in \Theta_i$) the relative orientation between i and k estimated at $t + \Delta t$. According to the *Cosine Rule*, the following equation can be derived:

$$cos\hat{\theta}_{ik} = \frac{d_0^2 + (T_p + \Delta t)^2 \hat{v}_{ik}^2 - r^2}{2d_0(T_p + \Delta t)\hat{v}_{ik}} \tag{7.6}$$

As $T_p + \Delta t > 0$, the projected link life time T_p between node i and k can be derived from Eq. 7.7 as:

$$T_p = \frac{\sqrt{d_0^2 cos^2\hat{\theta}_{ik} - \hat{v}_{ik}^2(r^2 - d_0^2)} + d_0 cos\hat{\theta}_{ik}}{\hat{v}_{ik}^2} - \Delta t \tag{7.7}$$

7.5.3 Link Availability Predictions

In [117] and [118], the authors assume that the distributions and the mean length of epochs are known to a mobile node. This assumption is unrealistic as the network configurations

are normally not known to a mobile especially to those just joined the network. With the proposed scheme of mobility estimations, a mobile can now approximate the epoch lengths and their distributions in networks of unknown configurations, without the help of any movement tracking systems. If a mobile node (e.g. i) has recorded sufficient number (e.g. over 30 [87]) of its past epoch time, it can estimate the CDF of its epochs for Eq. 7.4 and Eq. 7.5 as follows:

$$\mathcal{A}_i(x) = \mathcal{P}\{\psi \leq x\} = 1 - \frac{|\Psi_i^x|}{|\Psi_i|} \tag{7.8}$$

where the subset $\Psi_i^x = \{\psi \in \Psi_i \wedge \psi > x\}$. For the calculation of $L_2(T_p)$ using Eq. 7.5, an expected time ϕ at which first epoch change occurs within T_p (denoted as $\bar{\phi}$) can be also approximated in realtime as:

$$\bar{\phi} = \frac{1}{|\Psi_i^s| + |\Psi_k^s|}\Big(\sum_{\psi_i \in \Psi_i^s} \mathcal{F}(\psi_i, \tau_i) + \sum_{\psi_k \in \Psi_k^s} \mathcal{F}(\psi_k, \tau_k)\Big) \tag{7.9}$$

where the subset $\Psi_i^s = \{\psi \in \Psi_i \wedge \psi < T_p + \tau_i\}$, τ_i denotes the elapsed time of current epoch of i †, and $\mathcal{F}(\psi, \tau) = \psi - \tau$ $(\psi > \tau)$ or 0 $(\psi \leq \tau)$. The link availability estimation based on the realtime estimation of $\langle V, \Theta, \Psi \rangle$ is denoted as $L_{rt}(T_p, \Psi)$. We have $L_{rt}(T_p, \Psi) = L_1(T_p, \Psi) + L_2(T_p, \Psi)$.

Note that in a network where the distribution of epochs of a node is homogeneous, Ψ_i can be approximated by Ψ_k in the above equations, and vice versa.

7.5.4 Experiment 2

In a second experiment, the performance of link availability predictions is tested with mobility parameters (e.g. $\langle V, \Theta, \Psi \rangle$) estimated by the proposed scheme in networks with Random Waypoint or Random Walk mobility. For each estimated T_p the actual residual link life time (denoted as T_r) is measured. The resulting T_r/T_p is the actual link availability. Similar to [117],

†Denoted as *Curr_Epoch_Elapsed* in the pseudo code listed in Fig. 7.2.

let $L_{min}(T_p, \Psi) = L_{rt}(T_p, \Psi) - \varepsilon$ be the conservative link availability without considering the factors represented by ε. From m pairs of $(T_r/T_p, L_{min})$ of different values collected from simulations, ε is estimated as:

$$\varepsilon = \frac{1}{m}\sum_{j=1}^{m}(\frac{T_{r,j}}{T_{p,j}} - L_{min}(T_{p,j}, \Psi_{i,j})) \quad (i \in \mathcal{N}) \tag{7.10}$$

Fig. 7.6 and Fig. 7.7 plot the 10s averages [‡] of the actual link availability T_r/T_p and those of $L_{rt}(T_p, \Psi)$, as well as $L(T_p)$ with artificially given parameters [117] for Random WayPoint and Random Walk mobility, respectively. Note that the values of $L_{rt}(T_p, \Psi)$ and $L(T_p)$ that are bigger than 1 after adding ε are approximated as 1. As shown in both Fig. 7.6 and Fig. 7.7, the performance of the realtime prediction $L_{rt}(T_p, \Psi)$ is comparable to that of $L(T_p)$. In Fig. 7.6 when T_p is between 150s and 350s, the curve of $L_{rt}(T_p, \Psi)$ even has similar fluctuating tendency with that of actual results. In some cases when the actual link availability drops, as those observed at T_p over 350s in Fig. 7.6 or at T_p over 160s in Fig. 7.7, $L_{rt}(T_p, \Psi)$ does not match well with the reality. This is due to the fact that there are less data collected from the simulated network when T_p is large while ε is an average of many events making $L_{rt}(T_p, \Psi)$ more accurate for frequently occurring events. A similar phenomenon and its explanation can be found in [117].

7.6 Summary

In wireless mobile ad-hoc networks, the topological stability has significant impact on the quality of communications. Knowledge of the mobility of mobiles, such as their relative speeds and orientations as well as their epoch time, is important as these parameters are closely correlated with the frequency of topological reconfigurations in the network. Conventional ways of measuring the nodal mobility involves complex localization systems such as the costly GPS or

[‡]As $L_{rt}(T_p, \Psi)$ or $L_{min}(T_p, \Psi)$ are based on epoch information collected in realtime, instead of a fixed epoch distribution as in [117], their values are fluctuating as those of T_r/Tp do. 10s averages of these data are plotted to show their tendency while keeping a certain degree of accuracy.

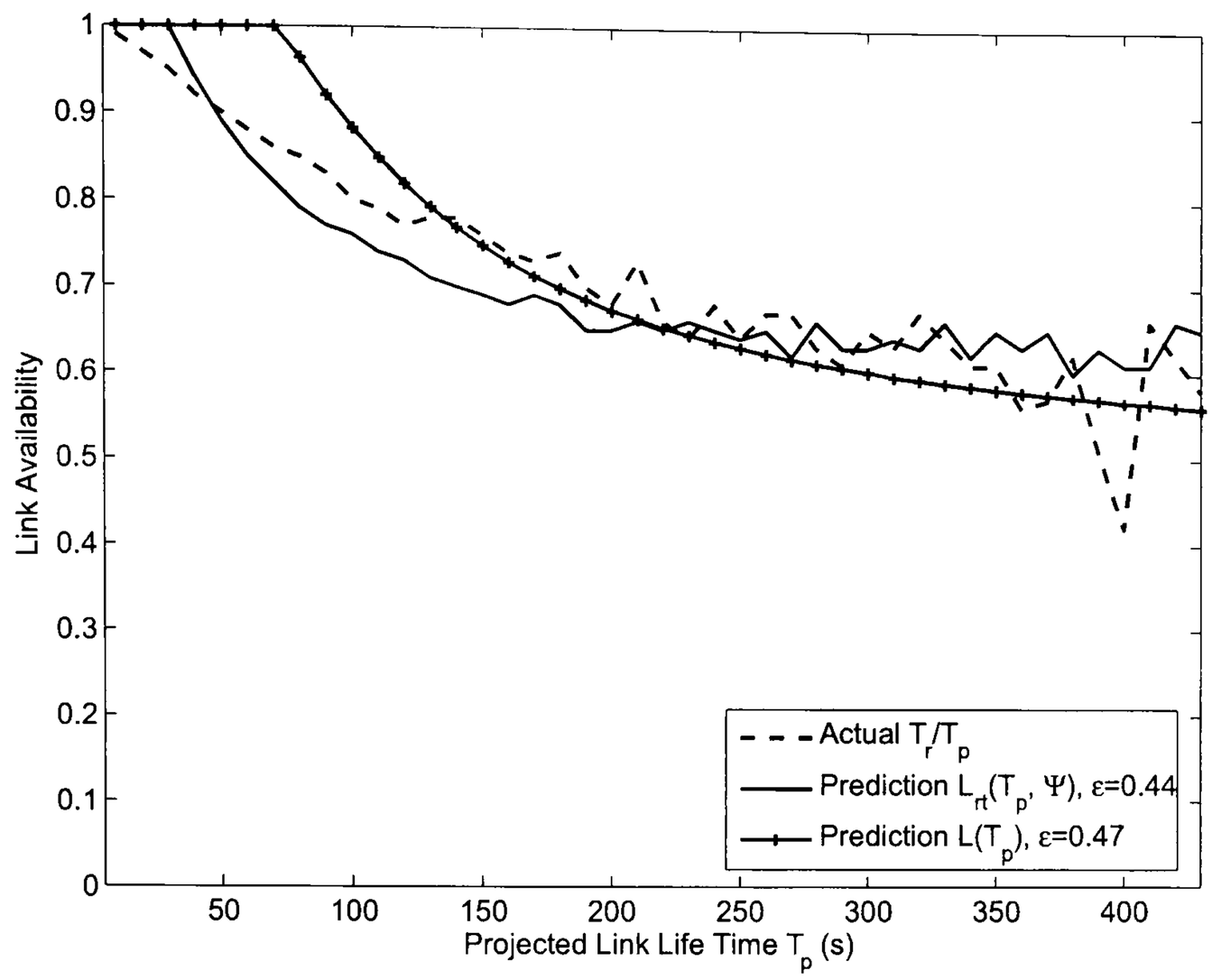


Figure 7.6: Predicted Residual Link Life Time Vs. Link Availability, Random WayPoint Mobility $\mathcal{L} = 800m$

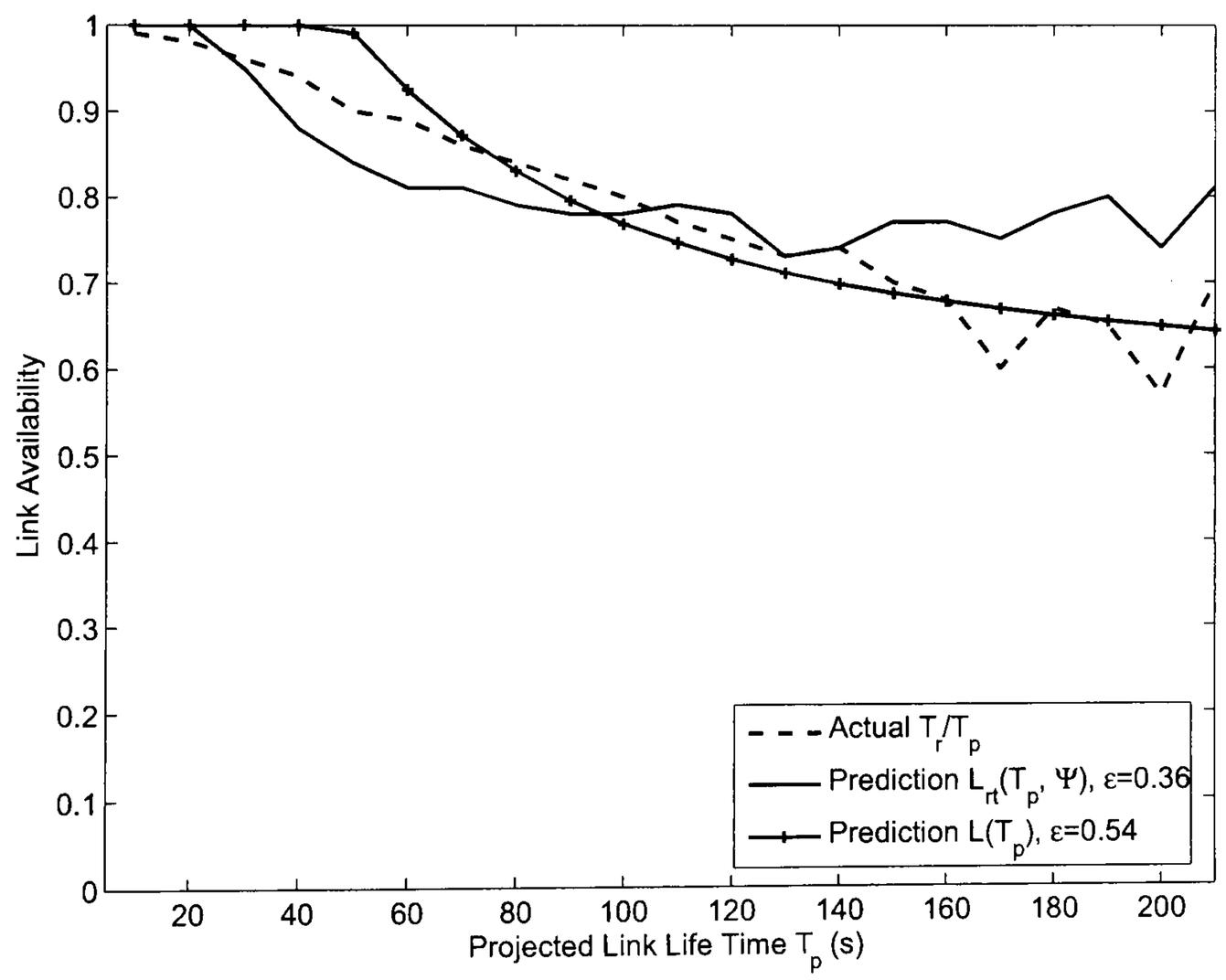


Figure 7.7: Predicted Residual Link Life Time Vs. Link Availability, Random Walk Mobility ψ= 45s

statistical approaches that rely on precise knowledge of the characteristics of signal propagation. In this Chapter, the scheme of range-based velocity estimations presented in Chapter 6 is extended to allow the estimation of the mobility of a mobile node (e.g. its relative speeds and orientations to its neighbours and its past epoch time), in realtime, using only the time-varying inter-node distance information. Represented by a *mobility triplet* $\langle V, \Theta, \Psi \rangle$, these mobility parameters are directly correlated to the stability of networks. In contrast to existing methods, the proposed scheme is less dependent on signal characteristics and is more robust in noisy environments. The performance of the proposed scheme is validated by computer simulations against an existing method in scenarios of various speed limits. As an application example, the proposed scheme is applied to an existing tool to provide mobility information for predicting link availability. Preliminary results from various network scenarios show that, with the proposed scheme, the tool achieved a comparable performance to when hard-coded parameters are used.

Chapter 8

Discussion, Future Work and Conclusions

8.1 Introduction

The last decade has witnessed the dramatic development of advanced wireless communication technologies as well as new networking concepts such as the Mobile Ad-hoc Networks (MANETs). A mobile ad-hoc network is an infrastructureless network that is comprised of mobile devices with temporary connections to their neighbouring nodes. Due to their unique characteristics (e.g. decentralized, multi-hop based packet delivery, unpredictable topology and limited computing/communication resources) communications in MANETs is more complex and difficult than in the wired Internet or cellular mobile/wireless networks. New effective and efficient networking strategies are necessary for MANETs to tackle the challenges that are not present in the wired/celluar IP networks and to achieve the performance expectations of communication services. This work emphasizes two fundamental aspects of mobile ad-hoc networking, namely, routing and mobility characterization. Without proper mechanisms for routing discovery, packet delivery in MANETs would be impossible. Existing routing protocols [6] [34] [30] [9] [85] have been shown to work well in delivering QoS sensitive traffic to their destinations in a MANET. However, the cost (in terms of resource consumption) is not under control in these protocols, which makes them unsuitable for networks where communications rely on scarce resources (e.g. bandwidth and power supply). An optimal routing solution would be one that can achieve the performance/QoS targets of the traffic with the least possi-

ble resource utilization. As the topology of MANETs is in constant reconfiguration, mobility awareness is also an important capability of the networking technologies, which can be used for choosing low mobility nodes as cluster heads, for estimating link/path life time and for prediction of communication quality in the current network settings. In order to measure the mobility of MANETs and the resulting performance, it is also necessary to define mobility metrics and to design algorithms for mobility measurement. Common mobility estimation algorithms and measures are based on either costly (in terms of energy/bandwidth usage) positioning systems or accurate knowledge of the properties of the wireless channel. This problem has motivated the development of resource-efficient and channel-independent mobility metrics and estimation methods for MANETs in the PhD project.

The main objectives are (1) To provide an understanding of the performance of existing ad-hoc networking technologies (e.g. routing and mobility characterization) in terms of QoS benchmarks (e.g. packet delivery ratio and end-to-end latency) as well as their resource utilizations (e.g. bandwidth overhead and energy consumption). 2) For ad-hoc networks of sparse and dense topology, to develop novel resource-efficient strategies for routing/service discovery with expected QoS and for characterization of the mobility of a mobile node/a whole ad-hoc network. 3)To Investigate the applications of these networking technologies to achieve QoS-adaptive communications (e.g. distributed adaptation of packet relay in sparse networks) and QoS predictions (e.g. to predict the loss rate of packet delivery and the availability of wireless links) in MANETs.

This Chapter discusses the main contributions and novelty of this work (Section 8.2), suggests future work (Section 8.4) and completes the dissertation with the conclusions (Section 8.5).

8.2 Contributions to Knowledge

The work presented in the dissertation has contributed to knowledge with the following:

1. **The development of a new "track-based" scheme for resource efficient hybrid service/routing discovery in dense MANETs**

   A critical problem facing existing routing or service discovery protocols is how to minimize resource utilization while preserving QoS promises for applications. The novelty of the "track-based" scheme is that it adopts a one-dimensional track-like structure for hybrid routing and service discovery. In comparison with existing hybrid protocols that are based on two-dimensional structures, the operations of the scheme require much less bandwidth and traffic overhead. The scheme is also adaptive to the network traffic and mobility leading to enhanced QoS provisioning. Preliminary results confirms that the "track-based" scheme is efficient in terms of bandwidth overhead, while providing QoS guarantees (e.g. querying delay and first attempt success rate) that are comparable to those of the existing methods. The work is described in Chapter 3.

2. **The development of a *store-and-forward* based scheme, Adaptive Multi-copy Routing (AMR), for energy-efficient and QoS guaranteed packet delivery in sparse MANETs with intermittently available wireless connections.**

   Relaying packets to their destinations in sparse MANETs covering a wide area demand a *store-and-forward* based strategy that can deal with the varying networking conditions that cannot be anticipated at the source for better resource-efficiency. Existing *store-and-forward* based protocols cannot achieve optimal performance as their delivery strategy is decided by the source. The proposed AMR scheme shifts the control of multi-hop relaying from the source to intermediate nodes. Information on the current status of the delivery progress is piggybacked on a packet header so that intermediate nodes in the scheme can independently adapt to the current networking conditions that cannot be sensed by the source and to exploit the situation for less traffic overhead while preserving QoS targets. The results show that the AMR scheme minimizes bandwidth and energy consumption whilst delivering packets within specific QoS targets (e.g. end-to-end delay

and packet loss). The work is presented in Chapter 4.

3. **The development of the concept and the associated algorithms of a new resource efficient "GPS-free" mobility metric for performance measurement in dense MANETs.**

   Most of the existing mobility metrics require the use of costly positioning systems (e.g. the GPS) for their measurements. The key novelty of the "GPS-free" mobility metric is that using only ranging information to neighbouring nodes, it is able to capture the inter-node relative movements in a 2D plane, in real-time. No costly (in terms of energy consumption) localization systems are required for the measurement of the metric. The proposed mobility metric can be used to approximate the network-wide average of inter-node relative speed. Several variants of this metric are also derived, which can be used for predicting of QoS measures (e.g. inter-group inter-meeting time and packet delivery rate) in networks with group and random mobility. To deal with the environmental noises in the ranging information, a calibration method is also developed to improve the accuracy of the metrics. The work is reported in 5.

4. **The development of two range-based velocity estimators for resource-efficient estimations of the relative velocity between mobile nodes in MANETs of sparse and dense node distribution. Based on the developed velocity estimators, the introduction of a range-based mobility estimation scheme for prediction of network reliability.**

   As the GPS-free mobility metric can only measure the network-wide average of inter-node relative speed, the range-based velocity estimators, namely *RVEs* and *RVEd*, are developed as a resource-efficient alternative for estimating of relative speeds/orienation between pairs of nodes. Based on the time-varying range information between a node and its neighbours, the *RVEs* and the *RVEd* methods are developed for estimation of relative velocity in sparse MANETs and dense MANETs (given the presence of a *departing node*), respectively. In comparison with existing methods, the velocity estimators

have major benefits, such as being robust in noisy conditions, independent of the characteristics of the wireless channel/signal, and efficient resource utilization as positioning systems are not needed. In order to allow mobiles to determine the applicability of the two estimators, an efficient algorithm for predicting the node density is also proposed, which has simple and unified form for both networks with uniform and non-uniform node distribution.

A new scheme for characterizing nodal mobility is also developed on the basis of the range-based velocity estimators. The scheme characterizes the mobility of a node with a triplet $\langle V, \Theta, \Psi \rangle$, which consists of elements that are correlated with the reliability of the wireless links and the network, i.e. a mobile node's speeds, orientations relative to its neighboring nodes and its historical epoch times. As an extension of the range-based velocity estimators, the scheme has similar benefits such as being reliable for channels of unknown properties, robust in noisy environment, and efficient utilization of resources, as the bandwidth/energy required for positioning is saved. The range-based relative velocity estimators are introduced in Chapter 6. The scheme for range-based mobility estimations and its application in link availability perdition are presented in Chapter 7.

8.3 Limitations of Current Works and Discussions

The work carried out in the project has a number of limitations that should be addressed in future studies.

- Simulation based performance evaluation

  The performance of the proposed protocol, algorithms and methods in this project are examined in simulated networks using the *ns-2* simulator. This approach has benefits of being fast, repeatable, easy to configure and customize. In a simulated network, many parameters such as node mobility, density and size of the geographical area are all con-

trollable. Simulation based tests are also much more economical than those based on emulation or physical implementation, which involves computing devices and communications interfaces. However, the creditability of the simulation based tests depends on the quality and accuracy of the simulation models used. For instance, the signal propagation models may fail to capture some aspects of the properties of the channel that may affect the operations of the protocols being tested. Consequently, simulation results of this project may not entirely reflect the performance of the developed algorithsms/protocols in reality.

- The appropriateness and applicability of the Mobility models

  Several parts of the project rely on an accurate model of the nodal movement. For example, the Adaptive Multi-copy Routing (AMR) scheme presented in Chapter 4 is based on the mobility model of the network to derive the distribution of inter-meeting time between contacts for distributed adaptation of relaying strategies. The method of node degree prediction described in Chapter 6 depends on an appropriate model of the arriving rate of neighbours at a node. In reality, the movement of mobile devices exhibit distinct characteristics that vary when they are in different environments, carried by different hosts or at different time of a day. As a result, ad-hoc networking technologies, such as the proposed AMR scheme and the algorithm of node degree estimation, may not perform as expected in the implementation if the mobility model they rely on does not reflect the reality. The reliability, accuracy and applicability of the existing mobility models are a major problem in this work and within the ad-hoc network research community in general.

- Lack of investigation into the integration of the developed networking technologies with MANET applications/services

  In this project, the performance of the protocols and algorithms developed are validated with network level QoS metrics (e.g. packet delivery rate, end-to-end delivery delay and normalized bias of estimates). But an investigation into the performance of QoS

sensitive applications and services (e.g. multimedia streaming) when they are integrated with the technologies developed remains. To present a clearer picture of the benefits of the developed networking protocols, it is necessary to perform quantitative analysis of the improvement contributed by these to the applications.

- Lack of consideration of other real world issues

  In addition to the specific reality-related issues raised above, there are other real world problems that have not been considered in the design of the proposed protocols/algorithms (or many other recent works in this area). Selfish or misbehavior nodes* in the MANET is a serious problem that may prevent the operation of many networking technologies. In the AMR scheme, the probability of the relaying node being a selfish node should be taken into consideration to improve its performance in real world implementations. The noncooperating nodes need to be considered also in the design of the range-based mobility metrics and velocity estimators, which require sharing of a node's ranging information to one-hop neighbours.

8.4 Suggestions for Future Work

8.4.1 General directions of future work

There are mainly three aspects of the research that can be improved and extended further in future work.

- Performance evaluation using emulation and real system implementations

  The problem of the simulation based performance evaluation can be addressed in the future by emulation. Emulation brings in some aspects of reality while keeping a certain degree of repeatability, configurability and other advantages of simulation. Emulation

*Nodes that are reluctant to join cooperative operations in the network because of some selfish reasons (e.g. for preserving its own energy) or being in an abnormal status.

can be considered as a compromise solution between simulation and physical implementation in a test bed. However, emulation is still not as popular as simulation in the ad-hoc network research community because of the difficulties in accurately emulating the mobility of mobiles as well as in recreating the desired number of nodes and the conditions of signal propagations.

In emulation based experiments, some networking parameters that are hard to recreate (e.g. nodal mobility, traffic congestion, queueing delay, etc.) are emulated using computer software whereas other modules are implemented in a real world system. For example, a metropolitan network can be emulated by linking several wireless devices with a computer server and creating the realistic traffic conditions in the server with simulated packet losses/delays caused by traffic congestions. By using emulations experiments are performed in a semi real environment, i.e. using real operating systems that are operating on real devices and running real applications. The algorithms and protocols developed in this project can be tested in an emulation environment to get more convincing results and to observe their performance when being implemented in real devices.

In order to study the behaviour of developed technologies in reality and to validate the simulation/emulation results, their physical implementations in real systems are eventually required as the research matures. The actual performance of the protocols and algorithms as well as accuracy of the simulation/emulation experiments can be fully tested as real functioning software. Furthermore, the collected results can be used to prove their usefulness for commercial applications.

- Adaptive, intelligent and self-tuning methods for mobility modeling

  The mobility of hosts such as human and vehicles (except those for public transport) are unpredictable in nature. The best solution to the inaccuracy and inapplicability of existing mobility models may not lie in more complex and more accurate models but an intelligent method that allows mobile nodes to capture the characteristics of nodal

mobility in the network and to update its knowledge by self-learning and self-tuning in different environments or scenarios. Such a method can be based on technologies such as neural networks, cognitive agents and evolutionary computing. Some on going work in this direction can be found in [122] [123] [124].

- Integration and interaction with MANET applications/services. In order to disclose the behaviour of the developed technologies in a real system and their effects on the upper layer applications/services, it is necessary to incorporate the proposed algorithms and protocols into a real world application, especially with one in which the provision of QoS guarantees is critical. The application of the developed technologies for an envisioned QoS sensitive Service, a Mobile Ad-hoc Grid in the scenario of E-Healthcare Emergency, is introduced in the next subsection.

8.4.2 Facilitating Mobile Ad-hoc Grid Computing for E-Healthcare Emergency

It would be also interesting to see an implementation of the developed protocols and algorithms to enhance the performance of the supporting mechanisms of a mobile ad-hoc Grid [125], e.g. service discovery, job submission or data sharing. The perspective of mobile ad-hoc Grids would benefit applications such as the e-healthcare emergency. Considering a medical emergency taking place in a hostile environment where access to high-performance computing/communication infrastructures is not available, a mobile ad-hoc Grid can be built instantly from a group of heterogeneous mobile devices/sensors for collaborative problem solving and coordinated resource sharing in order to accomplish complex and crucial medical missions (e.g. diagnosis and treatment) on site.

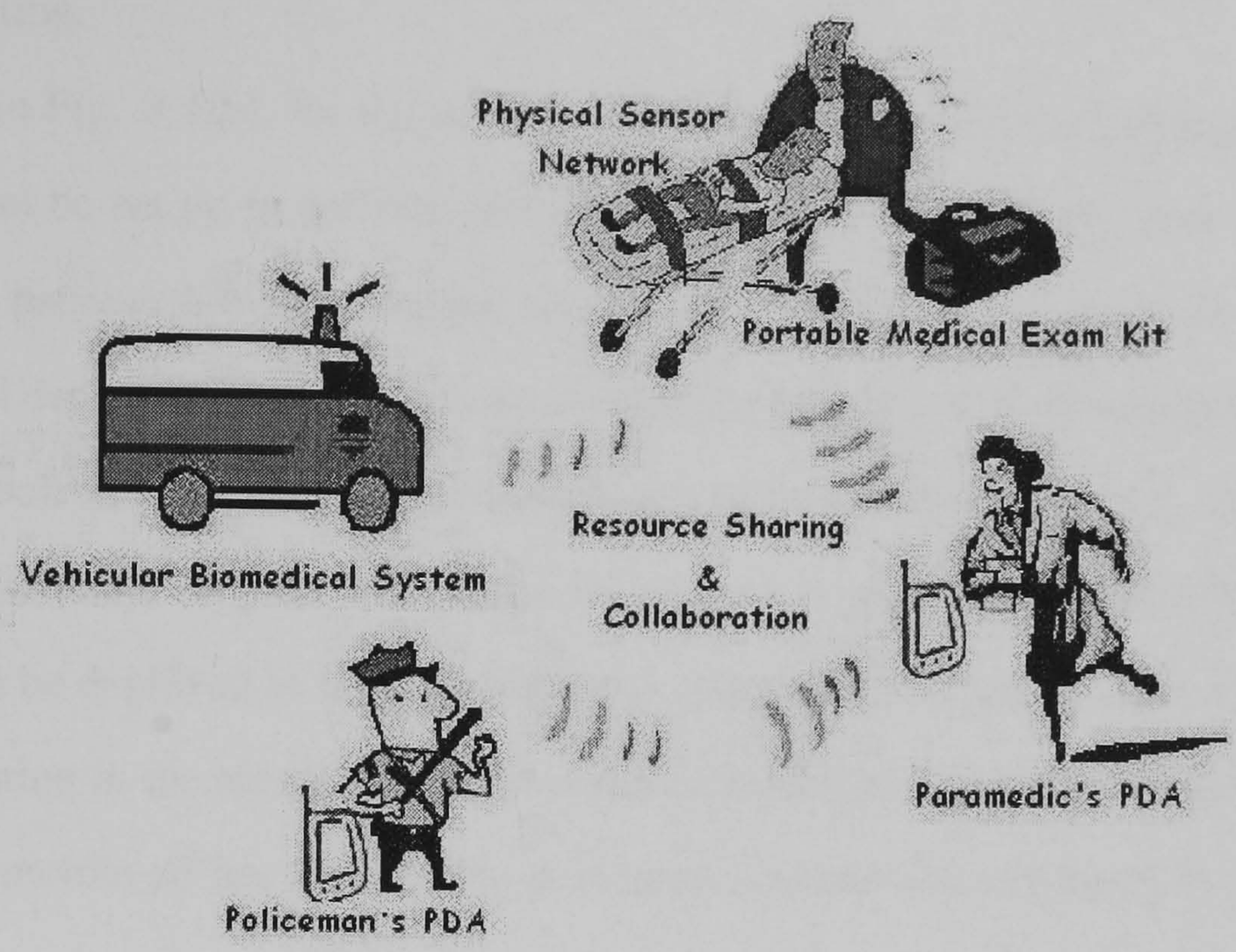
A graphic demonstration of the scenario

We present a scenario of healthcare emergency in Fig. 8.1 to illustrate how the technologies developed in this project can be utilized to address the challenges of constructing a mobile ad-hoc Grid that has a high demand on network reliability and QoS guarantees.

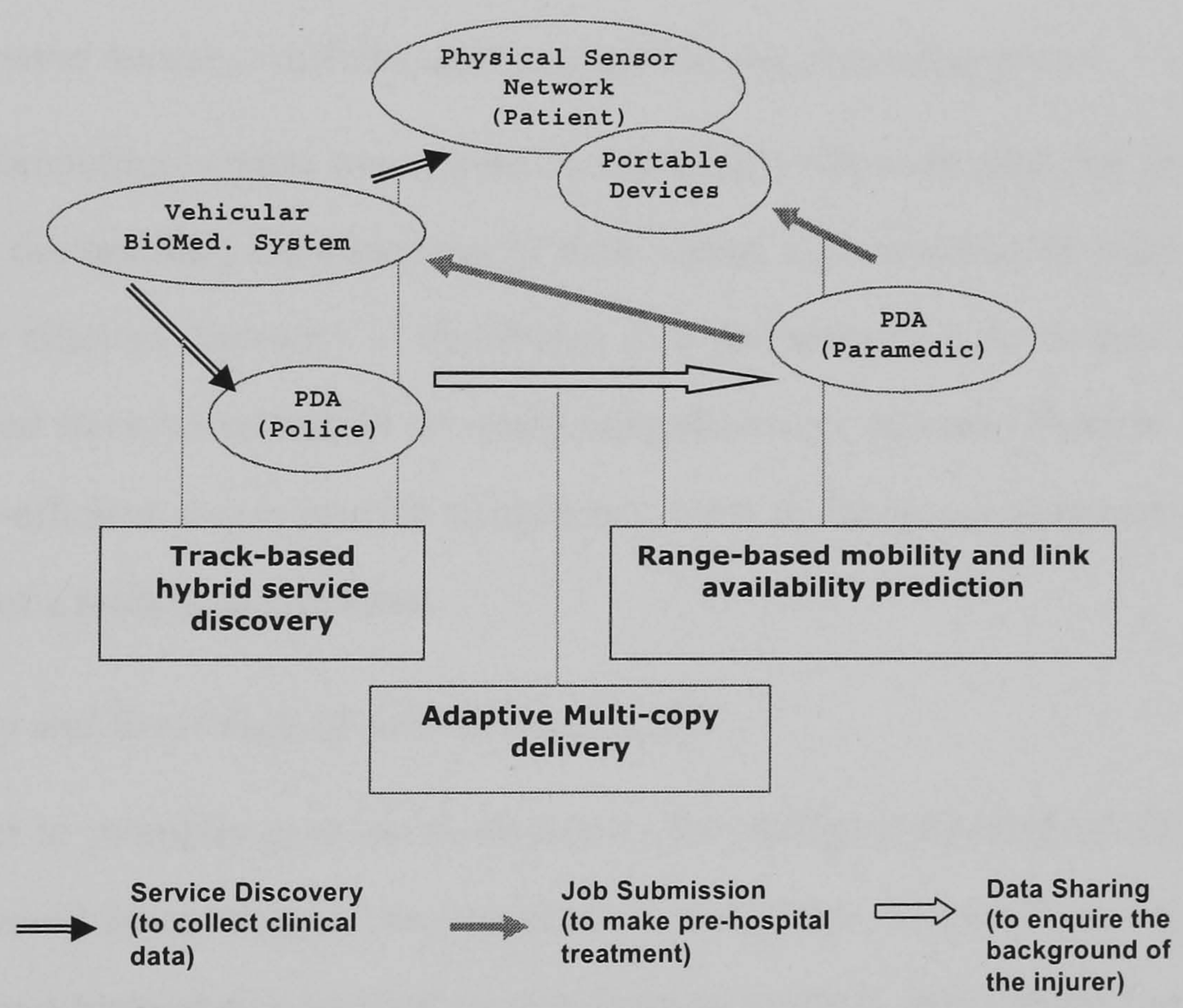
As depicted in Fig. 8.1(a), a car accident has happened in the middle of a highway. A man has been seriously injured and become unconscious. Police and ambulance have been called and coming to the scene of accident. By the time when an ambulance arrives, a policeman has arrived and collected the victim's personal or family information in his PDA. As soon as the ambulance and the paramedics arrive, a mobile ad-hoc Grid is constructed on the network of wireless devices (e.g. the police's PDA, the physical sensors attached to the injured person, a portable medical exam kit and the biomedical system in the ambulance) to process the collected information and to collaborate in various tasks for producing pre-hospital treatment suggestions.

Resource-efficient QoS provisioning for Grid computing in the Scenario

The ability to carry out pre-hospital diagnosis is of crucial importance to the injured person in the scenario, as "the appropriate medical intervention during the golden-window immediately following an accident significantly improves the chance of recovery" [126]. In order to give preliminary treatment suggestions in the scenario, an optimal solution is utilize Grid computing on top of the mobile ad-hoc network to process the physical data and the medical records of the injured person, which are distributed among the body sensors and wireless PDAs, with intelligent algorithms carried by the on-board biomedical systems. However, it is a challenging task for the mobile ad-hoc Grid to operate on top of the network of portable battery-powered devices in the wireless environment, as Grid computing was designed for wired reliable infrastructures. In future works, we will deploy the developed technologies to enhance the reliability (with QoS guarantees) of the ad-hoc network of resource constraint devices for the operations



(a)



(b)

Figure 8.1: Facilitating mobile ad-hoc Grid computing in e-Healthcare emergency . (a) a typical scenario of e-healthcare emergency: car accident, (b) the integration of the developed technologies with the system of mobile ad-hoc Grid computing.

of Grid computing.

As shown in Fig. 8.1(b), for the envisioned scenario of e-healthcare emergency, a mobile ad-hoc Grid can be set up to perform distributed data (e.g. vital signals from the biosensors attached to the patient) discovery and job scheduling (e.g. to analyzed received vital signals in several medical devices). These operations are sensitive to the QoS of the infrastructure and networking protocols in use. The proposed resource-efficient technologies (e.g. track-based service/routing discovery, adaptive multi-copy routing and range-based mobility/link availability prediction) can be deployed as building blocks to guarantee QoS and to support the operations of Grid computing in the scenario. We list at the following how the unique QoS requirements imposed by a mobile ad-hoc Grid in the envisaged scenario are addressed by the developed techniques.

- *On-demand access to distributed/roaming data and computing power*

  Grid computing expects transparent access to network-wide data and processing algorithms disregarding their location or their status, e.g. roaming or static. This would require efficient discovery of distributed data and services even in high mobility. The proposed track-based hybrid service/routing discovery scheme (Chapter 3) provides an energy-efficient way to provide on-demand access to data/processing power for the operations of a mobile ad-hoc Grid.

- *Latency and Error Rate of data in transmission*

  In order to promptly give initial diagnosis, the intelligent biomedical algorithms in the Grid would have stringent requirements of delay (for one-way/two-way communications) and bit/packet error rate for transmitting medical data. Some types of physical data would expect a minimum transmission rate, i.e. an accurate 12 lead over-sampled ECG set would require a transmission rate of 64kbps [126]. The provision of these QoS requirements in MANETs is challenged by the unreliability of wireless channels caused by various factors such as nodal mobility and signal interference, etc. To deal with this

problem, the proposed Adaptive Multi-copy Routing (AMR) (Chapter 4) can be deployed to increase the redundancy in packet delivery (e.g. for data sharing) and to guarantee the expected delay/delivery rate against erroneous and fluctuating links.

- *Reliable and adaptable communication*

  The operations of the Grid computing model were developed for being carried out on a stable and reliable communication infrastructure. As a result, there are not many mechanisms developed in the Grid computing model to cope with the unreliability of the underlying network, i.e. link breakages caused by relative movements between nodes. This problem can be addressed by the proposed range-based mobility metrics (Chapter 5) and estimators (Chapter 6 and 7) as well as the AMR routing scheme. For instance, the range-based mobility estimators can be used to examine reliability of the links/paths to a remote host in choosing a right candidate for submission of computing jobs. The AMR routing can be also deployed to improve the efficiency of reliable packet delivery by tuning the redundancy levels of packet delivery accordingly upon link failures. Therefore, the impacts of network failures on the operations of the mobile ad-hoc Grid can be minimized for a smoother diagnosis and treatment procedure.

8.5 Conclusions

Motivated by the challenging problem of preserving Quality of Service (QoS) requirements in mobile ad-hoc networks of resource-constrained wireless devices, the project was initiated to investigate and to develop resource-efficient networking strategies that can be used as light-weight alternatives to existing technologies to support a QoS-guaranteed and sustainable ad-hoc communication architecture. Specifically, the aim is to develop novel networking protocols and algorithms that can be used to facilitate QoS provisioning mechanisms (e.g. QoS prediction and QoS-adaptive communication) based on the limited computing power and communication capability of MANETs.

During the course of this project, several topics on mobile ad-hoc networking problems (e.g. service/routing discovery in networks of dense topology, packet delivery in networks of sparse topology and mobility estimations) have been addressed, starting with an investigation into the efficiency of existing technologies in terms of energy and bandwidth consumption. Based on in-depth analysis of the existing work, what follows are new solutions and improved designs, validated by theoretical analysis and computer simulations, that achieve significantly reduced resource utilization without compromising the expected Quality of Service (QoS).

The novelty and contributions of this project include a new one dimensional track-based structure for hybrid service/routing discovery, a new scheme of Adaptive Multi-copy Routing (AMR) for packet delivery over intermittently connected links, as well as the range-based mobility metrics, velocity estimators and mobility characterization scheme that capture the nodal/network-wide mobility in a two-dimensional plane, in real-time, without the use of costly localization systems or channel-dependent signal processing algorithms.

The outcomes of the project are new protocols and algorithms that can be used as building blocks (e.g. QoS guaranteed packet relay and link availability prediction) for a reliable and sustainable network architecture built on top of resource-constrained mobile devices or wireless sensors. The performance and behaviour of the proposed protocols and algorithms need to be further tested with physical systems, real nodal movements and commercial applications before a realistic implementation can be made for QoS sensitive scenarios, such as the e-healthcare emergency.

References

[1] J. Haartsen and S. Mattisson, "BLUETOOTH a new low-power radio interface providing short-range connectivity," *in Proc. of the IEEE Special Issue on Low-Power RF Systems*, vol. 88, no. 10, pp. 1651–1661, 2000.

[2] "IEEE 802.11 standard for wireless LAN, medium access control (MAC) and physical layer (PHY) specification," *IEEE Standards Department*, 1999.

[3] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: Current performance benchmarks and future potential," *in IEEE Communications Magazine*, vol. 43, pp. 129–136, February 2005.

[4] R. C. Qiu, H. Liu, and X. Shen, "Ultra-wideband for multiple access communications," *IEEE Communications Magazine*, vol. 43, pp. 80–87, 2005.

[5] S. Chen and K. Nahrstedt, "Distributed quality-of-service in ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1488–1505, August 1999.

[6] C. E. Perkins, E. M. Belding-Royer, and S. Das, "Ad-hoc on-demand distance vector (AODV) routing," *RFC 3561*, July 2003.

[7] I. Gruber, R. Schollmeier, and W. Kellerer, "Performance evaluation of the mobile peer-to-peer service," *in Proc. of IEEE International Symposium on Cluster Computing and the Grid (CCGrid'2004)*, pp. 363–371, April 2004.

[8] Z. J. Haas and M. R. Pearlman, "The Zone Routing Protocol (ZRP) for ad hoc networks," *Internet Draft, Mobile Ad-Hoc Network (MANET) Working Group, IETF*, 1998.

[9] A. Vahdat and David Becker, "Epidemic routing for partially-connected ad hoc networks," *Technical Report, Duke University*, April 2003.

[10] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Multi-copy routing in intermittently connected mobile networks," *Technical Report, University of Southern California*, 2004.

[11] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. & Mobile Comput. (WCMC)*, vol. 2, no. 5, pp. 483–502, 2002.

[12] L. M. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," *in Proc. IEEE Infocom*, vol. 3, pp. 1548–1557, April 2001.

[13] J. Aspnes, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, Y. R. Yang, B. D. O. Anderson, and P. N. Belhumeur, "A theory of network localization," *IEEE Transaction on Mobile Computing*, vol. 5, pp. 1663–1678, December 2006.

[14] A. El-Rabbany, "Introduction to GPS: The global positioning system," *ISBN 1580531830, Artech House Publishers, Boston*, January 2002.

[15] X. Zhang, J. K. Kurose, B. N. Levine, D. Towsley, and H. Zhan, ""'study of a bus-based disruption-tolerant network: mobility modeling and impact on routing"," *in Proc. ACM MobiCom*, pp. 195–206, September 2007.

[16] Z. Li, L. Sun, and E. Ifeachor, "Track-based hybrid service discovery in mobile ad-hoc networks," *the 16th Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC' 2005), Berlin, Germany*, vol. 2, pp. 1274–1278, September 2005.

[17] Z. Li, L. Sun, and E. C. Ifeachor, "Adaptive multi-copy routing for intermittently connected mobile ad hoc networks," *in Proc. of IEEE GlobeCom Technical Conference (GlobeCom' 2006), San Francisco, USA*, November 2006.

[18] Z. Li, L. Sun, and E. C. Ifeachor, "Energy-efficient packet relaying in sparse mobile sensor networks," *in Handbook of Wireless Mesh & Sensor Networking, George Aggelou (Editor), approved for publication by McGraw-Hill*, 2007.

[19] Z. Li, L. Sun, and E. C. Ifeachor, "GPS-free mobility metrics for mobile ad-hoc networks," *to be published by IET Proc. on Communications*, vol. 1, pp. 970–976, October 2007.

[20] Z. Li, L. Sun, and E. C. Ifeachor, "Range-based relative velocity estimators for networked mobile devices," *IEEE Transaction on Vehicular Technologies (under review)*, 2007.

[21] Z. Li, L. Sun, and E. C. Ifeachor, "Range-based mobility estimations in MANETs with application to link availability prediction," *in Proc. of IEEE ICC Conference (ICC'2007), Glasgow, UK*, pp. 3376–3382, June 2007.,

[22] CISCO Systems Inc., "Mobile ad hoc networking enhancements for router-radio links," *Production Literature*, July 2007.

[23] Proxim Wireless Corp., "Technical specifications of ORiNOCO 11b client PC card," *www.proxim.com, accessed on May 2007.*

[24] K. H. Wangand Baochun Li, "Group mobility and partition prediction in wireless ad-hoc networks," *in Proc. of IEEE ICC 2002*, vol. 2, pp. 1017–1021, October 2002.

[25] R. Bruno, M. Conti, and E. Gregori, "Mesh networks: Commodity multihop ad hoc networks," *IEEE Communications Magazine*, pp. 123–131, March 2005.

[26] I. Stoica, R. Morris, D. Liben-Nowell, D. R. Karger, M. F. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup protocol for internet applications," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 17–32, February 2003.

[27] M. Klein, B. Konig-Ries, and P. Obreiter, "Lanes - a lightweight overlay for service discovery in mobile ad hoc networks," *in Proc. of 3rd Workshop on Applications and Services in Wireless Networks (ASWN2003)*, July 2003.

[28] "Universal plug and play (upnp™) device architecture version 1.0," *UPnP™Forum*, June 2000.

[29] J. Veizades, E. Guttman, C. Perkins, and S. Kaplan, "Service location protocol," *RFC 2165*, June 1997.

[30] U. C. Kozat and L. Tassiulas, "Network layer support for service discovery in mobile ad hoc networks," *in Proc. of IEEE Infocom 2003, San Francisco, USA*, vol. 3, pp. 1965–1975, March 2003.

[31] C. Frank and H. Karl, "Consistency challenges of service discovery in mobile ad hoc networks," *Proc. of the 7th ACM international symposium on Modeling, analysis and Simulation of WIreless and Mobile systems*, pp. 105–114, October 2004.

[32] E. M. Royer and C. Toh, "A review of current routing protocols for ad-hoc mobile wireless networks," *IEEE Personal Communications Magazine*, pp. 46–55, April 1999.

[33] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," *Mobile Computing, Kluwer Academic Publishers*, vol. 353, pp. 153–181, 1996.

[34] Z. J. Haas, J. Y. Halpern, and L. Li, "Gossip-based ad hoc routing," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 479–491, June 2006.

[35] J. C. Liu, Q. Zhang, W. Zhu, and B. Li, "Service locating for large-scale mobile ad hoc networks," *Kluwer International Journal of Wireless Information Networks*, vol. 10, pp. 33–40, January 2003.

[36] I. Rubin, A. Behzad, H.-J. Ju, R. Zhang, X. Huang, Y. Liu, and R. Khalaf, "Ad hoc wireless networks with mobile backbones," *IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC 2004)*, vol. 1, pp. 566–573, September 2004.

[37] A. Helmy, S. Garg, P. Pamu, and N. Nahata, "Contact Based Architecture for Resource Discovery (CARD) in large scale MANETs," *Proc. of Int'l IEEE/ACM Parallel and Distributed Processing Symp.*, April 2003.

[38] V. Ramasubramanian, Z. J. Haas, and E. G. Sirer, "SHARP: A hybrid adaptive routing protocol for mobile ad hoc networks," *in Proc. of ACM MobiHoc Conference 2003*, pp. 303–313, June 2003.

[39] L. Wang and S. Olariu, "A two-zone hybrid routing protocol for mobile ad hoc networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, pp. 1105–1116, December 2004.

[40] S. Basagni, I. Chlamatac, V. Syrotiuk, and B. Woodward, "A distance routing effect algorithm for mobility (DREAM)," *in Proc. of The 4th Annual ACM/IEEE International Conference On Mobile Computing And Networking (Mobicom'98)*, pp. 76–84, October 1998.

[41] Y. B. KO and N. H. Vaidya, "Location-aided routing (LAR) in mobile ad-hoc networks," *Wireless Networks*, vol. 6, pp. 307–321, September 2000.

[42] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," *in Proc. of the 6th Annual ACM/IEEE International Conference On Mobile Computing And Networking (Mobicom'2000)*, pp. 243–254, August 2000.

[43] P. Jacquet, P. Muhlethaler, A. Qayyum, A. Laouiti, L. Viennot, and T. Clausen, "Optimized link-state routing protocol (OLSR)," *Internet Engineering Task Force RFC 3626*, October 2003.

[44] S. Merugu, M. Ammar, and E. Zegura, "Routing in space and time in networks with predictable mobility," *Georgia Institute of Technology, Technical Report, GIT-CC-04-7*, 2004.

[45] S. Jain, K. Fall, and E. Zegura, "Routing in a delay tolerant network," *in Proc. ACM SIGCOM Conference, Portland, USA*, vol. 34, pp. 145–158, October 2004.

[46] Z. Zhang, "Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: Overview and challenges," *IEEE Communications Surveys and Tutorials, 1st Quarter*, vol. 8, no. 1, pp. 1105–1116, 2006.

[47] J. Leguay, T. Friedman, and V. Conan, "DTN routing in a mobility pattern space," *in Proc. of WDTN Workshop in association with ACM SIGCOMM*, pp. 276–283, August 2005.

[48] A. Lindgreny, A. Doria, and O. Scheleny, "Probabilistic routing in intermittently connected networks," *SIGMOBILE Mobile Computng Communication Review*, 2003.

[49] Y. Wang, S. Jain, M. Martonosi, and K. Fall, "Erasure-coding based routing for opportunistic networks," *in Proc. of ACM SIGCOMM Workshop on Delay Tolerant Networking and related topics (WDTN-05)*, pp. 229–236, August 2005.

[50] Y. Liao, K. Tan, Z. Zhang, and L. Gao, "Estimation based erasure-coding routing in delay tolerant networks," *Microsoft Technical Report*, 2006.

[51] Q. Zheng, X. Hong, and S. Ray, "Recent advances in mobility modeling for mobile ad hoc network research," *in Proc. ACM Southeast Regional Conference, Alabama, USA*, pp. 70–75, 2004.

[52] E. M. Royer, P. M. Melliar-Smith, and L. E. Moser, "An analysis of the optimum node density for ad hoc mobile networks," *in Proc. of IEEE ICC'2001*, pp. 857–861, June 2001.

[53] C. Bettstetter, "On the connectivity of ad hoc networks," *The Computer Journal*, vol. 47, pp. 432–447, July 2004.

[54] J. Yoon, M. Liu, and B. Noble, "Random waypoint considered harmful," *in Proc. of INFOCOM, San Francisco*, vol. 2, pp. 1312–1321, April 2003.

[55] X. Hong, T. Kwon, M. Gerla, D. Gu, and G. Pei, "A mobility framework for ad hoc wireless networks," *In Proc. of ACM Second International Conference on Mobile Data Management (MDM 2001)*, January 2001.

[56] V. Davies, "Evaluating mobility models within an ad hoc network," *Masters thesis, Colorado School of Mines*, 2000.

[57] F. Bai, N. Sadagopan, and A. Helmy, "IMPORTANT: A framework to systematically analyze the impact of mobility on performance of routing protocols for adhoc networks," *in Proc. of IEEE INFOCOM*, Apr. 2003.

[58] A. Jardosh, E. Royer, K. Almeroth, and S. Suri, "Towards realistic mobility models for mobile ad hoc networks," *in Proc. ACM MobiCom03*, pp. 217–229, September 2003.

[59] T. Tugcu and C. Ersoy, "How a new realistic mobility model can effect the relative performance of a mobile networking scheme," *Wiley Journal on Wireless Communications and Mobile Computing*, vol. 4, no. 2, pp. 383–394, 2004.

[60] J. Yoon, B. Noble, M. Liu, and M. Kim, "Building realistic mobility models from coarse-grained traces," *in Proc. of the 4th International Conference on Mobile Systems, Applications, and Services (MobiSys'2006)*, June 2006.

[61] M. McNett and G. M. Voelker, "Access and mobility of wireless pda users,," *ACM Mobile Computing Communications Review (MC2R)*, vol. 9, pp. 40–55, April 2005.

[62] B. Kwak, N. Song, and L. E. Miller, "A mobility measure for mobile ad-hoc networks," *IEEE Commun. Lett.*, vol. 7, pp. 379–381, Aug. 2003.

[63] P. Johansson, T. Larsson, N. Hedman, B. Mielczarek, and M. Degermark, "Scenario-based performance analysis of routing protocols for mobile ad-hoc networks," *in Proc. of ACM/IEEE MobiCOM*, Aug. 1999.

[64] J. Boleng, W. Navidi, and T. Camp, "Metrics to enable adaptive protocols for mobile ad-hoc networks," *in Proc. of the Intern. Conf. on Wireless Networks (ICWN'02)*, pp. 293–298, Jun. 2002.

[65] N. Sadagopan, F. Bai, B. Krishnamachari, and A. Helmy, "Paths: Analysis of path duration statistics and their impact on reactive manet routing protocols," *in Proc. ACM MobiHoc03*, June 2003.

[66] L. Qin and T. Kunz, "Increasing packet delivery radio in DSR by link prediction," *Internatonal Conference of Hawaii System Science*, 2003.

[67] A. Sampath and J. M. Holtzman, "Estimation of maximum doppler frequency for handoff decisions," *in Proc. of IEEE Vehicular Technology Conference*, pp. 859–862, May 1993.

[68] P. Mohapatra,, C. Gui, and J. Li, "Group communications in mobile ad hoc networks," *IEEE Computer Magazine*, February 2004.

[69] L. H. II and R. Tafazolli, "A survey of qos routing solutions for mobile ad hoc networks," *to appear in IEEE Communications Surveys and Tutorials, 2nd Quarter*, 2007.

[70] M. Allman, V. Paxson, and W. Stevens, "TCP congestion control," *RFC 2581*, April 1999.

[71] "Methods for subjective determination of transmission quality," *ITU-T Recommendation P.800*, August 1996.

[72] M. Andrews, J. Cao, and J. McGowan, "Measuring human satisfaction in data networks," *in Proc. of IEEE INFOCOM 06*, 2006.

[73] E. Althaus, G. Calinescu, I. I. Mandoiu, S. Prasad, N. Tchervenski, and A. Zelikovsky, "Power efficient range assignment for symmetric connectivity in static ad hoc wireless networks," *Wireless Networks*, vol. 12, june 2006.

[74] R. Zheng, J. C. Hou, and L. Sha, "Optimal block design for asynchronous wake-up schedules and its applications in multihop wireless networks," *in IEEE Transactions on Mobile Computing*, vol. 5, September 2006.

[75] S. Capkun, M. Hamdi, and J. Hubaux, "GPS-free positioning in mobile ad-hoc networks," *Cluster Computing*, vol. 5, Apr. 2002.

[76] G. Sharma, R. Mazumdar, and N. Shroff, "Delay and capacity trade-offs in mobile ad hoc networks: A global perspective," *in Proc. IEEE Infocom'06, Barcelona, Spain*, April 2006.

[77] "JINI™ architectural overview, technical white paper," *Sun Microsystem*, January 1999.

[78] "UDDI version 3.0.2," *UDDI Spec Technical Committee Draft*, October 2004.

[79] M. R. Pearlman and Z. J. Haas, "Determining the optimal configuration for the zone routing protocol," *IEEE Journal on Selected Areas in Communications*, vol. 17, August 1999.

[80] T. Clausen, P. Jacquet, and L. Viennot, "Analyzing control traffic overhead versus mobility and data traffic activity in mobile ad-hoc network protocols," *ACM Wireless Networks journal (Winet)*, vol. 10, July 2004.

[81] "The network simulator - ns-2," *available on line at http://www.isi.edu/nsnam/ns.*

[82] Z. J. Haas, J. Deng, and S. Tabrizi, "Collision-free medium access control scheme for ad-hoc networks," *in Proc. of IEEE MilCom'99, Altantic City, USA*, October 1999.

[83] J. Leguay, T. Friedman, and V. Conan, "Evaluating mobility pattern space routing for DTNs," *in Proc. IEEE Infocom*, April 2006.

[84] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Single-copy routing in intermittently connected mobile networks," *in Proc. of IEEE SECON*, October 2004.

[85] T. Small and Z. J. Hass, "Resource and performance tradeoffs in delay-tolerant wireless networks," *in Proc. of WDTN Workshop in association with ACM SIGCOMM*, August 2005.

[86] R. Groenevelt, "Stochastic models for mobile ad hoc networks," *PhD Thesis, University Nice-Sophia Antipolis*, April 2005.

[87] D. C. Montgomery and G. C. Runger, "Applied statistics and probability for engineers (second edition)," *John Wiley*, 1999.

[88] W. Navidi and T. Camp, "Stationary distributions for the random waypoint mobility model," *IEEE Transactions on Mobile Computing*, vol. 3, pp. 99–108, January-March 2004.

[89] S. Cho and J. P. Hayes, "Impact of mobility on connection stability in ad hoc networks," *in Proc. of IEEE WCNC*, pp. 1650–1656, March 2005.

[90] Y. Han, R. J. La, and A. M. Makowski, "Distribution of path durations in mobile ad-hoc netwroks - PALM's theorem at work," *in Proc. of the 16th ITC Specialist Seminar*, August 2004.

[91] J. Y. Le Boudec and M. Vojnovic, "Perfect simulation and stationarity of a class of mobility models," *in Proc. of INFOCOM, Miami, FL*, pp. 2743–2754, March 2005.

[92] B. An and S. Papavassiliou, "An entropy-based model for supporting and evaluating route stability in mobile ad hoc wireless networks," *IEEE Commun. Lett.*, vol. 6, pp. 328–330, August 2002.

[93] E. Commisson, "The galilei project - galileo design consolidation," *Technical Documents of the GALILEO - European Satellite Navigation System*, October 2003.

[94] H. Wu, C. Wang, and N.-F. Tzeng, "Novel self-configurable positioning technique for multi-hop wireless networks," *IEEE Transations on Networking*, vol. 13, pp. 609–621, June 2005.

[95] Ekahau Inc., "Ekahau positioning engin 4.1," *Ekahau Product Datasheet, available on www.ekahau.com*, September 2007.

[96] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, pp. 54–69, July 2005.

[97] X. Cheng, A. Thaeler, G. Xue, and D. Chen, "TPS: A time-based positioning scheme for outdoor wireless sensor networks," *in Proc. IEEE (INFOCOM04), Hong Kong, China*, vol. 4, pp. 2685–2696, March 2004.

[98] K. H. Wang and B. Li, "Efficient and guaranteed service coverage in partitionable mobile ad-hoc networks," *in Proc. of IEEE INFOCOM*, June 2002.

[99] C. de Waal and M. Gerharz, "Bonnmotion: a mobility scenario generation and analysis tool," *http://www.cs.uni-bonn.de/IV/BonnMotion/*.

[100] J. Werb and C. Lanzl, "Designing a positioning system for finding things and people indoors," *IEEE Spectrum*, vol. 35, no. 9, pp. 71–78, 1998.

[101] B. Alavi and K. Pahlavan, "Modeling of the distance error for indoor geolocation," *in Proc. of IEEE WCNC*, 2003.

[102] G. Azemi, B. Senadji, and B. Boashash, "Mobile unit velocity estimation based on the instantaneous frequency of the received signal," *IEEE Transaction on Vehicular Technology*, vol. 53, pp. 716–724, 2004 2004.

[103] P. Samar and S. B. Wicker, "On the behavior of communication links of a node in a multi-hop mobile environment," *ACM MobiHoc*, Jun. 2004.

[104] D. Niculescu and B. Nath, "Position and orientation in ad hoc networks," *Elsevier journal of Ad Hoc Networks*, vol. 2, pp. 133–151, April 2004.

[105] R. Narasimhan and D. C. Cox, "Estimation of mobile speed and average received power in wireless systems using best basis methods," *IEEE Transactions on Communications*, vol. 49, pp. 2172–2183, December 2001.

[106] S. Mohanty, "VEPSD: A novel velocity estimation algorithm for next-generation wireless systems," *IEEE Transaction on Wireless Communications*, vol. 4, pp. 2655–2660, November 2005.

[107] T. G. Basavaraju, C. Puttamadappa, M. A. Goutam, S. K. Sarkar, and B. Ramesh, "Doppler location algorithm for mobile ad hoc networks," *in Proc. of Network and Communication Systems*, 2006.

[108] C. Shao, K. D. Mann, and J. C. Olivier, "Mobile speed estimation for TDMA-based hierarchical cellular systems," *IEEE Transactions on Vehicular Technology*, vol. 50, pp. 981–991, July 2001.

[109] B. Alavi and K. Pahlavan, "Modeling of the TOA-based distance measurement error using UWB indoor radio measurements," *IEEE Communications Letters*, vol. 10, pp. 275–277, April 2006.

[110] M. P. Wylie-Green and S. S. Wang, "Robust range estimation in the presence of the non-line-of-sight error," *in Proc. IEEE Vehicular Technology Conference (VTC'2001 FALL)*, vol. 1, pp. 101–105, 2001.

[111] J. McDougall and S. Miller, "Sensitivity of wireless network simulations to a two-state Markov model channel approximation," *in Proc. IEEE Globecom'2003,*, vol. 2, pp. 697–701, December 2003.

[112] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," *in Proc. of the 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pp. 80–91, June 2002.

[113] L.-H. Yen and C. W. Yu, "Link probability, network coverage, and related properties of wireless ad hoc networks," *in Proc. of the 1st IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS)*, pp. 525–527, October 2004.

[114] J. D. C. Little, "A proof for the queuing formula: L= λW," *Operations Research*, vol. 9, pp. 383–387, May-June 1961.

[115] A. Lozano, "Integrated channel assignment and power control in mobile wireless communications systems," *PhD Thesis, Standford University, USA*, November 1998.

[116] A. B. McDonald and T. F. Znabi, "A mobility-based framework for adaptive clustering in wireless ad hoc networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1466–1487, August 1999.

[117] S. Jiang, D. He, and J. Rao, "A prediction-based link availability estimation for routing metrics in manets," *IEEE/ACM Transaction on Networking*, vol. 13, pp. 1302–1312, December 2005.

[118] S. Jiang, Y. Liu, Y. Jiang, and Q. Yin, "Provisioning of adaptability to variable topologies for routing schemes in manets," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1347–1356, September 2004.

[119] T. Liu, P. Bahl, and I. Chlamtac, "Mobility modeling, location tracking, and trajectory prediction in wireless atm networks," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 922–936, August 1998.

[120] Z. R. Zaidi, B. L. Mark, and R. K. Thomas, "A two-tier representation of node mobility in ad hoc networks," *in Proc. of IEEE International Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, pp. 153–161, October 2004.

[121] N. Samaan and A. Karmouch, "A mobility prediction architecture based on contextual knowledge and spatial conceptual maps," *IEEE Transactions on Mobile Computing*, vol. 4, pp. 537–551, November/December 2005.

[122] A. Quintero, "A user pattern learning strategy for managing users mobility in UMTS networks," *IEEE Transactions on Mobile Computing*, vol. 4, pp. 552–566, November/December 2005.

[123] J. M. Franois, G. Leduc, and S. Martin, "Learning movement patterns in mobile networks: a generic approach," *in Proc. of European Wireless (EW'2004), Barcelona, Spain*, pp. 128–134, Feburary 2004.

[124] T. Nakano and T. Suda, "Self-organizing network services with evolutionary adaptation," *in IEEE Transactions on Neural Networks*, vol. 16, pp. 1269–1278, September 2005.

[125] Z. Li, L. Sun, and E. Ifeachor, "Challenges of mobile ad-hoc grids and their applications in e-healthcare," *in Proc. of International Conference on Computational Intelligence in Medicine and Healthcare (CIMED' 2005), Lisbon, Portugal*, July 2005.

[126] I. Widya, A. van Halteren, V. Jones, R. Bults, D. Konstantas, P. A. M. Vierhout, and J. Peuscher, "Telematic requirements for a mobile and wireless healthcare system derived from enterprise models," *Proc. 7th Int. Conference on Telecommunications (ConTEL 2003)*, vol. 2, pp. 527–534, June 2003.