

MOBILITY MODELS FOR AD-HOC NETWORKS

by

Robert J. Neilson

BSc. Simon Fraser University, 1987

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the School
of
Computing Science

© Robert J. Neilson 2004

SIMON FRASER UNIVERSITY
July 2004

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

APPROVAL

Name: Robert J. Neilson
Degree: Master of Science
Title of thesis: Mobility Models for Ad-Hoc Networks

Examining Committee:

Dr. Ramesh Krishnamurti
Chair

Dr. Joseph G. Peters
Senior Supervisor

Dr. Tiko Kameda
Supervisor

Dr. Richard Vaughan
SFU Examiner

Date Approved:

July 30, 2004

SIMON FRASER UNIVERSITY



Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author's written permission.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

In this research we study the effects of constrained node movement and route selection heuristics on the stability of routes in ad-hoc networks. Our results show that the choice of mobility model and the underlying shape of the simulation area have significant impacts on the measured route stability.

Our experiments are done through simulation, which is a common approach for studying routing and performance in ad-hoc networks. However, simulation is only a reasonable experimental technique if the underlying models and assumptions used in the simulation are representative of the scenario being modelled. In the domain of ad-hoc networking, the models typically used to simulate node mobility tend to be primitive and unrealistic. Consequently we introduce the *Constrained Path Mobility Model*, which is used throughout our simulations.

The Constrained Path model is compared against traditional models, indicating that many of the results claimed for ad-hoc routing protocols are simply not achievable in real-world scenarios. Through further investigation and study, we discover that the geometry of the corridors used in our models influences the ability of the routing protocols to discover and select stable routes.

Additionally, we find that heuristics that have been proposed for the purpose of discovering stable routes perform no better than common position-based approaches such as compass routing.

DEDICATION

To my family.

ACKNOWLEDGEMENTS

I would like to thank Joe Peters for his guidance, particularly in the areas of experimental design and communication systems. The many hours of discussion about networks and network research were invaluable in shaping and focussing this work. I would also like to thank Tiko Kameda and Richard Vaughan for their insightful comments and suggestions, which ultimately improved this body of work.

Most importantly, I would like to thank Kathy, Samantha, and Nicole for their continued support and encouragement, and for affording me the privacy, space, and time necessary to undertake this endeavour.

TABLE OF CONTENTS

Approval	ii
Abstract	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
List of Abbreviations	x
List of Symbols	xi
1 Introduction	1
1.1 Motivation.....	2
1.2 Contributions.....	4
1.3 Thesis Outline	5
2 Background	6
2.1 Mobility Models.....	6
2.1.1 Entity Mobility Models.....	8
2.1.2 Group Mobility Models	16
2.2 Routing Protocols.....	19
2.2.1 Associativity Based Routing.....	21
2.2.2 Position-based Routing.....	23
2.2.3 Shortest Path Routing	26
2.2.4 Other Route Selection Strategies	28
2.3 Summary of Related Work	30
3 Experimental Design	32
3.1 Method	32
3.2 Objective Functions	33
3.2.1 Route Availability.....	33
3.2.2 Route Stability	34
3.2.3 Route Length.....	35
3.3 System Parameters and Factors	36
3.3.1 Primary Factors.....	36
3.3.2 Mobility Model Levels	38
3.3.3 Route Selection Heuristic Levels.....	41
3.3.4 Node Density Levels.....	41

3.3.5	Simulation Area Topology Levels	42
3.3.6	Secondary Factors	42
3.3.7	Levels (Secondary Factors).....	47
3.4	Configuration of Experiments.....	47
3.4.1	Mobility Models.....	48
3.4.2	Routing Heuristics	49
3.4.3	Simulation Environment	51
4	Results of Experiments	52
4.1	Node Density: The Primary Factor?	52
4.2	Mobility Models, Routing Heuristics, and Route Availability	55
4.3	Mobility Models, Routing Heuristics, and Route Stability.....	58
4.4	Understanding Mobility Models	65
4.5	The Effect of Simulation Area Topology	66
4.6	Understanding Simulation Area Topology	68
5	Conclusions.....	72
5.1	Summary of Results	72
5.2	Future Work	72
	Bibliography	74

LIST OF FIGURES

Figure 1: Mobility pattern generated by the Random Walk mobility model	8
Figure 2: Mobility pattern generated by the Random Waypoint mobility model	9
Figure 3: Mobility pattern generated by the Boundless Simulation Area model	11
Figure 4: Example topology for a City Section mobility model	12
Figure 5: Example topology for a Graph-Based mobility model	13
Figure 6: Voronoi diagram created from eight location points.	14
Figure 7: Augmented Voronoi diagram showing all destination vertices	15
Figure 8: Advance of a group of nodes in the RPGM model	16
Figure 9: Group motion using the Column mobility model	18
Figure 10: Route stability using the principle of associativity	22
Figure 11: The forward progress of a node in position-based routing strategies	24
Figure 12: Route selection using both compass and greedy approaches	26
Figure 13: Constrained Path Mobility model	39
Figure 14: Constrained Path model showing examples of two different paths.	40
Figure 15: Physical layout of the Constrained Path model used in all simulations	49
Figure 16: Route availability at various densities, as influenced by M and H	55
Figure 17: Route availability with RPGM inter-group and intra-group data series	57
Figure 18: Stability of routes for various combinations of M and H	60
Figure 19: Stability of routes in mobility models	60
Figure 20: Route length in mobility models	60
Figure 21: Route availability versus density for various simulation area topologies	68
Figure 22: Route stability in rectangle topologies	70
Figure 23: Route length in rectangle topologies	70

LIST OF TABLES

Table 1: Objective functions used during this research	33
Table 2: Description of parameters that affect stability in ad-hoc networks.....	36
Table 3: Simulation area topologies used during second phase experiments	42
Table 4: Levels for secondary factors used in the experiments	47
Table 5: RPGM Parameters used in simulations.....	48
Table 6: Levels and Factors for Experiment 1	52
Table 7: Allocation of effects of D , M , and H	53
Table 8: Analysis of variance for initial availability, stability, and length results.....	53
Table 9 Factors and levels used detailed study of availability.....	55
Table 10: Relative effects of M and H on route availability	58
Table 11: Analysis of variance of effects for Route Availability.....	58
Table 12: Allocation of effects of M and H on Route Stability	63
Table 13: Analysis of variance of effects for Route Stability	63
Table 14: Allocation of effects of M and H on Route Length.....	64
Table 15: Analysis of variance of effects for Route Length	64
Table 16: Factors and levels used in study of simulation area topology.....	66
Table 17: Allocation of effects of D , T , H on availability, stability, and length.....	67
Table 18: Analysis of variance of effects of D , T , H on A , S , L	67
Table 19: Levels used for simulations involving topologies and densities.....	68

LIST OF ABBREVIATIONS

ABR	Associativity-Based Routing	21
AODV	Ad-hoc On-Demand Distance Vector	27
CARD	Contact-Based Architecture for Resource Discovery	29
CPM	Constrained Path Model.....	39
DBF	Distributed Bellman-Ford	26
DREAM	Distance Routing Effect Algorithm for Mobility.....	24
DSDV	Destination Sequenced Distance Vector	21
DSR	Dynamic Source Routing.....	9
FRESH	Fresher Encounter Search	30
FSR	Fisheye State Routing	27
GLS	Grid Location Services	24
GPS	Geographic Positioning Service.....	23
LANMAR	Landmark Routing	29
MFR	Most Forward Within Radius.....	25
NFP	Nearest with Forward Progress.....	25
OLSR	Optimized Link State Routing	26
QoS	Quality of Service	1
RABR	Routelifetime Assessment Based Routing	28
RPGM	Reference Point Group Mobility.....	16
RPM	Random Progress Method.....	25
RWD	Random Walk Destinations	19
SSA	Signal Stability Adaptive	28
TORA	Temporally Ordered Routing Algorithm	27

LIST OF SYMBOLS

<i>A</i>	average route availability in a network
<i>b</i>	amount of traffic generated by a node
<i>D</i>	number of nodes used in a simulation
<i>H</i>	routing heuristic used for route selection
<i>L</i>	average route length in a network
<i>l</i>	lifetime of a node
<i>M</i>	mobility model governing the motion of mobile nodes
<i>N</i>	number of nodes in a simulation
<i>p</i>	duration that nodes pause at their destinations
<i>r</i>	transmission range of a mobile radio
<i>S</i>	average route stability level
<i>s</i>	speed of mobile nodes
<i>T</i>	shape and size of the simulation area topology
<i>v</i>	type of communication services used by a node
<i>x</i>	size of the simulation area

CHAPTER 1

INTRODUCTION

Ad-hoc networks are self-configuring networks of mobile hosts (*nodes*) connected by wireless links. Each node may function as a router as well as a host. The nodes, being mobile, are free to move independently – resulting in a dynamic and rapidly changing network topology.

Ad-hoc networks do not rely on any fixed network infrastructure for communication; instead, network nodes organize and cooperate to create a multi-hop routing scheme that facilitates data transmission amongst themselves. This routing scheme must be constantly updated and adapted as the underlying network topology evolves.

Ad-hoc networks are typically built from low cost wireless devices such as notebook computers or handheld devices. These devices often have limited power sources and transmission range, and are free to refuse to relay data if they need to conserve power. Over time, new nodes may join the network, and existing nodes may leave (by powering off or by moving out of range).

Interest in ad-hoc network research is being driven in part by the promise of easily deployable, low cost networks for use in military operations, search and rescue activities, disaster recovery activities (fire, flood, earthquake), shop floor communications, or for general peer-to-peer communication in situations where a wireless Internet access point may not be readily available (conferences, classrooms, libraries).

Yet despite the large amount of ad-hoc network research that has taken place over the past 10 or so years, there are still many challenging problems awaiting researchers. Some of the general problems in need of new ideas and fresh approaches relate to scalability limitations, bandwidth constraints, power consumption, security concerns, and quality of service (QoS) mechanisms.

1.1 Motivation

Our research is motivated by the scalability limitations that are inherent in ad-hoc networks. Scalability refers to the ability of a network to maintain reasonable packet delivery levels as the new nodes join the network, causing the network to grow.

As ad-hoc networks grow, a number of problems emerge:

- There is an increase in data traffic in the network, as each new node presumably wants to participate in one or more data exchanges. This increase in data requires additional bandwidth in networks that are traditionally fairly limited in their bandwidth capabilities.
- There is an increase in the number of possible routes between source and destination node-pairs, all of which need to be established and maintained via control packets; this increase in the control packet overhead reduces the overall bandwidth that is available for data transmission.
- Since there are additional control and data packets traversing the network, the data relaying (routing) responsibilities of many of the nodes is increased, inducing a corresponding increase in power consumption.
- If the physical size of the network is constrained, such as in a room, and the network grows, the overall density of nodes must increase. As the number of nodes that are within range of each other increases, the network will experience a higher number of collisions –translating into end-to-end delay.

Some of these problems, such as power consumption and packet collisions, are best addressed at the MAC layer, which is responsible for efficient and reliable transmission of packets between neighbouring nodes. Other problems, such as efficient use of network bandwidth, are better addressed through routing protocols and effective network layer techniques.

Our interest is in reducing the amount of control traffic that is used to establish and maintain routes. Much work has been done in this area [18, 24, 28, 36, 39], with the typical approach being to develop new protocols that require less overhead to maintain a routing topology. In Section 2.2 we provide an overview of the more important

contributions in this area. In this research we chose to investigate a different approach, one which involves the exploitation of movement patterns exhibited by nodes to create routes that are longer lived (*stable*). By preferring stable routes over less stable ones, we expect to reduce the control packet overhead that is required to maintain routes.

This concept of route stability has been studied by Toh [52], who proposed a new class of ad-hoc routing protocol based on the concept of *associativity*. The concept of associativity states that two nodes have high associativity if they have been connected for a relatively long period of time, as defined by a threshold value A_{Th} . By selecting routes comprised of nodes with high associativity values, the resulting routes are likely to be more stable. Toh's results look very promising, although his study was constrained to a Random Waypoint [28] mobility model, and the evaluation of the performance of his protocol was limited to a comparison against the more traditional ad-hoc routing protocols.

Our research studies the concept of route stability using realistic mobility models and through comparison with routing heuristics rather than established routing protocols. Ad-hoc routing protocols tend to fall into one of two categories, *on-demand* protocols that discover and maintain routes as they are needed, and *proactive* protocol strategies that rely on periodic update messages between nodes to allow them to maintain current routes to other nodes in the network. It has been shown that, in general, on-demand protocols exhibit better overall performance than proactive protocols, as they are able to respond faster to topology changes and they result in less control packet overhead in the network [4, 10, 27]. Therefore, for purposes of our research, we have decided to focus on routing heuristics that are used by the various on-demand routing protocols.

We believe that the use of realistic mobility models is extremely important in the assessment of routing protocol performance. However, we have found that the Random Waypoint and Random Direction models are the most common approaches applied to individual node movement patterns, and Reference Point Group Mobility is commonly adopted to model group movements. These models, all of which are discussed in detail in Chapter 2, tend to be best suited to coordinated group movements (search and rescue, crowd control) or for extremely random situations (rock concerts, parties). These

applications of ad-hoc networks are interesting – albeit somewhat esoteric. Of greater interest to us are models that can be used to study the application of ad-hoc networks to everyday peer-to-peer civilian communications problems such as one might encounter at a university campus or in a small office environment. One of the distinguishing characteristics of these environments is that node movement is constrained to specific pathways and corridors within a building or terrain – a characteristic that we explore through simulation. Of specific interest is whether or not these constrained movement patterns exhibit good route availability and stability properties.

As mentioned in the introduction, ad-hoc networks are complex systems whose behaviour is influenced by a myriad of internal and external factors. In this research we study a subset of these factors to identify those that have the greatest impact on the availability and stability of communication routes between nodes in an ad-hoc network. By identifying and characterising the effects of significant factors, we hope to set the stage for the identification of guidelines and/or heuristics that can be exploited in the development or refinement of future ad-hoc network routing protocols.

1.2 Contributions

Our work compares the ability of routing heuristics to select stable routes in ad-hoc networks. We find that a heuristic that uses node associativity to select the next hop does not result in routes that are any more stable than routes selected by a heuristic that uses a position-based approach, such as the *compass direction* to the destination, as a basis for next hop selection.

We devise a new mobility model for ad-hoc networks – one that realistically constrains node movement to defined paths – and that can be used to model ad-hoc networks inside buildings or cities. Our model is similar to the *Graph-Based* model proposed by Tian et al [51], however our model supports variable width corridors and non-linear paths between destinations. The model, which we call the *Constrained Path Mobility Model*, is compared against traditional mobility models in our study of route availability and stability.

Lastly, we quantify the performance of ad-hoc networks in rectangles of various sizes. This work evolved from our investigation of the Constrained Path model, and clearly illustrates the effect that geographical constraints have on ad-hoc routing.

1.3 Thesis Outline

First, in Chapter 2, we provide a survey of related work, along with the background material that the reader needs to understand the techniques used in our experiments. Next, in Chapter 3, we describe the experiments, providing detailed discussion on the parameters, their levels, and the experimental methods employed. In Chapter 4, we present the results of our experiments, interpret the results, and discuss the associated implications to ad-hoc networking. Lastly, in Chapter 5, we summarize our results and discuss the direction we would like to take this work in the future.

CHAPTER 2

BACKGROUND

2.1 Mobility Models

New approaches to ad-hoc networking are difficult to study empirically. Imagine trying to recruit 100 or more people with wireless devices to participate in a controlled study of a network that may or may not work.

Instead, researchers make use of detailed simulations to study these networks, and to test and quantify potential network improvements. Fortunately there are many simulators available for modelling ad-hoc networks, including *ns-2*, *OPNET*, and *GloMoSim*.

Unfortunately, as Cavin, Sosson, and Shiper [7] show, there are significant divergences in the results produced by the various simulators – even for simple routing simulations like flooding.

In any case, the simulators require the researcher to define the scenario to be modelled by providing details about the nodes in the network, the data to be transmitted, and the mobility model to be used. When a simulation starts, nodes are placed at their assigned positions, and the mobility model computes their movement within the simulation area. For example, if a node is to move along the perimeter of a square, the mobility model will determine the next corner of the square that the node should move to and compute an appropriate direction and speed to this corner. The simulator will use this computed information to start the node moving.

Early ad-hoc network simulations relied on relatively simple mobility models, such as Random Walk. In this model each node is assigned a random direction, and it moves in this direction at a randomly chosen speed. The node stops and selects a new direction and speed after some fixed time interval or distance.

By far the most popular mobility model used in current ad-hoc network simulations is the Random Waypoint model. This model is similar to Random Walk, except that nodes

choose a destination rather than a direction, and they stop and spend some random amount of time at each destination they reach. It is the opinion of this author that the model is popular only because it is readily available with the open-source *ns-2* simulator. The reality is that Random Waypoint is not much better than Random Walk, as both require that nodes follow straight line paths to their intended destinations, travelling through any walls or obstacles that happen to be in the way. Furthermore, a study by Yoon, Liu, and Noble [56] shows that Random Waypoint produces unreliable results as it fails to reach a steady state in terms of instantaneous average node speed. This problem calls into question many of the results that are based on the Random Waypoint model.

Only recently have researchers sought to develop more realistic models [21, 26, 51] for use in their simulations. For a model to be realistic it needs to consider the characteristics of the scenario that is being modelled. For example, assume that we want to model a search and rescue operation. In this scenario nodes move independently of one another, but the nature of the movement of the participants is likely coordinated in some manner. Nodes are definitely not migrating towards random destinations or walking off in random directions. Certainly there will be some degree of randomness in the manner that each node moves, but this randomness is relatively small and is likely to be local to each node. Furthermore, nodes are likely to encounter obstructions and obstacles that constrain their movement, and around which they must navigate. Nodes will also change speed and direction regularly to adapt to their environment and terrain, and to accommodate any activities they may be performing as they move.

Later in this survey we will see that these characteristics are simply not addressed by the majority of mobility models. In particular, the constraints placed on movement and the ability for nodes to have independent local movement are rarely considered by the models.

In general, existing models can be categorized either as entity models, which attempt to model the movement of autonomous nodes within the simulation area, or group models, which attempt to model the coordinated movement of groups of nodes. The remainder of this chapter gives an overview of the more common entity and group

mobility models. For a complete discussion of other not so common models, we refer the reader to an excellent survey paper by Camp, Boleng, and Davies [5].

2.1.1 Entity Mobility Models

2.1.1.1 Random Walk

The Random Walk mobility model was discussed briefly in Section 2.1. It is a widely used model that is relatively simple to implement and places few constraints on the nodes. It is worth noting that the movement pattern resulting from this model is occasionally referred to as Brownian motion in ad-hoc network literature.

The model makes use of three constants, *minspeed* and *maxspeed*, which bound the range of speeds of the nodes, and *t*, which defines the amount of time that nodes travel before changing direction. Prior to the start of the simulation, the nodes are placed randomly about the simulation area and assigned an initial speed and direction. The speed is randomly chosen from the range [*minspeed*, *maxspeed*], and the direction is randomly chosen from the range $[0, 2\pi]$.

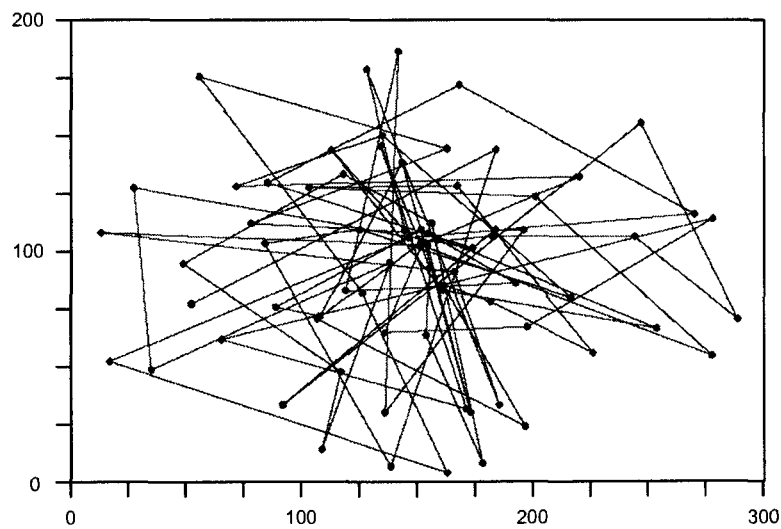


Figure 1: Mobility pattern of a node using the Random Walk mobility model, where distance moved is based on t .¹

When the simulation starts, the nodes begin moving (walking). When the simulation time reaches t , new speed and direction values are computed for each node, and the walk cycle

¹ Figure adapted from [5], pp. 4.

repeats. An alternative model uses a distance d instead of t . In this case the nodes continue moving in their prescribed direction until they have travelled the specified distance d .

If a node's movement causes it to reach the boundary of the simulation area, the node's direction is changed as if it "bounced" off the perimeter. Figure 1 illustrates possible node movement using the Random Walk mobility model.

This model's use is fairly limited as it is really only suitable for modelling scenarios in which the mobile nodes exhibit erratic and unpredictable movement patterns.

2.1.1.2 Random Waypoint

The Random Waypoint model was introduced by Johnson and Maltz in their study of Dynamic Source Routing (DSR) in Ad-hoc Networks [28]. In this model the nodes are initially placed randomly or assigned to pre-defined positions inside the simulation area.

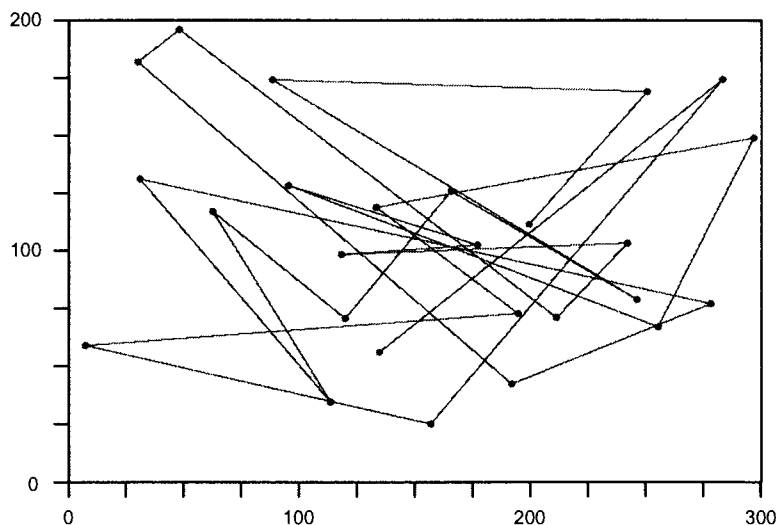


Figure 2: Mobility pattern of a node using Random Waypoint mobility model.²

When the simulation starts, nodes are assigned random destinations within the simulation area. Each node then moves towards its destination at a speed that is uniformly distributed within the range $[minspeed, maxspeed]$. When a node reaches its destination it pauses for a time interval p and is assigned a new destination and speed. Once the time

² Figure adapted from [5], pp. 5.

interval p elapses the node begins travel towards its new destination. This cycle repeats until the simulation ends. Figure 2 depicts node movement patterns using the Random Waypoint model. When comparing the motion of Random Walk (Figure 1) with the motion shown for Random Waypoint (Figure 2), we notice that the Random Walk model tends to keep the node centred roughly around its initial starting position, whereas the Random Waypoint model allows the node to roam more freely about the entire simulation area.

The pause time p is intended to represent the time that nodes spend at their destinations before moving on. The concept of destinations was added based on the assumption that nodes are typically heading somewhere, and not just walking aimlessly about. However, the model is still very primitive and is limited in the types of real-world scenarios that it can model.

The performance of the Random Waypoint model has been extensively studied and documented in the literature, most notably in [4, 29, 56].

2.1.1.3 Boundless Simulation Area

The Boundless Simulation Area model allows nodes to move freely about the simulation area. Instead of bouncing off the simulation area boundaries, nodes continue their motion on the opposite edge. This creates a simulation area topology in the shape of a torus.

Each node's motion is continuous, with frequent speed and direction changes being made at a constant interval Δt . Changes to the speed v and direction Θ are functions of the current speed and direction, as follows

$$v(t + \Delta t) = \min[\max(v(t) + \Delta v, 0), V_{max}],$$

$$\Theta(t + \Delta t) = \Theta(t) + \Delta\Theta$$

Where V_{max} = the maximum velocity of a node
 A_{max} = the maximum acceleration of a node
 α_{max} = the maximum angular change in direction of a node
 Δv = a random number in the range $[-A_{max} * \Delta t, A_{max} * \Delta t]$
 $\Delta\Theta$ = a random number in the range $[-\alpha_{max} * \Delta t, \alpha_{max} * \Delta t]$

The resulting movement pattern avoids the abrupt and sudden changes in direction and speed that occur in the Random Waypoint and Random Walk models. Combining this movement pattern with the unbounded simulation area results in continuous, unobstructed, not overly chaotic, node movement. This is illustrated in Figure 3.

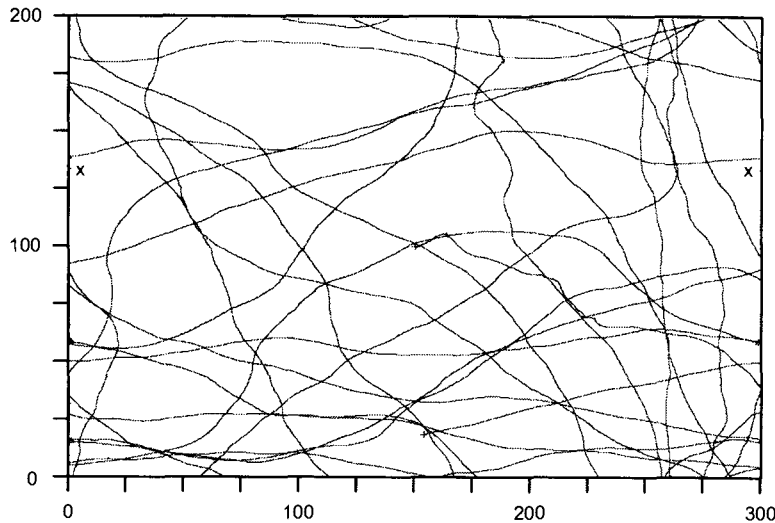


Figure 3: Mobility pattern of a node using the Boundless Simulation Area mobility model.³ Traces continue on opposite sides of the simulation area, as shown by the x symbols.

2.1.1.4 City Section

The City Section model is intended to model vehicular traffic along streets. There are streets, intersections, and speed limits. The streets are bi-directional, and there are no restrictions on the number of vehicles on any street. Vehicles do not interfere with one another, and there is no synchronisation or control at intersections. The intent is for the model to be constructed by the researcher such that it represents an actual section of a city.

The model is constructed using a planar graph in the form of a grid to represent a section of a city. The edges of the graph are straight lines, which represent the streets that vehicles can travel down. The vertices of the graph are intersections or corners of streets. The edges are assigned a length l and a speed s . Nodes moving along an edge are constrained to speed s . The graph is assumed to be connected.

³ Figure adapted from [5], pp. 9.

Initially, nodes (vehicles) are placed at random starting positions on the edges of the graph. When the simulation begins, the nodes are assigned random destinations (points on the edges of the graph), and the nodes move towards their destinations. Since there may be multiple routes to any destination, each node selects the route that minimizes its travel time.

When a node reaches its destination it pauses for a constant time interval, is assigned a new destination, and repeats the movement process.

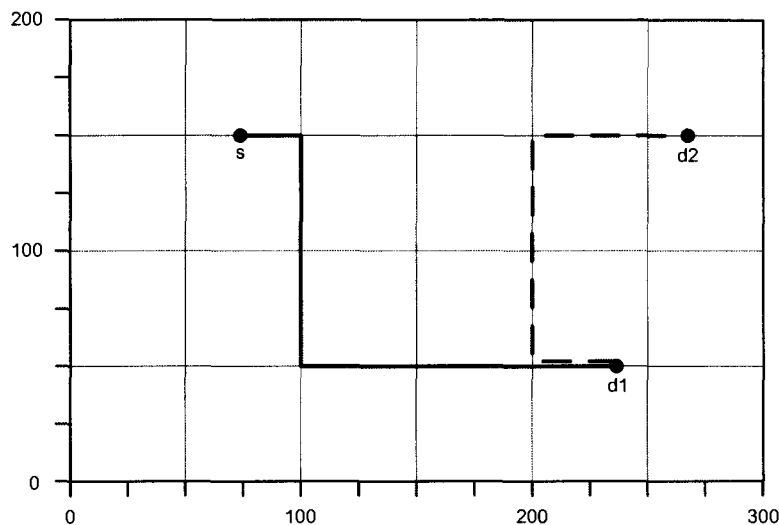


Figure 4: Example topology for a City Section mobility model. *s* represents the initial starting location of a node. The points *d1* and *d2* represent successive destinations. The path taken is the shortest in terms of time.

2.1.1.5 Graph-Based

The Graph-Based mobility model [51] is a variation of the City Section mobility model. It uses an arbitrary graph instead of a grid to represent the region being modelled. The model assumes that the vertices of the graph are the destinations, and the edges are the paths between destinations. The model can be used for vehicular or pedestrian traffic.

Initial node positions and destinations are randomly selected from the set of graph vertices. Each edge of the graph is assigned a weight, which represents the length of the edge. All paths are computed such that the distance travelled is minimized.

As with the City Section model, node movement is constrained to the edges that connect the destinations. This model provides for a reasonably accurate representation of

the pattern of movement that one would expect in a real-world scenario where the scale of the simulation is fairly large, such as a section of a city.

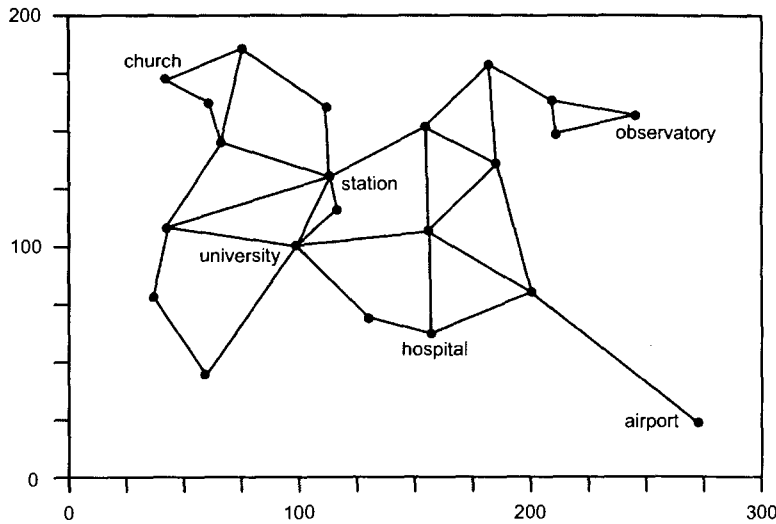


Figure 5: Example topology using the Graph-Based mobility model.

2.1.1.6 Obstacle

The Obstacle mobility model [26] is the most sophisticated of the graph type models surveyed. The authors were looking to construct a model that could depict movement scenarios encompassing obstacles and physical obstructions. The intuition is that people do not typically travel through obstructions in a random manner, rather, they use entrances such as doors or walkways. The authors of the Obstacle model argue that doors are typically located near the centres of the walls of buildings, not near the corners, and they reflect this in the model. Furthermore, they contend that the actual pathways taken by people are dictated by the obstacles they must navigate around as they migrate between destinations.

The Obstacle model assumes that radio signals cannot pass through obstacles. Nodes can only communicate with one another if they are within range, and, if they are both inside the same building or outside of buildings, and, if they are within line of sight of each other.

To construct a model of a scenario using the Obstacle approach, the researcher defines and places obstructions in a rectangular simulation area. This is accomplished by specifying the locations of the corners of each obstruction. These corners, or obstacle

vertices, are then used to construct a Voronoi diagram [9]. The edges of the Voronoi diagram become the pathways that may be traversed by nodes during the simulation.

A Voronoi diagram is defined by:

Assume a set of n sites $P = \{p_1, p_2, \dots, p_n\}$ in a two-dimensional Euclidian plane. The Voronoi diagram of P is a partition of the plane into n convex cells with one cell per location point, such that every point in a cell is closer to its location point than to any other location point.

For example, consider Figure 6. In this figure we have two squares, with eight location points (the vertices of the squares). The resulting Voronoi diagram creates eight regions, each containing one of the location points, as shown. In Figure 6 the location points are labelled $l1 \dots l8$, and the vertices of the resulting Voronoi diagram are labelled $v1 \dots v3$.

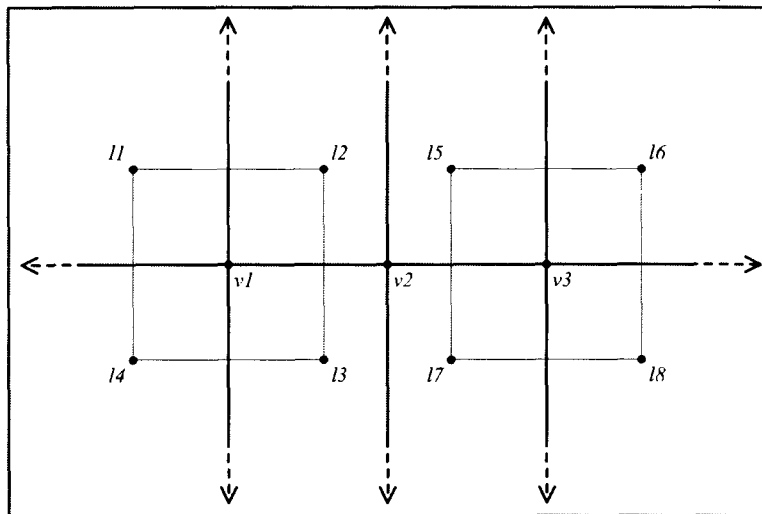


Figure 6: Voronoi diagram created from eight location points.

Once the Voronoi diagram has been constructed, it must be augmented before it can be used as the basis for a mobility model. Specifically, there is a need to define the destinations between which nodes travel, and to constrain the Voronoi diagram to the finite dimensions of the simulation area.

To account for the finite simulation area, all edges of the Voronoi diagram are *clipped* at the points where they intersect the simulation area boundary. These intersection points are added to the mobility model as additional graph vertices. Similarly, the points where

the edges of the Voronoi diagram intersect the obstacles are added as graph vertices. These intersections of obstacles and graph edges are intended to represent the doorways or entrances into buildings.

Figure 7 shows an augmented Voronoi graph with points $v1 \dots v6$ being the original vertices of the Voronoi diagram, points $c1 \dots c6$ indicating the vertices created by the clipping of the edges, and points $d1 \dots d6$ indicating the vertices created by the intersection of obstacles and Voronoi graph edges.

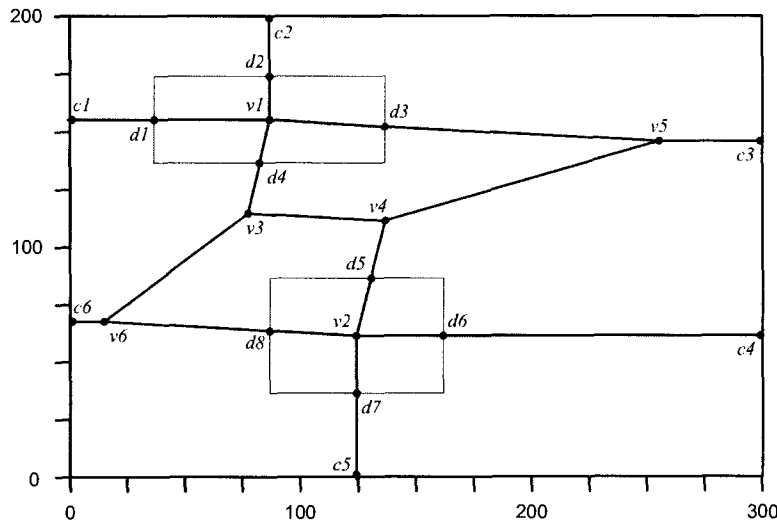


Figure 7: Augmented Voronoi diagram showing all destination vertices. All vertices are possible destinations for nodes. Nodes follow paths created by the edges of the graph.

For simulation purposes, node destinations are selected from the set of all vertices in the augmented Voronoi diagram. This includes the doorways, the ends of the clipped edges, and the original Voronoi vertices. The edges of the cells define the paths that the nodes traverse during the simulation.

Before the simulation starts, the sets of destinations (vertices) and pathways (edges) are calculated, and nodes are placed randomly at destinations. Once the simulation is under way, the nodes are each assigned a random destination. They travel towards their destinations at randomly selected speeds, taking the path of shortest distance.

2.1.2 Group Mobility Models

2.1.2.1 Reference Point Group Mobility

The Reference Point Group Mobility (RPGM) [21] model was designed to model scenarios that involve coordinated group movements, such as search and rescue or disaster recovery operations. As it turns out, the model is extremely flexible, and through clever configuration can be adapted to model many types of group as well as individual node movements.

The model is based on a hierarchical relationship of nodes and groups in which a group consists of one or more nodes, and a node can belong to only one group. Each group has a logical centre, which defines the movement (direction, speed, pause time) of all nodes in the group. The motion of the logical centre of each group is represented by a group motion vector, \overrightarrow{GM} , which is either pre-defined or assigned randomly.

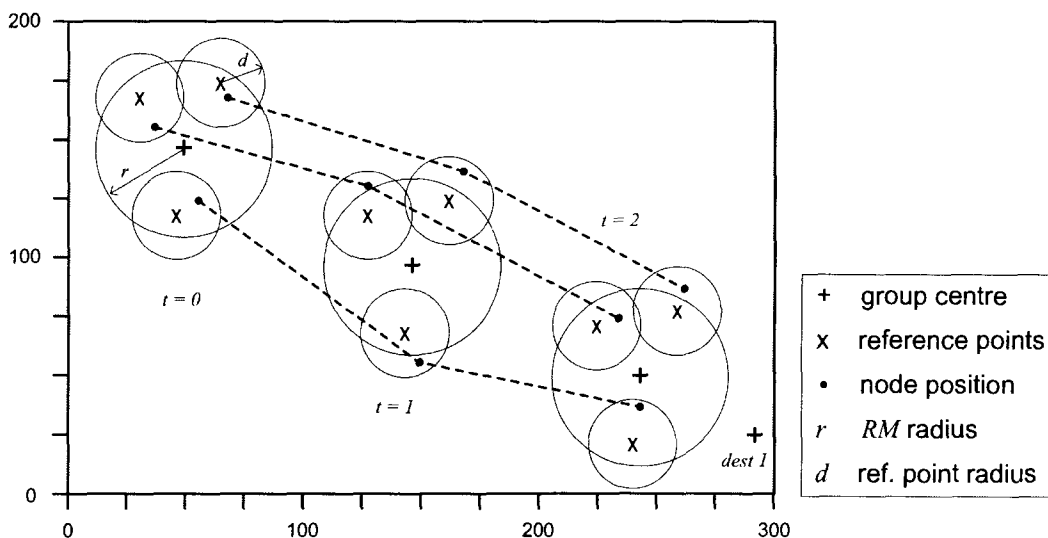


Figure 8: Advance of a group of three nodes over two time intervals in the RPGM model. When the group centre reaches *dest1*, the group will pause and a new destination will be selected. The radii r and d can be defined for each simulation to allow variation in node separation and random node movement.

Each node is assigned a logical reference point, the position of which is fixed relative to the group centre. Typically, the reference points for a group are uniformly distributed within some distance d from the group centre before the simulation begins, however they could be assigned predetermined positions. As the simulation progresses, nodes move

randomly within a small region around their assigned reference points. They move according to a random motion vector \overline{RM} which has a length that is uniformly distributed within some radius r centred on the reference point, and a direction that is uniformly distributed within the range $[0, 2\pi]$.

To coordinate the smaller individual node movements with the overall group movement, a constant time interval τ is assigned. During an interval τ , each reference point is advanced from its current position $RP(\tau)$ to its next position $RP(\tau+1)$, according to the group motion vector, \overline{GM} . Next, each node is assigned its new position by adding its random motion vector \overline{RM} to the new reference point $RP(\tau+1)$. After repeating this process enough times, the entire group will have advanced to its destination. Figure 8 shows the advancement of a group of three nodes over two time intervals.

By strategically assigning values to model parameters such as r , group size, reference point location, and \overline{GM} , all of the other group models presented in this chapter can be constructed. RPGM can also be used to construct some entity models. For example, the Random Walk model can be created by assigning one node per group and randomly generating \overline{GM} .

2.1.2.2 Column

The Column mobility model [45] was created to represent scenarios where people move in lines, such as marching soldiers, or people searching an area. An initial reference grid is defined by placing reference points (one per node) at successive offsets from an initial position, forming a line.

Nodes move randomly around their reference point (RP) according to a random movement vector \overline{RV} , using a Random Walk entity model that incorporates a constraint on the region the nodes can walk in (to ensure that the node position (NP) remains in the immediate vicinity of the reference point).

The reference grid moves according to an advance vector \overline{AV} , which defines the angle and distance to advance. The angle component of \overline{AV} is assigned randomly from the range $[0, \pi]$, which ensures that the motion is always in a forward direction. The

distance component of \overline{AV} is assigned randomly subject to the constraint that all reference points remain within the simulation area.

Reference points move to their new positions according to

$$RP(t+1) = RP(t) + \overline{AV},$$

and nodes move to their new positions according to

$$NP(t+1) = NP(t) + \overline{AV} + \overline{RV}$$

Figure 9 depicts a column of 5 nodes moving to a new location.

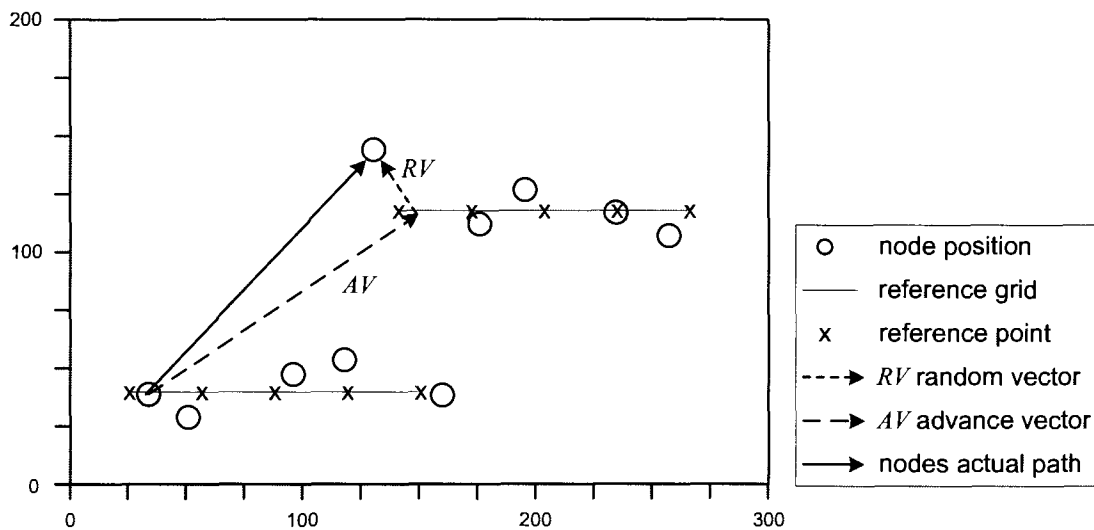


Figure 9: Group motion using the Column mobility model. The reference grid is advanced using AV , nodes follow the grid, with a small randomness in their path, as defined by RV .

The Column mobility model is not rigorously specified in the literature, and therefore it is left to the researcher to incorporate implementation decisions appropriate to the scenario being modelled. For example, the procedure for selecting destinations is not specified, nor is the behaviour of the group when a destination is reached. A reasonable assumption would be that once a new destination is assigned, the reference grid will be realigned to the direction of travel.

If destinations are not to be used, the researcher must define the behaviour when the grid reaches the simulation area boundaries.

2.1.2.3 Nomadic Community

The Nomadic Community mobility model [5, 45] allows groups of nodes to roam randomly around a single reference point, and to follow the reference point as it migrates between locations.

Initially, a reference point is established and all nodes start to move within the vicinity of the reference point using an entity model such as Random Walk. To keep the community together, the Random Walk destinations *RWDs* are limited to a radius r around the reference point. After a random time interval t has elapsed, the reference point is moved to a new randomly chosen position within the simulation area. Subsequent *RWDs* are assigned such that they are within r of the new reference point position, thereby forcing the nodes to “follow” the reference point.

This model may be useful for scenarios where groups of people are engaged in similar activities, and where the people in a group act independently – such as organized tours or recreational activities such as scuba diving.

2.2 Routing Protocols

The routing protocol’s role in any communication network is to define and maintain a strategy for sending messages using the nexus of paths that interconnect the nodes in the network. The resulting strategy should be efficient in terms of the network resources used, it should be sufficiently reliable so that messages have a reasonable chance of reaching their destination, and it should be robust so that it can adapt to changes in network topology – such as when links go down or when new nodes join the network.

These characteristics of routing protocols hold true for fixed-line networks, for traditional wireless networks, and for ad-hoc wireless networks. However, ad-hoc networks present routing protocol designers with unique challenges. First and foremost, ad-hoc networks are not subject to any centralized control. Each ad-hoc network is self-organizing, with no fixed infrastructure such as centralized naming services, location services, or registries which could be used to distribute routing information to the nodes.

Since ad-hoc networks are subject to rapidly changing topologies, the neighbour-set of a node is likely to change frequently, requiring the routing protocol to be highly

adaptive. These rapid topology changes result from the mobility aspects of the network as well as the poor reliability of the low power wireless links. Contrast this with a fixed-line network where neighbour-sets are likely to change only when a router fails or is taken offline for upgrades or maintenance. This constantly changing neighbour-set and the lack of centralized control have motivated researchers to invent new types of routing protocols for use in ad-hoc networks.

And invent they have! There have been so many ad-hoc network routing protocols proposed that it becomes difficult to keep track of them all. Despite the diversity of approaches, all reasonable protocols include three common phases:

1. Route Discovery

Prior to the route discovery phase, the routing protocol has no knowledge of the route from a particular source to a particular destination. The protocol must discover a route or set of routes to the destination. This is typically accomplished by flooding “where are you?” messages onto the network, and waiting for responses. Since flooding is extremely inefficient, many researchers have focussed their work on devising efficient strategies for route discovery [18, 24, 33, 36, 48, 55].

2. Route Selection

Once the routing protocol has identified a set of routes to the destination, it must decide which route to use. This necessitates a metric or heuristic that can be applied to ensure that the “best” route is selected. Since most of the popular ad-hoc routing protocols have their ancestry in fixed-line counterparts, the metric of choice is *shortest path*.

This appears to be particularly short-sighted, as we know that the purpose of selecting shortest paths in fixed-line networks is to reduce congestion and delay, thereby reducing the overall cost of forwarding a packet. Since the links are relatively stable, it makes sense to send packets along whichever route has the lowest cost.

However this simply does not hold true for ad-hoc networks. Consider the scenario in which a link in the shortest path breaks, and the routing protocol needs to resort to flooding to find an alternate route. The cost of this recovery operation will likely be much higher than if the protocol had selected a more stable, longer route in the first place.

Recently researchers have become aware of this inefficiency, and have been proposing protocols that make use of alternative route selection heuristics [12, 15, 20, 29, 52].

3. Route Maintenance

As with fixed-line protocols, a compelling alternative to re-discovering complete routes is to provide a mechanism by which existing routes can be “repaired” when they are affected by topology changes. All ad-hoc protocols do this to some extent, although some more than others. For example, the Dynamic Source Routing (DSR) [28] protocol discovers routes only as they are needed, and it exploits the nature of the 802.11 MAC layer to effect localized repairs and optimizations of links. In contrast, the Destination Sequenced Distance Vector (DSDV) protocol [42] relies on constant updates of neighbour information between nodes to maintain localized routing tables in each node.

Obviously any level of route maintenance induces a cost on the network, reinforcing our belief that proactive identification of stable routes during the route selection phase is critical to efficient protocol operation.

The next sections provide a brief overview of ad-hoc routing protocols, and discuss the more common techniques used for route selection.

2.2.1 Associativity Based Routing

The goal of Associativity Based Routing (ABR) [52] is to prefer long-lived routes over short-lived ones. ABR is motivated by the intuition that once a link has been in existence for some period of time, the nodes connected by that link are likely to be moving with similar speed and in a similar direction, and therefore are likely to remain in connected

for a relatively long period of time. The period of time that two nodes remain connected is used to determine the *associativity* of the nodes.

A node's *associativity count* with its neighbour changes over time, and is measured in units called *ticks*. ABR uses the associativity count in conjunction with a threshold A_{Th} to determine stable links, where a link is said to be stable if its associativity count is greater than A_{Th} . A_{Th} is an implementation-specific constant that is derived from other system parameters. For example, if $tick=1\text{sec}$, the transmission range of a node is 30 meters (radius), and the minimum speed is 3 m/sec, then A_{Th} would be set to 20. This value indicates the maximum time that a stationary node and a moving mode could be in contact.

When two nodes first move within range of each other, their associativity count is 0. The count remains at this level until they have been in contact for A_{Th} seconds, after which time the associativity count is increased by 1 for each subsequent interval *tick* that they remain in contact. To ensure that neighbours are aware of their presence, nodes periodically transmit *beacons* identifying themselves (one beacon every *tick* seconds). Associativity counts are reset to zero when nodes move out of range of each other.

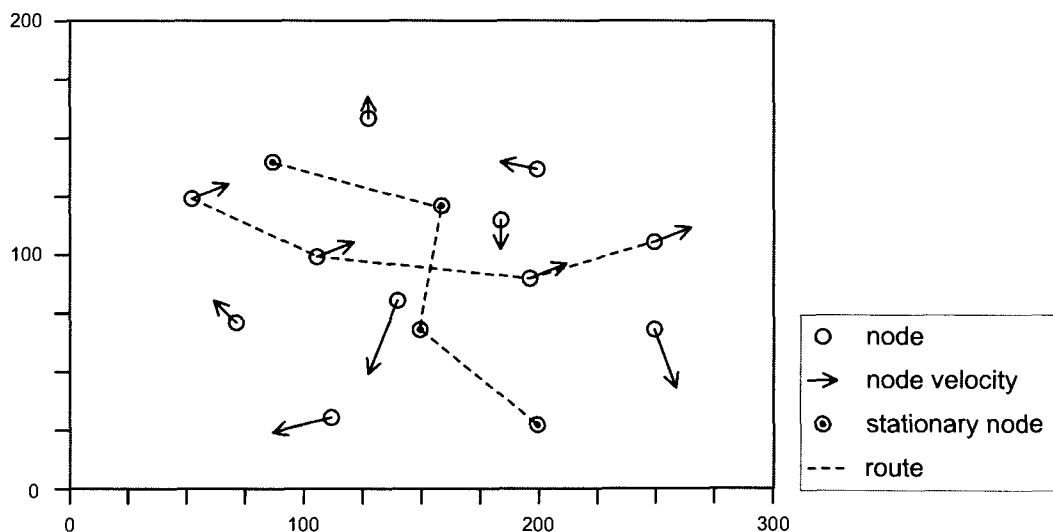


Figure 10: Two stable routes, based on the principle of associativity. In one route the nodes are stationary and in the other the nodes are travelling with identical velocity. The path will remain stable as long as no node makes significant changes to its speed or direction.

Figure 10 depicts two routes with stable links. One is comprised of nodes that are mobile, and the other is comprised of nodes that are stationary. The vectors in the diagram indicate the relative speed and direction of the nodes.

ABR is an on-demand protocol, whereby routes are discovered only when they are required. When a node s requires a route to another node d , it broadcasts a route request message to its immediate neighbours. If a recipient n of the request has previously seen the same request, n drops the request. Otherwise n appends its address, associativity count (for its link with the previous hop), as well as some information regarding the link delay and forwarding load. This modified request is then broadcast to all neighbours of n . This process continues until the request reaches d .

Once d receives a request, it waits some amount of time (to allow requests that followed other paths to arrive), and then selects the best path. The path selection algorithm looks for the path that has the highest aggregate associativity count H_A that does not violate any maximum forwarding load levels defined for the system. This additional forwarding level constraint is introduced to ensure that individual nodes do not assume excessive packet forwarding responsibilities. In the case where there are multiple routes with equal H_A , ABR chooses the route with lowest hop count. The route request for the selected route is then unicast back to s (the originator) via the selected route. Intermediate nodes update their routing tables with the selected route as they relay the route request back to s .

ABR also defines a route maintenance scheme through which broken routes can be reconstructed. This scheme uses the associativity counts of disconnected nodes to select a new partial route to reconnect them.

2.2.2 Position-based Routing

Position-based routing strategies assume that each node is aware of its geographical position in the network. This position could be absolute, such as might be provided by a Geographical Positioning Service (GPS), or it could be a position that is relative to the other nodes in the network. In [6], Capkun, Hamidi, and Hubaux describe a distributed positioning system that does not rely on fixed infrastructure or GPS.

To forward packets, nodes are not required to maintain detailed information about the positions of the other nodes in the network. Instead, position-based forwarding strategies require only that each node be aware of its own position, the positions of its one-hop neighbours, and the position of the destination. Forwarding is performed in a hop-by-hop manner, with each node making local forwarding decisions based on the above-mentioned position information and a predefined heuristic.

One of the challenges of position-based approaches is the need for efficient *Location Services*. Consider a node n_1 that needs to send a message to a node n_2 . If n_1 does not know n_2 's position, it will request it from the location service. The location service is aware of all the nodes positions through a registration process. This process requires that each node register its current position with the service on a regular basis. In ad-hoc networks, location services are distributed throughout the nodes in the network. Distance Routing Effect Algorithm for Mobility (DREAM) [2], Grid Location Services (GLS) [32], as well as homezone [47], and quorum approaches [17] are examples of distributed location services that have been proposed.

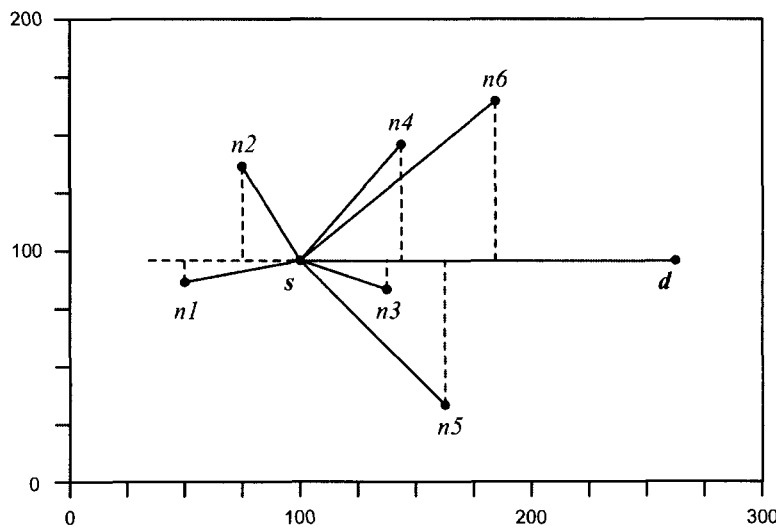


Figure 11: The forward progress of a node is the distance travelled along the straight line path from s to d . Node n_6 has made the most forward progress, and n_3 the least. Nodes n_1 and n_2 have made backward progress.

Armed with the position information described above, nodes make local forwarding decisions based on some heuristic. Most of the common position-based forwarding

heuristics are derived from the principle of *forward progress*, which is illustrated in Figure 11.

The Most Forward Within Radius (MFR) [49] algorithm requires that nodes make forwarding decisions such that the forward progress of the packet is maximized. If a node has no neighbours that would result in forward progress, the packet is returned to the previous sending node which selects its next best hop. If the node is returned all the way to the originator, it is assumed that no routes exist. The approach is loop free, and results in a reasonable hop-count, but it will not find routes that require the packet to make backwards progress.

The Random Progress Method (RPM) [15] takes a different approach. RPM assumes that the transmission power of a node can be adjusted, and that the probability of collisions is higher as the distance between neighbours increases. To account for this, RPM considers all neighbours that result in forward progress, randomly selecting the neighbour to use for each packet transmission. This strategy reduces the probability of collisions while still guaranteeing forward progress whenever such progress is possible.

In Nearest with Forward Progress (NFP) [22], the packet is sent to the *closest* neighbour that makes forward progress, and the transmission power is adjusted to the minimum required to send to this neighbour. This results in even lower levels of collisions than RPM.

Compass routing [31], illustrated in Figure 12, stipulates that a node n_l forward the packet to the forward progress neighbour n_i that is closest to the straight line between n_l and the destination d . More precisely, the next hop n_i is chosen such that the direction from n_i to d is closest to the direction from n_l to d , for all neighbours of n_l . This process is repeated until the destination or a local maximum is reached. Note that compass routing stipulates that a local maximum exists when there is no neighbour through which the packet can make forward progress. In this case it is assumed that there is no path to the destination.

A variety of greedy schemes, which are surveyed in [15], have also been proposed. A typical greedy scheme is the one proposed by Finn [13]. This scheme requires that a node n_l forward packets to the next hop that is *closest* to the destination. If there is no

neighbour that is closer to the destination than the $n1$, a search with radius h hops is performed to find a node that is closer to the destination than $n1$.

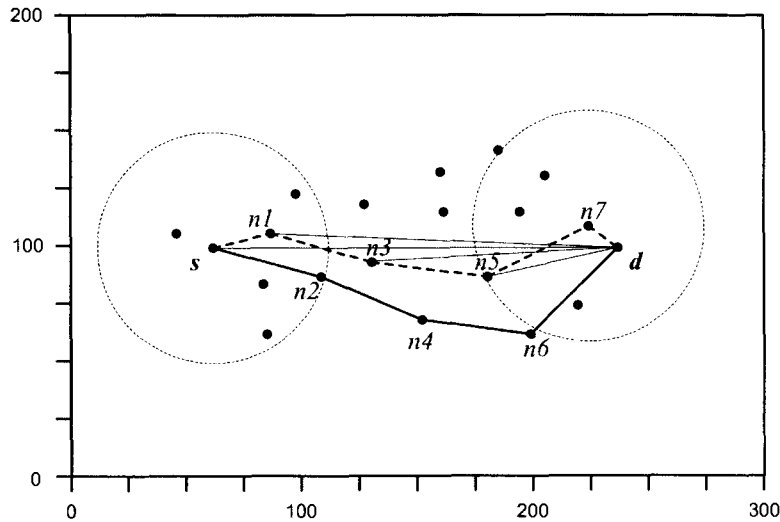


Figure 12: Comparison of route selection using both compass and greedy approaches. The route $n1, n3, n5, n7$ was selected using the compass method, while $n2, n4, n6$ was selected using the greedy approach.

2.2.3 Shortest Path Routing

In shortest path routing strategies the cost of a route is measured as its total hop count. Shortest path routing protocols attempt to discover and select routes that minimize this cost. Traditional multi-hop wireless networks often made use of the Distributed Bellman-Ford (DBF) [50] algorithm to find the shortest paths between nodes. Unfortunately this algorithm is known to suffer from slow convergence problems [46], and attempts to mitigate the problem in ad-hoc networks have had limited success [43].

The slow convergence problems associated with DBF type algorithms were addressed in fixed-line networks through the adoption of *link state* protocols such as OSPF [35] and IS-IS [37]. These protocols require that each router maintain a full or partial map of the network. When the state of a link changes, an advertisement is flooded throughout the network, and affected routers update their maps. A similar approach was proposed for ad-hoc networks. It is the Optimized Link State Routing (OLSR) [8] protocol, and it utilizes a *multi-point relay* mechanism to reduce the overhead caused by topology updates. Multi-point relay involves selecting a subset of nodes to distribute updates through, similar to a

multicast tree. In OLSR, each node creates a minimum set of multi-point relay nodes such that the set covers all nodes that are two hops away. Unfortunately the sheer volume of topology updates prevent OLSR from scaling well to larger ad-hoc networks [23].

The shortest path protocols that have proven effective in ad-hoc networks tend to be of the on-demand, or *source-initiated*, variety. These include protocols such as Dynamic Source Routing (DSR) [28], Ad-hoc On-Demand Distance Vector (AODV) [43], and Temporally Ordered Routing Algorithm (TORA) [39]. The underlying characteristic of these protocols is that they do not attempt to maintain up-to-date routes to all nodes in the network, as this has proven to be too costly. Instead they rely on the flooding of queries to discover new routes as they are needed. In the case of DSR, a node requiring a route broadcasts a query, placing its own address and the address of the destination in the query header. Each node that receives the query checks its local cache to see if it knows of a route to the destination. If it does not know of a route to the destination, it appends its own address to the header and re-broadcasts the query. If it has a route, it appends the route to the header and unicasts a route reply message back to the initiator. As the reply is relayed back to the originator, many nodes will “hear” the reply and will add the route to their local route cache. To send an actual data message along a known route, the sender uses source-initiated routing. This involves adding the entire predetermined route to the header of each packet. Intermediate relay nodes simply get the address of the next hop directly from the packet header, and forward the packet.

AODV works in a similar manner, although it does not rely on source-initiated routing. Instead, next-hop information for discovered routes is stored at each node, and is used for packet forwarding. AODV uses flooding of requests and unicast replies to discover routes. In contrast, TORA employs a strategy that relies on the controlled flooding of route reply messages. This results in a directed acyclic graph of known routes that is rooted at the destination.

Fisheye State Routing (FSR) [40] is another technique that adopts the shortest path metric for route selection. FSR employs a mechanism proposed by Kleinrock and Stevens [30] to reduce the amount of data required to view graphical information. In this context, the information nearer the centre of the image is presented in great detail, while the

information nearer the periphery contains less detail. The generalized notion is that the further you move from the point of interest, the less data is needed. This concept is applied to ad-hoc routing by selectively distributing routing information to nodes based on how far away they are. To accomplish this, FSR defines a number of *scopes*, and assigns nodes to scopes based on the number of hops they are from the source of the update. Nodes that are in closer scopes receive updates more frequently, and nodes that are further away receive updates less frequently. Messages containing data are sent between nodes using the only route that is known to the originator, as the originator assumes this to be the shortest path. As the message nears the destination the route is refined with the more accurate routing information. This approach succeeds in significantly reducing the overhead of control message updates. The biggest drawback of FSR occurs in networks with high mobility rates. In these cases the routes to far away destinations quickly become stale and need to be updated.

It is worth noting that many of the above mentioned protocols could use heuristics other than shortest path during the route selection process. AODV, for example, claims to support QoS via a mechanism that allows it to discard routes according to user-defined criteria, such as bandwidth or signal strength. In practice this feature is ignored in favour of shortest path route selection.

2.2.4 Other Route Selection Strategies

2.2.4.1 Signal Strength Approaches

Signal Stability Adaptive Routing (SSA) [11] is similar to ABR in that it attempts to distinguish *strongly connected* links from *weakly connected* links, and to use this heuristic to assist in route selection. A link is said to be strongly connected if it has been in existence for a minimum time t .

Another approach based on signal strength is Routelifetime Assessment Based Routing (RABR) [1]. In this approach, nodes monitor the signal strength of messages transmitted by their neighbours and attempt to predict when the signal strength will drop below a critical level. In this manner the expected lifetime of each link can be estimated, and the link with the longest projected lifetime can be preferred.

2.2.4.2 Hierarchical Routing Schemes

A variety of hierarchical routing schemes exist for ad-hoc networks [18, 24, 40]. These strategies are built upon the concept of clusters of nodes. Each cluster selects a representative as a *clusterhead*, and charges this clusterhead with distributing packets to nodes that are located outside the cluster. The clusterheads themselves may be arranged hierarchically to achieve an even higher degree of scaling.

The motivation for this strategy is two-fold. First, it reduces the overhead of distributing routing information to all nodes, and secondly, it allows for spatial reuse of radio channels by reducing interference between nodes that would otherwise attempt to propagate the same messages.

Hierarchical routing schemes do not result in shortest path routes, as messages are routed through the clusterheads which cannot provide shortest paths to and from all nodes within their respective clusters. In these schemes routes are not so much selected as they are imposed, and path stability may or may not be taken into consideration by the clusterhead selection algorithm.

2.2.4.3 Landmark Routing

In [41], Pei, Gerla, and Hong propose the Landmark Routing (LANMAR) strategy. LANMAR is intended for use in group mobility scenarios where the members of each group stay together most of the time. Within a group, nodes use FSR to manage their internal routing. Routing between groups is managed via a logical layer of *landmarks*. Group of nodes are uniquely identified by a subnet tag. Each group elects a landmark node, which advertises its subnet. When a message is to be sent to a node outside the group, the sender forwards the message towards the appropriate landmark. As the message nears the landmark, it is routed towards the final destination by means of FSR.

A similar approach, Contact-Based Architecture for Resource Discovery (CARD) [20], uses the *small world phenomenon* to maintain lists of *contacts* that can be used to route messages to remote communities. The small world phenomenon is the theory that everyone in the world can be contacted through a relatively short chain of social contacts. In mobile routing scenarios, contacts are made when nodes come within range of one another.

2.2.4.4 Encounter Ages

In the Fresher Encounter Search (FRESH) [12] approach, nodes maintain tables containing the time of their last encounter with each node they have previously been a one-hop neighbour with. Nodes discover their neighbours by overhearing messages when they are within range, or by hearing periodically sent “hello” messages.

To find a route to a remote destination, the source node searches for the node $n1$ near itself that has encountered the destination most recently. In turn, $n1$ searches for a node $n2$ that has encountered the destination even more recently. This process continues until the destination is reached.

In this approach the selected route is the only route that is discovered, and is based on the heuristic of “most recent encounter”. The rationale behind this approach is that the nodes that have had the most recent encounters with the destination are likely to be relatively close to the destination. The biggest benefit of FRESH is that it allows flooding to be avoided. However the approach is not guaranteed to find existing routes, and in the worst case it degenerates to full flooding. Furthermore, routes found using FRESH do not consider the stability or length of the resulting route.

2.3 Summary of Related Work

In [5], Camp, Boleng, and Davies provide an in-depth analysis of a number of mobility models. They conclude that the choice of mobility model is a significant factor in the comparison of routing protocols. This claim is further supported by our work. Camp et al also show that the RPGM model is extremely adaptable and capable of representing a wide variety of networking scenarios.

The recently introduced Graph [51] and Obstacle [26] models are examples of mobility models that constrain node movement to lines within the simulation area. Both papers conclude that spatial constraints have a significant impact on the performance of ad-hoc routing. In both cases, nodes follow the edges of an undirected graph.

Some work has been done relating to the stability of routes in ad-hoc networks. Toh [52] shows that the associativity property can produce relatively long-lived routes in ad-hoc networks. Turgut, Das, and Chatterjee [53] argue that stability requires prior

knowledge of the lifetime of a route. They investigate predictive routing approaches that select routes based on the expected time of route disruption. Gerharz, deWaal, Frank, and Martini [14] utilize a statistical analysis of link duration to predict the stability of links in an ad-hoc network. A similar approach was taken by Grober and Li in [16].

Some researchers [3, 24] have studied the stability of elected clusterhead nodes in hierarchical routing models, and others, such as [41], measure stability during the performance analysis of their proposed protocols.

The availability of routes in ad-hoc networks is studied in [57]. The authors use the relative velocity of nodes to predict link availability times.

CHAPTER 3

EXPERIMENTAL DESIGN

3.1 Method

For this study we have chosen to use a generalized full-factorial experimental design, as described by Jain [25]. This approach to experimental design involves:

- Selection of the attributes of the system that we wish to measure and/or optimize. These attributes are termed *objective functions*.
- Identification of the parameters of the system that are likely to influence the performance of the system being studied. We call these system parameters *factors*. Some of the factors are likely to have a larger effect on overall system performance, and we choose to vary these in our experiments. Other factors are held constant throughout the experiments. The parameters that are varied in the experiments are referred to as *primary factors*, and the factors that are held constant are referred to as *secondary factors*.
- Design of experiments to measure the performance of the system with the factors set to various values, which we refer to as *levels*. Initially, experiments are performed using a small number of levels for each factor, thereby allowing the researcher to determine the relative effects of different factors and combinations of factors.
- Once the most significant factors have been determined, detailed experiments are devised to study the behaviour of the system across a wide range of levels.

Finally, the results of the detailed experiments are analyzed, and the behaviour of the system is characterized.

The remaining sections in Chapter 3 discuss the objective functions, factors, experiments, and levels that have been used in this research.

3.2 Objective Functions

The typical objective functions used in the characterization of mobility models include the percentage of packets delivered, average packet delay, average path length, and control packet overhead [5, 21, 26, 52]. Of these metrics, only average route length is suitable for our experiments, as we do not simulate data and control packet transmission.

We feel that a study of mobility models must include some metric that reflects the partitioning and connectedness of the network as a result of node mobility. With this in mind we introduce *route availability* as a metric of interest. We note that this metric is also used in [57].

The final metric that we wish to consider is path life, which we call *route stability*. This metric is common in the study and characterization of link stability in ad-hoc networks [3, 14, 53].

A summary of the objective functions that we have chosen to study are listed in Table 1, and are discussed in detail in the sections that follow.

Objective Function	Symbol	Short Description
1. Route Availability	A	Proportion of time a route is available between a pair of nodes.
2. Route Stability	S	Length of time a selected path is available for use (<i>pathlife</i>).
3. Route Length	L	Number of links (<i>hops</i>) in a path between two nodes.

Table 1: Objective functions used during this research.

3.2.1 Route Availability

In a communication network, a route is a series of links and the routers that terminate them. A route is used to transmit data between a pair of nodes. There may be many routes between a pair of nodes or there may be no routes, as dictated by the topology of the network. The topology of an ad-hoc network is dynamic, with communication links continually being created and destroyed based on the movement pattern of the nodes.

What is Route Availability?

In these experiments we define *route availability* as the proportion of time that a pair of nodes is connected by at least one route such that the route can be discovered by the

routing protocol that is used in the experiment. During each experiment we record the total amount of time that the routing protocol is aware of a route between each pair of nodes. At the end of the experiment we calculate the route availability A in a network a as

$$A_a = \frac{2 \sum_{i=1}^N \sum_{j=i+1}^N (t_{ij}/t_s)}{N(N-1)} \quad \text{Equation 3-1}$$

Where $N =$ the number of nodes in the experiment
 $t_{ij} =$ the total time that nodes i and j were connected by one or more routes
 $t_s =$ the total number of seconds the simulation was run for

Why select Route Availability as an objective function?

Availability is a factor that is likely to be strongly influenced by the mobility model, as well as the routing heuristic. By identifying and adopting protocols and models that optimize availability, we can increase the overall performance of a network. Additionally, we anticipate a relationship between route availability levels and stability levels, and would like to investigate this possibility.

3.2.2 Route Stability

We define *route stability* as the length of time that a route exists after it has been selected for use. This definition of route stability is also referred to as *pathlife*. A route ceases to exist when data packets transmitted on this route are no longer able to reach their destination. Using this definition, a network a is said have higher stability than another network b if the average pathlife of the routes in a , S_a , is greater than the average pathlife of the routes in b ($S_a > S_b$).

Why is stability important?

The ability to identify and use stable routes in ad-hoc networks is highly desirable, as the process of discovering and selecting new routes incurs a large overhead in ad-hoc networks. By continually selecting routes that have a longer pathlife, the amount of control traffic in the network is decreased, thereby improving the bandwidth constraints

that are inherent in ad-hoc wireless networks. As an example of the types of bandwidth constraints that can occur, consider the recent study performed at the University of Maryland [54] that indicates that an 11 Mbps 802.11 network is constrained as low as 1Mbps throughput with only 10 nodes present.

The stability of a network, S_a , is calculated as the average life of all routes selected for use during the experiment, as defined by Equation 3-2.

$$S_a = \frac{2 \sum_{i=1}^N \sum_{j=i+1}^N (t_{ij} / r_{ij})}{N(N-1)} \quad \text{Equation 3-2}$$

Where $N =$ the number of nodes in the experiment
 $t_{ij} =$ the total time that nodes i and j were connected by one or more routes
 $r_{ij} =$ the total number of routes between nodes i and j

3.2.3 Route Length

Route length is defined as the number of links (also known as *hops*) in the route. It is included as an objective function to allow us to determine if a relationship exists between the length of a route and the stability of that route. Intuitively, we expect shorter routes to exhibit higher stability than longer routes, based on the assumption that when routes have more links there is a higher probability that one of the links will break due to node movement. The inclusion of *route length* as an objective function allows this assumption to be tested.

Furthermore, we hypothesized that some mobility models, such as Random Waypoint, induce significantly shorter routes than other models, such as RPGM. The inclusion of route length allows us to quantitatively study and compare the route lengths induced by the various mobility models.

How is route length calculated?

Our experiments measure the number of hops in each route that is selected for use by the routing protocol. We then average these measurements to obtain an average route length for the network.

The formula used to calculate the average route length L for a network a is

$$L_a = \frac{2 \sum_{i=1}^N \sum_{j=i+1}^N (h_{ij} / r_{ij})}{N(N-1)} \quad \text{Equation 3-3}$$

Where h_{ij} = the total number of hops in all routes between nodes i and j
 r_{ij} = the total number of routes between nodes i and j

3.3 System Parameters and Factors

There are many parameters that influence the behaviour of ad-hoc networks. Table 2 lists the parameters that were considered in the design of our models and experiments. The remaining sections in this chapter explain the relevance of these factors to the study of route availability and stability in ad-hoc networks. Specifically, we indicate which parameters will be held constant, and which will be varied during our experiments.

Parameter	Symbol	Description
1. Route Selection Heuristic	H	Heuristic used by the routing protocol to select the routes.
2. Mobility Model	M	Node movement pattern within the simulation area.
3. Node Density	D	Number of nodes in the simulation area.
4. Topology	T	Overall shape of the simulation, including obstacles.
5. Transmission Range	r	Range of the radio used to send packets.
6. Mobility Rate	s	Speed at which nodes move within the simulation.
7. Simulation Area	x	Total size of the simulation area in square meters.
8. Network Load	b	Amount of traffic generated by each node.
9. Node Lifetime	l	Amount of time a node spends in the simulation area.
10. Pause Time	p	Amount of time that nodes are stationary at destinations.
11. Service Model	v	Type of communications services a typical node uses.

Table 2: Description of parameters that affect stability in ad-hoc networks.

3.3.1 Primary Factors

The primary factors are those factors that are varied during the experiments. Our research proceeded in two phases. In the first phase, three factors were analyzed using a limited range of levels, with the goal of quantifying the contribution of each factor with the observed behaviour of the system. The factors studied in these initial experiments included:

1. *mobility model (M)*
2. *route selection heuristic (H)*
3. *node density (D)*

To ensure that dependencies between the factors were accounted for, the combined effects of these three parameters were calculated and considered in our analysis.

Why were these parameters selected as primary factors?

Our work focuses on the identification and analysis of factors that affect routing performance in ad-hoc networks. One of our hypotheses is that the mobility model, which is largely ignored in other studies, has a significant effect on routing performance. Another of our hypotheses is that associativity based metrics result in the selection of stable routes. Therefore both of these factors must be varied if we are to use the results of these experiments to test our hypotheses.

In [44], Royer, Melliar-Smith, and Moser show that node density is an important consideration in the evaluation of ad-hoc networks. Therefore we felt it necessary to examine the capabilities of the models and heuristics across a range of density levels before making any claims about their relative performance.

Density is an interesting factor, as it is one that cannot be easily controlled by network designers or protocol developers. If we are able to find some relationship between density and other factors that can be controlled, we may be able to exploit this relationship in the development of new protocols.

In the second phase of experimentation a fourth factor is introduced:

4. *simulation area topology*

When an ad-hoc network simulation is configured, the simulation designer must specify the physical size and geometry of the area that the mobile nodes move in. Typically this is a square or rectangle. This shape and size is referred to as the *simulation area topology*.

Simulation area topology was introduced as a factor based on the analysis performed during the phase 1 experiments, which indicated that the shape of the simulation area may be more significant than initially assumed. During this second phase, the simulations were run using a small set of topologies to determine if the topology has any effect on

performance. These experiments concluded that topology does have a significant effect on the objective functions. This resulted in more detailed experiments that incorporated a wider range of simulation area topologies and node densities.

All remaining factors were considered secondary and held constant at the levels indicated in Section 3.3.7. The choice of which factors to consider as secondary was driven by our desire to maintain generality in the experiments. Many of the secondary factors, such as b and v , require the introduction of fully functional routing protocols and data transmission capabilities. These protocols introduce a myriad of additional protocol specific parameters which could impact the performance of the system. Rather than enter into an engineering exercise that seeks to optimize each protocol, we chose to work at a general level and to look for trends that could influence the evolution of these protocols in the future. Section 3.3.6 provides additional detail about each of the secondary factors.

3.3.2 Mobility Model Levels

One of our goals was to use realistic values and models for our simulations. As discussed earlier, we do not feel that the Random Waypoint and Random Walk models accurately reflect the mobility scenarios that people are likely to encounter in every day life. However, they are the basis for almost all ad-hoc network simulation. As such we decided to include one of these models, to facilitate comparison with other approaches and to establish a baseline. Of the two models, *Random Waypoint* was selected because it results in a more uniform coverage of the simulation area, as illustrated in Figure 1 and Figure 2 (see Section 2.1).

The RPGM model was selected to represent the class of group mobility models. It is the most sophisticated group model, and facilitated the comparison of a variety of group scenarios. RPGM is configured with many parameters, which are discussed in Section 3.4 along with the configuration and design of the simulation software.

The quest for modelling realism demands the selection of a model that includes geographical constraints that restrict the movement of nodes. Candidate models such as the Obstacle and Graph mobility models did not exist when this research began. This lack of an existing, realistic mobility model motivated our decision to devise a new approach.

We call our model the *Constrained Path Mobility (CPM)* model. CPM allows a researcher to model scenarios that include open areas as well as the corridors that connect them. The open areas are the destinations that are used in the simulations. Unlike the graph or obstacle models, CPM does not require that pathways follow the lines of a graph, and it does not constrain the destinations to fixed points in the simulation area. Instead, CPM uses a series of connected, narrow rectangles to construct what we call *corridors*, and it uses circles to represent the open areas (destinations). The trajectory of each node's path through a corridor is uniformly distributed across the width of the corridor, and changes at a random interval t . Each node's speed is random between $[minspeed, maxspeed]$, and is assigned for each t .

The idea of open areas was adopted to allow us to model rooms in a building. Open areas are connected to corridors by vertices which define the *entrances* to the open areas (rooms). When a node reaches an entrance, it selects a random location in the room, travels there, and waits for some period p of time. When p has elapsed the node selects a new destination and resumes travelling.

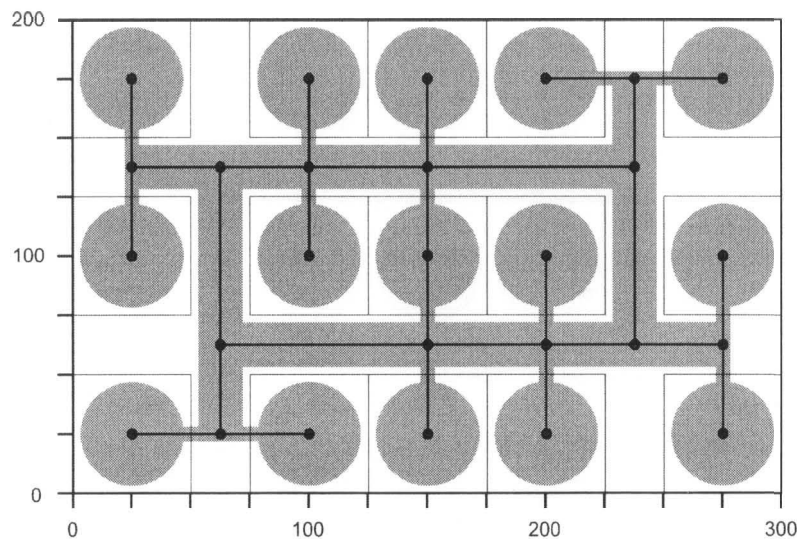


Figure 13: Constrained Path Mobility model. Shaded areas depict areas where nodes can travel. Circles indicate open spaces where nodes can congregate. The paths that nodes take are derived from the vertices and edges of the underlying graph.

The model is implemented as an undirected planar graph where the vertices represent entrances, the centres of open spaces (destinations), and the physical centres of corridor

intersections. When a vertex is to represent a destination, it is assigned a radius r that defines the region covered by the open space. Edges define the centre's of corridors, which constrain the paths that nodes can travel down. Each edge is assigned an l -value and a w -value to represent the length and the width of the corridor.

Figure 13 shows the layout of a building that is being modelled, indicating the resulting graph and regions that nodes may travel through. Figure 14 illustrates two example paths through the same building.

This model facilitates the study of the effects of constrained node movement without limiting the nodes to straight-line travel between point destinations. Nodes are no longer able to travel through walls or obstacles, and instead must travel down the defined corridors to reach their intended destinations. Since CPM is intended as a model for buildings or small-scale outdoor areas (e.g. parks or schoolyards), we decided to allow the radio signals to propagate through the obstacles – which would be office walls or trees and hedges. This behaviour is consistent with that exhibited by the radio frequencies used in 802.11 style networks.

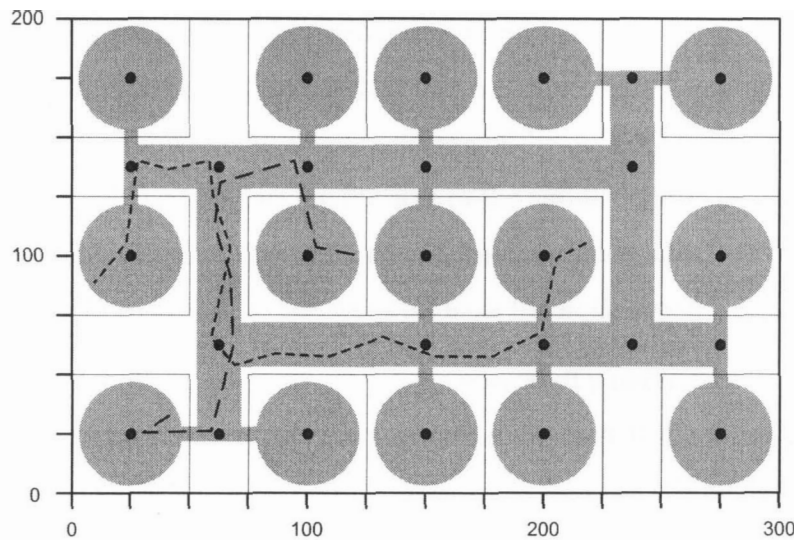


Figure 14: Constrained Path Mobility model showing examples of two different paths.

3.3.3 Route Selection Heuristic Levels

Chapter 2 provided an overview of the heuristics used for route selection in ad-hoc networks. From these, the *shortest path*, *compass*, and *associativity* heuristics were selected as representatives of the available choices.

Shortest path (minimum hop count) is by far the most common route selection heuristic, and was included to provide a baseline against which other heuristics could be compared. Compass routing is representative of the position-based heuristics, and has been shown to perform reasonably well in terms of both hop count and success rate (ability to discover routes). Finally, we felt that any study investigating stability would be incomplete without the associativity metric, which claims to facilitate selection of stable routes.

3.3.4 Node Density Levels

Node density is typically measured in nodes per unit of simulation area, for example: 10 nodes per 1000 square meters (m^2). However, this type of measurement is largely irrelevant when considered independently of the transmission range of the nodes. In order to perform meaningful simulations of ad-hoc networks, it is necessary to determine the number of nodes required to obtain a desired radio coverage level of the simulation area, thereby allowing the average degree of the network to be estimated. This can be accomplished by varying the size of the simulation area x , the transmission range r , or the number of nodes that are active in the simulation area. We have chosen to vary the number of nodes, which we refer to as the *node density*, or D .

A range of node density levels typical of a university campus are used in this study. The selected levels represent low density (such as the times when there are few students in the halls), medium density (such as an average school day), and high density (such as a day near the beginning of term when everyone is on campus).

The lowest level considered is $D = 10$ nodes, which represents a density of 1 node per $2000 m^2$ (based on the levels selected for the related secondary factors, to be discussed subsequently). The radio coverage of each node is $2827 m^2$, suggesting that these ten nodes could theoretically cover the entire simulation area and remain in contact with each other.

The maximum level selected is $D = 90$ nodes, or 1 node per 222 m². This is roughly 10 times as dense as the lowest level. When considered in the context of the values selected for r and x , the levels chosen for D should produce networks that have average degree from 1.8 (at $D=20$) to 10.3 (at $D=80$), when the nodes are uniformly distributed throughout the simulation area.

3.3.5 Simulation Area Topology Levels

Simulation area topology was introduced as a primary factor in the second phase of experimentation. It was decided to constrain the shape of the simulation area to a rectangle, and the overall area to be roughly 20000 m² - the same value used in phase 1. Therefore, the length and width of the simulation area were varied to create topologies ranging from squares to long narrow rectangles.

The narrowest rectangle considered is one where the radio coverage of a node spans the width of the rectangle, regardless of the position of the node in the rectangle. The complete set of topologies used is listed in Table 3.

Length (metres)	Width (metres)	Area (sq. metres)
142	142	20164
200	100	20000
300	67	20100
400	50	20000
500	40	20000
600	33	19800
700	29	20300
800	25	20000

Table 3: Simulation area topologies used during second phase of experimentation.

3.3.6 Secondary Factors

The following sections discuss the secondary factors and the constant values that were chosen for them. In all cases we attempted to select values that were representative of real-world scenarios, and that were consistent with the values used by other researchers.

3.3.6.1 Transmission Range

In ad-hoc networks, each node is equipped with a radio. The maximum distance that this radio can transmit a signal is the transmission range r of that node. In our experiments, all nodes have identical transmission ranges.

Current state-of-the-art radios for mobile 802.11 hosts are capable of transmitting very large distances (up to several kilometres). The costs of larger transmission ranges are:

- reduced lifetime of the node due to battery depletion
- network congestion due to collisions at the MAC layer

It is for these reasons that most ad-hoc networks are studied with relatively small transmission ranges.

Another important consideration for researchers is the relationship between transmission range r , node density D , and the size of the simulation area x . Obviously, the potential connectivity of the network increases as the degree of the network increases, and the average degree of an ad-hoc network is determined by the number of nodes that are within range of each other (neighbour-pairs). Any increase in r or D , or any decrease in x , will increase the number of neighbour-pairs, resulting in higher average degree. Since we have already chosen to vary D , r and x will be held constant.

The value $r=30$ meters, which is about the length of a large room or gymnasium, is used in these experiments.

3.3.6.2 Mobility Rate

The mobility rate s is the speed at which nodes move within the network. For this study, s is constrained to ad-hoc networks consisting of pedestrian traffic. This is accomplished by placing upper and lower bounds on the allowable movement rate s . For each node, the simulation software generates a mobility rate that is uniformly distributed within these bounds. Once a rate is assigned to a node, it is used until the node reaches a destination, a waypoint, or the end of one leg of a journey (models that use a time interval t for random motion can be said to have multiple legs in each journey, where a leg is the portion of the

path travelled in the interval t). A new rate is assigned for each destination, waypoint, or leg.

The speeds used in the experiments range from $s=1$ m/sec to $s=5$ m/sec. These are consistent with the speed of a very slowly moving person, to a fast running person.

3.3.6.3 Simulation Area

The simulation area x , refers to the actual size of the area that mobile nodes can move in during the simulation. As discussed in Section 3.3.6.1, the size of the simulation area x , the node density D , and the transmission range r , all work together to define the degree of the network at any point in time. With this in mind we chose to vary D , and to hold x constant.

This research is concerned with networks that are about half the size of an average city block. This size is interesting to us as it corresponds to the area that might be occupied by a single department in a company, the floor of a classroom building on a large campus, a small park, or a group of businesses in a small area of town. This constraint on the size of the network is justified by the observation that larger networks, such as those in urban settings, are likely to include fixed-line infrastructure to connect across longer distances.

A value of $x=20000$ square meters is selected for the experiments. Initially this will be in the form of a 100m by 200m rectangle. Further discussion of the simulation area topology was presented in Section 3.3.4.

3.3.6.4 Network Load

Network load, b , refers to the volume of data and control packets that are placed on the network by any node. As the load on the network increases, the network starts to experience congestion-related problems. In 802.11 style networks this typically shows up as collisions at the MAC layer, which occur when two or more neighbouring nodes try to send packets simultaneously.

Network load is a relevant factor in studies that focus on optimizing and/or characterizing throughput in mobile ad-hoc networks, as is the case with most protocol evaluations. However, our goals are different. This initial study is focussed on the affect

that mobility models have on the availability and stability of routes, regardless of any physical bandwidth constraints, MAC layer implementations, or radio technology used in the network.

To avoid the influences of these network engineering concerns, we make the assumption that nodes are capable of handling any amount of load that is placed on the network.

3.3.6.5 Node Lifetime

Node lifetime, l , refers to the amount of time that a node spends in the simulation area. In the simplest case, a node has infinite life – which suggests that it will remain in the simulation area until the simulation is complete. This is exactly what is done in these experiments, which set $l = T_s$ (the total simulation time), implying that all nodes are active participants in the entire simulation.

Other interesting ways to model node lifetime include:

- Constrain the node life to be a function of the battery power used by the node
- Allow the node to move out of the simulation area (in which case its “life has ended”)
- Constrain the lifetime of a node to a duration generated by a function (random or otherwise)

It is our intention to include these models of node lifetime in future mobility model simulations, so that we can try to understand the relationship between node lifetime, mobility models, and the resulting route availability and stability.

3.3.6.6 Pause Time

Pause time, p , defines the amount of time that nodes are stationary upon reaching their destinations. When $p=0$, nodes simply reach their destinations and continue on to their next destination. In most cases this is not a reasonable scenario, as the majority of people (nodes) migrate between destinations, spending some time at the destination before moving on.

An example would be a study of an ad-hoc networks at a university campus. We would likely find that nodes migrate between classrooms, recreation areas, and study

facilities – spending enough time at each destination to accomplish a task (eating lunch, participating in a class, etc.).

This notion of pause time is critical to the operation of ad-hoc networks, as routes through stationary nodes are generally more stable than routes through highly mobile nodes [52]. Knowing this, we designed our experiments to test the heuristics and models in situations that would not contain long-lived routes through stationary nodes. This involved limiting the pause time to $p=5$ seconds for all nodes in the simulation.

3.3.6.7 Service Model

The service model defines how the nodes in the network access network services (email, news, games, etc). Some examples of different service models are:

- a distributed model in which services are hosted by a subset of the overall nodes, and where these service providers move throughout the network
- a peer-to-peer model in which all nodes share information with other nodes, and where no single node acts as a primary server or coordinator of the services
- a centralized model where services are accessed through central servers, and where these service providers are most likely stationary

What does the service model have to do with stability and availability?

Consider the case of a centralized service as described above. If all nodes want access to this service, they have to find routes between themselves and the server. If the server is stationary, the routes are very directional. This undoubtedly affects the discovery of stable routes, especially when compared to something like the peer-to-peer model in which any node-pair is just as likely to communicate as any other node-pair.

Again, this factor is interesting but it is beyond the scope of the research being done in this paper. These experiments assume a peer-to-peer model in which each node-pair communicates with equal probability.

3.3.7 Levels (Secondary Factors)

The levels for all secondary factors discussed in the previous sections are listed in Table 4. Actual real-world ranges of levels are provided for factors where these levels are known.

Factor	Range of Levels (real-world)	Selected Level
5. transmission range t	up to 300m in closed environment	30m
6. mobility rate r	up to 50m/sec (vehicles) or 5m/sec (people)	1 to 5 m/sec
7. simulation area x	Unbounded	20000 sq. m.
8. system load b	Unknown	n/a
9. node lifetime l	Up to 1 day (depends on battery usage)	3600 seconds (same as T_s)
10. pause time p	Unbounded	5 seconds
11. service model v	Distributed, centralized, peer-to-peer	Peer-to-peer

Table 4: Levels for secondary factors used in the experiments.

3.4 Configuration of Experiments

All of the experiments were performed using computer simulations developed explicitly for this research. Initially, the *ns-2* simulation environment was considered, however it lacked support for most of the heuristics and models needed for the experiments. Since the decision had already been made to omit per-hop throughput and traffic analysis from our experiments, the creation of custom simulation software was deemed a viable alternative.

The resulting software is highly configurable, and implements all mobility models, topologies, routing heuristics, and factors needed for the experiments. Furthermore, the software is designed such that it can save node events to *ns-2* input files, ensuring that the models developed here can be used for a wide range of future research activities.

The following sections discuss the implementation of the various mobility models and routing heuristics.

3.4.1 Mobility Models

3.4.1.1 Random Waypoint Mobility Model

The Random Waypoint model implemented for the simulations is exactly as described in Section 2.1.1.2. The model takes three parameters, *minspeed*, *maxspeed*, and *pausetime*. These are set to 1 m/sec, 5 m/sec, and 5 seconds respectively, as discussed in Section 3.3.

3.4.1.2 RPGM Mobility Model

The RPGM model was constructed to be consistent with the model described in Section 2.1.2.1. Group destinations are assigned randomly within the simulation area, as is each node's initial starting position. Other parameters used for the simulation are listed in Table 5.

RPGM Parameter	Description	Value
<i>numgroups</i>	Number of groups in the simulation.	Variable
<i>nodespergroup</i>	Number of nodes in each group.	5
<i>maxrefsep</i>	Max separation of a ref point from the group centre.	10 metres
<i>maxnodesep</i>	Max separation of nodes from their ref point.	5 metres
<i>tick</i>	Mean duration of a small random node movement.	5 secs
<i>tickvar</i>	Max variation of randomly selected tick value.	± 1 sec

Table 5: RPGM Parameters used in simulations.

The number of nodes in each group was held constant at five nodes per group, and the number of groups is derived from the total number of nodes in the simulation. Using these values, a permanent route is guaranteed to exist between all members belonging to the same group. This was verified experimentally. The effect of this group model characteristic is that it inflates the average availability and stability rates in the network. With this in mind we decided to follow the lead of other researchers [5] and separate the results for *inter-group* and *intra-group* routing. In doing this we found that the *inter-group* results were more interesting and provided a better comparison against the results from the entity models. Additional details regarding the *inter-group* and *intra-group* results are provided in Section 4.2.

3.4.1.3 Constrained Path Mobility Model

The Constrained Path model was introduced and discussed in Section 3.3.2. However, for the purposes of the experiments it is necessary to model an actual scenario within the simulation area boundaries defined for the simulations. The model used in the simulations is intended to represent a section of a campus or office building. It consists of eight destinations (open areas) that represent fairly large, uniformly-sized rooms. The proximity of corridors to rooms allows for some relaying of packets through nodes that have moved close to the walls of the rooms, as the model assumes that the walls are not impervious to the radio frequency used by the nodes.

Figure 15 illustrates the actual CPM scenario implemented for our experiments. The shaded rectangles are the corridors that nodes travel down, and the shaded circles are the open areas that define the destinations for the nodes.

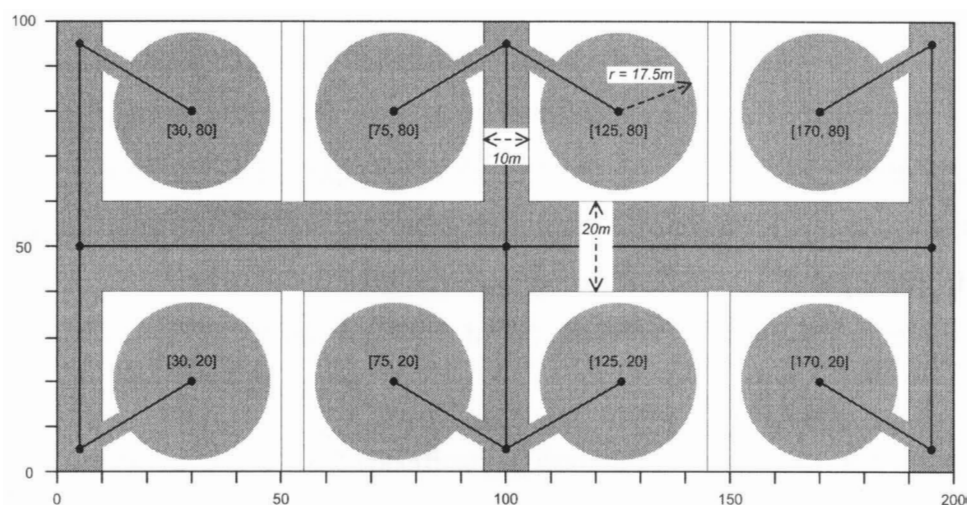


Figure 15: Physical layout of the Constrained Path model used in all simulations. Nodes move within the shaded areas. Destinations are within radius r of the centre of each open area.

3.4.2 Routing Heuristics

3.4.2.1 Shortest Path Routing

During the simulations, shortest paths are calculated using Dijkstra's algorithm. Most fixed-line routing protocols choose to adapt their routes to use new shorter paths as they become available. The reasoning is that a new shorter path reflects a change in network

topology, likely due to the failure or reconfiguration of a router, and therefore it is in the best interest of the network to adopt this newer route.

However, network topology changes are constant in ad-hoc networks. If the routing protocol was to adopt each new shortest path that appeared, pathlife would be very short and control overhead would be very high. Therefore it was decided to select paths and hold on to them as long as possible, until node mobility causes the route to become disconnected.

In reality, all credible ad-hoc routing protocols that employ shortest path route selection also make use of some form of route maintenance to keep routes alive. However our experiments define stability as the longevity of a single path, without any maintenance activities. This seems reasonable as route maintenance activities incur overhead just as route discovery activities do.

3.4.2.2 Associativity Based Routing

The associativity heuristic used in the simulations differs in a few minor ways from the protocol described in [52]. One of the differences is the lack of forwarding load and link delay information. Since the nodes are not sending actual data, it is not possible to take this information into consideration during route selection. Consequently, the route selection used in the simulations is based solely on associativity count, with hop count being used to break ties.

According to the ABR specification, destination nodes delay the route selection decision for some constant amount of time after the first request packet is received, to allow requests that traversed slower paths to arrive. In our implementation, all possible paths are considered before routes are selected. This has the side effect that if a route exists, it will be identified and selected, regardless of associativity levels.

Associativity requires the specification of a *tick* constant and a threshold A_{Th} . The value of these constants is based on cell size and speed. The value of *tick* was set to 1 second, and A_{Th} to 30 seconds – for all of the simulations.

3.4.2.3 Compass Routing

Compass routing was implemented as defined in Section 2.2.2. No mechanisms are provided to backtrack and recover when local maximums are reached, which means the heuristic will present lower availability levels than experiments that use the other heuristics.

The software assumes the existence of a location service, and calculates next hops based on the actual current positions of all nodes.

3.4.3 Simulation Environment

Simulations were always started using well known seeds for the random number generator, allowing individual experiments to be repeated if necessary. The random number generator used in the software is a multiplicative linear congruential generator that produces a sequence of pseudo-random numbers with a period of $2^{32} - 2$. A detailed study of this generator is available in [38].

All simulations were run for 3600 seconds (1 hour) of simulation time. Collection of statistics is not started until the 600 second mark of each simulation, to ensure that the random number generators, heuristics, and models are in a steady state.

All simulations were run on dedicated PC's. These are Pentium 4 machines with the Microsoft Windows XP operating system. Throughout the course of this research, close to 1000 simulations were run. The range of execution times for a single simulation was from approximately 2 seconds (5 nodes), to 4.5 hours (90 nodes).

CHAPTER 4

RESULTS OF EXPERIMENTS

4.1 Node Density: The Primary Factor?

Our first experiments were performed with the goal of comparing and quantifying the effects of node density (D), mobility model (M), and routing heuristic (H) on the objective functions defined in Chapter 3, these being route stability (S), availability (A), and length (L).

Does D have a greater influence on the objective functions than the other factors?

Simulations were run with each of the three factors set to the levels indicated in Table 6. A full-factorial model was used to test all 27 combinations of the various factors and levels. Each simulation was repeated five times using different seeds to minimize experimental error.

Factor	Level 1	Level 2	Level 3
D	20	50	80
M	Random Waypoint	Constrained Path	RPGM
H	Shortest Path	Associative	Compass

Table 6: Levels and factors used for experiment 1.

The results of the simulations were analyzed using the ANOVA method, and are summarized in Table 7. First order values indicate the portion of the result that can be attributed to individual factors, second order values indicate the effect of combinations of factors. For example, in Table 7 there is a second order combined effect DM , which has a value of 0.11. This indicates that 11% of the change in stability is caused by some interaction between density (D) and mobility model (M). The Error term indicates the proportion of the results that cannot be attributed to any of the factors, and therefore are assumed to be caused by experimental error. This term does not necessarily indicate that the results are poor – only that they cannot be attributed to the factors used in the experiments.

An *F*-test was applied to verify the statistical significance of the values, and is illustrated in Table 8. According to this test, all results from Table 7 that are representative of more than 1.0 percent of the system behaviour are statistically significant at a 90% confidence level. This includes the effects of individual factors as well as the more noticeable combined effects.

The results that are not significant (shaded in grey), include the combined effects *DH* and *MH*, as well as the second order effects. The reader will observe that the influence of these factors is negligible.

<i>Allocation of Effects</i>	Mean Availability	Mean Stability	Mean Route Length
Primary Effects	0.95	0.61	0.81
<i>D</i>	0.72	0.42	0.55
<i>M</i>	0.22	0.04	0.14
<i>H</i>	0.01	0.14	0.11
1st Order	0.03	0.19	0.18
<i>DM</i>	0.02	0.14	0.14
<i>DH</i>	0.00	0.01	0.03
<i>MH</i>	0.00	0.05	0.01
2nd Order	0.00	0.00	0.01
<i>DMH</i>	0.00	0.00	0.01
Error	0.03	0.20	0.01

Table 7: Allocation of effects of *D*, *M*, and *H*. A generalized full factorial analysis was used, with each experiment being repeated five times.

F-Table $\alpha=0.1$		F-Computed		
		<i>Availability</i>	<i>Stability</i>	<i>Length</i>
9.47	<i>D</i>	1509.31	147.42	3453.95
9.47	<i>M</i>	450.68	15.23	909.76
9.47	<i>H</i>	15.44	51.05	714.35
3.80	<i>DM</i>	23.17	23.84	432.71
3.80	<i>DH</i>	1.59	2.28	91.42
3.80	<i>MH</i>	1.80	8.17	39.19
2.36	<i>DMH</i>	1.23	0.17	9.40

Table 8: Analysis of variance of effects for initial availability, stability, and length results. Results that are not shaded are statistically significant at a 90% confidence level.

The results confirm that *D* has the largest overall effect on route availability, stability, and length. This result is not surprising for route availability, as higher node densities

necessarily result in a network with higher average connectivity. It is also worth noting that the range of levels selected for D vary by as much as 400%, indicating that a small change in D may not have a large effect on the performance of the network.

We also notice that the H has negligible effect on the overall availability, which can be attributed to the fact that all classes of routing protocols do a reasonable job of finding routes when routes are available. On the other hand, H does contribute significantly to the stability of the routes, as well as to the length of the routes. We will explore this relationship in more detail in the following section using the results from the 2-factor experiments to be described later.

The choice of M does not appear to have a large affect on route stability, however it does provide a significant contribution to the availability and length of the routes. This observation is important as it backs up our claim that mobility models have a significant effect on network performance. Furthermore, we notice that M is present in the first order effect DM , which is responsible for the largest contribution of all combined effects. This combined effect of density and mobility model suggests that the performance of certain models is somehow related to the density of the network. In Section 4.3 we will see that as the density increases, the effect of the different models on stability and route length becomes more pronounced.

In conclusion, we answer the question posed earlier in this section by claiming that D is the most significant of the factors tested so far – at least with the ranges of levels used for these initial experiments.

Of course D is something that we have little control over in real-world scenarios, and is a factor that is likely to span a wide range of realistic values. In most cases the number of nodes in a network cannot be pre-defined, and the best that can be done is to construct networks and protocols that work effectively across a wide range of densities.

4.2 Mobility Models, Routing Heuristics, and Route Availability

The experimental results documented in this section are based on a 2-factor, full factorial experimental design, with M and H as the factors. These experiments were created to answer the question:

What effects do M and H have on route availability at various node density levels?

To answer this question we ran the simulations from the previous section using an increased number of density levels, and graphed the change in A as D increases. The levels used for the simulations are shown in Table 9.

Factor	Levels used
D	10, 20, 30, 40, 50, 60, 70, 80, 90
M	Random Waypoint, Constrained Path, RPGM
H	Shortest Path, Compass

Table 9: Factors and levels used in the detailed study of availability.

Note that we did not include ABR as a routing heuristic in these new experiments, as both the *ABR* and *Shortest Path* algorithms produce the same availability results (they both guarantee to find a route if one exists).

Figure 16 illustrates the results.

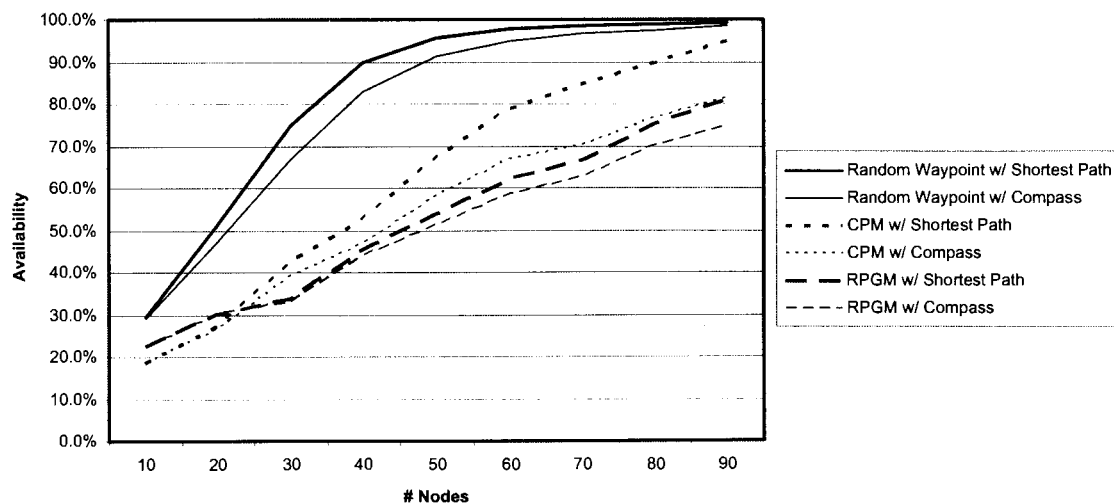


Figure 16: Route availability at various densities, as influenced by M and H .

The following trends are observed:

1. *Highest route availability levels occur in experiments that use the Random Waypoint model.*

As density increases, networks modelled using Random Waypoint exhibit a more rapid increase in availability than networks based on other mobility models. We expect this is caused by the random movement pattern that distributes the nodes more uniformly than other models, leading to a higher level of connectedness within the underlying network structure.

Certainly Random Waypoint is not a very realistic model. This is evident in our everyday lives where people simply do not move randomly, nor do they distribute themselves evenly within their environments. The most disturbing aspect of this result is that most ad-hoc network protocol studies seem to use this Random Waypoint mobility model, suggesting that the routing protocol performance reported in these papers is optimistic and unrealistic.

2. *Availability in simulations using the Random Waypoint model starts to level off at an average of 93% at a density of 50 nodes. The other models average 62% availability at this same density level.*

The rate of change of A in the Constrained Path and RPGM models is similar, and appears to be linear. What these models have in common is that they require groups of nodes to take similar paths through the simulation area. This suggests that availability is much lower in scenarios where nodes travel in similar patterns to one another. Such scenarios require a significantly larger number of nodes to achieve route availability levels equivalent to scenarios that model random independent node movement. For example, Figure 16 indicates that the first non-random scenario to achieve 90% availability uses the CPM model, and it reaches this level at 90 nodes. The Random Waypoint models achieved 90% availability with less than 50 nodes.

3. *RPGM models produce the lowest route availability.*

In part this may be due to the RPGM parameters we used. Initially we suspected that the low availability rates could have been caused by the exclusion of the intra-group

communication results. However, Figure 17 shows that except at very low densities, the intra-group results make very little difference. In any case, the RPGM model produces availability levels that are similar to those of a Constrained Path model, which suggests that it is more reasonable to use than Random Waypoint.

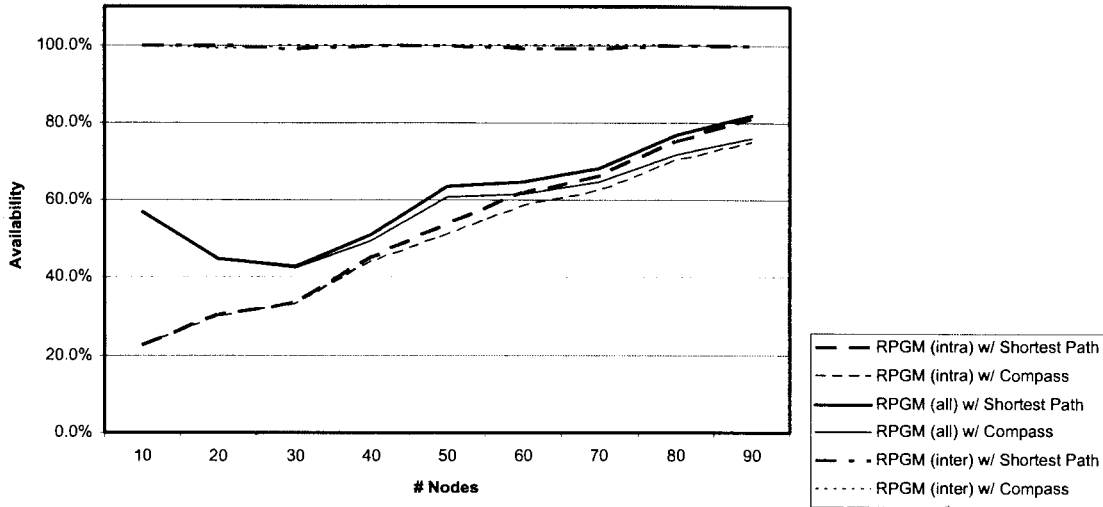


Figure 17: Route availability with separate RPGM inter-group and intra-group data series.

Based on the observations enumerated above, we conclude that the mobility model has a large effect on route availability, and the choice of routing heuristic does not appear to make much of a difference. To determine the specific effects of each of these factors, we performed a 2-factor ANOVA analysis for each of the densities that were simulated (from 10 to 90 nodes). Table 10 summarizes the ANOVA analysis results, and Table 11 indicates which results are statistically significant.

The first thing we notice in Table 10 is the extremely high Error value, 0.76, for the lowest density (10 nodes per 20000m²). This indicates that neither M nor H are the primary factors at $D = 10$. We conjecture that there is insufficient network connectivity to obtain reliable results.

We also notice that M has a consistent and dominant effect on route availability at all density levels. At lower densities the contribution of H to route availability is not significant, but as we reach higher densities H starts to contribute more, reaching 12% at 90 nodes per 20000m², suggesting that the choice of routing protocol is much more important in high density environments.

Availability Effects	Node Density								
	10	20	30	40	50	60	70	80	90
Primary Effects	0.24	0.83	0.94	0.93	0.91	0.95	0.89	0.92	0.88
<i>M</i>	0.24	0.82	0.94	0.92	0.89	0.92	0.83	0.84	0.76
<i>H</i>	0.00	0.01	0.01	0.01	0.02	0.03	0.05	0.08	0.12
1st Order	0.00	0.00	0.01	0.00	0.01	0.02	0.03	0.04	0.07
<i>MH</i>	0.00	0.00	0.01	0.00	0.01	0.02	0.03	0.04	0.07
Error	0.76	0.17	0.04	0.07	0.08	0.03	0.10	0.04	0.05

Table 10: Relative effects of *M* and *H* on route availability at various node densities.

F-Table $\alpha=0.1$		F-Computed (Availability)								
		10	20	30	40	50	60	70	80	90
9.46	<i>M</i>	5.70	88.97	400.02	252.89	192.10	612.12	155.41	371.55	267.13
9.46	<i>H</i>	0.00	0.55	4.96	3.66	5.31	22.91	8.73	34.15	41.41
3.81	<i>MH</i>	0.00	0.26	1.51	0.43	0.63	5.06	3.06	9.90	12.19

Table 11: Analysis of variance of effects for Route Availability. Results that are not shaded are statistically significant at a 90% confidence level.

In summary, we have shown that the choice of *M* has a significant effect on overall route availability at all node densities in an ad-hoc network, and that the Random Waypoint model provides an overly optimistic environment for protocol evaluation and testing. The heuristic, *H*, plays a much less significant role, but becomes increasingly important as the node density increases.

4.3 Mobility Models, Routing Heuristics, and Route Stability

Recall from Chapter 4.1 that while density accounted for roughly half the change in route stability, the routing heuristic and the combined effect of heuristic and density accounted for another 25%. The mobility model had some effect, but was not nearly as important to overall stability. In this section we study the effect that *M* and *H* have on stability by analyzing the results of simulations at many more density levels.

The questions we are investigating are:

What effects do different types of routing heuristics have on route stability?

How much does the mobility model contribute to overall route stability?

The factors and levels used in the simulations analyzed here are the same as those used in the study of availability in the previous section. The reader should refer to Table 9 for the exact factors and levels used.

Simulation results are presented in Figure 18, Figure 19 and Figure 20. These results illustrate the change in stability and route length as density increases.

The following trends are observed:

1. *Simulations using the Random Waypoint model produce a higher number of stable routes than simulations using the Constrained Path model, regardless of H.*

All of the results for Random Waypoint simulations show higher stability levels than the results from the Constrained Path simulations. An explanation of why this happens is included with the discussion of trend 4, below. However, we conclude, as we did for availability, that Random Waypoint models produce overly optimistic environments for the study of routing protocols.

2. *Compass routing produces the most stable routes of any given mobility model.*

The relative ordering of stability is constant between the models studied, with routes found by the compass heuristic being the most stable, and routes found with shortest path routing being the least stable. This result is a bit surprising, as we expected the ABR routing heuristic to produce significantly more stable routes than compass-based routing – at least at higher densities. Instead what we observe is almost identical stability levels and trends for compass and ABR routing, with compass being slightly more stable at all densities.

We explain this result by suggesting that the compass heuristic does not select many of the less stable routes because it does not find paths that route packets backwards – away from the final destination. The exclusion of these paths shows up in the lower availability rates for the compass heuristic (see Figure 16), in the high stability, and in average path lengths that are shorter than those in ABR. In comparison, ABR includes these paths that route backwards (and are more likely to be short-lived), thereby lowering the average measured stability.

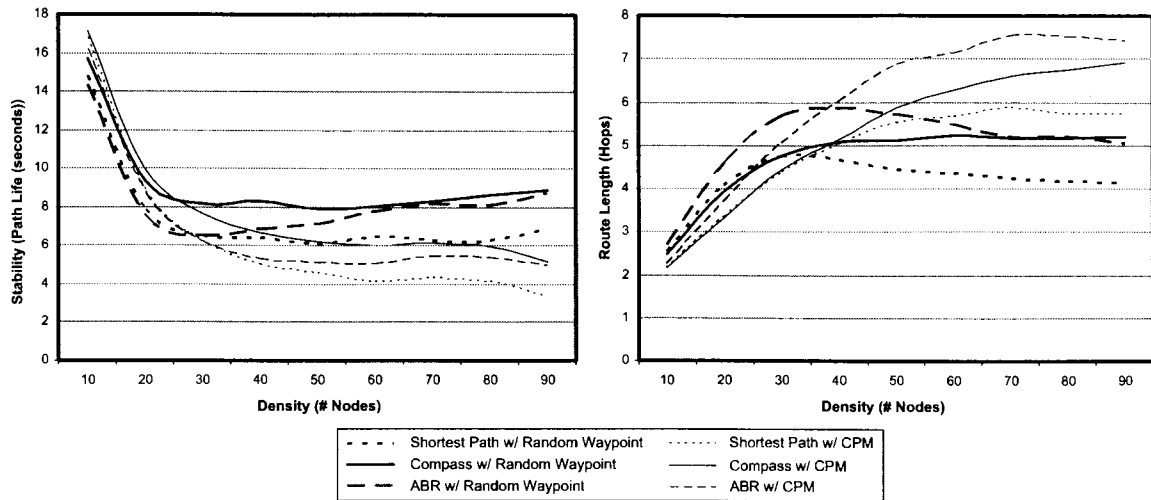


Figure 18: The graph on the left shows the stability of routes in the two entity models, for various combinations of M and H . On the right we have the corresponding route lengths.

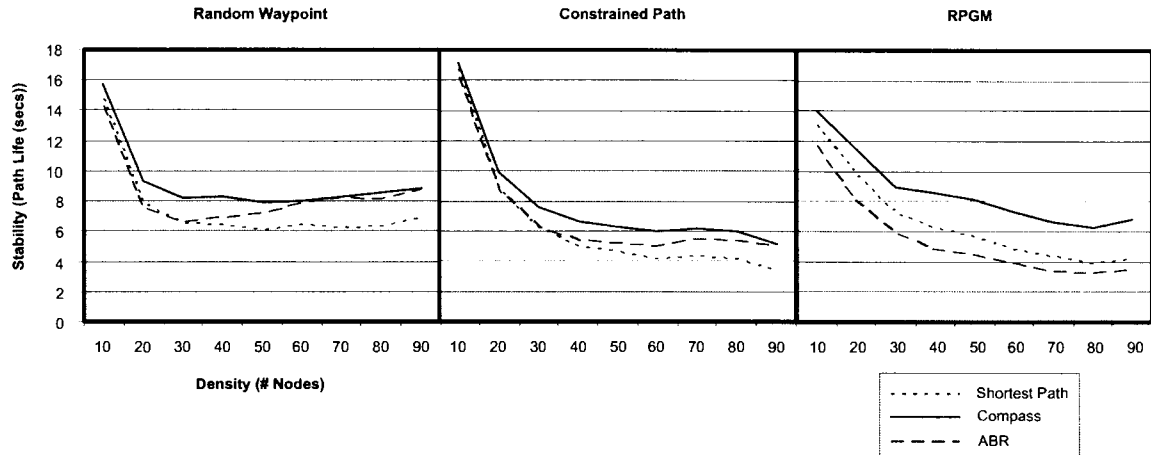


Figure 19: Stability of routes in each mobility model.

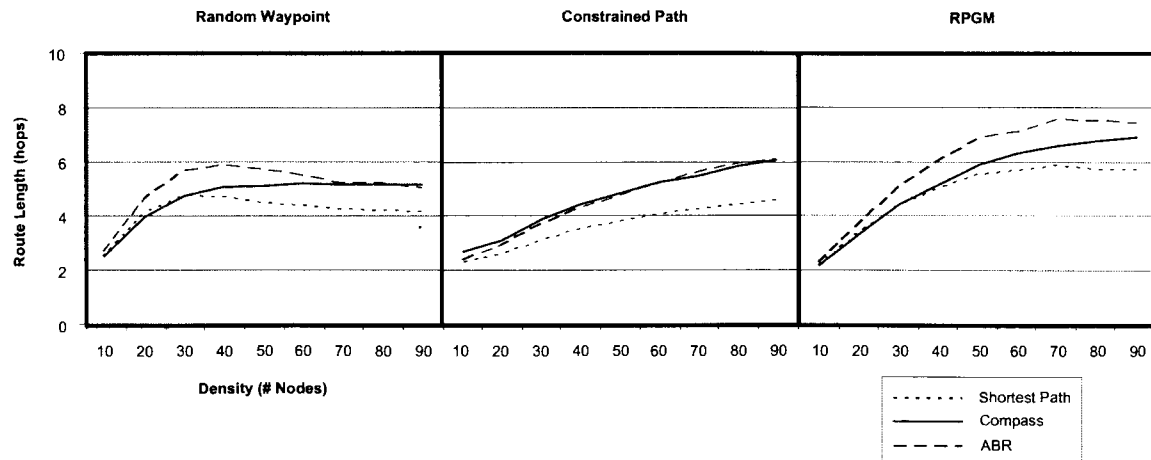


Figure 20: Route length in each mobility model.

3. *Models that produce shorter routes are generally more stable than models that produce longer routes.*

We note that in observation 2, routes found with shortest path routing are less stable than routes found with the other routing heuristics. However, this does not imply that shorter routes are always less stable than longer routes. If we consider both graphs in Figure 18, we see that all the routes in Random Waypoint simulations are shorter than routes in Constrained Path simulations, however, the stability of the Random Waypoint routes is higher than stability in Constrained Path models. We explain this paradox as follows:

- When the model is consistent for a set of simulations (i.e.: it is not a factor), the route selection heuristic has the largest influence on stability. Heuristics that consider stability as a criterion when they select routes are more likely to produce routes that have a higher degree of stability than routes found with heuristics that do not consider stability as a route selection criterion (i.e. shortest path). When the model is varied between experiments, different lengths of paths are produced. It is this change in route length between different models that increases or decreases stability.

In summary, routes found with shortest path do not have lower stability because of their length, they have lower stability because of how their routes were selected. The observable trend is that as route lengths increase, stability decreases, until such a time as the route lengths approach a steady state. When this occurs, the stability also approaches a steady state.

4. *Variation in path stability between best and worst cases is quite large.*

Figure 18 indicates that the highest average steady state route life (8.3 seconds) occurs in the simulation of compass routing in the Random Waypoint model. The lowest average steady state route life (4.1 seconds) occurs when we simulate shortest path routing in a Constrained Path model. This is an increase of more than 100%. To understand what is causing this large variation, we need to consider the nature of the M and H used in the experiments.

In the Random Waypoint model, nodes move in any direction, and can be located anywhere within the simulation region. Given that node movement is random, we assume that nodes will be fairly evenly distributed within the simulation area at any point in time. If this is the case, the diameter of the communication graph will be a function of the transmission range and the size of the simulation area. For a 100m x 200m simulation area that is completely covered by nodes with 30m transmission range, the maximum diameter is 8 hops (using shortest path). If the simulation area is not completely covered with nodes, but the communication graph is fully connected, we could have paths up to 16 hops long. In our experiments, the average steady state route length calculated from Random Waypoint simulation results is 5.0 hops, and the maximum measured route length is 13 hops.

In the Constrained Path model, node movement is restricted to physical corridors, and the links in the communication graph are constructed predominantly from nodes that are moving within these corridors. The actual CPM scenario implemented for our study has approximately 740 meters of corridors, with the furthest distance between any 2 nodes being 360 meters. This results in a maximum route length of 24 hops if the nodes happen to be spaced inconveniently. In our experiments, the average steady state route length calculated from CPM simulation results is 5.7 hops, and the maximum measured route length was 23 hops.

Furthermore, there are regions of the Constrained Path where nodes that are physically 45m apart must communicate through a corridor that is roughly 260m long. In the Random Waypoint model the route between these nodes could be as short as 2 hops, but it will not be less than 9 hops in the Constrained Path scenarios. Consequently we see longer route lengths and shorter route lives in simulations that use the Constrained Path model.

Another characteristic of the models that affects stability is the nature of node movement. In the restricted environment of the Constrained Path model, adjacent nodes are either travelling in the same direction down a corridor, or they are travelling in completely opposite directions. As a result we expect neighbouring nodes to exhibit very high or very low associativity, with a 50% chance of the

associativity being high or low. When considered in conjunction with the longer Constrained Path routes, we can see that there is a good chance that at least one node-pair in a route will have low associativity, and the route will break.

In the Random Waypoint model, associativity is not so polarized. We expect most pairs of neighbouring nodes have an average level of associativity, as their relative velocities are uniformly distributed. This implies that the weakest link in a Random Waypoint route is likely to exist longer than the weakest link in a Constrained Path route. Combine this with the shorter route lengths and it is not surprising that the Random Waypoint model exhibits significantly higher overall route stability.

So far in this chapter we have seen that compass-based routing produces the most stable routes, and that routes in Constrained Path models are significantly less stable than routes in Random Waypoint models. We now turn our attention to a brief discussion of the relative effects of M and H at the various density levels used for these experiments.

As with the study of availability presented in the preceding section, the ANOVA method is used to analyze the effects on stability and route length. Results are summarized in Table 12, and the analysis of the variance of the results is documented in Table 13.

Effects on Stability	Node Density								
	10	20	30	40	50	60	70	80	90
Primary Effects	0.11	0.41	0.58	0.60	0.54	0.61	0.62	0.71	0.70
M	0.09	0.13	0.04	0.17	0.23	0.37	0.41	0.50	0.55
H	0.02	0.28	0.54	0.43	0.31	0.24	0.21	0.21	0.15
1st Order	0.00	0.06	0.07	0.10	0.14	0.11	0.11	0.08	0.08
MH	0.00	0.06	0.07	0.10	0.11	0.11	0.11	0.08	0.08
Error	0.89	0.53	0.35	0.30	0.35	0.28	0.27	0.21	0.21

Table 12: Allocation of effects of M (Mobility Model) and H (Routing Heuristic) on Route Stability.

F-Table $\alpha=0.1$		F-Computed (Stability)								
		10	20	30	40	50	60	70	80	90
9.46	M	1.91	4.51	2.08	10.53	11.53	23.24	27.99	43.02	46.16
9.46	H	0.32	9.31	27.33	25.78	15.80	15.42	14.24	18.40	12.86
3.81	MH	0.03	0.96	1.79	3.04	2.82	3.41	3.67	3.25	3.54

Table 13: Analysis of variance of effects for Route Stability. Results that are not shaded are statistically significant at a 90% confidence level.

The results in Table 12 indicate that both M and H are significant factors affecting stability, although at lower densities (up to 40 nodes per 20000m²) the large error term prevents the results from being statistically significant. Once the density reaches 40 nodes per 20000m² the error term drops somewhat – although it is still quite high. This high error term suggests that a large percentage of the change in stability is not explained by either M or H .

Despite the error term, we notice a consistent trend emerging in the statistically significant results. This trend indicates that the relative effect of the mobility model increases steadily from 17% to 55%, while the relative effect of the routing heuristic decreases steadily from 43% to 15%. This suggests an inverse relationship between the effect of the heuristic and the effect of the model, with the model having a bigger impact on system performance at higher densities.

Unfortunately the combined effect of these factors, MH , is not significant – a result of the high error term.

Effects on Route Length	Node Density								
	10	20	30	40	50	60	70	80	90
Primary Effects	0.89	0.90	0.91	0.94	0.93	0.94	0.96	0.95	0.89
M	0.82	0.73	0.63	0.61	0.59	0.64	0.57	0.58	0.82
H	0.07	0.17	0.28	0.34	0.35	0.31	0.39	0.37	0.07
1st Order	0.07	0.08	0.06	0.04	0.03	0.04	0.03	0.02	0.07
MH	0.07	0.08	0.06	0.04	0.03	0.04	0.03	0.02	0.07
Error	0.04	0.02	0.03	0.01	0.04	0.02	0.02	0.03	0.04

Table 14: Allocation of effects of M (Mobility Model) and H (Routing Heuristic) on Route Length.

F-Table $\alpha=0.1$		F-Computed (Length)								
		10	20	30	40	50	60	70	80	90
9.46	M	386.01	760.14	412.73	749.64	282.89	495.30	664.74	375.64	386.01
9.46	H	34.34	176.12	181.67	416.41	167.53	237.67	456.56	237.95	34.34
3.81	MH	15.89	41.76	19.97	26.03	7.15	13.91	15.96	7.64	15.89

Table 15: Analysis of variance of effects for Route Length. All Results are statistically significant at a 90% confidence level.

Table 14 confirms that M is also the dominant factor affecting the route length, although H does come into play at higher densities. The rationale behind this was

discussed earlier in this section. Although it is not the dominant factor, the routing heuristic H is still important since it accounts for roughly 30% of the overall variation in route length.

4.4 Understanding Mobility Models

The results presented and discussed in Sections 4.2 and 4.3 indicate that the choice of M has a significant effect on the availability, stability, and length of routes in ad-hoc networks. This led us to pose the following question:

What underlying aspects of the mobility models could be responsible for the wide range of ad-hoc network performance observed in the previous experiments?

We suspect that the answer lies in the restricted manner in which the nodes are forced to move within the underlying simulation area. We have observed that availability and stability are lower in mobility models that have more restricted movement schemes, such as the Constrained Path or RPGM models. In Section 4.3, while discussing the results for the Constrained Path model, the concept of corridors was introduced. We extend this concept to RPGM by suggesting that since the nodes travel in groups, they are essentially travelling along a corridor – albeit one that was defined dynamically when the group was assigned its waypoint.

Regardless of whether a corridor is defined explicitly or dynamically, the end result is the same: groups of nodes have a restricted movement pattern that follows a pathway through their environment. From this observation we hypothesize:

Hypothesis: Restricting the movement of nodes within ad-hoc networks to narrow corridors will result in lower route availability and stability, and longer average route lengths.

To test this hypothesis we devised a set of experiments to characterize availability and stability in situations where the *topology* of the simulation area constrains the movement of nodes. These experiments are described in Section 4.5.

4.5 The Effect of Simulation Area Topology

To determine the effect of simulation area topology (T) on network performance, we chose to perform simulations similar to those from previous experiments, using (as much as possible) the same factors and levels listed in Chapter 3, with the following exceptions:

- The addition of simulation area topology (T) as a primary factor. Previous experiments set T to be a rectangle with dimensions $x=200\text{m}$ and $y=100\text{m}$, and a resulting area of 20000m^2 . We now consider different values for x and y , while keeping the overall area near 20000m^2 .
- The relegation of M to a secondary factor. This was done to prevent the mobility model from interfering with any effects caused by the new topologies. To accomplish this we set M to be a model that does not itself define any corridors, Random Waypoint.

The levels of the primary factors used for the simulation area topology experiments are listed in Table 16.

Factor	Level 1	Level 2	Level 3
<i>D (density)</i>	20	50	80
<i>T (topology)</i>	142x142	300x67	800x25
<i>H (heuristic)</i>	Shortest Path	Associative	Compass

Table 16: Factors and levels used in study of simulation area topology.

The objective functions being optimized are the same as before, namely A , S , and L . All simulations were repeated five times. The results were analyzed with ANOVA to determine the effects and combined effects of the primary factors. Table 17 summarizes the results, and Table 18 shows the computed F-values used in the analysis of variance. This test indicates that for these experiments, all non-zero effects are statistically significant.

Observe that the topology T has a large effect on the availability of routes in the network, even larger than the effect of density. The reader will recall from our earlier experiments that D had a much larger effect than any of the other factors. In the following sections this observation is investigated in more detail to ensure that it is not

simply a result of the limited number of levels that were selected for these preliminary experiments.

<i>Allocation of Effects</i>	Mean Availability	Mean Stability	Mean Route Length
Primary Effects	0.95	0.67	0.62
<i>D</i>	0.40	0.54	0.40
<i>T</i>	0.54	0.09	0.19
<i>H</i>	0.00	0.04	0.02
1st Order	0.05	0.29	0.37
<i>DT</i>	0.04	0.27	0.36
<i>DH</i>	0.00	0.01	0.01
<i>TH</i>	0.00	0.01	0.00
2nd Order	0.00	0.00	0.00
<i>DTH</i>	0.00	0.00	0.00
Error	0.01	0.04	0.01

Table 17: Allocation of effects of *D*, *T*, and *H* on mean availability, stability, and route length.

F-Table $\alpha=0.1$		F-Computed		
		<i>Availability</i>	<i>Stability</i>	<i>Length</i>
9.47	<i>D</i>	1006.03	785.37	3258.00
9.47	<i>T</i>	1344.24	129.43	1534.70
9.47	<i>H</i>	2.53	54.28	196.95
3.80	<i>DT</i>	111.12	198.58	1440.46
3.80	<i>DH</i>	0.49	6.71	41.96
3.80	<i>TH</i>	1.20	5.07	8.97
2.36	<i>DTH</i>	0.55	0.47	5.95

Table 18: Analysis of variance of effects of *D*, *T*, and *H* on *A*, *S*, and *L*. Results that are not shaded are statistically significant at a 90% confidence level.

Route stability and route length also appear to be directly affected by *T*, although not to the extent that *A* is. The reader will notice a significantly large combined effect *DT* for both *S* and *L*. This suggests a strong relationship between the node density and the topology of the simulation area. Again, this relationship will be explored in the sections that follow.

The final observation to be made regarding this analysis of effects is the relatively minor influence of the routing heuristic. We conclude from this that researchers ought to pay more attention to the shape of the simulation areas that are used in studies that have

the goal of quantifying the performance of routing protocols and heuristics in ad-hoc networks.

These results also support the claim that the mobility model has a major effect on ad-hoc network routing performance. In particular, this result establishes that it is likely the constraint on node movement imposed by the logical and/or physical corridors that is responsible for the differences in performance noted in Section 4.3.

4.6 Understanding Simulation Area Topology

To explain the effects of the simulation area topology, we ran additional simulations at the density levels used in the previous study of mobility models. The scope of these new experiments was limited to rectangle topologies with a wide range of length and width dimensions. Simulations were run for all combinations of levels listed in Table 19.

<i>Factor</i>	<i>Levels Used for Simulations</i>
<i>D</i>	10, 20, 30, 40, 50, 60, 70, 80, 90
<i>T</i>	142x142, 200x100, 300x67, 400x50, 500x40, 600x33, 700x29, 800x25
<i>H</i>	Shortest Path, Compass, ABR

Table 19: Levels used for simulations involving topologies and densities.

The trends for route availability are illustrated in Figure 21. All results are based on simulations involving shortest path routing.

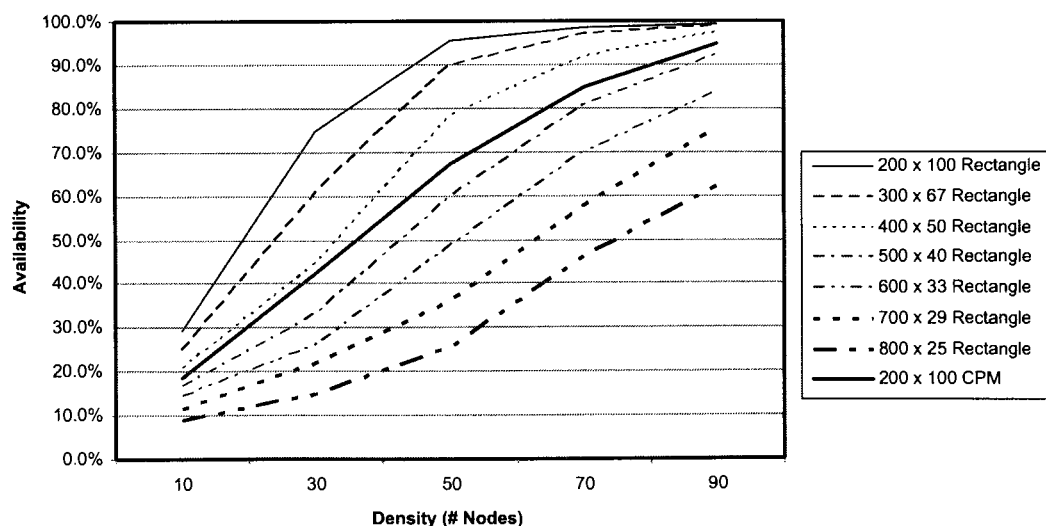


Figure 21: Route availability versus density for various simulation area topologies.

As expected, availability decreases as the topology becomes narrower. This is similar to the earlier results for mobility models that incorporate defined corridors – although the effect is much more pronounced in these simulation area topology experiments. This decrease in availability can be attributed to the reduced connectivity inherent in the narrower networks. With the lower degree of connectivity it only takes one node-pair to become disconnected to cause many routes to break, because many routes must use the same nodes to relay messages within the narrow simulation area. In the worst case rectangle (800x25) nodes are essentially forced into a line that follows the corridor. Portions of this line will be used to construct the various routes between nodes in the network. If any two adjacent nodes in the line move out of range of each other, there is a good chance that the network will become disconnected, invalidating many routes.

For comparison purposes Figure 21 includes a data series for the 200x100 Constrained Path model analyzed in Section 4.3. This CPM result has an availability that is similar to the availability of a 500x40 rectangle (which uses Random Waypoint). The movement pattern in the CPM model is certainly more complex than the movement through the 500x40 rectangle, however, it is worth noting that the corridors in our CPM scenario range from 10 metres to 40 metres wide, and have a total length of 740 metres.

In any case, the results in Figure 21 support the hypothesis that long narrow corridors induce lower availability rates in ad-hoc networks. To test the rest of the hypothesis (which states that narrow corridors induce lower stability and longer route lengths), the simulation results were analyzed for trends in stability and route length. Our results are documented in Figure 22 and Figure 23.

Figure 22 indicates that at low densities, stability is best in narrow topologies. As density starts to increase, the stability decreases. At higher densities the narrow topologies produce very unstable routes, regardless of the routing heuristic used to determine those routes. This explains the large combined effect that is evident in Table 17, and suggests that there are topologies that exhibit good routing characteristics for low density ad-hoc networks, and different topologies that allow higher density ad-hoc networks to perform well.

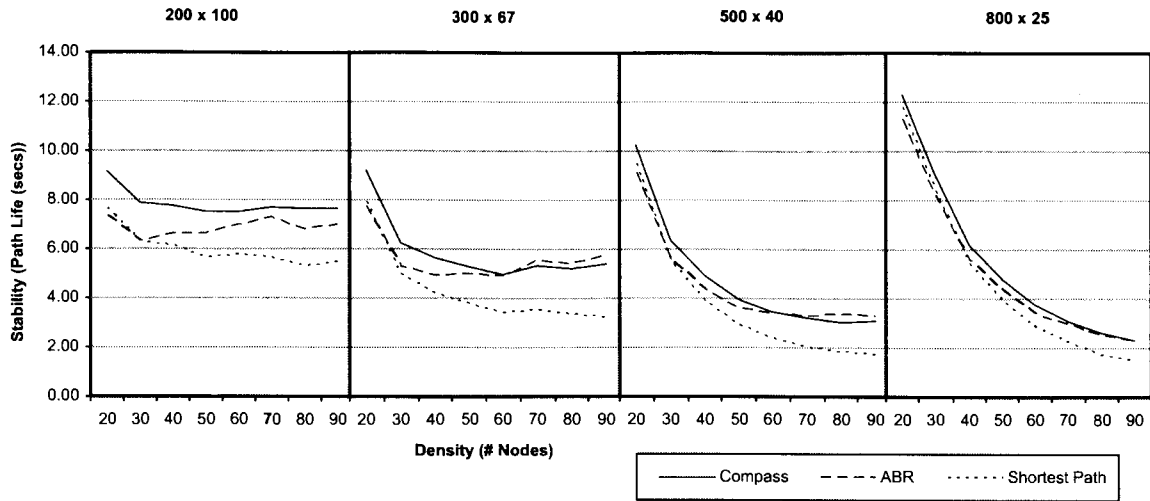


Figure 22: Route stability in rectangle topologies. Results are shown for compass, associative, and shortest path heuristics. Densities from 20 to 90 nodes per 20000m² are included.

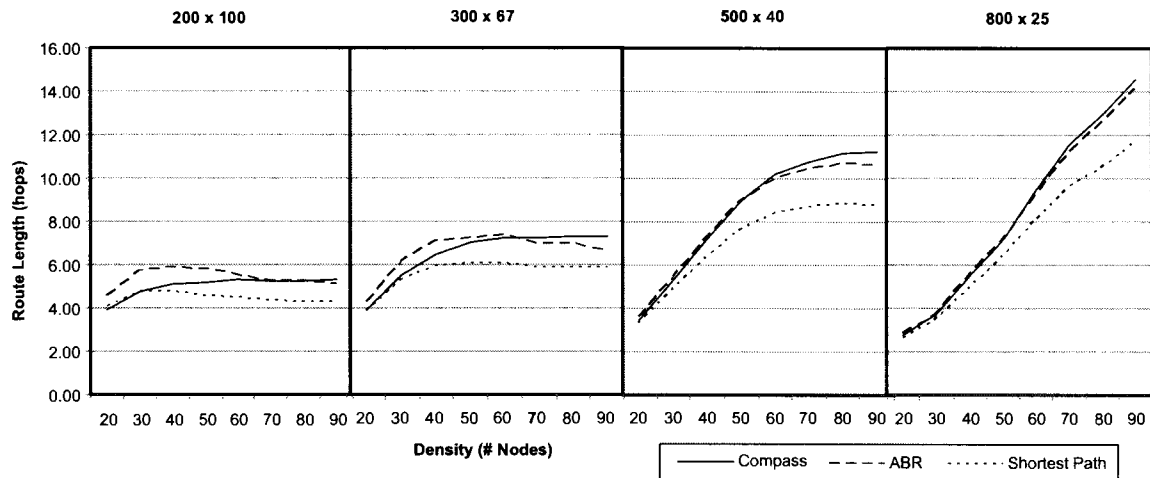


Figure 23: Route length in rectangle topologies. Results are shown for compass, associative, and shortest path heuristics. Densities from 20 to 90 nodes per 20000m² are included.

Figure 23 illustrates the extreme route lengths that can occur in high density, narrow topology networks. Recall that the lengths charted in these graphs are *mean* route lengths, with the maximum route lengths being much longer. This long route length and low stability suggest to the author that communication in high density, narrow topology networks is unlikely to be effective using the routing protocols that exist today due to the overhead required to establish and maintain routes that have a mean life of 1.43 seconds (worst case).

We now revisit our hypothesis from Section 4.4, which reads:

Hypothesis: Restricting the movement of nodes within ad-hoc networks to narrow corridors will result in lower route availability and stability, and longer average route lengths.

We conclude that the results documented and discussed in this section support the hypothesis, and from this we deduce that the performance observed when using the CPM model is likely a result of the restricted node movement induced by the corridors in the CPM model. We also note the relative importance of simulation area topology, which appears to affect the availability of routes in ad-hoc networks more than node density does.

CHAPTER 5

CONCLUSIONS

5.1 Summary of Results

In the simulation of ad-hoc networks, the availability and stability of routes is highly dependent on the mobility model used. Specifically, we have demonstrated that the Random Waypoint model provides an overly optimistic environment in which to test routing protocols, and should be avoided. More realistic models such as the Constrained Path model should be adopted and developed further to add credibility to the research that is being done in the field of ad-hoc networks.

We also find that the shape and size of the simulation area is an important consideration in ad-hoc network modelling, as long narrow corridors induce lower availability and stability, and longer route length. It is expected that this effect is partially responsible for the degradation in performance observed when models that constrain node movements are used.

The associativity heuristic used by ABR claims to discover routes that are inherently more stable than those found by other protocols. We found this to be true when compared against routes discovered using the shortest path heuristic, which is the basis of many of the existing ad-hoc routing protocols. However when compared against another relatively uncommon heuristic, *compass direction* to destination, the associativity rule did not live up to expectations.

5.2 Future Work

A number of interesting research directions emerge from this study. In particular, additional research should be performed to discover attributes of the constrained pathways that can be exploited in the development of communications protocols. As well, the Constrained Path model could be enhanced to include the notion of obstacles that

block node transmissions, and it should be benchmarked against other similar models, such as the Obstacle and Graph models.

To further study the Constrained Path model it should be adapted for use with a simulator such as *ns-2*, and the actual per-node relaying load and throughput should be quantified.

We also feel that additional investigation into the concept of associativity is required. It would be interesting to measure the associativity levels of paths found by other heuristics such as shortest path and or compass, to determine if it is the associativity effect that is leading to the path stability results observed in our experiments.

BIBLIOGRAPHY

- [1] S. Agarwal, A. Ahia, J. Singh, and R. Shorey, "Routelifetime Assessment Based Routing (RABR) Protocol for Mobile Ad-hoc Networks", *In Proc. IEEE International Conference on Communications 2000 (ICC'00)*, Vol. 3, pp. 1697-1701, 2000.
- [2] S. Basagni, I. Chlamtac, V. Syrotiuk, and B. Woodward, "A Distance Routing Effect Algorithm for Mobility (DREAM)", *In Proc. of the 4th Annual ACM/IEEE Int. Conf. on Mobile Computing and Networking (MOBICOM'98)*, pp. 76-84, Dallas, TX, USA, 1998.
- [3] C. Bettstetter and R. Krausser, "Scenario-based Stability Analysis of the Distributed Mobility-Adaptive Clustering (DMAC) Algorithm", *In Proc. of the 2nd ACM Symposium on Mobile Ad-hoc Networking*, pp. 232 – 241, Long Beach, CA, 2001.
- [4] J. Broch, D. Maltz, D. Johnson, Y. Hu, and J. Jetcheva, "A Performance Comparison of Multi-hop Wireless Ad-hoc Network Routing Protocols", *In Proc. of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)*, pp. 85-97, Oct. 1998.
- [5] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad-hoc Network Research," *Wireless Comm. and Mobile Computing (WCMC): Special Issue on Mobile Ad-hoc Networking: Research, Trends, and Applications*, vol. 2, no. 5, pp. 483-502, 2002.
- [6] S. Capkun, M. Hamdi and J. Hubaux, "GPS-Free Positioning in Mobile ad-hoc Networks", *In Proc. of the 34th Annual Hawaii International Conference on System Sciences (HICSS-34)*, 2001.
- [7] D. Cavin, Y. Sasson, and A. Schiper, "On the Accuracy of MANET Simulators", *In Proc. of the Workshop on Principles of Mobile Computing (POMC'02)*, ACM, Oct. 2002, pp. 38-43.
- [8] T. Clausen, P. Jacquet, "Optimized Link State Routing Protocol (OLSR)", *Internet Engineering Task Force (IETF)*, rfc3626, Oct. 2003.
- [9] M. de Berg, M. van Kreveld, M. Overmars and O. Schwarzkopf, "Computational Geometry: Algorithms and Applications", *Springer Verlag*, 2000.
- [10] S. Das, C. E. Perkins, and E. Royer, "Performance Comparison of Two On-demand Routing Protocols for Ad-hoc Networks", *In Proc. of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000)*", pp. 3-12, 2000.
- [11] R. Dibe, C. Rais, K. Wang, and S. Tripathi, "Signal Stability Based Adaptive Routing (SSA) for Ad-hoc Mobile Networks", *IEEE Personal Communications*, Feb. 1997.

- [12] H. Dubois-Ferriere, M. Grossglauser and M. Vetterli, "Age Matters: Efficient Route Discovery in Mobile Ad-hoc Networks using Encounter Ages", *In Proc. of the 4th ACM International Symposium on Mobile Ad-hoc Networking (MobiHOC 2003)*, pp. 257-266, Annapolis, MD, USA, June 2003.
- [13] G. Finn, "Routing and Addressing Problems in Large Metropolitan-scale Internetworks", *ISI Research Report*, ISU/RR-97-180, Mar. 1987.
- [14] M. Gerharz, C. de Waal, M. Frank and P. Martini, "Link Stability in Mobile Wireless Ad-hoc Networks", *In Proc. of the 27th IEEE Conference on Local Computer Networks (LCN) 2002*, pp. 30-39, Tampa, Florida, Nov. 2002.
- [15] S. Giordano, I. Stojmenovic, L. Blazevic, "Position-based Routing Algorithms for Ad-hoc Networks: A Taxonomy", *Ad-hoc Wireless Networking*, X. Cheng, X. Huang and D.Z. Du (eds.), Kluwer, Dec. 2003.
- [16] I. Gruber and H. Li, "A Statistical Approach to Forecast Link Availability in Ad-hoc Networks", *WLN*, Sept. 2002.
- [17] Z. Haas, "A new routing protocol for the reconfigurable wireless networks", *In Proc. of the IEEE Int. Conf. on Universal Personal Communications.*, Oct 1997.
- [18] Z. Haas and B. Liang, "Ad-hoc Mobility Management with Uniform Quorum Systems", *IEEE/ACM Trans. on Networking*, Vol. 7 Nr. 2, pp. 228-240, April 1999.
- [19] Z. Haas and M. Pearlman, "Determining the Optimal Configuration for the Zone Routing Protocol", *IEEE Journal on Selected Areas in Communications*, Aug. 1999.
- [20] A. Helmy, S. Garg, P. Pamu, and N. Nahata, "Contact Based Architecture for Resource Discovery (CARD) in Large Scale MANets", *In Proc. of 3rd IEEE/ACM International Workshop on Wireless, Mobile and Ad-hoc Networks (WMAN)*, pp. 219-227, April 2003.
- [21] X. Hong, M. Gerla, G. Pei, and C. Chiang. "A Group Mobility Model for Ad-hoc Wireless Networks", *In Proc. of the ACM International Workshop on Modeling and Simulation of Wireless and Mobile Systems (MSWiM)*, August 1999.
- [22] T.-C. Hou and V. O.K. Li, "Transmission Range Control in Multihop Packet Radio Networks", *IEEE Trans. on Communications*, Vol. 34 Nr. 1, pp. 38-44, January 1986.
- [23] Y. Hu and D. Johnson, "Implicit Source Routes for On-Demand Ad-hoc Network Routing", *In Proc. of the 2nd ACM International Symposium on Mobile Ad-hoc Networking & Computing (MOBIHOC'02)*, pp. 1-10, Long Beach, CA, USA, 2001.
- [24] A. Iwata, C. Chiang, G. Pei, M. Gerla, and T. Chen, "Scalable Routing Strategies for Ad-hoc Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Aug. 1999, pp. 1369-1379.
- [25] R. Jain, "The Art of Computer Systems Performance Analysis", *John Wiley & Sons*, 1991.
- [26] A. Jardosh, E. Belding-Royer, K. Almeroth, and S. Suri. "Towards Realistic Mobility Models for Mobile Ad-hoc Networks", *In Proc. of The 9th International*

Conference on Mobile Computing and Networking (Mobicom), San Diego, CA, pp. 217-229, Sept 2003.

- [27] 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 the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '99)*, pp. 195-206, Aug. 1999.
- [28] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad-hoc Wireless Networks", *Mobile Computing*, Chapter 5, pp.153-181. Kluwer Academic Publishers, 1996.
- [29] B. Karp, "Geographic Routing for Wireless Networks", *PhD Thesis*, Harvard University, 2000.
- [30] L. Kleinrock and K. Stevens, "Fisheye: A Lenslike Computer Display Transformation", *Technical Report*, UCLA, Computer Science Department, 1971.
- [31] E. Kranakis, H. Singh, and J. Urrutia, "Compass Routing on Geometric Networks". *In Proc. of 11th Canadian Conf. on Computational Geometry*, Vancouver, August 1999.
- [32] J. Li, J. Jannotti, D. De Couto, D. Karger, and R. Morris, "A Scalable Location Service for Geographic Ad-hoc Routing", *In Proc. of the 6th Annual ACM/IEEE Int. Conf. on Mobile Computing and Networking (MOBICOM 2000)*, pp. 120-130, Boston, MA, USA, 2000.
- [33] K. Mase, Y. Wada, N. Mori, K. Nakano, and M. Sengoku, "Flooding Schemes for Clustered Ad-hoc Networks", *IEICE Transactions on Communications*, Vol. E85-B Nr. 3, Mar. 2002.
- [34] M. Mauve, J. Widmer, and H. Hartenstein, "A Survey on Position-Based Routing in Mobile Ad-Hoc Networks", *IEEE Network*, Vol. 15 Nr. 6, pp. 30-39, 2001.
- [35] J. Moy, "OSPF Version 2", *Internet Engineering Task Force (IETF)*, rfc1583, Mar 1994.
- [36] S. Murthy and J. J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks", *ACM Mobile Networks and Applications Journal, Special issue on Routing in Mobile Communication Networks*, Vol. 1, No. 2, 1996.
- [37] D. Oran, "OSI IS-IS Intra-domain Routing Protocol", *Internet Engineering Task Force (IETF)*, rfc1142, Feb. 1990.
- [38] S. K. Park and K. W. Miller, "Random Number Generators: Good Ones are Hard to Find," *CACM 31:10*, Oct. 88, pp. 1192-1201.
- [39] V. Park and M. S. Corson, "Temporally-ordered routing algorithm (TORA) version 1: Functional Specification", *IETF*, Internet Draft, draft-ietf-manet-tora-spec-01.txt, August 1998.
- [40] G. Pei, M. Gerla, and T.-W. Chen, "Fisheye State Routing: A Routing Scheme for Ad-hoc Wireless Networks", *In Proc. of the IEEE International Conference on Communications*, pp. 70-74, New Orleans, June 2000.

- [41] G. Pei, M. Gerla, and X. Hong, "LANMAR: Landmark Routing for Large Scale Wireless Ad-hoc Networks with Group Mobility", *In Proc. of IEEE/ACM Workshop on Mobile Ad-hoc Networking & Computing (MobiHoc 2000)*, Boston, MA, Aug. 2000.
- [42] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers", *In Proc. of ACM SIGCOMM'94*, London, UK, Sept. 1994, pp. 234-244.
- [43] C. E. Perkins and E. Royer, "Ad-Hoc On-Demand Distance Vector Routing", *Proc. of IEEE WMCSA '99*, New Orleans, Feb. 1999, pp. 90-100.
- [44] E. Royer, P. Melliar-Smith, and L. Moser, "An Analysis of the Optimum Node Density for Ad-hoc Mobile Networks", *In Proc. of the IEEE International Conference on Communications*, Helsinki, Finland, June 2001.
- [45] M. Sánchez, "<http://www.disca.upv.es/misan/mobmodel.htm>", *Universidad Politécnica de Valencia*, Accessed January 2003.
- [46] A. Schmid and C. Steigner, "Avoiding Counting to Infinity in Distance Vector Routing", *Telecommunication Systems 19, Vol. 3-4*, pp. 497-514, Kluwer Academic Publishers, March – April 2002.
- [47] I. Stojmenovic, "Home Agent Based Location Update and Destination Search Schemes in Ad-hoc Wireless Networks", *Technical Report TR-99-10*, Computer Science, SITE, University of Ottawa, September 1999.
- [48] I. Stojmenovic, M. Seddigh, and J. Zunic, "Internal Nodes Based Broadcasting in Wireless Networks", *In Proc. of the 34th Annual Hawaii International Conference on System Sciences (HICSS-34)*, 2001.
- [49] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals", *IEEE Trans. on Communications*, Vol. 32 Nr. 3, pp. 246-257, March 1984.
- [50] K. Thulasiraman and M. Swamy, "Graphs : Theory and Algorithms", *Wiley-Interscience*, Feb. 1992.
- [51] J. Tian, J. Hahner, C. Becker, I. Stepanov, and K. Rothermet, "Graph-Based Mobility Model for Mobile Ad-hoc Network Simulation", *In Proc. of the 35th Annual Simulation Symposium*, April 14-18, 2002 San Diego, California. pp. 0337
- [52] C.K. Toh, "Associativity-Based Routing For Ad-Hoc Mobile Networks", *Wireless Personal Communications*, Vol. 4, No. 2, March 1997.
- [53] D. Turgut, S. Das and M. Chatterjee, "Longevity of Routes in Mobile Ad-hoc Networks", *In Proc. IEEE VTC 2001*, pp. 2833-2837, Spring 2001.
- [54] A. Vasan and A. UdayaShankar, "An Empirical Characterization of Instantaneous Throughput in 802.11b WLANs", *Univ. of Maryland*, Technical Report CS 4389.
- [55] Y. Yi, M. Gerla and T. Kwon, "Efficient Flooding in Ad-hoc Networks using On-Demand (Passive) Cluster Formation", *In Proc. of IEEE/ACM Workshop on Mobile*

Ad-hoc Networking & Computing (MobiHoc'03), Annapolis, Maryland, USA, June 2003.

[56] J. Yoon, M. Liu, and B. Noble, "Random Waypoint Considered Harmful", *In Proc. of IEEE INFOCOM*, April 2003, San Francisco, CA.

[57] D. Yu and H. Li, "Path Availability in Ad-hoc Networks", *In Proc. of the 10th International Conference on Telecommunications (ICT2003)*, Tahiti, Feb. 2003.