Routing Protocols for
Sparse Mobile Ad hoc Networks

Saeed Shahbazi
Department of Computer Science and Software Engineering
The University of Melbourne

Submitted in total fulfillment of the requirements of the degree of

Doctor of Philosophy

August 2011
Abstract

Delay/Disruption Tolerant Networks (DTNs) is a class of Mobile Ad hoc Networks (MANETs) in which the ubiquitous connectivity is challenged due to the lack of end-to-end routes between pairs of source and destination nodes. In DTNs, the mobile nodes are disconnected for significantly long periods of time. As a result, traditional routing protocols proposed for MANETs, which assume at least one steady path between pairs of source and destination nodes, are ineffective in DTNs. One approach to improve communication in DTNs is through gossip based protocols because these protocols do not rely on a fixed path. However, this approach suffers from poor delivery performance as it is dependent on opportunistic communication. Another approach is to control the movement of the mobile nodes and/or use special mobile nodes called ferry nodes. However, in some practical applications it is not possible to force the mobile nodes to change their trajectory and therefore it is desirable to have a more easily deployable and scalable protocol which does not depend on changing the trajectory of mobile nodes. In this thesis, we first review important challenges of DTNs and survey the existing applications of this domain. Furthermore, we provide a taxonomy of existing DTN routing protocols and discusses their strengths and weaknesses in the context of their assumptions and applicability in DTN applications.

We propose a novel routing protocol called Tag Based Routing (TBR) for DTNs which overcomes some of the limitations of existing DTN routing protocols. In TBR, the forwarding opportunities are increased by augmenting the DTNs with easily deployable and independent stationary relay nodes. Specifically, we propose an approach where mo-
bile nodes deposit/retrieve messages to/from known stationary passive locations in the geographic region, which have no attached battery and are thereby suitable for long term usage. Messages are delivered from a source by being deposited at one or more locations that are later visited by the destination. A proposed implementation of our approach using read/writable passive Radio Frequency Identification (RFID) tags, one per point location, is considered in this work. Passive RFID technology is desirable because it operates wirelessly and without the need for attached power.

We have conducted a comprehensive study of the performance of our protocol using simulations as well as analytical models under many different scenarios. We also compare our protocol with state-of-the-art approaches. As part of our study we present and implement a variety of mobility models as the delivery performance of DTN routing protocols is highly dependent on the mobile nodes’ mobility models. Our evaluation results show that our protocol is fault tolerant, high performance, and scalable, while it can achieve competitive message delivery performance in comparison to existing approaches in the literature.

The performance of TBR is highly dependent on the location of passive relay points. Relay placement is NP-hard problem hence it makes it a more complicated issue in DTNs. In order to improve the performance of TBR we demonstrate several techniques for optimizing the stationary relay node placement, namely relay pruning, probability based relay distribution, and a genetic algorithm based optimization and we show their effect on the performance of the proposed protocol as well as comparing their effectiveness with the existing placement techniques in the literature. Our simulation results show that an optimized relay placement technique can significantly improve the performance of TBR in both message delivery rate and average message delay. Further, we show that our genetic algorithm based optimization placement technique provides the best solutions to this problem.
Additionally, we propose efficient buffer management policies to increase the performance of our proposed protocol. Mobile nodes in TBR may store the messages for a long time in their buffer until they find a contact opportunity. Long term storage accompanied by message replication needs a high buffer overhead; specifically, passive relay nodes have a more limited buffer storage than mobile nodes which intensifies the buffer overhead. Therefore, using efficient buffer policies is necessary to decide which messages should be dropped in case of overflowing. We also show that traditional buffer management policies like drop-tail or drop-front fail to consider all relevant information in this context and are, thus, sub-optimal. Our simulation results show that our buffer management policies including queuing policies and forwarding strategies significantly improve the delivery performance of TBR.

The last chapters of this dissertation focus on deriving analytical expressions to model the performance of TBR. Firstly, we propose an analytical model to evaluate the performance of our proposed protocol and provide mathematical equations to calculate the performance of the protocol. The proposed analytical model can overcome the burden of simulation based evaluations in terms of both time and resource complexities. Secondly, we derive an analytical model that enables us to evaluate the performance of relay placement techniques in DTNs. We then use the proposed analytical model to optimize the relay placement techniques. Thereby, we propose a number of semi-heuristics approaches which utilize our model to optimize the placement. Our simulation results show that our analytical based placement approaches outperform the simulation based approaches in terms of data delivery performance.
This is to certify that:

i. the thesis comprises only my original work towards the PhD,

ii. due acknowledgement has been made in the text to all other material used,

iii. the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies, appendices, and footnotes.

Signed,

__________________________________
Saeed Shahbazi
4th August 2011
I would like to dedicate this thesis to my loving parents ...
Acknowledgements

It is a pleasure to thank all those who made this thesis possible. Particularly, I would like to express my deepest gratitude to my supervisors Dr. Aaron Harwood and Dr. Shanika Karunasekera for having given me the opportunity to work on this exciting research topic.

I must acknowledge the generous financial support that I have received from the University of Melbourne in the form of two scholarships (MIFRS and MIRS). In addition, both faculty of Engineering and Department of Computer Science and Software Engineering supported me with several travel grants during my PhD. Their supports allowed me to explore wonderful new opportunities and to achieve valuable experiences.

I would also like to acknowledge the financial support of National ICT of Australia (NICTA) who provided me a top-up scholarship and the opportunity to attend the conferences where I had publication.

Last, but by no means least, I would like to express my absolute appreciation to my beloved parents and my dearest friends who have supported me in every possible way during my studies. Their unconditional love, continuous belief, and spiritual support have sustained me for traveling thus far. Particularly, I would like to express my deepest gratitude to my dear friend, Masud Moshtaghi, for his support during my last years of studies.
# Contents

Nomenclature

1 Introduction
   1.1 Background and Context
   1.2 Motivation and Problem Statement
   1.3 Research Objective and Methodology
      1.3.1 MATLAB Simulation Environment
   1.4 Contribution of the Thesis
   1.5 Structure of the Thesis
   1.6 List of Publications
      1.6.1 Published Papers
      1.6.2 Papers under Review

2 Delay/Disruption Tolerant Networks and Routing Paradigms
   2.1 DTNs Research Background
      2.1.1 Our Definition of DTNs
   2.2 Applications of DTNs
      2.2.1 Vehicle-based DTNs (VDTNs)
      2.2.2 Delay Tolerant Networking for Sensor Networks
      2.2.3 Communication between rural zones in developing regions
      2.2.4 Deep space communication
   2.3 Routing in DTNs
      2.3.1 Contact Opportunities
      2.3.2 Evaluation Metrics for DTN Routing Protocols
      2.3.3 Classification of DTN Routing Protocols
## CONTENTS

2.3.3.1 Store-Carry-Forward ........................................ 28
2.3.3.2 Stationary Relay Points ................................. 34
2.3.3.3 Heterogeneous Approaches ............................... 35
2.3.4 Discussion .................................................. 36

3 Achieving Ubiquitous Network Connectivity using an RFID Tag-based Routing Protocol .......................... 40
3.1 Tag Based Routing Protocol ................................. 41
   3.1.1 Protocol Description .................................. 41
   3.1.2 Routing Protocol Design Principles in TBR .............. 45
3.2 Evaluation .................................................. 47
   3.2.1 TBR Implementations .................................. 47
   3.2.1.1 RFID MANET Implementation ....................... 47
   3.2.2 Simulation Methodology ................................ 50
   3.2.3 TBR Evaluation ........................................ 51
   3.2.3.1 Control Variables ................................ 51
   3.2.3.2 TBR Effectiveness ................................ 52
   3.2.3.3 Transmissions per Second ......................... 57
   3.2.3.4 Power Consumption vs. Number of Relays ........... 57
   3.2.3.5 Distribution of Message Latency .................... 58
3.2.4 Possibility of Using a Single Base Station to Cover the Area .................................................. 61
3.3 Comparisons to Existing Work ............................... 64
   3.3.1 Comparison to Epidemic Routing ........................ 64
   3.3.2 Comparison to Message Ferrying ........................ 67
3.4 Conclusion .................................................. 68

4 Heuristics for Improving TBR Delivery Performance under Different Mobility Models .................................. 69
4.1 Mobility Models ............................................. 70
   4.1.1 Entity Mobility Models ................................ 71
   4.1.1.1 Random Waypoint Model ........................... 71
   4.1.1.2 Random Walk Model ................................ 71
   4.1.1.3 Restricted Random Waypoint Model ............... 73
   4.1.2 Group Mobility Models ................................ 74

viii
## CONTENTS

4.1.2.1 Reference Point Group Mobility Model ........................................ 74
4.1.3 Discussion ......................................................................................... 74
4.2 Relay Placement Techniques .............................................................. 76
  4.2.1 Uniform Grid ................................................................................... 76
  4.2.2 Relay Pruning .................................................................................. 79
  4.2.3 Probability based Relay Distribution .............................................. 80
  4.2.4 Genetic Algorithm based Optimization ............................................ 81
  4.2.5 Discussion ....................................................................................... 82
4.3 Evaluation of Relay Placement Techniques using Various Mobility Models ................................................................. 84
4.4 Queuing Policies and Forwarding Strategies ......................................... 89
  4.4.1 Queuing Policies ............................................................................. 89
  4.4.2 Forwarding Strategies .................................................................... 91
    4.4.2.1 DDM - Delete Delivered Messages ........................................... 93
    4.4.2.2 INF - Intermittent Forwarding .................................................. 93
4.5 TBR Robustness .................................................................................. 94
4.6 Conclusion ........................................................................................... 97

5 An Analytical Model for Performance Evaluation of TBR ......................... 99
  5.1 Introduction ........................................................................................ 100
  5.2 TBR Assumptions .............................................................................. 100
    5.2.1 Performance Metrics .................................................................... 101
      5.2.1.1 Throughput ........................................................................... 101
    5.2.2 Random Round Point Mobility Model ....................................... 102
    5.2.3 Buffer Policy ............................................................................... 102
  5.3 Our Model .......................................................................................... 103
    5.3.1 System Model .............................................................................. 103
    5.3.2 Buffer Policy-Keep Old Messages .............................................. 103
      5.3.2.1 Performance Metrics .............................................................. 106
    5.3.3 Buffer Policy-Over Write Permitted .......................................... 107
      5.3.3.1 Throughput .......................................................................... 108
      5.3.3.2 End-to-end Delay ................................................................. 109
  5.4 1D Network ....................................................................................... 110
5.4.1 Simulation ................................. 110
  5.4.1.1 Buffer Policy-Keep Old Messages ........ 111
  5.4.1.2 Buffer Policy-Over Write Permitted ....... 112
5.4.2 Extended Model-Random Tag Placement ........ 113
5.5 2D Networks ................................. 113
  5.5.1 Simulation ................................. 115
    5.5.1.1 Buffer Policy-Keep Old Messages ....... 116
    5.5.1.2 Buffer Policy-Over Write Permitted .... 117
5.6 Multiple Tags ............................... 117
  5.6.1 Maximum Number of Tags .................... 117
  5.6.2 Performance Metrics ....................... 118
  5.6.3 Simulation ............................... 119
    5.6.3.1 Buffer Policy-Keep Old Messages ........ 119
    5.6.3.2 Buffer Policy-Over Write Permitted .... 119
5.7 Conclusion .................................. 119

6 Analytical Approach for Passive Stationary Relay Point Placement in DTNs 122
  6.1 Introduction ............................... 123
  6.2 Problem Statement ......................... 124
  6.3 Modeling Relay Performance ................ 125
    6.3.1 An Analytical Model for Relay Placement .... 125
      6.3.1.1 Conceptual Mobility Model .............. 127
      6.3.1.2 Expressions for DTNs delivery performance ... 127
      6.3.1.3 Objective Function ..................... 132
  6.4 Analytical based Relay Placement Approaches .... 132
    6.4.1 Simulated Annealing based Optimization .... 133
    6.4.2 Greedy Placement ....................... 133
      6.4.2.1 Complexity of the Proposed Techniques .... 134
  6.5 A Case Study ................................ 134
    6.5.1 Extending RRP to Random Waypoint Model .... 135
    6.5.2 Relay Position vs. its Performance ......... 135
    6.5.3 Experiments ............................ 137
<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5.3.1 Performance Evaluation</td>
<td>137</td>
</tr>
<tr>
<td>6.5.3.2 Comparison to Existing Simulation-based Approaches</td>
<td>139</td>
</tr>
<tr>
<td>6.6 Conclusions</td>
<td>139</td>
</tr>
<tr>
<td>7 Conclusion and Future Directions</td>
<td>141</td>
</tr>
<tr>
<td>7.1 Conclusion</td>
<td>141</td>
</tr>
<tr>
<td>7.2 Future Directions</td>
<td>143</td>
</tr>
<tr>
<td>References</td>
<td>158</td>
</tr>
<tr>
<td>A Dashed Region Area</td>
<td>159</td>
</tr>
<tr>
<td>B Genetic Algorithm based Optimization</td>
<td>163</td>
</tr>
</tbody>
</table>
# List of Figures

1.1 Example of dense and sparse MANET nodes as a result of distance from a base station. Sparseness may also arise when wireless signals are obstructed, when signal power constraints are applied and when node mobility is restrained. ........................................... 3
1.2 Mining scenario as an application of our protocol. ................. 8
1.3 TBR Structure Modules. .................................................... 14
2.1 Packet delivery ratio vs. link fraction to all potential contacts between each pair of nodes. ................................. 22
2.2 Our classification of DTN routing protocols. ....................... 38
3.1 Aspects of the Tag Based Routing model. ......................... 43
3.2 Message Delivery Rate vs. Average MN Speed. .................... 53
3.3 Message Delivery Rate vs. RN Transmission Range. ............ 54
3.4 Message Delivery Rate vs. RN Buffer Size. ........................ 54
3.5 Message Delivery Rate vs. RN Spacing. ............................. 55
3.6 Average Message Delay vs. Average MN Speed. .................. 55
3.7 Average Message Delay vs. RN Transmission Range. ............ 56
3.8 Average Message Delay vs. RN Buffer Size. ........................ 56
3.9 Average Message Delay vs. RN Spacing. ............................. 57
3.10 Transmissions per second required by a mobile device........... 58
3.11 The average relay hit number per second versus the number of relays placed on a regular grid. ......................... 58
3.12 CDF of message latency vs. various parameters .................. 60
3.13 PDF of message latency fit to a Generalized Extreme Value distribution. In this example, the distribution parameters are $k = -0.10493$, $\sigma = 24.1557$ and $\mu = 39.6301$. 

3.14 Network Inefficiency vs. Base Station Broadcasting Range. 

3.15 Average Message Delay vs. Base Station Broadcasting Range. 

3.16 Message Delivery Rate vs. Base Station Broadcasting Range. 

3.17 Deploying RNs is an alternative when it is not possible to deploy a Base Station in the center of an area. 

3.18 TBR’s CDF for message delivery rate versus message delay, (a) and (b). Note that these CDFs cumulate to the percentage of delivered messages. Comparisons to existing work are shown in (c) and (d). 

4.1 Traveling pattern of a MN using the Random Way Point mobility model (50 steps). Red square indicates the start location and blue square shows the end location of the MN. 

4.2 Traveling pattern of a MN using the Random Walk mobility model (50 steps). Red square indicates the start location and blue square shows the end location of the MN. 

4.3 Traveling pattern of 4 MN using the Restricted Random Way Point mobility model (50 steps). Red squares indicate the start locations and blue squares show the end locations of the MNs. 

4.4 Individual node movement. 

4.5 Traveling pattern of 3 MN in a group using the Reference Point Group Mobility model (50 steps). 

4.6 RN usage frequency distribution for 121 RNs. 

4.7 RN positions after pruning. 

4.8 Probability based relay distribution. 

4.9 Effect of non-uniform relay distribution. 

4.10 Effect of GA approach and probability based distribution on placement of RNs. 

4.11 An square area is partitioned to 4 equal sized sub-areas which has an overlap equals to 40% of area length. Node mobility is allowed only inside its respective sub-area.
4.12 The process of measuring the network performance for different placement techniques. .................................................. 86
4.13 Delivery performance of MOFO versus FIFO respecting to average MN speed. ....................................................... 91
4.14 Delivery performance of MOFO versus FIFO respecting to average RN buffer size. .................................................... 92
4.15 Delivery performance of MOFO versus FIFO respecting to the number of RNs. .......................................................... 92
4.16 By deleting the copies of previously delivered messages the performance of TBR is increased. We used the following setting: \( N_{tag} = 36, N_{node} = 10, X/Y_{min/max} = 1000m, d = 200, B_{node} = 1000, B_{relay} = 100, r = 50, S_{min/max} = [20, 30], \lambda = 0.5, \) and the mobility model used is random waypoint. Relays are placed on a regular grid. ................................. 93
4.17 By controlling the transmission number we can address the overhead/power consumption of TBR. ............................ 94
4.18 Message Delivery Rate vs. RN Failure Rate ........................................ 96
4.19 Message Latency vs. RN Failure Rate ....................................... 97

5.1 (a) The probability of reading from the tag converges to \( \frac{m}{2} \). (b) The probability of reading from the tag converges to \( \frac{m}{2-m} \). .... 106
5.2 The network has only one dimension with the length of \( l \) and the tag is placed at the beginning of the network which has a broadcasting range of \( r \). .................................................. 110
5.3 Performance metrics calculated based on the simulation and analytical model in 1D network. ................................. 111
5.4 The network has only one dimension with the length of \( l \) and the tag is placed at the beginning of the network which has a broadcasting range of \( r \), \( a \) is randomly chosen to place the tag at the location \( a + r \). .................................................. 112
5.5 (a) Node-tag angles. (b) Region, different cases are shown, the closest points would be selected to calculate the distance to the perimeter of the region. ................................. 114
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>The probability of hitting a tag placed at (3, 4)</td>
<td>115</td>
</tr>
<tr>
<td>5.7</td>
<td>Performance Evaluation based on the simulation and analytical model in a 2D Network</td>
<td>116</td>
</tr>
<tr>
<td>5.8</td>
<td>(a) The probability of hitting at least one tag by a mobile node.</td>
<td>118</td>
</tr>
<tr>
<td>5.9</td>
<td>Performance evaluation based on the simulation and analytical model using multiple tags placed in a 2D network.</td>
<td>120</td>
</tr>
<tr>
<td>6.1</td>
<td>Merge operation in the case of having a 2 slot relay buffer. $n$ is the order of messages in time, i.e., all messages tagged with $n$ in a slot were generated in the last round, all messages tagged with $n - 1$ in a slot were generated two rounds ago, etc.</td>
<td>131</td>
</tr>
<tr>
<td>6.2</td>
<td>Merge operation in the case of having a 3 slot relay buffer.</td>
<td>131</td>
</tr>
<tr>
<td>6.3</td>
<td>Message Delivery Rate VS. Relay ID.</td>
<td>136</td>
</tr>
<tr>
<td>6.4</td>
<td>Delivery performance of different solutions found by SA respecting to the different importance of hit probability and distance factors. Mobile nodes move according to RWP model.</td>
<td>138</td>
</tr>
<tr>
<td>6.5</td>
<td>Delivery performance of different solutions found by Greedy respecting to the different importance of hit probability and distance factors. Mobile nodes move according to RWP model.</td>
<td>138</td>
</tr>
<tr>
<td>A.1</td>
<td>(a) Node-relay angles. (b) Region, different cases are shown, the closest points would be selected to calculate the distance to the perimeter of the region.</td>
<td>160</td>
</tr>
<tr>
<td>B.1</td>
<td>Example genetic algorithm crossover and mutation.</td>
<td>165</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The design and evaluation of routing protocols for sparse and intermittently connected mobile ad hoc networks is a significant problem of importance in today’s dynamic communication environment. The lack of infrastructure and/or the lack of built-in long range communication systems in popular portable devices complicate data delivery in such networks. In this thesis, we show how the mobility of mobile devices can be used to store, carry, and forward data messages to their destination in sparse mobile networks. In particular, we augment these highly disconnected networks with a fixed, easily deployable infrastructure using stationary, passive relay nodes to accommodate the connectivity and increase the delivery performance of the network. Our work improves the ability to communicate in very extreme circumstances.

In this chapter, the context and problems are described, the motivation behind the thesis is discussed, the research objective and the methodology utilized in the thesis is provided, the thesis contribution is presented, and the structure of the thesis is summarized.

1.1 Background and Context

Recent demands for affordable, portable wireless communication and computation devices have resulted in an exponential growth of wireless networks. Zeng et al. [90] has briefly surveyed the evolution of wireless communications and how
the dream of ubiquitous communication between anyone, anywhere, at anytime became a reality. The major goal of wireless networks is achieving ubiquitous network connectivity without pre-existing wired infrastructure. Wireless networks are widely used for a variety of applications where it is difficult to wire the environment such as concrete buildings, trading floors, manufacturing facilities, warehouses, and historical buildings or where it may not feasible for the operational environment to accommodate a wired network, or where it is not economical/practical to build a wired network for temporary purposes, e.g., conference registration centers, campus classrooms, emergency relief centers, and tactical military environments.

One of the emerging classes of wireless networks is mobile ad hoc networks (MANETs), which became popular as a research topic since the mid- to late 1990s due to the growth of laptops and 802.11/Wi-Fi wireless networking. MANETs consist of mobile wireless devices, or nodes, that do not require fixed infrastructure for communication while in traditional wireless network (cellular networks), one or more central nodes which are called base stations or access points exists in the network allowing the communication between nodes to take place. Routing in traditional wireless networks is dependent on the base stations and it happens in two steps. In the first step, any source node forwards its messages to the central infrastructure through an access point, and the central infrastructure destines the messages in one hop to the destination node in the second step. On the contrary, MANETs use decentralized routing algorithms. In other words, routing decisions are made among the nodes in a multi-hop manner and each node can act as a router.

MANETs have many advantages over traditional cellular networks, which require the establishment of infrastructure. For example, MANETs can be setup on demand which makes them desirable to be deployed in places with no infrastructure such as military applications ([81, 97]). They are also good alternatives in the places where the communication infrastructure is damaged and a rapid deployment of a communication network is required such as disaster recovery, and search and rescue [45]. The other advantages of MANETs in comparison to cellular networks include their fault tolerance and unconstrained connectivity. A point of failure in MANETs does not necessarily lead to a break down of the whole
1.1 Background and Context

communication network because other mobile nodes can keep routing/forwarding messages. Moreover, as mobile nodes can join/leave the network on the fly, the scalability of the network can be higher and the connectivity constraints could be less in comparison to cellular networks.

MANETs can provide ubiquitous network connectivity for mobile devices outside the range of wireless base stations and cellular networks. This can occur, e.g., when users roam to areas that are difficult for wireless signals to reach (either because of distance or because of interference with surrounding objects), when the mobile devices have a power/design constraint that limits their transmission capabilities, or when the cost of providing wireless infrastructure is too high. In this case, mobile devices form network connections with each other and connections to Internet services can be made via the MANET, where at least one mobile device has a connection to the wireless infrastructure. Figure 1.1 shows how MANETs provide network connectivity for mobile devices outside the range of wireless base stations and cellular networks. It also shows different types of MANETs, dense and sparse.

Figure 1.1: Example of dense and sparse MANET nodes as a result of distance from a base station. Sparseness may also arise when wireless signals are obstructed, when signal power constraints are applied and when node mobility is restrained.
1.1 Background and Context

Devices in a MANET are able to move independently in any direction which leads to having a dynamic network topology. Likewise, they must act as a router to forward traffic unrelated to their own use. The fundamental challenge in building MANETs, therefore, is equipping each device with an effective routing protocol which allows them to continuously maintain the information required to properly route/forward traffic. The nodes’ mobility, e.g., their velocity, grouping and spatial constraints, is a central factor affecting the ability to route messages, that we consider in this thesis. Other factors include power usage, failure, and interference.

There are more extreme scenarios in which there might be no end-to-end route between each pair of source and destination nodes due to the small coverage area of the network, high node mobility, or power constraints. This kind of scenarios makes a new class of MANETs called delay/disruption tolerant networks (DTNs). DTNs arise naturally from applications such as wildlife tracking, vehicle-based disruption-tolerant networks, rural kiosks in developing countries, and underwater exploration and monitoring, or from fragility and failures in the network itself due to wireless radio range limitations, high node mobility, disasters/attacks, jamming and noise, and power failure. DARPA has recently funded many DTN projects because of the demand of creating better services/applications in DTNs.

DTNs are also referred to as sparse MANETs, extreme wireless networks, or intermittently connected networks in the literature. DTNs use decentralized routing algorithms, i.e., routing decisions are made among the nodes and each node acts as a router. If two nodes are within the broadcast range of each other and the link between them is up then we say they are connected. If a source node and a destination node are connected then the source can send a message directly to the destination, i.e., in one hop, otherwise communication can take place via intermediate nodes, i.e., using multiple routing hops. In the literature, the connected link is referred to as a contact. Traditional routing protocols proposed for MANETs are ineffective in DTNs. The reason for the poor performance of the early MANET routing protocols in DTNs is due to the assumption that the underlying network is connected. A

connected network here means that there exists at least one (possibly multi-hop) path between each pair of nodes and that that path exists for a long-enough period of time to allow a packet to traverse it. Furthermore, these protocols assume that if the path is disrupted it can be repaired or replaced in a relatively short time. The assumption of connectivity is clearly ineffective in DTNs because of the lack of instantaneous end-to-end paths in such networks which prevents establishing any routes to forward the data packets.

In order to overcome the lack of instantaneous end-to-end paths in DTNs, a routing protocol can use a store and forward paradigm. Thereby, a new class of routing protocols, referred as store-carry-and-forward [12; 34; 36], has emerged. This class of routing protocols exploits the mobility of the nodes in the network to forward the data packets by relaying packets to intermediate nodes. The intermediate nodes then keep the data and send it later to the final destination or to another intermediate node.

The currently implemented applications which use the store-carry-and-forward routing protocols include Princeton ZebraNet, MIT CarTel, Cambridge Haggle, UW Waterloo KioskNet, UMass DieselNet, MSR VanLAN, and NASA Interplanetary Internet. As an example, ZebraNet [36] system uses some tracking nodes, which are referred as collars, attached to animals and the goal is studying their movement patterns across a large, wild area. The attached nodes are responsible to deliver the logged data back to researchers operating as a peer-to-peer network. These nodes contain a GPS to locate the animals, memory to store the data, wireless transceivers, and a small CPU. As the base-stations at the researchers’ side are mounted on cars which move around at irregular intervals there is no fixed continuous coverage. Therefore, when animals meet each other, the attached nodes exchange their data with the hope of increasing the chance to deliver the data messages to the researchers’ cars. Further study is continued in Section 2.

Researchers showed that utilizing mobility in data delivery can significantly increase delivery performance in disconnected networks. For example, Grossglauser et al. [25] showed that using a store-carry-forward protocol, a source node can reach a constant per-node throughput capacity. Further, there are numerous studies, reported in Section 2, to design routing protocols which rely on
node mobility to forward the data packets. The performance of these protocols is sensitive to contact opportunities among the intermediate nodes. The important factor that separates these protocols from each other is having a higher delivery performance and using fewer resources in terms of power and storage capacities as well as having a lower overhead. Therefore, these studies are targeting trade-offs by either maximizing the data delivery rate or minimizing the delivery delay and resource consumption.

There are other approaches that reinforce connectivity on demand in DTNs by utilizing additional communication resources in the network. Examples of these resources include satellites, base-stations, unmanned aerial vehicles, etc. Satellite communication needs a direct view of receiver/sender to the satellite and is expensive as it requires all user stations such as handsets, portables, mobile stations, etc. to be equipped with specific operational units to allow the communication to take place. Further, satellite availability might be poor in noisy environments, e.g., when the operator is very close to large machineries. Blind spots is a well-known problem when using base-stations for communication due to the possible obstacles in the network. Additionally, deploying base-stations might be difficult in areas which are difficult to reach and which need time/resources to take place. In addition, base-stations usually need to have a wired connection to the backhaul.

Missed contact opportunities is a common issue in DTNs which can dramatically decrease throughput and increase delay in the network. A very new approach to compete with the traditional DTN routing protocols is augmenting the network with some low cost, easily deployable fixed relay nodes in order to increase the number of contact opportunities. We refer to this emerging class of protocols as stationary relay point protocols. In these approaches some stationary relay nodes are added to the network in order to improve connectivity. Moreover, this class of DTN routing protocols are suitable in cases where we cannot expect the mobile nodes to change their mobility pattern.

We categorize stationary relay point approaches as active and passive based on the type of relay nodes. If the relay node can initiate the communication

we refer to the approach as an active relay point; otherwise, we refer to it as a passive one. Further, in active approaches [6, 33, 96], stationary relay nodes have their own supply of power while in passive approaches as we propose in this thesis, they are mainly powered by the readers, i.e., mobile nodes. Also, in active approaches, the number of relays is less than the passive approaches while their broadcasting range is usually bigger.

1.2 Motivation and Problem Statement

DTNs suffer from poor delivery performance because of the lack of end-to-end routes between each pair of source and destination nodes. Existing routing protocols in the literature try to increase the delivery performance, decrease the usage of resources in terms of power and storage capacities, or reduce the overhead mainly for specific scenarios based on the traffic model or mobility pattern of the nodes. These protocols focus on existing mobile nodes in the network acting as mobile relays, i.e., intermediate nodes, which are both mobile and active. To the best of our knowledge nobody has studied the case of passive stationary relay nodes which is desirable for radio silent applications as relays do not initiate communication. The main problem that this thesis addresses is evaluating the potential improvement deriving from augmenting the DTNs with passive stationary relays. Passive relays can be easily deployed in an ad hoc manner into the network without using any attached source of power, e.g., batteries, for long-term usage as they are powered by the readers. Placing the passive relays can be done easily using aerial vehicles for difficult-to-reach areas, e.g., a disaster, or could be placed around a road for sparse vehicular networks. Additionally, they can be deployed in a large number to increase the contact opportunities among the nodes as they can be produced with a very low cost. Passive RFID tags that are commercially available is an example of a type of passive relay.

Consider a hypothetical scenario shown in Figure 1.2 in which miners need to communicate to each other in order to improve mining productivity, through the use of a DTN. Satellite communication fails in underground areas as there is no direct line of sight. Using base-stations in underground areas like mines/tunnels
is not practical due to short communication range, deployability, and cost effectiveness since UHF/VHF technology has a very short communication range in underground areas \cite{92} and there could be blind spots due to the lack of line of sight transmission path of a wireless signal. The line of sight is violated as there might be obstacles in underground areas (e.g., curvy tunnels or branched mines). Moreover, since mines might be expanding, e.g., in gold mines to explore new sources of gold, setting up new base-stations is time-consuming. For the above mentioned reasons and in order to arrive at a competitive protocol, we use passive RFID tags as relay nodes. Miners can communicate to each other through RFID tags. Further, they can use mobile vehicles as data mules to communicate with the base-station. Data mules would pick up their messages from the tags and also directly from the miners and later they deliver these messages to the base-station.

![Diagram of mining scenario as an application of our protocol.](image)

Using stationary relays in DTNs improves the performance of forwarding protocols and consequently the related applications. As an example, Banerjee \textit{et al.} \cite{5} show experimental results collected from the UmassDieselNet that adding a fixed relay to the network improves the packet delivery by 37\% and reduces the message delivery latency by more than 10\%. We show in this thesis that passive relays which have a lower radio range significantly improve the delivery performance of the DTNs.

W. Zhao \textit{et al.} \cite{95} proposed a scenario/application which further motivates our approach. They consider the scenario where a severe earthquake has taken
1.3 Research Objective and Methodology

place and consequently most of the facilities and infrastructure are destroyed. In such a disaster, an urgent solution to rescue people trapped in debris is vital. The rescue team and the victims cannot make a connected network because available devices such as cell phones or PDAs have a limited communication range and also they may not have access to any infrastructure. Unmanned aerial vehicles (UAVs) are used to share the information about the number and location of survivors and potential hazards, acting as ferries. They also help the rescue team coordinate. Distributed over the region by an aerial vehicle, a passive unconnected RFID infrastructure facilitates the communication among the mobile nodes and increases the performance of the network in terms of delivery as a result. Additionally, our protocol does not force any mobile nodes to modify their trajectories.

Therefore, in this thesis, in order to have better connectivity in DTNs, we consider design and analysis of a passive stationary relay point based protocol to deliver the messages. Apart from the hypothetical scenario introduced above there are more applications which can utilize our protocol. As an example, our protocol can be applied to DTN radio-silent applications which require detection prevention. In other words, in such applications, the relays should not initiate the communication and have short radio signaling range. As another interesting application, we can use our protocol in tracking stolen cars in disconnected areas by placing the relays over/near the roads and recording the passing cars' ID, event time, etc.

1.3 Research Objective and Methodology

The objective of this research is to propose, analyze, and evaluate mechanisms for achieving robustness for DTN routing schemes while also respecting performance and scalability. The research outcomes include detailed design and analysis of a routing system which is created specifically for DTNs that can effectively deliver messages across the network. The protocol as we show is also customizable to handle user preferences, i.e., the trade-off between having a more delivery rate and lower message delivery latency.
The key characteristic of our proposed protocol, called Tag-base Routing (TBR), is augmenting DTNs with an unconnected, fixed, and easily deployable infrastructure using stationary, passive relay nodes to increase the connectivity among the mobile nodes. An increased connectivity in the network results in an increased delivery performance of the network. In most cases we use $N_{relay}$ stationary relay nodes (RNs) over a region defined by its extents, $(X_{min}, X_{max})$ and $(Y_{min}, Y_{max})$ where $N_{node}$ mobile nodes (MNs) will move. Generally, mobile/relay nodes have limited buffer sizes which are merged when an MN encounters an RN. Therefore, MNs can communicate to each other through RNs resulting an increased connectivity. In fact, two MNs may never meet each other in a typical DTN but they can indirectly communicate via RNs. More details are provided in Section 3.1.

For ease in referring common terms used in different chapters of this thesis, Table 1.1 summarizes the notation. Units are always meters for distance, seconds for time, speed in meters/second, unless otherwise noted. Notation is further explained in the text as required.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{relay}$</td>
<td>number of relays</td>
</tr>
<tr>
<td>$N_{node}$</td>
<td>number of nodes</td>
</tr>
<tr>
<td>$X/Y_{min/max}$</td>
<td>extents of the geographic region (meters)</td>
</tr>
<tr>
<td>$d$</td>
<td>regular spaced distance between relays (meters)</td>
</tr>
<tr>
<td>$B_{relay}$</td>
<td>relay message buffer size (messages)</td>
</tr>
<tr>
<td>$B_{node}$</td>
<td>node message buffer size (messages)</td>
</tr>
<tr>
<td>$r$</td>
<td>relay range (meters)</td>
</tr>
<tr>
<td>$S_{min/max}$</td>
<td>extents of the mobile node speed (meters/sec.)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>mean arrival rate of messages over all nodes (msgs/sec.)</td>
</tr>
<tr>
<td>$t$</td>
<td>time (sec.)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>message delivery rate (msgs/sec.)</td>
</tr>
<tr>
<td>$L_{avg}$</td>
<td>average message delay (sec.)</td>
</tr>
<tr>
<td>$E$</td>
<td>$\rho/L_{avg}$, which is used as an objective function</td>
</tr>
</tbody>
</table>

To evaluate the proposed protocol, simulation and analytical studies are performed. Simulation study allows repeatable and fully controllable experiments to show the effectiveness of our protocol in different scenarios including different
placement techniques, traffic models, mobility models, number of nodes/relays, communication range, buffer sizes, buffer policies, and queuing strategies. Further, it gives the chance to compare our proposed protocol with existing protocols in the literature. The analytical model reduces the simulation burden; however, the need of simulation is still demanding as it is difficult to have a thorough analytical model for the whole system. We use different metrics to evaluate the performance of TBR. The main evaluation metrics are defined as follows:

**Message Delivery Rate**  The message delivery rate is defined as the ratio of the number of successfully delivered messages to total number of messages, i.e., identically generated messages by all source nodes:

\[
\rho = \frac{M_{\text{delivered}}}{M_{\text{total}}}. 
\]

A low \( \rho \) indicates that the buffer sizes are not large enough to handle the rate of messages in comparison to average delay experienced by a message to get from the source node to the destination node.

**Message Delay**  Since a typical destination node may receive multiple copies of a message, we define message delay, \( L_i \), for message \( i \), as the time between when a message is generated to the first time the message is received by the destination. The average message delay is then:

\[
L_{\text{avg}} = \frac{1}{M_{\text{delivered}}} \sum_{i=1}^{M_{\text{delivered}}} L_i. 
\]

Message latency is computed only over those messages that were successfully delivered. We account for other messages in terms of the Message Delivery Rate above. More metrics are further described in the text as required.

Additionally, we study how to improve the performance of our protocol using a variety relay placement optimization techniques and also enhanced buffer management policies. We also enhance the proposed protocol using analytically
enhanced relay placement techniques. Further, we facilitate the performance calculation of the network by proposing analytical equations.

1.3.1 MATLAB Simulation Environment

We implemented an event driven simulation of TBR in MATLAB®. We designed and developed a modular application which includes following main modules:

- **Relay Model** - This module defines the desired placement technique. Different options are chosen from our proposed placement techniques introduced in Section 4.2.

- **Genetic based Optimization Model** - We developed a genetic algorithm based optimization introduced in Appendix B which is used to optimize the relay model. In other words, this module is designed to let the system to learn about the spatial characteristics of the relays in order to optimize relay placement.

- **Message Model** - This module defines the desired message model. It has various options to specify the source nodes and sinks. Source nodes as well sinks can be chosen randomly or according to a predefined model. Messages are stored in an array sorted based on their arrival time. Each message has a data structure that can be used for analysis purposes.

- **Path Model** - Path model generates all the paths that a mobile node will traverse during the simulation given a time interval and the mobility model. Later, we use this model to find the interaction times between mobile nodes and stationary relay points.

- **Main Engine** - Main engine controls the end-to-end simulation. Given the relay, message, and path models, it would simulate the interactions between MNs and RNs. It starts from time 0 and ends at a given time t. Firstly, it calculates all the meeting times between MNs and RNs including the connect/disconnect time in a sorted order based on the events time. Afterwards, it starts processing all the events one by one from the first event
1.4 Contribution of the Thesis

This thesis provides the following contributions.

until the last one. Accordingly, it performs the proper actions related to each events. Before processing an event it obviously checks the next message arrival time. If there are any messages before the event happens it adds those messages to the related MNs’ buffer.

• **Merge Module** - This module merges MN’s buffer and RN’s buffer conceptually in a single buffer and then calls Buffer module to decide which messages should be kept in each buffers. The merge operation is described in Section 3.1.

• **Buffer Model** - Buffer model specifies the buffer management policies introduced in Section 4.4. This module would be called either when a MN generates a new message, i.e., in run-time when the Engine calls this module to allow a MN to add new messages, or when a merge between MN’s buffer and RN’s buffer happens.

• **Analyze Model** - This module defines the performance metrics used in this thesis and calculates the value of each metric given the simulation results.

• **Analytical Model** - We developed different modules for our analytical models. They include calculating the hit probability of a relay by MNs, introduced in Appendix A and different numerical simulation modules to verify the models.

• **Simulated Annealing Model** - This model is developed to minimizes an objective function with the method of simulated annealing. This module is independent of any other modules and it uses the Analytical model module to calculate the value of the given objective function in each iteration.

Figure 1.3 shows the structure of the main modules we used to evaluate TBR.
1.4 Contribution of the Thesis

- **Design and analysis of a routing protocol for DTNs** - As stated earlier we design and analyze a new routing protocol for DTNs called TBR in order to overcome existing issues in DTNs. Further, we evaluate the performance of the proposed protocol in terms of packet delivery rate and message latency, and the overhead of it in terms of number of transmissions in the network. Our simulation results show that TBR achieves competitive message latency and delivery rates in comparison to existing routing protocols in the literature.

- **Effect of mobility models on the proposed protocol performance** - Mobility models have a significant impact on the performance of routing protocols in DTNs. We analyze the performance of TBR for a number of well-known mobility models widely used in the literature. The mobility models are including both Entity Mobility Models and Group Mobility Models. We also introduce a new mobility model named restricted random waypoint in order to further study the effect of special mobility models on the performance of TBR. Our analysis shows that the performance of our protocol is highly dependent on the mobility models as they can increase/decrease the number...
of contact opportunities between nodes/relays.

- **Analytical modeling of the proposed protocol** - We propose an analytical model for the proposed protocol accompanied by closed-form expressions to calculate the delivery delay of data messages as well as throughput. This analytical model can be used by network designers to consider the effect of different strategies in deploying the network without the need of implementing a priori-prototype. Accordingly, the proposed analytical model facilitates meeting application requirements by avoiding the unnecessary waste of resources and time in measuring the cost dictated by simulation.

- **Relay placement techniques** - The location of the relay nodes plays an important role in the performance of the stationary relay point protocols. Therefore, we demonstrate several techniques for optimizing relay node placement, namely relay pruning, probability based relay distribution and a genetic algorithm which are based on simulation. A comparison between the techniques has been done through a comprehensive simulation and also with other existing placement techniques in the literature. We show how we can significantly improve the performance of our proposed protocol using a more effective placement technique.

- **Analytical modeling of relay placement techniques** - We propose an analytical model to study the effect of relay placement techniques on the delivery performance of the proposed protocol. We are aiming to figure out how we can configure the best relay nodes placement analytically in order to optimize the performance of the protocol and reduce the burden of wasting computational resources and time dictated by simulation-based approaches. Further, we study the way of using our analytical model to propose better strategies in placing relay nodes. Two different techniques are proposed in this thesis namely simulated annealing based optimization and greedy placement which are based on the proposed analytical model. These techniques use the proposed analytical model to optimize the relay placement solution. Also, we compare the effectiveness of our analytical-based placement techniques with simulation-based approaches. The analytical-based approaches
overcome the simulation-based ones both in terms of resource usage and effectiveness. Analytical-based approaches achieve a better performance than simulation-based approaches while they minimize the required amount of computational resources and time.

- **Buffer policies and queuing strategies** - We study another interesting problem to further improve the performance of our protocol by using more intelligent buffer policies and queuing strategies. Messages can be differentiate based on the probability that they could reach the destination and since relays/mobile nodes’ buffer is limited in terms of capacity we can chose the best messages in hope that they will be delivered to the destinations in the future with a higher chance.

## 1.5 Structure of the Thesis

The chapters of this thesis are organized as follows. Chapter 2 discusses DTNs and routing as an important issue in this kind of network. It also reviews the current state-of-the-art routing approaches for DTNs with a classification of the techniques used in this domain. In Chapter 3, we present a detailed description of the proposed protocol. A comprehensive simulation study is presented in this chapter to show the effect of different parameters such as node velocity, communication range of stationary relays, mobile node/relay buffer size, mobile node/relay number, etc on the performance of the proposed protocol. This chapter ends with a comparison of the proposed protocol with some of the state-of-the-art protocols in the literature. In Chapter 4, we first study the possible enhancement on our proposed protocol using relay placement techniques. We then study the optimization of buffer policy management techniques which further improve TBR delivery performance. In order to evaluate the effect of optimization techniques we used in this chapter, different node mobility models are proposed and used. Chapter 5 presents an analytical model in terms of network- and application-dependent parameters for our protocol. We show how to overcome the simulation overhead in terms of time and resources complexity to find the best configuration for different
applications. Chapter 6 shows how we analytically improve the performance of the proposed protocol using enhanced relay placement techniques. This chapter also presents competitive solutions to the simulation based approaches. Additionally, analytical equations are provided to facilitate the performance measurement of the network. Finally, Chapter 7 concludes the thesis and proposes future research.

1.6 List of Publications

The following publications were generated during the research contributed toward this thesis:

1.6.1 Published Papers


1.6.2 Papers under Review

Chapter 2

Delay/Disruption Tolerant Networks and Routing Paradigms

In this chapter we first review the important issues of Delay/Disruption Tolerant Networks (DTNs) and then we give our definition for DTNs based on the connectivity of the network. Additionally, we present the existing applications in this domain, which motivates the research on DTNs. Furthermore, routing in DTNs, as one the most important challenges in this domain is discussed and a taxonomy of routing approaches for DTNs is provided. Finally we review and classify the current state-of-the-art routing approaches in DTNs and then we discuss the scope of our work in this classification.

2.1 DTNs Research Background

Delay/Disruption Tolerant Networks (DTNs) are a class of wireless networks that suffer from lack of continuous connectivity. DTNs have been used in a variety of applications such as wildlife tracking in biology, civilian applications like vehicle-based disruption-tolerant networks and rural kiosks in developing
countries\footnote{68}, and scientific applications such as underwater exploration and monitoring\footnote{58}. In a DTN, there is generally no end-to-end route between a source and destination node due to aspects such as the small coverage area of the network, high node mobility, or power constraints.

As there is no fixed infrastructure in DTNs to deliver the messages from source nodes to destination nodes, mobile nodes have to participate in routing and they have to act as a router similar to mobile ad hoc networks (MANETs). The number of connections present in a MANET at any one time is an important topological property because such connections allow communication to take place. We categorize MANETs as DTNs depending on the degree to which connections are available. In DTNs, the number of nodes per unit area, or the node density, is small and the nodes do not connect frequently. The network may remain partitioned into individual nodes for relatively long periods of time. DTNs are also referred to as sparse MANETs, extreme wireless networks, or intermittently connected networks in the literature. DTNs need to use decentralized routing algorithms, i.e., routing decisions are made among the nodes and each node acts as a router. If two nodes are within the broadcast range of each other and the link between them is up then we say they are connected. If a source node and a destination node are connected then the source can send a message directly to the destination, i.e., in one hop, otherwise communication can take place via intermediate nodes, i.e., using multiple routing hops. In the literature, the connected link is referred to as a contact.

\subsection*{2.1.1 Our Definition of DTNs}

There are a few definitions of DTNs in the literature. For example, Perur \textit{et al.}\footnote{62} define sparseness of a MANET in terms of connectivity of the network, i.e., the probability that the network graph forms a single connected component. If the probability of a single connected component is less than 0.95, then it is referred to as a sparse MANET. From another perspective, we categorize a mobile network as a DTN for a given time interval, if the average number of contacts is less than 5% of all potential contacts, i.e., contacts between all pairs of nodes, over the
2.1 DTNs Research Background

interval. At this threshold the performance of the traditional MANET routing protocols, in terms of packet delivery rate, decreases to a level when their routing is unworkable.

To establish this threshold, we conducted an experiment in which we investigated the failure point of representatives of traditional MANET routing protocols, i.e., AODV [61], DSDV [60], DSR [35], and ADIAN [69], based on the sparseness of the network. To do this, we used NS2. We set a fixed number of mobile nodes, i.e., 20, while increasing the area within which they could move, thus increasing sparseness. The data traffic used in the simulation is constant bit rate (CBR) with a rate of 8 kbps. Mobile nodes move according to the random waypoint (RWP) mobility model with a pause time of 0 seconds and maximum allowed speed of 3 m/s. The simulation time is 1000 sec and radio range of the nodes is 250 m. In the Random Gossip protocol, each node picks a connected node at random and forwards the packet. The maximum number of possible contacts between each pair of nodes can be calculated as follows:

\[
\text{Max Contact No.} = \frac{N_{\text{node}} (N_{\text{node}} - 1)}{2}.
\]

The results are shown in the Figure 2.1. Figure 2.1 shows the performance of latter well known MANET routing protocols compared to random gossiping in terms of the fraction of delivered packets to total generated packets versus average percentage of available contacts between each pair of nodes. According to Figure 2.1, the other protocols converge to the performance of the Random Gossip protocol when the total contacts are about 5% of all 190 possible contacts between each pair of nodes.

The reason for poor performance of the state-of-the-art MANET routing protocols [15; 19; 35; 51; 57; 59; 60; 61; 69] in DTNs is that they need to establish a/multiple fully connected path(s) between communication endpoints for communication to be possible. This assumption of connectivity is clearly ineffective in DTNs because of the lack of instantaneous end-to-end paths in such networks.

\(^{1}\)The details of the experiment are given here for convenience and further explained in later chapters of this thesis.
Figure 2.1: Packet delivery ratio vs. link fraction to all potential contacts between each pair of nodes.
which prevents establishing any routes to forward the data packets. This results in the situation that they are no more effective than random gossiping. The nodes’ mobility, e.g. their velocity, grouping and spatial constraints, is a central factor affecting the ability to route messages, that we consider in this thesis. Other factors such as power usage, failure, and interference are out of the scope of this thesis.

2.2 Applications of DTNs

In this section, we present important applications of DTNs in detail and describe the related works in each application. Examples of such networking applications include, but are not limited to:

2.2.1 Vehicle-based DTNs (VDTNs)

When the vehicles in a vehicular network move more quickly or they go to areas with low connectivity then a VDTN results. Soares et al. [76] refer to a VDTN as a new network architecture which is based on the concept of DTNs aiming to handle non-realtime applications at low cost, under unreliable conditions, enabling connectivity in diverse scenarios, using vehicles to carry data between terminal nodes. Burgess et al. [12] introduce UMass DieselNet composing of 30 buses which are equipped with WiFi nodes as a VDTN. UMass DieselNet was designed as a testbed for DTN research and was enhanced for further research consisting of 40 buses [91]. Buses follow regular routes and can establish a pair-wise connection upon encountering each other to exchange data. Passengers and passersby can access WiFi nodes inside the bus and are able to opportunistically connect to the Internet for several seconds via an open fixed 802.11 access point that is attached to the Internet, e.g., offered by a cafe or in the bus garage. Further, they can send their messages to the encountering buses with the hope that they could deliver their messages on behalf of them. Additionally, there is a GPS device in each bus in order to record times and locations. Balasubramanian et al. [4] study the case
of search the Web from a bus using the same testbed. Other example of VDTNs include CarTel \cite{32} and the Drive-thru Internet platform \cite{54}. CarTel is designed to collect and deliver data from mobile units such as automobiles to a central portal while Drive-thru Internet platform is designed in order to provide network and Internet connectivity to mobile users in vehicles. *Delay tolerant vehicular sensor networks* is a sub-class of VDTNs which is designed for monitoring the physical world of urban/rural areas. Li et al. \cite{43} propose a technique to enable transferring data when vehicles are connected only intermittently in a Delay Tolerant Vehicular Sensor Networks. They use a real-trace driven simulator called Shanghai Urban Vehicular Sensor Network (SUVSNet) in where 4000 taxis and 1000 buses constitute a virtual vehicular sensor network in Shanghai metropolitan area.

### 2.2.2 Delay Tolerant Networking for Sensor Networks

There are many applications for monitoring and surveillance tasks, which are using a sensor network, where providing reliable data communication across the network is vital. Ho et al. \cite{29} show how using DTN concepts in sensor networks can provide such a reliable data communication, that a regular sensor network is unable to make, as well as mitigate the possible communication interruptions in a sensor network. DTN concepts also are employed in different mobile sensor networks. SeNDT \footnote{http://down.dsg.cs.tcd.ie/sendt/} (Sensor Networking with Delay Tolerance) is a project funded by Trinity College Dublin designed for environmental monitoring, e.g., water quality monitoring and road-side noise monitoring. *Delay tolerant underwater sensor networks* is another class of sensor networks that use DTNs concepts in many applications. Partan et al. \cite{58} propose a taxonomy of different underwater networking regimes and review the practical issues in underwater networks including DTNs. Wang et al. \cite{84} also study the Delay/Fault-Tolerant Mobile Sensor Network (DFT-MSN) in order to gather pervasive information by developing some data delivery schemes tailored specifically for DFT-MSN. Wildlife tracking for biology research is another interesting application using DTN concepts in sensor
networks. For example, ZebraNet \[36\] system uses some tracking nodes, which are referred as collars, attached to animals and the goal is studying their movement patterns across a large, wild area. The attached nodes are responsible for delivering the logged data back to researchers, operating in a peer-to-peer network. These nodes contain a GPS to locate the animals, memory to store the data, wireless transceivers, and a small CPU. As the base-stations in researchers’ side are mounted on cars which move around at irregular intervals there is no fixed continuous coverage; therefore, when animals meet each other, the attached nodes exchange their data in hope to increase the chance to deliver the data messages to the researchers’ car. Further, Small et al. \[75\] utilize DTN concepts for studying whale behavior by attaching a sensor to each whale and using some mobile nodes called info-stations as data collectors.

### 2.2.3 Communication between rural zones in developing regions

Developing regions are environments experiencing intermittent connectivity that need to use DTN paradigms. Demmer et al. \[18\] study the design of a system based on DTNs to make the communication between rural zones in developing regions possible, where conventional strategies may fail. DakNet \[11\] is another project which is based on DTN architecture to provide asynchronous Internet access to remote rural residents. In their system motorcycles and buses are responsible for gathering the messages from the users, e.g., users’ email and web search queries, and deliver them to the open access points. KioskNet \[68\] is also a project aiming to transfer data from cities with access to the Internet for smaller villages in rural areas.

### 2.2.4 Deep space communication

The Deep space network has different sites in three locations around the world which provides continuous contact with a distant spacecraft from the stations on
the Earth as the Earth rotates by hand-offs to each other periodically. DARPA funded NASA, MITRE and others to develop a proposal for the Interplanetary Internet (IPN) which is a great motivation for working on DTNs. Different applications are required on such an IPN which need to communicate with counterparts across disparate networking environments. These environments are suffering from different sets of physical and operational constraints and also wide variations in transmission latency. Burleigh et al. [13] introduced a DTN architecture as a solution for the IPN that operates in multiple disparate environments and also simplify the development and deployment of such applications.

Other applications include disaster recovery after a severe natural disaster where the infrastructure is significantly damaged, radio silent applications where links in the networks should be short in range and be shut down frequently for security reasons, and military ad hoc networks encountering continuous attacks to the links.

DTNs have their own specific characterizations which make them different to other type of wireless networks. These characterizations include but are not limited to intermittent connectivity, power constraint network elements, and large message delay [21]. All the latter issues make the routing scheme the most important factor in these kind of networks. An effective routing scheme allows the communication in the network to take place more efficiently to achieve applications’ specifications. Therefore, in this thesis we address the routing issues in DTNs. In the next section we review the research on routing protocols for DTNs and classify existing works in this domain. Finally we show the place of our work in this taxonomy and its importance in the literature.

2.3 Routing in DTNs

A pragmatic approach to overcome partitions in DTNs is by using longer transmission ranges and thereby maintaining persistent network connectivity [67]. Increased sparseness of the network then leads to increased power requirements, which is a clear shortcoming of this approach. Furthermore, using long radio range in some applications may not be possible or desirable.
In section 2.1, we showed that traditional mobile ad hoc routing protocols fail to address the existing challenges in DTNs due to the lack of any instantaneous end-to-end routes between source and destination nodes. However, if we let the mobile nodes interact with each other over time, they might be able to make a delayed end-to-end route. New routing protocols \[12; 34; 36\] have emerged to address the existing issues in DTNs. They are mainly based on store-and-forward strategies and exploit the intermediate nodes across the network as relays. Intermediate nodes then store the forwarded messages unrelated to their usage until they have an opportunity to forward them to another node. Thus, the data is incrementally distributed throughout the network, i.e., in the intermediate nodes, leading to facilitate the data delivery process.

Although store-and-forward techniques can overcome the delivery issues in DTNs, they dictate a high latency in the network as messages need to be muled by intermediate nodes for long periods of time until a forwarding opportunity happens. This delay has a trade-off with the rate at which contact opportunities are created by mobile nodes and network resources such as buffer storage for data muling and power for participating in forwarding the messages. Therefore, various DTN routing protocols in the literature have been proposed to increase the delivery performance, decrease the usage of resources in terms of power and storage capacities, or reduce the overhead while respecting to reduce the delay. However, their performance and complexity depend on network regime including mobility models of the nodes, traffic model, number of participating nodes, amount of resources in the network, flexibility to control the movement of the nodes, performance objectives to achieve, etc.

This section surveys and classifies various research works that have considered routing schemes for DTNs. Al Hanbali \textit{et al.} \[27\] classify these protocols based on the degree of knowledge of the mobile nodes about their future contacts with other mobile nodes. Specifically, their classification depends on whether these contact opportunities are scheduled, controlled, predicted, or opportunistic. From another perspective, as in this thesis we augment the DTNs with cheap and easily deployable stationary relay nodes, we categorize DTN routing protocols based on the possibility of using fixed stationary relay points. We further categorize each category based on more details which are presented in the next section.
2.3 Routing in DTNs

2.3.1 Contact Opportunities

The most important factor in message delivery performance of DTN routing protocols is the number of contact opportunities in the network. Therefore, different routing protocols for DTNs have their own strategies to increase the contact opportunities. Some protocols control the contacts between the mobile nodes. Some of others predict the contacts, taking advantage of nodes mobility history. Determining the probability distribution of future contact times, mobile nodes can choose a proper next hop in order to improve the end-to-end message delivery probability. Other approaches neither control any contact nor use any historic information about the contact opportunities which makes their deployability quite simpler and more scalable. These approaches use opportunistic contacts and forward the messages in each opportunity.

2.3.2 Evaluation Metrics for DTN Routing Protocols

In order to evaluate a DTN routing protocol performance we have to consider special evaluation metrics based on the specific characteristics of DTNS. In this section, we describe main evaluation metrics which are used in the literature to evaluate the performance of the DTN routing protocols.

- *Delivery Rate* - The delivery rate is defined as the ratio of the number of successfully delivered messages to total number of messages.

- *Message Delay* - The average required time between when a message is generated to the (first) time the message is received by the destination.

- *Power Usage/Transmission Count* - Power usage include all the energy consumption in the system such as required processing energy, transmission energy, flash read/write energy, radio start-up and shut-down energy, etc. However, the main criteria in the literature is transmission energy. To simplify the calculation, transmission count are generally measured as the power usage and each transmission is considered to require the same units of energy.
2.3 Routing in DTNs

- **Overhead** - The overhead mainly associated by the control overhead of the routing protocol in the literature. The control overhead can be calculated based on the total bytes of control messages transmitted by the routing protocol over the total bytes transmitted.

- **Route Length/Forward Count** - Route length which alternatively referred as hop count is the number of routers traversed by a packet between its source and destination. Also, a similar concept referred to as forward count is used in the literature to show the average number of relays used for one delivered destination to receive a data item.

- **Average Working Time** - The average working time of a host denotes how much time it spends on its own work instead of message transmission.

- **Message Copies** - Message copies is the average number of identical message replicas generated in the network.

2.3.3 Classification of DTN Routing Protocols

Recently there has been focus on developing routing protocols for DTNs. In this section, we divide these protocols into two major categories: *Store-Carry-Forward* paradigms and *Stationary Relay Points* approaches based on the possibility of using fixed stationary relay points. We further categorize store-carry-forward paradigms based on the possibility of controlling some/all mobile nodes’ mobility pattern. Additionally, we categorize stationary relay points approaches based on the characteristics of participating relay nodes.

2.3.3.1 Store-Carry-Forward

The class of Store-Carry-Forward paradigms exploits the mobility of nodes to buffer data packets during network partitions and forward them when connections become available. They are divided into two categories: *reactive* and *proactive* schemes. Reactive routing protocols ([77; 78; 83]) use mobility of the participating nodes to buffer and deliver messages across network partitions. While,
proactive routing protocols ([9; 24; 42; 86; 87; 95]) control the mobility of some nodes to accommodate disconnections. Both of these approaches exploit the mobility of nodes to buffer data packets during network partitions and forward them when the network is connected again.

**Reactive Schemes** Reactive approaches use mobility of the participating nodes to buffer and deliver messages across network partitions without forcing any mobile node to change its trajectory. This class of DTN routing protocols are differentiated based on the possibility of using flooding based approaches, historic information about the contacts, and social network based structure.

- **Flooding based Approaches** - Flooding based routing approaches are based on flooding message replicas over the network, i.e., replicating messages at forwarding opportunities, to increase the opportunity of delivering messages to the destination nodes. They are further categorized based on unlimited or bounded message replication. Vahdat *et al.* [83] propose an unlimited message replica routing protocol for partially-connected ad hoc networks called Epidemic Routing. In this protocol, every node exchanges all the data packets stored in its buffer while encountering with other mobile nodes until meeting the destination(s). A Bloom filter is used to reduce the space overhead while nodes are exchanging their data packets. Unlimited message replication can cause the protocols to waste valuable network resources leading to a poor scalability. Spyropoulos *et al.* [77] tried to improve the performance of unlimited flooding-based routing schemes such as Epidemic Routing by bounding the number of data packet replicas. In their protocol, source nodes are allowed to spray a limited number of identical messages to the network using distinct relays and wait until one of those relays meets the destination to perform a direct transmission. They further enhanced their approach by proposing spray and focus [78] where instead of direct transmission in a wait phase, relays can forward their copies to a newly encountered node based on an employed utility function, i.e., whenever an encountered node’s utility function value is higher than theirs a
2.3 Routing in DTNs

copy of the messages would be forwarded. Also, Tournoux et al. [82] propose a measurement-oriented variant of the spray-and-wait algorithm called DA-SW (density-aware spray-and-wait) that can tune the number of message replicas in a dynamic manner. Gao et al. [23] improve the Epidemic Routing scheme by reducing the data forwarding cost. They employ a multi-cast scheme to select the relay nodes considering the forwarding probabilities to multiple destinations simultaneously.

• History based Approaches - Some reactive approaches use historic information about node contacts, spatial information, etc., to calculate the delivery expectation of subsequent hops, indicating their probability of being able to successfully deliver a data packet. These approaches are including PROPHET [44], NECTAR [17], pattern-based MobySpace Routing [38], location-based Routing [40], and context-based Forwarding [52]. For example, Lindgren et al. [44] use historic information in terms of delivery expectation about next hops indicating their probability of being able to successfully deliver a data packet on their previous behavior to select the message carriers. Burgess et al. [12] improve the performance of the network by deleting useless data packets and scheduling packets for transmission to other peers. Boldrini et al. [10] propose a context-based protocol (HiBOp) which reduces resource consumption while increasing the performance of the PROPHET in terms of message loss rate and message delay. Balasubramanian et al. [3] study routing in DTNs as a resource allocation problem and propose a protocol to enhance the performance of a specific routing metric by using some heuristics. Yuan et al. [89] propose the Predict and Relay scheme which predicts the future contacts of specified nodes at a specified time. Then, source/intermediate nodes select a proper neighbor as the next hop to forward their messages using their estimations about the future contacts of their neighbors and the destinations. More recently, Lee et al. [39] propose a routing algorithm, called Distributed Max- Contribution (DMC) in which routing, scheduling, and message replication decisions are based on contemporarily available knowledge in the network.
2.3 Routing in DTNs

- Social Network based approaches - Other approaches try to make a network structure similar to social networks in order to find the nodes as message carriers with the highest centrality, i.e., the structural importance of the node, which typically have a stronger capability of connecting other network members together. The representatives of these approaches are SimBet [16] and BubbleRap [31]. Hossmann et al. [23] further evaluate SimBet and BubbleRap performance under real mobility traces.

**Proactive Schemes** Proactive approaches control the mobility of some nodes to accommodate disconnections. They are further classified as ferry-based and non-ferry-based whether they use a set of special nodes, called ferries, that are responsible for carrying data for all nodes in the network while non-ferry approaches do not use any ferries; however, they control the movement pattern of mobile nodes. Goldenberg et al. [24] use mobility control to reach optimality by moving relay nodes to their optimal positions. They propose two approaches: synchronous scheme and asynchronous scheme. In synchronous scheme, using a uniform distributed algorithm they allow the relay nodes to move to their optimal positions. This scheme relies on a globally synchronous operation mode. In the second scheme, they remedied this violation of the localized design requirement using an asynchronous algorithm. Li and Rus [42] propose the possibility of changing host trajectories to send messages in a disconnected ad hoc wireless networks. Using motion information of the destination node, they utilize cooperation of the intermediate nodes to deliver messages by asking them to modify their trajectory dynamically.

Zhao et al. [95] proposed a ferry-based approach called Message Ferrying (MF). Ferries act as a moving communication infrastructure for the network. They present two approaches called NIME and FIMF. NIME (Node-Initiated MF) uses a mode transition function for nodes including Go-to-work, Working, Go-to-ferry, and Send/Receive. Each node in the system acts based on a given algorithm to go to one of the mentioned modes and do the assigned task. In Go-to-ferry mode, the node calculates a shortest path to meet the ferry and then moves toward the ferry node. In Send/Receive mode, it exchanges messages with a ferry node and finally in Go-to-work mode, it moves back to its prior location.
2.3 Routing in DTNs

The ferry node follows a given fixed algorithm. It moves according to a predefined route and broadcasts Hello messages periodically. They also present different algorithms for node and ferry operations called Ferry-Initiated Message Ferrying (FIMF).

Wu et al. [86] propose a logarithmic store-carry-forward scheme through a hierarchical structure of trajectory for ferries that controls the number of relays which ends up with reducing average delay which was very high in MF and they also utilize some new technical issues like on-demand ferry solicitation, dynamic trajectory planning of ferries, rendezvous point placement, and adaptive ferry migration and load balancing to enhance the network performance. Tariq et al. [9] extend the MF approach by forcing ferries to follow fixed routes to overcome the problem of arbitrary movements of mobile nodes which cause losing precise location of nodes at any given time; therefore, they simplified the design of complex routes where the ferry can contact the nodes with certainty. To solve this problem they propose a method called Optimized Way-points (OPWP) which produces ferry routes that are in the form of an ordered set of way-points and waiting times at each of these way-points carefully. Yang et al. [87] extend the MF approach by proposing a mechanism to replace ferry nodes. Since the network operation relies on the ferries to provide connectivity in the whole network, they can be a single point of failure. Also, to keep a balance among all mobile nodes (with limited resources) in the network it may be desirable to rotate this role with others after a fixed duration. A summary of some existing routing approaches for DTNs can be found in [93].

Summary Store-carry-and-forward paradigms need either a contact between the source and destination nodes or a contact between an intermediate node, which is storing, carrying, and forwarding source node’s message replicas to other intermediate nodes/destination node, and the destination node. Accordingly, increasing contact opportunities between mobile nodes results in increased message delivery performance. Though, state-of-the-art store-carry-forward routing protocols increase contact opportunities by either utilizing ferry nodes or altering mobility trajectories of intermediate nodes, for many DTNs, mobility model of the nodes cannot be predicted nor modified. Further, the optimal trajectories
that mobile nodes should follow to reach the optimum delivery performance is difficult to find. A new class of routing paradigms referred to as stationary relay point approaches has emerged to improve the contact opportunities in the network by augmenting the network with additional low cost stationary nodes to relay the messages. In the next section, we describe them and detail existing approaches in this class.

2.3.3.2 Stationary Relay Points

In Stationary Relay Points approaches such as \[6; 33; 96\], by adding some stationary relay points, the connectivity of the network is increased and the performance of the network is improved as a result. This class of intermittently connected MANET routing protocols works best in circumstances such as when we cannot expect a ferry to move in an inhospitable terrain, or through the obstacles due to a disaster, etc., where missed contact opportunities may decrease network performance. Based on relay point type, we divide the class of stationary relay points into two categories: active and passive.

Active Approaches  Active stationary relay point approaches associate with the attached resource of power to the stationary relay nodes which can initiate the communication. Zhao, Chen et al. \[96\] propose an active approach for Stationary Relay Points. Relay nodes are called “throwboxes” and they are considered to be powered, wireless base stations that cannot interact with each other. However, they assume that relay nodes can initiate communication. This approach increases the network connectivity based on a fixed unconnected infrastructure using throwboxes. Furthermore, in their work, a mixed integer programming solution is proposed for the placement of throwboxes, which discretizes space and is NP hard. Therefore, a greedy heuristic is used to place throwboxes. The throwbox work is further analyzed in \[33\]. Although they introduce some powerful platforms to meet the requirements of throwboxes in DTNs, the possibility of making such an infrastructure may not be feasible or desirable. For example, in underground/underwater environment, the case of using solar panels to recharge throwboxes’ batteries is ineffective and may result in a short-term usage.
2.3 Routing in DTNs

Banerjee et al. [6] use different types of infrastructure to enhance DTNs, namely untethered relays, Infostations, and a wireless mesh. In the case of Infostations and a mesh, mobile nodes propagate data packets via an intra-network proxy with storage, which can avoid contemporaneous connections through the base stations; however, there is a trade-off between enhancing the network and the cost of the infrastructure. Cost, point of failures, blind spots, etc. are main issues in using either base-stations or Infostations. Use of untethered relays is similar to [96].

**Passive Approaches** There are different DTNs where providing active relay points is challenging due to their cost, short-term operation, etc. Additionally, the nature of radio-silent applications for DTNs need the stationary relay points to be passive and/or having short communication range. For the above mentioned reasons and also the motivation described in 1.2 we propose passive relay point approaches in this thesis, where stationary relay points are unable to initiate any communications and they are powered by readers, i.e., mobile nodes. Passive stationary relay points can be added to the network in an ad hoc manner without using any attached source of power, e.g., batteries, for long-term usage. A large number of passive stationary relay nodes can be deployed in the network as they are produced with a very low cost to increase the contact opportunities among the nodes. However, their broadcasting range is usually shorter than active approaches. An alternative for passive relay points is RFID tags as they are commercially available with low cost.

**2.3.3.3 Heterogeneous Approaches**

There are other approaches in which routing decisions are dynamic based on the network topology. We refer to this class of DTN routing as heterogeneous approaches. One of the heterogeneous approaches is clustered networks where the network posses both dense and sparse regions of connectivity at a same time. Piorkowski et al. [63, 64] study modeling both the mobility and the formation of clusters in these kinds of networks, where the nodes have a dense connectivity inside the clusters while the connectivity between the clusters is sparse. Ryu et
al. [66] propose a routing strategy for clustered networks called two-level Back-Pressure with Source-Routing algorithm (BP+SR) in which routing and scheduling within clusters are separated from the communications that occur across clusters using a different time-scaling strategy. Other DTNs that use heterogeneous approaches can hold different types of devices, e.g., handhelds, vehicles, sensors, etc., which makes a heterogeneous DTN. In such a network, mobile nodes have diverse characteristics and mobility patterns which make them behave differently in the network. Spyropoulos et al. [79] study the problem of routing in such networks and propose a routing protocol which is capable to identify the mobile nodes with the highest utility for routing purposes. Li et al. [41] study the case of Socially Selfish Delay Tolerant Networks in where nodes are willing to forward messages only to the nodes that are tied socially with. They propose a routing protocol called Social Selfishness Aware Routing (SSAR) to address the routing issue in such networks.

Table 2.1 presents a summary of the discussed DTN routing protocols as well as their main characteristics. The first feature describes the classification of the protocol in DTNs including the main class, sub-class, and possible sub-sub-class as indicated in Figure 2.2. The second feature shows possible further information about the protocol. The third feature indicates the nodes’ degree of knowledge about their future contact opportunities which are identified as being opportunistic, predicted, or controlled. The fourth feature presents the key performance metrics optimized by each proposal. The last feature shows the applicability of the protocol in radio silent usage. As shown in the table, TBR is the only protocol that can be effectively used in radio silent applications.

Figure 2.2 shows the classification of DTN routing protocols.

2.3.4 Discussion

Existing works in the literature are based on contact opportunities between mobile nodes which can debilitate network performance due to the occurrence of missed contact opportunities. Stationary relay point approaches can improve the
### Table 2.1: DTN Routing Protocols.

<table>
<thead>
<tr>
<th>Routing Protocol</th>
<th>DTN Type</th>
<th>more Details</th>
<th>Contact Opport.</th>
<th>Metric to Optimize</th>
<th>Radio Silent</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER[1]</td>
<td>Homog./SCF/Reactive</td>
<td>FUR</td>
<td>Opportunistic</td>
<td>Delivery Rate/Delay</td>
<td>No</td>
</tr>
<tr>
<td>Spray &amp; Wait[2]</td>
<td>Homog./SCF/Reactive</td>
<td>FLR</td>
<td>Opportunistic</td>
<td>Delivery Rate/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>Spray &amp; Focus[3]</td>
<td>Homog./SCF/Reactive</td>
<td>FLR</td>
<td>Opportunistic</td>
<td>Delivery Rate/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>M-SDM[5]</td>
<td>Homog./SCF/Reactive</td>
<td>FUR</td>
<td>Opportunistic</td>
<td>Delivery Rate/Delay/Forward Count</td>
<td>No</td>
</tr>
<tr>
<td>PROPHET[6]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>NECTAR[7]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate</td>
<td>No</td>
</tr>
<tr>
<td>MobiSpace[8]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Control Overhead</td>
<td>No</td>
</tr>
<tr>
<td>LBR[9]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Delay/Route Length</td>
<td>No</td>
</tr>
<tr>
<td>CBF[10]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate</td>
<td>No</td>
</tr>
<tr>
<td>HiBOp[12]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Delay/Buffer Usage</td>
<td>No</td>
</tr>
<tr>
<td>RAPID[13]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Delay</td>
<td>No</td>
</tr>
<tr>
<td>Predict and Relay[14]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Delay/Route Length</td>
<td>No</td>
</tr>
<tr>
<td>DMC[15]</td>
<td>Homog./SCF/Reactive</td>
<td>History based</td>
<td>Predicted</td>
<td>Delivery Rate/Delay/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>SimBet[16]</td>
<td>Homog./SCF/Reactive</td>
<td>Social Net. based</td>
<td>Predicted</td>
<td>Delivery Rate/Delay/Forward Count</td>
<td>No</td>
</tr>
<tr>
<td>BubbleRap[17]</td>
<td>Homog./SCF/Reactive</td>
<td>Social Net. based</td>
<td>Predicted</td>
<td>Delivery Rate/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>Goldenberg et al.[18]</td>
<td>Homog./SCF/Proactive</td>
<td>non-ferry based</td>
<td>Controlled</td>
<td>Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>Li and Rus[19]</td>
<td>Homog./SCF/Proactive</td>
<td>non-ferry based</td>
<td>Controlled</td>
<td>Delay/Average Working Time</td>
<td>No</td>
</tr>
<tr>
<td>MF[20]</td>
<td>Homog./SCF/Proactive</td>
<td>Ferry based</td>
<td>Controlled</td>
<td>Delivery Rate/Delay/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>Logarithmic SCF[21]</td>
<td>Homog./SCF/Proactive</td>
<td>Ferry based</td>
<td>Controlled</td>
<td>Delay/Hop Counts</td>
<td>No</td>
</tr>
<tr>
<td>OPWF[22]</td>
<td>Homog./SCF/Proactive</td>
<td>Ferry based</td>
<td>Controlled</td>
<td>Delivery Rate/Delay</td>
<td>No</td>
</tr>
<tr>
<td>Ferry replacement[23]</td>
<td>Homog./SCF/Proactive</td>
<td>Ferry based</td>
<td>Controlled</td>
<td>Delivery Rate/Power Usage</td>
<td>No</td>
</tr>
<tr>
<td>Throwbox[24]</td>
<td>Homog./SRP/Active</td>
<td>-</td>
<td>Opportunistic</td>
<td>Delivery Rate/Delay</td>
<td>No</td>
</tr>
<tr>
<td>Relays/Base-stations[25]</td>
<td>Homog./SRP/Active</td>
<td>-</td>
<td>Opportunistic</td>
<td>Delivery Rate/Delay/Trans. Count</td>
<td>No</td>
</tr>
<tr>
<td>TBR[26]</td>
<td>Homog./SRP/Passive</td>
<td>-</td>
<td>Opportunistic</td>
<td>Delivery Rate/Delay/Trans. Count</td>
<td>No</td>
</tr>
<tr>
<td>UBR[29]</td>
<td>Heterog./Heterog. Nodes</td>
<td>-</td>
<td>Predicted</td>
<td>Delivery Rate/Delay/Trans. Count</td>
<td>No</td>
</tr>
<tr>
<td>SSAR[30]</td>
<td>Heterog./Socially Selfish</td>
<td>-</td>
<td></td>
<td>Delivery Rate/Trans. Count</td>
<td>No</td>
</tr>
</tbody>
</table>

1Store-Carry-Foreward
2Flooding-unlimited Replicas
3Flooding-limited Replicas
4Stationary Relay Points
5Clustered Networks

---

1[83], 2[77], 3[75], 4[82], 5[23], 6[44], 7[17], 8[38], 9location-based Routing [40], 10context-based Forwarding [52], 11[12], 12[40], 13[3], 14[54], 15[54], 16[54], 17[31], 18[24], 19[12], 20[92], 21logarithmic store-carry-forward [86], 22[9], 23[87], 24[96], 25[6], 26our proposed protocol, 27[63], 28[64], 29Utility-based Replication [73], 30[30], 31[41]
2.3 Routing in DTNs

Figure 2.2: Our classification of DTN routing protocols.
performance of the network by creating a greater number of contact opportunities in DTNs through augmenting the network with easily deployable relay points. However, radio silent applications need to have a short communication range as well as passive operations. Active stationary relay point approaches proposed in the literature such as the work done by Zhao, Chen et al. [96] cannot satisfy these kind of applications. Moreover, active stationary relays need their own attached resource of power which can increase their cost and degrade their effective usage time. We found that a passive relay point approach not only is desirable to be used in a radio silent application but also is effective in DTNs because they can be deployed into the network in a large number as they are produced with a very low cost. Passive relay points also can be used in DTNs easily for long term usage as they are powered by readers without the need of any attached resource of power. Passive RFID tags as an alternative for passive relay points are commercially available. In the next sections, we propose a novel routing protocol that uses our new idea of using passive stationary relay points. We then evaluate the performance of our protocol through simulations and analytical models and compare it with the existing approaches in the literature. Our evaluation studies show that our approach outperforms the state-of-the-art DTN routing protocols in terms of message delivery rate and message latency. We also improve the performance of our protocol by optimizing the relay placement techniques to place the stationary relay points in strategic locations as well as proposing efficient buffer management policies.
Chapter 3

Achieving Ubiquitous Network Connectivity using an RFID Tag-based Routing Protocol

In this chapter, we present our DTN routing protocol called Tag Based Routing (TBR) which handles data delivery in disconnected networks using easily deployable passive wireless devices acting as stationary relays. TBR is designed to increase the contact opportunities among network nodes in order to increase the data delivery rate and reduce the message delay. We evaluate our protocol under various scenarios and also compare it with the the state-of-the-art routing protocols. Our simulation results indicate that our approach can achieve competitive message delay and delivery rates. In this thesis, we focus on our protocol performance and the issues related to security, privacy, etc., are outside of the scope of the thesis. This chapter is derived from our published work [70] and submitted work [74] on improving performance of DTNs through passive relay points.
3.1 Tag Based Routing Protocol

In this section, we introduce our protocol and also the evaluation metrics. Since the protocol was initially considered for passive RFID tags, we refer to our protocol as a Tag-based Routing or TBR Protocol.

3.1.1 Protocol Description

We consider a static distribution of $N_{\text{relay}}$ stationary relay nodes (RNs) over the region where $N_{\text{node}}$ mobile nodes (MNs) will move. The region is defined by its extents, $(X_{\text{min}}, X_{\text{max}})$ and $(Y_{\text{min}}, Y_{\text{max}})$. Each RN contains a message buffer of size $B_{\text{relay}}$ messages (each message has unit size), and each MN contains a message buffer of size $B_{\text{node}}$ messages. MNs know the positions of all the RNs and they can interact with an RN if it is within distance $r$ from the RN. Later in this section we show the technical aspect of how a MN can interact with an RN.

Messages are only transmitted between MNs and RNs and vice versa, never between MNs and other MNs. MNs contact with other MNs infrequently (by definition of a DTN) and thereby gains made through their contact is insignificant with respect to that gained by the contact with RNs. In other words, RNs are contacted much more frequently than MNs because the number of RNs is much larger than the number of MNs. If we included MN-to-MN communication then the overall performance of our system would never be worse but the power requirements would increase as MNs need to find each others. Finding RNs is not required, or at the very least is an off-line consideration, because the RNs in our proposal are placed according to a placement strategy.

Communication is only initiated by MNs since the RNs are passive. To consider MN-RN interactions we define relay connection and relay disconnection events in time. When a MN moves from a distance $> r$ to within a distance $r$ of any RN at time $t_1$ then we say the MN has connected with the RN at time $t_1$. When the MN moves, say at time $t_2$, outside distance $r$ of an RN that it is connected to, then we say the MN has disconnected from the RN at time $t_2$. In this way, given the paths for all MNs, we consider the set of distinct connection and disconnection events in time. To be complete, we also consider message arrival as
3.1 Tag Based Routing Protocol

an event in time. MNs move obliviously to each other and to any other aspects of the system that makes the implementation of our protocol simpler and more scalable as no MN needs to know extra information about other MNs.

Our protocol is now defined by what happens at each of the three events, also depicted in Figure 3.1(a):

**Relay Connection:** the connecting MN examines the RN’s buffer to see if any of the messages are destined to itself, if so then the message is said to have reached its destination at the time of the connection. The MN then merges its buffer with the RN’s buffer according to Algorithm 1.

**Relay Disconnection:** the disconnecting MN examines its state to see if it is still in the merge operation with the RN, if so then it terminates the merge operation.

**Message Arrival:** the MN puts the newly arrived message into its buffer, the oldest message, i.e., the one with the earliest arrival time, is discarded if the buffer is already full. Merging does not happen at this event (even though the MN may be within range of an RN).

Figure 3.1(b) provides a brief example of the movement of two MNs through a regular spaced field of RNs. MN A deposits a message on RNs 2, 3 and 1. Assuming MN B passes through RN 1 before MN A, but passes through RN 3 after MN A, then MN B retrieves the message from A at RN 3 and deposits the message in RN 4.

To describe MN-RN interaction we define a *merge operation* between a MN message buffer and an RN message buffer. Algorithm 1 shows the merge operation. According to Algorithm 1, the MN acquires some meta-data regarding the current messages in the RN’s buffer. Based on the messages in its buffer it decides to read/write messages one by one until it is disconnected from the RN or all messages are replicated. Therefore, the merge operation virtually combines both buffers into a single buffer with messages ordered according to their arrival times in the system. Then, the MN and RN both keep as many of the latest messages, i.e., the ones with larger arrival times, as can fit in their respective buffer. Older messages are discarded. In general, it is not possible for a MN to know at first face whether a message has been delivered to a destination or not (apart from the destination node itself), so delivered messages will continue to
3.1 Tag Based Routing Protocol

(a) Movement of a MN through the circular range of an RN.

(b) Example of two MNs, A and B, moving through a field of RNs spaced at regular intervals of distance $d$. In the example, MN A reaches RN 2 and 3, but not RN 1, before MN B. MN A deposits a message on these RNs, MN B retrieves the message from RN 3 and deposits it on RN 4.

Figure 3.1: Aspects of the Tag Based Routing model.
propagate until they are pushed out by newer messages. It is possible to increase the delivery performance by using more intelligent buffer policies, for example, Ma and Jamalipour [47] propose a fuzzy logic-based delivery framework called FLDF to facilitate the ranking of messages based on their delivery preference in the future to increase the message delivery rate.

**Algorithm 1** Merge Operation(MN,RN)

{MN and RN stand for mobile MN and relay MN respectively}

- \(S_1\) ← list of message IDs in MN’s buffer
- \(S_2\) ← list of message IDs in RN’s buffer
- \(S\) ← \([S_1, S_2]\)
- Sort\((S, \text{arrival time})\)
- \(counter\) ← 1

for all \(i\) ∈ \(S\) do
  if \(i\) \(\not\in\) \(S_1\) and \(\text{len}(S_1) > counter\) then
    Replicate \(msg\) \(i\) from \(RN\) to \(MN\)
    if MN’s buffer overflow then
      Evict oldest message from MN
  end if
  if \(i\) \(\not\in\) \(S_2\) and \(\text{len}(S_2) > counter\) then
    Replicate \(msg\) \(i\) from \(MN\) to \(RN\)
    if RN’s buffer overflow then
      Evict oldest message from RN
  end if
  if IsDisconnected\((MN, RN)\) then
    break;
  end if
  \(counter\) ← \(counter\) + 1
end for

Interactions are considered to be collision free; however, there could be two possible collisions. First, RNs could be placed in such a way that a MN falls into the communication range of them simultaneously. In order to avoid this kind of collision, MNs can label the RNs and send a request to communicate with the RN with the lowest ID first and then with the second lowest ID RN, and so on. Labeling the RNs is possible because the MNs know the positions of all
3.1 Tag Based Routing Protocol

the RNs. However, in most cases RNs are placed in such a way that this kind of collision is highly unlikely to be occurred. Second, two or more MNs may try to communicate with the same RN simultaneously. To handle this kind of collision we need collision detection and recovery/avoidance techniques in wireless networks \[85\] which is outside the scope of this thesis. However, since two MNs by the definition of a DTN, are highly unlikely to access an RN simultaneously, the impact of this kind of collision is negligible.

As no messages can come back to the same buffer after leaving a buffer as we are using a first-in-first-out (FIFO) buffer management policy, the protocol can avoid routing loops.

3.1.2 Routing Protocol Design Principles in TBR

In designing TBR we have considered important design principles mentioned in \[88\] to achieve a scalability which is quite demanding in DTNs:

• **Compartmentalization**: A very common issue in designing a scalable routing protocol is localizing problems or failures, which is called network compartmentalization. In TBR, MNs and RNs can be added to the network on the fly; therefore, based on the demand we can increase the number of RNs, replace them, or increase their buffer size to accommodate a higher traffic load. From another point of view, if an RN fails to store/forward messages to encountering MNs, which may occur due to the hardware/software failure, signaling problems, power constraints, etc., then the delivery could still be continued by other participating RNs. Therefore, the failures in a specific part of network will be prevented from spreading over the entire network.

• **Making Proper Trade-offs**: Interaction between MNs and RNs costs energy. Additionally, adding RNs in a typical network may cost money; however, in designing a routing protocol, making proper trade-offs is necessary according to the application. We have sacrificed cost in terms of money and energy to improve the performance of the network (i.e., in comparison
3.1 Tag Based Routing Protocol

to Store-Carry-Forward protocols) and to have a higher applicability. Further, in our protocol, the stationary relay points do not require their own power supply and this can lead to a greater network life time. In TBR MNs never communicate messages directly to each other, rather they communicate messages only via these stationary relay points which are in known places. This helps to save energy because the transmission sub-system can be powered off when MNs are not in the communication range of the RNs, which is a challenging issue in DTNs as the MNs are battery powered. Banerjee et al. [5] show that using 802.11 radio to search for contacts in a DTN devotes 99.5% of the total energy of MNs just to find other MNs, which results in a short network lifetime. Additionally, the possible communication between MNs is, by definition of a DTN, infrequent and therefore its effects are negligible. Although we augment the network with a passive unconnected infrastructure, MNs still can make a network on the fly without a priori connected infrastructure.

- **Reducing Burdens of Routing Information Processing:** To reduce routing processing burdens one should consider three issues including: 1) placing the routing intelligence in strategic spaces over the network; 2) reducing the routing complexity; 3) reducing the effects of route flapping, i.e., a router alternately indicates different routes to reach a destination node in quick sequence. We have considered these issues in TBR. As we will discuss in future sections/chapters, we have studied the possible ways to increase the efficiency of TBR by placing more RNs in more strategic locations where the probability of visiting RNs by MNs is much more than other locations. Therefore, the intelligent routing tries to put RNs at those points which allows delivering more data packets when we have a limited number of RNs to be distributed over the network. Our proposed routing protocol is very simple as the MNs only communicate with the RNs and no route between source and destination nodes are established. Additionally, since TBR does not rely on any specific RNs to deliver messages, a faulty RN cannot cause route flapping which may occur due to the hardware/software failure, unreliable connections (which is common in MANETs), and configuration errors...
3.2 Evaluation

In this section, we first describe a possible implementation of our proposed protocol using passive RFID tags and then present the methodology that we used to design and develop our simulations. Further, we evaluate the proposed protocol using simulation in different scenarios, e.g., different number of RNs, MNs, speed, buffer sizes, etc., and also we compare our protocol with the state-of-the-art routing protocols in the literature. Moreover, a deeper study of TBR evaluation under the effect of different mobility models and different relay placement techniques is provided.

3.2.1 TBR Implementations

In our simulation results the proposed protocol is considered for passive RFID tags, however, it would be equally valid to consider the results of our work on a network where the point locations are other type of wireless devices with a capability of storing messages, if such an infrastructure was feasible for the application. A comparison between different implementation of Stationary Relay Point approach is shown in Table 3.1. The network designer can choose between them based on the application.

3.2.1.1 RFID MANET Implementation

Radio-frequency identification (RFID) is an automatic identification method, which is based on remote data storage/retrieval using devices called RFID tags or transponders. An RFID tag is a module which is used for identification purposes using radio-waves. In this work we have considered its use for storing/retrieving messages in DTNs. Some commercially available tags in the industry can be read from a distance of several meters away and beyond the line of sight of the reader;
Table 3.1: Stationary Relay Points approaches.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>WiMAX Base Station</td>
<td>$24000</td>
<td>stand-alone/high power</td>
<td>500m-4km</td>
<td>&gt;0.005m³</td>
<td>capable</td>
</tr>
<tr>
<td>Active RFID tag</td>
<td>&gt;$20</td>
<td>stand-alone/med. power</td>
<td>up to 100m</td>
<td>≈3⁻⁵m³</td>
<td>capable</td>
</tr>
<tr>
<td>throwbox</td>
<td>$100 – $300</td>
<td>stand-alone</td>
<td>up to 250m</td>
<td>from 1.12⁻⁴m³</td>
<td>capable</td>
</tr>
<tr>
<td>Passive RFID tag</td>
<td>$.07 – $.20</td>
<td>powered by reader</td>
<td>up to 35m</td>
<td>10⁻⁸ to 4.8⁻⁴m³</td>
<td>unable</td>
</tr>
</tbody>
</table>

- [http://www.gaorfid.com/](http://www.gaorfid.com/)
- [Zhao et al.](http://www.willow.co.uk/Stargate_Datasheet.pdf) introduce Intel Stargate as a device which meets throwbox requirements. Stargate specification is available at [http://www.willow.co.uk/Stargate_Datasheet.pdf](http://www.willow.co.uk/Stargate_Datasheet.pdf)
3.2 Evaluation

For example, Mojix[^1] has recently developed a way to dramatically increase the range of passive RFID tags up to 200m, opening up many new applications for low-cost tags. Some tags are currently able to store kilobits of information; e.g., Fujitsu reported the development of a 64KByte High-Capacity FRAM RFID Tag[^2] in 2008. RFID tags can have larger memory size of up to 1MByte[^3]. RFID technology is rapidly improving and in the near future we expect to see more powerful RFID tags, which can be set-up more efficiently, with a wider range and higher memory capacity.

The RFID tags which are used in our work would be placed in predefined locations on the application area. They do not utilize any transducers, Global Positioning System (GPS), or routing modules for performing data delivery. Therefore, they can be easily deployable with low cost. Each tag would typically be made up of three parts: (i) an integrated circuit which is responsible for storing/processing information, modulating and demodulating the signals; (ii) an antenna to receive and transmit the signal; and (iii) a power unit. Passive RFID tags, that we consider in this work, are powered by the incoming signal and so they do not initiate communication.

From another point of view, MNs must be equipped with RFID readers. Recently, there has been significant attention in developing mobile RFID reader devices with Wi-Fi Network support and GPS[^4]. This makes it possible for the applications where the mobile station cannot approach the RFID tags by itself, by using a smaller unit equipped with the mobile reader that can do the task of reading/writing from/to the mobile station from/to the tags.

RFID technology is desirable because it operates wirelessly and without the need for attached power. This makes its deployment relatively easy and sustainable. It would be equally valid to consider the results of our work on a network where the point locations are wireless base stations (with no connections between the base stations), if such an infrastructure was feasible for the application. The increased range of a wireless base station compared to the range of a

[^1]: http://www.mojix.com/products/
[^4]: http://rfid.net/product-listing/reviews/176-csl-cs101-handheld-reader
3.2 Evaluation

read/writable RFID tag would serve to increase the effectiveness of our protocol in terms of delivery ratio and message latency.

3.2.2 Simulation Methodology

We implemented an event driven simulation of TBR in MATLAB\textsuperscript{®}. We assumed that we have an ideal wireless link layer that causes no collision with other transmissions during any transmission session. This simplification does not significantly affect our results because in a typical DTN, the density of MNs in a specific area is by definition insignificant (as discussed earlier in Section 3.1). Similarly, we assume communications from an RN to an MN and vice versa are instantaneous. In a real implementation, this assumption may not hold for some kinds of applications where MNs move relatively rapidly or cannot pause for the time taken to operate on an RN. However, there has been focus on designing low delay passive RFID reader systems which can accommodate this issue \cite{2, 94}. As an example, E-Toll RFID tags are widely used in Australia on all toll roads where drivers can pay tolls automatically without stopping\footnote{http://www.rta.nsw.gov.au/usingroads/etoll/tag/index.html}. Additionally, in Algorithm \ref{algorithm1} we showed how we can prioritize the messages based on their arrival time, to be exchanged between RNs and MNs for the time that they are being connected, using meta-data related to the messages. We leave the issue of read/write delay to future work.

Our simulation starts at time $t = 0$ with all RN buffers and MN buffers empty. MNs are placed randomly and messages arrive with exponentially distributed random inter-arrival times, at a global mean rate of $\lambda$, such that each MN generates messages at a mean rate $\lambda/N_{\text{node}}$. Each message has a single destination node picked uniformly at random from the MNs (not including the source node). The simulation is run for a fixed period of time, $t_{\text{max}}$.

In the following section, we study the effect of different node mobility models and different relay placement strategies for our protocol. Furthermore, we evaluate our protocol when using many small RNs and when using a super RN which is placed in the middle of the area.
3.2 Evaluation

3.2.3 TBR Evaluation

In this section we aim to evaluate the performance of TBR in general and the effects of its different setting parameters on the performance of the network in particular. The independent variables are including MN speed, MN buffer size, RN buffer size, RN transmission range, and RN spacing. We use the metrics introduced in Section 1.3 to evaluate the network performance. In addition, we show the required power consumption of MNs in terms of the required transmission number per second versus some of the independent variables. We use Random Waypoint (RWP) mobility model which is further studied in Section 4.1 in order to compare our results with the available results of the state-of-the-art approaches in the literature. We use a simple technique for placing the passive relay points which places the relays on a regular grid with a predetermined distance to each other. In Chapter 4 we show how we can significantly improve the performance of our protocol using a better relay placement technique. The queuing policy used for evaluation is fixed and it follows a first-in-first-out (FIFO) mechanism. Later, we show that we can further improve the performance of TBR using better buffer management policies. Finally, we present the distribution of message delivery delay using the proposed protocol.

3.2.3.1 Control Variables

We use the following settings unless otherwise noted: \( X_{max} - X_{min} = 1000, Y_{max} - Y_{min} = 1000, d = 50 \) (\( \Rightarrow N_{relay} = 441 \)), the average speed of MNs is 5 (\( S_{min} = 2.5, S_{max} = 7.5 \)) with no pause time in the Random Way Point model, \( B_{node} = 1000 \), \( B_{relay} = 100 \), and \( r = 5 \). These settings are mainly chosen based on a hypothetical environment monitoring application where mobile nodes are some mechanical devices with small amount of memory, i.e., 1000 messages, and passive relays are passive RFID tags which can hold 100 small messages. All messages have equal size. In most cases we have provided all simulations for \( N_{node} = 10 \) with 0.5 message/second per node (\( \lambda = 5 \)). RNs are placed on a regularly spaced grid.
3.2 Evaluation

3.2.3.2 TBR Effectiveness

Figures 3.12-19 show the effect of Node Speed, RN Buffer Size, RN Transmission Range, and RN Spacing on the delivery performance of the proposed protocol. According to the latter figures, in most cases the message delivery rate of our protocol is higher when there are less MNs in the network while the message latency is higher. The reason is that more MNs in the network leads to more generated messages in the network and the probability that MNs overwrite the RNs is higher as a result; therefore, messages have a lower probability to stay in the RN’s buffer for a longer time.

Figure 3.6 shows that increased MN speed significantly reduces the message delay. As the MN speed increases the MN meets more RNs per message, however, the MN may also meet the same RNs more than once and so higher speeds does not continue to decrease message delay with the same rate. Figure 3.2 shows the corresponding delivery rates. Similarly, meeting RNs sooner greatly increases the delivery rate because messages have a smaller chance of being dropped if they exist in more buffers. Figures 3.3 and 3.7 show that increased RNs transmission range also significantly increases the message delivery rate and reduces the message delay; however, the speed up in delivery performance is not as high as the effect of increased average MN speed. In general, with a larger RN transmission range, there is a larger chance that a MN meets more RNs per unit time which results in meeting more RNs per message and meeting RNs sooner.

Figures 3.4 shows that an increasing RN buffer size increases the message delivery rate much higher than the latter parameters. As the RN buffer size increases, the MNs can replicate more messages, spread over more RNs, and replicated messages can exist in more buffers for a longer period of time; therefore, messages could be delivered to the destination nodes with a higher chance instead of being dropped at an early stage. However, according to Figure 3.8 increased RN buffer size can lead to an increase in the average message delay. There are two factors which affect the average message delay: (1) an increased RN buffer size allows the messages to stay in the buffers for a longer time and those messages can be delivered in a longer period of time rather than being dropped when the buffer size is smaller, which thereby increases the average message latency;
(2) an increased RN buffer size reminds that more buffers can hold the same messages which leads to a decrease in the average message delay. According to Figure 3.8(b) (1) has a greater impact when the delivery rate is low while (2) has a greater impact when the message delivery rate is approaching 1.

Figures 3.5 and 3.9 show the delivery performance of the proposed protocol versus RN spacing, i.e., the total number of RNs in the regular array, as given in Section 3.1. Significant variation in delay does not occur until the number of RNs is quite small, i.e., RN spacing is quite big, and the RNs are near the perimeter of the region. Message delivery rate is significantly affected by the RN spacing and the number of MNs. A smaller number of RNs leads to a significantly decreased rate. For large numbers of MNs, there is a greater total arrival rate of messages and so buffers are exhausted more quickly.

In some cases when a large number of packets are lost, the delivery latency is improved as message latency is computed only over those messages that were successfully delivered.
3.2 Evaluation

Figure 3.3: Message Delivery Rate vs. RN Transmission Range.

Figure 3.4: Message Delivery Rate vs. RN Buffer Size.
3.2 Evaluation

Figure 3.5: Message Delivery Rate vs. RN Spacing.

Figure 3.6: Average Message Delay vs. Average MN Speed.
3.2 Evaluation

Figure 3.7: Average Message Delay vs. RN Transmission Range.

Figure 3.8: Average Message Delay vs. RN Buffer Size.
3.2 Evaluation

3.2.3.3 Transmissions per Second

Figure 3.10 shows the average transmissions per second that a mobile device uses as a result of TBR, for various RN spacing and speeds. The power requirement is directly proportional to the number of transmissions (because each transmission has constant power requirement). Slower moving nodes and fewer RNs both lead to fewer transmissions.

3.2.3.4 Power Consumption vs. Number of Relays

Passive nature of relay nodes in our protocol requires communication cost from the mobile nodes. Assuming that RFID readers on average need equal amount of energy for each interaction with relay nodes, e.g., Klair et al. present a table showing RFID readers power consumption, Equation shows the relation between the number of relays and the required power consumption for reading/writing the
3.2 Evaluation

![Figure 3.10: Transmissions per second required by a mobile device.](image)

where \( P_i[H] \) is the probability of hitting the relay \( i \) by a node per second and \( E_{r/w} \) is the average required energy for each interaction between nodes and relays. \( E_{r/w} \) depends on the storage space of the passive relay nodes and the traffic load.

In prior work [71], we showed how to calculate the average hit probability of a relay node by a mobile node using a conceptual mobility model. Further, in [72] we showed how we can extend this conceptual mobility model to RWP.

In order to see the average number of relay hits by mobile nodes per second, we ran an experiment where \( X_{max} - X_{min} = 1000, Y_{max} - Y_{min} = 1000 \), the average speed of mobile nodes is 5 \( (S_{min} = 2.5, S_{max} = 7.5) \) with no pause time in the Random Way Point model, \( r = 5 \), and \( N_{node} = 10 \) while we change the number of relay nodes placed on a regular grid. Figure 3.11 shows the corresponding results. According to Figure 3.11 by placing 2500 relays on a grid, on average 1.26 of nodes would hit the relays per second. If we place 400 relays on average every 5 seconds

\[ \bar{E} = N_{node} \sum_{i=1}^{N_{relay}} P_i[H] E_{r/w}, \]  

(3.1)
3.2 Evaluation

one mobile node would meet a relay and if we assume $E_{r/w} = 180\text{mwatt}$ for the case of Skye Module M1-Mini introduced in [37] then we need $36\text{mwatt/sec}$ for the read/write operations in the system.

![Graph showing average relay hit number per second versus the number of relays placed on a regular grid.]

Figure 3.11: The average relay hit number per second versus the number of relays placed on a regular grid.

In our simulation results we assume the worst case where the transmission happens at hit time when the distance between the end points equals the maximum transmission range of the tags. This assumption assures that the required power consumption would never be worse than our simulation results as if the mobile node gets closer to the relay it may require less power for communication. In other words, to calculate the required power for transmission we consider the largest possible distance between end points.

3.2.3.5 Distribution of Message Latency

Figure 3.12 shows the cumulative distribution function of message latency, for delivered messages, versus various parameters. Anecdotally, we have fit these distributions to the Generalized Extreme Value distribution using MATLAB `dfittool`: 
3.2 Evaluation

![Figure 3.12](image)

Figure 3.12: CDF of message latency vs. various parameters.

\[ f(x; \mu, \sigma, k) = \frac{1}{\sigma} \left( 1 + k \frac{x - \mu}{\sigma} \right)^{-\frac{1}{k}} e^{-\left(1 + k \frac{x - \mu}{\sigma}\right)^{-1/k}}, \]

for \(1 + k \frac{x - \mu}{\sigma} > 0\).

\[ 1 + k \frac{x - \mu}{\sigma} > 0. \]

The Generalized Extreme Value distribution was the only distribution that consistently fit the data over various ranges of parameters that we tested. Figure 3.13 shows an example message latency probability density function fit to the distribution. Further analytical work is required to substantiate this relationship.
3.2 Evaluation

Figure 3.13: PDF of message latency fit to a Generalized Extreme Value distribution. In this example, the distribution parameters are \( k = -0.104493, \sigma = 24.1557 \) and \( \mu = 39.6301 \).

3.2.4 Possibility of Using a Single Base Station to Cover the Area

In this section, we study the possibility of using a single super RN placed in the middle of the area instead of using many small RNs distributed over the area. Instead of using many small RNs, one can use a few super RNs to achieve similar delivery performance; however, there is a trade-off between the power usage and using longer radio range. In this section, we evaluate the the delivery performance of a super RN, i.e., a base-station, placed in the middle of area versus of its broadcasting range and later we show the effectiveness of using the base-station in comparison to using multiple regular RNs introduced in this thesis. The results show that using the latter alternative we achieve better performance in terms of power usage.

In order to evaluate the performance of using a single base-station placed in the middle of the region we ran an experiment. In the experiment we chose same configuration as scenario 1 in Section 4.3 except \( N_{\text{relay}} = 1 \) and \( B_{\text{relay}} = 1600 \). Our objective function is to maximize \( \rho \) and minimize \( L_{\text{avg}} \) simultaneously. Conceptually, network inefficiency could be defined vice versa given as \( L_{\text{avg}}/\rho \). The goal is then to minimize network inefficiency. Figure 3.14 shows that an increased
broadcast range of the corresponding base station improves the performance of
the network in terms of increasing message delivery rate and decreasing message
delay together. As an example, for the RWP mobility model, when the broadcast
range of the base station approaches 55\,m, the performance is almost the same as
the GA in Table 4.2. The covered area by RNs is equal to 942.48\,m^2 (16 \times \frac{3}{4}\pi 5^2)
but the covered area by the base station is 7127.49\,m^2 (\frac{3}{4}\pi 55^2); therefore, the
required area to be covered by the base station is 7.5625 times larger. If we as-
sume that to cover each unit of area we need the same amount of energy, then to
achieve the same delivery performance using a base-station we need 750\% more
power than using multiple RNs. In addition, Figures 3.16 and 3.15 show the
delivery rate and average end-to-end delay versus base station range respectively.

In some applications, such as military, relying on a single base station may
end up with a point of vulnerability which may be attacked and, if destroyed,
terminate the network operation. Additionally, in some scenarios such as ring like
areas, it is not possible to place a base station in the middle of the area which is
unreachable. Alternatively, we can place the RNs all over the rings (Figure 3.17).
3.2 Evaluation

Figure 3.15: Average Message Delay vs. Base Station Broadcasting Range.

Figure 3.16: Message Delivery Rate vs. Base Station Broadcasting Range.
3.3 Comparisons to Existing Work

3.3.1 Comparison to Epidemic Routing

In prior work, A. Vahdat et al. [83] presented a routing protocol called Epidemic Routing (ER). In ER, every host acts as a carrier to distribute application messages and whenever hosts meet each other they exchange all the messages together; therefore, after a while all the messages will spread over the network. In this section we will show how we can improve ER performance reported in [83] using TBR considering that we have involved a number of stationary RFID tags in the network.

Figure 3.18(b) shows the CDF of message delay for various transmission ranges in TBR. We use the same parameters used for evaluation in ER as following: the area size is $1500 \times 300$, the average speed of MNs is $10$ ($S_{\text{min}} = 0$, $S_{\text{max}} = 20$). The message rate is $1$ ($\lambda = 1$) and $d = 50$ ($\Rightarrow N_{\text{relays}} = 217$). The buffer size for all the RNs and MNs is 1000 messages.
3.3 Comparisons to Existing Work

Figure 3.18: TBR’s CDF for message delivery rate versus message delay, (a) and (b). Note that these CDFs cumulate to the percentage of delivered messages. Comparisons to existing work are shown in (c) and (d).
### 3.3 Comparisons to Existing Work

<table>
<thead>
<tr>
<th>Range</th>
<th>ER $\rho$</th>
<th>TBR $\rho$</th>
<th>$L_{avg}$</th>
<th>TBR $L_{avg}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>250(^a)</td>
<td>100</td>
<td>100</td>
<td>0.2</td>
<td>7.8</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>100</td>
<td>12.8</td>
<td>15.5</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>100</td>
<td>153.0</td>
<td>42.7</td>
</tr>
<tr>
<td>25</td>
<td>100</td>
<td>100</td>
<td>618.9</td>
<td>90.6</td>
</tr>
<tr>
<td>10</td>
<td>89.9</td>
<td>100</td>
<td>44829.7</td>
<td>248.9</td>
</tr>
</tbody>
</table>

\(^a\)This network is a DTN by our definition (average number of connections is more than 20\% and our definition requires less than 5\%)

Most of the messages are delivered by time 100, when $r \geq 25$. In ER, this phenomenon happens only when the transmission range, i.e., the wireless range between MNs, is more than 100. Furthermore, according to Table 3.2, ER results show that it is not scalable due to its weak performance for small transmission ranges, where e.g. it takes 44829.7 seconds to deliver a message, on average. ER has a better performance when we have a high density network, i.e., 50 MNs with transmission range of higher than 100 in an area $1500 \times 300 m^2$.

If we scale the size of network from $1500 \times 300$ to a larger size or reduce the number of participating MNs, using even a larger transmission range, TBR will outperform the ER approach. In addition, the message delivery rate of TBR will be higher in very sparse networks.

We have also investigated the performance of TBR and ER by using different MN/RN buffer sizes. According to Table 3.3, when the transmission range is 50 meters, TBR has a higher delivery performance than ER. This case is more significant when we have a limited buffer size.

In ER, they also presented the CDF for bounded resource consumption, i.e., bounded buffer sizes. In this case, the amount of buffer space in the MNs is limited. Assuming all the MNs and RNS in TBR have the same amount of buffer space and all the parameters are the same as earlier with a transmission range of 50 for all the RNS, the results for message delivery rate are plotted in Figure 3.18(a).
3.3 Comparisons to Existing Work

Table 3.3: TBR vs. ER (different buffer size).

<table>
<thead>
<tr>
<th>Buffer size</th>
<th>ER ρ</th>
<th>TBR ρ</th>
<th>L_{avg}</th>
<th>L_{avg}</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>100</td>
<td>100</td>
<td>147.3</td>
<td>40.6</td>
</tr>
<tr>
<td>1000</td>
<td>100</td>
<td>100</td>
<td>148.7</td>
<td>41.6</td>
</tr>
<tr>
<td>500</td>
<td>100</td>
<td>100</td>
<td>149.2</td>
<td>43.8</td>
</tr>
<tr>
<td>200</td>
<td>99.6</td>
<td>100</td>
<td>152.0</td>
<td>46</td>
</tr>
<tr>
<td>100</td>
<td>95.2</td>
<td>99.5</td>
<td>157.5</td>
<td>41.9</td>
</tr>
<tr>
<td>50</td>
<td>79.7</td>
<td>94.1</td>
<td>148.2</td>
<td>41.4</td>
</tr>
<tr>
<td>20</td>
<td>50.2</td>
<td>73.1</td>
<td>129.5</td>
<td>35.2</td>
</tr>
<tr>
<td>10</td>
<td>29.3</td>
<td>53.2</td>
<td>98.9</td>
<td>34.3</td>
</tr>
</tbody>
</table>

3.3.2 Comparison to Message Ferrying

In prior work, W. Zhao et al. [95] presented a Message Ferrying (MF) approach for data delivery in sparse MANETs. They described two approaches called NIMF and FIMF. We chose the NIMF (Node-Initiated MF) to be compared with our work based on their results reported in [95]. We ran our simulation based on the same parameters reported in [95] as following: the area is 5000 × 5000, there is 1 ferry with average speed of 15 m/s, S_{min} = 0, S_{max} = 5, λ = 1.25, d = 50 (= N_{relays} = 10201), N_{node} = 40 and r = 20. Note that from a comparison point of view the transmission range of the RNs is much less than the transmission range of the MNs in NIMF.

Figures 3.18 (c) and (d) show the effect of MN/RN buffer size on average message delay and message delivery rate for different approaches including TBR, NIMF, and ER. The delivery rate in TBR is significantly better than other approaches, because of the added buffer capacity of the RNs, which makes it suitable for bounded resource lossless applications. The message delay in TBR seems to be high for small buffer size but we believe that this higher delay is due to having more messages delivered because there are some messages in the network which are delivered after a long time and in NIMF they are dropped. This phenomenon is intensified in ER which has a low delivery rate (according to Figure 3.18(c), by increasing buffer size in ER/NIMF, the delay becomes higher which confirms our
3.4 Conclusion

In this chapter, we studied the applicability of using point locations as “mailboxes” for storage/retrieval of messages to facilitate ubiquitous network connectivity. A practical implementation of our work could use low cost, tiny, unattached-power RFID tags. Our approach is an alternative, for achieving competitively low message delay and high message delivery rates, to existing methods that rely on altering/controlling some of the MNs’ movements. We designed and analyzed a new protocol called TBR (Tag-based Routing) in which MNs communicate to each other only via passive RNs such as passive RFID tags. Our results show that TBR is effective at expanding the connectivity of DTNs, over a broad range of parameter values. Our TBR protocol provides lower message delay and higher message delivery rates than existing methods including NIMF and ER. In the next chapter, we study TBR to develop more advanced relay placement techniques and buffer management policies to increase its performance.
Chapter 4

Heuristics for Improving TBR Delivery Performance under Different Mobility Models

In chapter 3, we proposed TBR as an alternative routing protocol in DTNs which outperforms state-of-the-art routing schemes in the literature. In this chapter, we first study possible enhancements on our proposed routing protocol, TBR, through different relay placement techniques. We also evaluate the improvement on the performance of TBR for different mobility models including Entity Mobility Models and Group Mobility Models. As the performance of TBR is highly dependent on the participating stationary relay nodes’ locations, we demonstrate several techniques for optimizing the stationary relay node placement, namely: relay pruning, probability based relay distribution, and a genetic algorithm (GA) based optimization. We show their effect on the performance of the proposed protocol as well as comparing their effectiveness with the existing placement techniques in the literature. Our simulation results show that our protocol outperforms the
4.1 Mobility Models

Mobility is a natural phenomenon in DTNs. Studying the mobility models in DTNs is important because it has a direct impact on the performance of network protocols. Further, to evaluate the performance of TBR, we need to implement network scenarios in which MNs need to move based on a specific pattern. Therefore, in this section, we will describe some widely used mobility models which are later used to evaluate TBR.

There have been many mobility models proposed so far in the literature. Camp et al. [14] classifies mobility models for ad hoc network research into two categories: entity mobility models and group mobility models. In entity mobility models, the movement of one MN is independent of all other MNs, whereas in group mobility models, a set of MNs move in a group. The widely used representatives of entity mobility models are Random Walk Model and Random Waypoint Model [7, 8], while Reference Point Group Mobility (RPGM) Model [30] is a well-studied representative of group mobility model. Although, some
of the above mobility models are artificial, they are widely used in the literature [83, 85] to evaluate the performance of routing protocols. Therefore, in this thesis we use them to evaluate the performance of our protocol and also to compare our protocol with some of the state-of-the-art protocols which use similar mobility models. However, there is other work that uses traced-based simulation. For example, N. Banerjee et al. [6] use traces from a bus route model to evaluate the effectiveness of their work.

4.1.1 Entity Mobility Models

In this class of mobility models, MNs move independently to others. We detail some of the widely used entity mobility models in this section and further we propose a new mobility model called Restricted Random Waypoint Model.

4.1.1.1 Random Waypoint Model

In this model, every MN chooses a random destination called waypoint in the simulation area and moves toward it in a straight line with a constant speed which is selected randomly from the inclusive interval $[S_{\text{min}}, S_{\text{max}}]$. After approaching the destination, it stays there for a certain period of time (i.e., a pause time) and then it repeats the procedure. Figure 4.1 shows the traveling pattern of a single MN using the Random Waypoint Model.

4.1.1.2 Random Walk Model

In this model, every MN chooses a new random direction in the polar angle interval $[0, 2\pi]$ and moves from the current location toward its new location with a speed value of $s$ which is selected randomly from $[S_{\text{min}}, S_{\text{max}}]$. Therefore, it is completely memory-less from speed and direction points of view [4] and it can be used in vehicular and large-scale environments. There are two options for each MN to choose the new movement: after a constant time interval $t$ or a constant distance traveled $d$. When the MNs hit the boundaries they are wrapped to the
4.1 Mobility Models

Figure 4.1: Traveling pattern of a MN using the Random Way Point mobility model (50 steps). Red square indicates the start location and blue square shows the end location of the MN.

Figure 4.2: Traveling pattern of a MN using the Random Walk mobility model (50 steps). Red square indicates the start location and blue square shows the end location of the MN.

other side of the simulation area from where they continue their trip [28]. Figure 4.2 shows the traveling pattern of a single MN with a constant time interval.
4.1 Mobility Models

4.1.1.3 Restricted Random Waypoint Model

The Random Waypoint Model is the most widely studied mobility model in MANETs. We propose a restricted version of the Random Way Point model where every MN can move only in a restricted area. Figure 4.3 shows the traveling pattern of 4 typical MNs where the area is divided to 4 sub-areas with 40% overlap in each sub-area; however, we can divide the area based on the application to different sections. According to Figure 4.3, in two sub-areas there is only one MN while in other sub-areas there have two and no MNs respectively.

4.1.2 Group Mobility Models

There are situations in DTNs where a set of MNs move together. In this section, we detail the most widely used representative of the group mobility models, called Reference Point Group Mobility.
4.1 Mobility Models

4.1.2.1 Reference Point Group Mobility Model

The Reference Point Group Mobility (RPGM) model is the most general group mobility model. In this model, each group of MNs has a random motion and inside of each group, each individual MN has a random motion as well [30]. In this model, each group has a virtual center and group movements are represented by the locus of the center. The virtual center moves according to a group motion vector, $\vec{GM}$, Randomly chosen or predefined, and every MN adds a random motion vector, $\vec{RM}$, to its reference point. Therefore, MN movement equals the sum of two components: $G\vec{M} + R\vec{M}$ in which $G\vec{M}$ is motion of reference point (RP) and $R\vec{M}$ is relative random motion around reference point (Figure 4.4). RPGM is a very effective framework and maybe adapted to model various scenarios, like disaster recovery, battlefield, and convention scenario. Figure 4.5 shows the traveling pattern of five MNs as a group using RPGM.

4.1.3 Discussion

Delivery performance of routing protocols in DTNs is highly dependent on node mobility models; hence, studying mobility models is important in this domain. Mobility models represent spatial and temporal characteristics of mobile nodes.
including changes on their location, velocity, and acceleration over time. In this section we studied these mobility models and categorized them in two major classes including entity mobility models and group mobility models.

Table 4.1 represents a comparison between the mobility models we used in this thesis. According to Table 4.1 random walk mobility model has sharp and sudden turns while nodes are handled by wrapping to the inside of the region when encountering the borders. This prevents its applicability to real applications. However, its simplicity and its evenness to cover all the locations makes it a good alternative for evaluating the performance of protocols. Random waypoint model is the most widely used mobility model in the literature; accordingly, it can be used in order to compare the performance of a protocol with the state-of-the-art protocols. It also is a building block to make other mobility models such as restricted random waypoint and RPGM which we use in this thesis. However, lack of regular node movement and memory-less behavior of this mobility model is an issue in modeling. Density wave is another drawback of this mobility model as mobile nodes tend to move in the middle of the region than corners. Restricted random waypoint model is built on the random waypoint model which has a smoother wave density. It can be used in real applications such as agricultural...
4.2 Relay Placement Techniques

Placement of RNs plays an important role in our protocol as its performance is dependent on their positions. RN placement is already studied in VDTN [22], however, in this section, we propose different relay placement strategies which are developed in this thesis.

Under the random way point model, Figure 4.6 shows a sample frequency distribution of the messages carried by the different RNs in a uniform grid, where RNs are deployed based on a uniformly spaced RN array with distance $d = 100$ between RNs, i.e., 121 RNs, and it shows (as expected) that RNs at the center of the region have a higher probability of being met than RNs placed at the perimeter, hence the RNs around the center are responsible for delivering more messages than the ones at the edges. Considering this non-uniform message load among RNs, it is natural to consider what the optimal position of $N_{\text{relay}}$ RNs would be in order to minimize $L_{\text{avg}}$ and/or to maximize $\rho$. We evaluate different relay placement schemes as described below.

4.2.1 Uniform Grid

In this scheme, RNs are placed on a regularly spaced grid with known distance $d$ between two neighbor RNs, such that the RN coordinates are $(X_{\text{min}} + i d, Y_{\text{min}} +$
### Table 4.1: Mobility Models Comparison.

<table>
<thead>
<tr>
<th>Mobility Model</th>
<th>Class</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>RWP</td>
<td>entity</td>
<td>most commonly used mobility model in the literature/ building block for developing a variety of other mobility models</td>
<td>lack of regular movement modeling/density wave/memory-less movement behaviors</td>
</tr>
<tr>
<td>Random Walk</td>
<td>entity</td>
<td>simple implementation/covering all locations</td>
<td>sharp and sudden turns/unrealistic wrapping</td>
</tr>
<tr>
<td>Restricted RWP</td>
<td>entity</td>
<td>covering more locations than RWP/ applicable to more applications than RWP</td>
<td>density wave in some areas</td>
</tr>
<tr>
<td>RPGM</td>
<td>group</td>
<td>realistic mobility model</td>
<td>temporal and spatial dependency to other nodes</td>
</tr>
</tbody>
</table>
Figure 4.6: RN usage frequency distribution for 121 RNs.
4.2 Relay Placement Techniques

![Figure 4.7: RN positions after pruning.](image)

\[ N_{\text{relay}} = \left( \left\lceil \frac{X_{\text{max}} - X_{\text{min}}}{d} \right\rceil + 1 \right) \left( \left\lceil \frac{Y_{\text{max}} - Y_{\text{min}}}{d} \right\rceil + 1 \right). \]

Although this placement technique is not optimal, for simplicity we mostly consider uniform grid placement as a baseline to evaluate the proposed protocol.

### 4.2.2 Relay Pruning

This is a simple pruning scheme where a certain percentage of RNs are removed based on the number of messages delivered by the RNs when placed in a regular grid. RNs responsible for delivering least number of messages are removed until the specified percentage of RNs are retained. The remaining RNs will still remain
at the original grid locations. Figure 4.7 shows a sample RN distribution after 40% of 121 regularly spaced RNs are pruned based on this strategy.

### 4.2.3 Probability based Relay Distribution

By analyzing the number of messages delivered by different RNs, when placed in a regular array, RNs can be redistributed based on the probability distribution of the messages carried by the RNs (higher RN densities in areas where the probability of a message being carried is high, and lower RN densities in areas where the probability of a message being carried is low). We have implemented this scheme using an efficient re-sampling scheme given in Table 3.2 of [65], for repeating points, followed by jittering the points by adding Gaussian noise. Figure 4.8 shows one realization of a random distribution of 121 RNs that follow a probability distribution derived from a message frequency distribution.

Figures 4.9(a) and (b) show the average message latency and delivery rate
4.2 Relay Placement Techniques

(a) Average Message Latency.

(b) Delivery Rate.

Figure 4.9: Effect of non-uniform relay distribution.

compared to a regular array with the same number of RNs for the two schemes described above. For this experiment, we used $N_{\text{node}} = 10$, $r = 5$, $B_{\text{relay}} = 100$, $\lambda = 5$, i.e., 0.5 message/second per node, $X_{\text{max}} - X_{\text{min}} = 1000$, $Y_{\text{max}} - Y_{\text{min}} = 1000$, $S_{\text{min}} = 5$, $S_{\text{max}} = 15$, $B_{\text{node}} = 1000$, and MNs move according to the RWP model with no pause time. In the relay pruning scheme 40% of the RNs are pruned based on the load carried. It can be seen that both non-uniform relay distribution schemes outperform the regular relay placement in terms of message latency and message delivery rates. Note that relay pruning shows an equivalent performance to a regular array but uses less RNs.

4.2.4 Genetic Algorithm based Optimization

To compare the previous heuristics to a more general heuristic, we have implemented a basic genetic algorithm (GA). We did not intend to undertake an extensive study of GA solutions, but rather we intended to compare previous heuristics. GAs have been used before to find the optimal location of base-stations (transmitters) in order to satisfactorily cover subscribers [26, 49, 80]. There are many
4.2 Relay Placement Techniques

other non-linear optimization techniques, e.g., simulated annealing, neural networks, etc., that could be explored but we leave them to future research. For these reasons, we have referred the details of our GA approach to Appendix B. Our discussion assumes a general knowledge of GA terminology that can be found here [50].

We implemented a genetic algorithm, using MATLAB Genetic Algorithm and Direct Search Toolbox, to search for the locations of \( N_{\text{relay}} \) RNs in the region with the following objective function:

\[
\text{minimize} \left[ \frac{L_{\text{avg}}}{\rho} \right].
\]

Figure 4.10 shows the effect of the GA solution compared to the probability based distribution on placement of the same number of RNs. For this experiment, we used \( N_{\text{node}} = 7 \), \( r = 5 \), \( B_{\text{relay}} = 100 \), \( \lambda = 3.5 \), i.e., 0.5 message/second per node, \( X_{\text{max}} - X_{\text{min}} = 1000 \), \( Y_{\text{max}} - Y_{\text{min}} = 1000 \), \( S_{\text{min}} = 5 \), \( S_{\text{max}} = 10 \), \( B_{\text{node}} = 1000 \), and MNs move according to the RWP model with no pause time. The GA solution consistently pushes the RNs closer together, into the center of the region. In this figure, GA is shown to be the superior relay placement among the other methods in both latency and delivery rate. We undertake a comprehensive simulation study, reported in section 4.3, for different node mobility models introduced in section 4.1 and also we will show how our GA based placement technique outperforms state-of-the-art placement techniques in the literature.

4.2.5 Discussion

In this section, we proposed different relay placement techniques in order to enhance the performance of TBR. Some placement techniques are independent of mobility models; consequently, their performance can be degraded in different scenarios. However, other placement techniques can adaptively find an optimal solutions given the mobility model. Further, others such as GA-based optimization are able to find the optimal solution through multiple non-optimal solutions in parallel. In order to evaluate the performance of each placement techniques,
4.2 Relay Placement Techniques

(a) Probability based tag distribution.

(b) GA based distribution.

(c) Average Message Latency.

(d) Delivery Rate.

Figure 4.10: Effect of GA approach and probability based distribution on placement of RNs.
4.3 Evaluation of Relay Placement Techniques using Various Mobility Models

Section 4.3 represents a comparison between the latter placement approaches under the effect of various mobility models introduced in Section 4.1.

4.3 Evaluation of Relay Placement Techniques using Various Mobility Models

In this section, we study the effect of different relay placement strategies on the performance of TBR. In this study, we also consider the effect of mobility models mentioned in Section 4.1 on TBR effectiveness. We compare our placement strategies with the ThrowBox algorithm proposed in [96]. In ThrowBox, base stations are greedily placed one by one so as on a grid of possible places in such a way to maximize the network throughput. In other words, given the traffic model and the mobility model of MNs, all possible positions for a base station are considered. After placement of one base station, the algorithm continues to the next base station excluding the previously taken positions. Since our optimization to find the optimal position, is based on $E$, in the case of ThrowBox we used the same objective function for optimization, which is different to the original objective function, i.e., increasing the delivery rate, that they used in their work.

The simulation methodology is as described in Section 3.2.2. We also used the following configuration to evaluate the efficiency of each placement technique: the average speed of MNs is 20 ($S_{\text{min}} = 10$, $S_{\text{max}} = 30$) and there are 16 RNs in the network ($\Rightarrow N_{\text{relay}} = 16$). We show using a relay placement optimization technique we can still achieve competitive message delivery performance even using low number of RNs, i.e., 16. Other settings are as Section 3.2.3.1. Simulation time is 5000 seconds and each result is based on 10 different runs to reduce the existing variation in the simulation. This configuration is chosen based on the results of previous experiments reported in Section 3.2.3. In RPGM, there are 3 different groups moving together. One group includes 4 MNs while the others include 3 MNs. In the Random Walk model, a constant time interval $t = 10$ seconds
4.3 Evaluation of Relay Placement Techniques using Various Mobility Models

Figure 4.11: An square area is partitioned to 4 equal sized sub-areas which has an overlap equals to 40% of area length. Node mobility is allowed only inside its respective sub-area.

is chosen to choose the new movement. In addition, in Restricted RWP model, the area is partitioned to 4 equal sized sub-areas which have an overlap equal to 40% of the area length (Figure 4.11) and the MNs are distributed uniformly in these sub-areas.

Furthermore, to evaluate the effect of relay placement techniques on TBR, two scenarios are studied (Figure 4.12). In scenario 1, TBR performance is measured based on the same instances of traffic and mobility model which are used in each relay placement strategy. The results are shown in Table 4.2. According to Table 4.2, the GA approach exhibits the best performance among other approaches and the ThrowBox approach has the best second performance. The reason that the GA outperforms the ThrowBox approach is that the GA approach searches the area for a set of placement solutions than searching for a single RN position at a time as in ThrowBox. Placing the relays on a regular grid and the relay pruning
4.3 Evaluation of Relay Placement Techniques using Various Mobility Models

![Figure 4.12: The process of measuring the network performance for different placement techniques.](image)

approaches have the worst performance as they place the relays on a grid with no respect to the mobility models of the nodes. Accordingly, they do not place many relays in strategic locations. Although the relay pruning approach trims some of the RNs responsible for delivering the least number of messages, its performance is very poor to make it a good alternative. The probability based distribution approach performs much better than the regular grid placement and the relay pruning; however, its performance is much less than the GA approach.

In scenario 2, new instances of traffic and mobility models are used in relay placement techniques. This scenario is more general to be used as it is independent of the traffic model. The results are shown in Table 4.3. According to Table 4.3, the GA approach is the best in ranking among other placement techniques except for the Random Walk mobility model where ThrowBox competes with the GA. For the same reason mentioned above, the GA approach can learn spatial characteristics of the relays in a more effective way than other approaches for all mobility models except the Random Walk which results in having a better performance than other approaches. More investigation is required to find out why the GA approach is unable to effectively learn the Random Walk characteristics. One hypothesis to describe this is related to the limited iteration number of learning process used in the GA. As the Random Walk has the highest random behavior among the other mobility models, by adding a random traffic model the problem of learning spatial characteristics of the relays becomes more complicated and needs more observations.

The results in this section are based on the specific cases as the genetic algorithm placement technique takes excessive computation to complete.
### Table 4.2: Comparison of tag placement techniques (same traffic/mobility instance).

<table>
<thead>
<tr>
<th></th>
<th>RWP</th>
<th>Restricted RWP</th>
<th>RWalk</th>
<th>RPGM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>$L_{avg}$</td>
<td>$E_a^b$</td>
<td>$N_{r}^b$</td>
</tr>
<tr>
<td>Reg. Grid</td>
<td>.122</td>
<td>192.5</td>
<td>6.34</td>
<td>16</td>
</tr>
<tr>
<td>Relay Pru.</td>
<td>.113</td>
<td>168.3</td>
<td>6.71</td>
<td>6</td>
</tr>
<tr>
<td>Prob. bas.</td>
<td>.286</td>
<td>153.9</td>
<td>18.6</td>
<td>16</td>
</tr>
<tr>
<td>GA</td>
<td>.424</td>
<td>100.1</td>
<td>42.4</td>
<td>16</td>
</tr>
<tr>
<td>Throwbox</td>
<td>.396</td>
<td>114.7</td>
<td>34.5</td>
<td>16</td>
</tr>
</tbody>
</table>

$^a \times 10^{-4}$

$^b N_{relay}$

4.3 Evaluation of Relay Placement Techniques using Various Mobility Models
### Table 4.3: Comparison of relay placement techniques (different traffic/mobility instance).

<table>
<thead>
<tr>
<th></th>
<th>RWP</th>
<th>Restricted RWP</th>
<th>RWalk</th>
<th>RPGM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>$L_{avg}$</td>
<td>$E_a^a$</td>
<td>$N_r^b$</td>
</tr>
<tr>
<td>Reg. Grid</td>
<td>.121</td>
<td>197.1</td>
<td>6.14</td>
<td>16</td>
</tr>
<tr>
<td>Relay Pru.</td>
<td>.111</td>
<td>172.3</td>
<td>6.44</td>
<td>6</td>
</tr>
<tr>
<td>Prob. bas.</td>
<td>.267</td>
<td>155.4</td>
<td>17.2</td>
<td>16</td>
</tr>
<tr>
<td>GA</td>
<td>.349</td>
<td>115.7</td>
<td>30.1</td>
<td>16</td>
</tr>
<tr>
<td>Throwbox</td>
<td>.348</td>
<td>120.6</td>
<td>28.9</td>
<td>16</td>
</tr>
</tbody>
</table>

$^a10^{-4}$

$^bN_{relay}$
4.4 Queuing Policies and Forwarding Strategies

In Section 4.3, we showed through simulation study that a better placement technique can significantly increase the performance of TBR under the effect of different mobility models. In our study, we use a simple first in first out (FIFO) mechanism to deal with buffer constraints. Relay nodes in TBR have a limited buffer capacity which directly limits the delivery performance.

An effective buffer management policy can increase the performance of routing protocols via selection of more important messages to remain in the buffers for a longer time, increasing their chance to be delivered. In this section, we introduce and evaluate different queuing policies and forwarding strategies in order to enhance the performance of TBR. Queuing policies can handle the selection of important messages in an overflow situation while forwarding strategies can manage the number/order of forwarding the messages. In this section, we show that traditional buffer management policies like drop-tail or drop-front fail to consider all relevant information in this context and are, thus, sub-optimal through simulations.

4.4.1 Queuing Policies

In a typical DTN, the mobile nodes are disconnected for a relatively large period of time to any other mobile/relay nodes. During this time, due to their limited buffer size they have to evict some messages to accommodate new messages. Additionally, when a mobile node forwards/reads some messages to/from a relay node, both mobile and relay nodes may have to evict some other messages from their buffer to give space to the incoming messages. Therefore, a queue management policy is quite demanding to define which messages should be dropped, if the buffer is full, when a new message has to be accommodated. In this section, we introduce two different queue management policies which are studied in the literature and also we evaluate their effects on the performance of TBR.

We evaluated TBR in the prior chapter based on the first in first out (FIFO) queuing Policy. In FIFO queuing Policy the message that first entered the queue is the first message to be evicted from the queue. This policy is easy to be
4.4 Queuing Policies and Forwarding Strategies

implemented but it can be inefficient as it does not consider other useful information about the probability of a message reaching the destination. Therefore, we implement another queuing Policy called evict most forwarded first (MOFO) in which nodes/relays keep track of the number of times each message has been forwarded and based on this information they evict those messages that have been forwarded the largest number of times. Consequently, nodes/relays can provide a higher chance of forwarding to the messages that have been forwarded fewer times.

We ran different experiments to evaluate the performance of TBR using MOFO and also to compare it with FIFO. In these experiments we use the following settings: \( N_{tag} = 36, N_{node} = 10, X/Y_{min/max} = 500m, B_{tag} = 100, \quad B_{node} = 1000, r = 20, \quad \lambda = 0.5, \quad S_{min/max} = [20, 30] \), i.e., the average node speed is randomly chosen from the latter range, and the node mobility model used is RWP. We also use the probability based relay optimization placement technique. This technique is quite easy to use for simulation purposes while it can place the relays in more strategic locations in comparison to a grid placement technique leading to highlight the changes in the TBR performance.

Figure 4.4.1 shows the results of the first experiment. In this experiment we vary the average mobile node speed. According to Figure 4.4.1(a), MOFO outperforms FIFO in terms of message delivery rate; however, it has a higher message delay which is shown in Figure 4.4.1(b). MOFO provides a higher delivery rate than FIFO as it allows more messages to reside in the buffers for a longer time leading to an increase in the forwarding opportunity of more non-identical messages to the destinations. This is shown in Figure 4.4.1(a) where MOFO can provide up to 20% more delivery rate than FIFO when mobile nodes move faster than 10m/s. However, there is a trade-off between increasing the message delivery rate and having increased message delay. Therefore, MOFO is more suitable for delay tolerant applications by focusing on increasing the delivery rate than decreasing the message delay.

Figures 4.4.1 and 4.4.1 further confirm that MOFO can provide a higher message delivery performance in comparison to FIFO under the effect of relay buffer size and relay number varying, respectively. Figure 4.4.1 shows the results of the second experiment when the TBR delivery performance is shown using
4.4 Queuing Policies and Forwarding Strategies

Figure 4.13: Delivery performance of MOFO versus FIFO respecting to average MN speed.

MOFO and FIFO subject to relay buffer size. For the very low relay buffer size, e.g., $B_{\text{relay}} = 10$, the message delivery of both queuing management policies approaches equality. In this case, during the required time by a MN to visit an RN, the number of messages generated by MNs can be much more than the capacity of the relay buffer. Therefore, on each visit, MNs may overwrite the relays with their own newly generated messages. Figure 4.11 shows the results of the last experiment when the TBR delivery performance is shown subject to the number of participating relays. According to Figure 4.11 MOFO can approach 100% message delivery rate using 120 relays while FIFO’s delivery rate is hardly more than 80%.

4.4.2 Forwarding Strategies

In Section 4.4.1 we showed that using an appropriate queue management policy, message delivery performance can be improved. However, there are some situations such as power management that requires considering the number/order of message forwarding. Further, due to finite buffer sizes, it is desirable to not occupy buffers with messages that have a low chance to reach their destinations.
4.4 Queuing Policies and Forwarding Strategies

![Figure 4.14](image_url)

**Figure 4.14**: Delivery performance of MOFO versus FIFO respecting to average RN buffer size.

![Figure 4.15](image_url)

**Figure 4.15**: Delivery performance of MOFO versus FIFO respecting to the number of RNs.
4.4 Queuing Policies and Forwarding Strategies

Figure 4.16: By deleting the copies of previously delivered messages the performance of TBR is increased. We used the following setting: \( N_{tag} = 36, N_{node} = 10, X/Y_{\text{min/\max}} = 1000m, d = 200, B_{node} = 1000, B_{relay} = 100, r = 50, S_{\text{\min/\max}} = [20, 30], \lambda = 0.5, \) and the mobility model used is random waypoint. Relays are placed on a regular grid.

Therefore, the order in which the nodes forward the messages to the relays as well as the number of messages to be forwarded are important factors affecting the performance of a routing protocol in DTNs. The policy that is used to handle these issues is called a forwarding strategy. In the rest of this section we present different forwarding strategies and we compare their effectiveness.

4.4.2.1 DDM - Delete Delivered Messages

In this policy, destination nodes delete the copies of previously delivered messages from encountered RNs. Figure 4.16 shows that using this simple technique, the performance of TBR in terms of message delivery rate is increased while the latency is almost at the same level.

4.4.2.2 INF - Intermittent Forwarding

In this approach, mobile nodes exchange their messages with every second relay on their journey. We refer to this approach as Intermittent Forwarding (INF).
4.5 TBR Robustness

INF is useful for dynamic policies as the mobile nodes’ battery level getting low, they can switch to this mode. Additionally, using this strategy the number of transmissions in the network is decreased by the factor of 50%. In order to compare INF policy with the regular policy used in TBR we ran an experiment based on the same settings as in 4.4.1. Figure 4.4.2.2 shows the corresponding results. According to Figure 4.4.2.2, the message delivery rate is decreased by a factor of less than 10% and message delay increased by almost 35%; however, the total number of transmissions in the network decreased by the factor of 50%.

![Figure 4.17: By controlling the transmission number we can address the overhead/power consumption of TBR.](image)

4.5 TBR Robustness

In the last sections we showed possible enhancements on the delivery performance of TBR using efficient relay placement techniques and buffer management policies. In this section, we evaluate the robustness of TBR against faulty nodes using our placement techniques. Robustness is a key characteristic of a routing scheme in DTNs as network failure is a natural issue in this domain. A failure in a part of network should not severely limits the performance of the network. Though
passive RFID tags, as an attractive alternative for passive RNs, do not need any attached supply of power and also RFID technology does not require line-of-sight reading, there are situations that the RNs fail in communication. For example, consider the situation where passive RNs are spread by unmanned aerial vehicles over the region and for some reasons some of them are damaged and cannot be read by MNs. In that case, our protocol should be fault-tolerant, meaning that failures in RNs should not significantly affect its performance. To evaluate the robustness of the TBR, we show an experiment based on various placement techniques. In our experiments, the network is subjected to permanent RN failures of random locations.

Our experiment is based on three different placement techniques namely regular (uniform) grid, probability based relay distribution, and GA based optimization. After placing the RNs by each placement technique, some RNs are chosen independently at random, based on the fault threshold, to fail permanently. We use the same settings as in Section 3.2.3.1 except: $X_{\text{max}} - X_{\text{min}} = 500$, $Y_{\text{max}} - Y_{\text{min}} = 500$, and $d = 100$ ($\Rightarrow N_{\text{relay}} = 36$).

Figure 4.18 shows the results in terms of message delivery rate. According to this figure, the performance of TBR in terms of message delivery rate on average is never less than 50% of the minimum of its maximum performance for each placement techniques subject to permanent failure of half of the RNs. For example, the delivery rate when no RN fails on average is 40.27% and when on average half the RNs fail is 27.56%. Also, at the point when on average 90% of the RNs fail permanently the protocol is still working with more than 30% of its normal delivery rate. For example, in the case of probability based relay distribution, the delivery rate is 12% while the normal performance at the point of no failure is 23%. The most affected technique is GA based optimization since in this placement technique all the RNs have a more significant role in delivering messages and a failure in some of them can drop the performance more than other placement techniques.

Figure 4.18 shows the changes in the message delay subject to RN failure. According to this figure, GA based optimization has the lowest delay in all cases. However, there is no dramatic changes in the delay for any placement techniques as non-faulty RNs have almost the same throughput, i.e., number of delivered
Figure 4.18: Message Delivery Rate vs. RN Failure Rate
4.6 Conclusion

In this chapter, we studied TBR to improve its delivery performance using more advanced relay placement techniques and more enhanced buffer management policies in order to increase its effectiveness. We first represented various mobility messages per time. Delivered messages at any time in this experiment reach their destinations almost in the same time. However, the variation in GA based optimization is much lower than the other techniques when the fault rate is increasing. RNs have a higher chance to be placed at non-strategic locations by regular grid and probability based optimization techniques than using the GA optimization technique; accordingly, a failure in any RNs placed in a more strategic place has a more significant impact than a failure in other RNs which leads to an increased variation in delay in comparison to GA based optimization as fault rate increases.

Figure 4.19: Message Latency vs. RN Failure Rate
models including entity mobility models and group mobility models in order to evaluate TBR more comprehensively and also in order to evaluate the effectiveness of proposed techniques to enhance TBR under different scenarios. Then we showed heuristic methods for placing stationary relay nodes over the geographic region, including relay pruning, a probabilistic distribution based on relay node utility, and a genetic algorithm based optimization approach that attempts to minimize the message delay/delivery rate ratio. The genetic algorithm based optimization was shown to be a superior relay placement technique under effect of different proposed mobility models. Afterwards, we proposed different buffer management policies including queuing and forwarding strategies to further enhance TBR. Our simulation results showed that our queuing policies and forwarding strategies introduced in this chapter can significantly improve the delivery performance of TBR. Finally, we evaluate a key characteristic of TBR, robustness, in DTNs. The simulation results showed that TBR is robust against of faulty relays. In the rest of this thesis, we study TBR to develop a rigorous analytical framework to study the TBR evaluation as well as more advanced analytical based relay placement techniques to increase its effectiveness.
Chapter 5

An Analytical Model for

Performance Evaluation of TBR

The simulation results obtained in Chapters 3 and 4 represent the effectiveness of our proposed protocol at expanding the connectivity of DTNs and how it can outperform the exiting approaches in the literature. In order to validate our simulation studies as well as finding a general model to describe the behavior of TBR, in the rest of this thesis, we propose different analytical models. First of all, in this chapter, we analytically study TBR and provide mathematical equations to calculate the performance of the protocol based on the given settings. The proposed analytical model overcomes the burden of simulation based studies in terms of both time and resource complexities. This chapter is derived from our published work [71] on modeling performance evaluation in sparse mobile ad hoc networks. Secondly, in Chapter 6 we propose a generic analytical model to study the performance of relay placement techniques and we utilize this model to enhance the performance of heuristic placement techniques presented in Section 4.2.
5.1 Introduction

Based on physical characteristics of the network and application properties, it is desirable to choose the optimal routing protocol/network configuration to facilitate meeting application requirements. By measuring the performance metrics that represent the cost sustained in utilizing a protocol, one can select the optimal protocol/network configuration before deploying the network. There are some realistic simulation based approaches like [46] to calculate cost metrics in DTNs. However, this kind of approaches requires further computational resources and time to measure the cost in comparison to analytical approaches.

In this chapter, we present an analytical model in terms of network- and application-dependent parameters for TBR. In our analytical model, we show how we can model average throughput and average end-to-end delay using the latter protocol where mobile nodes communicate indirectly through small wireless read/writable RFID tags which act as mailboxes. We then verify our model through numerical computation. Further, we extend our protocol by proposing different buffer policies and then we present an analytical model for each buffer policy. To ease the understanding of our model, first we present a simplified protocol and then we show how we can extend it.

The remainder of the chapter is organized as follows: Section 5.2 presents a detailed description of our protocol, Section 5.3 describes our analytical model, Sections 5.5 and 5.4 include the simplified version of our protocol in a 1D network and in a real network respectively, Section 5.6 extends our model for placing multiple tags in the network, and Section 5.7 concludes the chapter.

5.2 TBR Assumptions

In this section, we describe the assumptions we made to analyze TBR. In chapter 5.3 we described the TBR protocol. In TBR messages are only transmitted between nodes and tags and vice versa, never between nodes and other nodes. Communication is only initiated by nodes since the tags are passive. To describe node-tag interactions for the analytical model, we define a write operation from
a node to a tag message buffer, a read operation from a tag message buffer to a destination node message buffer, and a hit event in time. When a node presents within distance $r$ of some tag at time $t$ then we say the node hits the tag at time $t$. In this way, given the paths for all nodes, we can consider the set of distinct hit events in time. We assume that nodes move obliviously to each other and to any other aspects of the system.

When a hit occurs, the connecting node examines the tag’s buffer to see if any of the messages are destined to itself, if so then the message is said to have reached its destination at the time of the hit. The node then writes as many messages as fit in the tag’s buffer.

The write operation happens based on the buffer policy which is described in Section 5.2.3. The read operation only happens when there is a message for the encountering node. Since the source node is not able to know whether a message has been delivered to a destination or not, it will continue to propagate its messages (if there is more than one tag in the network) hoping that the destination will pick up the messages later.

5.2.1 Performance Metrics

Besides the global evaluation metrics used in this thesis we further evaluate our protocol using the metric defined below in our analytical model.

5.2.1.1 Throughput

From an application’s perspective, throughput is a measure of how much useful (application-level) data is sent over a connection per unit time. We measure throughput as the ratio of the number of successfully delivered messages to the time:

$$\varrho = \frac{M_{\text{delivered}}}{t},$$

$t$ is referred as round in the next of this chapter and is described in Section 5.2.2.
5.2 TBR Assumptions

A low $\rho$ indicates that the tag buffer size and tag broadcasting range are not large enough to handle the rate of messages in comparison to average delay experienced by a message to get from the source node to the destination node.

5.2.2 Random Round Point Mobility Model

In order to propose a model for the performance of the network, we need a conceptual mobility model which is independent of any scenarios. This mobility model should be described in discrete events in order to observe the behavior of the network/placement techniques. Therefore, we propose a mobility model where time is considered as discrete rounds, numbered $1, 2, \ldots, n$. In each round, each node chooses a random destination in the region and will appear at that destination in the next round. Therefore, all nodes are assumed to have moved from their old position in the previous round to their new position in the current round. However, we do not consider node velocities, nor the spatial-temporal relationships such as the order of events that happen as the nodes move. We refer to this model as the Random Round Point (RRP) model. Although RRP is a simple mobility model used in this paper for our evaluation. As we will show in the next chapter, it is possible to derive more complicated models based on that such as Random Way Point (RWP).

5.2.3 Buffer Policy

We describe two different buffer policies for our protocol. In the first policy, messages are locked in the tag buffers until the destination node encounters the tags. In other words, a write operation only happens when the tag buffer is empty. This policy is best for lossless applications. In the second policy, older messages kept in the tag message buffer are overwritten with the new messages. This policy is more useful when data loss is tolerated and the system needs real-time performance.
5.3 Our Model

The most fundamental factor in our model is to calculate the probability of hitting a tag by nodes. This section shows how we can express the throughput and average end-to-end delay based on the latter factor.

5.3.1 System Model

In this section we present our assumptions to derive our model:

1. There are two independent mobile nodes, namely node A and B, which are communicating using our protocol.

2. Mobile nodes move within a bounded region according to the Random Round Point model and communicate through RFID tags placed at predefined/random locations with the broadcast range of $r$.

3. Tags have a single buffer which can keep a single message at a time.

4. Node A generates traffic for node B.

5. There is always at least one message from any source node (i.e. node A) at any instant of time.

Table 5.1 contains a symbol for each event and a description.

In the rest of this chapter we assume that the probability of hitting a tag by a node, $P_n[H]$, is:

$$P_n[H] = m,$$

(5.1)

(later we show how we can calculate $m$).

5.3.2 Buffer Policy-Keep Old Messages

Having only one tag in the network, if we use the policy of keeping old messages (KOM), a mobile node can only write to a tag when the tag is empty and the
node is in the range of the tag. Then the probability of writing a message in a tag message buffer at round \( n \), \( P_n[W] \), can be expressed as follows:

\[
P_n[W] = P_n[H]P_n[E] = mP_n[E],
\]

(5.2)

where \( P_n[E] \) is the probability of the tag buffer being empty.

Similarly, the probability of reading a message from the tag buffer can be expressed as follows:

\[
P_n[R] = P_n[H](1 - P_n[E]) = m(1 - P_n[E]).
\]

(5.3)

\( P_n[E] \) can be expressed as the previous state of the buffer. If the buffer was empty in the previous state, no write should happen (i.e. the source node should not hit the tag) and if it was full, a read should happen in the previous state to make it empty (i.e. the destination node should hit the tag). Therefore, we can express \( P_n[E] \) as follows:

\[
P_n[E] = P_{n-1}[E](1 - P_{n-1}[H]) + (1 - P_{n-1}[E])P_{n-1}[H].
\]

(5.4)
Using the Equation 5.1 we have:

\[ P_n[E] = P_{n-1}[E](1 - 2m) + m. \]  

(5.5)

The closed form of the Equation 5.5 is:

\[ P_n[E] = \frac{1}{2} + \frac{1}{2} (1 - 2m)^{n-1}. \]  

(5.6)

It is possible to approximate \( P_n[R] \). Since:

\[ 0 \leq c = \lim_{n \to \infty} P_n[R] \leq 1, \]

using Equations 5.3 and 5.6 we have:

\[ \lim_{n \to \infty} P_n[R] = \lim_{n \to \infty} \left( \frac{m}{2} - \frac{m(1 - 2m)^n}{2} \right). \]

Since \( 0 < m < 1 \) then \(-1 < 1 - 2m < 1\), and:

\[ -1 < \lim_{n \to \infty} (1 - 2m)^n < 1. \]  

(5.7)

Therefore:

\[ \lim_{n \to \infty} P_n[R] = \frac{m}{2}. \]  

(5.8)

Figure 5.1(a) shows the probability of reading a message by the destination node using KOM buffer policy. According to Figure 5.1(a), all the probabilities will converge to the value of \( \frac{m}{2} \) which verifies Equation 5.8.
5.3 Our Model

Figure 5.1: (a) The probability of reading from the tag converges to $m^2$. (b) The probability of reading from the tag converges to $m^2 - m$.

5.3.2.1 Performance Metrics

In this section, we show the analytical expressions to calculate the performance of the protocol in terms of two application level metrics: throughput and delay. These metrics are expressed in terms of parameters influenced by both the mobility model and communication protocol characteristics.

**Throughput** To show an analytical expression for throughput, we sum up all the probabilities of reading a message at each round and divide the result by the number of rounds, therefore we will have:

$$\varrho = \frac{1}{n+1} \sum_{i=0}^{n} P_i[R],$$

where $n$ is the number of rounds we consider for evaluation purposes. Based on the Equation 5.3, the closed form for the Equation 5.9 is:

$$\varrho = \frac{m}{2} + \frac{(1-2m)^n - 1}{4(n+1)};$$

(5.10)
5.3 Our Model

therefore, using Equations 5.10 and 5.7 we have:

\[
\lim_{n \to \infty} \theta = \frac{m}{2}.
\] (5.11)

**End-to-end Delay** The end-to-end delay between a source and a destination node, \( L_{\text{avg}} \) can be abstractly characterized as:

\[
L_{\text{avg}} = \sum_{n=0}^{\infty} n \prod_{i=1}^{n-1} (1 - P_i[R]) P_n[R],
\] (5.12)

where \((1 - P_i[R])\) is the probability that no read happens at round \(i\).

5.3.3 **Buffer Policy-Over Write Permitted**

In this section, we propose another buffer management policy called Over Write Permitted (OWP) where the source node can overwrite the tag message buffer with its newest messages (i.e. using the second buffer policy in Section 5.2.3). In this policy, some messages would be dropped when the source node overwrites the tag message buffer. Later, we show the analytical expressions for drop rate as well. Since the source node can always overwrite the tag and at each round it has a new message then:

\[
P_n[W] = P_n[H].
\]

Using Equation 5.1, we have:

\[
P_n[W] = m.
\] (5.13)

Since the writing process can happen in any round, ignoring the previous state of the buffer, if a write happens then the buffer becomes non-empty. Therefore, considering the previous state of the tag, if it is not empty, there should be a read to make it empty and in any case no write should happen to make sure that the
5.3 Our Model

buffer remains empty:

\[ P_n[E] = (P_{n-1}[E] + (1 - P_{n-1}[E])P_{n-1}[H])(1 - P_{n-1}[W]). \]

Using the same model as in Equations 5.3 and 5.13, \( P_n[E] \) can be expressed as follows:

\[ P_n[W] = m. \quad (5.14) \]

The closed form of the Equation 5.14 is:

\[ P_n[E] = \frac{m - 1 - (m - 1)^{2n-2}}{m - 2}. \quad (5.15) \]

Using 5.3 and 5.15 we can approximate the probability of read as follows:

\[ \lim_{n \to \infty} P_n[R] = \lim_{n \to \infty} \left( m - m \times \frac{m - 1 - (m - 1)^{2n-2}}{m - 2} \right). \]

Therefore:

\[ \lim_{n \to \infty} P_n[R] = \frac{m}{2 - m}. \quad (5.16) \]

Figure 5.1(b) shows the probability of reading a message by the destination node using the latter buffer policy. According to Figure 5.1(b), all the probabilities will converge to the value of \( \frac{m}{2 - m} \) which verifies Equation 5.16.

5.3.3.1 Throughput

Since source nodes are allowed to overwrite the tag message buffer, the probability of reading by the destination node will increase and the throughput increases as a result. However, some messages are dropped in tags due to the overwrite process and the number of dropped messages can be calculated as the number of reads
5.3 Our Model

in \( n \) rounds or alternatively:

\[
\text{No. of dropped messages} = \varrho(n + 1).
\]

Using the same model as Equation 5.9, we show the closed form of the throughput:

\[
\varrho = \frac{m}{2 - m}.
\]  \( (5.17) \)

5.3.3.2 End-to-end Delay

Since the messages can be overwritten by the source node in any round, average end-to-end delay can be expressed as follows:

\[
L_{\text{avg}} = \sum_{n=1}^{\infty} \sum_{k=1}^{n} kP_k[W'] \prod_{i=1}^{n-1} (1 - P_i[R]) P_n[R],
\]  \( (5.18) \)

where \( P_k[W'] \) is the probability that the latest write was \( k \) rounds ago. Since \( P_k[W'] = m(1 - m)^{k-1} \), then

\[
\sum_{k=1}^{n} kP_k[W'] = \frac{1 - (1 + nm)(1 - m)^n}{m}.
\]  \( (5.19) \)

Based on Equations 5.18 and 5.19 we express the average end-to-end delay as:

\[
L_{\text{avg}} = \sum_{n=1}^{\infty} \frac{1 - (1 + nm)(1 - m)^n}{m} \prod_{i=1}^{n-1} (1 - P_i[R]) P_n[R].
\]  \( (5.20) \)
5.4 1D Network

As a simple example of our model, we place one tag at the beginning of a line with the length of $l$ and we suppose mobile nodes are moving on that line. Section 5.4.2 shows how we can extend our model to the case when the tag is placed randomly on the line. Additionally, in Section 5.5 we will extend the model to allow nodes to move randomly in a 2D network.

Figure 5.2 shows the network environment. Given the range of the tag as $r$ and the length of the region as $l$, then the probability of hitting the tag by any mobile node at round $n$, $P_n[H]$, can be expressed as follows:

$$P_n[H] = \frac{r}{l}. \quad (5.21)$$

This is the simplest case to calculate $P_n[H]$. Replacing $m$ by $\frac{r}{l}$ in Section 5.3.2.1 we can calculate cost metrics introduced in this chapter based on the buffer policy we are using.

5.4.1 Simulation

We verify our analytical model with numerical simulation. In the numerical simulation, we place one tag in the middle of the region and change its communicating range from 0 to size that covers whole the region and then we measure the
throughput and average end-to-end delay. The numerical simulation is performed for both KOM and OWP buffer policies.

### 5.4.1.1 Buffer Policy-Keep Old Messages

Figure 5.3 shows the comparison between the numerical simulation results and the results obtained through the Equations 5.10 and 5.12 for performance metrics. According to Figure 5.3 (a), throughput is increasing when tag range increases. The maximum throughput is 0.5 using this policy. Figure 5.3 (b) shows the delay is decreasing when tag range increases. The minimum average end-to-end delay is 2 rounds.

Figure 5.3: Performance metrics calculated based on the simulation and analytical model in 1D network.
5.4 1D Network

5.4.1.2 Buffer Policy-Over Write Permitted

In this section, we present the simulation results for the Section 5.3.3. Figures 5.3(c) and 5.3(d) show both the results from simulation and the analytical model. The maximum throughput is 1 using this policy and the minimum average end-to-end delay is 1 round. The average end-to-end delay is much less than the other buffer policy.

5.4.2 Extended Model-Random Tag Placement

In this section, we extend our model by placing the tag randomly on the line. Figure 5.4 shows the new placement of the tag. Obviously, \( a \) is randomly chosen.

Figure 5.4: The network has only one dimension with the length of \( l \) and the tag is placed at the beginning of the network which has a broadcasting range of \( r \), \( a \) is randomly chosen to place the tag at the location \( a + r \).

\[ P_n[H] \] is dependent on the previous location of the node. If the node was in the left side of the tag and it presents to the right or inside the tag in the next round, it has to hit the tag. Similarly for the case that it was on the right and targeting the left or inside the tag. Also, if it stays inside the tag in the next round, we say it still hits the tag. Therefore, \( P_n[H] \) can be expressed as:

\[
P_n[H] = P_{n-1}[L]P_n[M] + P_{n-1}[L]P_n[G] + P_{n-1}[M]P_n[M] + P_{n-1}[G]P_n[M] + P_{n-1}[G]P_n[L], \tag{5.22}
\]
5.5 2D Networks

where \( \mathbb{P}_n[L] \), \( \mathbb{P}_n[M] \), and \( \mathbb{P}_n[G] \) are the probability of a node being in the left outside of the tag, inside the range of the tag, and the right outside of the tag respectively.

If we extend the Equation \(5.22\) then we have:

\[
\mathbb{P}_n[H] = \frac{a}{l} \left( \frac{2r}{l} + \frac{a}{l} - \frac{2r}{l} - a \right) + \frac{2r}{l} \left( \frac{2r}{l} - \frac{a}{l} - 2r \right) + \frac{a}{l} \left( l - 2r - a \right) - \frac{a}{l} l.
\]

(5.23)

Replacing \( m \) by the new value of \( \mathbb{P}_n[H] \) in Section \(5.3.2.1\) we can calculate cost metrics introduced in this chapter based on the buffer policy we are used.

5.5 2D Networks

In this section, we extend our model for a 2D network. We show an equation to calculate the probability of hitting a tag. The throughput and average end-to-end delay can be obtained from the hit probability using a similar approach as we had in Section \(5.4\).

Given a node at point \((x, y)\) in the region, we can compute the probability, \( P_{hit}(x, y, i, j, r) \), that it hits a tag centered at \((i, j)\) with range \( r \) during its next round. All of the destination points in the region are equally likely. Some of those destinations will give rise to a hit; in Figure \(5.5\) (a), these points are in the shaded region. Let \( \delta = \sqrt{(x - i)^2 + (y - j)^2} \) be the distance from the node to the tag center. If \( \delta \leq r \) then a hit occurs with probability 1. Otherwise, a hit occurs with a probability given by the fraction of total region that gives rise to a hit.

We consider the node’s movement given by a direction angle, \( 0 \leq \phi < 2\pi \).

Let \( 0 \leq \alpha < 2\pi \) be the direction from the node to the center of the tag and the angle of integration is \([\alpha - \theta, \alpha + \theta]\), where

\[
\theta = \tan^{-1} \left( \frac{r}{\sqrt{\delta^2 - r^2}} \right).
\]
Figure 5.5: (a) Node-tag angles. (b) Region, different cases are shown, the closest points would be selected to calculate the distance to the perimeter of the region.

We found the area of the dashed region in Figure 5.5 (a) to be:

\[
SD = S_{(x,y)(x,Y)(X,y)} - S_{(x,y)(X,y)C} - S_{(x,y)(x,Y)D} - 2S_{PM(x,y)} + 2S_{MP}.
\]

Appendix A gives the details.

The total area of the network is

\[
S_{total} = (Y - y_0)(X - x_0).
\]

From the above we find:

\[
P_{hit} = \begin{cases} 
\frac{SD}{S_{total}} & \delta > r, \\
1 & \text{otherwise.}
\end{cases}
\]
5.5 2D Networks

Figure 5.6 shows the hit probability of a tag placed at location (3, 4) in a $10 \times 10m^2$ network. The closer points to the circle indicate a higher chance of hitting the tag.

![Figure 5.6: The probability of hitting a tag placed at (3, 4)](image)

Having the probability of hitting the tag, $P_{hit}$, it is possible to calculate $\varrho$ and $L_{avg}$ with a similar approach as we had in Sections 5.3.2 and 5.3.3.

5.5.1 Simulation

In this section we will verify our analytical model through simulation.
5.5.1.1 Buffer Policy-Keep Old Messages

To achieve a reasonably accurate comparison, we place the nodes and the tag randomly in 200 different locations in a $1000m \times 1000m$ network and then we select 5000 different destinations for mobile nodes on the next round to compare the simulation and analytical models.

Figure 5.7 (a) shows the comparison between the simulation results and the results obtained through the Equation 5.10 for throughput in 2D networks. Figure 5.7 (b) shows the comparison between the simulation results and the results obtained through the Equation 5.12 for average end-to-end delay in 2D networks. As we expect, the simulation results verify the model.

![Graphs showing throughput and delay vs. tag range for both simulation and analytical models with and without keeping old messages.](image)

Figure 5.7: Performance Evaluation based on the simulation and analytical model in a 2D Network.
5.6 Multiple Tags

5.5.1.2 Buffer Policy-Over Write Permitted

In this section, we present the simulation results for the Section 5.3.3 in a 2D network. To achieve a reasonably accurate comparison, we use the same setting as Section 5.5.1.1 to compare the simulation and analytical models. Figures 5.7 (c) and 5.7 (d) show both the results of simulation and analytical models. As we expect, the simulation results verify the model.

5.6 Multiple Tags

In this section, we extend our model by placing multiple tags in the network. Suppose we place \( k \) tags namely \( \text{tag}_1, \text{tag}_2, \ldots, \text{and tag}_k \) in the network. If we define the probability of hitting at least one tag in the entire network as:

\[
P_n[H] = 1 - \prod_{i=1}^{k} (1 - P_n[H_i]),
\]

we should make sure that all tags can be independently hit by nodes. As an example, we place 2 tags in a network sized 1000 m \( \times \) 1000 m. Figure 5.8 (a) shows the probability of hitting at least one of the tags by a mobile node. According to Figure 5.8 (a), when tags range gets bigger (more than 100 m), they are overlapping to cover the area and they cannot independently be hit by nodes as a result (i.e. our model is not correct for the ranges bigger than 100 m in this example).

5.6.1 Maximum Number of Tags

The maximum number of tags that could be placed in a network in such a way that can be hit approximately independently by a mobile node in each round is dependent on the broadcasting range size of the tags and the network size. To find the latter number, we consider the case of placing tags on a regular grid. The range of tags is the same for all tags and equals 0.05 meter. According to Figure 5.8 (b), the probability of hitting multiple tags is very small if the tag
5.6 Multiple Tags

Figure 5.8: (a) The probability of hitting at least one tag by a mobile node.

Spacing is greater than or equal to 100m while when tag spacing is 50 meters (i.e. we place 400 tags) in a 1000m × 1000m network, this probability is around 66%. We found the maximum number of 50 tags to be placed in the network (with the above settings) that gives a small probability of overlapping which is less than 2%.

5.6.2 Performance Metrics

Assuming that tags are placed in such a way that they can be hit independently in each round by nodes, we can define the probability of a read in the entire network as:

\[ P_n[R] = 1 - P_n[\tilde{R}] , \]

where \( P_n[\tilde{R}] \) is the probability of no read happening at round \( n \), as follows:

\[ P_n[\tilde{R}] = \prod_{i=1}^{k} (1 - P_n[R_i]) . \]

Calculating the probability of each \( P_n[R_i] \) using the equations introduced in Section 5.3.2 \[ 5.3.3 \] and 5.5 we can measure the throughput and average end-to-
end delay. Therefore, for those applications such as search and rescue in radio silent mode where the traffic load is not very high, we can place a limited number of nodes (e.g. 50 in the example in Section 5.6.1) and use our model to measure the performance metrics.

### 5.6.3 Simulation

#### 5.6.3.1 Buffer Policy-Keep Old Messages

In this section, we compare the simulation and analytical models for the case of placing multiple tags in the network using keeping old messages buffer policy. Figures 5.9 (a) and 5.9 (b) show this comparison. Our model is fine when the tag range is up to 15m.

#### 5.6.3.2 Buffer Policy-Over Write Permitted

In this section, we compare the simulation and analytical models for the case of placing multiple tags in a network using over write permitted buffer policy. Figure 5.9 (c) and 5.9 (d) show this comparison. Our model is fine when the tag range is up to 15m.

### 5.7 Conclusion

In this chapter, we proposed an analytical model to express the throughput and average end-to-end delay of a DTN where mobile nodes are communicating through our proposed communication protocol. This model can be used to provide expressive performance metrics in order to facilitate finding the specification of operating environment in network deployment process. Our model removes the burden of excessive simulation time and resources to find the desired results. Specifically, we can use the model for a real time adoption facing changing
### 5.7 Conclusion

Figure 5.9: Performance evaluation based on the simulation and analytical model using multiple tags placed in a 2D network.
environmental conditions. Our analytical model has been verified by the simulation. In the next chapter, we present a generic analytical model to study the performance of relay placement techniques and we utilize this model to enhance the performance of heuristic placement techniques presented in Section 4.2.
Chapter 6

Analytical Approach for Passive Stationary Relay Point Placement in DTNs

In previous chapter we presented an analytical model to evaluate the performance of our proposed protocol, TBR, and validate our simulation results. In this chapter, using the proposed analytical model concepts, we present a generic analytical model to study the performance of relay placement techniques. We utilize this model to enhance the performance of heuristic placement techniques presented in Section 4.2. This chapter is derived from our paper [73] on placement of passive stationary relay nodes in DTNs.
6.1 Introduction

As mentioned earlier, TBR works on the basis of augmenting DTNs with easily deployable stationary relay nodes, making an unconnected infrastructure to facilitate the message delivery by increasing forwarding opportunities. Relay nodes are capable of downloading, storing, and forwarding the messages from/to the mobile nodes. In Chapter 4 we showed that a better relay node placement can significantly improves the performance of TBR. The simulation results showed that the improvement can reach to a factor of 250% in both packet delivery and message latency. Zhao et al. [96] considered the relay placement as NP-hard and it was later proved by Farahmand et. al [22]. The proposed approaches for relay placement in the literature are based on heuristics to find an approximated optimal solution. Zhao et al. [96] proposed the first relay placement in DTNs based on a greedy algorithm. We proposed different heuristics for relay placement in Section 4.2. These simulation based approaches are highly sensitive to the input scenario as well as the initial setup. Moreover, finding the impact of different parameter values on the system performance is a time-consuming process and also statistical fitting of the simulation results is not a trivial task. In order to overcome these issues, in this chapter we propose an analytical model to describe the performance of the network based on the position, communication range, and buffer size of the relays. This model is a function of the mobility pattern of the nodes; therefore, we provide a case study to evaluate the effectiveness of our model based on the random waypoint mobility model. In order to use the proposed model for placing the relays effectively, we utilize two heuristic approaches. The first approach is based on optimization of the network performance using simulated annealing and the second one relies on a greedy approach to find the best location for each relay one at a time. Our simulation results show that our approaches outperform the simulation based approaches in terms of data delivery performance.

The remainder of this chapter is organized as follows: Section 6.2 details the assumptions about relay nodes, presents our network model, defines our performance objectives, and describes the problem statement of this chapter. Section 6.3 proposes the analytical model used to describe the trade-offs between the
6.2 Problem Statement

The objective of this chapter is to analytically study the placement techniques to find a model which can be used to optimize the heuristics for placement techniques introduced in Section 4.2. The state-of-the-art placement techniques are based on the simulation. Zhao et al.\[96\] propose an iterative greedy algorithm to place the relays one at a time. In their approach, the relays are placed in the position that achieves the most capacity in terms of data rate. However, in their approach when a relay is deployed, it is not possible to change its location; therefore, their approach is not optimal as they cannot utilize the possible effects of relay spacing on the performance of the network. Placing relays very close to each other can lead to forward redundant messages. We also proposed different relay placement techniques in Section 4.2 namely tag pruning, probability based tag distribution, and genetic algorithm based optimization. Two latter approaches consider the placement of relays all together; therefore, they take the effect of distance between each pair of nodes into account. All mentioned approaches above are based on simulation; however, these simulation based approaches are highly sensitive to the given scenario. Moreover, simulation-based relay placement techniques require significant computational effort to search the space of possible locations to find an optimal solution.

In the design phase of the network, we may expect to answer the following questions: How many relays are needed to achieve an expected performance? How big should the buffer size of the relays be given an expected performance? How big
should the range of relay communication be to satisfy an expected performance? Using simulation-based approaches, a designer needs to do a massive evaluations based on many different scenarios to find out the answers. Therefore, we are motivated to find a way to answer these questions more quickly/easily using a general analytical model.

The network model used in this chapter is similar to Section 3.1. We assume that all relays have sufficient energy for the entire network life. This assumption is realistic by using passive RFID relays as they are powered by the reader. Therefore, our objectives are to propose an analytical model for relay placement techniques regardless to the nature of relays. Table 6.1 shows a basic comparison between representatives of active and passive relay nodes, i.e., throwboxes and passive RFID tags respectively.

### 6.3 Modeling Relay Performance

In this section, we propose our analytical model for the relay placement techniques in DTNs. The model is first developed with a single relay and then we extend it to the case of multiple relays. Analytical equations are proposed to find the overall performance of the network based on the position, communication range, and buffer size of the relays.

#### 6.3.1 An Analytical Model for Relay Placement

In this section, our goal is to derive analytical equations for message delivery performance based on the relays’ buffer size, communication range, and position. We consider a DTN consisting of $N_{\text{relay}}$ relay nodes and $N_{\text{node}}$ mobile nodes and derive analytical equations for the number of dropped and delivered messages by the relay. As the delivery performance is dependent on the mobility model, we propose a conceptual mobility model and later we show how we can extend this conceptual model to real mobility models.
### Table 6.1: Stationary Relay Points approaches.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Passive RFID tag</td>
<td>$0.07 - $0.20</td>
<td>powered by reader</td>
<td>up to 35m</td>
<td>from $10^{-8}m^3$ to $4.8^{-4}m^3$</td>
<td>unable</td>
</tr>
<tr>
<td>Throwbox</td>
<td>$100 - $300</td>
<td>stand-alone</td>
<td>up to 250m</td>
<td>from $1.12^{-4}m^3$</td>
<td>capable</td>
</tr>
</tbody>
</table>

---


*b* Zhao *et al.* [96] introduce Intel Stargate as a device which can meet throwbox requirements. Stargate specification is available at [http://www.willow.co.uk/Stargate_Datasheet.pdf](http://www.willow.co.uk/Stargate_Datasheet.pdf)
6.3 Modeling Relay Performance

6.3.1.1 Conceptual Mobility Model

We are using the same conceptual mobility model as the previous chapter introduced in Section 5.2.2 to analyze the placement techniques and evaluate the performance of the proposed protocol.

6.3.1.2 Expressions for DTNs delivery performance

The overall performance of the network in terms of message delivery can be described by the average message delivery of each relay per unit of time. Knowing the average hit time of a relay by a mobile node, we can calculate the message delivery rate of a relay per unit of time. In the other words, if we increase the number of delivered messages by each relay and decrease the contact time between hitting a relay by mobile nodes, we can increase the overall network performance. Therefore, the overall delivery performance of the DTN can be described as follows:

\[ \rho = \sum_{i=1}^{n} \frac{\text{Number of Delivered Messages}_i}{\text{Hit time}_i}, \]

where \( n \) is the number of relays in the network. Next we derive analytical equations to express Hit time and Number of Delivered Messages for each relay.

All the messages in our model have a limited life time in the network and they can be described as the following observations:

1. Once a message, let’s say \( x \), has left a buffer, it never comes back to the same buffer because we are using FIFO approach to discard the messages on demand.

2. After generating at least \( B_{node} \) messages after message \( x \) being generated by a source node, \( x \) will be discarded from the node buffer.

3. After a time \( t1 \), no MN would buffer message \( x \) as they have limited buffers which gets exhausted after a while by new messages.
4. After a time \( t_2 \), no MN would exchange message \( x \) with the relay even if it is available in their/relay’s buffer because the maximum number of messages that an MN can exchange with a relay equals to the smaller buffer size between the MN and the relay. Therefore, if \( x \) gets pushed back far in the larger buffer, it has no chance to be exchanged with the smaller buffer anymore.

The above observations motivated us to model the performance of the network based on the position of the relays, their communication range, and relay buffer size.

From a relay’s point of view, mobile nodes hit the relay and exchange their buffer with its buffer at discrete times. At each hit event there are newer messages from nodes to be written in the relay buffer and also there are messages in the relay buffer which are destined to the mobile nodes. In order to let newer messages to be written to the relay buffer, the relay may have to evict some messages. In this model, we consider FIFO in the case of evicting. Therefore, knowing the number of dropped and delivered messages per round by a relay, we can model the relay performance on average in a steady state.

To find the number of delivered/dropped messages from a relay point of view in each round, we use the hit event and later we show how to calculate the average hit time; therefore, by dividing the number of delivered/dropped messages during the hit time by the hit time, we can calculate the delivery rate and drop rate of a relay in each round. To do this, we assume that mobile nodes do not hit the relay again unless a number of other mobile nodes hit the relay changing all messages in the relay buffer; therefore, on average \( \frac{1}{N} \) of the messages in the relay buffer is destined to the encountering node in the steady state. This assumption is realistic if \( B_{\text{relay}} \) is not very big. Therefore, the number of delivered messages is as follows:

\[
\text{Number of Delivered Messages on Hit} = \frac{1}{N} B_{\text{relay}}. \quad (6.1)
\]

Knowing the number of messages to be written to the relay on each hit, the
number of dropped messages is as follows:

\[ \text{Number of Dropped Messages on Hit} = M_w - \frac{1}{N} B_{\text{relay}}, \]  

(6.2)

where \( M_w \) is the number of messages to be written to the relay buffer. Equations related to \( M_w \) are reported in Table 6.2. Later, we show how we derive these equations.

<table>
<thead>
<tr>
<th>( B_{\text{relay}}(\lambda \text{Hit time}) )</th>
<th>( M_w(\lambda \text{Hit time}) )</th>
<th>assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq 1 )</td>
<td>( B_{\text{relay}}/\lambda \text{Hit time} )</td>
<td>the same mobile node cannot hit the relay for 2 sequential rounds</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{3}{2} )</td>
<td>the same mobile node cannot hit the relay for 2 sequential rounds</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>the same mobile node cannot hit the relay for 2 sequential rounds</td>
</tr>
<tr>
<td>4</td>
<td>( \frac{5}{2} )</td>
<td>the same mobile node cannot hit the relay for 3 sequential rounds</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( k )</td>
<td>( \frac{k+1}{2} )</td>
<td>the same mobile node cannot hit the relay more than once for ( \lceil \frac{k+1}{2} \rceil ) seq. rounds</td>
</tr>
</tbody>
</table>

To find \( M_w \) on each hit, we need to find the average time between sequential hits and the average message generation rate by mobile nodes. In an area of \( X \times Y \), the probability of hitting a relay by a node using RRP mobility model in the next round is defined as

\[ \mathbb{P}_n[H] = m. \]  

(6.3)

The details are available in the previous chapter. Note that this model is proposed only for RRP which is extensible to RWP as we will show later in this chapter; however, for other mobility models, one may need to model this hit probability to find the value of \( m \).

This probability is a constant value, i.e., \( m \), based on the position and communication range of the relay. Since in each round, the relay might be independently
hit or not hit by any mobile node, the hitting probability of a relay by any node is described by a binomial experiment and has a geometric distribution. Therefore, the average number of rounds needed to hit the relay can be described as follows:

\[
\text{Hit time} = \frac{1}{\mathbb{P}[\text{Hit}]} = \frac{1}{1 - \mathbb{P}[\text{NoHit}]).
\]  

(6.4)

On the other hand, the \(\mathbb{P}[\text{NoHit}]\) is the product of the probability that each individual mobile node does not hit the relay and can be shown as follows:

\[
\mathbb{P}[\text{NoHit}] = \prod_{i=1}^{N} (1 - \mathbb{P}_i[H]).
\]

On average, after the \(\text{Hit time}\), each mobile node would generate \(\lambda \times \text{Hit time}\) messages. If we define the unit size of any buffer as \(\lambda \times \text{Hit time}\), then the number of messages to be written by any mobile node to the relays has a direct impact with the size of the relay’s buffer. If we assume that the buffer size of nodes is bigger than the buffer size of relays, then Table 6.2 shows the number of messages to be written to the relay versus the relay buffer size on each hit.

If \(B_{\text{relay}} \leq \lambda \text{Hit time}\), then any node would generate a number of new messages more than or equal to \(B_{\text{relay}}\) during a round and in the merging operation of buffers, the mobile node would overwrite the whole buffer of the relay. We prove the case of \(B_{\text{relay}} = 2\lambda \text{Hit time}\) and then we extend it for the other cases. If \(B_{\text{relay}}\) has 2 units of size, i.e., \(\lambda \text{Hit time}\), then after \(\text{Hit time}\), on average the encountering node has one unit messages newer than the messages in the relay buffer. Therefore, when they exchange their buffers, the relay has to push back one unit of messages to allow the newer messages to be written. Also, since we assume the encountering node is a different node than the previous node that visited the relay, then on average half of the messages in a unit has to be discarded to allow the exchanging process to write the newer messages. This process is shown in Figure 6.1. Figure 6.2 shows an extended scenario for the case of having a bigger relay buffer. According to Figure 6.2, an MN overwrites \(2/3\) of the relay
6.3 Modeling Relay Performance

buffer. We found that on average a number of messages equal to \( \frac{k+1}{2} \lambda \text{Hit time} \) needs to be discarded to allow the new messages to be written in a relay with \( B_{\text{relay}} = k \lambda \text{Hit time} \) and with the assumption that no identical node can hit the relay more than once for \( \frac{k+1}{2} \) sequential rounds.

![Figure 6.1: Merge operation in the case of having a 2 slot relay buffer.](image1)

\( n \) is the order of messages in time, i.e., all messages tagged with \( n \) in a slot were generated in the last round, all messages tagged with \( n-1 \) in a slot were generated two rounds ago, etc.

![Figure 6.2: Merge operation in the case of having a 3 slot relay buffer.](image2)

From Equation 6.2 and Table 6.2, the number of dropped messages per round could be derived based on the \( B_{\text{relay}} \), \( \lambda \), and \( \text{Hit time} \). Supposing \( \lambda \) is a fixed value, then the performance of each relay is only based on its \( B_{\text{relay}} \) and \( \text{Hit time} \). \( \text{Hit time} \) is also dependent on the relay communication range and its position.
6.4 Analytical based Relay Placement Approaches

6.3.1.3 Objective Function

In general if we minimize the $\prod_{i=1}^{n} \bar{Hit\ time}_i$ where $n$ is the number of relays, the probability of hitting the relays by nodes is increasing and the delivery performance based on our model is increasing as a result. If we only increase this probability, relay placement techniques may place the relays very close to each other or even on top of each other which can result in storing redundant messages in the same region leading to poor performance. Therefore, we consider another factor to keep the relays far away enough from each other to make sure that relays are not overlapping. Later, we will show how we can control the effect of each parameter used in our objective function. Based on the heuristics discussed above we propose the following objective function:

$$\text{loss} = 1 - \left( \prod_{i=1}^{n} \mathbb{P}_i[H] \right)^{\alpha} \bar{D}^{\beta},$$

(6.5)

where $n$ is the number of relays, $\mathbb{P}_i[H]$ is the probability of hitting relay $i$ by any node during a round, $\bar{D}$ is the average distance between each pair of relays, and $\alpha$ and $\beta$ are used to control the importance of their factors. The best value of $\alpha$ and $\beta$ to optimize the delivery performance of the network is based on the mobility model. In the next section, we provide a case study to find the best values for $\alpha$ and $\beta$ as well as evaluating the effectiveness of our model.

In order to optimize this objective function, we need a heuristic approach to search the problem space. Later, we propose different placement techniques to optimize relay placement.

6.4 Analytical based Relay Placement Approaches

Simulation based approaches are highly sensitive to the given scenario. Moreover, they dictate computational and time complexities to search the space of possible locations to find an optimal solution. In this section, in order to overcome these issues, we use the analytical equations we proposed in Section 6.3 to develop
two different placement techniques to find the best relay deployment. The first technique is based on optimization of the network performance using simulated annealing and the second one relies on a greedy approach to find the best location for each relay one at a time.

### 6.4.1 Simulated Annealing based Optimization

We use Simulated Annealing (SA) technique which is utilizing our proposed analytical model as an objective function. SA is a generic probabilistic meta-heuristic mainly used for global optimization problems. Given an objective function, it can find the global optimum value for that function in a large search space. It is mainly based on finding an acceptably good solution in a fixed amount of time, rather than the best possible solution.

Relays are initialized in a regular grid with a distance of $d$ from each other. Later, SA minimizes the proposed objective function according to Equation 6.5. In order to evaluate the effectiveness of our approach we use our case study to model random waypoint. The results are represented in Section 6.5.3.1. Further, we compare the proposed approach with the simulation based approaches in the literature in Section 6.5.3.2.

### 6.4.2 Greedy Placement

Greedy relay node placement for DTNs has been proposed in [96]. We use a similar algorithm using our analytical model equations. In this approach we place the relays one by one as the first relay would be placed in the position that minimizes the value of the Equation 6.5. The next relay would be placed in the same manner until all the relays are placed. Table 6.3 shows the algorithm giving a set of locations, i.e., $h$.

The experimental results are represented in Section 6.5.3.1. Further, we compare the proposed approach with the simulation based approaches in the literature in Section 6.5.3.2.
6.5 A Case Study

Table 6.3: Greedy algorithm.

<table>
<thead>
<tr>
<th>Line</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>$x_i = 0, i \in L$;</td>
</tr>
<tr>
<td>2:</td>
<td>for $t = 1$ to $N_{relay}$ do</td>
</tr>
<tr>
<td>3:</td>
<td>for all $i \in L$ and $x_i = 0$ do</td>
</tr>
<tr>
<td>4:</td>
<td>Compute the value of loss function when a relay is placed at $i$;</td>
</tr>
<tr>
<td>5:</td>
<td>end for</td>
</tr>
<tr>
<td>6:</td>
<td>Let $h$ be the location that achieves the min. value of loss function.;</td>
</tr>
<tr>
<td>7:</td>
<td>$x_h = 1$;</td>
</tr>
<tr>
<td>8:</td>
<td>end for</td>
</tr>
</tbody>
</table>

6.4.2.1 Complexity of the Proposed Techniques

Greedy algorithm searches on $L$ possible locations to place $N_{relay}$ relay nodes. On each iteration it calculates the objective function which needs finding $L$ different probabilities. Consequently, its complexity is $O(L^2 \times N_{relay})$. On the other hand, simulated annealing based optimization calculates the objective function on each iteration for each relay nodes, which needs $L \times N_{relay}$ operations in total for each iteration. As simulated annealing terminates its operation after $K_{max}$ rounds even if it cannot find a good enough solution, at the worst case our simulated annealing based optimization has the complexity equals $O(L \times K_{max} \times N_{relay})$. In general finding the optimal solution to this problem has been shown NP hard in other work, and our solutions are heuristics.

6.5 A Case Study

In this section, we show how we can customize our analytical model when mobile nodes move according to RWP mobility model. We leave deriving other mobility models from our conceptual mobility model to future work. In this section, first we show how we can model RWP based on RRP and then we demonstrate some experimental results to show the relation between the relay position and its performance in terms of delivery. Finally, the simulation results are provided to study the effectiveness of the proposed analytical model and we compare the
proposed approaches which are developed based on our analytical model with the simulation based approaches in the literature.

### 6.5.1 Extending RRP to Random Waypoint Model

In this section, we show how we can approximate RWP with RRP. In RWP model, every node chooses a random destination called waypoint in the simulation area and moves toward it in a straight line with a constant speed which is selected randomly from $[S_{\text{min}}, S_{\text{max}}]$. After approaching the destination, mobile nodes stay there for a certain period of time, i.e., a pause time, and then they repeat the procedure.

In order to extend RRP to RWP we calculate the average length of a random line (i.e., the straight line from the source location to the destination) in a square area of a $lm^2$ which is $0.521405dm$. The average length of a round in seconds then can be obtained as follows:

$$AvgRoundtime = 0.521405 \times \frac{l}{\text{avgSpeed}}.$$ 

Knowing the average round time, we can approximate RWP (or even other mobility models) using RRP. By multiplying the number of rounds by the average round time, the model can be described based on seconds. Therefore, on average we can approximate how long it takes for a node to travel from a point to another point.

### 6.5.2 Relay Position vs. its Performance

In RWP model, the frequency of hitting the relays placed in the middle of the area is higher than the ones placed around the corner. From the equations proposed in Section 6.3, if the buffer size of any relay in the network has the same number of slots in terms of $\lambda Hit\ time$, then the performance of the relay is independent of its position. In other words, if we place a relay in the middle of the area with a RWP
mobility model, the probability of hitting the relay is higher than a relay placed at the corner of the area. Therefore, Hit time of the relay in the middle is much less than the relay placed at the corner. If we select the same number of slots for each relay buffer based on their Hit time (note that the size of slots for different relays are different), in order to get the same performance from relays in terms of message delivery rate, the relay placed in the middle needs a smaller buffer size than the other relays because the size of its slots is smaller than the other ones. Figure 6.3 shows the message delivery rate of each relay versus the ID of the relay which verifies this observation through the simulation. The simulation setting used is as follows: $N_{\text{relay}} = 49$, $N_{\text{node}} = 10$, $X/Y_{\text{min/max}} = [0 \ 1000]$, $d = 150$, $r = 5$, $S_{\text{min/max}} = [15 \ 20]$, and $\lambda = 5$ (i.e., 0.5 msg/sec on average for each node). Mobile nodes move based on the RWP mobility model.

![Message Delivery Rate vs. Tag Number](image)

Figure 6.3: Message Delivery Rate VS. Relay ID.
6.5.3 Experiments

In this section, we propose our experimental results using RWP to evaluate the performance of the proposed techniques. Moreover, in this section, we compare the performance of our techniques with the simulation based techniques in the literature.

6.5.3.1 Performance Evaluation

In this section, we show the experimental results to evaluate the performance of our analytical based techniques. The results for the Simulated Annealing based optimization are presented followed by the results for the Greedy based algorithm would be presented. We have implemented our methods using MATLAB®. The simulation setting used is as follows: $N_{\text{relay}} = 9$, $N_{\text{node}} = 10$, $X/Y_{\text{min}/\text{max}} = [0 \ 1000]$, $d = 1000/3$, $B_{\text{relay}} = 500$, $B_{\text{node}} = 1000$, $r = 5$, $S_{\text{min}/\text{max}} = [15 \ 20]$, $\lambda = 5$ (i.e., 0.5 msg/sec on average for each node), and $t = 10000$. Mobile nodes move based on the RWP mobility model.

Simulated Annealing based Optimization Figure 6.4 shows the results for delivery performance of the network in which relays are placed with the solution SA found considering that mobile nodes move according to RWP.

According to the Figure 6.4 the best solution on average is when the $\alpha = 0.01$ and $\beta = 0.1$.

Greedy Placement In this part, we investigate the performance of Greedy based algorithm. Figure 6.5 shows the results for delivery performance of the network in which relays are placed with the solution Greedy algorithm found while mobile nodes move according to RWP. The objective function to maximize the overall performance is based on minimizing the value of Equation 6.5.

According to the Figure 6.5 the best solution on average is when $\alpha = 0.01$ and $\beta = 0.1$. 

137
6.5 A Case Study

Figure 6.4: Delivery performance of different solutions found by SA respecting to the different importance of hit probability and distance factors. Mobile nodes move according to RWP model.

Figure 6.5: Delivery performance of different solutions found by Greedy respecting to the different importance of hit probability and distance factors. Mobile nodes move according to RWP model.
6.5.3.2 Comparison to Existing Simulation-based Approaches

In this section, we compare the results from simulation based approaches introduced in Section 4.2 namely Regular Grid, Relay Pruning, probability based relay distribution, and Genetic Algorithm based optimization (GA) as well as the placement technique proposed by Zhao et al. [96] as we refer to it as throwbox placement with our analytical based approaches proposed in Section 6.4. We used the following settings to evaluate the efficiency of each placement technique: the area size is $500 \times 500$, the average speed of mobile nodes is 20 (i.e., $S_{\text{min}} = 10$, $S_{\text{max}} = 30$), $\lambda = 0.5$, $N_{\text{relays}} = 16$, $B_{\text{relay}} = 100$, $B_{\text{node}} = 1000$, $r = 5$, and $N_{\text{node}} = 10$. Simulation time is 10000 seconds and each result is based on 10 different runs to reduce the existing variation in the simulation. Table 6.4 shows these results.

<table>
<thead>
<tr>
<th>Placement Technique</th>
<th>$\rho$</th>
<th>$L_{\text{avg}}$</th>
<th>$L_{\text{avg}}/\rho$</th>
<th>$N_{\text{relay}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reg. Grid</td>
<td>10.3</td>
<td>202.1</td>
<td>19.45</td>
<td>16</td>
</tr>
<tr>
<td>Tag Pruning</td>
<td>10.1</td>
<td>172.3</td>
<td>17.01</td>
<td>7</td>
</tr>
<tr>
<td>Prob. based</td>
<td>26.7</td>
<td>155.4</td>
<td>5.82</td>
<td>16</td>
</tr>
<tr>
<td>GA</td>
<td>34.9</td>
<td>115.7</td>
<td>3.31</td>
<td>16</td>
</tr>
<tr>
<td>Throwbox</td>
<td>34.8</td>
<td>120.6</td>
<td>3.46</td>
<td>16</td>
</tr>
<tr>
<td>Analyt. Greedy</td>
<td>38.6</td>
<td>98.7</td>
<td>2.56</td>
<td>16</td>
</tr>
<tr>
<td>Analyt. SA</td>
<td>38.9</td>
<td>96.3</td>
<td>2.47</td>
<td>16</td>
</tr>
</tbody>
</table>

According to Table 6.4, the analytical-based SA and Greedy give a better performance in comparison to simulation based approaches. Also, Greedy and SA improve the design phase by avoiding excessive simulations.

6.6 Conclusions

In this chapter, we proposed an analytical model to study the effect of relay placement techniques on the delivery performance of DTNs. In the proposed
analytical model there is no prior knowledge about the contact opportunities and traffic model among the relay and mobile nodes and mobile nodes are assumed to know the position of the relay nodes which are stationary. Mobile nodes can communicate to each other through the relay nodes and relay nodes never initiate any communications. Our model is dependent on the mobile nodes’ mobility pattern, and we considered the case when the mobile nodes move according to the random waypoint model. Further, we provided two different placement techniques based on our analytical model which are using heuristic search algorithms namely simulated annealing and greedy to place the relays. The first approach optimizes the placement problem considering the effect of all relays together while the other one considers an iterative placement of relays one at a time. The chapter showed that our analytical based approaches outperform the simulation based approaches in the literature. There are still some open questions that we leave for future work such as studying real scenarios when there are bandwidth constraints, delivery failures, etc., between mobile and relay nodes. Moreover, we want to study the required number of relay nodes given an expected delivery performance value for the network. Additionally, we will study our approach for other mobility models.
Chapter 7

Conclusion and Future Directions

7.1 Conclusion

Our research on design and developing an effective routing protocol for delay tolerant networks (DTNs) addresses the open technical issues of how to efficiently deliver messages in such disconnected networks that in general suffer from lack of continuous network connectivity. We specifically focused on routing issues in DTNs. This thesis studied various DTN routing techniques, and proposed a new protocol that improves delivery performance of the DTNs using the idea of augmenting the DTNs with low-cost easily deployable passive stationary relay nodes.

We found that the number of contact opportunities is the most important factor towards increasing the delivery performance of the DTN routing schemes. However, one has to consider the trade-offs between increasing contact opportunities in DTNs and routing assumptions and efficiency such as controlling or not controlling the mobile nodes, traffic patterns, power consumption, routing protocol complexities, high scalability, fault tolerance under dynamic environment, and high delivery performance in terms of message delivery rate and message delay.

Our DTN routing protocol called Tag Based Routing (TBR), introduced in Chapter 2 was presented and evaluated for handling data delivery in disconnected networks using easily deployable passive wireless devices acting as stationary relays. We found that the nature of our protocol which is based on cooperation of
mobile nodes/stationary relays makes our protocol to effectively increase contact opportunities while to be robust to cope with the highly dynamic and unpredictable environment in DTNs.

We evaluated TBR and its performance thoroughly via comprehensive simulations and analytical models. As part of our study we presented and implemented a variety of entity and group mobility models in order to evaluate the performance of our protocol under various scenarios. Our studies indicate that our approach can achieve competitive message delay and delivery rates while it outperforms existing state-of-the-art routing protocols in the literature.

Passive relay nodes’ characteristics including their location and buffer size have a high influence on the TBR performance. Accordingly, at first, several techniques for optimizing the stationary relay node placement namely relay pruning, probability based relay distribution, and a genetic algorithm based optimization, were demonstrated in order to investigate how to further improve the delivery performance of the proposed protocol. Our studies show that using a better placement technique, the performance of TBR can be further improved by a factor of 250% in both message delivery rates and average message delays.

Secondly, we studied effective buffer management policies in order to address the relay buffer size constraints. Our simulation results show an optimized buffer management policy by considering more relevant information in the context can significantly increase the performance of TBR in comparison to traditional buffer management policies like drop-tail or drop-front which only consider temporal information related to the messages.

To verify our simulation studies, we developed analytical models in Chapters 5 and 6 to model our proposed protocol performance without the need of pre-running simulations. Mathematical equations were provided that can indicate performance of the protocol based on given settings. The proposed analytical model can overcome the burden of simulation based studies in terms of both time and resource complexities while proposing a more general behavior model independent of any specific scenarios.

Moreover, we utilized our analytical models to analytically evaluate the performance of relay placement techniques in DTNs in order to further improve placement techniques. In Chapter 6, we proposed two different techniques which
are based on the proposed analytical models. The simulation results show that our analytical based placement approaches outperform the simulation based approaches in terms of data delivery performance by more than 30%.

The summary of the overall results we presented in this thesis are:

- Achieving a competitive routing scheme in DTNs is practical by augmenting the network with low-cost easily deployable passive stationary relay nodes. An implementation for such relay nodes is passive RFID tags. RFID technology is desirable because it operates wirelessly and without the need for attached power. This makes its deployment relatively easy and sustainable. Our evaluation results based on the passive RFID tags implementation have shown acceptable message delivery performance in terms of both message delivery rate and delay and it is specifically desirable for radio silent applications.

- There are trade-offs between the configuration of DTNs and the performance of the routing protocols. We showed these trade-offs through simulations and analytical models. For example, we showed the relationship between communication ranges of network elements (NEs), their buffer capacity, the number of participating NEs, etc. These relationships are important because they can be used to design and develop DTNs effectively.

- Robustness and efficiency is essential in DTNs as the environment is highly dynamic, partitioned, and unpredictable; hence, enhancement and optimization of DTN routing protocols is crucial. We proposed effective relay placement techniques as well as effective buffer management policies in order to significantly increase the performance of our proposed protocol. Moreover, we proposed a rigorous analytical framework to find the optimal relay placement which can be used to further increase TBR performance.

## 7.2 Future Directions

During our research, we identified the following topics of particular interest to DTNs:
7.2 Future Directions

1. **Real Life Applications and Scenarios.** This thesis was based on the hypothetical applications and scenarios which are in general provided in the literature. In order to commercialize our thesis, we are looking to make a real world application in which real observations would be studied. As an example, we studied our protocol under the effect of traditional mobility models but not in a real scenario having real traces. Other examples include possible constraints that can happen in a real scenario such as bandwidth constraints, delivery failures, etc., between mobile and relay nodes.

2. **Power Complexities.** TBR has been shown to minimize the extra overhead of mobile node power consumption to find other nodes in the network as nodes never communicate messages directly to each other, rather they communicate messages only via these stationary relay points which are in known places. Banerjee et al. [5] show that using 802.11 radio to search for contacts in a DTN devotes 99.5% of the total energy of mobile nodes just to find other nodes which is leading to have a short network lifetime. However, a full study on power issues and constraints in TBR is still remaining for future work.

3. **TBR Enhancement and Optimization.** Several studies presented in this thesis were done in order to optimize and enhance TBR functionality and performance. There are still open areas that we believe could lead to an increase in the performance of TBR. For example, utilizing a meta-knowledge scheme among the mobile/relay nodes can help forwarding/storing messages more efficiently. In our future work, we would like to utilize distributed artificial intelligence techniques to TBR and make it more robust under the dynamic environment as well as increasing its performance.

4. **Security Issues.** In TBR we assumed that all mobile nodes cooperate in message delivery. However, in a real application, there might be situations where selfish nodes may deny to distribute other mobile nodes’ messages. In other situations, malicious nodes can dramatically affect the performance of TBR. Malicious nodes may attempt to reduce network connectivity by
pretending to be cooperative but in effect not delivering any data messages. This situation can be intensified if they attempt to overwrite relays’ buffer with their own messages threatening the delivery performance of the protocol. In our future work, we propose different mechanisms to effectively detect malicious nodes and find methods to balance their side effects.
References


147


REFERENCES


REFERENCES


[75] Tara Small and Zygmunt J. Haas. The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there
is a way). In MobiHoc ’03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing, pages 233–244, New York, NY, USA, 2003. ACM.


In this section we show how to calculate the probability of hitting a relay by a mobile node. Given a node at point \((x, y)\) in the region, we can compute the probability, \(P_{hit}(x, y, i, j, r)\), that it hits a relay centered at \((i, j)\) with range \(r\) during its next epoch. All of the destination points in the region are equally likely. Some of those destinations will give rise to a hit; in Figure A.1 (a), these points are in the shaded region. Let \(\delta = \sqrt{(x - i)^2 + (y - j)^2}\) be the distance from the node to the relay center. If \(\delta \leq r\) then a hit occurs with probability 1. Otherwise, a hit occurs with a probability given by the fraction of total region that gives rise to a hit.

We consider the node’s movement given by a direction angle, \(0 \leq \phi < 2\pi\). Let \(0 \leq \alpha < 2\pi\) be the direction from the node to the center of the relay and the angle of integration is \([\alpha - \theta, \alpha + \theta]\), where \(\theta = \tan^{-1}\frac{r}{\sqrt{\delta^2 - r^2}}\). We found the area of dashed region in Figure A.1 (a) to be:

\[
SD = S_{\square}^{(x,y)(X,Y)(X,Y)(X,Y)} - S_{\triangle}^{(x,y)(X,Y)C} - S_{\square}^{(x,y)(X,Y)D} - 2S_{\cap MP}^{PM(x,y)} + 2S_{\cap}^{MP},
\]

where the \(\square\) and \(\triangle\) indicate that the element shown by the coordinates is a square or a triangle respectively. \(S\) stands for the area of region. For example, \(S_{MP}^{\cap}\) means the area of the segment \(MP\). To calculate the area of the dashed region, we break it up to the smaller areas and calculate the area of each component individually. Then, we sum up the results to find the total area of the
dashed region. Therefore, we divide the dashed region to 4 sub-areas as follows:

\[
S_{\bigcap MP} = \frac{1}{2} r^2 \left( \frac{\pi}{2} - \theta - \cos(\theta) \right),
\]

\[
S_{\triangle PM(x, y)} = \frac{1}{2} r (\delta - r) \cos(\theta),
\]

\[
S_{\triangle(x, y)(X, Y)C} = \frac{x_1 y_1}{2}, \quad S_{\triangle(x, y)(X, Y)D} = \frac{x_2 y_2}{2}.
\]

To calculate the value of \(x_1, y_1, x_2,\) and \(y_2,\) we use a function \(\beta(\phi; x, y)\) that provides the distance from \((x, y)\) to the perimeter of the region in the direction of \(\phi.\) Based on that, we can define:

\[
x_1 = \frac{1}{2} \beta(\alpha - \theta) \abs{\cos(\alpha - \theta)},
\]

\[
y_1 = \frac{1}{2} \beta(\alpha - \theta) \abs{\sin(\alpha - \theta)},
\]

\[
x_2 = \frac{1}{2} \beta(\alpha + \theta) \abs{\cos(\alpha + \theta)},
\]
\[ y_2 = \frac{1}{2} \beta (\alpha + \theta) \text{abs}(\sin(\alpha + \theta)) . \]

The \( \beta \) function is determined by the bounding region. In a rectangular region given by coordinates \( x_0, x_1, y_0, y_1 \), we consider six possible (not mutually exclusive) cases (Figure A.1 (b)).

We also consider nine different cases to calculate the area of bounding square region \( S_{(x,y)(x,Y)(X,Y)(X,y)} \), which is presented in Table A.1 (\( a_1 = \alpha - \theta, a_2 = \alpha + \theta \)).

From the above we find:

\[
P_{\text{hit}} = \begin{cases} 
\frac{SD}{S_{\text{total}}} & \delta > r, \\
1 & \text{otherwise}. 
\end{cases}
\]
\[
S_{(x,y)(x,y)(X,Y)(X,y)} = \begin{cases} 
(x_1 + x_2)(Y - y) & a_1 \geq 0, a_1 \leq \pi/2, a_2 \geq \pi/2, a_2 \leq \pi, \\
(X - x)(y_1 + y_2) & a_1 \geq 0, a_1 \leq \pi/2, a_2 \leq 0, a_2 \geq -\pi/2, \\
(x + x_1)(Y - y + y_2) - x_1 y_2 & a_1 \geq 0, a_1 \leq \pi/2, a_2 \leq -\pi/2, a_2 \geq -\pi, \\
(x + x_2)(y + y_1) - x_2 y_1 & a_1 \geq \pi/2, a_1 \leq \pi, a_2 \leq 0, a_2 \geq -\pi/2, \\
x(y_1 + y_2) & a_1 \geq \pi/2, a_1 \leq \pi, a_2 \leq -\pi/2, a_2 \geq -\pi, \\
(x_1 + x - x)(y_2 + y) - x_1 y_2 & a_1 \leq -\pi/2, a_1 \geq -\pi, a_2 \geq \pi/2, a_2 \leq \pi, \\
(x_1 + x_2)y & a_1 \leq -\pi/2, a_1 \geq -\pi, a_2 \leq 0, a_2 \geq -\pi/2, \\
(X - x)(y_1 + y_2) & a_1 \leq 0, a_1 \geq -\pi/2, a_2 \leq 0, a_2 \geq -\pi/2, \\
(X - x + x_2)(y_1 + Y - y) - x_2 y_1 & a_1 \leq 0, a_1 \geq -\pi/2, a_2 \geq \pi/2, a_2 \leq \pi, \\
\max(x_1, x_2)\max(y_1, y_2) & \text{otherwise.}
\end{cases}
\]

Table A.1: \( S_{(x,y)(x,y)(X,Y)(X,y)} \).
Appendix B

Genetic Algorithm based Optimization

In our GA placement approach, the genome is represented as a sequence of relay coordinates:

\[ [x_1, y_1, x_2, y_2, \ldots, x_n, y_n] \]

where \( x_i, y_i \in [0..1] \) (0..1 is a normalized coordinate). A single population was used of 120 genomes, with each genome initialized by placing relays uniformly at random in the region. The number of elite genomes was set to 10 and the search was run for 100 iterations. Through some preliminary experiments we determined some GA parameters values that improved GA performance. We did not do an exhaustive search over the entire parameter space. We leave this to future work. The mutation and crossover functions were customized for our problem. We hypothesized that the structure/topology of the geographic relationships between relays is important in terms of performance. In order to allow the GA to learn about spatial characteristics of relay placement, we identified regions of relays using a breadth first tree construction based on the Delaunay graph [53] defined by the genome. Initially, we simply chose relays at random, rather than using a breadth first tree approach and the resulting GA performance was comparably poor.

Figures B.1 (a) and (b), show two genomes, each consisting of 50 relays. The Delaunay graph of the relays is shown using light lines. A random breadth first tree is constructed by choosing a relay at random and forming a breadth first tree
that consists of the required number of relays. An example set of relays in such a tree is shown using solid dots for each genome.

The child shown in Figure B.1 (c) is constructed from the selected trees in genome 1 and genome 2. In practice, for our crossover function, the number of trees and the number of relays in each tree selected from genome 1 is randomized, i.e., we choose a small number of random trees from genome 1. The resulting number of relays (and number of trees) selected from genome 2 is constrained so that the total number of relays in the child equals $N_{\text{relay}}$. If genome 1 and 2 have identical selected relays (as in the case when they have common ancestors) then one of the identical relays is replaced with a point selected at random in the region.

As an example representing crossover operator, assume the following two genomes:

\[
g_1 = [x_{1,1}, y_{1,1}, x_{1,2}, y_{1,2}, \ldots, x_{1,n}, y_{1,n}]
\]

\[
g_2 = [x_{2,1}, y_{2,1}, x_{2,2}, y_{2,2}, \ldots, x_{2,n}, y_{2,n}]
\]

therefore, the result of crossover operator would be a new genome as follows:

\[
g_3 = g_1 \times g_2 = [x_{3,1}, y_{3,1}, x_{3,2}, y_{3,2}, \ldots, x_{3,n}, y_{3,n}]
\]

Our mutation operation similarly selects a random breadth first tree (in practice consisting of up to 5% of the relays). For mutation, the selected relays are replaced with relays chosen at random in the region. Figure B.1 (d) is a mutation of genome 2. Note that a “hole” appears where the selected relays were, since the new relays are less likely to appear in the selected region (for a small number of selected relays and hence a small selected region).
Figure B.1: Example genetic algorithm crossover and mutation.