Universität Stuttgart

Fakultät Informatik, Elektrotechnik und Informationstechnik

Master Thesis Nr. 3177

Design and Implementation of an Efficient Message Dissemination Protocol for Public Sensing

Amir Bannoura

Studiengang:	Information Technology
Prüfer:	Prof. Dr. Kurt Rothermel
Betreuer:	M.Sc. Patrick Baier
begonnen am:	18. April 2011
beendet am:	18. October 2011
CR-Nummer:	C.2.1, C.2.2, C.2.3, C.2.4, E.1

Institut für Parallele
und Verteilte Systeme
Abteilung Verteilte Systeme
Universitätsstraße 38
D-70569 Stuttgart

Abstract

This thesis focuses on message dissemination and information distribution for large size messages. Our work considers different message dissemination approaches that depend on the communication mechanism to distribute messages. We use short range communication (ad hoc) to disseminate messages among the mobile nodes. Also, we use long range communication (GPRS) to transmit messages from the sever to the mobile nodes.

Our main aim is to produce an approach that is effective. This means delivering the message to all the mobile nodes that are located in the area of interest. In addition, message dissemination process must be efficient in order to conserve the mobile nodes limited resources.

Our work discusses three different approaches of message dissemination. The first approach is a naive approach in which a message is delivered directly from the server to all the nodes using GPRS communication. The second approach is a static approach that consists of two parts. First, messages are sent to a subset of mobile nodes using GPRS. The selection of this subset is based on the idea of the set cover problem. Then, the chosen mobile nodes broadcast ad hoc messages to their neighbors. The third approach is the mobility approach which has a geographic aware sensing capabilities. It is built around the concept of matching problem to cover the segments that form the area of interest. There are some similarities to the static approach. First, the nodes subset is chosen according to the mobile nodes' information status and the geographic information of their locations. Messages are sent using GPRS to this subset. Later, the mobile nodes move around to cover the area by broadcasting ad hoc messages.

Finally, we present the evaluation of the three approaches. We compare between them from the point view of message delivery rate and energy consumption. Also, the evaluation shows how message size has an effect on the total amount of energy that is consumed.

Acknowledgements

I would like to take this opportunity to thank the Distributed Systems department represented by Prof. Kurt Rothermel for giving me the chance to write my Master Thesis. A deep appreciation for Dr. Frank Dürr and his research group for allowing me to be a student researcher within their group.

Special thanks to the DAAD for their support during the last two years. You were always there guiding us through our studies, giving us the opportunity to be here and letting us get to know new cultures and people.

The support that I received from my family, friends in Germany and around the world is unimaginable. I hope that I was up to your expectation. Thank you all for your love and trust.

Last but not least, this work won't come into existence without the support and guidance of my supervisor Patrick Baier. Thank you for your remarks during the course of this work and for your sense of humor and our talks about sports.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Technology is developing in a rapid speed which has a continues impact on our daily life. The major effect is in the hardware area where production of small chips and sensors of low cost increased. Such that every new device is shipped with these sensors. The introduction of sensors in mobile phones helped in transforming the way how people contact each other. It changed the way how information is exchanged. Currently, companies are racing in developing new smart devices that have several sensors included for different purposes such as measuring speed, temperature and their ability to provide Global Positioning System (GPS). This kind of development of small inexpensive sensors that is implemented in smart devices opened the path for context-aware public sensing.

1.1.1 Public Sensing

Context-aware public sensing [17] is the process of gathering contextual information about the users' activities and the surrounding environment using mobile devices. An example of public sensing is citizen noise pollution monitoring [31] which allows the people to provide a rating about the level of the noise and tagging its location in urban environment. In context-aware public sensing, different sensing scales that reflect the type of the sensing in an environment are available. These scales depend on the user behavior and his willingness to share information. There are three types of sensing scales:

- **Individual sensing:** This type of sensing provides the user with a generated data about his own activities. This data is not shared with any other party. It is only for self interest information. UbitFit Garden [18] is an example of self interest application that gathers data for personal use which helps the users to monitor their personal health and fitness.

- **Group sensing:** An individual user participates in activities that concern other users through information sharing. These users form a group that shares common goals and activities in the same interest subjects. An example of group sensing is GarbageWatch [1] where users participate in improving the recycling process of the garbage.
- **Community sensing:** When there are a massive number of individual users who participate in the sensing process. Then, information gathering becomes useful. The usage scale of this type assures the users that the available information is of value to them. An individual might be interested in knowing that a city has some street segments with high traffic density to avoid taking these streets. This will help him to find a detour to his destination.

The sensing scales raise an important issue regarding the involvement and participation of the individual users. There are two sensing methods that affect the way how data is gathered and disseminated to be available to different users. The first sensing method is participatory sensing [14] that aims on letting the individual user be largely active in the process of sensing the surrounding environment. WeTap [2] is a participatory sensing project that allows users to map, identify and describe the public fountains for water drinking. On the contrary, opportunistic sensing [15] aims to make the user involvement limited in data gathering in which his participation will be passive. Pothole Patrol (P^2) [3] is an opportunistic sensing project that allows cars to monitor the roads surfaces and report to a server.

The use of the sensing scales will be employed to design an efficient message dissemination protocol for transmitting messages to mobile nodes.

1.1.2 Message Dissemination

Message dissemination is defined as the process of sending a message to the public without receiving any direct feedback. Companies tend to use message dissemination to reach a lot of people to promote and advertise for their products. The mechanisms to reach this large number of people must be cost efficient in order to achieve valuable results in terms of the resources that are available for the use.

In addition, users like to be up to date on specific topics. As an example, a football fan likes to see match highlights of his favorite football club. The message content size that the fan receives is very large. The delivery of such messages may deplete the available resources which makes it difficult to receive any others.

The requirements for designing an efficient message dissemination protocol depend on the content type of the message which is disseminated. Normal text promotion ads received from companies consume less energy than the match highlights received by the football fan. Message sizes define the time and the amount of resources needed to disseminate the message. In this sense, we want to employ the public sensing scales to find a mechanism to disseminate messages in an efficient manner.

Use of Public Sensing in Message Dissemination

Public sensing is used to gather information regarding the surrounding environment. The information is employed to create a knowledge set about the area and the mobile nodes. The knowledge about the location of one mobile node gives an estimated distance how far the next mobile node is located. In addition, the knowledge about the area provides a way in which it predicts the mobile nodes movement. This knowledge helps to develop a message dissemination protocol to transmit large messages efficiently by reducing the amount of energy consumed.

1.2 Research Objectives

Our research focuses on designing a protocol that is capable of sending large sized messages to all the nodes that are located in the area of interest. The idea is to disseminate a message and deliver it by all the nodes that are located in the area of interest. Figure 1.1 explains a naive solution to deliver a message to all the nodes that are located in the area of interest.

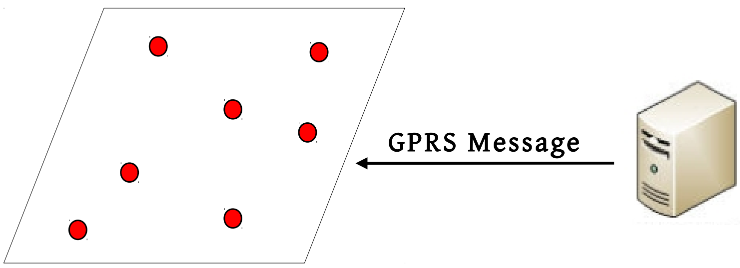

Figure 1.1: *Message Transmission*

Although, the naive solution sends the message using GPRS to all the nodes in the area. However, message delivery using GPRS is costly. The amount of energy consumed on the mobile nodes is large. Our goal is to design a protocol that can achieve the following characteristics:

- **Energy Efficient:** The limited nodes' resources force us to find a way to reduce the total amount of energy consumed on the mobile nodes when a message is disseminated.
- **Message Delivery Effectiveness:** Our aim is to increase the message delivery rate. The protocol will transmit the message to all the nodes in the area of interest.

There is a trade off between message delivery rate and the total amount of energy consumed by the mobile nodes using the naive solution. Different solutions are proposed to achieve a balance between both characteristics by using a hybrid communication techniques. However, there are many constrains prevent us from achieving that balance.

1.3 Research Challenges

Designing a message dissemination protocol for wireless mobile nodes is a challenging task due to the several constrains that limit the nodes from being able to communicate and exchange messages. Part of these constrains discussed in [10]:

- **Energy Efficiency:** The nodes have a limited resources to consume. Communication and message delivery must not take a long time that leads eventually to consume the nodes' resources.
- **Scalability:** The number of nodes that is located in one area might be a large amount. The designed protocol must adapt to such amount.
- **Mobility:** The mobility of the nodes effects the designing of the protocol. The fast movement of the mobile nodes gives short periods for message routing and delivery. The protocol must cope with the movement of the nodes.
- **Communication Techniques:** Nodes usually communicate with each other using ad hoc communication due to the fact that MANETs have no centralized infrastructure. The nodes might not be able to connect to other nodes for long period of times, which disrupts the network connectivity. However, Nodes can use GPRS communication to deliver messages but it consumes large amount of energy.

In addition to the above constrains, the environment where the nodes move imposes other restrictions. These restrictions complicate the nodes message delivery:

- **Area Size:** The larger the area the more number of nodes that it can accommodate. Small size areas can be found for example in city centers, which increases the period that the nodes have to exchange messages.
- **Node Density:** An Increase in the number of nodes in an area increases connectivity but decrease the speed of the nodes. This affect the nodes' ability to cover other parts in the area and limits message delivery.

These constrains limit our goal in delivering messages to all the nodes in the area of interest with the least amount of energy consumed.

1.4 Thesis Structure

This thesis consists of seven chapters including the introduction. Chapter 2 gives an overview of several approaches related to our work that concern message routing and delivery. A description of the system model and location model explained in chapter 3. Chapter 4 clarifies the concept of the three different message dissemination approaches: the naive, the static and the mobility approach that are designed in this thesis. A comparison between these approaches is given at the end of the chapter. An explanation about the simulation environment and the protocol implementation is provided in chapter 5. The result of our work is presented in chapter 6. Finally, a complete summary and a future look on the enhancement of the protocol are discussed in chapter 7.

Chapter 2

Related Works

2.1 Overview

The introduction presented an overview of the public sensing, message dissemination and the challenges that limit our work. The availability of context-aware sensing helps to gather different kind of information from the surrounding environment. The type of environment defines the way how nodes communicate with each other. In a highly dynamic environment, the nodes benefit from the gathered data to provide a common ground for communication among the nodes. This data helps to predict when the nodes might meet or how long one node carries the data until the exchange of the data occurs.

There is an increase interest in the development of communication techniques among the nodes in highly dynamic environments that are characterized by communication discontinuity. This led to shift the research toward Delay Tolerant Networks (DTNs) [16][39]. DTNs address the issue of network connectivity continuity problem. DTNs are characterized by:

- Connectivity discontinuity
- High packet loss rate
- Long time propagation

Some examples of DTNs include:

- Terrestrial Mobile Networks: Communication between Satellite and nodes might be disrupted due to the mobility of the nodes on the ground.
- Military Ad-Hoc Networks: Communication might be intentionally disrupted because the communication occurs in a hostile environment.

In addition, various research projects investigate in DTNs among those are: The Delay-Tolerant Networking Research Group [4] and PodNet [5]. In PodNet, the main aim is to develop a distribution system for communicating user-generated broadcast content among mobiles.

2.1.1 DTNs Structure

Several routing protocols are proposed to study the problem of DTNs. These protocols fall under two categories described in [39] based on the different types of DTNs.

- Deterministic Time Evolving Networks
- Stochastic Time Evolving Networks

Figure 2.1 provides an overview about the protocols that are discussed by DTNs. The protocols of deterministic time evolving networks are discussed in section 2.2 and the stochastic time evolving networks are discussed in section 2.3. Eventually, section 2.4 summarizes DTNs and compares the protocols to our approach.

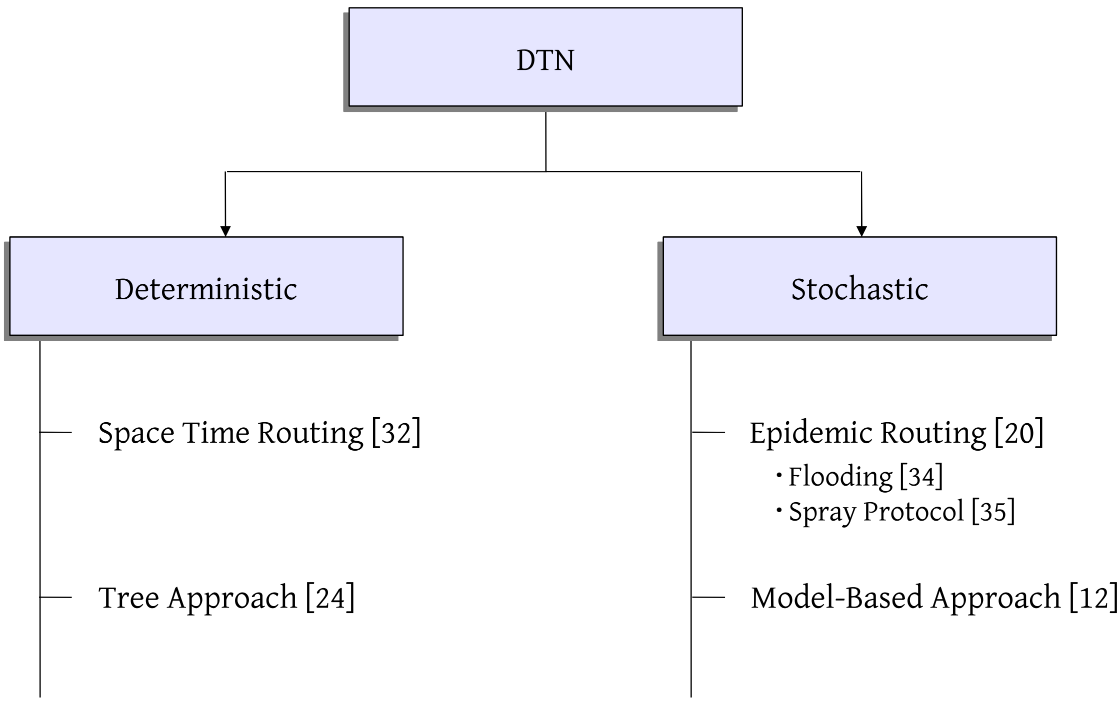

Figure 2.1: *DTNs Categories and Protocols*

2.2 Deterministic Time Evolving Networks

The future topology of this type of networks is deterministic and known or at least capable to be predicted [39]. These networks consider the knowledge of nodes' movement and link connections among each other. This section covers two approaches: space time routing discussed in section 2.2.1 and the tree approach discussed in section 2.2.2.

2.2.1 Space Time Routing

We consider both time and space dimension of the network topology in space time routing [32]. Furthermore, we introduce the concept of evolving graphs [21]. Evolving graphs have the ability to capture the networks' dynamic characteristics. These networks are time-varying networks which are formed at specific instances of time when one node is adjacent to another.

Definition: An evolving Graph is defined as a graph G(V,E) and a sequence of subgraphs $S_G = G_1, G_2, \ldots.., G_T$. The subgraphs indicate the network connectivity at any given time t. Then, the system yields g=(G,S_G) where each subgraph is an evolving graph.

Network Connectivity

The nodes v_1 and v_2 create a communication link between each other when node v_1 enters the geographic communication range of node v_2. A link is created only for that instance of time. When node v_1 leaves the communication range that is covered by v_2 the link between the two nodes is not active any more.

Message Delivery

The nodes' continues movement within the system makes it difficult to deliver the message directly from the source to the destination. Intermediate nodes carry the message until they come within the communication range of other nodes. The time that a node keeps carrying the message is ought to be a delay in message delivery. Sometimes the nodes that carry the message might not come into contact with another node which in turn lead the nodes to drop the message. In this sense, messages are lost and not delivered to their intended destinations.

Figure 2.2 represents an evolving graph. The graph has six nodes V = {A,B,C,D,E,F} and it's edges E are marked with time instance t. At the time t, there exist active links among the nodes. These active links between the nodes form subgraphs. Any node can forward the message to another node that is connected to it by an edge during that time t.

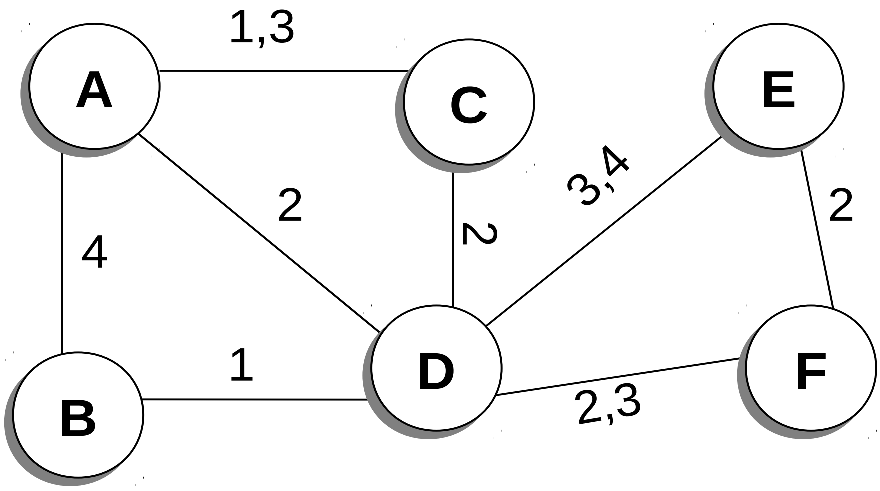

Figure 2.2: *Time-Space Evolving Graph*

A message can be forwarded only if the time instance isn't in a decreasing order with respect to the path that was previously passed. The path from D to F that passes through E is not a valid path for message delivery. Since the time instance between E and F is 2 and the link is established between D and E during the time 3, 4. This indicates that the link between E and F is already not active anymore when the link is established between D and E.

Assume that we want to deliver a message from B to E. There are two valid paths for message delivery. The first path is B-D-E and the second path is B-D-F-E. Although, the number of hops for the second path is larger than the first path. However, the delay in delivering the message is less. Message is delivered in the first path at time 3 whereas in the second path it is delivered at time 2. The shortest path isn't always the optimum path for message delivery. The messages are forwarded to the nodes without any knowledge about their movements. This issue is addressed in the tree approach that is presented in the next section.

2.2.2 Tree Approach

This approach depends heavily on the construction of characteristic profiles [24]. Characteristic profiles include the knowledge of each individual node's movement and time periods. Initially, the knowledge of node's characteristic profile is not known for all the nodes. Nodes exchange their characteristic profile to gain knowledge on how to forward a message. Based on this knowledge a tree is constructed in which a path is chosen to deliver messages from the source to the destination.

Tree Construction

The tree is constructed starting from the source node as a root. The other nodes are added as children to the tree assuming the message chooses the path that the node lies on. Unvisited nodes only added to the tree in order to avoid recursions. When the tree is constructed, a depth-first search is used to find the appropiate path that it will follow. The search could use several metrics in choosing the path. The message could be delivered with minimum number of hops or reducing the time delay of message delivery.

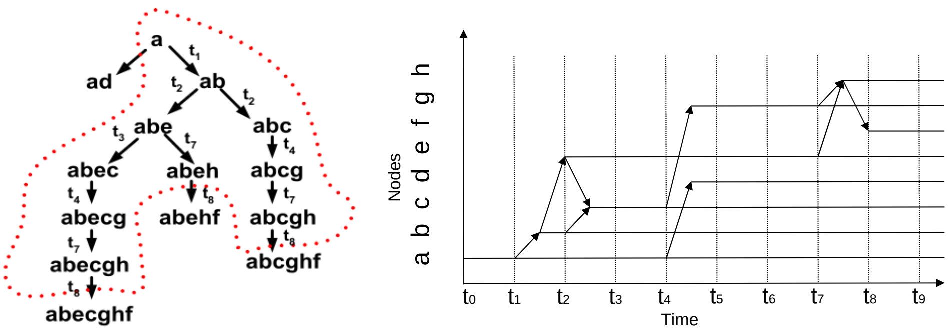

Figure 2.3: *Computed Path Tree*

A network contains V = {a,b,c,d,e,f,g,h}. An Edge E marked with time t indicates the connection time between the nodes. Now, let the source be a and the destination be h. Figure 2.3 on the left side [24] represents a tree. The source node a has a global knowledge about the paths that the message can travel to reach the destination h. Those paths were built using nodes' characteristic profiles that are exchanged during the connection time in figure 2.3 on the right side. An arrow symbolizes that the nodes can exchange a message at that period of time. It means that two nodes are in the communication range of each other. At period [t_1,t_2], node a connects with node b. At period [t_2,t_3], there are two available paths. The message either is delivered by node c or node e. The process of adding nodes to the tree continues until all the nodes are added. Again, no node is added to the tree more than once in order to avoid recursions.

The dashed red contour in figure 2.3 on the left side surrounds the available paths that could be chosen to deliver a message to the final destination. Any metric can be used such as shortest time delay or minimum hop count. Although, the time delay is the same if we choose any path in this example but the number of hops that the message travels is different. The minimum number of hops that a message travels is four through the path a → b → e → h.

2.3 Stochastic Time Evolving Networks

Stochastic time evolving networks are networks where the future topology is not known and can only be estimated [39]. In the previous section 2.2, deterministic networks routed the message based on some knowledge about the path that it has to travel to reach a destination. In Stochastic networks, messages are carried within intermediate nodes until they reach their destination. The path to the destination is not known. In this section, we discuss two approaches for stochastic networks. Section 2.3.1 discusses an epidemic approach where the intermediate nodes carry a message until the message is received at the destination. Section 2.3.2 gives an explanation on a model based scenarios where the nodes move in a certain pattern not in a random way.

2.3.1 Epidemic Routing Approaches

Since the knowledge about the paths is not available. Epidemic routing [20] depends on the intermediate nodes to carry the message. Then, the nodes move around and come across other nodes. The number of nodes that carries the message is extended. Although, the aim is to deliver a message to a specific node. Epidemic routing goal is to minimize the number of intermediate nodes that carries the message to the final destination and decreases the time delay of message delivery. However, increasing the number of intermediate nodes provides better results specifically in message delivery. But the problem is that it increases the total amount of resources consumed. In the epidemic approach, two protocols discuss the matter of message delivery through intermediate nodes. The first is flooding [36] and the second is the spraying protocol [35].

Flooding

The fastest way to disseminate and spread a message is to let the intermediate nodes flood it to their neighbors. The message is received by another number of nodes and they also can flood the message again. The way the epidemic approach work is that two nodes within the communication range of each other exchange the messages which are already not seen by anyone of them. As an example, node A carries messages with ids (1,2,3) and node B carries messages with ids (1,2,4). Both of the two nodes exchange message 3 and

4. Eventually, both nodes carry messages (1,2,3,4).

A consideration to the node's buffer size is important in order to accommodate the number of messages that are disseminated in the network at any given time instance. The buffer size must at least be equal to the number of messages flooded to ensure a higher message delivery rate. Otherwise, a message delivered to a node that has a full buffer would lead to drop the message that might be intended to that specific node. The evaluation results of the simulation in [36] using an infinite buffer size in comparison with finite buffer size resulted in the fastest message delivery with a delivery rate of 100%.

In addition, each message is associated with a hop counter. The larger the hop counter is the more nodes the message is delivered. However, the total amount of resources consumed within the network increases while the message travels more nodes.

An optimized flooding protocol (OFP) [33] is proposed to minimize the number of nodes that broadcast a message. OFP is a location-aware approach that uses a strategic nodes' locations to retransmit the message. It depends on minimizing the number of circles that covers the cells. In this way, not all the nodes are required to retransmit the message, only those nodes that lie within the transmission range of the other nodes and minimize the number of circles in the area. Figure 2.4 [22] explains the broadcast situation and shows the minimum number of circles that covers an area. Notice that only the nodes who retransmit the message are located at the edge of the communication range of the node that broadcast it. Only allowing those nodes to retransmit the message will cover more area and reduce the amount of resources required to transmit the message.

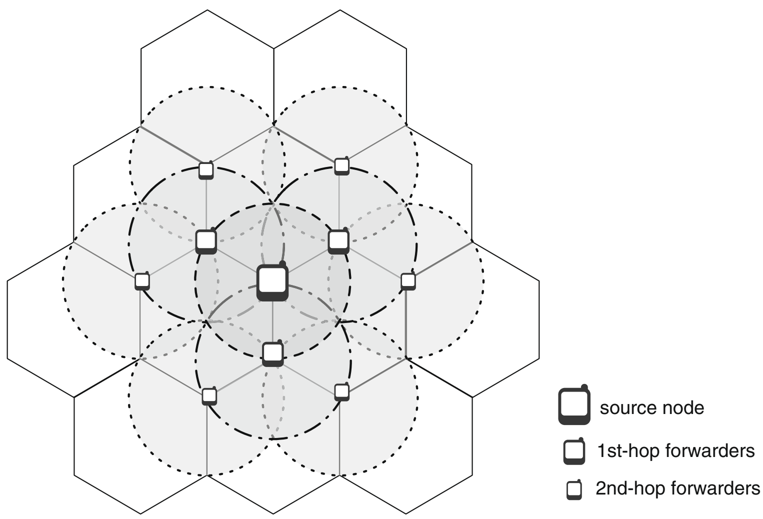

Figure 2.4: *OFP Principle*

Spraying Protocol

In spraying routing [35], it is assumed that the location of the final destination is known to the source. In this sense, the system includes a tracking manager that is responsible for tracking nodes' locations. The nodes register with the location manager responsible for the cell which the nodes are located in. An update of the nodes' location positions occurs each time the nodes move between the cells. The system contains a combination of switches and end points. These switches are intended for routing the message to its final destination.

Spray routing forwards messages within the neighborhood of the last known location of a node's destination endpoint [35]. In this way, it limits the process of flooding the message to all nodes. The message is forwarded to the nearest neighbor cell of the destination location using unicast communication. Then, a broadcast occurs to the message to a defined number of cells. The spray routing uses hybrid of unicast and broadcast communication. However, it is ambiguous matter on how the first nodes are chosen to receive the message using unicast communication. In addition, the protocol includes two parameters:

- Depth: Indicates the number of hops a switch is far from the destination switch cell before broadcast occurs. A message is sent in a unicat fashion to a switch that is away X hops from the destination cell. This X is the hops number of switches to reach the destination cell switch.
- Width: Indicates the cell level number of neighbors in which the message will be sprayed. This number determines the number of levels to be sprayed in order to deliver the message to the destination because the destination node can move from its current cell due to its mobility property.

Figure 2.5 explains the way message disseminates using the spray routing approach. At the beginning, the message is unicasted to one of the gray cells. The depth value in this case is 1 because one hop is required to reach the white cell where the final destination last known location. The switch in the white cell sprays the gray cells with the message to cover the node's movement after the message is received at the switch in the white cell. The value of the width in this case is also 1 because we have one level of neighbors.

The values of the depth and width depends largely on the nodes mobility. High mobility indicates a larger values might be needed. Those values are determined based on the information history generated on the switches such as time of nodes' connection and maintainability within one cell.

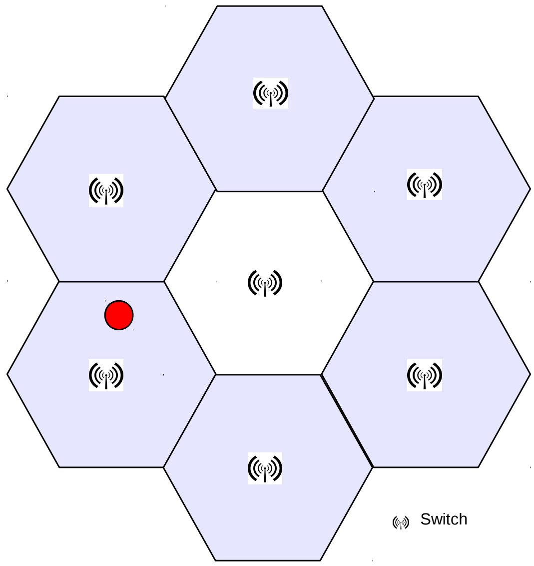

Figure 2.5: *Spray Routing Approach*

2.3.2 Model-Based Approach

Model-Based Routing (MBR) [12] approach realizes real world models. The nodes' movement is based on some known patterns. These patterns reflect the nodes' movements along a certain path. It is assumed that the nodes are either vehicles travel on streets or pedestrians walk on the sidewalks. MBR depends on the following information to route a message to a node:

- Location Model: It contains geographic information about the location area such as streets, buildings and intersections.
- User Profile: The benefit of user movement patterns reveals the probability of being in a specific location at a specific time.

The available knowledge of user profiles reduces forwarding the message to intermediate nodes. A message is forwarded to high probability nodes that travel to the destination or near by locations. Although, location models are available most of the time for a specific region. However, the problem is that the certainty of the nodes meeting each other reduced in low density areas. Also, the nodes' behavior exhibit different speed frequencies. Pedestrians walk in a low speed compared with vehicles. Thus, the meeting time of the pedestrians is larger than the vehicles. This issue makes the network topology changes very fast giving vehicles a very limited window to exchange messages.

2.4 Related Works Conclusion

2.4.1 Summary

This chapter gave an overview of some of the protocols used for message dissemination and routing. It explained the delay tolerant networks and the challenges associated with these networks in disseminating a message to reach a destination. Several protocols were discussed and categorized under two categories: deterministic and stochastic time evolving networks. The difference between these two categories is that the future topology in the deterministic network is known or capable of being predicted. Whereas the future topology in the stochastic network is not known or only estimated. Decisions on a message dissemination and path selection by protocols depend on the available information gathered from the surrounding environment as well as the history and profile characteristics of the mobile nodes.

2.4.2 Comparison to our Approach

Most of the protocols discussed in this chapter focus on finding a solution on how to disseminate a message using ad hoc communication. These protocols didn't brought the attention to how the server chooses the subset of nodes that receive the message using unicast transmission before its broadcast to the other nodes. The only protocol discussed this matter is spray protocol (see section 2.3.1). Similar to spray protocol, our approach depends on hybrid of communication short and long communication ranges. However, we provided a precise technique in choosing the subset of mobile nodes that receive the message using GPRS.

Our Approach provides a way to check if a message is delivered by the neighboring nodes or not. In contrast to the flooding protocol (see section 2.3.1) that broadcast messages even if all the nodes have delivered the message. Our approach is designed to flood the area if at least one node didn't deliver the message. Through our approach we limit flooding the message in the network and we reduce message duplication in mobile nodes.

Model-Based Routing (see section 2.3.2) realizes a real world scenarios. These scenarios provide the geographic information of a specific region. The MBR benefits from the information in knowing the geographic structure of the map which helps the message routed in an efficient manner. Also, our approach benefit from the geographic information by determining on which segment the node is moving and the number of connected neighboring segments. This information is used as the core in our approach implementation.

The knowledge of nodes' movement pattern could reveal the possibility when nodes come in contact with each other. Both space-time approach (see section 2.2.1) and tree approach (see section 2.2.2) explain the predictability of the nodes meeting at a certain time. These approaches use the characteristics behavior of the mobile nodes to determine the routing paths to reach the final destination. The characteristics behavior of the mobile nodes are included in our approach to know the position of the mobile nodes and the direction of their movement because these characteristics provide the server the help it needs for the selection mechanism of the subset of mobile nodes.

Chapter 3

System Design

Our main purpose in designing a system is to make it reliable and extendable to accommodate the increasing number of messages that will be delivered. This requires a set of dedicated and concrete system components. Section 3.1 explains our system model that is required for building our protocol. Section 3.2 describe the location model that explains the location structure in general.

3.1 System Model

Our approach depends on a combination of communication techniques. It uses a long and a short range communication. In addition, some hardware components are proposed to facilitate the process of message dissemination. We assume that the hardware components are reliable and no failure occurs to the system components. Thus, our system does not consider any hardware components replication. Our approach is built around two main components a server and mobile nodes. Theses two components are tightly connected with each other to facilitate the message delivery.

We design our protocol to have the capability of delivering a message of a large size. The efficiency of our system depends on the total energy consumed in delivering a message. Our interest is to deliver large message size and reduce the total energy that is consumed. We need to find a suitable way to use the available nodes' resources efficiently. The hybrid communication techniques mentioned above try to adjust the consumption of energy among the nodes.

The nodes don't include a queue to store several message to process them later. It only handles each message at a time. We reduce the complexity of several messages being routed within the network. This reduce network congestion by minimizing the number of messages in the network.

A Single geographic area is our main interest. The aim is to let all the mobile nodes that are located in this area deliver the message. When they move out of the area, they are not considered anymore for message delivery. A node tracking to other cells or geographic areas is not possible because no handover between the cells is implemented. It means that the node's information isn't available out of the geographic area. It is very complex process to follow each node to the neighboring cells because of that this work focuses on a single area.

3.1.1 Server

Our system includes a central server that monitors the mobile nodes. It has the ability to communicate with the nodes using a long range communication. Also, it has a knowledge about the number of mobile nodes circulating in the area where the message dissemination happens because the mobile nodes subscribe themselves to the server providing their initial positions. Furthermore, the server contains a map where the nodes move. The map is divided into several components called segments. These segments are generated for example using GIS data sets [6]. These sets provide a realistic simulation of cities. They describe the location area and the map structure where the mobile nodes move (see section 3.2.1).

3.1.2 Mobile Nodes

Mobile nodes move in a continuous manner on the street segments. Thus, our choice is to consider moving vehicles rather than pedestrians (pedestrians carrying mobile devices). The point is that vehicles are in a continuous movement and the speed is ranged between 6.7 m/s to 18.7 m/s in our scenarios. Another reason for choosing moving vehicles is due to the incorrect mapping of pedestrians nodes to the street segments (see section 3.2.2). The nodes have the ability to identify their locations using Global Positioning System (GPS). When the server queries a node location, the node will be capable to provide the server on which segment it is located.

The mobile nodes use a combination of communication techniques. They use a short range communication to broadcast the message to neighboring nodes using ad hoc communication. Also, they use long range communication such as UMTS or GPRS to acknowledge the delivery of a message by the nodes.

3.2 Location Model

The knowledge of the geographic information about a specific area makes it easier to determine whether two nodes meet or not. Location models cover attributes such as street segments, intersections and building blocks. Moreover, algorithms such as Geographic Routing Protocols take advantage of the location information to make routing techniques more efficient [23]. Section 3.2.1 covers the map structure that we work with. Section 3.2.2 describes the mapping failure of mobile nodes to street segments.

3.2.1 Map Structure

The segments that construct the map are one of the important attributes in defining the location of different mobile nodes. The knowledge of this structure and the characteristic profile of the mobile nodes define the way how a message is exchanged in a specific area. The behavior of the mobile nodes such as speed and position is a factor that comply with the map structure. The area where the nodes move force the nodes to adjust their pattern in order to make it possible for a message to be delivered.

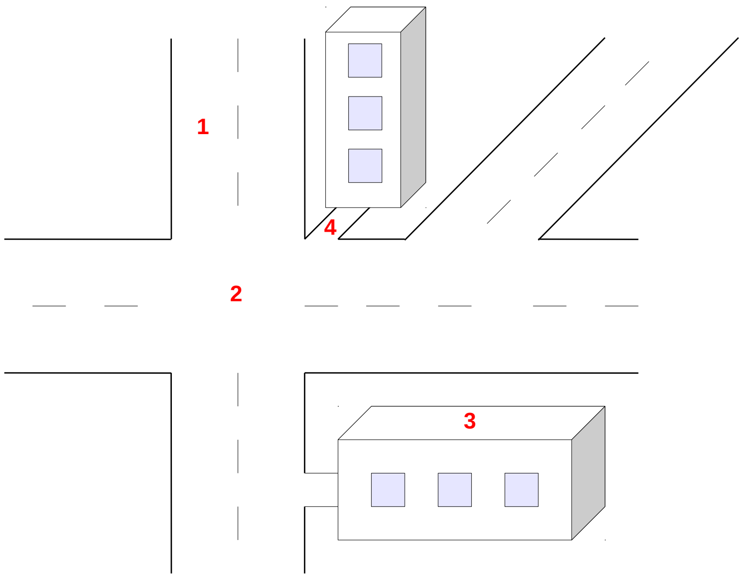

Figure 3.1: *Map Structure*

Figure 3.1 describes a simple map structure. The map constitutes of several components that are numbered as the following:

1. **Street segments:** The mobile nodes are able to travel in both directions.

2. **Street intersections:** A connection of several segments. At this point, the mobile nodes change their directions.

3. **A building:** Pedestrians have the ability to enter these buildings. The mobile nodes are considered in an idle state inside.

4. **A pathway:** These small segments connect larger segments to the buildings. The pathway might have an effect on the node position (see section 3.2.2).

Not only these components are available, but also there are several others such as rivers, sidewalks, traffic lights, base stations and subway stations. All of these components affect the way how nodes move and messages disseminate.

3.2.2 Incorrect Location Mapping

Mobile nodes are distributed over the map which includes the different structure described in the previous section. The positions of the mobile nodes are not directly fixed on the street segments. They are positioned in a short distance to the street segments as in figure 3.2. There is a need to map each node to a segment. In order to do that, the node is assigned to a segment that has a minimum shorter distance to that segment. The problem is that at some instances the minimum distance between a node and a segment is less than another segment which might disrupt the continuity of the mobile node movement.

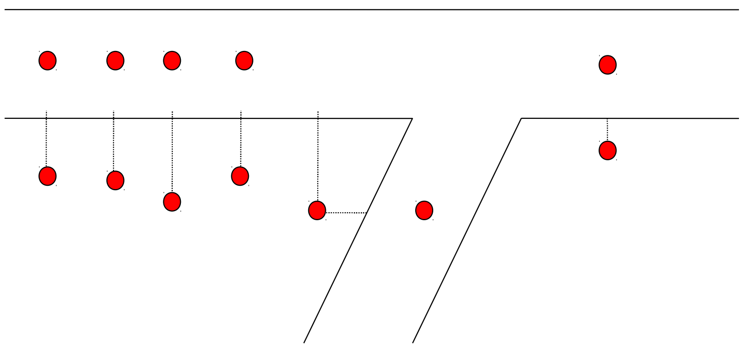

Figure 3.2: *A Node Mapping*

A node moves from the left to the right as in figure 3.2. A node is mapped to the segment when the distance to that segment is minimized. The first four points of the node are correctly mapped to the horizontal segment. The fifth point maps itself to the slopped segment because the distance to that segment is less than the horizontal one. The sixth point maps itself again to the horizontal segment. In order to minimize this problem, we consider only moving vehicles because vehicles are mapped to street segments not pathways. All of the pathways lead to the buildings and only pedestrians enter these buildings. The effect of incorrect mapping reflected in the following ways:

- Failure to pick the right subset of nodes from the complete set when deciding on an appropriate message dissemination approach.
- The wrong positioning of nodes can lead to let two nodes that are supposed to meet never have the chance to meet. This decreases message delivery rate because the node will not exchange the message.

Chapter 4

Conceptual Design

4.1 Overview

The hardware components and location model described in the previous chapter define a structural model that supports the design of message dissemination protocol. Multiple approaches provide different methods in disseminating a message. The designed concept behind a message dissemination depends on the outcome that we are planning to achieve.

Our main aim is to transmit a message to all the mobile nodes that are located within an area of interest. There are two methods to transmit a message to the mobile nodes. The first method uses GPRS and the second method is a message broadcast using ad hoc communication. Figure 4.1 describes the dissemination methods of a message to the mobile nodes in an area of interest. The basic approach uses GPRS to disseminate a message to all the mobile nodes in the area of interest. Other advance approaches allow the mobile nodes to broadcast the message using ad hoc communication to their neighbors.

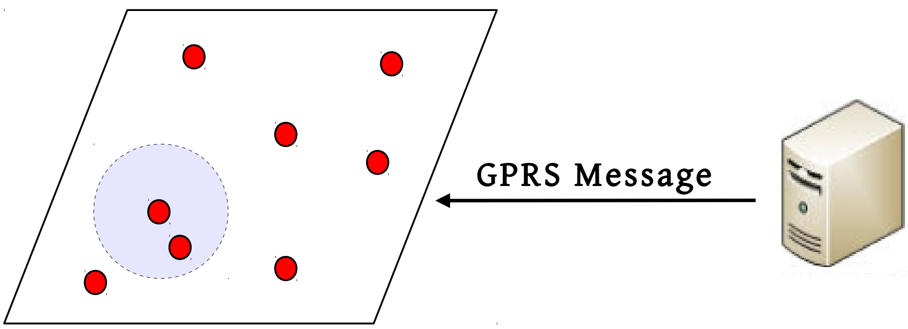

Figure 4.1: *Message Dissemination Methods*

In this chapter, we present three different message dissemination approaches. These approaches depend on the communication method that is used to disseminate the message. Also, they depend on the movement of the mobile nodes within the area of interest. These approaches are: a naive approach described in section 4.3. The second approach is a static approach that is explained in section 4.4. Section 4.5 provides the conceptual design of the mobility approach that we designed. Some basic concepts in section 4.2 are provided to clarify the common design characteristics of the three approaches.

4.2 Basic Concepts

The dissemination approaches use the server as the core component to disseminate a message to the mobile nodes. However, the server needs to know how many mobile nodes are located in the area of interest to make sure that the message is transmitted to all of them.

These basic concepts are shared among the three message dissemination approaches. The characteristics of each approach are separately discussed in later sections.

4.2.1 Area of Interest

It is required to know the location of the mobile nodes that are registered to the server. Since, the mobile nodes that are located within the boarders of the area will deliver the message. Any mobile node is located outside the area will not be allowed to deliver the message. The dynamic nature of the mobile nodes allows them to move freely entering and exiting the area of interest. Thus, it is required to know the location of the mobile nodes before message transmission occurs. In this case, the mobile nodes register their information to the server to provide the server with a sufficient knowledge about the expected number of mobile nodes who will deliver the message.

4.2.2 Nodes Subscription

Mobile nodes register themselves to the server to give the server the necessary knowledge about the expected number of mobile nodes in order to deliver the message to all of them. The registration process of the mobile nodes to the server is to make sure that the message is delivered by all the mobile nodes that are located in the area of interest which eventually fulfill our goal which is message delivery by all the mobile nodes.

Figure 4.2 explains the process of subscribing the mobile nodes to the server. The server initiates the process of mobile nodes subscription. The server sends a request message to

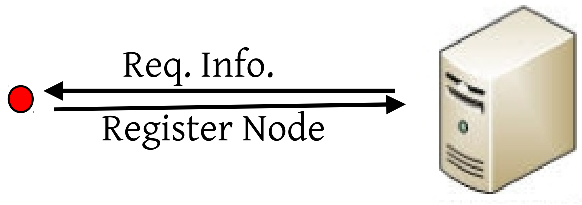

Figure 4.2: *Node's Subscription*

identify all the mobile nodes that are located in and near the area of interest who have the possibility to enter it. The mobile nodes provide their position using GPS. The knowledge of the number of nodes that registers to the server determines the maximum number of nodes that is expected to deliver the message. The server requests the node id and the initial positions of the mobile nodes. The mobile nodes register their information on the server before message dissemination takes place.

The following parameters are provided at the server when the mobile nodes are registered.

Symbol	Parameter
client.number	An incremental number assigned to the nodes when they register at the server
client.xCord	An initial x coordinate of the node
client.yCord	An initial y coordinate of the node

Table 4.1: Node Subscription Parameters

The subscription process is the same in the three different approaches. Thus, the process of calculating the amount of energy that is consumed on the mobile nodes is the same in the three approaches because of that we don't add the energy that is consumed in the subscription process in our calculation.

4.2.3 Message Transmission and Acknowledgment

Mobile nodes registration provides the server with the exact amount of mobile nodes that are located in the area. According to that, the server can send messages expecting that all the mobile nodes that are registered on the server deliver the message.

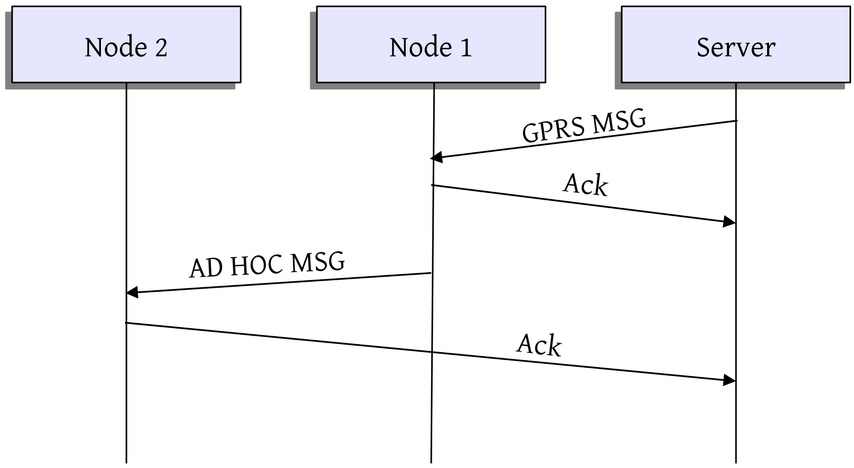

Figure 4.3: *Message Delivery and Acknowledgment*

However, the server needs to know if a mobile node delivered the message. An acknowledgment is sent to the server after a mobile node delivers a message indicating a successful reception of a message. The number of acknowledgment that is received at the server indicates how many mobile nodes delivered the message.

Figure 4.3 describes two methods on how messages are transmitted. The nodes send an acknowledgment to the server indicating the delivery of the message. The first node sends an acknowledgment after receiving the message by GPRS. Whereas, the second node receives the message through an ad hoc broadcast. Then, it sends the acknowledgment to the server.

4.2.4 Time Out Clock

Message transmission takes place until a specific time is reached. At this point in time, message broadcast is stopped. The server checks how many nodes have delivered the message through the number of received acknowledgments. The server sends a message requesting the position of the mobile nodes who didn't deliver the message. In case the mobile nodes are located within the borders of the area of interest, they send the server their position. The server sends the message to the mobile nodes so that they deliver it after the server receives the position of the mobile nodes.

The functionality of the clock indicates the right time to stop message transmission. This time is defined by the application according to the analysis of the performance of the dissemination protocol taking into account the broadcast time of message transmission because the density of mobile nodes differs in an area in the early morning from late at night.

4.3 The Naive Approach

4.3.1 Naive Approach Concept

The naive approach uses a direct unicast communication to send a message to all the mobile nodes that are located in the area of interest. Unicast communication uses GPRS or UMTS to transmit a message from the server as in figure 4.4. The cost of transmitting a message through one of these methods is very expensive in term of energy consumption. Although, the delivery rate using the naive approach is 100% when the mobile nodes are located in the area of interest at the transmission time. However, our interest is to decrease the energy that is consumed on the mobile nodes. Figure 4.4 explains message dissemination in the naive approach.

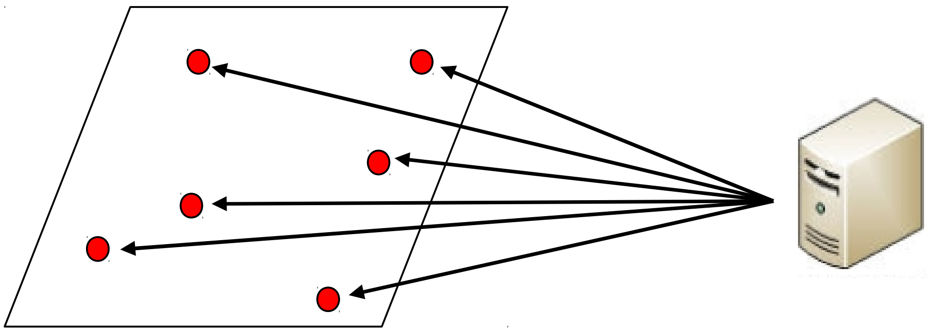

Figure 4.4: *Naive Approach Message Dissemination*

The naive approach considers every mobile node as a separate entity. The benefit behind considering every mobile node as s separate entity is that every mobile node has different characteristics such as their speed and direction. Message delivery using the naive approach is the best practice to address the issue of the environment heterogeneity because the only thing is considered in the naive approach is the position of the mobile nodes regardless of any characteristic that the mobile nodes displays.

On the other hand, the mobile nodes don't take advantage of their ability to communicate and exchange information with each other. Since, they communicate directly to the server using GPRS. The context-aware information that is gathered from the surrounding environment is not used to facilitate message dissemination among the mobile nodes. The disadvantage of using GPRS over ad hoc communication is that energy consumption on mobile nodes increases.

The process of message dissemination in the naive approach starts when the server requests the mobile nodes position in the area of interest. Every mobile node that is located in the area of interest provides its location. The server transmits the message to all the nodes that provided their positions. The mobile nodes confirm message delivery by sending an acknowledgment to the server. (see section 4.6.1 for naive approach scenarios).

4.4 The Static Approach

The naive approach uses GPRS only for message transmission. On the contrary, the static approach uses both GPRS and ad hoc communication to disseminate a message in the area of interest. In addition, the context-aware information that is gathered from the surrounding environment becomes more valuable for knowing the location of the mobile nodes related to other mobile nodes' locations. Since, mobile nodes use their locations to form neighborhoods where they can disseminate messages. Beside, the static approach decreases the relative total energy consumed compared to the naive approach because of the use of ad hoc communication.

4.4.1 Static Approach Concept

The idea of the static approach is to get the number of mobile nodes that are in the ad hoc communication range of another mobile node. Figure 4.5 explains the process of message dissemination in the static approach.

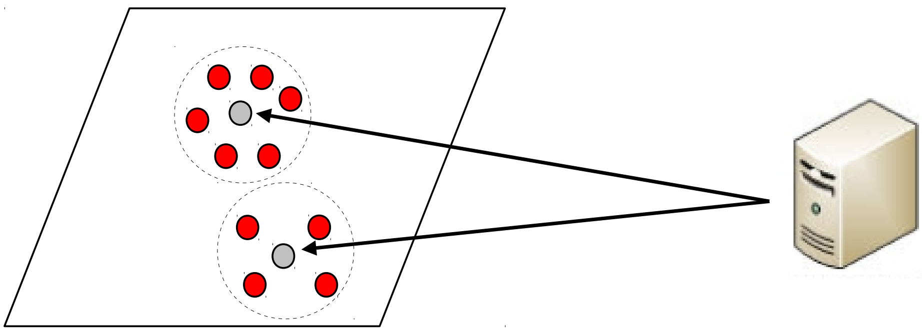

Figure 4.5: *Static Approach Message Dissemination*

The mobile nodes are grouped into circles that cover the mobile nodes that are located in their communication range. At the beginning, the mobile nodes are sorted descending from the mobile node that covers the maximum number of neighboring mobile nodes to the mobile node that covers the least number of mobile nodes. A message is sent from the server using GPRS to the first node in the list after sorting the mobile nodes. The mobile node delivers the message and broadcast it to the neighbors that are within its communication range. The mobile nodes that deliver the message are removed from the list. The processes of message delivery continues until all the mobile nodes at the area of interest are covered either by ad hoc or GPRS communication.

The aim of the static approach is to minimize the number of circles that cover the mobile nodes which means minimizing the number of GPRS message transmission that is required. Figure 4.6 is an example of the message dissemination of static approach. The mobile nodes are covered with four circles. Each circle contains a number of mobile nodes that are covered with ad hoc communication. The red mobile nodes are located within the communication range of the gray nodes. The gray mobile node receives the message using GPRS and broadcast it to the neighbors.

There are more circles than the four that are depicted in figure 4.6. Each node has its own communication range and it can form a circle. The minimum number of circle that cover all the nodes at the time of the broadcast is four. Minimizing the number of circles means minimizing the number of messages transmitted using GPRS.

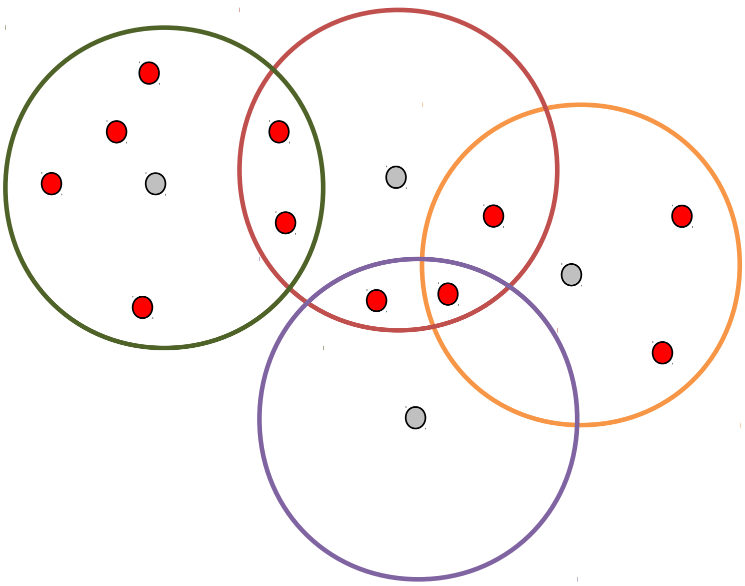

Figure 4.6: *Static Approach Message Dissemination Example*

The number of mobile nodes in figure 4.6 is 15 mobile nodes. They are distributed and covered by four circles. The green circle contains 6 nodes, the red circle contains 5 nodes, the orange circle contains 4 nodes and the purple circle contains 2 nodes. The gray nodes are not counted part of the circles because they deliver the messages when they receive it through the GPRS.

The following steps occur to cover the mobile nodes in the example:

1. Sort the nodes based on the maximum number of nodes that a circle can cover.
   Green: 6, Red: 5, Orange: 4, Purple: 2.
2. Send a message using GPRS to the gray node in the first circle in the list. Broadcast the message and cover the shared elements.
   Red: 3, Orange:4, Purple: 2.
3. Repeat step 1: Orange: 4, Red: 3, Purple: 2.
4. Repeat step 2: Red: 1, Purple: 1.
5. If the elements are equal, send the GPRS message to any gray node.
6. Finally, we are left with one gray node that receives the final GPRS message.

The static approach considers all mobile nodes in an idle state at the time when the message is broadcasted. The reason behind this is to be able to identify which mobile nodes communication range cover more nodes in the area. The static approach broadcast a message using ad hoc communication only once for every mobile node that forms a circle.

The static approach is based on the idea of the set cover problem. A classical problem in complex theory and one of karp's 21 NP-complete problems [27]. It is implemented in the static approach to minimize the number of nodes that delivers the message using GPRS. Also, it makes use of ad hoc communication in broadcasting the message to the neighboring nodes. Message broadcasts reach more nodes than sending unicast GPRS messages. The static approach benefits from the set cover problem by maximizing the number of nodes that is covered each time a message is broadcasted.

4.4.2 Set Cover Problem

Definition: Let U defines a universal set where |U| = m. A family F is a subset of U where F $\subseteq 2^U$. A sub collection of elements within the set (S $\in$ F) covers the universal set U, if

$$\mathrm{U} = \bigcup_{S \in F} S$$

The aim is to find a minimum size subset such that $C \subseteq F$ which covers all the elements in the universal set U.

$$\mathrm{U} = \bigcup_{S \in C} S$$

The following example illustrates the above definitions:

U = {0,1,2,3}

F_0 = {0,1} , F_1 = {1,2} , F_2 = {2,3} , F_3 = {0,3}

You can clearly see that the union of the sub collection C that belongs to the families gives the universal set U. Although, you can achieve the universal set U by combining only the families F_0 and F_2 or F_1 and F_3. The cardinality of the minimum cover set equals 2. The result is minimized to only 2 sets but the result is not unique.

Set cover problem is used in several important applications in the networking environment. In optical networks [34], the set cover problem is used to construct a spanning tree minimizing the number of wave lengths available to be used on each edge. Another application is spreading wireless antennas in a specific area. This needs planning the antennas geographic distribution in order for the antennas to cover a larger geographic area.

IBM experiment [38] used set cover problem to find computer viruses. The number of viruses which was known around 5000. The number of sets consisted of 9000 substrings of 20 or more consecutive bytes from viruses. It was enough to search 180 substrings to find known computer virus.

Our focus is to apply the set cover problem to disseminate a message to the mobile nodes that are located in an area of interest. Taking into consideration the distance and position of the mobile nodes in the geographic area.

4.4.3 Sensor Cover Problem

Sensor Cover Problem [13] extends the idea of set cover problem by considering the points that are positioned in the xy plane. Sensor cover problem describes the duration of the time needed to maintain cover of the region. Also, it describes covering for one and two dimension. We are interested in the two dimension that contains overlapping of the arbitrary set as in figure 4.7.

Figure 4.7 shows how many arbitrary sets are needed to cover the mobile nodes in the plane. However, the problem is that in the set that contains only one node would be cost

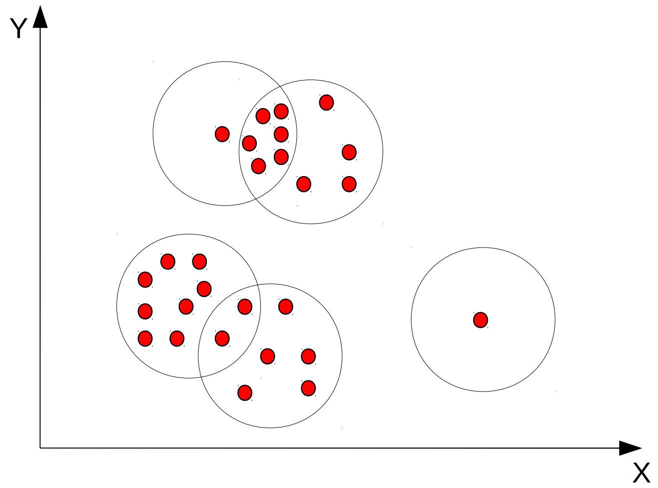

Figure 4.7: *Sensor Cover Problem*

consuming trying to cover the region. Our interest is to maximize the overlapping between the sets. The two sets that shares six mobile nodes in their intersection are more energy efficient than the one set that contains one mobile node. Since, the two sets cover 11 mobile nodes and only two of them are doing the process of covering while the other 9 only receive a message when broadcast occurs.

Set Cover Problem Complexity

Set Cover problem is one of the NP-hard problems. It is difficult to find a solution for solving such a problem in a polynomial time. One could find a solution if P = NP [19]. The running time of the NP-hard problems can take up to thousands of years, so it is assumed that there is no efficient solution for these kind of problems. This leads to end up finding an approximate solution to the optimal solution. An approximate solution for set cover problem is a greedy algorithm that finds a solution in the polynomial time which gives a near optimal solution for every instance. The approximation solution using the greedy algorithm achieves $O(\log_e n)$ [26] in comparison to the optimal solution. In general, there is no other approximation algorithm in polynomial time achieves a better result unless P = NP.

4.4.4 The Greedy Algorithm

In the previous section, a set cover problem was described. We showed that it is NP-hard problem which means no optimal solution is found in a polynomial time. The idea is to find an approximation of the optimal solution.

An attempt to find an approximate solution is to use a greedy algorithm. It was first presented by Johnson [26]. In his paper, he examined polynomial time heuristic approximation algorithms for the NP-Complete problems.

Definition: Greedy Algorithm picks the subset $S \in F$ that contains the largest number of elements. It does that at each step until all the elements in the universal set U are covered.

Algorithm 1 describes the greedy algorithm and how it works. The algorithm begins with an empty set C in order to cover all the elements in the universal set U. We repeat the steps at each time until all the elements in the universal set U are covered. A set S is selected from F which covers the largest number of elements in U. U subtracts the shared elements that are contained in S. Then, C adds the covered elements from S to its set. The algorithm terminates when all the elements in U are covered. We are interested to compare the results that are obtained from this approximation in comparison with the optimal solution.

Algorithmus 1 *The Greedy Algorithm*

Input: Universal set U, Family F of subsets of U.
Output: Set cover C.

$C \leftarrow \emptyset$
while $U \neq 0$ **do**
  Select S $\in$ F that maximizes $|S \cap U|$
  $U \leftarrow U - S$
  $C \leftarrow C \cup \{S\}$
end while
return C

Assume the universal set contains 10 elements and these elements are distributed over 5 different family sets. This examples explains how the greedy algorithm works.

U = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} , F = {S_1, S_2, S_3, S_4, S_5}

S_1 = {1, 2, 3, 6, 7, 9} , S_2 = {1, 2, 3, 4, 5} , S_3 = {4, 5, 7} ,

S_4 = {5, 8, 10} , S_5 = {6, 7, 8, 9, 10}

Figure 4.8 shows the above example for easier analysis.

The choice to cover the elements that are contained in the set depends on the greedy algorithm. First, the greedy algorithm chooses to cover the set S_1. The elements that are shared between S_1 and the other different sets are removed from these sets. After covering the elements in set S_1, the remaining sets contain:

S_2 = {4, 5} , S_3 = {4, 5} , S_4 = {5, 8, 10} , S_5 = {8, 10}

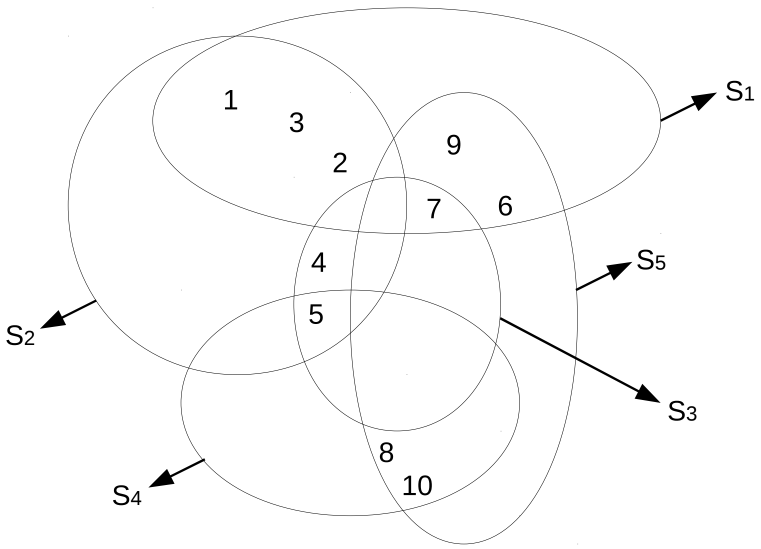

Figure 4.8: *A Greedy Algorithm Set Cover Example*

Again, the algorithm chooses the set that contains the largest number of elements which is S_4. When the elements in S_4 are covered, S_5 is automatically covered since $S_5 \subset S_4$. The remaining S_2 and S_3 contains the same element which is 4. Pick any set out of them to cover the last element in the universal set.

The optimal solution covers all the elements in the universal set uniting S_2 and S_3. The size of the optimal solution is 2 which is compared to the approximate solution that needs a size of 3. The approximate solution combines S_1, S_4, S_2 or S_3 to cover all the elements in the universal set.

The greedy algorithm implements the message dissemination under the static approach. The example in figure 4.8 explains the concept of message dissemination delivered by the mobile nodes that are located in the area of interest. Since, the optimal solution is difficult to achieve. The greedy algorithm is the best approximation to cover all the mobile nodes in the area of interest. The main goal behind the greedy algorithm is to minimize the number of sets that covers the mobile nodes which leads to minimize the number of GPRS message transmission through.

4.5 The Mobility Approach

The static approach considers mobile nodes in a constant state when message dissemination occurs. The advantage use of mobile approach is that it uses the fact that mobile nodes move in the area of interest which cover more segments. This process can help to further decrease the number of messages that is delivered using GPRS communication because message is disseminated using ad hoc communication.

4.5.1 Mobility Approach Concept

The mobility of the nodes can cover more segments than a node in a static position that doesn't move. It will help to increase the chance that a node get in contact with other nodes in its travel which will give the nodes a chance to exchange messages. Thus, our mobility approach depends on the knowledge of both:

- The characteristics behavior of the mobile nodes
- The connected segments that form the area of interest where the mobile nodes move

A subset of mobile nodes is chosen at the beginning to deliver a message using GPRS. The concept of choosing this subset differs from the static approach. The server makes a decision about which mobile nodes receive the message. However, the server needs to acquire some information about the mobile nodes movement and the segment on which the mobile node is located. The knowledge of the node's mobility and behavior provides more chances to predict the direction and the next position of the node.

The geographic location of the mobile nodes is the major factor in determining the subset of the mobile nodes. The knowledge of the segments that are given in the server help in determining whether the nodes are located on the same segment or they are located away from each other. Figure 4.9 shows an example of the connected segments.

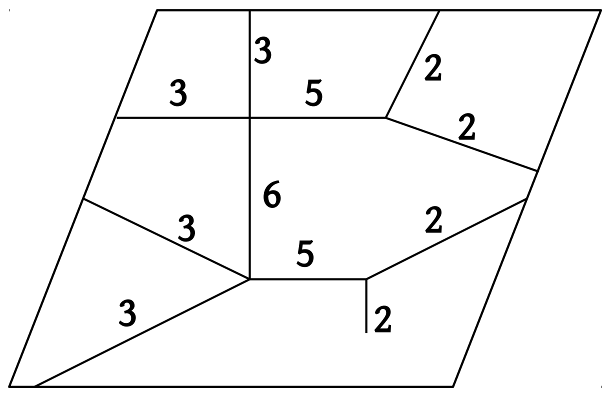

Figure 4.9: *Connected Segments of the Area of Interest*

The number that is located near each segment indicates the number of segments that is connected to that segment. The server benefit from the knowledge of the connected segments to formulate the concept of message transmission when mobile nodes are located on these segments. The segments are sorted in a list according to the number of connected segments from the higher number of connected segments to the lower number of connected segments. The priority is to send a GPRS message to a mobile node that is located on a segment that is connected to larger number of segments. First, the server requests the mobile nodes positions. It checks the mobility of the nodes because the GPRS message transmission is sent to a mobile node that is in continues movement. Static nodes will receive the message through ad hoc communication or through GPRS when message broadcast stops.

After obtaining the position of mobile nodes, the server compares these positions with the segments that are arranged descending in the list. If there is a mobile node that travels on the first segment in the list, the priority is to send a GPRS message to that mobile node. The connected segments to the segment where the mobile node is located are covered and no other mobile node is allowed to receive a message through GPRS unless the message broadcast is stopped and a GPRS occurred to transmit a message to those nodes who didn't deliver it. Figure 4.10 explains the example of message transmission under mobility approach and how segments are covered.

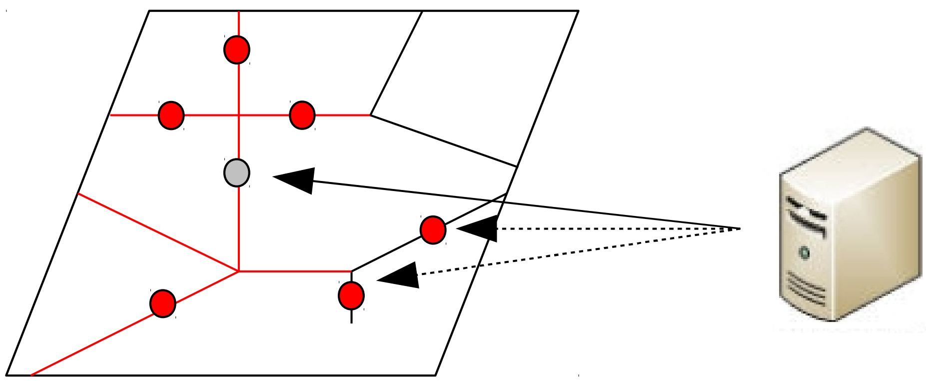

Figure 4.10: *Message Dissemination and Segment Cover*

The server checks the mobile nodes movement and on which segments they are located. First, the gray mobile node is located on a segment which is connected to six other segments. The server sends a message using GPRS to a mobile node that is located on a segment that has more connected segments. Since, the aim is to decrease the number of GPRS transmissions. The mobile nodes that are located on the six connected segments including the segment where the mobile node received the GPRS message are marked in red. Mobile nodes that are located on these segments are not allowed to receive a message through GPRS unless the broadcast message stops and the mobile nodes didn't deliver the message using ad hoc communication. The second step would be choosing another mobile node that isn't located on the covered segments. There are two possibilities to transmit a message using GPRS to one of the mobile nodes that are indicated in figure 4.10. One mobile node receives the message through GPRS and the second one receive it through ad hoc communication. Then, the segments are covered in order not to allow any mobile node receive the message through GPRS on these segments.

Although, we are interested in transmitting a message using GPRS to the mobile node that is located on a segment that is connected to the higher number of segments. However, the distribution of mobile nodes in the area of interest might differ at the time of the broadcast. The mobile node might be located on a segment that is not connected to the higher number of segments as in the example in figure 4.11. The mobile nodes who receive the message using GPRS are different than those who receive the message in figure 4.10.

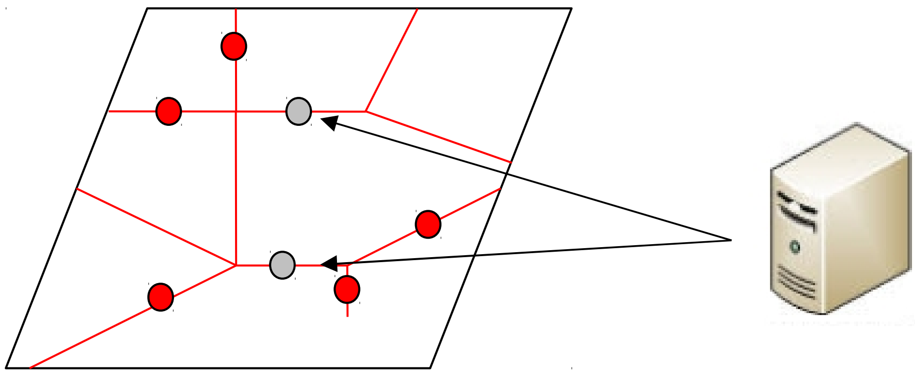

Figure 4.11: *A Variant Message Dissemination and Segment Cover*

The gray mobile nodes are located on segments that are connected to five other segments. Although, there is a segment that is connected to six segments but there is no mobile node located on that segment. Figure 4.11 is to show that message transmission using GPRS to a subset of nodes depends on the mobile node movement in the first place. Then, it depends on the structure of the connected segments that forms the area of interest.

The choice of the segment where the mobile node is located depends on the idea of the matching problem. Matching problem is used to compare and cover the connected segments to the one that the mobile node is located on. The next section describes how the matching problem is used to decrease the GPRS messages by allowing the mobile nodes move and disseminate a message using ad hoc communication.

4.5.2 The Matching Problem

The purpose of the matching problem is to identify the connection between the segments that form the area of interest. The matching problem provides the server with a knowledge about the connected segments. In this sense, the server identifies which mobile nodes are located near each other on which segments. This process will help to reduce the number of GPRS transmission to the mobile nodes that are located on the connected segments. It is possible by allowing only one GPRS transmission to a mobile node located on these connected segments. The rest of the mobile nodes that are located on the connected segments will receive the message through ad hoc communication because the communication range of one mobile node will cover the connected segments. Thus, it won't be a necessary to send several GPRS messages to mobile nodes that are located near each other.

Definition: Let G=(V,E). A set M $\subset$ E is matching when two edges in M are not adjacent and do not share the same vertex.

Definition: A maximal matching is a matching in a graph that contains a minimum number of edges.

Definition: A maximum matching is the largest number of edges a matching can contains.

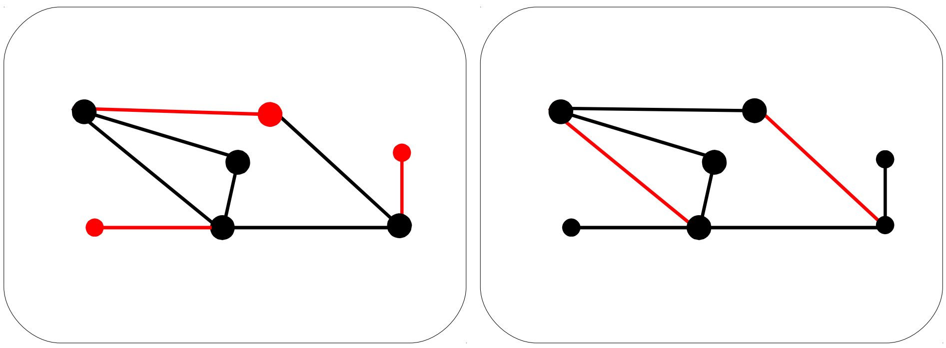

Figure 4.12: *Maximum vs Maximal matching in a Graph*

Figure 4.12 shows the difference between the maximum and the maximal matching edges. The red edges are the matching edges. The left side of figure 4.12 shows the maximum number of edges that can be obtained from the matching graph which is 3. However, the right side of figure 4.12 gives the maximal matching of the graph with only 2 edges.

The application of the matching problem is crucial in our approach. The difference between the maximum and the maximal matching is very important regarding GPRS transmission because the maximal edge matching reduces the number of edges. The mobile nodes that are located on the maximal matching edges receive the minimum number of GPRS. On the contrary, mobile nodes that are located on the maximum matching edges receive the maximum number of GPRS. The aim is to reduce the number of GPRS messages.

Algorithmus 2 *Maximal Matching* (G, V, E)

```
M ← ∅
while ∃ E to add to M do
   Select E that has no shared vertex in M
   M = M ∪ E
end while
return M
```

A maximal matching of a graph can be achieved using a greedy matching algorithm. The algorithm implies an incremental edge addition. This edge has no shared vertex with the matching edges in M. The process continues until all edges are covered. The maximal matching can be found in a polynomial time. For further information about matching theory refer to [30]. The maximal number of edges matched means the smallest number of possible GPRS transmissions.

Matching segments are formulated based on the mobile nodes locations. The set of matched segments is formed when the server requests the position of the mobile nodes and on which segment they are located. The optimal solution is to achieve maximal matching segments. In case this is not possible, an approximation of the matching segment is given according to the mobile nodes location.

Matching problem has multiple applications. Construction of high speed switches [29] need buffer management in order to avoid packet loss. Assume there are multiple packet destined to the same output. A contention will occur at the output port. The use of matching algorithm will organize the packets transmissions.

The matching segments where the mobile nodes are located is the first part that covers our approach. The second part depends on the path that the mobile nodes can travel. The probability that a mobile node travels a path reduces the number of GPRS transmissions.

4.5.3 Path Probability

The matching problem described in the previous section gives an optimal solution on the number of segments that are covered. However, the total segments that are covered depend on the nodes' locations and their characteristics behavior. Usually, the nodes are distributed across the area. Some of these nodes are static and the others are in a continues movement. Our aim is to deliver the message through GPRS to a node that is in the move which increases the chance of a node to contact other nodes while moving in the area.

Beside, more areas are covered when a node moves.

Our interest is to cover the maximal matching edges where the nodes are located which will reduce the number of GPRS transmission to the mobile nodes that are located on these edges. This will give a chance for ad hoc message dissemination to take place in order to achieve the maximum benefit out of the nodes movement. The server screens all the nodes after receiving their previous and current positions. Each node will be located on one of the segments. An algorithm sort the segments descending based on the number of connected segments to the segment where the node is located. It means the node has several path possibilities to travel. In this case, the probability of each path the node travels decreases.

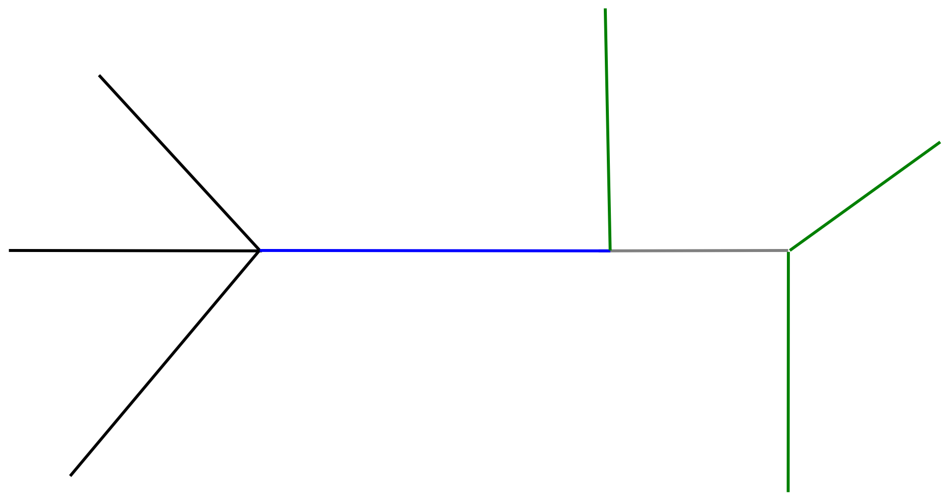

Figure 4.13: *An Example of Paths Probabilities*

Figure 4.13 provides an example on paths probabilities that each node can travel. The segments are colored to differentiate the number of the neighboring connected segments. The nodes have different probabilities in choosing which segments it travels. According to figure 4.13:

- The blue segment is connected to 5 segments. A node has a path probability equals 0.20.
- The gray segment is connected to 4 segments. A node has a path probability equals 0.25.
- The black segment is connected to 3 segments. A node has a path probability equals 0.33.
- The green segment is connected to 2 segments. A node has a path probability equals 0.50.

An increase in the number of connected segments lower the probability of the path that a mobile node can travel. However, choosing the node that has the least path probability means a decrease in the number of GPRS transmission required because of the matching problem which covers the connected segments. It doesn't allow a node that is located on the matched segments to receive a message through GPRS. According to the example in figure 4.13, the node that lies on the blue segment and moving in any direction is chosen to receive a message through GPRS. The five segments that are connected to the blue segment are covered according to the matching problem. Another GPRS is required to transmit the message to one of the two green segments at the right hand side to complete the matching segments.

A GPRS transmission to the nodes that are located on the blue segment and one of the two connected green segments means a maximal matching is achieved. The solution that is achieved would be an optimal solution to decrease the number of GPRS transmission.

The number of message transmission through GPRS increases when choosing to send a message to a node that is located on a segment that is connected to less number of segments. According to the above example, choosing the vertical green segment which covers the blue and gray segment to send a GPRS message to the node that is located on it yields the need to send two other GPRS messages. These GPRS messages are sent to the nodes that are located on one of the green segments at the right hand side and one of the black segments.

Path probabilities are tightly connected to the matching problem. The choice of the nodes that has less paths to travel decreases the number of the matched segments. Our aim is to increase the number of the matched segments that are included in our set in order to decrease the number of GPRS messages. This is achieved by choosing a higher path probability a node can travel. The number of connected segments that is included in the matched sets is increased. Then, the message dissemination using GPRS decreases.

4.5.4 Optimized Flooding

Mobility approach applies a control broadcast mechanism in order to limit the number of mobile nodes that floods the neighborhood. The process of checking mechanism applied in the mobility approach to check whether a node delivered a message consumes energy. The mechanism applied in the mobility approach allows the static mobile nodes broadcast only once. Any mobile node in the communication range receives the message. A continues broadcast in the same spot is not effective because the mobile nodes in that area received the message. Mobile node movement and broadcast is more effective because there is a possibility that the mobile node comes across others who didn't deliver the message.

Figure 4.14 has similar segment structure as figure 4.13. The node that is located on the blue segment travels to the right. The communication range of this node covers the node that is located on the vertical green segment. Although, the node travels toward the gray segment but the nodes that are located near the edges and lie in the communication range of other nodes can deliver the message. The node that is located on the black segment can get in contact with another node when it moves around the area or a node comes in contact with it. In case the node didn't deliver the message before the broadcast stops. It receives the message through GPRS.

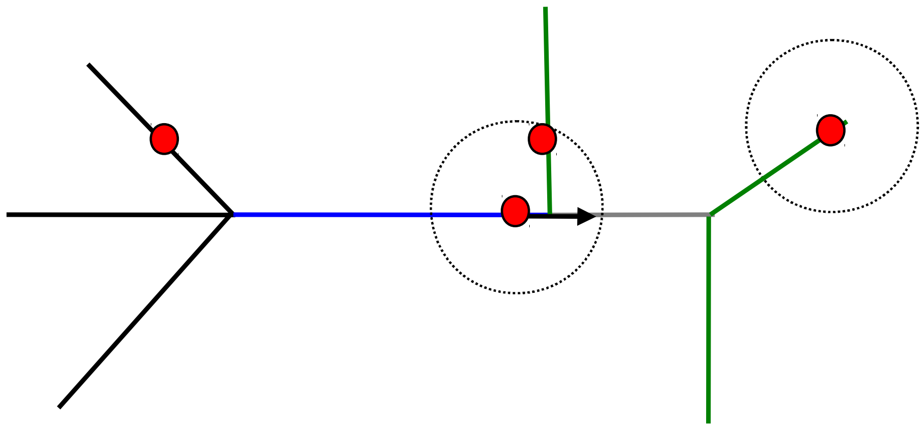

Figure 4.14: *Communication Optimization*

The mobility approach is designed to allow every mobile node broadcasts the message after delivering it. At a specified periods of time the mobile nodes broadcast the message. The problem is the number of messages broadcasted increases in the network. The checking mechanism that is applied allows the mobile nodes saves resources through sending a check message that has a small size compared to the actual message that may contain video content with large sizes. The mobile node that is located on the end of the green segment in figure 4.14 keeps sending a checking messages to discover any mobile node that didn't deliver the message. That mobile node broadcast a message in the neighborhood once it found other mobile nodes that didn't deliver the message.

Some mobile nodes are static that don't move from their locations. It would be a waste in nodes' resources if they are allowed to keep checking their neighborhood. The mobility approach is designed to make use of the movement of mobile nodes to cover more segments in the area of interest. The mobile nodes that are located in the same positions are allowed to broadcast one time. The concept behind it is to allow the mobile nodes that are in the communication range of the static mobile node deliver the message (see OFP in section 2.3.1). The nodes that are in continues movement are allowed to check for available nodes that didn't deliver a message every specified period of time. This process decreases the number of messages which is broadcasted. It is an efficient process in saving nodes' resources and decrease the relative total amount of energy that is consumed.

4.6 Message Dissemination Scenarios

The previous sections described the concept of message dissemination approaches. Each approach has different scenarios according to the concept that the approach is built on. Also, the movement of the mobile nodes in the area of interest at the transmission time of the message gives different extension on how message is disseminated. These scenarios are described in the following subsections.

4.6.1 Naive Approach Scenarios

The high mobility environment affects the reliability of a message delivery in the naive approach. Although, a message is sent directly to the nodes without any intermediate carrier. However, the problem is that not all the nodes that are subscribed to the server are located within the area of interest at the time of a message transmission. There might be several nodes reside out side the map. After transmission takes place some nodes enter the area.

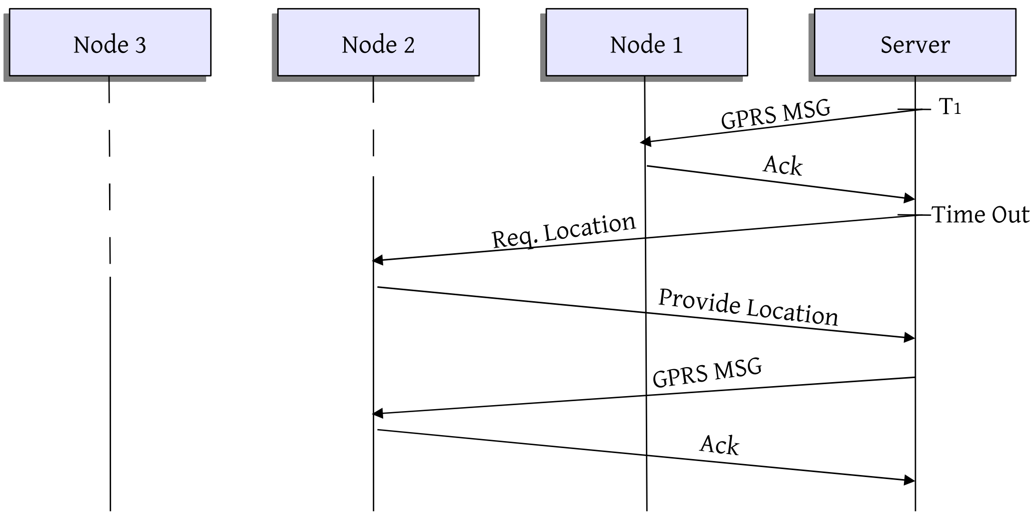

Figure 4.15: *Naive Approach Scenarios*

The nodes in this approach don't use ad hoc communication to broadcast the message.

These nodes have to wait until the time out clock signals and a check occurs to determine message delivery by the nodes.

The example in figure 4.15 contains three mobile nodes. Each of these three nodes enters the area at a different time instances. The continuous line emphasizes that the node is located in the area. At the time of transmission T_1, the first node is located in the area of interest from the beginning. It delivers the message and replies with an acknowledgment. The second node enters after the transmission T_1 occurs, but it stays in the area of interest until the checking time takes place. The server checks the nodes location who have not delivered the message and sends the message to them. The server knows the mobile nodes who delivered the message from those who didn't deliver the message from the acknowledgment that the mobile node sends after delivering the message. The main difference between requesting nodes' location information at the time out clock from requesting the nodes' location information at the beginning as in figure 4.2 is that the number of nodes at time out is different. The energy that is consumed in calculating the position of the mobile nodes at the time the clock times out is included in the calculation of the total amount of energy that is consumed. The number of nodes that don't deliver the message before the time out clock reached is different among the three message dissemination approaches because of the introduction of the ad hoc communication in the static and the mobility approaches.

The third mobile node enters the area of interest after the message is disseminated and the time out clock is reached. The delivery of a message to this mobile node is impossible. It can receive a new message when the broadcast takes place again if it stays in the area of interest.

4.6.2 Static Approach Scenarios

The static approach considers the constant mobile nodes at the time the message is broadcasted. Despite this consideration, the mobile nodes' high mobility within the network might affect the virtual forming of the sets. The virtual sets are formed for a specific time instance. If the nodes are in a continuous movement, this may lead to an increase of the number of messages sent through GPRS because there would be an increase in the number of virtual sets. Figure 4.16 explains how the nodes movements increase the number of GPRS.

The two gray nodes receive the message through GPRS. Then, they broadcast the message to the neighbors covering all the nodes that are in the area. However, a node from the orange set moves outside the communication range of the gray node just before the transmission of GPRS occurs. The message which is broadcasted isn't delivered by that node.

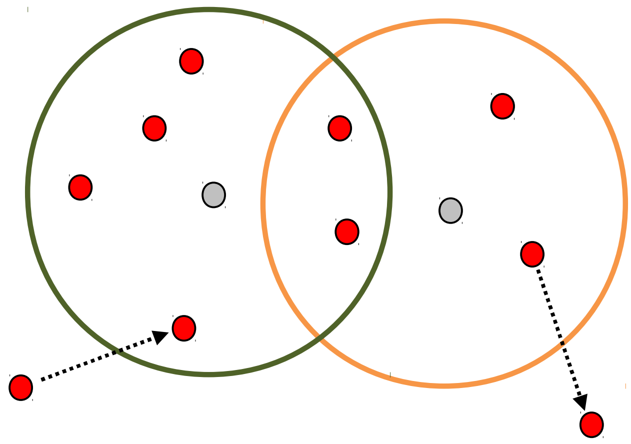

Figure 4.16: *Nodes Mobility in Virtual Sets*

There is a need for a third GPRS message to be delivered by the node that left the orange set. The node forms a new set by itself which has the capability of broadcasting the message to any node in it's communication range. The need for another GPRS transmission increases the total amount of energy consumed.

The node movement outside the orange set not necessary means that the node stayed within the area of interest. The node may move out side the borders. In this case, there is no need for another GPRS transmission. Its movement doesn't effect the total amount of energy consumed. However, it affects the message delivery rate. The number of nodes which delivers the message decreases. Unless, the node enters the borders again before the time out clock is reached. When the time runs out, a GPRS transmission occurs and the node delivers the message.

In high mobility networks, the time delay of a message transmission must be decreased to the minimum. Otherwise, the configuration of the virtual sets is changed. Assume that the gray node moves, then the configuration sets in figure 4.16 are not correct. A new sets must be chosen to cover the nodes. If no new sets are chosen, then the number of GPRS transmissions increases. The benefit of using ad hoc communication to broadcast the message is not achieved.

Also, the nodes have the possibility to enter the communication range of another node similar to the node that enters the green set. The entrance of the node doesn't effect the virtual set configuration. Unless, there are several number of nodes that enter the borders. The issue of these nodes that they won't be all covered in the communication range of the two sets. This will lead to transmit another GPRS message to one of the mobile nodes that covers the rest of the mobile nodes in order to deliver the message to all the nodes in the area of interest.

4.6.3 Mobility Approach Scenarios

Mobility approach depends on the movement of mobile nodes within the area of interest. It is important to transmit a message to a node that moves within the borders of the area. It will be a waste of the nodes' resources if a message is transmitted to a node that travels toward the borders. The chances that a node get in contact with another node decrease.

Assume the node that is located at the end of the segment that has an arrow in figure 4.17 travels towards the borders. If there is another node that is located on the same segment and isn't in the communication range of that node. Then, the chances that a node receives the message through ad hoc communication decrease. In case, there is no other nodes carry the message come into contact with the node. Then, the node will receive the message through GPRS which will increase the number of GPRS messages transmitted when the time out clock indicates that the broadcast is stopped.

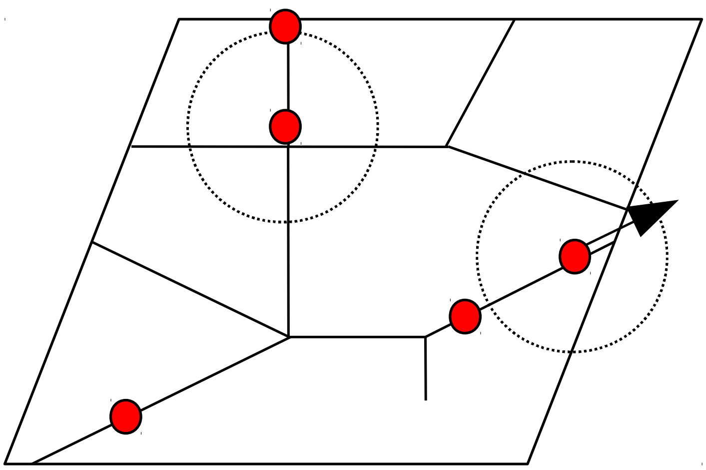

Figure 4.17: *Mobility Approach Scenarios*

Nodes that enter the geographic area can be considered as a candidate to receive the message through GPRS transmission. However, to consider a node as a candidate. It has to enter the area from a segment that is not already located in the set of the matching segments. Furthermore, the node has to continue moving in the same direction away from the borders. Although, the node can be considered as a good candidate but some time there are multiple nodes reside on the same segment. One of these nodes might be already chosen to receive the message through GPRS. The node that entered the area will receive the message when another node broadcasts it. The node in figure 4.17 that is located on the vertical segment enters the area from the above direction. However, there is already a node received the message through GPRS transmission. The node that entered the area of interest can now receive the message through ad hoc broadcast.

The advantage of the mobility approach over the static approach is that the nodes are moving around. The chances of delivering a message using ad hoc broadcast increase. Sometimes, there won't be a need for a message transmission using GPRS when message broadcast stops. The mobile nodes may cover the whole area through their movement.

The outcome of the mobility approach is reflected in reducing the number of GPRS transmissions because energy consumption for GPRS transmission is larger than ad hoc broadcast. For more information about energy consumption (see table 6.1).

4.7 Conceptual Design Summary

This chapter gave an overview about the protocol design concepts. The concept of three different approaches were discussed. A basic approach which is the naive approach. The second approach which is the static approach that considers the nodes mobility as static at the time of message transmission. The mobility approach depends on the nodes mobility property when broadcasting a message. The three approaches share some common properties among them (see section 4.2). Although, they share these characteristics properties but the design concept for each of them is different.

The naive approach uses only GPRS to transmit the message from the server to the nodes. The nodes that are within the borders of the map deliver the message. A request for every node position is required to transmit the message to that node. Despite that we consider static nodes at the time of message transmission. A control over the mobility of the nodes is impossible. The nodes are able to move outside the borders. Due to their mobility, the delivery rate is reduced. A consideration to the geographic location is not considered in the naive approach. Also, the structure of the map isn't important for message delivery in the naive approach.

The static approach uses both GPRS and ad hoc communication. This approach depends on the concept of set cover problem. A subset of nodes is virtually formed based on the number of nodes that is located within the ad hoc communication range of each node. Similar to the naive approach, the nodes are only considered static at the broadcast time. Only the node that forms the virtual subset uses ad hoc communication to broadcast the message to the neighbor nodes. The nodes mobility has an effect on how the virtual sets are formed to cover the mobile nodes that are located in the area of interest. The node that is not included in any set forms a set by itself which requires additional GPRS transmission to deliver the message.

The design of the mobility approach depends on two parts. The first part is the nodes' characteristics behavior. This approach benefit from the nodes' mobility which covers wider area and come into contact with multiple other nodes. The second part depends on the segment matching problem. The knowledge of the segments structure is important to avoid GPRS transmissions to the nodes that are located on adjacent segments. Since the nodes can move around, ad hoc broadcast might cover the rest of the nodes that are located on an adjacent segment to the node that receive a message through GPRS transmission.

In general, we try to reduce the number of GPRS transmissions to reduce the energy consumption on mobile nodes in the mobility approach. The number of ad hoc messages increases compared to GPRS. However, the use of optimized flooding and checking mechanism in ad hoc communication limit the number of the messages broadcasted. It is implemented to allow only the moving nodes to broadcast every specified time periods. The static nodes are limited to broadcast once to cover the nodes that are located in their communication range.

The following table provides a general comparison among the three approaches.

Description	Naive approach	Static approach	Mobility approach
Mathematical Algorithm	None	Set cover problem	Matching problem
Communication methods	GPRS	GPRS , ad hoc	GPRS , ad hoc
Considers node mobility at transmission time	No	No	Yes
# Ad hoc broadcasts	0	= # of minimum virtual sets	unlimited
# GPRS transmissions at broadcast time	= # of nodes within borders	= # of minimum virtual sets	Depends on nodes characteristics and matching segments

Table 4.2: Dissemination Approaches Comparison

Chapter 5

Implementation

The previous chapter presented the conceptual design behind our approach. It described the concepts on how our approach is implemented. This chapter presents the simulation environment which we used to implement our protocol. Section 5.1 describes the simulation tool that is used to design the protocol. An explanation about our protocol implementation is given in section 5.2. We will give an overview on the design of the classes that the protocol consists of.

5.1 Simulation Environment

The Design of a real system needs a lot of time to analyze and understand the system model. Simulation tools try to simulate the behavior of the real system. It is a matter of designing a model and try to test the result before constructing it. The cost of simulating an environment is less expensive than trying to test it on real components. There are several tools available to simulate the network communication such as OMNeT++ [7] and OPNET [8].

We implemented our protocol using network simulator (version 2) (ns-2) [9]. Ns-2 [25] is a discrete event driven network simulator used for the purpose of studying the dynamic nature of network communication. The simulation supports the research in the network area. It helps to design new models and provides a comparison of the different kind of protocols. Also, it provides support for wireless and wired communication networks. In our case, it is used to simulate the dynamic nature of network topologies by focusing on ad hoc mobile communication. Section 5.1.1 explains the architectural design of ns-2.

5.1.1 Ns-2 Architecture

Ns-2 consists of two languages: C++ and Object-oriented Tool command language (OTcl). The use of the two languages depends on the distinct requirements of the simulator. The simulator needs to operate over large data sets where the simulation run time speed is important because of that C++ is part of ns-2. However, the OTcl is flexible for fast changing the configuration environment scenarios compared to C++. It is easier to code in OTcl but it runs slower than C++.

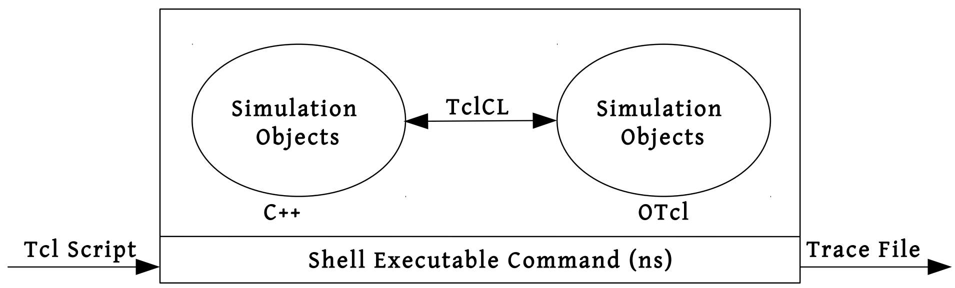

Figure 5.1: *Basic NS Architecture*

Figure 5.1 [25] shows the basic architectural structure of ns-2. The shell executable command receives the configuration of the simulation in a tcl script file. This tcl file contains a simulation configuration that will setup the simulator with a scenario. The simulation generates an output trace file that can be used to plot the results or to create an animation to the scenario. The analysis of the trace file is interpreted using NAM (Network AniMator) or XGraph. The TclCL that binds C++ and OTcl is an interpreter interface which is discussed in section 5.1.4.

5.1.2 C++

C++ classes form the backbone and the kernel of the ns-2. It is used to manipulate bytes and design protocols. Furthermore, it is used to process packets. We need to recompile the code each time we modify the code in the c++ classes. It is efficient, powerful and fast for run time speed of protocol simulation. However, it is slow for editing, adjusting and manipulating the code.

The standard ns-2 contains set of functionality and objects. These objects are available within the ns-2 to be integrated and implemented in designing protocols. In addition, a custom implementation of additional C++ objects can be written to extend the functionality whenever there is a need to simulate an unimplemented object.

Ns-2 includes several objects that are implemented using C++ such as agents. Agents are implemented in ns-2 and attached to the nodes. Their functionality is to create, destroy and route packets. Ns-2 contains several agents such as TCP agent and UDP agent. However, we extended the implementation to include our own agent. We implemented a message dissemination protocol agent (MDPAgent) that we are going to use in our protocol. The next section describes the second language that the ns-2 consists of which is OTcl.

5.1.3 OTcl

OTcl is an object-oriented programming language. It is an extenstion to the Tcl (Tool Command Language). The OTcl is integrated into the ns-2 to act as the front end. It is used for faster configuration of scenario files and parameter changes. Although, it is slow in the run time but the code can be adjusted fast according to the scenarios that we are seeking. In addition, it is flexible which allows it to be integrated with many languages and deployed on different platforms.

The user can interact with the ns-2 simulator to build topologies and scenarios using the OTcl. OTcl is an interpreter which means there is no need to recompile the code each time we run a scenario. For further information and programming examples about ns-2, C++ and OTcl refer to the lecture notes of Altman [11].

5.1.4 TclCL Interface

Tcl Classes (TclCL) are classes that behave as an interface that binds both OTcl and C++ classes. These bounded classes are interpreted classes for OTcl and compiled classes for C++. Figure 5.1 shows a link between OTcl and C++. The classes are not considered part of the interpreted classes nor the compiled classes when the classes are not linked together. Thus, they are considered standalone classes. The corresponding relationship between the bounded classes is one-to-one relationship.

Figure 5.2 shows how the classes of OTcl and C++ are bounded together in a one-to-one relationship. The classes that are not bounded in OTcl and C++ are standalone classes. The hierarchy of both of the classes that are bounded together belongs to the interpreted and compiled classes.

The TclCL consists of six classes that are written in C++ . These classes provide the interface between the OTcl and C++. As an example, they can be used for a global access to the compiled hierarchy from the interpreted hierarchy. These classes are: class Tcl, class InstVar, class TclObject, class Tclclass, class TclCommand, and class EmbeddedTcl. See

[25] for a complete description about the functionality of each class.

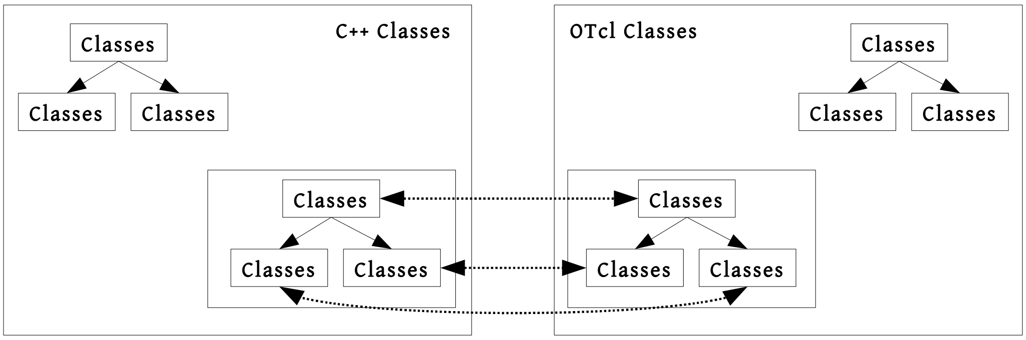

Figure 5.2: *OTcl and C++ One-to-One Relationship Classes*

5.2 MDP Architecture

The previous section gave an overview on the simulation environment. The benefit use of OTcl and C++ is to implement our message dissemination protocol (MDP). In chapter 4 the conceptual design provided an overview of the three approaches. Each of these approaches is implemented in the system by creating C++ classes and the corresponding OTcl scenario configuration file.

This section will present the main classes and components in the system and how they are implemented with each other to produce a functional protocol. The classes explain the structure of the MDP protocol. The description of the classes architecture will be presented using Unified Modeling Language (UML). We will implement the system from the basic architectural implementation. Then, start building upon this architecture the three different approaches that we already discussed in the conceptual design chapter.

5.2.1 A Basic Architecture

The main components of the system are a server and mobile nodes (see section 3.1). These components build the central system of message communication and dissemination. The mobile nodes use both communication techniques (GPRS and ad hoc). The architecture implements a C++ class that is MDPAgent. This class inherits the communication functionality from a base class in ns-2. The server class implements the functionality of sending and receiving messages without inhering any classes from ns-2.

Figure 5.3: *A Basic Class Diagram*

- MDPServer is the main class in the system. The server class coordinates all communication to and from the mobile nodes. Also, it provides the necessary support for a message routing.
- MDPAgent is a class that operates to provide the position of the mobile nodes in the the system. It makes the nodes' position available whenever the server class requests them. Beside, MDPAgent gives the mobile nodes the ability to map their positions in the map. In addition, it manages the ad hoc communication between the nodes. Also, it provides a message transmission functionality to the server.

5.2.2 The Naive Approach Architecture

The naive approach uses only the two essential classes that are implemented in the basic architecture. Figure 5.3 shows the basic architecture that consists of the two classes. The MDPServer class and the MDPAgent class. The naive approach uses the MDPServer class to send a message to all the nodes through GPRS and receive an acknowledgment message from the mobile node. The mobile nodes use the MDPAgent to send the acknowledgment message to the server. Although, the MDPAgent class inherits the functionality of the BaseAgent as the ad hoc communication. However, the direct message sending from the mobile nodes to the server is implemented in the MDPAgent class.

This process provides a direct communication between the two classes. No additional classes are required for the implementation of the naive approach. An extension to these classes is implemented in the static and the mobility approaches.

5.2.3 The Static Approach Architecture

The static approach operates to transmit a message from the server to a subset of mobile nodes in an area. The approach extends the functionality of the MDPServer class to include a decision mechanism to which node to receive the GPRS message. This mechanism depends on the mobile nodes functionality of ad hoc communication which is implemented in the MDPAgent class. Each mobile node is inserted in a vector with the neighbor nodes that its communication range can cover. These vectors are provided to the server to make a decision to which node to transmit the GPRS message.

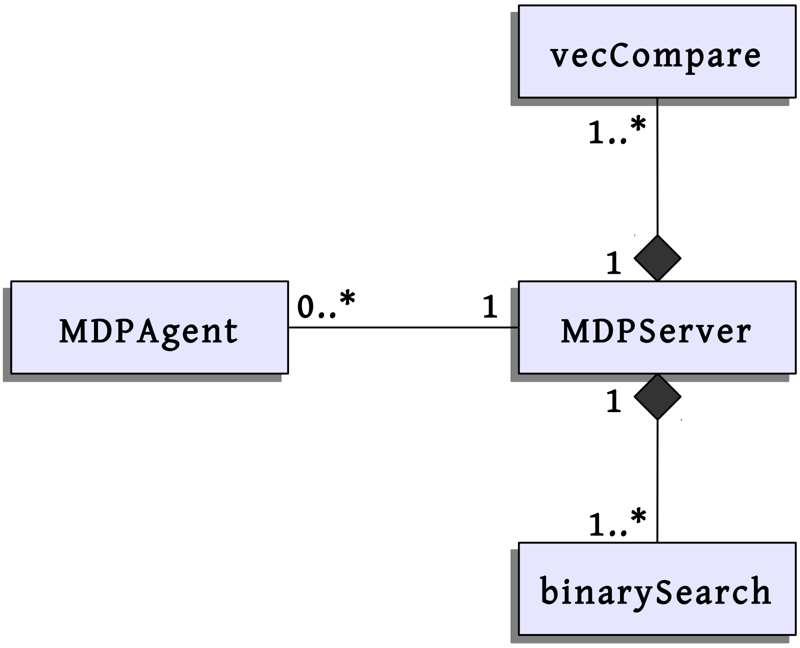

Figure 5.4: *Static Approach Class Diagram*

Figure 5.4 shows the extension of the basic architectural class diagram:

- vecCompare: An implementation structure to compare the sizes of the vectors that are generated through the grouping of a node and its neighbors that are covered by the ad hoc communication of one of the nodes. vecCompare is combined with a sort function to sort the vectors according to their sizes.

- binarySearch: The nodes that are contained in one vector might be found in other vectors because there is an intersection between the ad hoc communication range of the mobile nodes. The binarySearch structure operates to find these shared nodes and remove them when a node delivers the message.

The extension of the basic architecture functionality implements the decision procedure the MDPServer operates with. MDPServer is extended to include a sorting mechanism to the vectors that are grouped through MDPAgent class. The vectors are sorted based on their sizes from the largest to the smallest. Another extension is an implementation of a binary search mechanism. This extension is included to search for the nodes that delivered the message and removes those nodes from the other vectors that their communication range cover the intended node. The following example explains how the two type structures operate along with the MDPServer class:

- Sort the vectors based on their sizes from the largest to the smallest:

<1, <2, 3, 4, 5> >

<7, <2, 4, 6> >

<6, <7, 8> >

- The first number is the node that it's communication range cover the other nodes. The second part is the set of vectors that are covered by the ad hoc communication of the first node. A GPRS message is sent to node 1. Node 1 broadcast the message using ad hoc communication.

- A binary search occurs to find the nodes <1, 2, 3, 4, 5> in the other vectors and removes them.

<7, < 6 > >

<6, <7, 8> >

- Sort the vectors again and sends the GPRS message to node 6 which its ad hoc communication eventually covers node 7.

5.2.4 The Mobility Approach Architecture

The static approach architecture presented how we implemented the approach so that some of the nodes receive a message through GPRS transmission. It is the same for the mobility approach that uses the central class MDPServer to make a decision. However, the decision will be based on the knowledge of the map structure and the nodes mobility. In addition, there are some classes that are implemented to accommodate and support the MDPServer decision procedure.

The MDPServer class is associated with two other classes. DataList and CheckSegment classes that are related to segment information gathering. Figure 5.5 presents the class Diagram of the mobility approach.

- CheckSegment: This class stores the segment that the node is located on and all the connected segments. In addition, MDPServer checks whether the segment that the node is located on is within the matching segments set (see section 4.5.2). The MDPServer sends a message to the node that is located on a matching segment.

- DataSegment: It stores the node id, the segment where it is located and the number of the neighboring segments.

- DataList: The stored data from the DataSement are pushed into the list. The data in the list are sorted according to the largest number of connected segments. To provide a maximal matching when sending a GPRS message (see section 4.5.2).

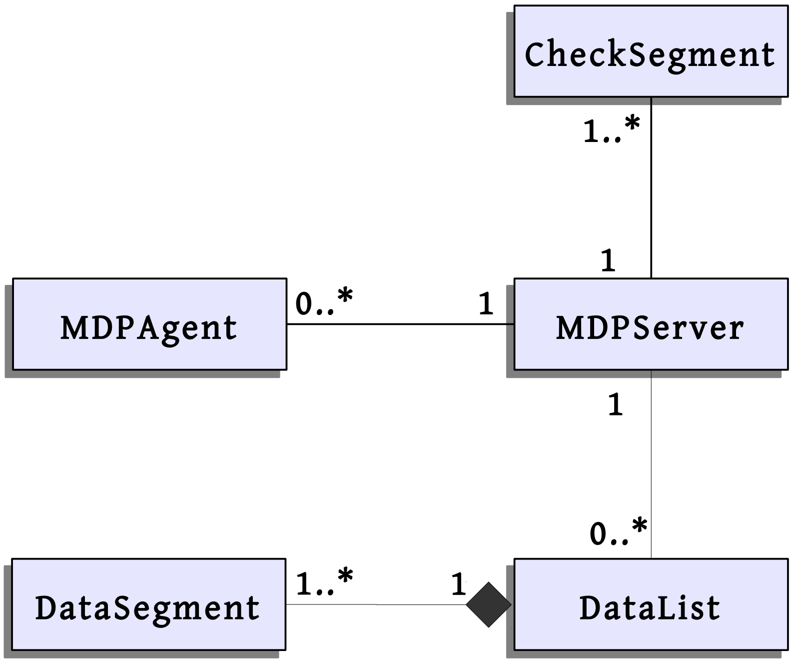

Figure 5.5: *Mobility Approach Class Diagram*

The following example provides an explanation how the mobility approach architecture is implemented to deliver a message to the nodes using GPRS.

The MDPServer requests node information from the MDPAgent class. The parameters that are requested are node id, the segment location and the number of the neighboring nodes. It pushes these information on the list sorting them according to the third parameter.

$$<1, 90, 8>$$
$$<2, 89, 7>$$
$$<3, 10, 5>$$

The first entry in the DataList has segment id equals 90. The MDPServer checks this segment in the SegmentClass to know if it is in the matched set or not. It adds this segment and all the neighboring ids to the matched set if the segment is not already in the set.

$$<90, <89, 91, 92> >$$

The second entry now in the list contains segment id 89. However, segment 89 is already part of the matching set. It moves to the third set and the process continues until all the map's segments are covered. The idea is to get the maximal matching segments (see section 4.5.2).

5.3 Implementation Summary

The complete MDP architecture of the three different approaches is displayed in figure 5.6.

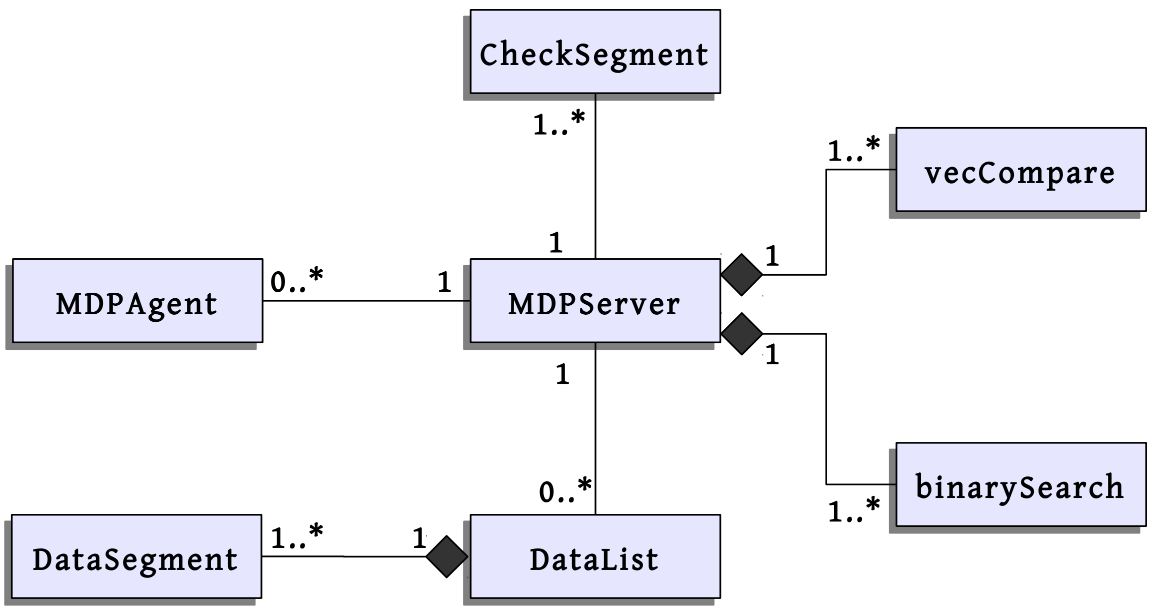

Figure 5.6: *MDP Architecture*

The protocol is implemented to accommodate the different message dissemination approaches. Each approach implements a class that suits their functionality. Figure 5.6 shows the complete MDP architecture where the two core classes are the MDPServer class and the MDPAgent class.

The naive approach uses the two central classes: MDPServer and MDPAgent to transmit the message directly from the server to the mobile nodes. The static approach needs more functionality for the process of decision procedure at the server to determine the right node subset that receives the message through GPRS. The extension of binarySearch and vecCompare are implemented to support the decision procedure in the static approach which is implemented according to the set cover greedy algorithm. The mobility approach uses the three implementation classes of the segment knowledge because the mobility approach depends heavily on the structure of the area of interest for the choice of the right node subset for GPRS message transmission. These classes extensions are used to provide the matching mechanism of the segments in the area of interest. (For deep architectural analysis of the approaches refer to each of the approaches architecture in the chapter).

Chapter 6

Evaluation

We presented the principle design and implementation of our protocol in the previous chapters. This chapter presents the evaluation results. The goal of this simulation is to present how effective and efficient the result that we achieved.

Mobile nodes have limited resources because of that we are interested in evaluating how much energy is consumed when implementing the three different message dissemination approaches. In addition, we are interested in evaluating the effectiveness of message delivery on the mobile nodes.

Section 6.1 describes the simulation setup that prepares the simulation environment. Then, section 6.2 provides an explanation about the evaluation metrics and parameters that are used to calculate the energy consumption for each approach. An overview of the evaluation results is given in section 6.3. Finally, a summary about the evaluation of our protocol is provided in section 6.4.

6.1 Simulation Setup

6.1.1 Tcl Configuration File

The message dissemination protocol suggests three different approaches to disseminate a message to the nodes that are located in the area of interest. A Tcl configuration file is the interface that provides the necessary inputs to create a simulation scenarios for these approaches. The configuration file defines different properties that shape the simulation scenarios. Listing 6.1 shows the input properties that the tcl configuration file contains.

```
set opt(mm) Chicago2kmx2km.map   # Location Model File
set opt(gx2)     2000;           # X Dimension of Topology
set opt(gy2)     2000;           # Y Dimension of Topology
set opt(nn) 718;                 # Number of Mobile Nodes
set opt(adhoc) on;               # Activate Ad Hoc Communication
set opt(sc)   ../ns.mobility;    # Traffic Model File
set opt(approach) 2;             # Message Dissemination approach
set opt(startTime) 300;          # Message Broadcasting Start Time
set opt(stopTime) 350;           # Message Broadcasting End Time

#logTime
for {set i 295} {$i < 370 } {incr i 5}  {
    $ns_ at $i "$segment_layer_ log";
}
```

Listing 6.1: Tcl Simulation Properties

Area of Interest

The area of interest where the nodes are located is part of Chicago, Illinois map. The area of interest has a size equals to 2 km^2. This area is generated using UDel Models [28]. UDel Models are simulation tools for realistic simulation of urban wireless networks.

Mobile Nodes

The number of mobile nodes that circulates in the area of interest is 718 nodes. One message is expected to be delivered by all the 718 nodes that are located in the area of interest. The mobile nodes use ad hoc communication property to enable them to send messages to each other.

Message Dissemination Approach

The configuration file contains an input parameter that defines the type of message dissemination approach. There are three input numbers that can be assigned to opt(approach) according to which dissemination approach is used:

- **0:** The naive approach

- **1:** The static approach
- **2:** The mobility approach

Simulation Times

The simulation defines different time instances that sequence the occurrence of events in the simulation scenarios. These time instances are:

- **startTime:** Indicates the start time of a message transmission. At this point, the message is transmitted using GPRS communication to the nodes that are chosen according to the dissemination approach.
- **stopTime:** At this point, the message transmission is stopped and the server checks the nodes who have delivered the message.
- **logTime:** A log is defined every instance of time in specific period to capture the energy consumption and the number of nodes that transmits and delivers a message.

6.1.2 Trace File

UDel-Models generate an input file for ns-2 that describes the pattern of nodes mobility at different time instances. The UDel-Models simulate the traffic flow based on the general realistic observation of the data that is gathered from the traffic flow in a specific hour. The time that the mobile nodes enter the area of interest is randomly defined based on their arrival to a street segment.

Listing 6.1 contains the file ns.mobility that provides the simulation traffic pattern of the mobile nodes. This file defines an initial coordinates for the mobile nodes as the following:

$node_(327) set X_ 1939.99
$node_(327) set Y_ 476.72
$node_(327) set Z_ 0.00

In this example, node 327 is assigned to initial X and Y coordinates due to the fact that the map where the simulation occurs is provided as a X/Y plane. For that reason, all mobile nodes assign zero to the Z coordinate. In addition, the trace file includes further information about the mobility of the nodes and their future X and Y coordinates when time passes. For example, the node 327 moves from its initial location at second 15 to a new location. Also, its speed is maintained on 8.227699 m/s to change its position on the Y coordinate at second 16.

$ns_ at 15.00 "$node_(327) setdest-udelmobility 1942.99 569.31 0.00 8.227699"
$ns_ at 16.00 "$node_(327) setdest-udelmobility 1942.99 577.54 0.00 8.227699"

Figure 6.1 explains the mobile node's movement in a trace file. Node 327 moves from its position at second 15 to a new position at second 16. The distance that the mobile node travels equals 8.227699 m.

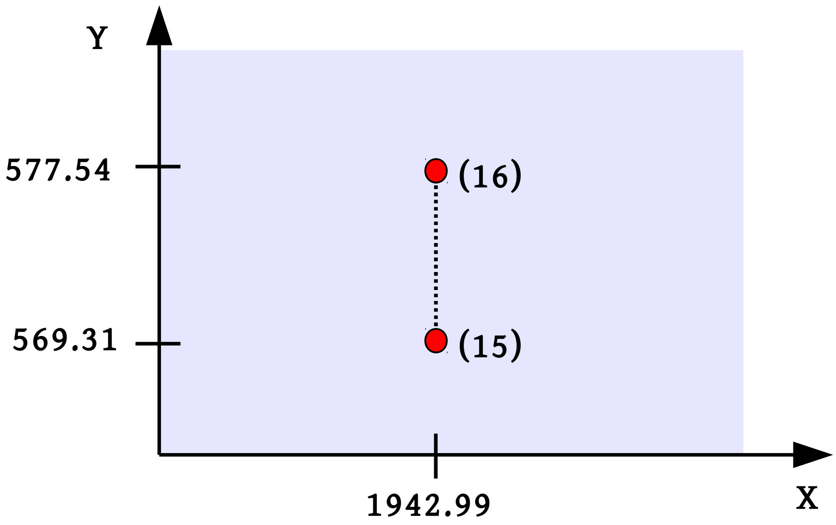

Figure 6.1: *Mobile Node Movement*

6.2 Evaluation Metrics

6.2.1 Parameters

The parameters for evaluating the message dissemination protocol depend on the objectives that we are seeking to achieve. Mostly, our concern is to produce an efficient messaging mechanism that decreases the consumption of the mobile nodes resources. The process to conserve mobile nodes resources means to find a mechanism that consumes less energy and increases message delivery rate.

Most of the energy consumed on mobile nodes occurs due to the misuse of the communication techniques. Each approach sends the message according to the communication techniques that it defines. There is a large difference between message transmission using GPRS or message transmission using ad hoc communication from the point view of energy consumption. The energy consumption is calculated only for the mobile nodes because they contain limited resources. The energy that is consumed on the server is not included in our calculation. Table 6.1 [37] provides the values of the energy that is consumed on the mobile nodes. These energy values are for a message that has a size of 1000 bit.

Parameters	**Energy [mJ]**
GPS	
Position Fix	75
802.11b at 1 Mbps (broadcast rate)	
Send (1000 Bit)	2
Receive (1000 Bit)	1
GPRS	
Send (1000 Bit)	80
Receive (1000 Bit)	40

Table 6.1: Energy Model

There are different operations that are used to calculate the amount of energy consumed during message dissemination. These parameters use table 6.1 for the value of the energy that is consumed. These parameters are:

- **GPRS_recv:** The mobile nodes receive a message using GPRS transmission.
- **GPRS_ack:** The mobile nodes sends an acknowledgment to the server using GPRS confirming the delivery of the message.
- **ADHOC_sent:** The mobile nodes broadcast a message in the neighborhood using ad hoc communication.
- **ADHOC_recv:** A neighboring mobile node receives a message through ad hoc communication.
- **ADHOC_check_sent:** The mobile nodes broadcast a check message in the neighborhood to checks whether a message is delivered in order to limit the flooding of the messages.
- **ADHOC_check_recv:** A number of mobile nodes receive the check message who report to the node who initiated the check message if the intended message isn't delivered.

- **GPS_count:** The Server requests about the position of the nodes to send a message. The initial position request about the nodes positions is equivalent among the three approaches. This count considers only the position of the nodes that didn't deliver the message when the message broadcast is stopped.

6.2.2 Energy Consumption Calculation

The energy values in table 6.1 are for a message of a size of 1000 bit only. An increase in a message size means the amount of energy consumed on mobile nodes increases. The total amount of energy consumed is calculated according to the following equation:

$$\begin{aligned} ec &= \#\, GPS\_count \times 75 + \#\, ADHOC\_check\_recv + \#\, ADHOC\_check\_sent \times 2 \\ &\quad + MsgSize \times \#\, ADHOC\_recv + MsgSize \times \#\, ADHOC\_sent \times 2 \\ &\quad + MsgSize \times \#\, GPRS\_recv \times 40 + \#\, GPRS\_ack \times 80 \end{aligned}$$

The equation implies the possibility to compare the efficiency of the three implemented approaches in our protocol by investigating the impact of the message size on the amount of energy consumed. In addition, each approach uses different parameters according to the design architecture of the approach.

The naive approach uses only GPRS communication. Then, the total amount of energy that is consumed will not include the ad hoc communication parameters in our calculation. The only parameters that are included in the equation are GPRS communication and GPS.

$$ec = \#\, GPS\_count \times 75 + MsgSize \times \#\, GPRS\_recv \times 40 + \#\, GPRS\_ack \times 80$$

The static approach uses both communication techniques GRPS and ad hoc. However, the static approach considers the nodes in a constant manner. There is no mobility consideration for the mobile nodes when the message is delivered. The subset nodes that are chosen flood the message in their neighborhood without sending an ad hoc check. The equation excludes the checking mechanism messages. Then, the equation of energy consumption is given as:

$$\begin{aligned} ec &= \#\, GPS\_count \times 75 + MsgSize \times \#\, ADHOC\_recv + MsgSize \times \\ &\quad \#\, ADHOC\_sent \times 2 + MsgSize \times \#\, GPRS\_recv \times 40 + \#\, GPRS\_ack \times 80 \end{aligned}$$

The mobility approach uses all the parameters that are provided in the original equation. It provides the checking mechanism to avoid a continuous flooding of the message among the nodes which might lead to consume all the available resources on the mobile nodes.

6.3 Results

The two metrics that evaluate the performance of our protocol are the total amount of energy consumed and the message delivery rate on the mobile nodes. The total amount of energy consumed depends on the message delivery approach that is used (see section 6.2.2). In addition, the message size has an effect on the amount of energy consumed. Our goal is to produce a message dissemination protocol that considers the size of the message that is disseminated. The evaluation results of the three approaches that are implemented focus on the relation between the message sizes and the total amount of energy consumed when time passes. The evaluation is done for four different message sizes (1000,2000,5000,10000) bits. A comparison between the three approaches to evaluate the total amount of energy that is consumed takes place showing the efficiency of the message dissemination approaches.

6.3.1 Energy Efficiency

Our aim is to maximize the efficiency and resource saving of mobile nodes when the message size increases. Figure 6.2 shows the total amount of energy consumed of the three approaches for a message size of 1000 bits. In this case, the mobility approach doesn't deliver messages in an efficient manner due to the fact that the size of the check message equals the size of the intended transmitted message. These check messages incur more energy than the static approach that broadcast messages without any checking mechanism.

On the other hand, figure 6.3 represents an increase in the size of the message to 2000 bits. However, the mobility approach energy consumption is still higher than the static approach but provides better results than a message of 1000 bits. It is clear that the size of the message affects the energy that is consumed.

Notice that the naive approach and the static approach are constant at the beginning because the GPRS message transmission occurs at one point in time. Although, the static approach considers ad hoc communication but ad hoc broadcast happens only once when the node receive the message at time 300. The mobility approach transmits GPRS messages as the naive and the static approaches. However, the mobility approach is designed to broadcast ad hoc messages when the mobile nodes change their locations.

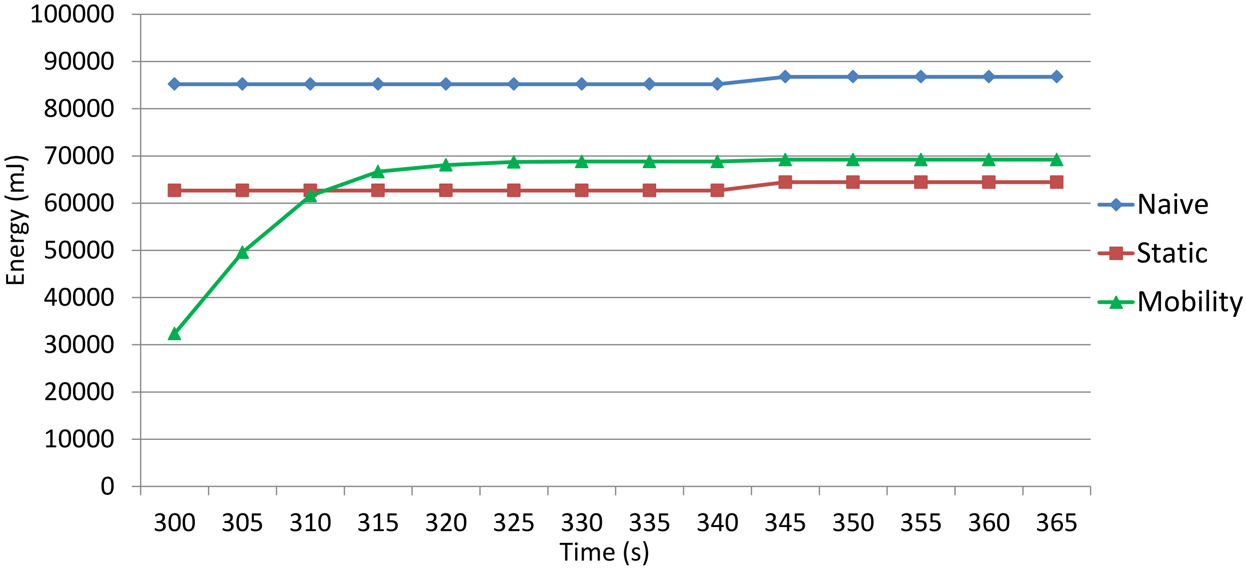

Figure 6.2: *Energy Consumption for a Message of Size 1000 Bits*

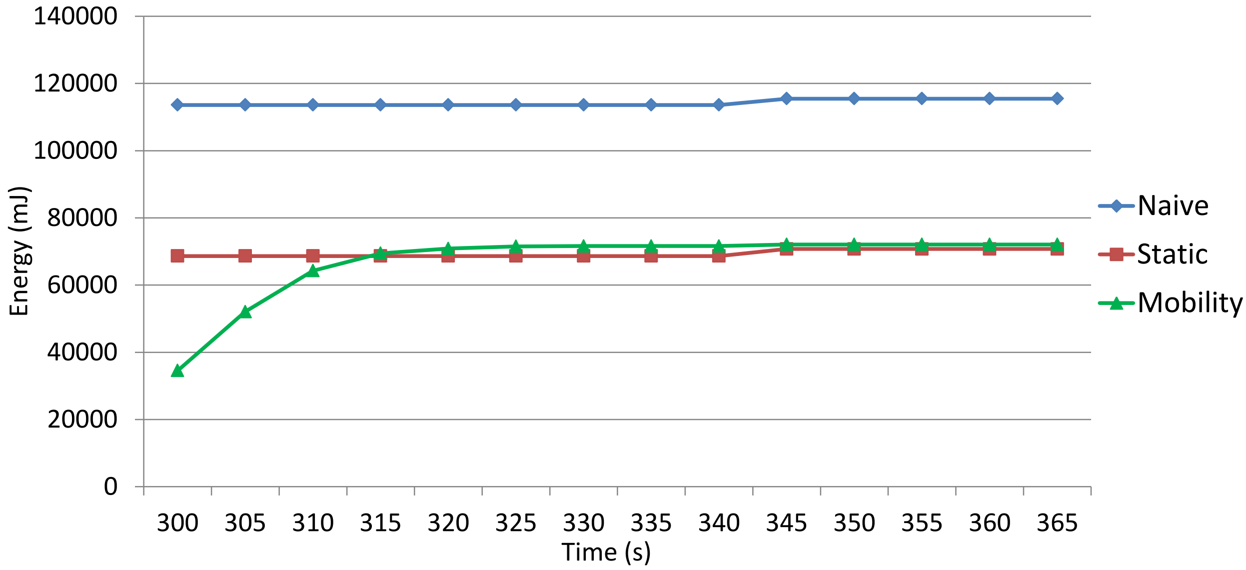

Figure 6.3: *Energy Consumption for a Message of Size 2000 Bits*

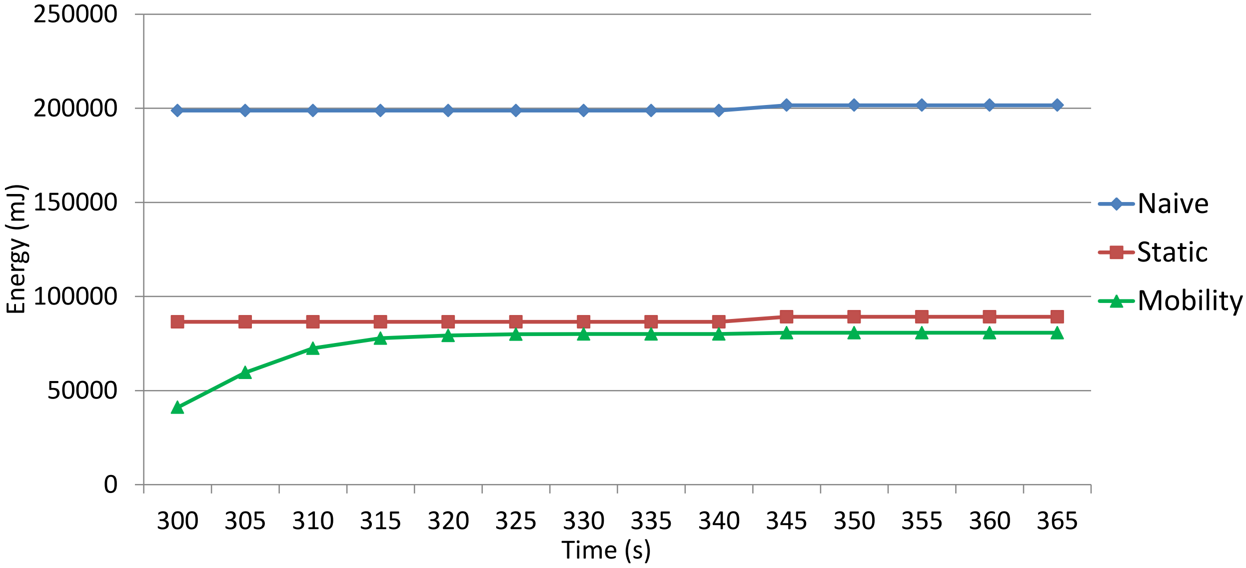

Figure 6.4: *Energy Consumption for a Message of Size 5000 Bits*

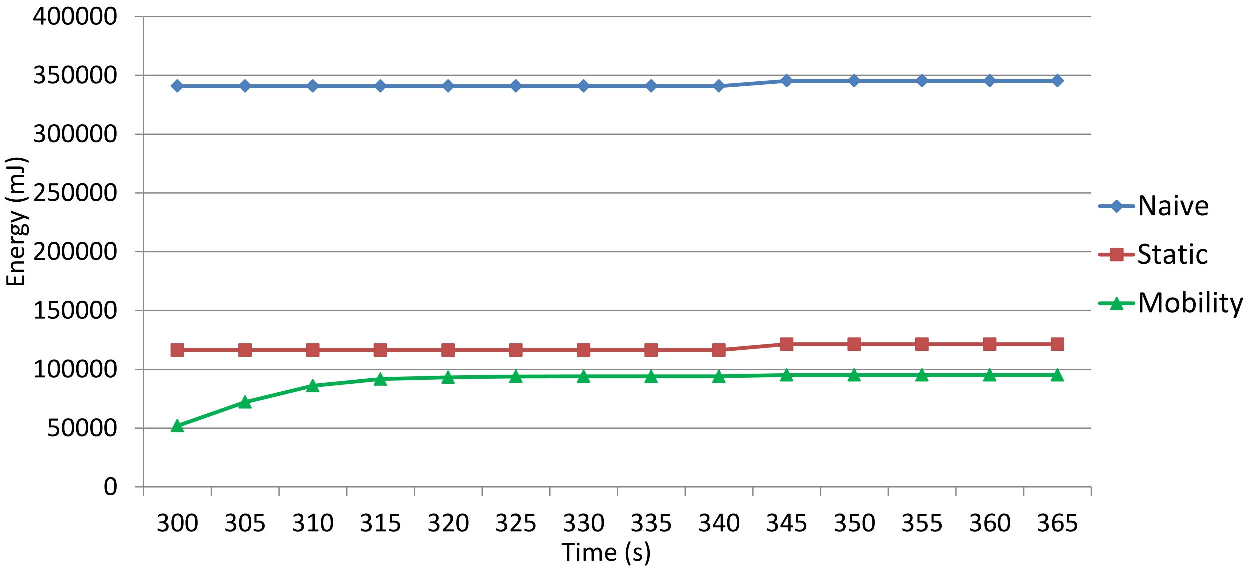

Figure 6.5: *Energy Consumption for a Message of Size 10000 Bits*

The evaluation of the mobility approach gets better when the message size increases. Figure 6.4 and 6.5 provide clear improvement of the mobility approach over the naive and static approaches from the point view of energy consumption. This improvement is due to the fact that the number of messages that are delivered using GPRS in mobility approach is less than the messages that are delivered in the other two approaches. The equation that is explained in section 6.2.2. Part of it includes the message size and the impact is in the difference of the energy values. The number of nodes that delivers a message through GPRS makes the difference in energy consumption between the three approaches.

- MsgSize × ADHOC_recv
- MsgSize × ADHOC_sent × 2
- MsgSize × GPRS_recv × 40

The total amount of energy that is consumed in the naive approach is always higher than the static and mobility approaches because the naive approach uses only GPRS communication. The energy consumption for GPRS is much higher than the energy consumption using ad hoc communication (see table 6.1 for comparison).

6.3.2 Message Delivery Rate

The evaluation of nodes' message delivery depends on the total number of nodes that delivers a message. The thing is that mobile nodes have a limited time frame to send the message to other nodes due to the dynamic network. Since, dynamic networks differ from time to time according to the movement and location of the mobile nodes in the area.

The goal of the message dissemination protocol is to achieve 100% delivery rate. All the nodes in the area of interest will deliver the message. Figure 6.6 shows the three different approaches that disseminate the message. The total amount of nodes is 718 and the time of the message broadcast is 300. The naive approach achieves 100% delivery using GPRS. Also, the static and mobility approaches achieve delivery rate of 100%. Although, they use a combination of communication to deliver the message.

The result which achieved delivery rate 100% at time 300 may not achieve the same delivery rate when message broadcast time changes because at the time the message transmission occurs the nodes might be located outside the area of interest and they might not enter again. However, the implementation of our protocol makes it possible to achieve 100% message delivery rate. Table 6.2 displays different broadcast times and the number of message delivered. We can see that we achieved at every point in time a delivery rate of 100%.

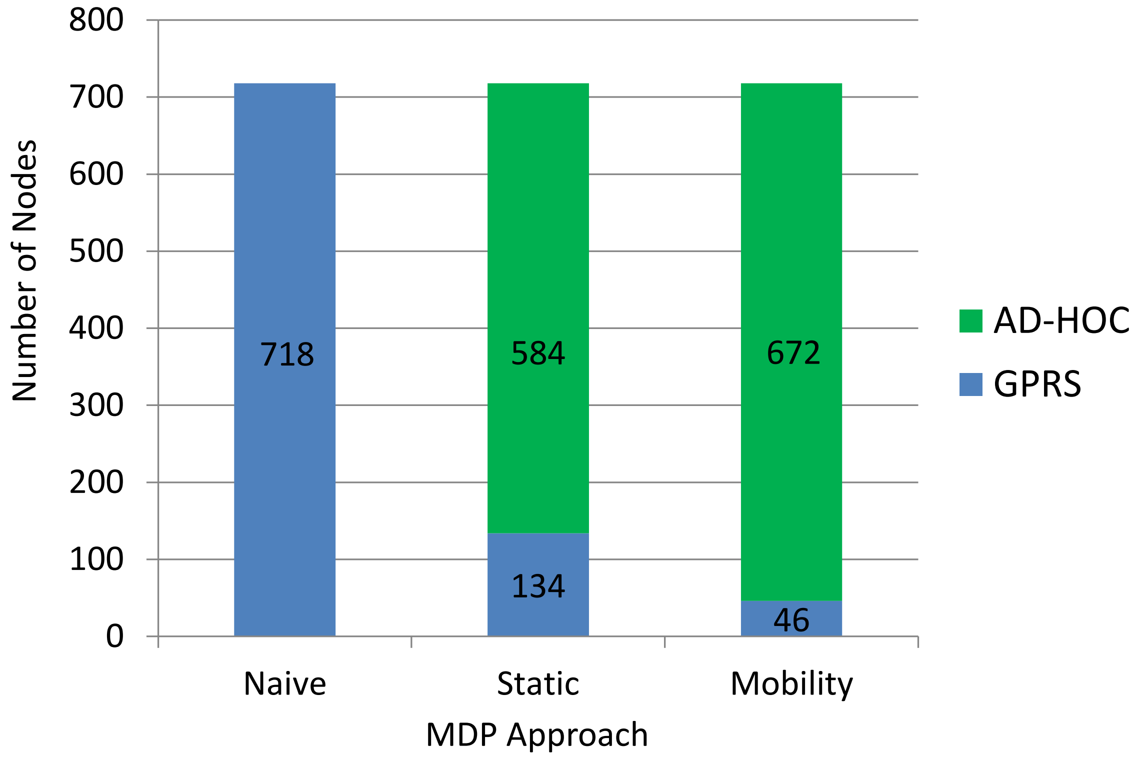

Figure 6.6: *Message Delivery at Time 300*

Time	Naive Approach			Static Approach			Mobility Approach		
	Message Delivery			Message Delivery			Message Delivery		
	GPRS	AD HOC	$\sum$	GPRS	AD HOC	$\sum$	GPRS	AD HOC	$\sum$
100	718	0	718	132	586	718	45	673	718
200	718	0	718	130	580	718	40	678	718
300	718	0	718	134	584	718	46	672	718
400	718	0	718	132	586	718	41	677	718
500	718	0	718	131	587	718	44	674	718
600	718	0	718	134	584	718	36	682	718
700	718	0	718	129	589	718	39	679	718
800	718	0	718	132	586	718	42	676	718
900	718	0	718	139	579	718	40	678	718
1000	718	0	718	134	584	718	37	683	718
1100	718	0	718	131	587	718	44	674	718
1200	718	0	718	133	585	718	40	678	718
1300	718	0	718	132	586	718	37	681	718

Table 6.2: Statical Analysis of Message Delivery

6.3.3 Mobile Nodes Density

The number of mobile nodes that is located in the area of interest has an effect over the total amount of energy that is consumed. An increase in the number of mobile nodes does not necessary mean an increase in the energy consumed because it depends on the dissemination approach that is used. The connectivity between the mobile nodes increases when the number of mobile nodes increases. The connectivity property between the nodes has no effect in the naive approach because this approach uses only a direct communication with the server through GPRS which means an increase in the number of mobile nodes increases the amount of energy that is consumed. On the contrary, the use of ad hoc communication to exchange messages among the mobile nodes decreases the number of GPRS messages needed which leads to decrease the relative energy consumed. Figure 6.7 shows the amount of energy that is consumed for a message of size 10000 bit when the number of mobile nodes increases in the area of interest.

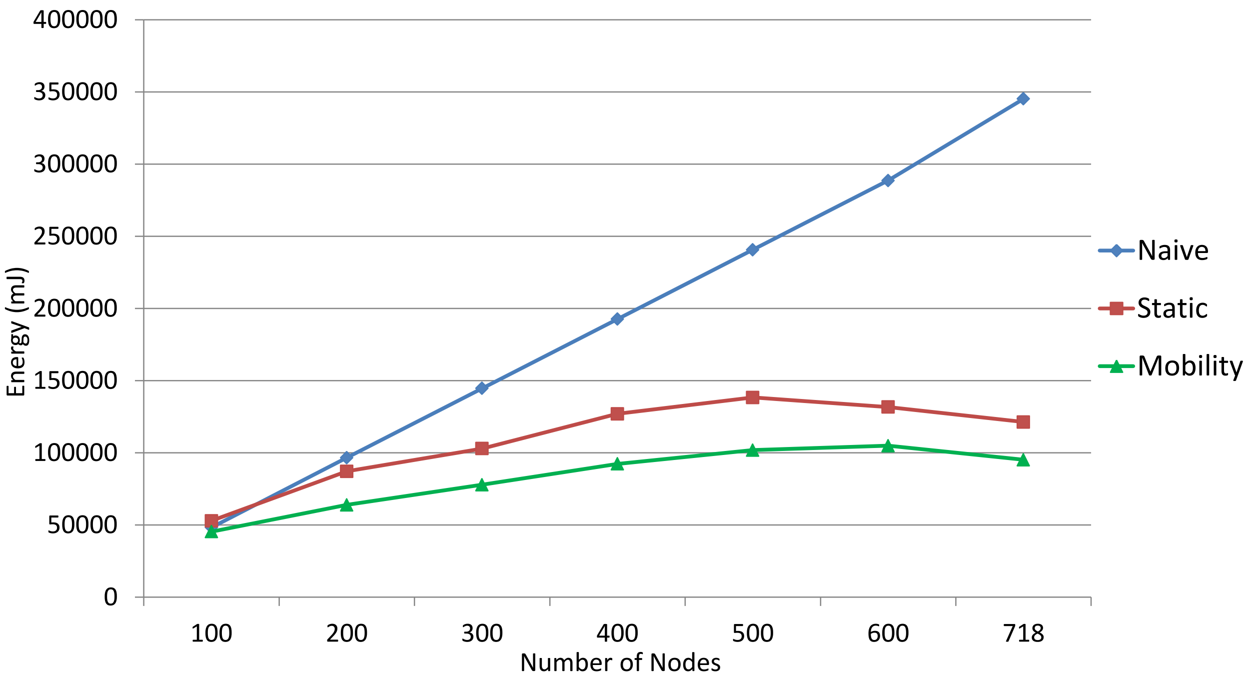

Figure 6.7: *Energy Consumption for Different Number of Nodes*

The evaluation emphasizes the importance of the ad hoc communication over GPRS from the point view of energy consumption. The results of the naive approach clearly show the increase of the total energy consumed when the number of mobile nodes increases. The result of that continues increase is due to the use of GPRS only in the naive approach. On the other hand, the message is transmitted to a subset of mobile nodes in the two other approaches through GPRS. Then, the message is disseminated using ad hoc communication which will result in a decrease in the energy consumption compared to the naive approach. However, notice the increase in energy consumption in the static approach when the number of nodes equals 500. The same as in the mobility approach when the number of nodes equals 600. The total amount of energy consumed is at its maximum because the connectivity between the mobile nodes is decreased which will allow messages to be transmitted using GPRS. There is a decrease in the amount of energy that is consumed when the number of mobile nodes increases because the message is broadcasted using ad hoc communication since there is an increase in the connectivity between the mobile nodes. decreases.

6.4 Evaluation Summary

This chapter described the evaluation of the implementation of the message dissemination protocol. It explained the difference between the three message dissemination approaches from the point view of energy consumption. An equation is introduced to calculate the amount of energy that is consumed on the mobile nodes when a message is delivered (see section 6.2.2). Also, the chapter gave an overview of the parameters that are used to evaluate the message dissemination protocol (see section 6.2.1). The evaluation showed that message delivery rate that was achieved equals 100%. However, each approach consumes different energy according to the communication mechanism that is used because the energy that is consumed through GPRS is more than ad hoc broadcast (see table 6.1). In addition, the message size has an effect on the amount of energy that is consumed. Figure 6.8 displays the evaluation results that are achieved for energy consumption for a message that is transmitted using the three dissemination approaches and the effect of the message sizes on the energy.

The mobility approach is less energy efficient than the static approach when the size of the message is small. However, the increase in message sizes results in improving the mobility approach efficiency over the static and the naive approaches. The mobility approach can be used efficiently in disseminating large size messages such as video highlights and visual ads compared to small text messages.

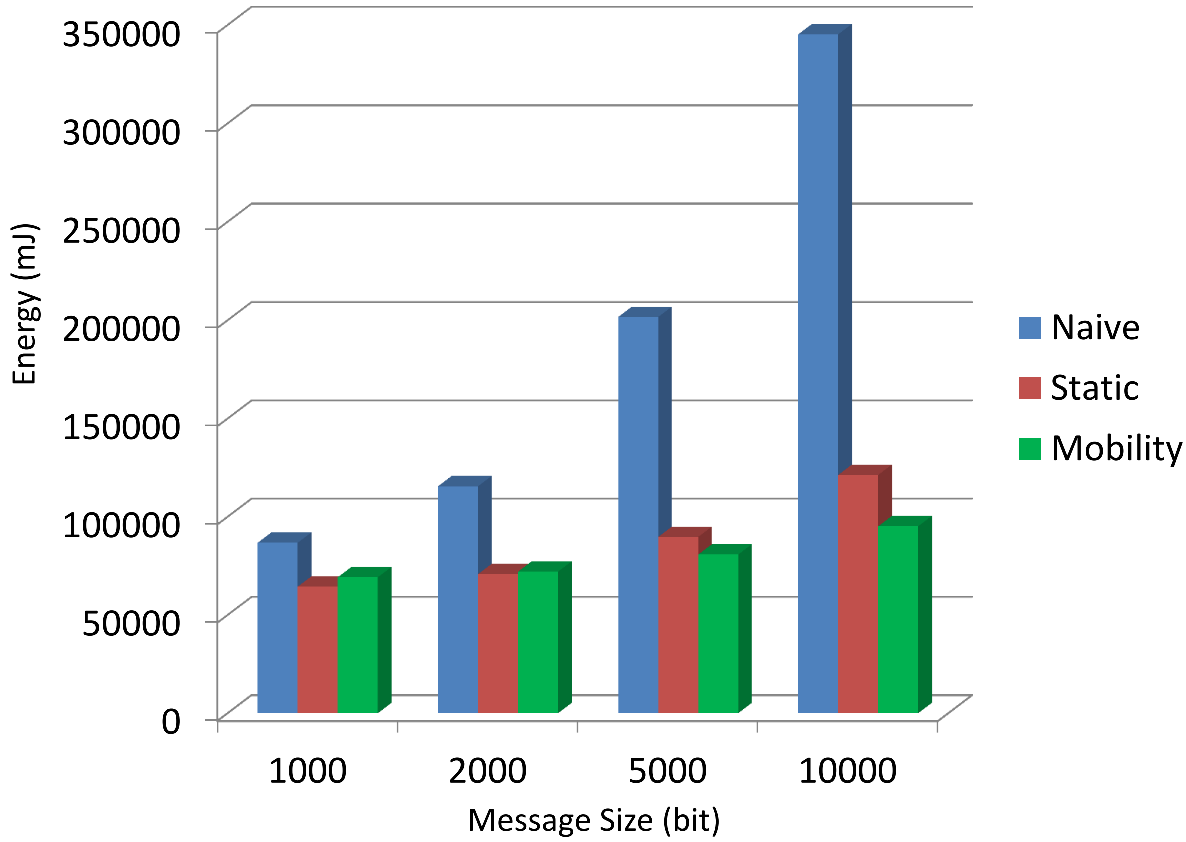

Figure 6.8: *Energy Consumption for Different Message Sizes*

Chapter 7

Summary and Future Work

7.1 Summary

This thesis deals with the design and implementation of an efficient message dissemination protocol for public sensing. The protocol is designed to disseminate large size messages to mobile nodes in an area of interest. The protocol focuses on two characteristics: energy efficiency of the message disseminated to conserve the available resources on mobile nodes. Effectiveness of the message delivered to all the mobile nodes that are located in the area of interest.

The protocol implements three separate approaches for message dissemination. The naive approach, the static approach and the mobility approach. Each one of them depends on different factors. One of the factors is the communication technique that is used in disseminating a message. The naive approach uses only GPRS to transmit a message directly from the server to all the mobile nodes. The naive approach is the basic solution for message delivery. The other two approaches the naive approach and the mobility approach are more sophisticated in term of transmitting the message through GPRS to a subset of nodes. In addition, they integrate ad hoc communication to broadcast messages between the nodes. At the beginning, the two approaches choose to transmit the message to a subset of nodes through GPRS. However, the requirement for choosing such subset is different among the two approaches. The static approach uses a greedy algorithm of the set cover problem to choose the smallest subset of nodes to minimize message transmission through GPRS. The choice of the nodes in the subset maximizes the number of nodes that their ad hoc communication can cover. The final approach is the mobility approach that uses knowledge about the structure of area and mobile nodes behavior such as mobile node direction and movement. The implementation of the area structure is based on the matching problem that gets the maximal street segments which are covered by mobile nodes movement.

The result of the simulation investigates the total amount of energy consumed and the message delivery rate. Message delivery rate achieved equals 100%. The amount of energy consumed among the three approaches differs according to the size of the message which is transmitted. The evaluation reveals that increasing the size of the message transmitted the better the mobility approach performs the lower the relative energy consumed compared to the naive approach and the static approach.

7.2 Future Work

Our implementation of the mobility approach uses the behavior of the mobile nodes and the knowledge of the street segments. The characteristics of the mobile nodes are not fully utilized. There is a possibility to utilize the mobile nodes in a way to use their speed, movement orientation and the distance that it covers to reach a specific point on the map in order to disseminate a message. In addition, the knowledge of the map structure still can be used. Figure 7.1 explains how to benefit from this concept to disseminate a message at specific location.

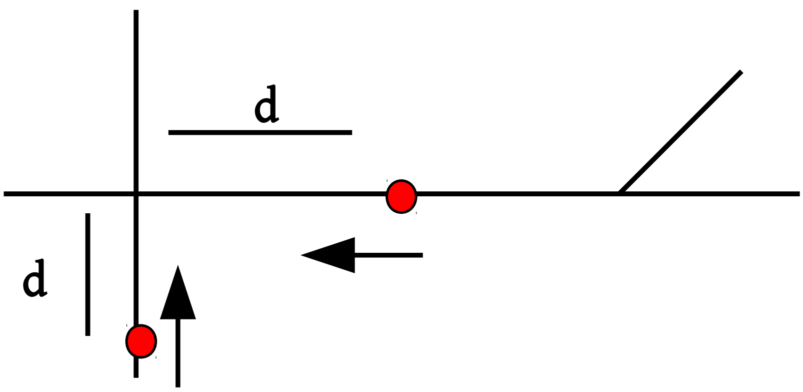

Figure 7.1: *Message Dissemination at Specific Location*

Usually, mobile nodes have more probability to meet at streets intersections because there is a density of mobile nodes who stop to change their direction or cross the traffic lights. It is possible to implement the dissemination protocol to broadcast a message when mobile nodes reach the intersection. It ensures that there are more mobile nodes near by to receive the message. Also, it provides another way to choose the subset of the mobile nodes who receive the message through GPRS because knowing the exact location of the mobile nodes at specific time facilitate the choice of the subset.

Assume that the two nodes in figure 7.1 move according to the direction of the arrows. Then, we can calculate for both of the mobile nodes the time it needs to reach the intersection:

$$Time = Distance\ (d)\ /\ Current\ Speed$$

It is assumed that the mobile nodes continue their movement on the street segments with the same speed until they reach near the intersection. Later, the calculated time is added to the current time clock and check if the two times are equal or in some time range that allows the two mobile nodes meet and exchange messages through ad hoc communication.

$$Possible\ Meeting\ Time = Time + Current\ Time$$

According to that a decision is made to send the message through GPRS to one of the nodes or to both nodes in case they don't meet.

Bibliography

[1] CENS, Participatory sensing/Urbansensing Project, http://research.cens.ucla.edu/, 23. August 2011.

[2] WeTap, http://we-tap.appspot.com/, 11. September 2011.

[3] Pothole Patrol, http://cartel.csail.mit.edu/doku.php?id=p2_pothole_patrol, 11. September 2011.

[4] The Delay-Tolerant Networking Research Group. http://www.dtnrg.org/wiki, 01. August 2011.

[5] The PodNet project. http://podnet.ee.ethz.ch/, 01. August 2011.

[6] Geographic Information System (GIS). http://www.gis.com/, 16. August 2011.

[7] OMNeT++, http://www.omnetpp.org/, 5. September 2011.

[8] OPNET, http://www.opnet.com/, 5. September 2011.

[9] The Network Simulator - ns-2, http://isi.edu/nsnam/ns/, 5. September 2011.

[10] K. Akkaya and M. Younis. A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks*, 3:325–349, 2005.

[11] E. Altman and T. Jimenez. *NS Simulator for begineers.* December 2003.

[12] C. Becker and G. Schiele. New mechanism for routing in ad hoc networks. *4th Plenary Cabernet Wksp.*, Oct. 2001.

[13] A. L. Buchsbaum, A. Efrat, S. Jain, S. Venkatasubramanian, and K. Yi. Restricted strip covering and the sensor cover problem. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '07, pages 1056–1063, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.

[14] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava. Participatory sensing. In *In: Workshop on World-Sensor-Web (WSW'06): Mobile Device Centric Sensor Networks and Applications*, pages 117–134, 2006.

[15] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson. People-centric urban sensing. In *The Second Annual International Wireless Internet Conference (WICON*, pages 2–5. IEEE Computer Society Press, 2006.

[16] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss. Delay-Tolerant Networking Architecture. RFC 4838 (Informational), April 2007.

[17] G. Chen and D. Kotz. A Survey of Context-Aware Mobile Computing Research. Technical Report TR2000-381, Dept. of Computer Science, Dartmouth College, Nov. 2000.

[18] S. Consolvo et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1797–1806, New York, NY, USA, 2008. ACM.

[19] S. A. Cook. The complexity of theorem proving procedures. In *Proceedings of the Third Annual ACM Symposium on the Theory of Computing*, pages 151–158, New York, 1971.

[20] A. J. Demers, D. H. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. E. Sturgis, D. C. Swinehart, and D. B. Terry. Epidemic algorithms for replicated database maintenance. In *PODC*, pages 1–12, 1987.

[21] A. Ferreira. Building a reference combinatorial model for manets. *IEEE Network*, 18(5):24–29, 2004.

[22] B. Garbinato, A. Holzer, and F. Vessaz. Context-aware broadcasting approaches in mobile ad hoc networks. *Computer Networks*, 54(7):1210–1228, 2010.

[23] L. J. García Villalba, A. L. Sandoval Orozco, A. Triviño Cabrera, and C. J. Barenco Abbas. Routing protocols in wireless sensor networks. *Sensors*, 9(11):8399–8421, 2009.

[24] R. Handorean, C. D. Gill, and G.-C. Roman. Accommodating transient connectivity in ad hoc and mobile settings. In *Pervasive*, pages 305–322, 2004.

[25] T. Issariyakul and E. Hossain. *Introduction to Network Simulator NS2*. Springer, 2009.

[26] D. S. Johnson. Approximation algorithms for combinatorial problems. In *Proceedings of the fifth annual ACM symposium on Theory of computing*, STOC '73, pages 38–49, New York, NY, USA, 1973. ACM.

[27] R. Karp. Reducibility among combinatorial problems. In R. Miller and J. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, 1972.

[28] J. Kim, V. Sridhara, and S. Bohacek. Realistic mobility simulation of urban mesh networks. *Ad Hoc Networks*, 7(2):411–430, 2009.

[29] Y. Li, S. Panwar, and H. Chao. Exhaustive service matching algorithms for input queued switches. In *High Performance Switching and Routing, 2004. HPSR. 2004 Workshop on*, pages 253 – 258, 2004.

[30] L. Lovász and M. Plummer. *Matching theory.* Annals of discrete mathematics. North-Holland, 1986.

[31] N. Maisonneuve, M. Stevens, M. E. Niessen, P. Hanappe, and L. Steels. Citizen noise pollution monitoring. In *Proceedings of the 10th Annual International Conference on Digital Government Research: Social Networks: Making Connections between Citizens, Data and Government*, dg.o '09, pages 96–103. Digital Government Society of North America, 2009.

[32] S. Merugu, M. Ammar, and E. Zegura. Routing in space and time in networks with predictable mobility. Technical report, Georgia Institute of Technology, 2004.

[33] V. Paruchuri, A. Durresi, and R. Jain. Optimized flooding protocol for ad hoc networks. *CoRR*, cs.NI/0311013, 2003.

[34] T. Shuai and X.-D. Hu. Connected set cover problem and its applications. In *AAIM'06*, pages 243–254, 2006.

[35] F. Tchakountio and R. Ramanathan. Tracking highly mobile endpoints. In *WOWMOM*, pages 83–94, 2001.

[36] A. Vahdat and D. Becker. Epidemic routing for partially-connected ad hoc networks. Technical report, Duke University, 2000.

[37] H. Weinschrott, F. Dürr, and K. Rothermel. Efficient capturing of environmental data with mobile rfid readers. In *Mobile Data Management*, pages 41–51, 2009.

[38] D. Williamson. Lecture notes on approximation algorithms. Technical Report RC 21273, IBM, February 1999.

[39] Z. Zhang. Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: Overview and challenges. *IEEE Communications Surveys and Tutorials*, 8(1-4):24–37, 2006.