Escola d'Enginyeria de Telecomunicació i Aeroespacial de Castelldefels

UNIVERSITAT POLITÈCNICA DE CATALUNYA

MASTER THESIS

TITLE: Survey of Opportunistic Network Schemes for Message Forwarding

MASTER DEGREE: Master in Science in Telecommunication Engineering & Management

AUTHOR: Sandra Rodríguez Aranguren

DIRECTOR: Roc Meseguer Pallarès

DATE: November 28th, 2013

Title: Survey of Opportunistic Network Schemes for Message Forwarding

Author: Sandra Rodríguez Aranguren

Director: Roc Meseguer Pallarès

Date: November 28th, 2013

Overview

The aim of this Master Thesis is to present the investigation results of the different routing schemes that are used in opportunistic networks, the so-called opportunistic routing schemes. In special the routing scheme with mobile infrastructure DataMULEs is subject of an exhaustive study.

Opportunistic routing schemes are from great interest and the subject of several studies, as they offer the possibility to communicate networks that cannot be successfully connected using traditional routing schemes. They allow communications in very sparse networks and also where traditional end-to-end transmissions are not possible.

This work begin with the description and characteristics of opportunistic networks and opportunistic routing schemes, it is also presented an extensive description of the classification for these schemes. Furthermore, in this work an important interest is given to the routing schemes that used mobile infrastructures in its communication, such as Message Ferrying and DataMULEs.

Finally the main focus is given to two study cases of the DataMULEs routing scheme, they were in detail analyzed and their features and functionality were addressed to summaries the benefits.

Título: Survey of Opportunistic Network Schemes for Message Forwarding

Autor: Sandra Rodríguez Aranguren

Director: Roc Meseguer Pallarès

Fecha: 28 de Noviembre del 2013

Resumen

El objetivo de esta Tesis de Maestría es ofrecer los resultados de la investigación llevada a cabo sobre los diferentes esquemas de enrutamiento usados en las redes oportunistas, y que son llamadas de igual forma esquemas de enrutamiento oportunista. Así mismo, el esquema de enrutamiento con estructuras móviles "DataMULEs" es de especial interés para este Tesis y será tratado a fondo en el contenido de la misma.

Los esquemas de enrutamiento oportunistas son de gran interés y el objetivo de de muchos estudios técnicos, ya que ellos ofrecen la oportunidad de comunicar redes, que no se podrían comunicar usando los esquemas de enrutamiento convencionales. Estos esquemas permiten el enrutamiento de los datos a través de redes dispersas en una gran área geográfica y donde sus enlaces sufren de constantes desconexiones, es decir, donde las transmisiones convencionales, con conexión constante de origen a fin, no se posibles de llevar a cabo.

El presente trabajo comienza con la definición de lo que significa redes y esquemas de enrutamiento oportunista así como las características más importantes de ellos, también se presenta una clasificación de los diferentes esquemas de enrutamiento y sus sub-divisiones. Para los esquemas de enrutamiento con estructuras móviles como "Message Ferrying" y "DataMULEs" se dará una descripción más exhaustiva.

Finalmente el mayor foco de estudio será para dos casos de estudio de los esquemas de enrutamiento de "DataMULEs", estos serán objeto de un completo análisis y se presentarán sus características, funcionalidades y beneficios.

INDEX

INTRODUCTION 1

CHAPTER 1. OPPORTUNISTIC NETWORK AND ROUTING 2

1.1 Opportunistic Network 2
1.1.1 Concept 2
1.1.2 Opportunistic Network Features 3
1.1.3 Opportunistic Network Architecture 3

1.2 Routing Background and Overview in Opportunistic Networks 4
1.2.1 Routing in MANET and in Opportunity Network 5
1.2.2 Layers and Routing Solutions in Opportunistic Network 5

1.3 Opportunistic Routing Approach its Basic and Operations 6
1.3.1 Opportunistic Routing Concept 6
1.3.2 General Process Flow in Opportunistic Routing Schemes 8

1.4 Challenges in Opportunistic Routing Schemes 8

1.5 Objective and Evaluation Metrics for Opportunistic Routing Schemes 9

CHAPTER 2. OPPORTUNISTIC ROUTING SCHEMES 11

2.1 Opportunistic Routing Schemes Classification 11
2.1.1 Routing Schemes without Infrastructure 12
2.1.2 Routing Schemes with Infrastructure 13

2.2 Without Infrastructure: Dissemination Based Schemes Classification 14
2.2.1 MV Scheme (Meetings-Visitings) 15
2.2.2 Epidemic Scheme 16
2.2.3 Network Coding 17

2.3 Without Infrastructure: Context Based Schemes Classification 18
2.3.1 CAR Scheme (Context-Aware Routing) 18
2.3.2 MobySpace Scheme 19

2.4 With Infrastructure: Fixed Infrastructure Schemes Classification 20
2.4.1 Infostations Scheme 20
2.4.2 SWIM Scheme (Shared Wireless Infostation Model) 21

2.5 With Infrastructure: Mobile Infrastructure Schemes Classification 22
2.5.1 MF Scheme (Message Ferrying) 22
2.5.2 DataMULEs 24

CHAPTER 3. DATA MULES ROUTING SCHEME 26

3.1 Introduction 26

3.2 Overview and Definition 26

3.3 Comparison to other Schemes 27

3.3.1 Mobile Infrastructure Schemes ... 27
3.3.2 Sparse Wireless Sensor Network Data Collectors ... 28

3.4 DataMules Three-Tier Architecture ... 28

3.5 DataMULEs Scheme Benefits and Limitation ... 30
3.5.1 Benefits ... 30
3.5.2 Limitations ... 31

CHAPTER 4. STUDY CASES AND CONCLUSIONS ... 32

4.1 Study Cases ... 32
4.1.1 DataMULEs for Energy Efficient Study Case ... 32
4.1.2 DataMULEs Store Management and Opportunistic Data Collection Study Case 36

4.2 Results Analysis ... 41
4.2.1 DataMULEs for Energy Efficient Result Analysis ... 41
4.2.2 DataMULEs Store Management and Opportunistic Data Collection Result Analysis ... 43

CHAPTER 5. CONCLUSIONS ... 47

BIBLIOGRAPHY ... 49

INTRODUCTION

Routing schemes represent an important issue in network communications, they have the task to transport the data to its destination, to achieve these routing schemes must find and choose the most convenience path that will drive them to the final destination. Routing schemes are mostly separated between traditional and opportunistic routing schemes. In traditional routing schemes the searching and selection of the path is made before the data is sent; in opportunistic routing schemes the searching and selection of a path to destination is made as the data is transmitted.

For opportunistic networks, where links between source and destination are very variable and suffers from frequent disruptions, traditional routing schemes are not appropriate; they are not robust enough to support the transmission of data in such unpredictable condition as they assumed that must be end-to-end path to its destination. Opportunistic routing scheme overcome these problems taking advantage of the broadcast features of wireless communications, here the next hop to destination is decided dynamically as the data is transmitted and taking into account the current conditions of network, the broadcast feature allows to transmit the data to multiples nodes in one transmission, thus having more possibilities to find the next best hop close to destination. Thus, opportunistic routing schemes offer many benefits concerning to the data transmission through a challenge network; they increase the robustness of the data transmission, reduce the number of reliable transmissions, create strong links from the combination of weak links and they are able to take advantage of unexpected circumstances that occur in the network in favor of the transmission. For all of these, are suitable for applications that tolerate high error rates and delays.

All of these aspects will be addressed in detail in the following chapters, describing the basis and background of the opportunistic routing schemes to follow of a structural classification of these schemes. The rest of the work will be focus in the research of the schemes with mobile infrastructure schemes and in special an intense study of DataMULEs.

Chapter 1. OPPORTUNISTIC NETWORK AND ROUTING

In this chapter will be introduced the concepts of opportunistic networks and routing. First will be explained the concept of opportunistic networks and which are their distinct network structure in its architecture, as well some of its main features will be presented. Then will be describing the challenges and solutions in routing suchlike networks, so then the concept and features of an opportunistic routing could be introduced, here will be explained how a basic opportunistic routing scheme works and what are the challenges that such a routing scheme must overcome. Finally some common aims and metrics for opportunistic routing schemes are giving.

1.1 Opportunistic Network

1.1.1 Concept

Opportunistic Network is an evolution of the Mobile Ad hoc NETworks (MANETs) [1] and for most of authors Opportunistic Networks are considerate as a sub-class of Delay/Disruption Tolerant Networks (DTNs) with a more flexible environment [2][3]. For simplicity and generalization purpose, throughout this document, we use the term Opportunistic Network for both DTNs and Opportunistic Networks.

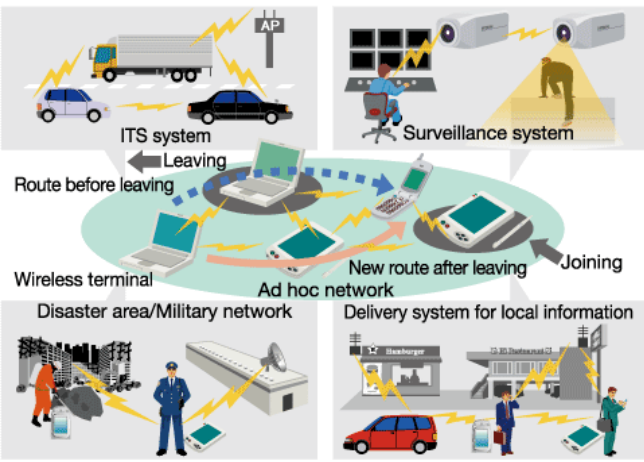


Figure 1. Mobile Ad-hoc Network Structure Example [4].

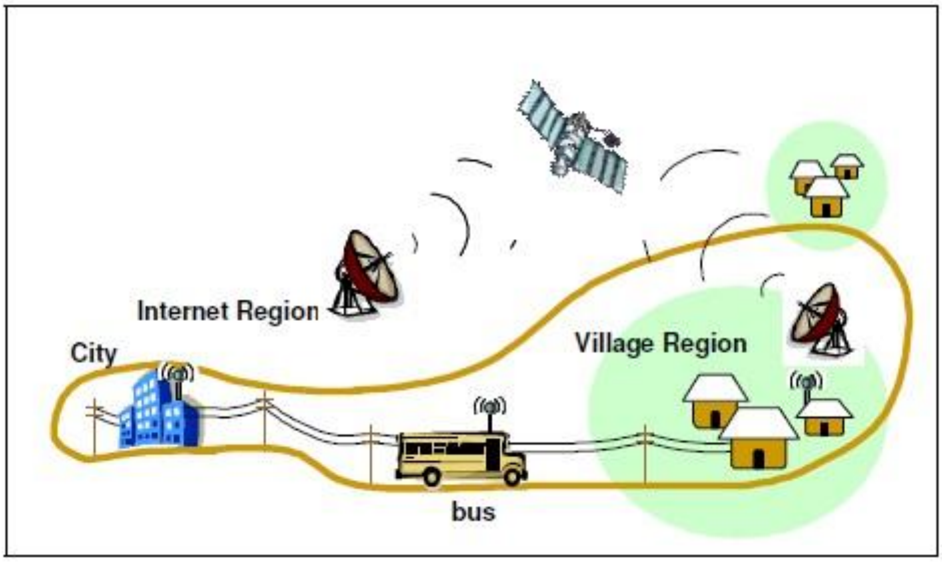


Figure 2. Delay/Disruption Tolerant Networks Structure Example [5].

Depending on the capabilities and characteristics of the nodes that form the network an Opportunistic Network can be described as a network where nodes are not connected all the time and are spare in an area without a direct path between them.

1.1.2 Opportunistic Network Features

From the articles [6] and [3], the following features can be addressed as the most distinguished of an Opportunistic Network:

- Infrastructure-less networks.
- Network contacts are intermittent.
- Link performance is highly variable or extreme.
- Frequently a complete path between source and destination does not exist or is very unstable and break quickly.
- Nodes are in constant movement and shut down periodically to save energy, these provoke constant changes in the network's topology.

1.1.3 Opportunistic Network Architecture

Opportunistic Networks have a very dynamically challenging topology, which make harder its process of forwarding the data. This challenging topology is considerate to be separated into several network partitions. This network partitions are the result of the intermittent connectivity features of opportunistic networks where links between its nodes are periodically disconnected; these regions appear every time a node is not available to communicate with others, maybe due to the node moving away of the communication range or the node turning off its power for save energy reason causing.

According to Authors in [3] the typical architecture in an opportunistic network is a structure where the devices in the different regions of the network are enable to interconnect to each other by forwarding the message with a store-carry-forward switching mechanism approach; this switching mechanism is carry out by the intermediate nodes by overlapping a new protocol layer “the bundle layer protocol” Figure 3, which is on top of heterogeneous region-specific lower layers of the Transmission Control Protocol (TCP) layer or its similar for transport protocol layer for opportunistic routing. The bundle protocol layer behaves as a convergence layer protocol to provide an easier transition to the other layers. The protocol stack is shown in the next figure [3].

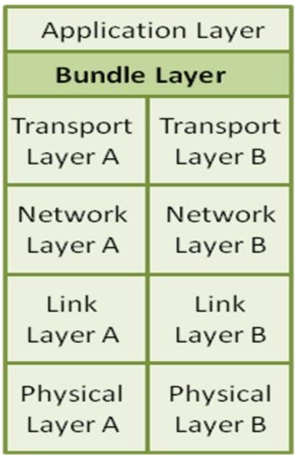


Figure 3. Protocol Stack [3].

In this architecture every node is consider as an independent component of the whole network, the node uses its bundle layer to be carry out the forwarding of packet to its destination. The bundle layer allows the nodes to perform as a host, router or gateway and applied the store-carry-forward switching mechanism. How it is explained in [3] as a router, the node bundle layer can store, carry and forward the entire bundles (or bundle fragments) between the nodes in the same region. But if what is needed to be done is messages transfer across different regions the node will perform as a gateway, a gateway can forward bundles between two or more regions and may optionally be a host, so it must have persistent storage and support custody transfers

1.2 Routing Background and Overview in Opportunistic Networks

The most relevant aspect in network communications is the way how a message/packet is forwarded through the network. The application of specific routing schemes are required to make possible transmissions in a network, owing to the fact that routing schemes have the task to find and choose a suitable path for the communications to take place.

The routing schemes techniques applied for forwarding in a certain network will depend on the characteristics and peculiarities of the network itself. Before getting into details on what and how an opportunistic routing is, we will review to its premises.

1.2.1 Routing in MANET and in Opportunity Network

As explained before due to the different features of an Opportunistic Network compare to a MANET, it is not viable to use routing schemes designed for MANET network for forwarding into Opportunistic Networks. The routing scheme use in an Opportunistic Network must be more sophisticated and robust to be able to manage the difficulties existing in a challenge network, that is, really few information about its topology and current status.

As it is explained in [2], routing schemes in Opportunistic Network have an asymmetric routing behavior that relies on the mobility of the mobile nodes to improve the possibility of encounters, also helps to create these encounters possibilities with other nodes that could be on the path to the destination. Routing schemes in MANET have a symmetric routing behavior that relies more on continuous end-to-end connectivity with a multi-hop behavior.

	Routing in MANETs	Routing in Opportunistic Networks
End-To-End Connectivity	Contemporaneous	Frequently Connected
Delivery Delay	Short	Long
Transmission Reliability	High	Low
Routing Behavoir	Symmetric	Asymmetric

Table 1. Difference between Routing in MANETs and Opportunistic Networks [2].

The table above (**Table 1.**) shows a lower reliability transmission and a longer delivery delay for routing in Opportunistic Network compared to the routing in MANET. This is due to the asymmetric and symmetric routing behavior of its routing schemes.

1.2.2 Layers and Routing Solutions in Opportunistic Network

In the past Section we explain the architecture structure of an opportunistic network and we explained how due to the topology and features of an opportunistic network, in contrast with traditional networks, a new protocol layer is used for transmission purpose. Besides each layer of the opportunistic networks have its own routing protocols or schemes to carry out the forwarding process.

The routing schemes use for The Network Layer are basically divided depending on the number of copies of a message forwarded by the node a single or multiple copies of the messages; when a routing scheme uses a single copy of the message to forward it, will be called a forwarding-based approach and when a routing scheme uses multiple copies of the message to forward it, will be called a flooding-based approach.

The Transport Layer of an opportunistic routing cannot use the existing traditional transport layer routing schemes (ex. TCP), these traditional routing schemes are not suitable for an environment where frequent disruption is a norm and end-to-end paths are typically not available. Some routing schemes propose for this layer are the Licklider Transmission Protocol (LTP) provides retransmission-based reliability, the TCP-based convergence layer (TCPCL) uses to link two bundle nodes and the Saratoga protocol, a rate-based UDP file transfer protocol that can also be used to transfer bundles. These routing schemes will not be further studied in this work, for detail information [3].

The Bundle Layer is a distinctive feature in opportunistic networks, which is responsible for storing, carrying and forwarding the data in the network. Routing schemes are group in unicast, multicast and anycast bundle delivery; multicast and anycast delivery approaches are used when there is more than one destination in contrast to unicast approach where there is only one destination. For purpose of our work we will study in more detail the routing schemes in the unicast approach.

The Application Layer, some traditional applications are delay-tolerant but they are not able to take advantage of the delay-tolerant nature of opportunistic networks, thus having a negative performance under these networks. This is due to that the routing scheme applied is not adequate for the network. To solve it some authors suggest amount other things the use of SMTP proxies to hide the disruptions between end users, or a web proxy by extending the World Wide Web Offline Explorer (WWWOFFLE). For more information refer to [3].

1.3 Opportunistic Routing Approach its Basic and Operations

1.3.1 Opportunistic Routing Concept

The traditional routing schemes (*best-path approach*) do not performance well in these environments; since it tends to forward the package to the best path known before transmission; this causes many package retransmissions and high loss rates. Where in Opportunistic Routing schemes the next hop to destination is selected after the package is transmitted, proving to be a good solution for these networks; opportunistic routing schemes mitigate the impact of high-loss-rate channels by exploiting the broadcast nature of wireless transmissions [7].

Opportunistic routing is an integrated routing and MAC technique for multiple-hop wireless networks that realizes some of the gains of cooperative diversity on standard radio hardware. It is based on broadcast transmissions to attempt to find the best path to destination using for each packet a dynamic selection of the next hop before it is transmitted. The protocol can use long radios links that have high loss rates combining multiple weak links into a strong link and take advantages of transmissions that reach unexpectedly near or unexpectedly far which provides the capability to adapt itself to very dynamic topologies.

For each destination a set of next-hop candidates are selected and ordering according to a certain criteria, then the sender broadcasts the packet and among thus nodes that hear the transmission start a process on which they must agree on their identities and choose a forwarder so the one closest to the destination is selected to forward the packet.

This technique is used to increase the probability that at least one potential relaying node receives the packet, reduces the number of transmissions for a reliable delivery as it avoids retransmission, resulting in an increase of the throughput [8]. It is intended for loss-sensitive application and avoided where the delay needs to be optimized.

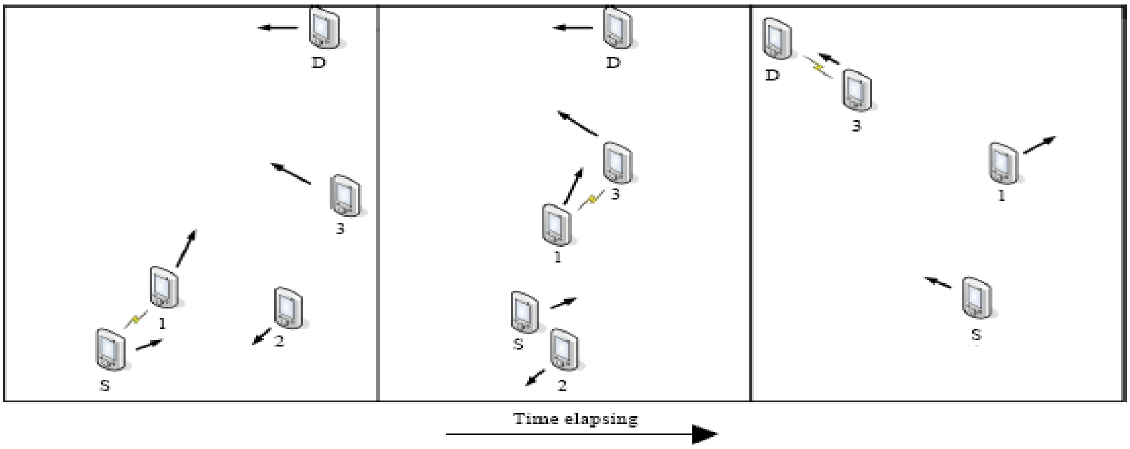


Figure 4. Time evolving in Opportunistic Networks [9].

The figure above illustrates a package transmission in a standard opportunistic routing scheme. The package is sent to its destination (D) from the source (*S*); the source broadcasts the package to all nodes in its range ratio, then through a dynamically particular process the next hop closest to destination is selected, this process repeats until the package finally arrive to its destination with the help of the intermediate nodes (1-5).

In general opportunistic routing is supported by the collaboration of intermediate nodes that retransmit the messages from the source to the final destination. These protocols take advantage from broadcast transmission that allows multiple nodes to receive the same data packet with just one transmission. Then, the receptors of the packet coordinate to elect one of them as the next transmitter. It is important to remark that this election is supported by the current propagation characteristics so it differs from one packet to another. Thus, opportunistic routing leads to a load balancing while it increases the robustness of the multihop wireless network as multiple receptors are potential relays [8].

1.3.2 General Process Flow in Opportunistic Routing Schemes

The specific steps that a particular routing scheme will flow for forwarding a packet to its destination will greatly vary from protocol to protocol, but we can summarize the most distinguish steps that every opportunistic routing scheme will flow in the transmission process, as it is described [8].

A sequentially forwarding process includes the following steps:

- Candidate Selection: a set of nodes is selected according to the protocols specification criteria, called as forwarding candidates, they may be ordered in a list according to some criteria in the second phase. The source informs them including the IDs of the candidates in the packet header.
- Candidate Priority Assignment: here the presented forwarding candidates are ordered in a list according to chosen appropriateness metrics and they will act as relying nodes. The nodes periodically measure its metrics parameters in order to keep the list updated. Usually these metrics could be related to the MAC layer as for example the probability loss.
- Data transmission: opportunistic routing schemes in general are supported by the transmission of broadcast packets, thus multiple neighbor-nodes can receive the packets, but for some unicast opportunistic routing schemes the best forwarding node is specified in the next hop field of the packet and the other candidates receive the packet by eavesdropping.
- Receiver coordination: Among the forwarding candidates that receive the data packet, just one of them should be the relaying node for the current packet. The elected node will be also responsible for confirming the data reception at the MAC layer. The goal of the procedure is that the selected node should be the highest-priority relay that has successfully received the packet.

1.4 Challenges in Opportunistic Routing Schemes

As mentioned before routing in opportunistic network is a challenging task, mainly due to the constant movement of its nodes and the intermittent connectivity features, there is not a direct path between source destination and that is why traditional routing are not able to support the forwarding process. To overcome this problems and in order to communicate with each other the nodes in an opportunistic network implement the store-carry-forward approach, where they must take any change of contact with its nodes neighbors and try to forward the information further in its path, and these are the main challenge in an opportunistic routing: the contact opportunity and the node storage. And it is explained in [3]:

- **Contact:** Due to the node mobility or the dynamics of wireless channel, a node might make contact with other nodes at an unpredicted time. Since contacts between nodes are hardly predictable, they must be exploited opportunistically for exchanging messages between some nodes that can move between remote fragments of the network.
- **Storage constraint:** As described above, to avoid dropping packets, the intermediate nodes are required to have enough storage to store all messages for an unpredictable period of time until next contact occurs. In other words, the required storage space increases a function of the number of messages in the network. Therefore the routing and replication strategies must take the storage constraint into consideration.

1.5 Objective and Evaluation Metrics for Opportunistic Routing Schemes

Authors in [2] propose some factors in order to achieve an effectiveness performance of an opportunistic routing:

- **Bandwidth**: This factor determines the number of messages that can be transmitted at each encounter opportunity. In addition, to transmit messages according to a corresponding priority is beneficial to utilize the bandwidth.
- **Buffer Space**: The sufficient buffer space is essential for the carried messages; since they would be buffered for a long period time until the upcoming encounter opportunity is available. In light of this, to discard the least important message due to buffer space exhaustion is beneficial to utilize the buffer space.
- **Energy**: An Opportunistic Network device often has limited energy and cannot be connected to the power supplier easily. Hence, the routing algorithms, which transmit few messages and perform less computation, are more energy efficient.

In [2] are addressed some metrics that can be use to evaluate the opportunistic routing schemes:

- **Delivery Ratio**: It is given by the ratio between the number of delivered messages and the number of generated messages.
- **Overhead Ratio**: It is given by the ratio between the number of message transmissions required for delivery and the total number of messages delivered.
- **Delivery Delay**: It is given by the time duration between the messages generation and their delivery.

The routing objective provides a tradeoff between maximizing the delivery ratio and minimizing the overhead ratio. On one hand, the ideal case of delivering the message before its given lifetime with the lowest overhead ratio is to keep this message until the destination is in proximity. While on the other hand, the effective approach to maximize the message delivery ratio is to relay this message at each encounter opportunity taking into account the candidate node selection. Although it is expected that the applications of Opportunistic Networks are inherently tolerant to the long delivery delay, this does not mean they would not benefit from short delivery delay.

Chapter 2. OPPORTUNISTIC ROUTING SCHEMES

This chapter is entirely focused on the classification of opportunistic routing schemes. Parting from the base that this work will be developed for unicasting schemes in the first part of the chapter will be explained the classification that was chose. In the next section will be explained the upper part of the classification and a description of its several parts is giving. The next sections will contained a basic description of every opportunistic routing scheme of the chosen classification; also it will be explained in general terms how these schemes work.

2.1 Opportunistic Routing Schemes Classification

Currently exists many classification for opportunistic routing algorithms depending of the author of the study, a first classification is the one given by authors in [6] where they explain a classification where the existing routing algorithms can be classified in unicasting, multicasting and anycasting issues; nevertheless the aim and focus of this study will be the unicasting schemes, where the message is delivered to its unique destination. For more information about the others classifications refer to [6].

The classification for unicast schemes is made according to the schemes design and characteristic like route design and location deployment. According to [1] a classification for the unicast opportunistic schemes is for those algorithms designed for completely flat ad hoc networks, the so called Schemes without infrastructure and the algorithms in which the ad hoc networks used some kind of infrastructure, these are called schemes with infrastructure.

In the followings sections will be explained the different schemes with infrastructure and without infrastructure and their subdivisions. Thus the aim of this work is the study of the schemes with mobile infrastructure; these will be studied with more detail.

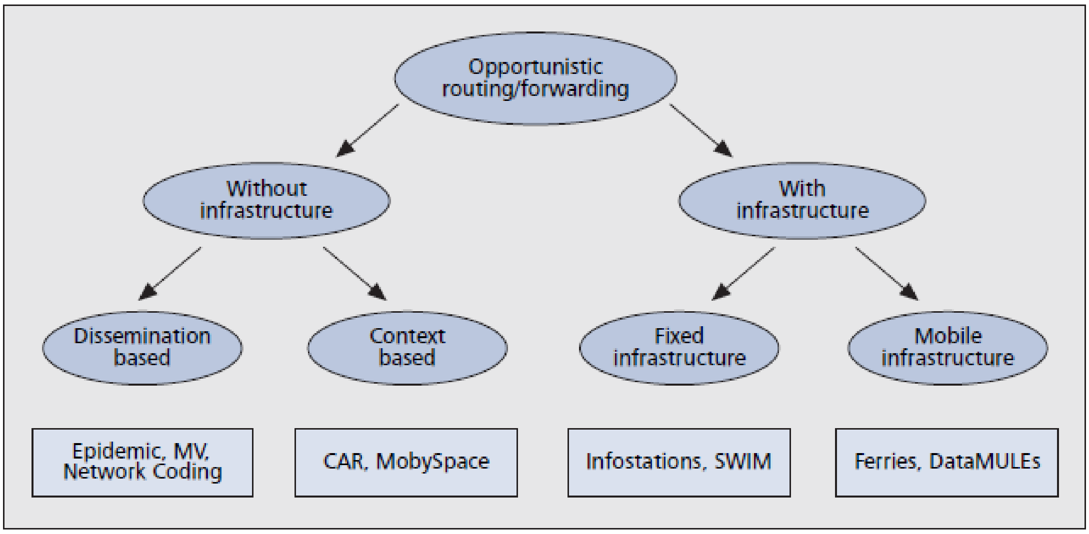


Figure 5. Taxonomy of Routing Schemes for Opportunistic networks [1].

2.1.1 Routing Schemes without Infrastructure

These schemes are designed for completely flat ad hoc networks and do not use any kind of infrastructure to forward the message throughout the network. They can be subdivided into dissemination-based and context-based routing schemes [1].

Dissemination-based algorithms are essentially forms of controlled flooding, and differentiate themselves by the policy used to limit flooding. Context-based approaches usually do not adopt flooding schemes, but use knowledge of the context that nodes are operating in to identify the best next hop at each forwarding step [1].

2.1.1.1 Dissemination Based

Some of the most distinguished features for these sub-division schemes presented in [1] are:

- Delivers a message to destination by diffusing it all over the network (flooding).
- No knowledge of a possible path or appropriate next-hop node.
- Well performance in highly mobile networks with frequently contacts opportunities.
- Tend to limit message delay.
- Consume a lot of resources.

- High contention and network congestion.
- Increase the network capacity by limiting the number of hops of the spreading radius or by limiting the message copies present at the same time.
- Network-coding-based-routing outperforms flooding as it is able to deliver the information with fewer messages in the network.
- Limits messages' flooding by exploiting knowledge about direct contact with destination nodes.

2.1.1.2 Context Based

Some of the most distinguished features for these sub-division schemes presented in [1] are:

- Exploits information about the context in which nodes are operating to identify next hops.
- More reduction of message's duplication compare to dissemination.
- Tend to increase message's delay in the delivery process.
- The computational cost is higher.
- Nodes need to maintain a state to keep track of the utility value.

2.1.2 Routing Schemes with Infrastructure

These are schemes that use some kind of infrastructure to deliver the message in an opportunistic form to its destination. A subdivision is made according to the type of infrastructure they rely on; in this case there are fixed infrastructure schemes and mobile infrastructure schemes. Special nodes of a fixed infrastructure are located at specific geographical points, whereas nodes of a mobile infrastructure move around in the network following either predetermined known paths or completely random paths, which will be collaborated in the forwarding of the message [1].

2.1.2.1 Fixed Infrastructure

Some of the most distinguished features for this sub-division schemes presented in [1] are:

- The message is sent only when a base station belonging to the sender node is reachable.
- Base stations are generally gateways towards less challenged networks.
- Two variations to the protocol are possible: one in which node-to-base-station communications is allowed. Another is in which node-to-base-station and node-to-node communication are allowed.
- Node-to-base-station communications experience high delay in delivery message.

2.1.2.2 Mobile Infrastructure

Some of the most distinguished features for this sub-division schemes presented in [1] are:

- Nodes of the infrastructure are mobile data collectors.
- The routes that the nodes follow can be predetermined or arbitrary.
- The nodes gather the messages from the nodes that pass by.
- In exclusive node-to-carrier communications, the special carrier-node is the only entities allow of delivering the messages.
- In exclusive node-to-carrier communications the carrier-node help to increase connectivity in sparse networks and also isolated node can be reached.
- In communications where is allow to nodes to communicate to carriers and ordinary nodes the delivery of the messages is accomplish by both.

2.2 Without Infrastructure: Dissemination Based Schemes Classification

In the following section will be addressed a detail description and procedure of some of the schemes belonging to this classification.

2.2.1 MV Scheme (Meetings-Visitings)

The following information is an overview of the features and procedures that defines the MV routing scheme; this information was taken from the work paper [10].

2.2.1.1 Features:

- Forwarding of messages from a mobile source to a stationary destination.
- Exploits movement structure by learning the peer’s motion patterns.
- These periodic patterns in peer’s movements (structures) are used to estimate the probability of delivery for a specific message and peer.
- The information of participants meetings and their visits to locations is used for routing and buffer allocation.
- Making informed routing decisions by estimating the probability of a particular message being delivered by a given peer.
- Uses autonomous agents to compensate for a mismatch between available capacity and demand, that is, they are able to adapt the bandwidth and latency requirements of a network with the movement of network participants and traffic flows.
- Improves the performance of disruption-tolerant networks.

2.2.1.2 How it works:

When a peer A meets another peer B, they perform a message exchange through a number of steps. First, A gives to B a list of the messages A carries as well as their destinations. Each message is also annotated by A with A’ likelihood of delivery according to the formula we derive below. A receives the same list from B and calculates the likelihood of delivering B’s messages. A now sorts the unioned lists by likelihood of delivery, removes its own messages, and also deletes all messages that B has a higher likelihood of delivering. A then selects the top n messages remaining, and requests from B all the messages that are not already stored. Note that MV calculates an estimation of delivery likelihood assuming an infinite buffer at each peer and limits the number of hops that are required in practice.

2.2.2 Epidemic Scheme

The following information is an overview of the features and procedures that defines the Epidemic routing scheme; this information was taken from the work papers [11] [12].

2.2.2.1 Features:

- Distribute application messages to hosts called carries, within connected portions of a network.
- Node mobility is used to put in contact carriers with another connected portion of the network.
- Deliver messages with a high probability of reaching their destinations.
- It maximizes message delivery rate and minimizes message delivery latency by placing an upper bound on message hop count and per-node buffer space.
- It poses high aggregate resource consumption.
- Delivery of messages to arbitrary destinations with minimal assumptions regarding the underlying topology and connectivity of the underlying network.
- Relies upon transitive distribution of messages through ad hoc networks.
- Use a hash table indexes a list of messages that are originated and buffering by the host.

2.2.2.2 How it works:

A source, S, wishes to send a message to a destination, D, but no connected path is available from S to D. S transmits its messages to its two neighbors, C_1 and C_2, within direct communication range. At some later time, C_2 comes into direct communication range with another host, C_3, and transmits the message to it. C_3 is in direct range of D and finally sends the messages to its destination. The hash table used indexes the list of messages, keyed by unique identifier associated with each message, also each host stores a bit vector, called the summary vector that indicates which entries in their local hash table are set.

2.2.3 Network Coding

The following information is an overview of the features and procedures that defines the Network Coding routing scheme; this information was taken from the work paper [13].

2.2.3.1 Features:

- Significantly reduces the overhead of probabilistic routing algorithms.
- Nodes code the information over the contents of several packets they receive.
- Nodes send packets with linear combinations of previously received information.
- It has a parameter that controls with which probability the reception of an innovative packet causes the node to send a packet.
- The memory available at a node constrains the number of information units a node can code over.
- Network coding is much more robust against packet loss than others algorithms.
- Allows the dissemination of information with a high probability at a lower overhead.
- It is affected by the generation size.

2.2.3.2 How it works:

An intermediate node sends a linear combination $\sum_i g_i x_i$ of the packets x_i it has received or wants to send, where all coefficients g_i and the data x_i are interpreted as numbers in a finite field. When a node receives m combinations of n packets, it can decode them provided that the set of combinations has a rank n, for m=n this occurs with a probability close to 1 if the coefficients g_i are chosen randomly and independently at every node.

2.3 Without Infrastructure: Context Based Schemes Classification

In the following section will be addressed a detail description and procedure of some of the schemes belonging to this classification.

2.3.1 CAR Scheme (Context-Aware Routing)

The following information is an overview of the features and procedures that defines the CAR routing scheme; this information was taken from the work paper [14].

2.3.1.1 Features:

- It is based on intelligent placement of messages.
- Achieve efficient and timely delivery of message by the evaluation and prediction of context information.
- Hosts are not aware of its or others hosts absolute geographical location, they only have logical connectivity information about its position.
- CAR uses predicted future values for the context.
- It only creates a single copy of each message.
- High level of scalability thanks to the almost constant value of the overhead in terms of the number of messages exchanged regardless of buffer size.

2.3.1.2 Hoe it works:

Each host calculates its delivery probabilities based on the prediction of future values and its composition. The calculated delivery probabilities are periodically sent to the other hosts in the connected cloud for updating. Each host maintains a logical forwarding table of tuples describing the next logical hop and its delivery probability for all known destinations. The updates are sent only when the evolution of the mobile scenario follows a certain trend. If a host does not store any information about the message recipient, it sends the message to the host in the cloud that has the highest mobility. If the carrier, while moving, meets a host with a higher delivery probability, the message is transferred to the host with higher delivery probability.

2.3.2 MobySpace Scheme

The following information is an overview of the features and procedures that defines the MobySpace routing scheme; this information was taken from the work papers [15] [16].

2.3.2.1 Features:

- Uses a high-dimensional Euclidean space (MobySpace) constructed upon node's mobility patterns.
- The nodes are represented by the coordinates that correspond to their probability of being found in each possible location.
- The messages are routed in a virtual space defined by the basis of mobility patterns.
- In the virtual space there is no notion of neighbor.
- Nodes are directly connected in a dynamically way to other nodes that are close or very far from each other.
- Nodes opportunistically take advantage of connections that promise to advance bundles toward the destination.
- The routing decision rely on the notion that a node is a good candidate for taking custody of a bundle if it has a mobility pattern similar to that of the bundle's destination.
- The efficiency of the virtual space tool may be limited if nodes change too rapidly their habits.
- Reduce bundle delay and through lower communication costs.

2.3.2.2 How it works:

For each of the nodes there is a well defined probability of finding that node at each of the N locations, which is based upon historic information regarding contacts that the node has already had. This set of probabilities is a node's mobility pattern, and is described by a MobyPoint, in an N dimensional MobySpace (Euclidean virtual space). The MobyPoint is the coordinates of the mobility pattern of a node. The bundles in the MobySpace are routed by sending them to nodes having mobility patterns that are successively closer to the mobility pattern of the destination.

2.4 With Infrastructure: Fixed Infrastructure Schemes Classification

In the following section will be addressed a detail description and procedure of some of the schemes belonging to this classification.

2.4.1 Infostations Scheme

The following information is an overview of the features and procedures that defines the Infostations routing scheme; this information was taken from the work papers [17].

2.4.1.1 Features:

- Offer geographically intermittent coverage at high speed, since data can often tolerate significant delay.
- It placed isolated, low cost, low power, short range base stations (Infostations), in locations that offer easy proximity to users and equip these with enough bandwidth.
- Infostations share a single block of the spectrum because they are separated
- Users can connect to the network in the vicinity of ports (or Infostations), which are geographically distributed throughout the area of network coverage.
- Provides strong signal quality to small disjoint geographical areas.
- Suitable for applications, which can tolerate significant delay.
- It is used for messaging and information systems.

2.4.1.2 How it works:

Infostations are simple devices and can be organized in different ways. Individual Infostations (placed in highways, streets and airports) may function independently, providing specialized gateways to the Internet and accessing remote servers for information and messages. The access between the Infostation and gateway could be wired or dial-up, depending on traffic and latency demands.

2.4.2 SWIM Scheme (Shared Wireless Infostation Model)

The following information is an overview of the features and procedures that defines the SWIM routing scheme; this information was taken from the work papers [18].

2.4.2.1 Features:

- It is a modification of the traditional Infostation model by offering geographical intermittent coverage at high speeds.
- Allows additional improvement in the capacity-delay trade off through a moderate increase in the storage requirements.
- Allows information to travel trough the network, in addition to only the user itself.
- The information reaches the Infostation by sharing (replicating and diffusing) itself in the network, using the mobile nodes as physical carries.
- Information packet is discarded from the nodeswhen the node reaches a Infostation and the packet is offloaded to the network.
- The Infostation's carrier stores the identities of packets that were offloaded and reject them if they are offloaded again to the Infostation.
- Significantly decreases delays due to the increased redundancy as the packets are allowed to spread throughout the mobile nodes, but the consumption of network capacity increases.
- The model is well-suited to delay-tolerant applications as biological information acquisition systems. In this case for whale data acquisition.

2.4.2.2 How it works:

The SWIM stations will be placed on buoys, floating in the water or in birds flying around the area. Each whale will have a RF for data transmission and a tag. Whale's tag will not only store its own data, but also information from other whales transmitted whenever they are within range transmissions. Several stationary SWIM station will be distributed within the transmission range of the known whale paths. Data collection of one day can be considered one packet. Each packet carries a distinct identifier, so that a whale or a SWIM station will reject packets which have already been received in the past. Each packet is stamped with a time stamp indicating its creation time and the required Time-To-Live. When the Time-To-Live of a packet expires, it is discarded from the whale's tags. The collected information is divided into separate packets, which are stored in the memory of the whale tag. As a whale surfaces and exposes its RF antenna, it may transmit to a close-by SWIM station, usually one within the line of sight. Anytime one whale comes within transmission range of another whale; the tags on the two whales share information by exchanging the stored information packets. Anytime a whale comes within transmission range of a SWIM station, the information from the whale tag is transmitted to the station and erased from the tag. When the time stamp on the packet expires, the packet is discarded from the tag's memory. After receiving and storing the information from whales, the SWIM stations would transmit the information to shore, either by coordination with another SWIM stations or directly to a satellite whenever it pass by.

2.5 With Infrastructure: Mobile Infrastructure Schemes Classification

In the following section will be addressed a detail description and procedure of some of the schemes belonging to this classification.

2.5.1 MF Scheme (Message Ferrying)

The following information is an overview of the features and procedures that defines the MF routing scheme; this information was taken from the work papers [19].

2.5.1.1 Features:

- It is a mobility-assisted approach, which utilizes a set of special mobile nodes, called message ferries for communication between nodes in the network.
- Introduce non-randomness in the movement of the nodes to help deliver data.
- Improve data delivery and energy efficiency.
- The nodes are classified as messages ferries (ferries) or as regular nodes based on their roles in communication.
- Ferries take responsibility of carrying messages among other nodes, while regular nodes do not have this responsibility.

- The use of ferries allows the communication between distance nodes that are out of range.
- Depending on whether ferries or nodes initiate non-random movement there are schemes: NIMF (Node-Initiated MF) and FIMF (Ferry-Initiate MF).
- Most of the communication involves short range radios, only in FIMF long range radios are used for small control messages.

2.5.1.2 How it works:

NIMF (Node-Initiated MF), a node operates in 4 modes: WORKING, GO_TO_FERRY, SEND/RECV and GO_TO_WORK. A node is initially in the WORKING mode and moves according to its assigned task. The trajectory control mechanism of the node determinates when it should proactively move to meet the ferry for receiving and sending messages, the ferries moving in its specific route periodically broadcast HELLO messages to nodes. The node respond to the Hello message with an Echo message so the ferry and the node can identified each other then the node enters to GO_TO_FERRY mode when it decides to go to the ferry and approaches it. When the node detects the ferry is within its transmission range, the node enters the SEND/RECV mode and exchanges messages with the ferry. After transmission ends the ferry move out of range and the node enters the GO_TO_WORK mode to return to its location prior to the detour. Upon return to the prior location, the node enters the WORKING mode. When the node meets the ferry unintentionally the node switch from SEND/RECV to WORKING directly. *FIMF (Ferry-Initiate MF)*, a node can be in two modes: DISASSOCIATED and ASSOCIATED. A node is initially in the DISASSOCIATED mode when there is no a service request from the ferry. The notification control mechanism determinates whether the node should send a service request message to the ferry. After sending a request message to the ferry, the node enters the ASSOCIATED mode and waits for the interaction with the ferry. When a node is in the ASSOCIATED mode, notification control determines when to send a Location_Update message to notify the ferry of the node’s new location. In both modes, the node may exchange messages with the ferry if it is close and receives a HELLO message from the ferry. After interaction with the ferry, the node returns to the DISASSOCIATED mode. The ferry operates in two modes: IDLE and WORKING. Initially the ferry is in the IDLE mode and follows a specific default route. It periodically broadcast its location information to nodes via long range radio. Upon the reception of a service request message from node, the ferry switches to the WORKING mode where the ferry maintains a set of nodes H that have request service and tries to meet these nodes to relay messages. When a request is received, the ferry updates H, computes a new ferry route and adjusts its movement to follow the new route. When the ferries arrives at the location of a node reported in its request or update messages, the ferry assumes that it has finished the visits with the node and removes it from H. When H becomes empty which means the ferry has visited all requesting nodes, it returns to the default route and enters the IDLE mode. In both modes, when the ferry comes close to a node, the ferry may exchange messages with the nodes.

2.5.2 DataMULEs

The following information is an overview of the features and procedures that defines the MF routing scheme; this information was taken from the work papers [20].

2.5.2.1 Features:

- Exploit mobiles entities present in an application scenario, in this case the use of mobiles nodes present in the sensor field as forwarding agents.
- Its objective is to collect data from sensors and deliver it to an access point in the infrastructure.
- The mobile entities called MULEs (Mobile Ubiquitous LAN Extensions) carry data from sensor to access point.
- The MULEs are assumed to be capable of short-range wireless communication.
- MULEs exchange data with sensors and access point as they pass by, buffer this data and drop it off at an access point.
- Requires less transmission power, but latency increased as the sensor waits for MULEs to deliver its data.
- Connectivity is provided by adding an intermediate layer of mobile nodes to the existing sensors layer and access-points layer.
- Extends the lifetime of the network by minimizing the communication responsibility of the resource-constrained sensors.
- It is limited to non real time applications which have mobility.

2.5.2.2 How it works:

A MULE continuously sends out a discovery message to detect a nearby sensor. At each sensor there is a queue of generated data that has not been delivered. When a MULE is in the sensor’s range the queue is served, so the data is transfer from the sensor’s queue to the MULE. The sensor then waits for the next MULE arrival event to transfer the data. Then the MULE transfer the data to an access point, the interaction between the MULEs and the access-points can be modeled on exactly the same principles as with the sensors.

Chapter 3. Data MULEs Routing Scheme

The aim of this chapter is to give a more detailed analysis and study of the opportunistic routing scheme DataMULEs, as it is the chosen schemes to analyze on this work. Here will be explained all the important aspects of this scheme, starting with a definition and comparison to other related schemes; also the relevant features of its architecture will be described. The chapter ends with an overview of the benefits and limitations that this schemes offers.

3.1 Introduction

As explained before, the “Routing Based on Mobile Infrastructure” approach is based in an infrastructure where its nodes are mobile data collector. As noted in [1], these mobiles nodes are described as carriers, supports, forwarders, ferries, mules etc. They are distributed in the network area and move with or without a predetermined path, and on this movements they collect, store and forward messages to its destination or next hop. In general and basic terms this is the functionality of the DataMULEs routing scheme.

DataMULEs are suitable for collecting and transmitting data in a wireless sensor network (WSN). The WSN is composed by a large number of sensors that are sparse in a large area. These kind of networks introduce problems to the communicate between each other sensors, making the transmission of the data to a specific destination which is usually a central controller (base station or access point) a challenge for the system; since in a sparse network the sensors are located far away from each other the communication process required a high level of energy consumption, this reduce the lifetime of sensors and in this way the lifetime of the network.

3.2 Overview and Definition

The routing scheme DataMULEs help to overcome the communication problems, helping in the increase of the network’s connectivity and guaranteeing that isolated nodes can be reached [1]. It uses mobiles entities usually nodes called Mule (Mobile Ubiquitous LAN Extensions), as they “carry” data from sensors to access point [20], as forwarding agents. The forwarding of the messages is carried out by opportunistic contacts between the mules through the wireless interface; they also display random pattern movements that are not predictable.

DataMULEs can be implemented in a large range of application as environmental monitoring, traffic monitoring, animal migrations and others application where the WSNs are isolated and unattended. Depending on the application the mobiles mules can be chosen as animals, humans, cars, ships, etc.

This scheme introduced a three-tier architecture that connects spare sensors within a WNS; for this a third new layer is added just between to the two existing layers of sensors and access points, in this new layer is where the mobile mules are placed. DataMULEs' architecture collect the sensor data in the WSN using the presence of these mobile entities "MULEs", the mule's task is to pick up the data from sensors when they are in close range to it, then buffer it, and finally drop off the data to its designated wired access point [21].

The three-tier architecture consists of a top tier with access points set at convenient places, also a middle tier consists of the mobile mules that are mobile nodes with unknown movements, and finally a bottom tier consists of sensors that are randomly distributed across a region [6].

3.3 Comparison to other Schemes

3.3.1 Mobile Infrastructure Schemes

As it is explained in Chapter 2, we mentioned DataMULEs and Message Ferries with its two versions (NIMF and FIMF) as mobile infrastructure approaches which aim is, as it name indicates, to use its moving structures to collect the data and forward it to its destination.

In the survey article [2] and with the information obtained from [19] and [21] a comparison among unicast schemes with mobile infrastructure assistance is giving, from this comparison can be obtained **Table 2**., in which is possible to compare DataMULEs features with the features of the two other versions of Message Ferries scheme: NIMF (Node-Initiated MF) and FIMF (Ferry-Initiate MF).

Routing Scheme	Infrastructure Movement	Assistance Behavior Controlling	Buffer Space
DataMULEs	Deterministic Movement	None	Limited
NIMF	Deterministic Movement	None	Limited
FIMF	Deterministic Movement	Controlling the movement based on request	Limited

Table 2. Comparison among unicast schemes with mobile infrastructure assistance. [2]

Based on the information founded in the mentioned surveys, the table above shows some features for these schemes; for instance they use as infrastructure the same deterministic movement, also their buffer space is limited for all the three schemes. The only difference found for these schemes is the way they control their assistance behavior; in DataMULEs and NIMF any control of the movement of their assistances (mules and ferries) is used as they (mules and ferries) just move randomly through the network until they find an opportunity to transmit. Unlike to DataMULEs and NIMF the FIMF approach control the movement of its ferries, as the ferries must adjust their trajectory as they received service request as explained in chapter 2.

3.3.2 Sparse Wireless Sensor Network Data Collectors

Besides the DataMULEs there are others schemes to collect data that uses the approach of a sparse wireless sensor network, these are the Base Station approach scheme and Ad-hoc Network approach scheme. As it is mentioned in [21] the Base Station approach consists of a few numbers base station scatter through the whole network area, here every sensors have the possibility to directly communicate to the base station, which will be the closest to each sensor; for the Ad-hoc Netowrk approach the number of sensors are big enough so it is possible to create an ad hoc network with them, these sensors transfer the data through the ad hoc network by multi-hop routing to the wired base station.

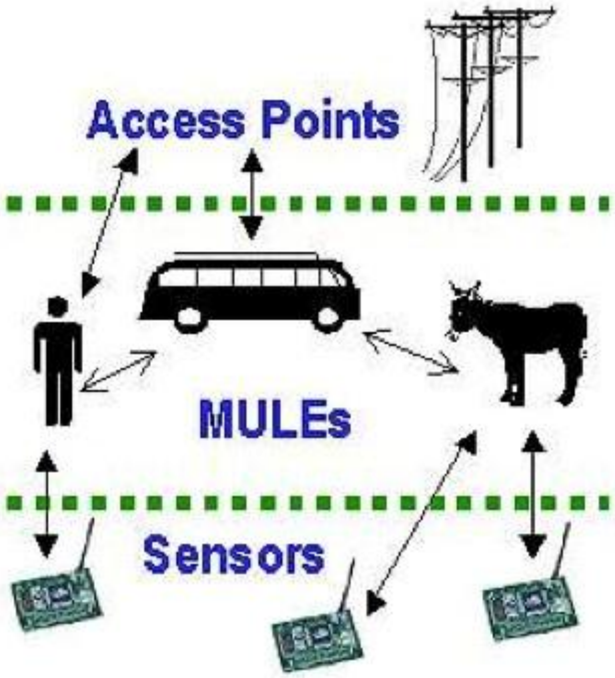
As it can be appreciate in the next **Table 3**. DataMULEs gets an overall better result than its similar approaches of the Sparse Wireless Sensor Networks Schemes. In terms of efficiency DataMULEs performs the best for the best price, and even with high latency. Its perform overcome that from the other mentioned approaches as in DataMULEs the sensor power and infrastructure cost are low, which are important features to considerate in order to implement any application, and as long as the application to implement is not a real time application, where the latency can make the communication not possible.

	Performance Metrics			
Approaches	**Latency**	**Sensor Power**	**Data Success Rate**	**Infrastructure Cost**
DataMULEs	High	Low	Medium	Limited
Base Stations	Low	High	High	Limited
Ad-hoc Network	Medium	Medium-low	Medium	Limited

Table 3. Performance of Different Approaches for Data Collection in Sparse Wireless Sensor Network [21].

3.4 DataMules Three-Tier Architecture

As mentioned before the DataMULEs architecture adds a new layer to the conventional sensor and access point layers of a WSN, this architecture is called a three-tire architecture. It provides wide-area connectivity for a sparse sensor network by exploiting the use of mobile entities (mules) [21]. These mules can be humans, animals, vehicles, ships or any other moving agent.



Figure#. DataMULEs three-tier architecture [20].

As it is explained in [20] and [21] the DataMULEs three-tier architecture is comprised by three tiers or layers:

• A Upper Tier:
In this level the layer consists of a set of wired access points or data repositories, which are set in convenient locations in the network area. They are equipped with Internet/Network's connectivity; enhance power, storage and processing capabilities. These devices are used to offload the data collected and stored by the mules, they communicate with a central data warehouse that enables them to synchronize the data that they collect, detect duplicates, as well as return acknowledgment to mules (ack maybe necessary to ensure reliability of data for certain applications).

• A Middle Tier:
This intermediate layer consists of mobile agents, named "MULEs", which move around an area covered by sensors to gather their data and transfer it to the access points. A mule has the responsibility to discover sensors and access points and to transfer the data between them.

Thanks to the mobile mule nodes this layer provides the system with scalability and flexibility for a relatively low cost. Its most important feature is that in comparison to the sensors mules are equipped with a larger storage capacity, renewable power, and have the ability to communicate with sensors and networked access points, as they are in proximity to each other. Mules are assumed to be unpredictable mobile entities whose mobility patterns cannot be calculate or predict in advance. However as a result of their motion, they collect and store data from the sensors, as well as deliver acks back to the sensor nodes. In addition, in some cases mules can communicate with each other to improve system performance.

• A Bottom Tier:
This layer is occupied by sensor nodes, which are distributed randomly across the network area. They are stationary sensors with not mobility features and periodically perform data sampling from the surrounding environment, which will depend on the application purpose. Sensors, provides data, communicate via a short-range radio, and have limited power and memory; that is why they cannot perform a lot of work, in order to save energy.

The main advantage of this architecture is the potential of large power savings by sensors because communication now takes place over a short range [20], but they suffer from a high latency.

3.5 DataMULEs Scheme Benefits and Limitation

As presented in the work article [20], an overview of the benefits and limitation of the DataMULEs scheme.

3.5.1 Benefits

- Energy Efficient: Substantial energy is saved because sensors communicate over a short range. Moreover, there are no hotspots in the network, as sensors do not forward data for other sensors.

- Spatial Reuse: The three-tier architecture of DataMULEs exploits spatial reuse of bandwidth by using short-range communication without losing long term connectivity and avoids radio communication complexities such as collisions.

- No routing overhead: In contrast to ad-hoc networks, this architecture does not have any routing scheme overhead for sensors.

- Robustness: Performance degrades gracefully as mules fail. Any single mule failure does not lead to a disconnected network. The primary effect of a mule failure on the overall system is a slight increase in latency as there are now fewer mules to pick up data. In contrast, in an ad-hoc network failure of few critical nodes might lead to a disconnected network.

- Scalable: The DataMULEs three-tier architecture is easily scalable as deployment of new sensors or mules requires no network reconfiguration.

- Simplicity: The data routing aspect of the mule architecture is very simple and extremely lightweight for the sensors. This is important because sensors are the bottleneck of the system. This architecture does not require any synchronization or location information; an assumption made by many approaches. This help to reduce the costs of the implementation.

3.5.2 Limitations

- Latency: The DataMULEs architecture has high latency and this limits its applicability to real time applications (although this can be mitigated by collapsing the mule and access- point tiers).

- Best-effort delivery: Data delivery in the basic architecture is best effort; delivery is not guaranteed. The system requires sufficient mobility. For example, mules may not arrive at a sensor or after picking the data may not reach near an access-point to deliver it. Also, data may be lost because of radio-communication errors or mules crashing. To improve data delivery, higher-level protocols need to be incorporated in the mule architecture.

Chapter 4. Study Cases and Conclusions

This chapter will be focused in the analysis of some works done about the DataMULEs scheme. First will be described how the chosen studies were done, and which kind of consideration were taken into account. After this, and explanation and analysis of the result is giving. Finally some conclusions and new ideas for future works are presented.

4.1 Study Cases

4.1.1 DataMULEs for Energy Efficient Study Case

This study case and subsection is entirely based on the research made in [20]. This study focuses in to find a solution to the problem of energy efficient for data collection in a sensor network. The approach used in this study exploit the nature of the DataMules architecture in term of the mobile nodes present in the sensor field as forwarding agents.

In this work the authors introduced an analytical model that they used to understand the key performance metrics such as data transfer, latency to the destination and power. The parameters used in this work are: sensor buffer size, data generation rate, radio characteristics, and mobility patterns of mobile nodes.

4.1.1.1 Problem description:

For the authors the energy problem in this architecture is due to three things. One is that the energy needed to transmit data over one hop is quite a lot due to the distance between sensors, which can be large in some cases. Another reason is that in this architecture network sensors have to not only send their data, but also forward data for other reasons. The third reason for the energy problem given by the author is the routing hotspots near the access point, where sensors that are near to the access point have to forward many more packets and drain their battery much more quickly. All of these cause a huge demand in the energy consumption from the sensors, which could limit the network performance and design cost.

4.1.1.2 Model Contributions

This study provides a contribution for the understanding of DataMULEs scheme in terms of the modeling and comparison with ad-hoc networks. Here some its most important contribution:

- An analytical model is presented; it is based upon queueing theory that helps to understand the relationship between performance metrics and system parameters. The authors characterized the model's performance along with three dimensions: data transfer rate, latency, and energy requirements at the sensors. This model incorporates system parameters such as sensor data generation rate, buffer size, sensor duty cycle, radio characteristics such as range and capacity, MULE velocity, MULE mobility model, etc.

- Detailed performance simulations to validate the analytical model studied and to gain finer understanding, with comparisons between the results.

- An analytical discussion of the benefits of the MULE architecture over ad-hoc networks both qualitatively and quantitatively using simulation. The model proves some advantage in the implementation of these networks as in comparison to the ad-hoc network, this is in term of operational lifetime and other related features.

- An efficient discovery process of sensors is addressed by using a low duty cycle at the sensor and this is incorporated in the analysis. A novel discovery mechanism is discussed that permits significantly lower duty cycles while at the same time has very little impact on performance.

4.1.1.3 Model Characteristics

The authors address the following characteristics for their study case:

- The model application context is focused on sensor networks unlike previous work to this, where the focus was towards mobile ad-hoc networks.

- This model tries to maximize sensor network lifetime by reducing the communication energy required at the sensors.

- The architecture used introduces MULE explicitly and encompasses Zebranet (for more exact information on ZebraNet project [22]) like scenarios, and focus on analytical modeling, energy efficient discovery and comparison with ad-hoc networks, which previously to this work were unaddressed.

- The DataMULEs architecture is as was presented in section 3.4 with the typical three tiers architecture: lower tier (sensors), middle tier (MULEs), upper tier (Access points).

- In this model MULEs do not communicate with each other.

- In this model the Access points are the eventual destination of sensor data. They are used to offload the data collected by and stored in the MULEs.

- Depending on the scenario, tiers in the model's architecture can be collapsed onto one device, increasing the applicability of our architecture that is sensors can be mobile so the sensor and the MULE tier could be mapped to the same device, also if MULE(s) have Internet connectivity they can act as an access-point, combining the MULE and access-point tiers.

4.1.1.4 Model Description

The primary component of the model is a queue of generated data, which has not been delivered at each sensor. In queuing theory terminology, generation of new data at a sensor corresponds to an arrival at the sensor's queue. The buffer size of the sensor defines the capacity of the queue. If the buffer is full then any new data is dropped. The queue is served whenever a MULE is in a sensor's range.

The arrival of a MULE in a sensor's range is considered as a discrete event. This event causes transfer of data from the sensor's queue to the MULE. The sensor then waits for the next MULE arrival event to transfer the data. Thus, the time between two MULE arrivals dictates when the queue is served.

The amount of data that can be transferred on a MULE arrival event is a random variable and depends on factors such as, the time the MULE is in the communication range of sensor. However, in the model this is taken as a fixed quantity.

The interaction between the MULEs and the access-points can be modeled on exactly the same principles. The authors' model only focuses on interaction between sensors and MULE, as they consider this to be the primary bottleneck of the system.

The MULE discovery process for this model is as follow: a sensor needs to discover a nearby MULE to be able to offload its data. In this architecture the prime responsibility of discovery is placed on the MULE, as the objective is to minimize the load on sensors. A MULE continuously sends out a discovery message to detect a nearby sensor. This requires a sensor to listen for discovery messages. Since listening consumes as much power as receiving, as it was proved in [23], the duty cycle at the sensor need to be reduce, which leads to a tradeoff between minimizing the listen energy and maximizing the probability of rendezvous with a passing MULE. The model implemented proves to offer a good impact of sensor duty cycle, by increasing the contact time the sensors can operate at low cycles without substantially affecting the performance.

4.1.1.5 Performance metrics and parameter

Metrics used in this study case:

- *Data success ratio (DSR):* This measures the effectiveness of data delivery.

- *Latency:* This is the average time taken by data to reach access-points from the time of its generation.

- *Communication energy:* it is considered both, the average energy consumed per sensor as well as the worst-case consumption, which dictates the network lifetime.

Parameters used in this study case:

- *Sensor related:* The data generation rate (λ) defines the average amount of data that a sensor is generating. This directly affects the buffer requirements at the sensor.

- *MULEs related:* The primary aspect is to determine when MULEs come into the communication range of a sensor. The MULE arrival within a sensor's range is modeled as a discrete event. The key parameter is the distribution of time between two MULE arrivals at a sensor. Our model abstracts out these complexities by assuming the knowledge of inter-arrival distribution. MULEs buffer size is another parameter, but for the purposes of this paper we assume that MULEs have sufficiently large buffers.

- *Access point related:* The important aspect here is the distribution and the number of access-points. This affects how frequently a MULE visits an access-point to deliver data.

- *Radio related:* The radio parameters affect the amount of data that can be transferred as a MULE passes by a sensor. We use a radial model for the radio, i.e. sensors and MULEs can communicate if they are within a distance *r*.

Our approach is to identify a few basic parameters that are sufficient to characterize the performance metrics. These basic parameters are: (1) sensor data generation, (2) sensor buffer size (*SB*), (3) amount of data transferred between a MULE and a sensor, denoted by *K* (4) MULEs arrival at a sensor and (5) a MULE's visit to access-points. The effect of other parameters can be understood by first studying how they change one or more of the basic parameters and subsequently studying how the performance is affected by the change in basic parameters. For example, the impact of increasing MULE velocity on performance can be examined in two steps. First, by examining the impact of increasing MULE velocity on the basic parameters. In this case, it increases the MULE arrival rate at the sensors/access-points and decreases K (see Section 5.4). Second, the analytical model is used to analyze the affect on performance due to the changes in these basic parameters. The effects of sensor duty cycle are modeled in a similar manner.

4.1.1.6 Assumptions Model

The following list provides a summary of assumptions and key notational symbols made in this study case:

- The MULEs arrival process at a sensor is a renewal process $\{S(t), t \geq 0\}$, where $S(t)$ is the total number of MULEs that have visited the sensor up until time *t*. The renewal assumption means that the inter-arrival times (time between arrivals of two MULEs) are independent and identically distributed.
- At a given time only one MULE interacts with a given sensor and vice-versa. Also, we assume that when a MULE visits a sensor no other sensor is nearby (and contending for service). This is reasonable because our networks are sparse. This assumption is verified using detailed simulations.
- Sensors are identical. Although not essential, we will assume that sensors are not mobile for ease of exposition.
- The data generation process at a sensor is a renewal process.
- The queueing discipline is FCFS. The data that is generated first is picked up first.
- MULEs have sufficiently large buffers.
- Data transmission does not incur any loss. The only loss is due to sensor buffer overflow.
- The queueing system is stable and only the stationary (time independent) probabilities are considered. These are the probabilities as *t.*

4.1.2 DataMULEs Store Management and Opportunistic Data Collection Study Case

This study case and subsection is entirely based on the research made in [24]. In this study the authors analyzes the problems of store management in DataMULEs networks and propose some ideas to solve it.

In order to find a solution to the storage management problem this work focuses on modeling the limited memory spaces of a WSN's (Wireless Sensor Network) sensor nodes as a distributed storage system.

4.1.2.1 Problem Description

As DataMULE's architecture is based on disconnected WSNs, these are separated into multiple isolated groups and do not have network connectivity to outside world; and for that reason is necessary to dispatch some mobile mules to visit the isolated WSNs from time to time in an opportunistic way to relay the sensing data back.

The problem is that sensors' memory spaces are limited and data collection from isolated WSNs to mules and then to BSs (Base Stations) relies on opportunistic communications in the sense that contact between these entities is occasional, storing and collecting higher-priority data is necessary and therefore the authors considerate this a critical issue to be addressed and to be analyzed.

4.1.2.2 Model Contribution

- The main problems analyzed in this model are the data storage management in each isolated WSN and the opportunistic data collection between these entities.
- A work that addresses and analyzes distributed storage strategy for isolated WSNs with opportunistic communications using mobile mules.
- Possible solutions to improve the data quality of collected packets in extreme sparse and resource-limited distributed WSNs.
- A novel concept to model the distributed storage of a WSN that could facilitate virtualizing the data management among sensor nodes' buffers.
- An opportunistic data exchange model, which could improve the data quality of collected packets.
- Implementation strategies in a real sensor platform that verifies the feasibility of the model.

4.1.2.3 Model Characteristics

Some of present the following characteristics:

- A network scenario with three components is considerate for this work. The three components are: (i) some static but disconnected WSNs, (ii) some mobile mules with uncontrollable mobility and (iii) some statics base stations accessible by mules.

- Mules have their own designated routes or destinations that are not under control of this model and therefore communications must rely on opportunistic contact between these entities.
- There are for types of communications in the model: (i) inter-communication within a WSN, (ii) opportunistic WSN-to-mule communications, (iii) opportunistic mule-to-mule communications, and (iv) opportunistic mule-to-BS communication.
- It is assumed that each BS has connectivity to the external world, thus the goal is to deliver the sensing data to BSs.
- Distributed Storage Management (DSM) and Opportunistic Data Exchange (ODE) strategies are the solutions propose to solve the storing and data collecting problems.
- The Distributed Storage Management (DSM) strategy is based on a novel shuffling mechanism similar to heap sort that allows nodes to exchange sensory data with neighbors efficiently in a distributed manner.
- The Opportunistic Data Exchange (ODE) strategy is based on a utility model that guides two mules to exchange data that would lead to a higher reward, this keep data with higher priorities closer to the sink.
- DSM uses a mesh-like structure to facilitate data exchanges.
- ODE drives a probabilistic model to guide the mule-to-mule data exchanges so as to maximize the expected reward of delivering sensing data to BSs.

4.1.2.4 Model Description

The model consists of a heterogeneous WSN and inside this network three different components can be identified: (i) some static but disconnected WSNs, (ii) some mobile mules with uncontrollable mobility, and (iii) some static base stations (BSs) accessible by mules.

All the isolated WSNs in the network are composed of some static sensor nodes, or simply nodes, these nodes are used to monitor the environment and periodically generate reporting packets with information of the actual network state, transmission sates and others; furthermore these nodes can also generate simply packets with messages. Sensor nodes are homogeneous and each has the same number of storage spaces. Multi-hop routing is supported in each WSN, but due to the network features, where WSNs are isolated and far from each other, WSNs rely on mobile mules to visit and carry their sensory data to their destination.

The model assumes that each isolated WSN has a designated sink node that will be visited by mules occasionally; it is also assumed that the movements of these mules are uncontrollable; that means, mules have their own routes or destinations which are not under control of the model system presented. Mules are considerate to be hikers, taxis, buses, and animals, any moving entity that is in the area network.

The model system presents four types of communication forms, these are: (i) inter-node communications within a WSN, (ii) opportunistic WSN-to-mule communications, (iii) opportunistic mule-to-mule communications, and (iv) opportunistic mule-to-BS communications. All these entities (sensor nodes, mules, and BSs) have the same communication interface.

The process of communication is as follow: a mule may stop by a WSN at any time and leave at any time. Therefore, communications only happen by opportunity; an opportunistic communications happens when two entities have communication contact. During these communication contacts, the sink should relay more important sensing data to the mule first. Any two entities can communicate with each other if they are within each other's transmission distance. Therefore, a piece of sensing data needs to be delivered from its originating sensor node to its sink, from the sink to a mule, from a mule to perhaps multiple mules, and finally from a mule to a BS. It is assumed that packets generated by sensor nodes are prioritized according to their importance (distance to the sink node) and an aging process is also applied to the priority of a report.

The data storage strategy for WSNs and the opportunistic data collection strategy among mules are applied to this communication scenario to overcome the multiple relaying activities that must be accomplished.

4.1.2.5 The DSM Strategy

The Distributed Storage Management (DSM) strategy is a distributed solution based on a shuffling mechanism. The DSM propose for this model by the authors consider only one isolated WSN and focus on its storage management problem, and then all WSNs will follow the same strategy to store their sensing data.
They assumed the existence of a predefined region nearby a particular WSN's sink called the Buffer Area (BA), a predefined region nearby its sink, which is defined as those nodes within 10 hops from the sink. The set of nodes in BA will act as a distributed storage system to store sensing data for the WSN when no mule is visiting it. When a mule arrives, these nodes will forward their data to the sink following some rules.

It is also assumed that each node knows its distance to the sink as well as its neighbor set. The process to obtain the distance between nodes and sink is easily achieved by sending a simple broadcast from the sink node; then to obtain the neighbor set, all the nodes inside the WSN exchange hello messages among them. All static sensor nodes will try to forward their packets toward nodes in BA at any time. Nodes not in BA will forward their packets to BA, while nodes in BA will observe neighbors' states and exchange packets with each other, if necessary.

Regarding the storage spaces in BA as a distributed storage system, the model's goal is minimize the dropping of packets in BA, and if the dropping of packets is unavoidable, the first packet to drop is the one with lower priority and so on and so forth. Another goal is to facilitate mobiles mules to collect data, for this, packets with higher priorities should be stored closer to the sink. To facilitate the determination of neighbors' priority for its classification each node will announce the priority of its packet when a new packet is generated.

DSM will eventually stop in an in-order status, the process is the following: a packet is stable if it is stored in a certain node and will not be exchanged with other nodes' packets, until a mule arrives or new packets with higher priority are being generated; each packet will become stable in finite time, which means that DSM will eventually stop, and when this happens each node is in-order. DSM utilizes the rich mesh links in a WSN to exchange packets. Higher-priority packets have more chances to stay closer to the sink.

4.1.2.6 The ODE strategy

The Opportunistic Data Exchange (ODE) is applied among mules so that more important data could be delivered to BSs. For ODE the authors have designed a strategy for each of the three types of opportunistic communication present in this model (WSN-to-mule, mule- to-mule, and mule-to-BS).

For **WSN-to-mule communications**, after DSM has already arranged packets as "in-order" in BA, a simple best-effort uploading strategy will achieve the data-collection goal; for this, when a mule arrives at the sink of a WSN, the sink will try to transmit as many packets in BA to the mule as possible until it loses the contact with the mule. By broadcasting an UPLOAD message, the sink will trigger data transmission from downstream nodes toward itself in a greedy way. After the sink makes sure the reception of a packet by the mule, it can drop the packet so as to make a space for subsequent packets. Once the sink loses the contact with the mule, it will broadcast a FINISH_UPLOADING message to trigger our DSM in the WSN.

The **mule-to-mule communications** designed for this model is as follow, each mule may has different probabilities to contact with BSs in the future, so a utility-based packet exchange strategy based on the contact probability will guide data exchanges between mules so as to maximize the reward of packets arriving at BSs. When two mules (*u* and *v*) have contact the ODE strategy follows these steps:

1. Each mule will consider whether copy packets from itself to another by sorting its packets based on the current utility of packets.
2. Mule *u* considers the packet "*p*", which has the highest utility and which has not been considered yet. Also, *v* considers the packet "*q*", which has the highest utility and which has not been considered yet. Then both *u* and *v* compute the benefits respectively.
3. If copying *p* makes a higher benefit, *u* copies *p* to *v*; otherwise, *v* copies *q* to *u*. Due to the CSMA channel used in this model, copying both packets *p* and *q* in this step is not allowed. Also, since the contact duration between mules is unpredictable, a best-effort copy policy is applied.
4. The packet that is copied is marked as ''considered''. If *u* and *v* are still within

each other’s communication range, go to step 2. Otherwise, stop.

For **mule-to-BS** communications, when a mule has contact with a BS a best-effort uploading strategy is applied by the mule to transmit packets of higher utility first.

4.2 Results Analysis

In the following subsections the results of study cases for Energy Efficient and for Store Management will be presented. These results are based on the ones given by the authors in their articles [20] and [24].

4.2.1 DataMULEs for Energy Efficient Result Analysis

The authors of this study focused their results on some key metrics aspects; which were based on their analytical study (which can be reviewed in detail in [20]). The authors performed simulations and examined the effects of varying metrics as μ (MULE Arrival Rate), *SB* (Sensor Buffer) and *K* (Data Transferred in One Interaction) as well the effects of the different mobility models applied to the network. They also compared the MULE’s results of varying sensor density in metric like energy ratio and hotspot ratio with the obtained in a traditional ad-hoc network. The results analyzed is as follow

4.2.1.1 MULE Arrival Rate μ *and Sensor Buffer* SB *Metrics*

Here simulations were carried out varying the metrics parameters μ (MULE Arrival Rate) and *SB* (Sensor Buffer) and examined their effects on average sensor buffer occupancy, data success ratio and on latency.

The simulations proved that when μ (MULE Arrival Rate) is increased the average buffer occupancy of the sensor decrease; these simulations also showed that the *SB* (Sensor Buffer) is not so relevant, as changing the sensor buffer size does not have much effect on buffer occupancy; only in the case when the MULE arrival rate is small which causes excessive load on the system. This performance results can be explained as follow: when the arrival of the MULEs become more frequently there is less amount of data generated between two consecutive arrivals.

As well, the increase of μ (MULE Arrival Rate) made the DSR (Data Success Ratio) to increase sharply until it gets close to one. This is because when μ is large, the buffer occupancy decreases and therefore less data are dropped. The DSR is also higher, when *SB* is larger. This is expected because when *SB* is large, less data is dropped. In general, one can increase DSR by either increasing μ or *SB*.

To determinate the effect of this two metrics on latency, it was assumed that K was sufficiently larger than the SB; with this assumption the results showed that the queuing delay is simply the residual life of the MULE arrival process, which decreases as μ, is increased, also *SB* has no impact on latency.

4.2.1.2 Data Transferred in One Interaction K Metric

In this case as the authors chose for simulations a large SB of 1MB the DSR is always close to one.

For the average sensor buffer occupancy and latency the simulations results showed that when K is small the average sensor buffer occupancy and latency are large. As it is explained in (30) this is because a sensor cannot transfer all the data in the queue to a MULE during a single contact, which made the average buffer occupancy to increase, as well as the latency also increased because a data unit has to wait for multiple MULEs to arrive before it can be served. In the other hand when K is large, both average sensor buffer occupancy and latency decrease sharply, this is just until certain limit of K, where it further increase will not have effects in the performance of average sensor buffer occupancy and latency. This is because K only needs to be large enough so as to absorb the occasional burst in the sensor buffer.

The next table showed an overview of the impacts in performance due to the increase of the metrics: *μ*, *SB* and *K*.

	Performance Metrics		
Parameter	**Buffer Occupancy**	**DSR**	**Latency**
μ	decrease	increase	decrease
SB	no effects	increase	no effects
K	decrease	increase	decrease
λ	increase	decrease	increase

Table 3. Effect of metrics increase on performance [20].

4.2.1.3 Mobility Models

This study considered four mobility models: (1) Random waypoint (2) Random Walk (new direction is chosen on reaching a street intersection) (3) Deterministic (MULEs arrive at fixed interval) (4) Poisson arrival. The results showed that for all mobility models as μ increases, the DSR increases and the latency decreases. The performance is best when the MULE arrival is deterministic and worst under the manhattan model. The performance of random-waypoint model closely matches that of poisson model.

4.2.1.4 DataMULEs vs Ad-hoc network

Varying the sensor density in the network and examining its impact on the metrics average energy ratio and hotspot ratio the authors proved the benefits of using mules' schemes when compared to the ad-hoc network. The average energy ratio was defined as the average energy consumed at a sensor in the ad-hoc network to the energy consumed in the MULE architecture. As well the hotspot ratio was defined as the ratio of hotspot usage in ad-hoc network to the hotspot usage in the MULE architecture, the hotspot usage is the maximum energy consumed by any sensor.

They proved that when the sensor density is low, the MULE architecture has less average energy consumption; this is because with few sensors the average distance between two sensors is large and the communication energy increases; this changes as the sensor density increases and saturate with the average energy ratio. And even the performance worse as sensor density get higher, mules' architecture is more efficient as traditional ad-hoc network, due to mule's avoidance to multiple hops.

The results obtained in hotspot ratio experiments are similar as the previous on average energy ratio; they indicated that MULE architecture experienced a much longer life-time than ad-hoc networks.

4.2.2 DataMULEs Store Management and Opportunistic Data Collection Result Analysis

In this study authors focus their results on the performances of the two strategies DSM and ODE applied to the network and under certain parameters.

To analyze the DSM strategy, the authors' examined the performance result of a WSN in terms of its priority distribution after the DSM is applied. They also examined the performance of DSM in two different topologies, mesh structure and tree structure, thus to understand the influence of the topology in the data collection.

In the case of the ODE strategy, the authors examined the performance varying three parameters and their impact on three different metrics; and thus obtain some notions on how ODE has an impact on the information quality and how they are improved by this strategy.

4.2.2.1 DSM Performance Results

Using two different topologies, mesh structure and tree structure, several experiments were executed to prove the benefits of the DSM strategy. Authors varied the numbers of nodes in BA and examined the effects on the average priority of packets at nodes in BA. As explain before (section 4.1.2.5) BA is the buffer area. The results showed that although the hop distance from each node to the sink is same in the shortest-path spanning tree of the mesh and the mesh itself, using a mesh structure has potential to collect higher-priority packets than using a tree structure, which can be explained due to mesh feature that allows much more directions of data exchanges, which will keep the higher-priority packets in BA.

In another experiment they varied the mule's visiting period and the contact duration. With this, they proved that a longer visiting period means that more important packets may be generated/collected during two consecutive visits. Also it was proved that a longer contact duration means that less important packets also have a chance to be collected.

They also examined the effect on the performance as the BA's size (hops to sink) changes, the results showed that the top 1/3 area in the network are mostly occupied by high-priority packets as BA is defined as five hops or more from the sink. As the BA is getting larger the number of packets exchanges (transmission overhead) also increase with a slightly decrease at the beginning. This is because packets with lower priorities have to travel longer to reach the BA when the BA is relatively smaller (3–4 hops). As the BA becomes larger (more than five hops), the cost of packet exchanges inside the BA is more dominant as before, and the transmission overhead is increasing again. Another factor as BA's size change is the average number of dropped packets, decreases as BA becomes larger, and this is because a larger size of BA can keep more packets and avoid dropping packets when they arrives BA.

To study the effects of network density has on the performance results, the number of sensor nodes in the field was varied. The results showed that increasing the packet arrival rate does not increase the overall transmission overhead proportionally; this is because once the BA has collected sufficiently important packets, the competition cost within BA will drop rapidly. It was also proved that a denser network will cause a higher cost of packet exchanges because sensor nodes have more neighbors to facilitate data exchanges for the new packet.

Finally, the DSM convergence time (minimal time for all packets in BA to be in order) is studied, for this different data arrival rates were considered. Here they varied the number of nodes and the size of BA; results showed that a dense network incur a longer convergence time because more number of packet exchanges are performed, besides the convergence time is not proportional to the data arrival rate, because packet exchanges in the network at a higher data arrival rate will be triggered frequently, also the convergence time slightly decreases first and then increases again as the BA’s size increases. The convergence time slightly decreases first and then increases again as the BA’s size increases. This is because in a relative small size of BA (from 3 to 7 hops) lower-priority packets have to move by more hops to compete against the higher-priority packets in BA. When the network has a relative large size of BA (from 8 to 13 hops), much more packet exchanges will incur in BA so the convergence time becomes longer.

In general, the gaps of convergence time between different data arrival rates shrink as BA’s size increases, because a larger BA provides more storage spaces to allow more concurrent packet exchanges.

4.2.2.2 ODE Performance Results

The ODE strategy is analyzed in terms of varying some network parameters and checking their impact in some metrics chosen by the authors. The network parameters are: contact duration between mule-to-mule and between mule-to-BS, the meeting period with the BS, and the aging period of packets. The metrics chosen to examined the ODE’s performance are: (1) packet delivery ratio which is the ratio of the total number of packets collected by the BS to the total number of packets collected by mules, (2) utility delivery ratio which is the ratio of the total utility of packets collected by the BS to the total utility of packets collected by mules, and (3) the total utility of packets collected by the BS.

For the contact duration variation the authors noticed the following results. As the contact duration increased the ODE packet delivery ratio also increase and is better as it is compared to the other scheme Greedy; the different between the two schemes are not so big in the beginning but as the contact duration increases the different is more notable. Besides, they examined the utility delivery ratio and noticed that the results from ODE in contrast to the ones from Greedy, presented younger packets collected and have fewer number of copies. The other parameter analyzed when varying the contact duration was the total utility of packets collected by the BS here again ODE showed better results as Greedy, as the contact duration increase ODE improved much more amount of information quality as the Greedy.

The other aspect examined in this study was the impact on the performance of the meeting period between the mules and the BS. Results showed that as the meeting period increase the packet delivery ratio also increase, and in comparison to Greedy the ODE results presented at the beginning were slightly better but as the meeting period increase both ODE and Greedy presented almost exactly results. Regarding to the analysis of the utility delivery ratio it was proved that the utility delivery ratio decreased as the meeting period increased; furthermore it was proved that with a larger meeting period mules have fewer opportunities to meet BS, and besides it, ODE as opposite to Greedy was able to collect younger packets and with fewer number of copies, and ODE’s performance was better too even with a very low frequency of meeting. The results for the total utility of packets collected by BS showed also that as the meeting period increase the total utility of packets decease, and also as this happened ODE obtained always better as the Greedy.

As for the analyzed varying the aging period of packets it was proved; first, the packet delivery ratio was more or less constant as the aging period increase, and both ODE and Greedy get similar results with ODE having a little better performance, which is due that ODE avoided copying too many higher-priority packets. The metric utility delivery ratio slightly increases as the aging period increase with ODE performing better than Greedy. Finally it was shown that as aging period increase the total utility of packets collected also increase, and in greater proportion for ODE as it is compared to Greedy.

Chapter 5. Conclusions

This work has provided an intense and up to date overview of the opportunistic routing schemes, which includes it background and definition, an extensive description of the classification of these schemes, where each schemes is explained from the perspective of its features and functionality. Also comparison with other type schemes is given.

It can be summarized that under challenge wireless networks with intermittent contacts and where the sate of network is constantly changing and an end-to-end path to destination may never exist, opportunistic routing schemes offer a significant improve in performance. These schemes use the broadcast techniques of wireless communication to send data to several possible next hops in just one transmission and taking into account the current condition of the network situation, all of these help to avoid unnecessary retransmission and create robustness in the communication against disconnections. This also produces some drawbacks to these schemes as additional delay in messages delivery and the error rate also increase, therefore these schemes are to be applied in not real time application, which are delay-tolerant in nature and these drawbacks do not cause major trouble to the communication.

The classification used in this work divide the routing schemes according to the strategies they use for their transmissions, that is, the main classification is between the routing schemes that use some kind of infrastructures assistance in their transmission process and the routing schemes that do not use them. With all routing schemes described and their procedures explained it was noticed that schemes with mobile architecture as Message Ferrying and DataMULE present the more interesting features and opportunities, as they offer the possibility of creating opportunistic communications, this is because their approach is concerned on the deployment of their elements in the network that allows the creation of new and opportunistic route inside the network, this approach and the use of mobile agents to carry the data through the network to its destination improve the encounters opportunities and thus the capacity and connectivity.

For this thesis, the DataMULE scheme was chosen as the subject of a more detail study. DataMULE uses randomly moving entities called mules to opportunistically collect and transmit data through the network, this approach turn to overcome the main problems of wireless sensor networks sparse in big areas. With its particular three-tier architecture it effectively uses data mules to enable communications between the other two tiers of the architecture. The implementation of this scheme give the advantage of low energy consumption and easy deploy, the big problem that affects this scheme is the high latency levels, for this reason DataMULEs are considerate to be suitable for delay-tolerant application.

In relation to the two DataMULE study cases presented in this work; it can be concluded that the study case related to energy efficient proved that using the present of mobile mules as forwarding agents reduce the demand of energy in sensors and thus increase the lifetime of the network, despite this advantage this DataMULE architecture is only suitable for non real time applications which have mobility, being in this case one of the most effective scheme to take into account. The store management and opportunistic data collection study case concluded that the application of the two proposed strategies DSM and ODE improve the results of the network communication; their results showed that DSM is able to collect optimal data priorities and ODE is able to collect fresher and fewer duplicate packets; therefore DSM and ODE could efficiently work together to collect important data through isolated WSN networks; this strategies could be used in application for outfields and back countries scenarios.

This Master Thesis provided a broad overview of the actual state of the opportunistic routing schemes and in special an intense research of the DataMULE scheme, addressing their/its features, advantage and disadvantage. The aim of this work is to offer help and a study material to the people interested in the study and research of such schemes.

Bibliography

[1] L. Pelusi, A. Passarella, and M. Conti, "**Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks**," *IEEE Commun. Mag.* vol. 44, N°. 11, pp. 134 –141, 2006.

[2] Y. Cao and Z. Sun, "**Routing in Delay/Disruption Tolerant Networks: A Taxonomy, Survey and Challenges**," *IEEE Commun. Mag.,* 2012.

[3] C.M. Huang, K. Lan, and C.H. Tsai, "**A Survey of Opportunistic Networks**," *IEEE Computer Society*, 2008.

[4] "**(Research Themes 2007) Ad Hoc Networks**," *Hitachi R&D Research Site*, http://www.hitachi.com/rd/portal/research/yrl/07/soc_02.html, last reviewed: November 11th 2013.

[5] "**DTN Network Model Evolution**," *Protogenist Blog*, http://protogenist.wordpress.com/2012/07/18/dtn-network-model-evolution/, last reviewed: November 12th 2013.

[6] Z. Zhang, "**Routing in Intermittently Connected Mobile Ad Hoc Networks and Delay Tolerant Networks: Overview and Challenges**," *IEEE Communication Surveys*, vol. 8 N° 8, 1st Quarter 2006.

[7] Z. Zhong and S. Nelakuditi, "**On the Efficacy of Opportunistic Routing**," *In Proc. of IEEE SECON*, Jun. 2007.

[8] A. Triviño-Cabrera and S. Cañadas-Hurtado, "**Survey on Opportunistic Routing in Multihop Wireless Networks**," *International Journal of Communication Networks and Information Security.* vol. 3, N°. 2, August 2011.

[9] L. Liu and Z. Chen, "**Data forwarding in Opportunistic Networks**," *Information Technology Journal 9 (2)*, 2010.

[10] B. Burns, O. Brock and B. N. Levine, "**Mv Routing and Capacity Building in Disruption Tolerant Networks**," *In Proceedings of IEEE INFOCOM*, vol. 1, pp. 398–408, 2005.

[11] A. Vahdat and D. Becker, "**Epidemic Routing for Partially-Connected Ad Hoc Networks**". *Technical Report CS-2000-06*, Duke University, July 2000.

[12] L. Song and D. F. Kotz, "**Evaluating Opportunistic Routing Protocols with Large Realistic Contact Traces**," *Technical Report*, Institute for Security Technology Studies (ISTS), Dartmouth College, September 2007.

[13] J. Widmer and J.-Y. Le Boudec, "**Network Coding for Efficient Communication in Extreme Networks**," *Proc. ACM SIGCOMM 2005 Wksp. Delay Tolerant Networks*, Philadelphia, PA, Aug. 22–26, 2005.

[14] M. Musolesi, S. Hailes, and C. Mascolo, "**Adaptive Routing for Intermittently Connected Mobile Ad Hoc Networks**," *Proc. 6th IEEE Int'l. Symp. World of Wireless, Mobile and Multimedia Networks (WoWMoM 2005)*, Taormina-Giardini Naxos, Italy, June 13–16, 2005.

[15] J. Leguay, T. Friedman, and V. Conan, "**Evaluating Mobility Pattern Space Routing for DTNs**," *Proc. IEEE Infocom 2006*, Barcelona, Spain, Apr. 2006.

[16] J. Leguay, T. Friedman, and V. Conan, "**DTN Routing in a Mobility Pattern Space**," SIGCOMM'05 Workshops, Philadelphia, USA, August 2005.

[17] D. Goodman *et al.*, "**INFOSTATIONS: A New System Model for Data and Messaging Services**," *IEEE VTC'97*, vol. 2, pp. 969–73, May 1997.

[18] T. Small and Z. J. Haas, "**The Shared Wireless Infostation Model — A New Ad Hoc Networking Paradigm (or Where There is a Whale, there is a Way)**," *Proc. 4th ACM Int'l. Symp. Mobile Ad Hoc Networking and Computing (MobiHoc 2003)*, Annapolis, MD, June 1–3, 2003.

[19] W. Zhao, M. Ammar, and E. Zegura, "**A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks**," *Proc. 5th ACM Int'l. Symp. Mobile Ad Hoc Networking and Computing (Mobihoc)*, ACM Press, pp. 187–98, May, 2004.

[20] S. Jain et al., "**Exploiting Mobility for Energy Efficient Data Collection in Wireless Sensor Networks**," *ACM/Kluwer Mobile Networks and Applications (MONET)*, vol. 11, no. 3, pp. 327–39, June 2006.

[21] R. Shah, S. Roy, S. Jain, and W. Brunette, "**Data mules: modeling a three-tier architecture for sparse sensor networks**," in *IEEE SNPA '03*, Anchorage, Alaska, USA, 2003.

[22] P. Juang, H. Oki, Y. Wang, M. Martonosi, D. Rubenstein and L. Peh, "**Energy-Efficient Computing for Wildlife Tracking: Design Tradeoffs and early Experiences with Zebranet**", in *ASPLOS-X*, October 2002.

[23] B. Chen, K. Jamieson, H. Balakrishnan and R. Morris, "**Span: An Energy-Efficient Coordination Algorithm for Topology Maintenance in Ad Hoc Wireless Networks**", in *Mobicom* , pp. 85–96, Rome, Italy, July 2001.

[24] B. Chen, K. Jamieson, H. Balakrishnan and R. Morris, "**Opportunistic Data Collection for Disconnected Wireless Sensor Networks by Mobile Mules**", in *Ad Hoc Networks Journal*, National Chiao Tung University, Taiwan, 11 Januar 2013.