JIMMA UNIVERSITY

INSTITUTE OF TECHNOLOGY

SCHOOL OF COMPUTING

Data Forwarding Using Interest Relationship in Delay Tolerant Social Networks

Andualem Chekol Zerihune

A THESIS SUBMITTED TO THE SCHOOL OF COMPUTING OF JIMMA UNIVERSITY IN PARTIAL FULFILMENT FOR THE DEGREE OF MASTERS OF SCIENCE IN COMPUTER NETWORKING

March, 2017

JIMMA UNIVERSITY

INSTITUTE OF TECHNOLOGY

SCHOOL OF COMPUTING

Data Forwarding Using Interest Relationship in Delay Tolerant Social Networks

Andualem Chekol Zerihune

Advisor: Ahmedin Mohammed Ahmed (PhD)

Co-Advisor: Kebebew Ababu Yitayih (MSC)

APPROVED BY EXAMINING BOARD:

Name	Signature
1. **Dr.Ahmedin Mohammed Ahmed , Advisor**	____________________
2. **Mr. Kebebew Ababu Yitayih , Co-Advisor:**	____________________
3. **Dr. Dereje Yohannes, External Examiner:**	____________________
4. **Mr.Fissha Bayu, Internal Examiner:**	____________________

Dedicated to:

Amsale Woldie and Chekol Zerihune

Acknowledgments

It is with great honor to thank my advisor Dr. Ahmedin Mohammed Ahmed for his valuable inspiration, insightful comments and unreserved guidance. I would also like to thank my co-advisor Mr. Kebebew Ababu for providing insightful comments and feedbacks. The door to his office was always open whenever I ran into a trouble spot or had a question about my research or writing.

Besides, I would like to thank the following people who helped me in the making of this work with their generous support; Geletaw Behayilu, Embet Kassa, Zerihun Olana, Amanuel Mihret, Abrham kassa, Samuel Maru, Teshager Tilahune, Muhammed Adane and I am indebted to my many of my university colleagues to support me during my MSc thesis. I must express my very profound gratitude to my parents and to all my family members for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Finally, to all the scholars who are keen to share and put their effort to make the world open for science and technology, particularly, an online discussion forums, The ONE mailinglist.

Abstract

Access to information and digital services on the go has grown with the proliferation of mobile devices and various types of wireless networks. This trend brings an opportunity for individual to access digital contents, share resources and communicate any where any time facilities. However, the connectivity coverage rate did not go hand in hand to the trends of mobile phone penetration rate, and there is connectivity coverage disparity among developed countries and developing ones. This motivates that infrastructureless networks such as Delay Tolerant Networks(DTNs)to play an alternative communication approach and to be leveraged in different application scenarios.

The characteristics of this kind of networks (i.e., intermittent connectivity and transient end to end path) create a challenging problem in routing data. For example, to deliver a given message towards destination node/groups of nodes, the forwarding scheme uses a store-carry-forward and utilizes different heuristics to select a potential forwarder node/intermediate node. This thesis focuses on those routing schemes that consider social properties of individual nodes, and aims to design forwarding algorithms as a means to improve network performance. Existing works overlooked interest relationships among users to enhance forwarding performance. Therefore, the main objective of the study is: designing social-aware based opportunistic forwarding solutions to attain a better forwarding performance. Therefore, Interest Relationship Based Forwarding Scheme in Delay Tolerant Social Networks (DTSNs) named as IntReF is proposed. It is motivated by the fact that combining two or more social similarity in terms of user interest, and its relationships in accordance with human interest dynamics showed fat-tailed probability distributions in human daily activities.

To evaluate the performance, an extensive simulations experiment is conducted. IntReF is compared against two selected schemes Bubblerap and Epidemic. The performance results shows IntReF achieves a better forwarding performance in terms of delivery ratio, average latency, as result compared to bubblerap and Epidemic. The simulation results demonstrate that IntReF outperforms bubblerap and Epidemic with higher message delivery ratio and maintains a reduced average latency. Moreover, IntReF can outperforms around 24.66% delivery ratio in average in terms of varied value of buffer size, message time to live (TTL) and simulation duration parameters. Simulation evaluation results towards varied value of buffer size, message time to live (TTL) and simulation duration support the effectiveness of the scheme.

Keywords: Delay Tolerant, Forwarding, Interest Relationship, Opportunistic, Social Properties

Contents

List of Figures

List of Tables

Acronyms

CPU Central Processing System

DTSN Delay Tolerance Social Networks

DTN Delay Tolerance Networks

GPS Global Position System

GSM Global System for Mobile Communications

3G third-generation

4G fourth-generation

GUI Graphical User Interface

HBA Human Behaviour Aggregation Algorithm

ICN Information Centric Network

ICT Information and Communication Technologies

TCP/IP Transmission Control protocol Internet Protocol

MANET Mobile Ad hoc Networks

MB Mega Byte

ONE Opportunistic Network Environment simulator

SF Scale-free Property

SMS Short Message Service

SNA Social Network Analysis

SW Small-world Property

TTL Time To leave

Wi-Fi Wireless Fidelity

Chapter 1

Introduction

1.1 Motivation

Mobile devices (e.g., mobile phone, Smart phone, and personal digital assistance) penetration becomes a daily news event in the world. The number of individual who access digital resources via mobile devices increases exponentially. For instance in the world, the number of individual in mobile cellular subscription has grown significantly [1]. This trend brings an opportunity for individual to access digital contents, share resources and communicate anywhere anytime to ensure accessibility of facilities.

However, the connectivity coverage rate did not go hand in hand to the trends of mobile phone penetration rate, and affordability is also questioned in developing country[2]. Moreover, Connectivity is concentrated in developed countries, and sparse in developing countries, where 78% of the population is online compared to 32%, respectively. Developing countries are home to 94% of the global off line population [1]. In addition, in developing countries perspective, addressing this huge gap with only infrastructure based network is not cost effective in the area where the density of the population is sparse. In addition, some digital contents and services are limited to a specific proximity of a given region, which are only required in that locality without highly dependent on the infrastructure based network.

In general, "Digital Divide" has been defined as the gap between those that have Internet access and those that do not. However, the ability to meaningfully participate in and benefit from the Internet is also dependent on the ability to use and have confidence in the technology, and it is mainly affected by various factors and barriers in including relevance and readiness barrier that hinder adoption of Inclusive Internet.

Therefore, an infrastructureless wireless networks, particularly opportunistic networks [3] are gaining huge attention from the research community to play an alternative communication approach and to be leveraged in different application scenarios. In this kind of networks, nodes (e.g., people who carry the device) come into contact with each other opportunistically and communicate wirelessly using Wi-Fi and Bluetooth enabled mobile devices. Moreover, the underlying principle in this network is based on the human to human interaction via the mobile devices by exploiting the relationships and social features of the human, to design and build effective and efficient communication protocols (e.g., routing and forwarding algorithms).

Delay tolerant Social Networks (DTSN) is a kind of opportunistic networks based on communication mechanisms where communicating nodes in the networks has a transient end to end connectivity, because of the reason that social-aware intermittent connectivity occurred due to frequent mobility of the nodes. Communication between and/or among nodes in this kind of network is achieved using a store-carry-forward paradigm in pair-wise fashion. This is a paradigm where nodes store the forwarded messages in to the persistence storage space and carry and exchange up on encounter and/or contacts. Relying on this principle, an effective and efficient data routing and forwarding strategies is required. During data forwarding from source to the destination of node or group of nodes, different algorithms are devised for the relay node selection, when nodes encounter [4, 5, 6, 7].

Although, DTSN is being used in the aforementioned scenarios it has paramount challenges and problems that has to be addressed properly towards data forwarding operations. Due to the reason that it has a major role for the realization of this network in various application scenarios.And various solutions aiming to solve routing problems in specific scenarios to improve various network performance objectives. This research work is concerned to investigate one main intriguing question among the research community; how to model the relationship between interests of the nodes and its impact on data forwarding scheme performance especially on the delivery ration, delay and overhead. Therefore, the main objective of this study relies on proposing a data forwarding scheme that considers social features (i.e., interest and interest relationships between users) to improve forwarding performance in the network.

In this case, this work is motivated by the fact that combining two or more social similarity in terms of user interest and its relationships results in a significant performance improvement. For instance, according to Zhou *etal.* [8], human interest dynamics showed fat-tailed probability distributions, which indicates the dynamics among individuals interests, conveys that individuals tend to return to highly ranked interest with relatively larger probability and stay in these interests in human daily activities. Additionally, occasional contacts, intermittent connectivity, involvement of power-constrained devices and possible nonexistence of end to end paths are common characteristics of this kind of network. A single or combined effects of the aforementioned characteristics creates a challenging problem in routing and forwarding data.

1.2 Statement of the Problem

In DTSNs, where social properties of the mobile nodes exploited to design effective and efficient data forwarding schemes, for instance, in this paradigm people with similar interest meet with one another in a higher probability than people with dissimilar interests, like people (e.g., classmates or colleagues) with the same interests usually get together to talk about and share their interests [5]. Therefore, relying on mobile node (i.e., people who carry the mobile device) interest property to design and improve the performance of data forwarding schemes in DTSNs have a paramount benefits.

However, most of the forwarding algorithms which rely on Interest-based approaches did not examine the apparent relationship between mobile user Interests and its impact on the efficiency of the performance of data forwarding schemes in this paradigm and the interrelationship between user interests and its impact on the efficiency of data dissemination has not been explored sufficiently [9].

Added to that, and a more recent work Int-tree [10], examined the relationship between interests of mobile node users and its impact on the performance of the data forwarding schemes. In Int-tree, relay selection and forwarding decision are critical to be made by current node based on certain routing strategies. However, they focused only on the inclusion relationship type, and community based forwarding strategies using community density and social tie metrics. More specifically, for intra community (inside the community) forwarding strategies, a given node measures social tie between the encounter node and a node which has higher value is selected for a better relay.

However, dead end and blind spot problems happen in the forwarding decision process when encounters nodes compares the utility value. For instance, in a situations where the utility value (the measured social tie value) of relay node and the neighboring node less than to the source node (sender in this case), or else the relay node is invisible to the source node (dead end and blind spot problems respectively) and the message stuck on the node itself and delivery ratio and average delay affected most [11]. Therefore, we are concerned to address the mentioned problems. In this thesis, we investigate and examine the following research questions. How to model the relationship between interests of the nodes? How to figure out its impact on forwarding scheme based on performance objectives such as delivery ratio, overhead ratio and average latency?

1.3 Objectives

General Objective

- The general objective of the study is to propose social-aware interest relationship based forwarding scheme in Delay Tolerant Social Networks.

Specific Objectives

- Identify the different types of opportunistic routing approaches.
- Understand and examine the existing opportunistic routing taxonomies.
- Exploit knowledge about users interests to understand its impact on forwarding performance.
- Examine and model relationship between mobile node user interests.
- Examine and Employ social-aware utility functions.

- Implement a social-aware and content-based forwarding approaches under DTSNs simulation environment.

- Test and evaluate the performance of the proposed scheme through simulations to show that it provides statistically significant forwarding performance improvement.

1.4 Methods

A. Literature Survey

A literature survey is a key component in the process of conducting a research. Its significance to the body of a research is crucial. Due to the reason that it provides an up-to-date understanding of the topic and its significance; it helps to identify the methods used in previous research on the topic; facilitate to work out how to answer the questions and indeed, what questions need to be asked with identifying topics for further investigations for the gaps identified. And it helps to relate and compare the achievement of the proposed research to prior research towards a given objectives. Moreover, it supports the originality and the relevance of the conducted research from prior.

Taking into consider the aforementioned importance of literature survey, to achieve some of the objectives of this thesis, extensive literature survey is conducted using different academic resources (e.g., scholarly books, research papers, technical reports and others) to have a comprehensive understanding and examining towards forwarding with replication schemes in DTSNs, and particularly forwarding with replication using Social properties and relations. Besides in order to develop and implement, the forwarding scheme, graph-based representation of the network, to form a social graph, and relevant social network analysis techniques and algorithms to obtain a strong utility functions that streamline forwarding decisions (i.e., Social properties and relations between nodes in the network) is also examined.

B. Design and Implementation

Simulation based approaches is a cost-effective method to design, implement, analyze and evaluate network protocols and is an important tool in computer networking research. Practical approaches (with testbed and real production environment) often have to fight against the complexity of the scenarios and high costs of using hardware. Simulations can efficiently investigate on the parameter space with respect to mobility model, message load, communication patterns and dependencies among protocols in a scalable way. Furthermore, documentation about the outputs and performance reports of the protocols under investigated ease and produced in an improved quality. Due to the reason that it provides a higher chance to control and filter irrelevant results such as noise.

The proposed model and algorithms are designed based on the stated objectives, and its implementation conducted with simulation based approaches that are suitable for forwarding and routing messages in DTSNs. Due to the reason that this networks has some unique features. In this thesis, the design and implementation of the proposed scheme limited and rely on simulation based approaches because, in one hand using simulation based approaches has benefits, on the other hand the characteristics of DTSNs demand a scalable approach that helps to streamline realistic mobility model, ease of using different network application scenarios and reduced model development and analysis time.

C. Evaluation of the Proposed Work

Simulation performance results and outputs are only of relevance and implied on if a precise modeling and evaluation has been done. Without it there is no any other means to confirm its validity whither it reflects the underlying scenario, also, how the proposed scheme represents the system under investigated? Due to this verification and validation process is critical. Therefore, either comparing to the existing system or following deterministic test case to ensure the validity and verification of the proposed scheme is key activities that should be conducted.

Therefore, using DTSNs simulation based experiment environment; the performance of the proposed forwarding scheme is evaluated and compared with the existing benchmark routing and forwarding algorithms based on different performance evaluation metrics (i.e., delivery ratio, average delay and average overhead ratio), by taking in to account different simulation parameters such as message TTL, node buffer size, simulation duration and others are calibrated and considered.

1.5 Scope

The scope of this thesis is limited to designing and implementing forwarding scheme with replication using Social properties and relations. Data forwarding scheme using interest relationship in delay tolerance social networks. The forwarding scheme, graph-based representation of the network, to form a social graph, and relevant social network analysis techniques and algorithms to obtain a strong utility functions that streamline forwarding decisions (i.e., Social properties and relations between nodes in the network) are designed. Therefore, designing social-aware combined with content features opportunistic forwarding scheme, in order to attain a better forwarding performance objectives(e.g.,delivery ratio,overhead ratio and average latency).

However, Security and privacy issues require and need in a separate study. And all nodes are assumed to be cooperative and willing to involve in the forwarding process although we will take in to account a limited node a buffer size in the performance evaluations. And various malicious properties of node that may occur in the network is not take in to account.

1.6 Application of Results

We build forwarding algorithms that facilitate and forward message exchanges in opportunistic networks e.g., delay tolerance social networks where infrastructure free network can be utilized as alternative communication mechanisms. Village communication, enabling connectivity alternative way to overcome digital divide in the sparsely populated regions in developing world. Sharing digital data and resources with geographically limited proximity (e.g., traffic information, local news feeds and message communication in neighborhood level).

1.7 Organization of the Thesis

The remainder of the thesis is organized as follows. We discuss Literature review on background and recent developments about Social Properties, Mobility, Social-Aware and Networks Delay Tolerance Routing and Applications in Chapter 2. We discuss basic forwarding process and common categories with routing schemes into different kinds of groups according to aspect of forwarding approaches. And discussion of applications of Delay tolerance social networks in to Developing region application scenarios and others services and applications. In Chapter 3, We summarize Related Works of forwarding schemes and their taxonomies. Particularly, replication based approaches can be further subdivided in to social similarity and social oblivious based on the the type of heuristics which they utilize for replication. In Chapter 4, we discuss towards the proposed scheme IntReF with its components. Also discuss Preliminaries. In chapter 5, discussion of demonstration of effectiveness of the proposed scheme proceeds with its analysis of results. And performance result of the proposed scheme with comparisons of other benchmark forwarding schemes conducted and examined. In Chapter 6, we conclude and discuss future works of the thesis.

Chapter 2

State of the Art: Background and Recent Developments

Wireless network has become one of the most pervasive core technology enablers for a variety of computing and communications applications and services. It enables users to communicate and access applications and information without rely on connecting wires directly. It is the connections between users and information sources without the use of wired communication links directly. In general as depicted in fig.2.1, wireless network work in two modes, infrastructure based and infrastructure-less mode. In the case of infrastructure based, it rely on a base station that coordinate the transmissions among nodes(e.g., Wireless Access Point and Radio tower),and ranging from third-generation/fourth-generation (3G/4G) cellular mobile network , broadband access and indoor/outdoor WiFi networks are some of the examples of this type of networks. Since the focus of this thesis is not towards infrastructure based wireless networks we limit the discussion only infrastructure less wireless network type.

In the case of infrastructure-less, the network has formed in ad hoc way that nodes (e.g., wireless hosts and mobile nodes) come in to contact with each other and communicate wirelessly using communication link (e.g., Wi-Fi and Bluetooth) with out rely on some kind of base stations that coordinate the transmissions among the nodes. And the communication coverage area is only depends on the number of nodes in the network and the type of wireless communication link.

This Multi-hop, infrastructure-less wireless networks fall into several categories, depending on the characteristics of the network and the underlying assumption that the communication protocols rely on. Those are mobile Ad hoc networks (MANETs), Social Aware Network(SAN) and Delay Tolerance Networks (DTNs). Since there is no base station in these networks, and nodes may have to relay messages among several other nodes in order to reach a destination. And nodes may also be mobile, with connectivity changing among nodes. Therefore, the nodes themselves must provide for services such as routing and forwarding, address assignment, and other functionality. And because of this the development of communication protocols for such networks is challenging and is the subject of much ongoing research [12, 13].

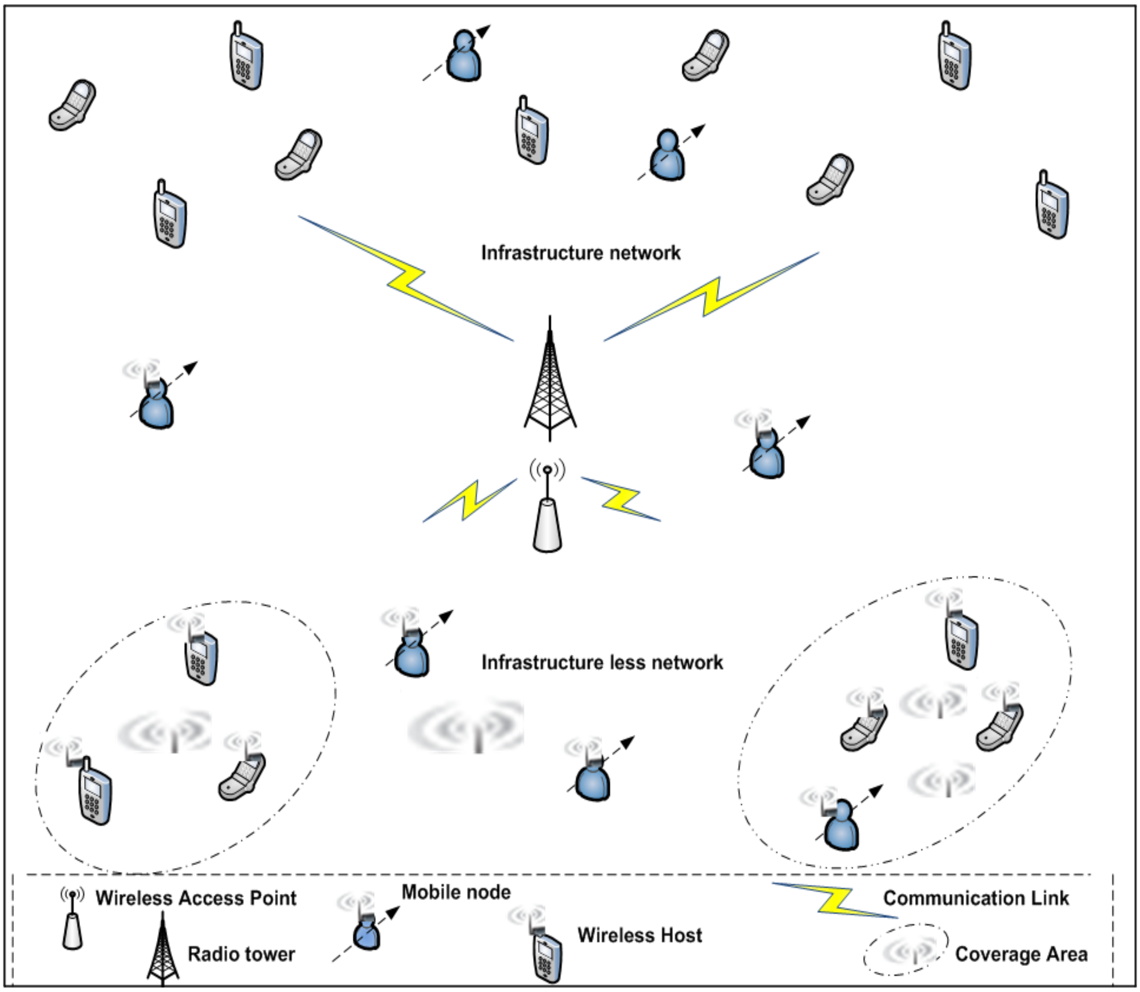


Figure 2.1: Elements of Wireless Networks[13]

Although mobility of nodes apparent in MANETs and DTNs, for the former, there is at least one end-to-end communication path between any source and destination node pairs at a certain point of time, whereas for the later, the availability of end-to-end communication path between any source and destination node pairs at a certain point of time is transient. This creates difficult problems for communication protocols that fit for MANETs do not able to function and provide all the functionality in DTN scenario. Therefore, taking in to account the static network information, for instance, social information of the nodes in the network has gained much attention to design routing and forwarding algorithms in DTSNs. Since this kind of information tends to stable over time, and do not highly coupled with the underling topology in the network, routing and forwarding algorithms relying on this information do not highly influenced by the dynamic nature of DTSN (i.e., mobility), and has a better performance result compared to dynamic network information.

2.1 Social Properties

In this section, we describe social network analysis approach to model networks and various properties and metrics that are being used to characterize and quantify the underlying network.

Social network analysis(SNA) approach inspired from graph theory is being used to analyze various forms of networks in different field of study (e.g., sociology, biology, computer science, and many other fields). A network is a collection of objects or individual entities in which some pairs of these entities are connected by links. And it consists of a pattern of interconnections among a set of things. This interconnection is represented by a graph and it is a robust approach. Depending on the setting and available information, different forms of relationships or connections can be used to define links. It is the study of relationships between entities and on the patterns and implications of these relationships based on the graph which represent the underlying network.

A graph is defined as a set of individual entities, termed as nodes, with certain pairs of these entities, connected by links called edges. For example, the graph in Figure 2.2, has 4 nodes labeled A, B, C, and D, with D connected to each of the other three nodes by edges, and C and A connected by an edge as well.

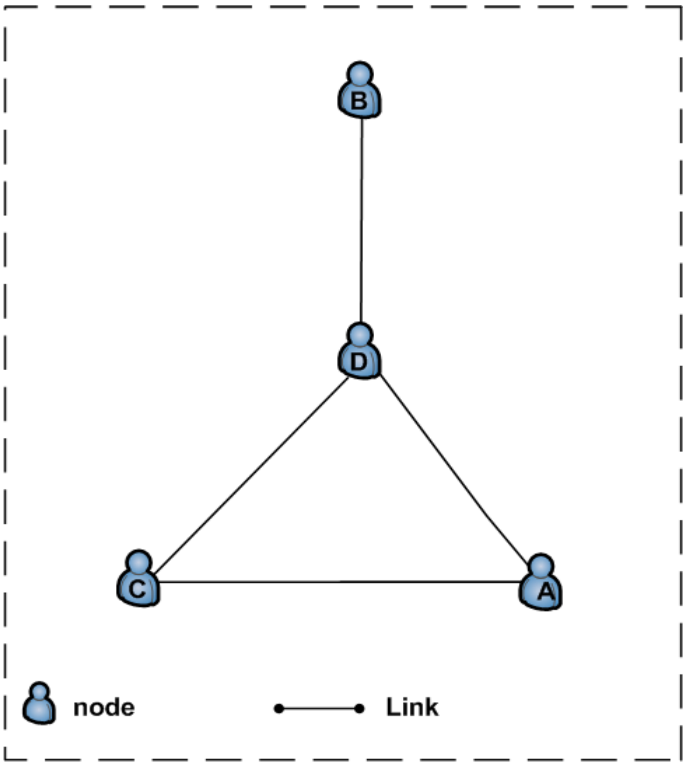


Figure 2.2: Example of graph with four nodes

In the case of DTSNs, the interaction between nodes in the network represented by individual nodes to its respective edge between nodes, and the edge represent the link between them. This notion different forms of relationships or connections has different level of strength of link, social strength or tie can be quantified based on different inputs (contact frequency, contact duration, and other social properties of nodes) to better understand the relationships between encounter nodes. And using SNA, different structural properties and metrics of the network are measured, in order to understand the characteristics of the network being analyzed to streamline various levels of decisions based on the specified objectives. For instance, in opportunistic network, exchange of messages among nodes, sending a message from source to destination or destinations requires, identification and selection of potential forwarder node towards the intended destination should be decided based those properties and metrics.

2.1.1 Potential Social Metrics

- **Centrality:**

It is a social metrics that measure the relative importance of a given node within the graph. A node in the network with a central position has a critical structural importance for connecting other members of nodes in the network with a stronger capability than the other nodes. A Centrality of a node can be various forms and measured in different ways, thus the most commonly used potential metrics, centrality with its variants are summarized as shown in table 2.1.

- **Similarity:**

Another category of potential social metric is similarity. This is a measurement that helps to signal about the commonness of a node or groups of nodes towards the intended destination or destinations. It can be measured with various features of nodes that found in the network, such as interest information, location and many more. For instance, in a social network, it is observed that individual often befriend other who has similar interests. And often this occurrence also called homophily [14], it is an event, where, for example, there are four nodes in the network (A, B, C, D) and C is a common friend to A and B, compared to D, the likelihood of being a friend between A and B is higher than D being a friend to either A or B.

In opportunistic network perspective, this social metric can be used as strong signal to send a message towards destination, by the assumption that similar nodes in the past are likely to interact in the future which is some kind of prediction heuristics to streamline message transmission among nodes.

2.1.2 Structural Social Properties

- **Small-world Property**

In most setting of a graph, a length of path is the number of hops something travels though the network. And this path can be a few hops or many. In graphs formed from different human traces most of the nodes pairs can reach each other via short paths. This property of the graph, named as Small-world Property(SW), has been explored and used as a strong signal for the existence of short routes/path between most of the nodes pairs in the network [17].

- **Scale-Free Property**

Keeping Small-world Property holds true in some kind of networks (e.g., delay tolerance social networks), and representing through a graph which consists of number of nodes connected each

Table 2.1: Summaries of potential social metrics

Metrics	Descriptions
Degree centrality [15]	The number of tie (link) a node has. it can be consider also counting the nodes neighbors or peers. To how many neighbors a node can directly deliver a message to.
Betweenness centrality [15]	How many pairs of nodes would have to pass through the node in order to reach one another in the minimum number of hops? This is measure of betweenness centrality metric conveys. The minimum number of shortest paths that pass through the node.
Closeness centrality [15]	It is the measure of the shortest path between a node and all other reachable nodes. It can be considered as a measure of how long it will take information to spread from a given node to other nodes in the network.
Similarity [16]	It is a measurement that helps to signal about the commonness of a node or groups of nodes based on some features or characteristic between nodes. And streamline message transmission towards destination.

other with links, does the number of connection each node has a certain kind of distribution? Do some nodes have popularity in terms of number of connection towards other nodes? Popularity is a property of a node in the given graph which characterizes its importance and significance in respect to other nodes to be used as hub, a node which has more connection than others.

To understand and describe this notion of popularity of a given nodes in terms of its degree in a graph, and its existence in the network, we describe some of the commonly used probability distribution functions (probability density functions) and later explain scale-free networks properties. One can ask the following very basic question that motivate and leads to the concept of probability distributions as follows: *As a function of k degree, what number of nodes on the network(graph) has at least k degree?* The larger value of k indicates greater popularity.

Normal distributions: The basic fact about normal distributions is quantified by two quantities, the mean value and the standard deviations around this mean. In other words, any quantity that can be viewed as the sum of many small independent random effects will be better approximated by the normal distributions. As shown in figure 2.3, values in the normal distribution, the probability of observing a value that exceeds the mean by more than a given constant values times the standard deviation decreases exponentially in a given constant values.

Although the normal distributions is widely used, which can be quantified around the mean, there are some kind of networks with some properties that normal distributions do not fit and describe well, e.g., distribution of number of friends in a social network, the distribution of number of node that has at least k degree in a graph formed from different opportunistic human traces.

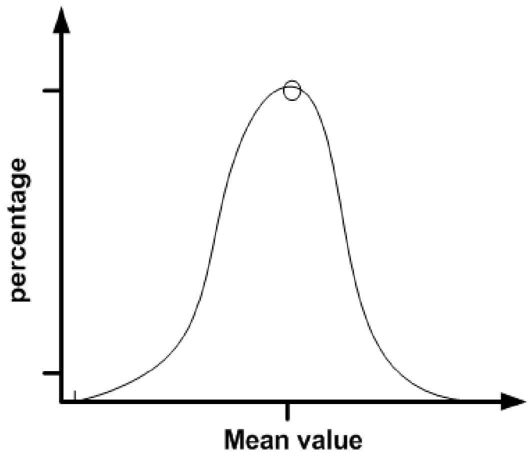


Figure 2.3: Values in normal distribution

Power law: Another kind of probability distribution that describe some properties of network, for instance, the distribution of number of node that has at least k degree in a graph. It is observed that the distribution of number of node that has at least k degree in a graph formed from different opportunistic human traces has shown that there are quite few nodes have a higher popularity (nodes having more degree k), whereas, majority of the nodes have low popularity, which is approximated by power law distribution as depicted in figure 2.4. It is a function that decreases as k to some fixed power α, as shown in equation in eq.2.1 [18].

$$P(k) = \frac{c}{k^{\alpha}} \tag{2.1}$$

where c is some constant, α is typically between 2 and 3, due to the reason that a power law only normalize $\alpha > 1$

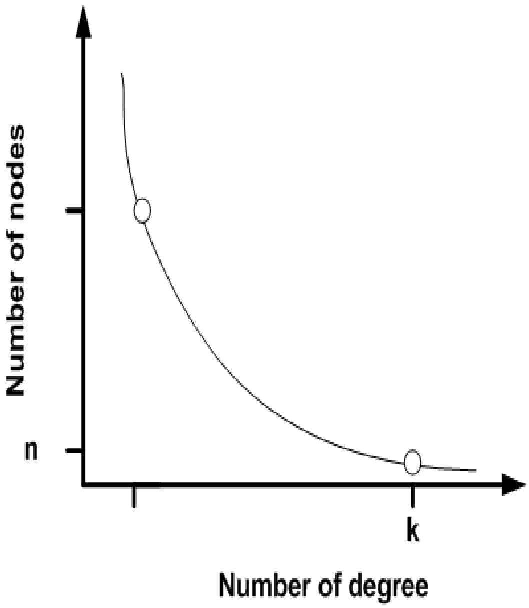


Figure 2.4: Power law distribution of popularity

This property of a given network, where, a significant number of nodes (i.e., hubs) with a degree higher than the average degree found in the network and its occurrence quite probably which is approximated by power law distribution is called scale free properties, and often network which has this kind of properties also know as scale free networks.

2.2 Mobility

Mobility is the actual movement trajectory of nodes and their locations. In wireless network which has mobile node to form the network and communicate with each other while on the go, the underlying network protocols (e.g., routing algorithms) is affected by mobility behaviors of the nodes. In order to understand its effect and observe the behaviors (i.e., performance evaluation and model validation), there are two approaches; either deploys the actual network in real life or follow simulation based approach. Due to economic feasibility (e.g., expensive cost of devices), and scalability issues related flexible features to fine tune a range of application specific setting and parameters, simulation based techniques are a suitable approach to deploy the network and streamline performance evaluation and model validation rather than deploy the actual network in real life. Moreover, Performance evaluation and model validation through simulations have been an essential part of research in the domain of computer networks[19][20].

In simulation-based approach to abstract and simulate actual movement trajectory of nodes, there are two mechanisms to dictate how the nodes move during the simulation. These are real human connection traces also know as real world trace and synthetic mobility model.

2.2.1 Real Human Connection Trace

Real mobility traces refer to the realistic human mobility and encounter information, which have been collected in experiments to represent the actual movement trajectory of nodes and their locations by providing mobile devices to the human users. From daily movement patterns of a given selected groups of individual for example, traveling to the office, visiting a mall, and returning back to home to actual movement trajectory of cars in a city has been collected. One can find various real world trace dataset available on the web site of CRAWDAD Project(from http://crawdad.cs.dartmouth.edu). For a survey about the trace-based mobility models, readers can be referred to [19].

Since the real-world traces capture real human behavior by keeping track of the movement pattern of a human being, using it in a given simulation to simulate in various realistic application scenarios is very essential to represent the actual conditions that exhibited in the real network deployment. However, due to its limited nature of scalability, unable to change specific setting and parameters, for example, one cannot increase the number of nodes that participate in the trace. Moreover, it may not be possible to use the trace in all kinds of scenarios, for instance, how one can use the existed trace which consist of selected groups of people (e.g., people participating in a conference or working in a research laboratory) to simulate a post-disaster scenario. Since scenarios of this kind requires different behaviors than the behaviors of normal conditions. With this using synthetic mobility model is also crucial.

2.2.2 Synthetic Mobility Model

It is mimicking the actual movement patterns of nodes in a given scenario through mathematical model by utilization various physical law and principle without taking in to account the actual movement trajectory of nodes and their locations. In a given simulation the synthetic mobility model itself dictates how the nodes should move. This kind of mobility can be applied and used in various field of study for different applications scenario, due to the reason that its scope is only limited by to what extent its mathematical model is accurate to represent the scenario in question.

In regards to performance evaluation and model validation through simulations in computer network, in order to simulate different scenarios with different movement of node in a given simulations, using synthetic mobility model is vital. Without it how one can simulate in a given wireless network consists of cars, pedestrians, and bus with each has different mobility behaviors in a given city, for example, to evaluate performance of message exchange using some forwarding scheme. There are several synthetic mobility models have been proposed. Some of them are Random Waypoint mobility(RWP)model, Reference Point Group mobility model(RPGM)and Map-based movement models(MBM)[19].

A. Random Waypoint Mobility Model

In this kind of synthetic mobility model, with in the terrain size of simulation area nodes are provided with a randomly selected coordinates. Nodes use the coordinates and move to the destination at constant speed, with some stops, and continue after it gets new random coordinates. Nodes keep continue this process in the whole simulations. Nodes follow RWP model has zig-zag trajectory path. However, because of its high randomness, this mobility model cannot be use to reflect human mobility patterns[21].

B. Reference Point Group Mobility Model

This is a group based mobility model where the movement of a node depends on the mobility of other nodes. In Reference Point Group Mobility Model (RPGM), at first all the nodes divided into different groups, and each with a group leader. And each leader in a group guides the movement of mobile node within the group, also determines the mobility attributes (e.g., the direction of path and speed). Individual mobile node follows a randomly movement based on the pre-defined reference points which is determined by the whole movement of the groups. Relying on Random Waypoint Mobility Model, the features of RPGM implemented. Compared to RWP, RPGM mobility model represents in better way to reflect reality[22].

C. Map-based movement Models

Considering a scenario given in a wireless network consists of cars, pedestrians, and bus with each has different mobility behaviors in a given city, for example, to evaluate performance of message exchange using some forwarding scheme. Simulating more realistic scenario like this can be possible with help of Map-based movement model.

In this Mobility model either the simple MBM or a more advanced Shortest Path Map-Based movement Model(SPMBM), the movement of nodes constrained by predetermined path on the map which is used to simulate the scenario. Different node groups can be configured to use only certain parts of the maps which can prevent, e.g., cars driving on pedestrian paths or inside buildings. And also to guide for example, bus can only follow specified schedule roads which we found in various cities in real life[20].

In the case of simple Map-based model, nodes are placed randomly on the map area and follow the path a segment forward until it reach the end of the segment, as stops point or hit an intersection point where different path segments branched. Starting from the intersection, nodes again can select randomly another path segment except the one that it came from. With these nodes continues its journey.

A more advanced map-based model, Shortest Path Map-Based movement Model places the nodes on the map in the same way as simple map based movement model do, except that fact that nodes selection of certain destination in the map for all the nodes uses Dijkstras shortest path algorithm to find the shortest path to the destination.

2.3 Social-Aware Networks

2.3.1 Emerging Paradigms in Social-Aware Network

With the proliferation of mobile devices, particularly smart mobile devices and various types of wireless networks, access to information and digital services on the go has grown. And this brings a convergence trend among all the entities participated in the network.

Due to the convergence of interaction and relationship among users, mobile devices and electronic system in the daily routines of human life, understanding and exploring about this interaction and relationship between entities gains greater attention from different field of study. This structure of entities with its interaction and relationship among various entities form a network called Social Aware Network. Such kind of network defines the behavior of entities and helps to understand different relation among them.

Social Aware Network(SAN) is formed from nodes (e.g., people who carry the device) come in to contact with each other opportunistically and communicate wirelessly using Wi-Fi and Bluetooth enabled mobile device. Moreover, the underlying principle in this network is based on the human to human interaction via the mobile devices by exploiting the relationships and social features of the human, to design and build effective and efficient communication protocols (e.g.,routing and forwarding algorithms). It overall structure and convergence depicted in fig.2.5

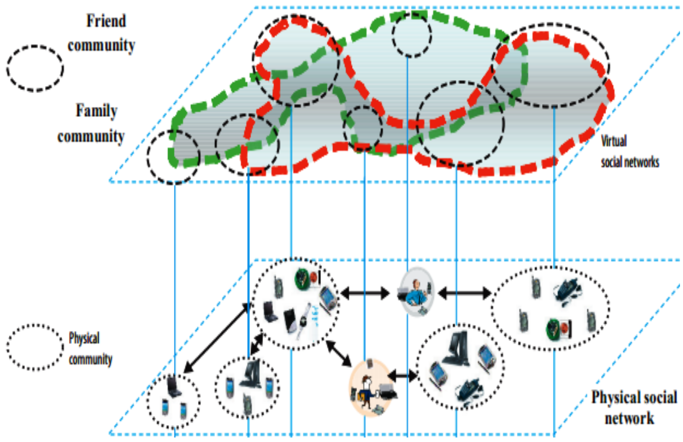


Figure 2.5: Structure and convergence of Social Aware Network[3]

In this figure, the physical social network comprises of various physical community which can be represented by groups of nodes(e.g., people who carry the device, and vehicle) come in to contact with each other opportunistically and communicate wirelessly using Wi-Fi and Bluetooth enabled mobile device for different use cases. But the virtual social network (e.g., friend and family community) are formed based on the social relationships between mobile carriers. One can observe that the convergence of two different network levels (The physical social network and the virtual Social Network). One is the projection of the other. There is a fusion and essential relationship exists between them.

The virtual social network overlay on the physical social network. This is a notion that the virtual social network actual does not exist in the network, but it is a major factor to guide the various operation in such a network than the physical social network. It can be used to design different networking solutions by exploit the social properties of network nodes (particularly mobile devices). It has paramount advantages than relying on properties of physical social network (mobility behaviors of nodes). Due to the reason that this social properties and relation do not change over time, and this make virtual social network comparatively less volatile than physical social network. Moreover, it does not highly coupled with the underling topology in the network, and it is a strong signal to design various network solutions[3].

2.3.2 Overview of Social Aware Network

In social aware networks similar with other opportunistic network method (e.g., DTN) follow a storecarryforward paradigm in pairwise fashion to provide communications between mobile devices "with high in-dependency" on the infrastructure network, and relying on short range communication technologies such as Wi-Fi and Bluetooth. And the social properties into consideration to design different network solutions. As shown in Fig.2.6 the work flow and its major activities is depicted[23].

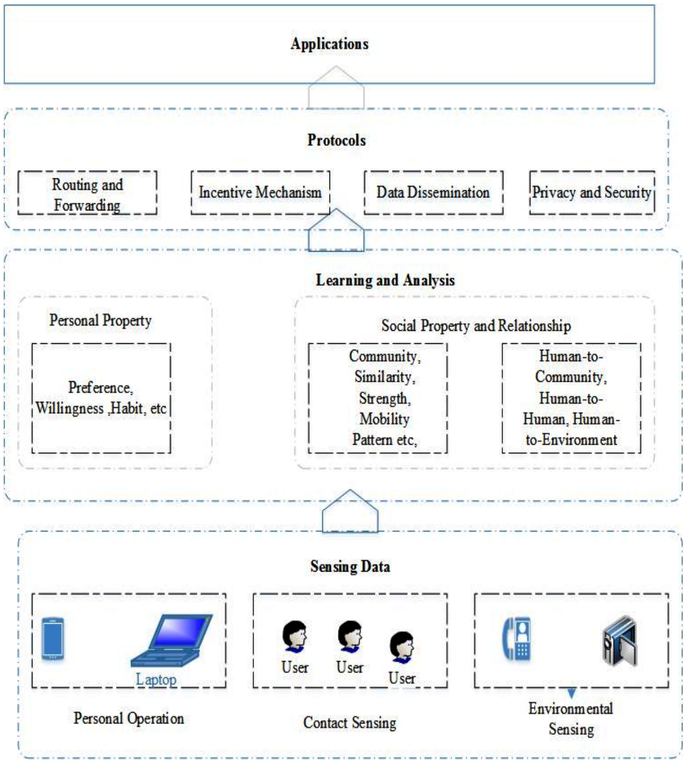


Figure 2.6: Overview of Social Aware Network Architecture[23]

In such a network, Sensing data, learning and analysis, protocols and applications are the major tasks that carried out to achieve different objectives. For instance, one can follow these steps to design a data forwarding scheme that streamline message exchange in a given network application scenarios.

At the sensing phase, social properties (e.g., contact duration and interest information) of individual nodes is collected and stored, and following that, the learning and analysis stage, use different methods, e.g., social network analysis technique to obtains and formulate essential metrics and relationships among the nodes in the network. Later the protocols, forwarding scheme in this case, uses to this potential social metrics and relations to proceeds message exchange among the nodes in the application scenario being used. In this thesis, inspired by these features of Social Aware Network and combined with characteristics of Delay Tolerance Network, the proposed data forwarding scheme is relying on.

2.4 Delay Tolerant Social Networks

2.4.1 Routing Algorithms

Delay Tolerance social networks are an integration of features of Social aware network(social properties and relations) with delay tolerance Network characteristics. It is a networks of self organizing wireless nodes connected by multiple time-varying links, and where end-to-end connectivity is intermittent. In DTN mentioned characteristics associated with the problems are addressed by using a store-carry and forward paradigm [24, 25, 26]. This is a technique, where the message from the source node moved to the next node which is used as a hop to the destination node. When the message traverses along the path, there is a persistent storage place in each node which is used to hold the message in a definite time before it is forwarded to the next hop. The main concern of data routing strategies is to decide whether to forward the data to the counterpart when two nodes encounter. Different schemes are devised for the relay selection.

Due to intermittent connectivity, routing protocols based on the knowledge of end-to-end paths perform poorly, and numerous opportunistic routing algorithms have been proposed instead. In order to carefully use the available resources and reach short delays, many protocols perform forwarding decisions using locally collected knowledge about node behaviour to predict which nodes are likely to deliver a content or bring it closer to the destination as for instance, the forwarding process illustrate in figure 2.7

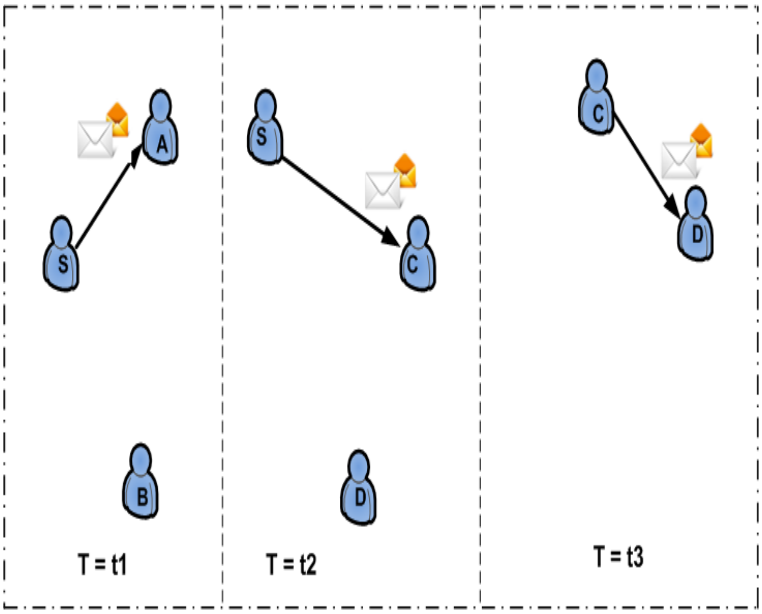


Figure 2.7: Example of opportunistic data forwarding processes. In the figure, "S" indicates the source node of the messages and "D" denotes its destination

Moreover, one of the fundamental goal of any communication network is to have the data or messages delivered to their corresponding destinations. Depending on the underlying network characteristics, the data or messages are required to be routed or relay along in the network with different routing approaches. In delay tolerance social network for instance, there are different groups of routing approaches in order to achieve distinct goals. With this, as shown in figure 2.8 the routing approaches grouped in to three major taxonomies.

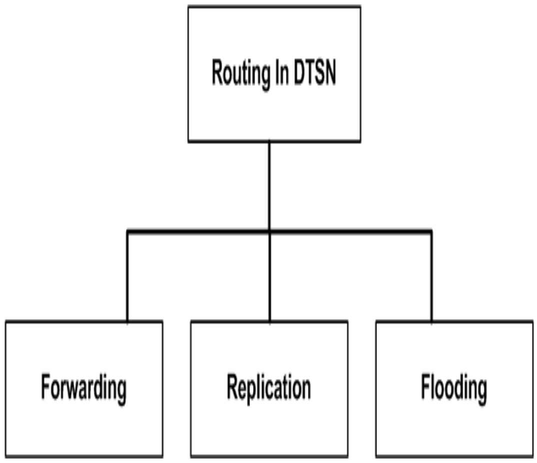


Figure 2.8: Routing Scheme taxonomies in DTSNS

A. Forwarding

In this category of forwarding approaches where only a single copy of a message exists in the network and message is forwarded towards the intended destination. Some of the proposed schemes in these categories are Direct Delivery [27] and First Contact[28]. Direct Delivery, the node that carries a given messages delivers directly to the destinations in other words the source node itself delivers its messages to the destination. Whereas in the case of First contact forwarding schemes, node forwards a message to the first node that encounter, and other node repeat the process until the message reaches its intended destination.

Although single-copy forwarding approaches aiming to optimize the utilization of network resources (i.e., storage space, overhead, and battery life), forwarding performance objectives, delivery ratio is very small and average latency become very long. Due to this performance drawback, these kinds of forwarding approaches are uncommon in DTSNs.

B. Flooding

In this forwarding and routing scheme category where a node carries a message and replicates/copies the messages to all encountered nodes, without taking in to account any kind of heuristic to select a node to replicate the message towards destination. Epidemic[29] is one of the most common proposed flooding based routing scheme which model message replication similar to the spreading of epidemic diseases in real life. Due to unlimited message coping behaviors, the scheme has both advantages and disadvantages. Forwarding performance objectives for instance, delivery ratio maximized and average latency is reduced during sending a message towards a destination, because, out of the replicated messages the chance that one of the message reaches to the destination is higher. However, a lot of network resource is consumed, for example the overall storage requirements become high in such networks.

C. Replication based

Compared to flooding approaches, replication based scheme either use by limiting the number of message copies being generated by each node through some configurable limits during message exchanging, to control replication process or nodes taking in to account some kind of heuristic to select among encounters node to replicate the message towards destination. Due to this characteristic, this approach maintains optimal forwarding performance objectives without incurring extra consumption of various network resources.

Spray-and-wait[30] is a proposed scheme that control the number of messages copies created. It sets a maximum limit on the number of possible replications of a message. And the scheme follows different techniques to deliver the message towards the destination with spray and waiting phases. In spray phase, a message is forward to at most some fixed number of different relay nodes. And if the given messages do not delivered at spray phase, waiting phase continues and some relay node deliver the message to the destination using direct delivery approach.

PRoPHET[31] a proposed scheme which tries to estimate which node has the highest chances of being able to deliver a message to the final destination based a heuristic, node past encounter history. And detail discussion on this approach with various heuristic is left for Chapter three.

In summary, both flooding and replication refer to the process where a transmitting node sends a copy of the message to the receiving node and itself keep the original message. And there is multiple copies of a message can exist in the network. Such approaches are used to enhance the chances of message delivery because of a transient end to end path apparent in the network. On the one hand, single-copy forwarding approaches aiming to optimize the utilization of network resources. On the other hand, flooding and replication-based approaches aim to optimize delivery probability. Both flooding and replication are multi-copy forwarding but the former has unlimited in coping data or message, and the later uses some kind of heuristics to which node a given message should be copied or replicated or sets maximum limits on the number of message being replicated.

The heuristics that guide (i.e., selecting a node and replicate the message to) replication of message during forwarding process mainly rely on network behavior in which each individual node exhibits in the network. The behavior of the network is described in stochastic terms, based on users mobility and social behavior, to better represent the dynamics of networks (i.e., how networks evolve overtime)[32, 33].

Moreover, single-copy forwarding has the advantage of using network resources(i.e.,storage space, overhead, and battery life) properly, but may end up taking too long to deliver the messages. Whereas flooding approaches have advantages of granting to deliver the messages fast, but may end up taking too much network resources. Replication based approaches, placed in between the two to optimize network performance goals and network resources.

2.4.2 Characteristics of DTSNs

Delay tolerance social networks (DTSN) are a kind of communication methods where there is transient end to end connectivity between the sender nodes to the receiver, destination node. In this kind of networks, intermittent connectivity, long or variable delay, asymmetric data rates and errors rate are common characteristics exhibited. Moreover, addressing the mentioned problems using the existing TCP/IP protocols stack is very difficult and challenging task. However, in DTN the above mentioned characteristics associated with the problems are addressed by using a story-and forward paradigm [34, 24, 35]. In DTSN, routing and forwarding a message among the nodes brings a challenging problem, since intermittent connectivity is common among the nodes. The problem of selecting which contacts to carry messages and when represents an instance of the DTSN routing problem. As depicted in fig.2.9 in DTSNs, the main characteristics that exhibited and challenge the design of opportunistic message forwarding are:

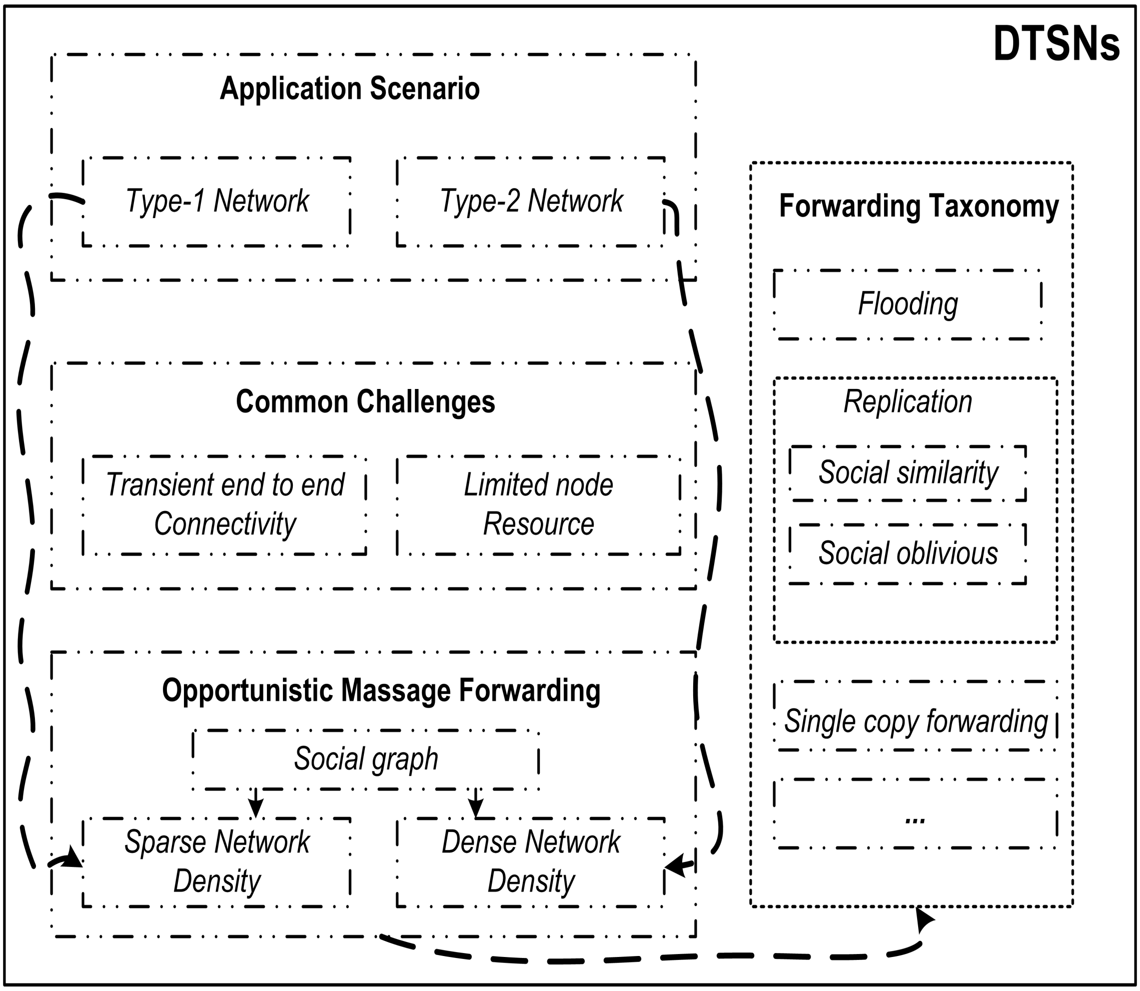


Figure 2.9: Characteristics of DTSNs and main operations

transient end to end connectivity and limited node resource. Each proposed forwarding solutions has a limited expectation of the mention characteristics. And network application scenario in DTSNs, is mainly characterized by the density of the network, where it has an influence on the underlying network model, social graphs, which represent the structure of the network. For instance, if the density of the network is low or smaller enough, it is called sparse network and/or development region network (type-1). Otherwise if density is dense enough, it is called urban scenario(type-2), as show in fig. 2.9.

Moreover, the doted arrow indicates that, each application scenario requires a unique set of forwarding schemes. In other words, it attempts to show, which forwarding schemes is best suitable for each application scenario. And the proposed forwarding algorithms rely on the social properties(i.e., Inter contact time, contact duration and interest information) which can be inferred from the mobility pattern of the mobile user based on contacts among nodes. A contact between any two pair of nodes is defined as an event that starts when nodes come within the transmission range of one another, and continues until they get reach out from each others transmission coverage range. For example, the time duration for which two nodes stay in contact is the contact duration(or contact time).

Besides, the dynamic nature of human behavior [33] has an impact on performance,and considering this features can also helps to streamline design of forwarding solutions either in type-1 network or type-2 network application scenario.

This story-carry-and forward method, as shown in figure 2.10, it is used in this kind of networks, where the whole messages are moved (forwarded) from a storage place on one node to a storage place on another node along a path that eventually reaches the destinations. The storage places can hold messages in a finite amount of time, based on the properties of the message setting and the underlying buffer management approaches. These storages places are called persistence storage. Furthermore, by moving whole messages in a single transfer, the message-forwarding technique provides network nodes with immediate knowledge of the size of messages, and therefore the requirements for intermediate storage space and retransmissions bandwidth.

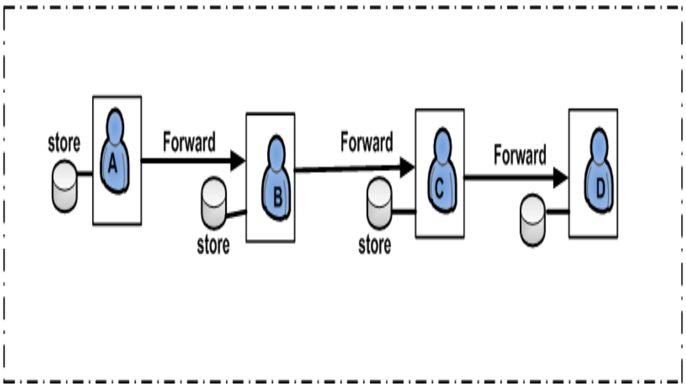


Figure 2.10: story-carry-and forward method

Moreover, in this kind of networks a message is not simply fragments of packets, rather it is bundle of actual messages with its fields as we do say packet header in conventional network. For example, some of the field a given massage have: *from, to, id, size,time Received ,time Created, content Type*, and others fields.

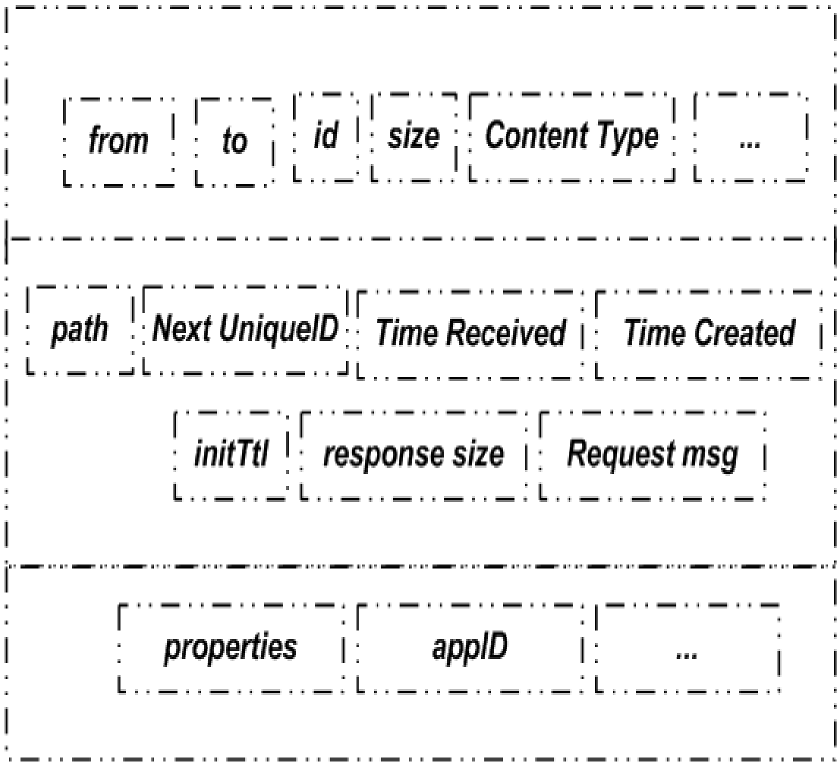


Figure 2.11: Message field in DTSNs

As shown in figure 2.11, *id*: field refers to identifier of the message, *size*: indicates size of the message in (bytes), *path*: field maintain list of nodes this message has passed. And *time Received*: field the time this message was received and *content Type*: refers to type message (i.e., interest of a given node towards a given content) *properties*: this field is container for generic message properties. Note that all values stored in the properties should be immutable because only a shallow copy of the properties is made when replicating messages. And *appID*: indicates application id of the application that created the message.

2.4.3 Applications

This section describes application scenarios where DTSNs can be used and a suitable application for opportunistic and delay tolerant networks [36, 37]. And the application scenarios can be grouped in two categories. A) Developing region application scenarios B) Others Services and applications.

A. Developing Region Application Scenarios

In developing countries perspective because of economic factors (e.g., no fixed infrastructure networks in some part of remote village rural area) and addressing this gap via infrastructure based network(e.g., the expensiveness of large steel towers backhaul link) only is not cost effective in the area where the density of the population is sparse. In addition, some digital contents and services are limited to a specific proximity of a given region, which are only required in that locality without highly dependent on the infrastructure network. Therefore, DTSN like networks could be an alternative solution and enable various application use cases. For example, village communication , social message exchange application and various services has a paramount advantages. Mostly the messages are non-real-time, and could be tolerable with some time interval in terms of hours or days.

Some of representative village Communication Network proposed rely on delay tolerant network that uses wireless technology to provide asynchronous digital connectivity services are daknet [38], kiosks[39].

In daknet for instance, the system consists of village Wi-Fi enabled kiosks, mobile access point-equipped vehicles and Internet access points, mostly far away towns from the village. These wireless devices transmit data via point-to point short links. Local users who live around in the village sends message in to Wi-Fi enabled kiosk , and during mobile access points (e.g., Mounted on and powered by a bus, a motorcycle, or even a bicycle) reach within the range of village Wi-Fi enabled kiosk, it exchange message (i.e., uploading and downloading). And carry the massages and transport with a physical means of transportation until it found an Internet access point, and then the message accessed from village Wi-Fi enabled kiosks delivered to Internet access points. With this fashion daknet system provide delay tolerate village communications to streamline message exchange services to provide last mile access. Inspired by daknet, kiosks is a similar projet that provide asynchronous village Communication Network.

In addition to the physical limitations of wireless infrastructure (mainly due to distance) to provide last mile access, there is also a growing disparity between those who can afford and not, because affordability and data plan cost add extra barriers for the average individual not to connect and consume data, let alone the people who are far from the connectivity coverage. Therefor, DTSN has attracted considerable attention, to employs social ties between nodes to offer data transmission services and used an alternative solution for the mention challenges. For instance, Social Network Service goose [40] and Social-based application D-Book [41] are some of the proposed applications in social message exchange services.

In Goose , it is short message service-style social network service that enables users to send both text and voice messages for non-digital literate users (i.e., users having difficulties in understanding a complex user interface), and strengthens the daily life message exchanges within the local inhabitants.

Depend on the available underlies network infrastructure (e.g., GSM) or taking social encounter among the nodes in the network the application streamlines message exchange within the users. For example, in the unicast case, a user uses Goose to send a message to a given user over GSM network. However, the application uses social encounters and epidemic algorithms for message dissemination (in the broadcast case) in the situation where no available network infrastructure exists.

Goose architecture is based on four components: the Network Manager, the Forwarding Manager, the Contacts Manager and the User Interface. The overall structure of Goose is shown in figure 2.12

Similar to Goose, D-Book, it is Social information-based application which provides social message exchange services(i.e., creating, modifying, and sharing profiles) among users in the network. In the application each user maintain a profile information which contains basic information, contact information and interests. Based on interest each user has, users can search for other users which have common interest. And enables users to send messages to other users and subscribe to the profile of other users based on interest.

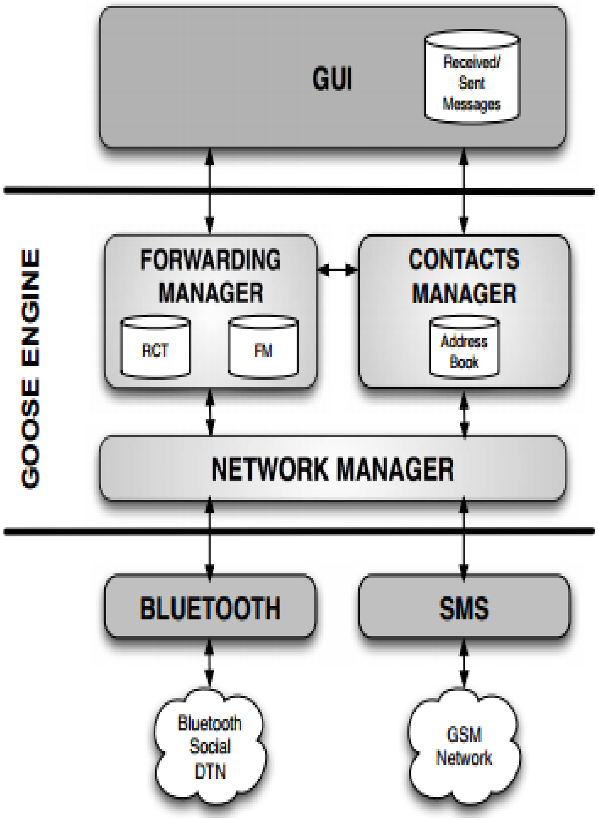


Figure 2.12: Goose architecture[40]

B. Others Services and Applications

In various domains of applications (i.e., Wildlife tracking and monitoring, disaster area networks and vehicle Ad-hoc Network) digital contents and services are limited to a specific proximity of a given geographical region, which are only required in that locality without highly dependent on the infrastructure network (e.g., traffic information among Vehicle , local news feeds and notifications services). Therefore, with this classes of applications opportunistic networking can be utilize an alternative solution.

In Wildlife Tracking and monitoring applications scenario, for example, zebranet [42], is used to track movement patterns of zebra by relying on peer-to-peer wireless Ad-hoc Network approaches. In zebranet system, each zebra mounted on a tracking nodes, collars attached to the necks of zebras and form network on large size of terrain under study. When two zebra encounters, the collars mounted on each zebra logs and exchanges information particularly tracks of locations information. And this kind of information keep exchanges among other zebras up on opportunistic encounters.

Moreover, Wildlife researchers drive across the area under study with a base stations to collect tracks of locations information. The base stations received information from the nearby zebra when they came with proximity. With this the researchers gained movement patterns of zebra population over a specified period of time (e.g., three days). For the nodes to log location information and exchange with other nodes up on encounters, it has positioning system (GPS), a flash memory, a wireless transceiver and a low-power CPU to perform all the operations.

There is also a similar approaches like Wildlife tracking and monitoring zebranet approaches in underwater delay tolerant communication network, SWIM [43], which is used to monitor the whales in the ocean. In this system, there is radio tag implanted on the whale and a SWIM base station floating on the water. When a whale nearby to the radio frequency range of an other whale up on encounters, it exchange the data about whale and the information being diffused with other whale in this fashion. And finally, when a whale nearby to the surface of the water to the frequency range of the floating base station on the surface of the water, it delivers (i.e., offloading) the information being collected so far to the floating base station.

Moreover, the characteristics of the environment which the network being used is challenging because, it requires nodes work loner period of time without human intervention and needs to have a better power source. Therefore, there should be an effective solution from hardware to algorithms design to cope with the nature of environment where this kinds of network deployed.

2.5 Digital Divide towards Developing Countries

The focuses of this Section is to overview and discuss about the Comprehensive definition of Digital Divide, multi dimension barriers that hinders the realization of inclusive Internet for everyone, and a model used for Comprehensive measurement to score the overall status of Inclusive Internet towards developing country.

2.5.1 Comprehensive Definition of Digital Divide

The number of individual who access digital resources via mobile devices increases exponentially. For instance in the world [1], the number of individual in mobile cellular subscription has grown significantly. This trend brings an opportunity for individual to access digital contents, share resources and communicate any where any time communications facilities.

However, the connectivity coverage (e.g., 3G/4G) rate did not go hand in hand to the trends of mobile phone penetration rate, and coverage of connectivity is concentrated in developed countries, and sparse in developing countries, where 78% of the population is online compared to 32% respectively. Developing countries are home to 94% of the global off line population [1]. That is what exclusion is all about which implies majority of the world population is offline in one way or another. In other word, it is safe to say that our world right now is not inclusive enough for everyone to access Internet also called inclusion, particularly in developing world perspective.

Digital divide, traditionally defined as the gap between those who can access digital information and service over the Internet, and those who can't. The assumption that individual can get online via mobile network as far as they live within the sufficient mobile network coverage do not valid any more; Since the disparity is not only created because of the connectivity coverage size and accessibility of the infrastructure network (i.e., excluding individual who are far from the connectivity coverage). Rather there are multi dimensional factors and barriers that fuels scale of this disparity in developing country perspective.

With this in mind there is comprehensive definition about Digital Divide. It is defined as the ability to meaningfully participate in and benefit from the Internet is also dependent on the ability to use and have confidence in the technology, and it is mainly affected by various factors and barriers, Availability, Affordability, Relevance and Readiness. This definition distilled with two main aspects: In one hand, it looks at the conditions to access information and communication technologies (ICT) from the supply side of digital infrastructures and services; On the other hand, it focuses on the levels of internet use, the motivations and abilities in using ICTs and internet services[2][44].

2.5.2 Multi-Dimension Barriers towards Inclusive Internet

In this Section, we take the following scenario in order to motivate and discuss multi-dimension barriers towards the realization of Inclusive Internet. Imagine that a coffee farmer in Ethiopia wants to access and use open internet to sell way more of his coffee products, A student in Congo using Wikipedia for free to study for her exams, and an expecting mother in Liberia looking up health information and how to safely raise her child. Taking in to account scenarios of this kind, how does a farmer for instance, coping all the barriers he face and be able to use internet confidently? How can an expecting mother really find relevant information tied to her realistic scenario?

With this in mind the key reason why individuals are not online and could not be able to participate fully in to the digital world challenged by various multi-dimension barriers as depicted in figure 2.13. There are four general categories. These are:

- **Availability:** It concerns to physical access to infrastructure network where individual live within the sufficient mobile network coverage. And a measurement about the quality and breadth of availability of infrastructure required for access. It is mainly quantified through the size of network coverage, number of Internet users and the average speed of the communication link, and these measurements are being used as an indicator for the over whole status of a given country towards availability.

- **Affordability:** It is all about the minimum cost that require for an individual to get connected with different ways to access, and be able to use internet. In this regard a cost to access ranges from cost of data plan (pre and post paid tariff) to cost of internet enabled handset. And these are some of the indicator to measure the affordability, and to what extent individual can afford based on its economic level. With this significant number of individuals fall behind off line, not because they can't access the internet because they do not live within range of sufficient physical infrastructure, rather they can't afford it.

- **Relevance:** As the term relevance conveys itself, the significance and leveraging of various application and services of Internet tied in to the daily life of user is also a common denominator for an individual to create a demand and use Internet. It can be different for different group of users. For individuals what value do they get being connected to the internet? It is a measure of relevance towards an individual in terms of useful services and contents and the availability of local content. Some of the indicators are availability and accessibility of basic information in the local language(s), local language(s) keyboard on devices and e-Government services in the local language(s). In this case although the infrastructure there and they can afford it, in reality majority of the population are fall behind offline. Due to the reason that no relevant information and services or insufficient content available in their primary language tied to directly in to their daily life, consider that mother in Liberia or that farmer in Ethiopia.

- **Readiness:** Compared to the other barriers towards inclusive Internet, this one is very broad. Due to the reason that it is tied to various socio-economic problems, from literacy (both basic and digital) to gender equality to access and use internet. In a broad sense, it is a measure of the ability and how safely use Internet. It ranges from the capacity to use mobile device and access internet to various application and services online. level of literacy, level of web accessibility, female access to the Internet, and existence of government e-inclusion strategy are some of the indicator to measure the status of readiness of a given country.

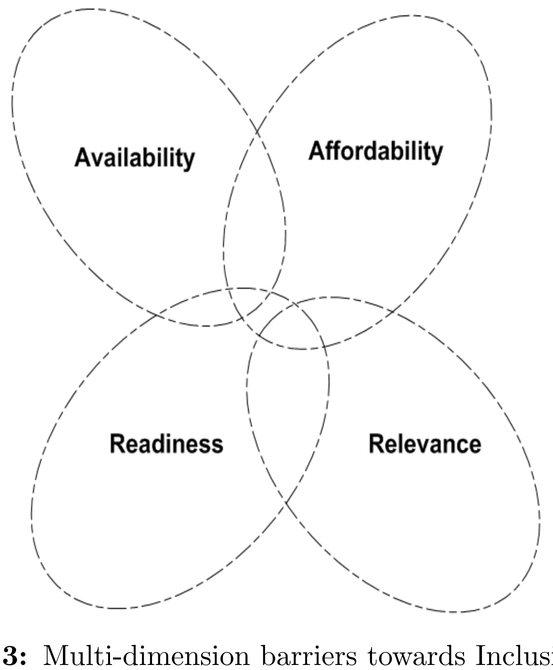


Figure 2.13: Multi-dimension barriers towards Inclusive Internet

Consequently, because of the comprehensive characteristics of digital divide, its solutions and mitigation do not have one size fit approach for all the barriers that hinders the adoption of inclusive internet towards a given country. And the solution should be carried out in a multi-stake holder approach, because of the reason that each barrier demands different tasks and activities but in a coordinated fashion.

This multi-stake holder approach, to achieve bridging digital divide or the realization of inclusion, all the activities and actions taken by government, commercial firms and civic organizations in a country has to be coordinated and implemented in a holistic approach to make this a reality. For this to happen, one can ask, how the overall score of status of inclusive internet for a given country computed in order to know its digital divide status. A model to score the overall status of Inclusive Internet for a given country, as shown in figure 2.14, and explaining top down, firstly, taking a measure of indicators in each category, for instance, indicator for affordability (i.e., cost to access internet in terms of pre and post paid tariff and cost of internet enabled handset with regards to income level of group of population) to score the status of affordability towards a given country. And the score of others category are measured in the same way by taking in to account respective measure of indicators. At this stage one can find that to what extent each barriers contribute towards for people falling behind offline. Added to that, for example, between relevance and readiness there is some kind of inter relationships as denoted by connecting curved arrow which indicates that it is not a mutual exclusive event[2][44].

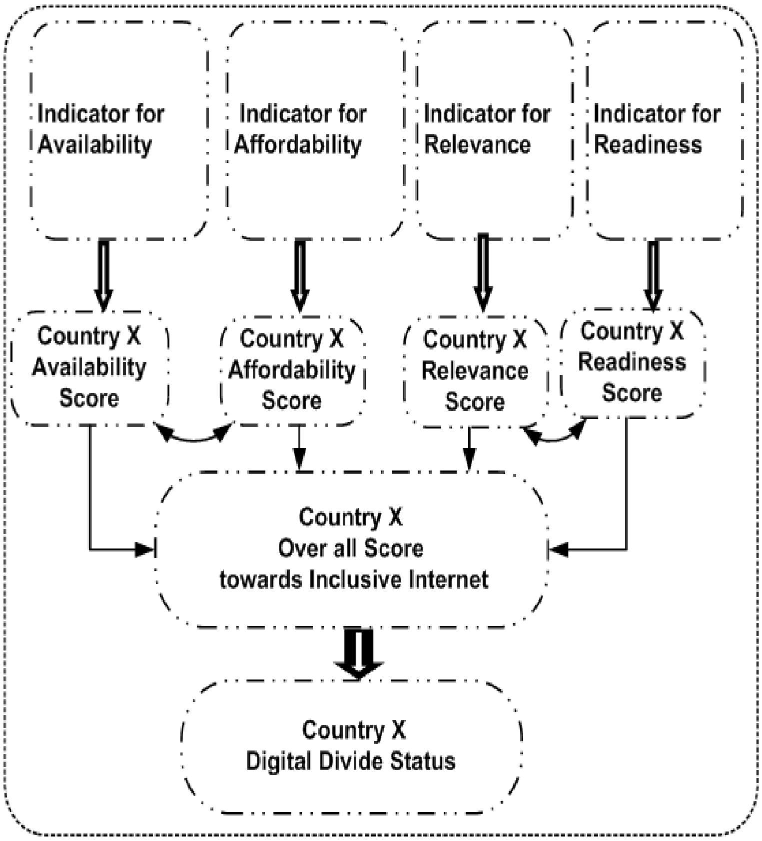


Figure 2.14: Model to score the overall status of Inclusive Internet for a given country

Secondly, one can follow this approach to compute the score of the overall status of inclusive internet towards a given country, by taking in to account a weighted average of the status of each categories score. Based on the result obtained and infer from it, knowledge about the status and level of digital divide towards a given country is obtained.

In conclusion, developing a specific strategy to address bridging digital divide and fostering of internet access among group of population, who are falling behind offline requires a comprehensive approach by different stake-holders. And one can use this model to optimize the strategies based on the result of the score towards each barrier having an impact on inclusion in general.

2.6 Open Issues

One of the fundamental goal of a communication network is to have the data or messages delivered to their corresponding destination or group of destinations. Depending on the underlying network characteristics, the data or messages are required to be routed or relay along in the network with different routing approaches. However, delivering the data itself is not worthy enough, if the underlying communication protocols (e,g., Routing and Forwarding scheme) perform its operations without aims to achieve effectiveness and various performance objectives.

In delay tolerance social network for instance, there are different groups of routing approaches in order to achieve distinct goals towards performance enhancement. However, the performance of routing and forwarding algorithm highly depend on different design considerations and assumption that one can take in to account. Let alone the common Characteristics (see fig.2.9) that challenge the underlying communication protocols operation and performance. Due to this there are open issues that require further exploration and investigation. These are:

- Improvement of performance and efficiency towards routing and forwarding scheme by exploring the comprehensive use of social metrics.

- Exploring an effective approach to collect, and store routing information (i.e., social information to be used for routing decisions) to design forwarding scheme with an optimal and acceptable network resource conservations (e.g., reducing energy usage and minimal buffer occupancy) for the mobile node.

- Forwarding and routing performance may be affected when the selfish behavior of nodes is considered. Because a node may discard the message received from due to resource constraints (i.e., energy, buffer space, and bandwidth) or malicious property. Exploring novel cooperative schemes and mitigation techniques with various approaches is a key component for the realization of DTSN in a real world. Moreover,most Proposed routing and forwarding schemes are designed for considering only in infrastructureless network, but it has paramount benefits if the scheme can work in both network environments (infrastructureless and infrastructure) such as dense city situations. Designing forwarding scheme that consider both network mode can bring an extra realism for the realization of opportunistic network in to reality. Besides handling of privacy and security problems.

Chapter 3

Related Works

3.1 Introduction

In this Chapter, we discuss about various approaches in forwarding algorithms in DTSNs. Delay tolerant networks have been proposed for more than one decade [25, 45]. Researchers focus on the data routing, one primary issue in DTSNs, and many studies have been carried out to handle the data delivery in the intermittent connected environment.

In Chapter 2, the routing approaches grouped in to three major taxonomies. However, replication based approaches can be further subdivided in to social similarity and social oblivious, based on the characteristics of each node in the network(e.g., users mobility pattern and social behavior). And this social similarity metrics can be represented through user interest, community and node popularity features. The overall forwarding taxonomies are depicted in figure 3.1.

The static network information, for instance, social information of the nodes in the network has gained much attention to design routing and forwarding algorithms in DTSNs. Since this kind of information tends to stable over time, and do not highly coupled with the underling topology in the network, routing and forwarding algorithms relying on this information do not highly influenced by the dynamic nature of DTSN (i.e., mobility), and has a better performance result compared to dynamic network information.

Moreover, forwarding schemes rely on Social similarity approaches improve the performance of different operations of data forwarding in DTSN. And in papers [46, 6], explained that social aware routing and forwarding algorithms are performed better than social oblivious routing algorithms. Due to the reason that forwarding schemes based on social relationships which is represented through social similarity are less volatile than forwarding schemes rely on mobility behavior of nodes where mobility patterns of nodes in the network changes now and then in the network and cause different operations of data forwarding computation unstable.

Moreover, in another aspect of the social features in the literature that streamline forwarding algorithms design is a community structure, when a group of users with social links, common interests, and similarities tend to interact with each other more frequently than those in other groups. These groups are called communities. In this community based forwarding algorithms, since nodes in the same community members frequently meet than different community, the forwarding strategies follows two distinct approaches to forward the data. These are intra-community and inter-community approaches.

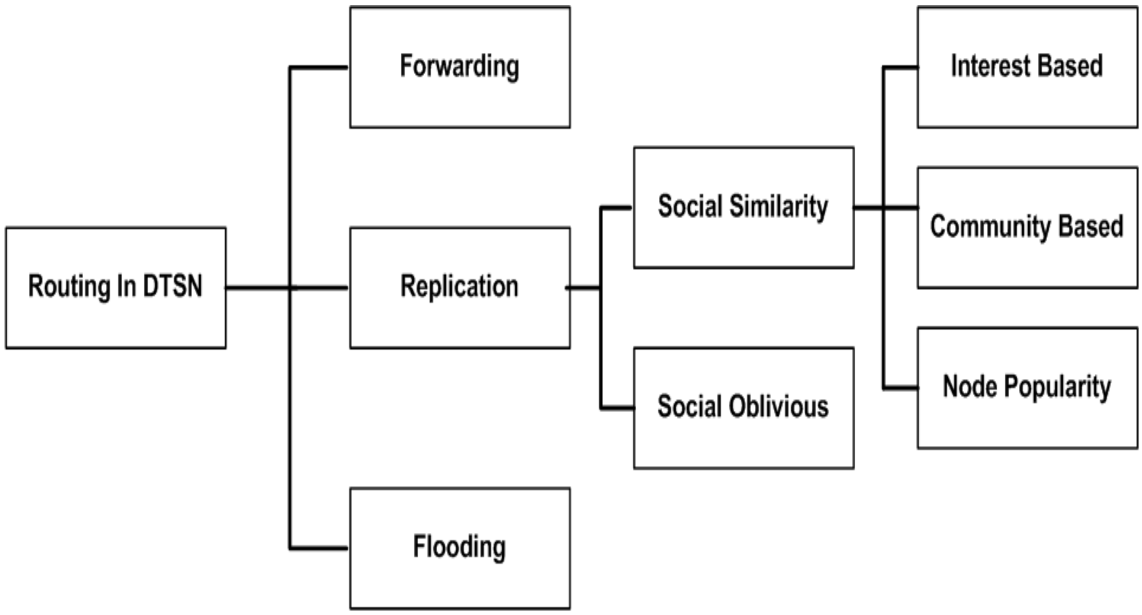


Figure 3.1: Forwarding Taxonomies in DTSNs

3.2 Prominent Examples

PRoPHET (a probabilistic replication based routing and forwarding scheme), Bubblerap (a social aware replication routing and forwarding scheme) and Epidemic (flooding based routing) are the most well known routing and forwarding algorithms that inspire most proposed routing and forwarding schemes in DTSNs. These schemes are used as benchmark routing algorithms to compare and evaluate different proposed forwarding scheme. Moreover, each benchmark scheme represents group of proposed scheme in given taxonomy (see, fig.3.1) to which it was designed for; to achieve a given objectives.

3.2.1 PRoPHET

In social oblivious categories for example, encounter based forwarding algorithms use encounter information of nodes in the network up on directly contact for data routing strategies. This kind of schemes forward messages according to node contacts frequency, and choose the node with higher contact probability as the relay for data delivery. For instance, a work that inspire most of the schemes in DTSN is Prophet [31], the Probabilistic Routing Protocol using History of Encounters and Transitivity used predictability for data delivery as the metric for relay selection. It has a probabilistic metric that is calculated based on the encounter pattern of node(e.g., frequency of contacts between nodes). This probabilistic metric, delivery predictability is a strong signal to deliver a message to a destination based on its past contact patterns and computed by the following equations from eq.3.1 to eq.3.3.

$$P_{(AB)} = P_{(AB)old} + (1 - P_{(AB)old}) \times P_{init} \tag{3.1}$$

$$P_{(AB)} = P_{(AB)old} \times \gamma^{\kappa} \tag{3.2}$$

$$P_{(AC)} = P_{(AC)old} + (1 - P_{(AC)old}) \times P_{(AB)} \times P_{(BC)} \times \beta \tag{3.3}$$

And these equations used to compute and update delivery predictability utility, its aging factor and the property of transitivity among the encounter nodes.

Moreover, $P_{(AB)}$ and $P_{(AB)old}$ are the current and previous delivery predictabilities respectively of node A towards node B over some contact patterns. And P_{init} is initialization constant. κ and γ are parameters aging constant and the number of time units expired since the last update of this predictability respectively. And β is a scale constant to what extent transitivity property affect delivery predictabilities utilities.

In summary, the scheme uses delivery predictability utility of a node to sending message towards a destination by selecting a node which has a higher utility value as a relay node. In addition, transitivity property used to select a relay node towards to the destination in the condition where contact patterns among nodes have transitivity. However, there is a high communication overhead due to an exchange of encounter information between the nodes.

3.2.2 Bubblerap

In bubblerap[32], a community based routing algorithms that forward message in pocket switch network. By using k-clique community detection algorithms, the schemes find the node to which community they belong. To forward the message from source to destination at any given point in time, the schemes compute the centrality value (i.e., betweenness centrality, it is metrics which qualify the level of relevance of the node in the network for message routing. For each node which it comprises of local value and the global value and their rank compared to the neighbor node. As shown in fig 3.2 once the messages bubbles from the source node, it traverses from one node to another node based on the global ranked centrality value of each node which has higher rank chosen to rely until a node which is belong to the same community where the destination node is found and then, using local centrality value, and hence once a member of the desired community receives the message, it uses local ranking of the other nodes in the community to further disseminate it until the message reaches to its corresponding destination node.

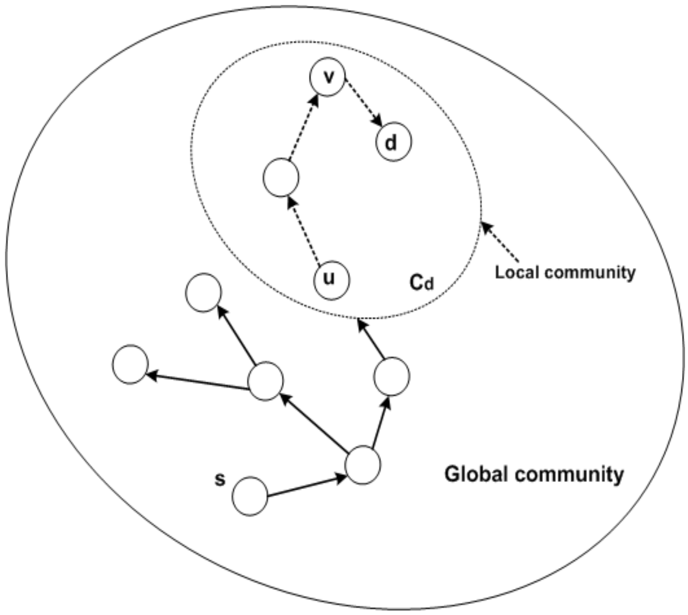


Figure 3.2: Message forwarding process in the Bubble Rap[32]

In similar way community based epidemic forwarding, Localcom[47], which used an extended clique distributed approaches which only use local information of the nodes to form and detect the community structure. And exploiting the virtual link between a pair of neighbor nodes to represent the relationships and finds at least one path with the maximum of k-hop distance between them. During message forwarding between nodes in same community, also called intra community forwarding strategies, LocalCom, used similarity metrics, which composed of closeness relationship between nodes that represent the minimum average separation period in terms of a given time, and some irregularity of separation period during at the same time reflected by the variance of separation period. The minimum average separation period is deduced from the frequency of contacts between neighbor nodes and duration of contacts, which helps to formulate similarity metrics. And messages forwarding operation done, based on high similarity metrics value and short hop count distance.

In the case of inter-community forwarding strategies, the scheme used flooding based approaches via bridge (i.e., these are kind of nodes which has a direct neighbor relationship between different communities.) nodes, where two or more communities are connected. Moreover, selecting which bridge nodes to facilitate forwarding between communities requires static and dynamic pruning mechanism to reduce redundancy because of flooding, and to limit the number of bridge nodes used. And message forwarding proceeds by using betweenness centrality value of the bridge node, from the current community where the message is originated to other neighbor community bridge node. Using the virtual link, the bridge node forwards the message to all local community members nodes where the bridge node is belong.

LABEL[48], a community based routing algorithms that apply social properties to forward message. In this approach, nodes in the network assigned label which helps to identify the affiliations. When Nodes encounters each other, exchange labels, and compare it towards label of the destination nodes. And the messages forwarded if the destination nodes belongs within the same community of the encounter node. However, this kind forwarding strategies has a difficulty when the destination node is not within the same community. And this brings low network performance (i.e., delivery ratio, average delay). However, these approaches suffer with the overhead of community formation and a drawback in regards to average delay since to form and have as table view of the community, takes more time.

3.2.3 Epidemic

Epidemic[29] is one of the most common proposed flooding based routing scheme which model message replication similar to the spreading of epidemic diseases in real life. Due to unlimited message coping behaviors, the scheme has both advantages and disadvantages. Forwarding performance objectives for instance, delivery ratio maximized and average latency is reduced during sending a message towards a destination, because, out of the replicated messages the chance that one of the message reaches to the destination is higher. However, a lot of network resource is consumed, for example the overall storage requirements become high in such networks.

3.3 Positioning of Contribution

In social aware based forwarding algorithms, the forwarding strategies rely on and examine the social properties of the nodes in the network when they encounters. Several schemes have been investigating the exploitation of social relationships as well as individual interests through social similarity in order to improve network performance.

In this categories, Mei *et al.*[49], based on the observation that peoples movement is affected by their interests, and thus propose a social-aware and stateless routing(SANE) which represent interest profile of a node by K-dimension vector space and the angle between two vectors (e.g., two encounter nodes in this case and K-indicates the number of interest profile) is used as a measure of divergence between the vectors by using cosine similarity metric (i.e., identical vectors value is 1.0 , and 0.0 for orthogonal) to measure the social tie of nodes to predict future contact opportunities. And message will only be forwarded to the node if the cosine similarity between encountering nodes is larger than a threshold. However, peoples movement would not be always reflected through interests, therefore, more node properties and characteristics utilized to measure social tie strength in order to predict contact opportunities among nodes in various network application scenario.

For instance, in BEEINFO [9], community structure is formed based on the user interest, and data forwarding strategies in the intra community case is done using the social tie metric and community density metric is used for inter community case. Moreover, the scheme inspired by food forging of bee, also called bee colony algorithm. It uses this algorithm to have an awareness capability to the context of the environment by taking in to account social properties of nodes in the network, similar to the nectars forging of bees. A similar work Int-tree [10], examined the relationship between interests of mobile node users and its impact on the performance of the data forwarding schemes , since relay selection and forwarding decision are critical to be made by current node based on certain routing strategies. It categorize the relation among users interests in to three types. Those are inclusion, cross-layer, and intersection via layer based concepts. However, they focused only the inclusion relationship type, and community based forwarding strategies using community density and social tie metrics. And for intra community (inside the community) forwarding strategies, a given node measures social tie between the encounter node and a node which has higher value is selected for a better relay. However, both this schemes have drawback in average latency.

In Dlife[50] a weighted contact graph represents the dynamics of the social structure. The scheme relying on the social properties and its dynamics nature over certain periods of time for specified number of nodes. By using weighted contact graph that represent the social structure and its dynamics behaviors, they studies how does the social structure evolve over some amount of time period. Moreover, they formulate TEDC(time evolving contact duration) that captures the evolution of social interaction among pairs of users; and TEDCi(time evolving contact duration importance) that captures the evolution of users importance, based on its node degree and the social strength towards its neighbors. Both utility functions measures and quantify contact duration and the importance of a node towards its neighbors. And the forwarding algorithm as follows when nodes encounters and comparing with its social strength towards the destination, nodes which has a higher relationship to the destination receives the message, due to the assumption that there is a greater chance for the encountered node to meet the destination in subsequent contacts. If the relationship to the destination is less or unknown, the scheme replicate the message to the encountered node which has a higher importance than the carrier.

In work [51], this proposed routing algorithms relying on the observations that nodes in the networks having common interest have more contact opportunity than nodes having disjoints interest. And have a higher more connection in the social graph in opportunities network. The scheme proposed to address congestion problems and to reduce message overhead and average delay during forwarding, occurred in the network that impacts the performance. However, grouping nodes based on the number of common interest it has limit the membership of nodes to a few interest social groups, which affects the scalability of the scheme.

3.4 Summary

In summary, replication refers to the process where a transmitting node sends a copy of the message to the receiving node and itself keep the original message. And there is multiple copies of a message can exist in the network. Such approaches are used to enhance the chances of message delivery because of a transient end to end path apparent in the network. It is indeed both a multi-copy forwarding approaches (flooding and replication-based approaches) aim to optimize delivery probability. However, in the case of flooding, there is an unlimited message coping, and a replication-based approach uses some kind of heuristics to select to a given node to replicate message for.

The heuristics that guide (i.e., selecting a node and replicate the message to) replication of message during forwarding process mainly rely on network behavior in which each individual node exhibits in the network. Forwarding scheme s which based social properties and relation as a heuristics, achieves a statistical significant performance results, but still there are room for further performance enhancement by incorporating some additional social metrics. IntReF (Interest Relationship Based Forwarding Scheme in Delay Tolerant Social Networks) is proposed. It belongs to in this category of forwarding schemes.

IntReF taking to account the comprehensive use of social metrics and knowledge about users interest aims to achieve further performance enhanced results. Its message forwarding operations proceeds as follows; nodes exchanges message within the group by selecting a node which has a higher time evolving contact to interest social strength value. And message exchange between nodes in different groups by selecting nodes which has a higher forwarding capability value during nodes up on encounters. Knowledge about interest relationship conveys to which group a given node belongs to. Related work summary of schemes are depict in table 3.1.

Table 3.1: Related works summary

Routing algorithms	Taxonomy	Interest based	Community	node popularity features
Epidemic[29]	Flooding	×	×	×
PRoPHET[31]	Social oblivious with replication	×	×	×
BubbleRap[32]	Social aware with replication	×	✓	✓
LABEL[48]	Social aware with replication	×	✓	×
BEEINFO[9]	Social aware with replication	✓	✓	✓
Dlife[50]	Social aware with Replication	×	×	✓
Local com[47]	Social aware with Replication	×	✓	✓
Int-tree[10]	Social aware with Replication	✓	✓	×
IntReF our proposed work	Social aware with Replication	✓	×	✓
SANE[49]	Social aware with Replication	✓	×	×
ONSIDE[51]	Social aware with Replication	✓	✓	×

Note: ✓ - Satisfies the condition × - not satisfies the conditions

Chapter 4

IntReF: Interest Relationship Based Forwarding Scheme

4.1 Preliminaries

4.1.1 Network Application Scenario

Delay tolerant social networks can be illustrated as a graph based on different behaviours of nodes and interaction structures among individual units or nodes [16, 46]. And the interaction is information exchange among nodes in the network. And social graph is the crucial tool to model this interaction.

A social graph is a graph-based representation of the network, where nodes corresponds individual entities and edges represents social ties between nodes. In this thesis, we use social properties(contact duration, interest information) to characterize the network graph. We model DTSN as a social graph G = (V, E, W) where V is the set of mobile nodes in the network, the set of social tie is represented by E and the set of tie strength (weight) is depicted by W. The social links indicate the social relations between two nodes and the weight of a link suggests the social strength and/or intensity. In this kind of networks, the social graph for a particular node consists of a set of nodes and ties connected directly or indirectly to this node. And it facilities the exchange of message among nodes in the network.

There is also vast amount of data is produced and consumed every day, from different application to individual mobile devices. And data sharing and/or exchanging among /between users are a growing demand in most application environment. And from observation as well as its apparent characteristics, there is a data category, and different interested parties have varied level of demand either to consume or produce the data. Moreover, there is also an inherited phenomenon to observe in various application use case that users are grouped in to different subjects, and exchange data within the group and among/between them. For instance, people in university campus with different group of subjects (e.g., law school students, computer science 1st, 2nd, and staff) exchange information based on their related interest.

As depict in fig.4.1, group of nodes relying on data categories, represented by related interest of the nodes (i.e., Interest Relationship) in our network application scenario. The dynamic nature of the network as well as its environment awareness by the node (understanding the context where the node is), and message exchange have two phases, forwarding within the group, and between groups as show in fig.4.1. In leveraging DTSNs to achieve various goals, to mitigate and by pass expensive cost of data mobile communication, or to bring an alternative communication approaches in the application scenarios.

In this thesis, there is an implicit assumption in the design of forwarding algorithms that all nodes in the network are cooperative and they are willing to work together in the data delivery process. Although the characteristics of some nodes in the network non-cooperative which is termed as selfishness behaviors.

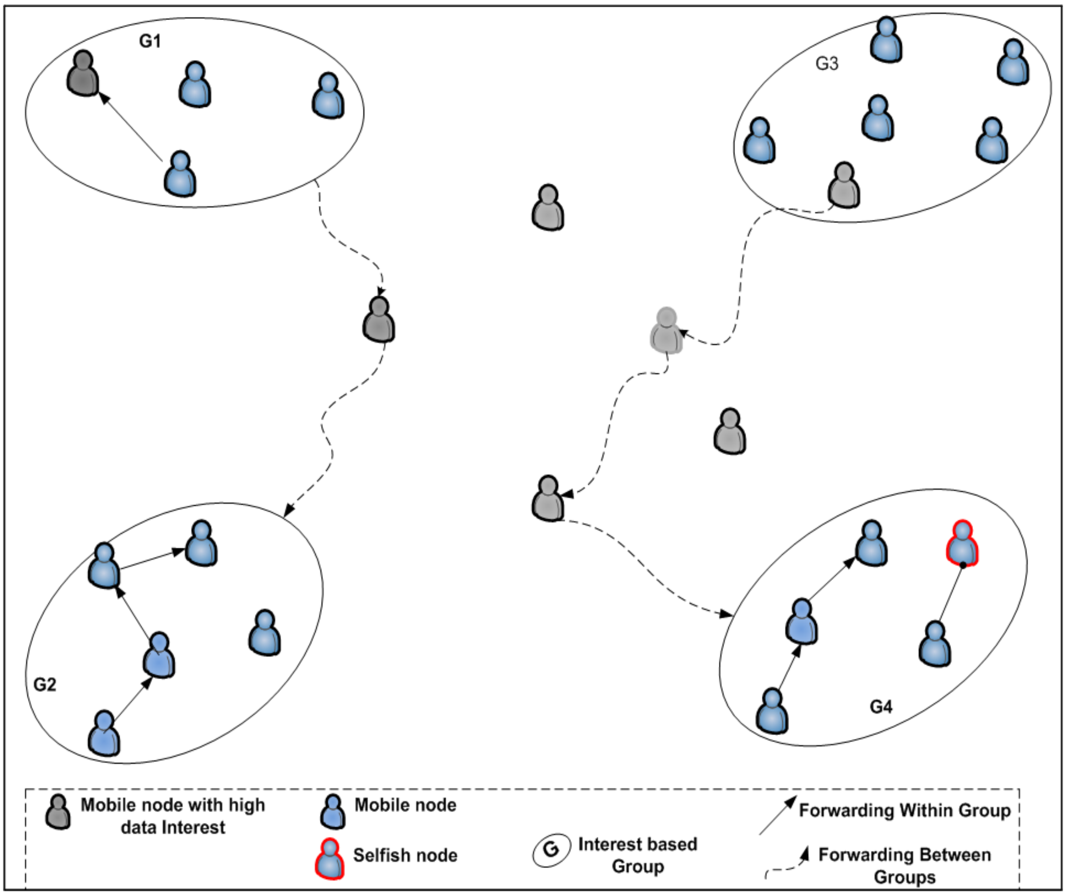


Figure 4.1: Conceptual framework,with network application scenario

4.2 Components

In this Section, we describe the main components of the proposed forwarding scheme, IntReF represents Interest Relationship Based Forwarding Scheme as depicted in fig. 4.2. IntReF has three components. Those are social graph formation component, interest relationship component and forwarding strategy component. And discussion follows in bottom to up.

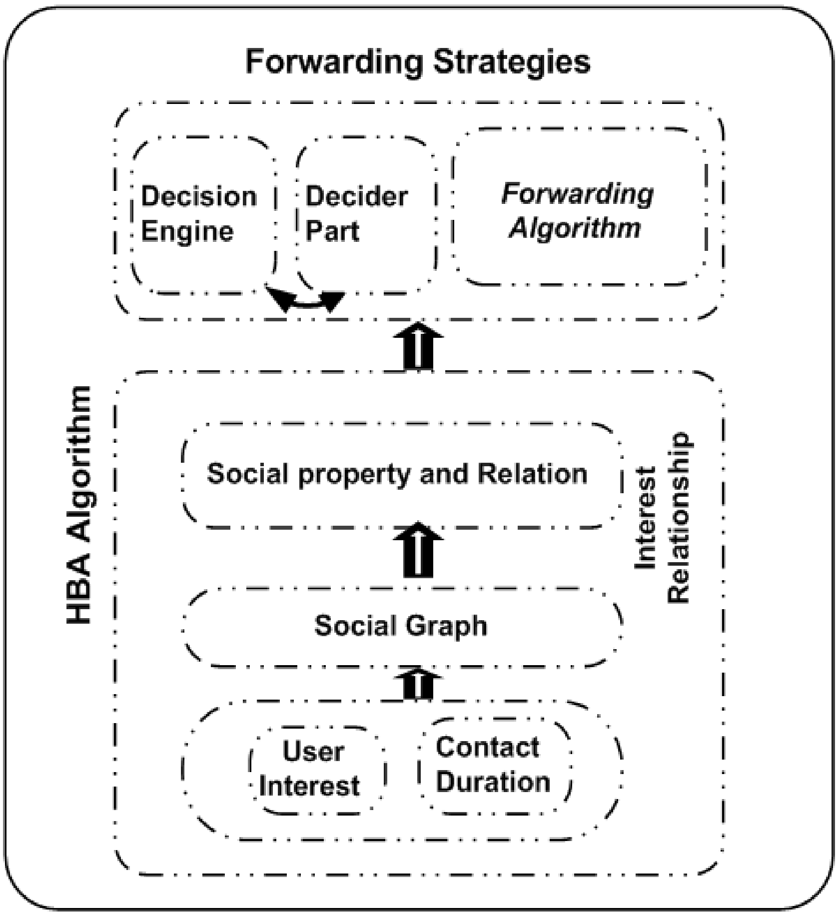


Figure 4.2: IntReF architecture

4.2.1 Social Graph

Social network analysis(SNA) methods can be used to examine various state information and properties of nodes in the network to design efficient routing and forwarding schemes in DTSN. This technique can retrieve inherent relationships and structure of mobile nodes and infer essential properties and metrics of a node which can be used for a better signal to streamline forwarding decisions (i.e., selecting a relay node via centrality, social similarity and tie-strength).

Moreover, the application of SNA, to represent social behaviors and interaction structures among nodes in the network through a social graph is vital. And to better cope and represent the dynamic nature of peoples social behaviors with social graph structure in different time period in order to formulate social similarity and other properties.

In this thesis, we rely on and adapt human behavior aggregation algorithm [52] to form a social graph that reflects varying strength of relationships among nodes in the network in different time periods. This algorithm form the social graph between any two pairwise contacts of nodes with contact duration towards node interest over daily sample, in this algorithm perspective there are 24 time Slot or daily sample of which each has one hour long(i.e., which is an interval time in which nodes in the network have a similar behavioral patterns), and taking contact duration in to account in order to measure different levels of social strength between pair of nodes. And equations computing social strength is depicted in eq.4.1 and eq.4.2.

A. Social Property and Relation

Taking social graph as an input and applying Social network analysis technique in order to formulate social property and relation among nodes in the network. The quantification of the relative importance of nodes and different levels of relationship between nodes using social property and relation streamline forwarding decisions in different levels. Three used Social property measures are contact to interest, interest density and forwarding capability. A combination of this Social property metrics can be used for a better signal during forwarding operation to determine which contact of a node has the strongest social relationship towards destination.

B. Social Strength

We borrowed from [52], to measure different levels of social strength, total time connected to interest, the average total time connected to interest, and time evolving contact to interest(T_e).

The time evolving contact to interest utility function measures the contact duration of a node with a given interest in different time slots. It measures the social strength of a node to the other nodes for a particular interest given by contact duration. And it facilitate message exchange between nodes within interest social group in our network application scenario.

In order to determine time evolving contact to interest utility function, it follows three steps: at first, measuring total time connected towards interest, and secondly computing the average total time connected to interest, and finally formulating T_e.

For total time connected towards interest(TCT) consider that a node A has N contacts with an other node B having an an interest I_1 with some contact durations($CD_{(AB)k}$) within an interval time period, and the total connected time to towards interest with each k contacts given by the sum of all each contact duration, taking this information in to account, and using cumulative moving average, the average total time connected to interest(SoT_{AB}) computed as show in eq. 4.1:

$$SoT_{AB} = \frac{TCT_{AB} + (j-1)SoT_{j-1}}{j} \tag{4.1}$$

where, TCT_{AB} is total time connected towards interest , $(j-1)SoT_{j-1}$ is the previous social strength, and j is the current day.

Finally, T_e utility function which measures the social strength between node A and node B within different time slots ΔTi(i.e., i=1,2,3) in different days is formulated as show in eq. 4.2:

$$\mathsf{T}_{(e)AB} = \sum_{k=0}^{i+t-1} \frac{t}{t+k-i} \times SoT_{(AB)k} \tag{4.2}$$

where, k- some sample in the same day and t-represents the total number of daily samples

4.2.2 Interest Relationship

A. Definitions

User Interest: Interest of a user is a specific type of service message the user is interested in. A user may have several different interests, and each interest is associated with a weight, standing for the degree of liking. In this thesis work, interest of the users and its type can be considered in various ways, some of those are: Item type, Message type (content type). Moreover, granularity difference among users interest is apparent. This different granularity between users interests (i.e., some interest very general, and others so specific and detail) for example, having an interest just about a car, and car with its model and other attributes reflects different granularity between users interest. In this study, we take granularity at coarse grained level (in the sense of generality in other words), since knowledge about users interest is required to group nodes, and message forwarding among nodes streamlined based on interests in the network.

Interest Relationship: In this thesis, we define interest relationship, the apparent relationship among interests of a given node A, which has multiple distinct interests for example, I_1, I_2, I_3, I_4, and I_5, I_n. Inspired by dynamics of interest and its scaling in human daily activities; the relation among multiple interests of a node is represented through ranking size of each interest, based on the associated weight to each interest standing for the degree of liking. And conceptual, this multiple interests of the node do not apparent at the same time all at once, and individual tend to stay in to a particular interest which has highest rank size.

B. Discussion of Interest Relationship

Recently work Int-Tree [10], examined the relationship between users interests, and categorize the relation among users interests in to three types. Those are inclusion, cross-layer, and intersection via layer based concepts. Indeed, characterizing multiple interests of a certain node and their relationship in this way has so many advantages, but from the characteristics of nodes which found in opportunistic network, modeling relationship of interests via tree structure raises some issues. For instance, there is a high mobility of nodes, maintaining the tree structure creates significant network overheads and also keeping its consistency is a difficult task, unless and other wise some central server manages this task, but this assumption is unlikely to be found such a server in opportunistic network environment in general, except with some application scenario, for instance, conference managing platform as in [10].

However, in this thesis, inspired by human dynamics research, which has uncovered astonishing statistical characteristics such as, scaling in human daily activities. For instance, in the work [8], human interest dynamics showed fat-tailed probability distributions, which indicates the dynamics among individuals interests, conveys that individuals tend to return to highly ranked interest with relatively larger probability and stay in these interests.

Since there is a dynamics among individuals interests (e.g., node with multiple interests in the network), we represent the dynamics through interest relationship definition, and rely on it, nodes in the network tend to group based on per interest of a given node which has a characteristics of higher frequency of occurrence to be a higher ranked interest (i.e., an interest which has a higher degree of liking by the node itself, compared to the other interests of the node has.) In other words, at a given daily sample intervals, nodes having similar ranking to a given particular interest among the set of interest they have, tend to group in the same interest social group.

With this definition various interest social group is formed based on the given interest with its ranking size. Moreover, grouping nodes based on per interest of the node relying on interest relationship, assure the scalability of our solution (i.e., node can be a member of many more interest social group based on the dynamics of the interests it has over a given daily sample.) whereas, grouping nodes based on the number of common interest it has limit the membership of nodes to a few interest social groups.

C. Interest Relationship Model

In our interest relationship model, each node with its distinct multiple interests are considered. The behaviors of user to its multiple interests can be represented via a bipartite graph as depicted in fig. 4.3 And if any behaviors happen of a user to a given interest, there will be an edge between them. The weight of the edge between user and its interest node can be measured by the degree of a likeability of a user to an interest.

The value of degree of liking of a given interest by the node ranges from 0 to 1, where 0 denotes the degree of likeability of the interest is none and 1 indicates there is maximum degree of likeability to the given interest. When the degree of liking within (0, 1), it means node has some degree of liking to that interest. Furthermore, each node in the network is represented by the interest, which already has a highest weight over the given daily sample. And in each encounters, node exchange this interest, and used to identify whether they have the same interest or not.

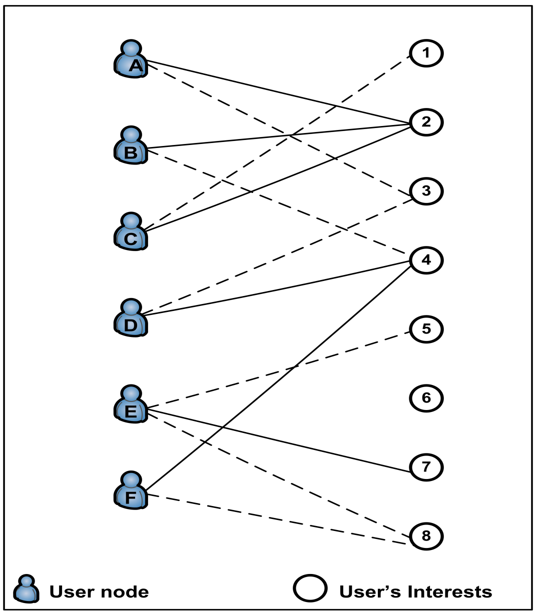


Figure 4.3: characterizing node with multiple interests and their relationship

To elaborate modeling of user with multiple interests and its relation, in fig. 4.3, there are six nodes (A, B, C, D, E, and F) with different multiple interests (I_1, I_2, I_3, I_4, I_5, I_6, I_7, and I_8) and each node has at least two interests. And their ranking size to a given particular interest is different and based on this apparent interest relationship; nodes tend to form a different interest social group. The higher the weight of the edge between user to its interest node, which is a solid line, as shown in fig. 4.3, it signals there is a higher ranking size, and nodes represented by that particular interest.

For instance, nodes A, B and C have a similar ranking size (i.e., highest weight) towards interest I_2. Whereas nodes D and F have a different ranking size towards interest I_2 compared to the mentioned node (e.g., node A), but a similar ranking size towards interest I_4. And node E and others would have a similar ranking size towards interest I_7. With this, nodes tend to group in to various interest social groups based on the given interest with its ranking size.

Table 4.1: List of variables and notations

variable	Description
A,B,C...	nodes
T	content type of Message M
ρ_t	total interest density
ρ_s	density over similar interest
ρ_f	density towards different interests
Aρf	Average density towards different interests
T_e	Time evolving contact to interest
F_C	Forwarding capability
r	weight of a particular interest
N	total number of contacts over a given time window
n	number of encounter nodes towards to same interests

D. Interest Density

Since message exchange is based on interest, the more interest a node has, the better it serves others nodes matching these interests, and it is the assumption considered in this work. And these nodes can be used as a potential bridge node to facilitate message exchange between nodes among different interest social groups.(forwarding among the groups) as depicted in fig. 4.1

Interest density of a node measures a rate of contacts of a node to different interests and it indicates the possibility of having more shared interest with the others nodes, and intuitively, if a node has higher interest density value, it shares more common interest with other nodes, thus can act as "interest hub node" and more likely to traverse message towards destinations. In this work, "interest hub node" is introduced to measure a power of a node concerning different interests to exchange messages among different interest social groups.(forwarding among the groups)

This density is not an intrinsic property of a node. It can change depending on how the node encounters with the others nodes, and this total interest density is reported both the sum of density towards same interest, and different interest in a given time period (i.e., a time slotΔTi). Since density towards different interests reflects the power of a node concerning different interest during encounter (contacts) in the network among the nodes. With this definition, for instance, in fig.4.3 the density of node A towards different interests (ρ_f) is the result of difference between total interest density(ρ_t) and density to same interest (e.g., sum of weight B, 2 and weight C, 2). Therefore, for a given contacts N a node has, this features of a node computed at the end of a time period Δ Ti is given by eq.4.3

$$\rho_t = \rho_s + \rho_f$$
$$\rho f = \sum_{j=1}^{N} r_j - \sum_{i=1}^{n} r_i \tag{4.3}$$

where, $\rho_t = \sum_{j=1}^{N} r_j$ **and** $\rho_s = \sum_{i=1}^{n} r_i$

Finally to represent a rate of contacts of a node to different interests, ρf is divided by N. And the values of ρf bounds in between 0 and 1. And both values indicates the extreme cases where either $\rho_t = \rho_s$ or $\rho_s = 0$ respectively. In opportunistic network perspective, these two cases are rarely found, the reason that, in one hand, since nodes are mobile, there is a chance of encountering nodes with different interests ($\rho_t \neq \rho_s$), and in the other hand, some nodes can have common interest with other nodes ($\rho_s \neq 0$).

Therefore, the value of ρf , lie in between 0 and 1 exclusive, which is a good signal to identify a better "interest hub nodes". To have the average value of density towards different interests (Aρf) for a given node with in different time(previous and current value), we take cumulative moving average of ρf_j and illustrated in eq.4.4

$$A\rho f_i = \frac{\rho f + (j-1)A\rho f_{j-1}}{j} \tag{4.4}$$

where, ρf is from eq.4.3 , $(j-1)A\rho f_{j-1}$ is the previous density towards different interests and j is the current day.

E. Forwarding Capability

Forwarding capability is the ability of node to perform forwarding operations during message exchange among the nodes in the network through a set of measurable features. It is considered that this capability of a node is a collection of various aspects of node properties (i.e., contact duration and rate of contacts towards interests) for a better forwarding performance, and helps to relay message to traverse towards destinations.

Moreover, this comprehensive nature can be quantified by adding feature together to represent forwarding capability of the node. From analysis and comparison of social aware forwarding approaches with and/or without content knowledge in to the forwarding approaches respect to performance in [53], it is noticed that there is a remarkable phenomenon that hybrid (or mixed) forwarding schemes often exceeds social-based approaches not mixed with content knowledge in terms of forwarding performance in different settings.

However, combining multiple forwarding strategies together may introduce the complexity of forwarding design (complexity of algorithms). In this regard concerning to our schemes, the information nodes maintains and exchange (i.e. interest information) is small enough and mobile nodes process limited information, which saves resources.

Therefore, the trade-off to be paid for combining multiple forwarding strategies for the improvement of forwarding performance is not too complex. Consequently, a comprehensive forwarding strategy, forwarding capability is defined in terms of density towards different interests and time evolving contact to interest metric.

This forwarding strategy takes the linear combination(sum) of the time evolving contact to interest and density towards different interest, and the former measures social strength of a node to the other nodes for a particular interest given by contact duration, and the later measures the power of a node concerning to different interests. This forwarding capability of node A at a given time period defined as follows in eq.4.5

$$F_C = \intercal_e + A\rho f \tag{4.5}$$

Where, $\intercal_e$ and Aρf are from eq. 4.2 and eq. 4.4 respectively.

4.2.3 Forwarding Strategies

A. Forwarding Algorithm

The forwarding process in our scheme, IntReF, proceeds as follows: during nodes encounter, they exchange their interest which has a similar ranking size among the list of interests they have. This interest has a highest weight within a given time period which helps encountered nodes aware themselves whether they have same similar ranking towards interest or not. Because having similar ranking towards interest indicates they are in the same social interest group.

Once node get knowledge about interest relationship, forwarding of message exchange within a group facilitate via the node which has a higher time evolving contact to interest utility value towards message destination, based on the utility value measured using in eq.4.2. Whereas, message exchange across different social interest groups using forwarding capability value of node, and node which has a higher value of forwarding capability could be selected to a potential forwarder. In our scheme node which has higher forwarding capability value called as an "interest hub node" and can be selected as a relay node for potential next forwarder, based on the utility value measured using in eq.4.5. And the pseudocode show the Forwarding Algorithm process.

For the sake of simplicity, in this example nodes A and B are the nodes that encounter each other, and message M has its content type. Forwarding strategies are described from node A perspective as a message $\boldsymbol{M}$ carrier.

Algorithm 1 Pseudocode of forwarding Algorithm

```
Given a message M with its T in the buffer of node A
for each node B encountered by node A do
// simple interest match
if T==I_B then
   A.replicateTo(B, M)
else if check within the group then
   if T_(e)B(B,M) > T_(e)A(A,M) then
      A.replicateTo(B, M)
   else
      // Forwarding among the group
      if F_(C)B(B,M) > F_(C)A(A,M)  then
         A.replicateTo(B, M)
      end if
   end if
end if
end for
```

B. Decision Engine

In order to perform forwarding based on the interest relationships of users, the decision engine: keep track of the contact duration to interest and degree of liking towards interest of each encounter between nodes; store this information, and facilitate message exchanging based on the response from the decider part.

Events for example, change of connection between encounter nodes (e.g., A and B) set to be up, the collection of message is examined to determine if a massage M should be sent to this new encounter node B, this to happen, nodes exchange forwarding information and updates themselves, here comes the decision engine interacts and request to the decider part, should message M be replicated (send) to node B? Hence decider part return value of true or false based on the forwarding strategies employed. A true value indicates the message should be added to the message store for forwarding and a false value indicates the message should be discarded.

C. Decider Part

This component is responsible for making decisions whether replication should occur or not. It also interacts with the decision engine to get relevant information(i.e., interest relationship, social strength and interest density) in order to perform forwarding decision up on nodes encounters/contact in the network.

4.3 Summary

Modeling and representing various network application scenarios in DTSNs streamline with graph based representations. In such a network different behaviors of nodes and their interaction represented via a social graph. In this graph, a set of nodes as the vertex of the graph and the interactions or contact represented via a set of edge and its strength by weights. With this graph representation, each node maintains its relevant forwarding information to facilities message exchange among nodes in the network.

Moreover, there is an inherited phenomenon to observe in various application use case that users are grouped in to different interest, and exchange data within the group and among/between them. Considering application scenario of this kind and used in delay tolerance social networks to guide message exchanges between nodes. And helps to design data forwarding scheme. The proposed forwarding scheme guided by this kind of application processes. Forwarding scheme would follow either message exchange within the group forwarding operation based on the message properties towards destination node/destinations or message forwarding between different groups. Mobile node with higher interest properties would aid to forward message among the group.

In line with this preliminaries inspiration an Interest Relationship Based Forwarding Scheme (IntReF)is proposed. The proposed scheme has three major components. Each component has its own set of task to operate and provide relevant information to the next components on top of it. Social graph component takes an input social information (interest and contact duration) to form the social graph, and formulate various social metrics and relation. In this component formation of social graph and formulation of various social strength guides by human behavior aggregation algorithm through social network analysis technique is carried out.

Then interest relationship component continues to formulate information about nodes group by taking in to account nodes interest relationship dynamics which is inspired by the existence of fat-tailed probability distributions, with this each node selects and represent itself by the interest which has this features. Knowledge about this interest conveys to which group a given node belongs to. Moreover, this component takes computed social strength value as an input from the first component, and combined another social metrics interest density to formulate a forwarding capability utility function that aids message forwarding among the groups.

IntReF third component is forwarding strategy. This is the component which guides the actual data forwarding operations. To achieve this operations, sub component decision engine maintain forwarding information about encountered mobile node and actual forwarding decision performed. And sub component decider part provides decision information based on the request up on receiving from the engine. More importantly the interaction between decision engine and decider part connected via forwarding algorithm component with its flow through the pseudocode shown. And nodes exchanges message within the group by selecting a node which has a higher time evolving contact to interest social strength measurement. And message exchange between nodes in different groups by selecting nodes which has a higher forwarding capability value during nodes up on encounters.

Chapter 5

Demonstration of Effectiveness

In this chapter, we describe the process of evaluation and analysis of results. Real human connection Mobility Trace scenario considered: to study how our proposed work cope with mobility level, from high mobile to near static nodes and to observe the impact of network load.

5.1 Performance Evaluations

This section starts by presenting the evaluation methodology and experiment settings, then followed by the discussions and results in the considered scenarios.

5.1.1 Simulation Setup

Using the ONE (Opportunistic Network Environment simulator) simulating software [20]. The ONE simulator is specifically designed for DTN routing, and it allows creating scenarios upon different synthetic movement models and helps to configure real human connection traces. we implement and evaluate performance of the proposed work, IntReF. we use following performance evaluation metric, delivery ratio, average latency and average costs with regards to node buffer space, time to live, and simulation duration time parameters and preliminary report depict in data figures. we compare the result of our work to the benchmark routing and forwarding algorithms Epidemic [29], and selected social ware scheme Bubblerap [32].

Performance Evaluation Metrics

In each experiments, we compare the performance of protocols based on the following metrics:

- **message delivery ratio:** It is the ratio of successfully delivered messages to the total number of unique messages created in a given period.
- **overhead ratio:** It is the ratio of relayed messages (delivered messages excluded) and delivered messages, reflecting the ratio of message replicas propagated into the network.

- **average latency:** It is the average time between the time a message is generated and the time it is delivered successfully.

Simulation parameters

Experimental setting and scenario most common aspects that are taken into account by each of the proposals for performance assessment, such as the number of nodes, meeting time (i.e., contact time) and time between such meetings (i.e., inter-meeting time), area size, message size, network load (i.e., number of generated messages), message TTL(time units), size of node buffer, and mobility model. As shown in table 5.1, simulation experiment parameters configurations and its default value settings are configured.

Table 5.1: Simulation parameters

simulation parameters	values	Default values
Simulation time	200000 s $\sim$ 1036800 s	1036800 s
number of runs	10	-
warm up	1000 s	-
Network area	100*100 m^2	-
Waiting time destination	0 $\sim$ 120 s	Random
Event interval	25 $\sim$ 35 s	Random
message size	500 $\sim$ 1024 KB	Random
message TTL	600 min $\sim$ 3600min	1440 min
node buffer	2M $\sim$ 20M	2M
number of nodes	25 $\sim$ 50	36

- **Simulation time:** It is parameter that configured to run the simulation experiment how many simulated seconds to simulate.
- **warm up:** It is the simulated seconds from the start, and Length of the warm up period. During the warm up the period new events are discard.
- **Waiting time destination:** Defines how long nodes should stay in the same place after reaching the destination of the current path. A new random value within the interval is used on every stop. Default value is 0,0. And minimum and maximum of the wait time interval (seconds).
- **Network area:** It is the size of the simulation world in meters (width, height). It is the terrain size where node are bound to move freely.

- **message TTL:** Time To Live (simulated minutes) of the messages created by this host. Nodes check every one minute whether some of their messages' TTLs have expired and drop such messages.
- **node buffer:** Size of the nodes' message buffer (bytes). When the buffer is full, node can't accept any more messages unless it drops some old messages from the buffer.
- **number of node:** It is the number of nodes that participate in the simulation runs.

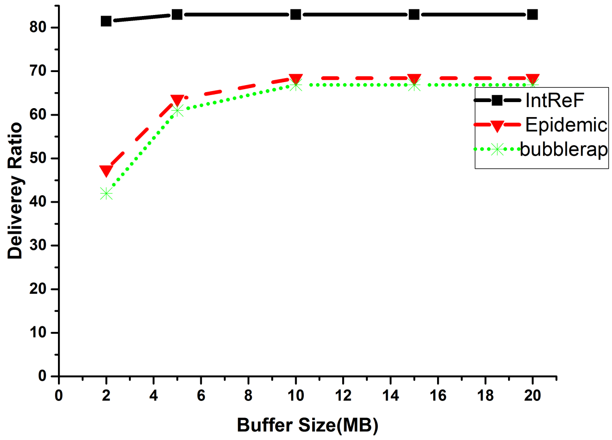
5.1.2 Dataset

The data set used and applied in our work is Cambridge [54], available on the web site of CRAWDAD(from http://crawdad.org/cambridge/haggle/20060131). This data set contains 36 individual in daily activity with their mobile device. Besides the contacts between mobile devices with connection action, its status (up or down), and contact duration among each encounter devices for number of days. Therefore, for this simulation experiments, we consider this data set scenario that represent development region network application scenario, which reflects sparse network density. Since the characteristic of applied data set has 26.38 average node degrees based on the analysis tool Gephi v0.8.2 [55], which indicates contacts between devices is sparse.

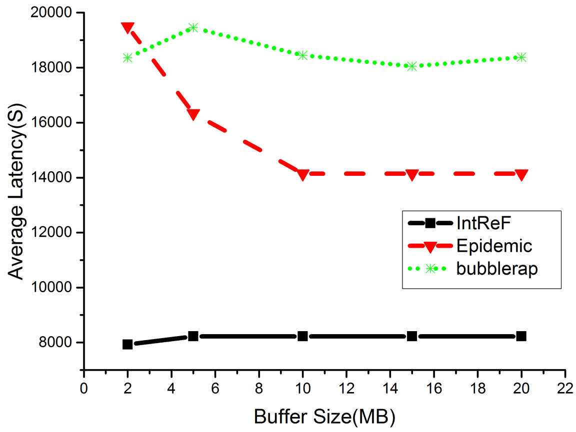
5.1.3 Analysis of Results

We evaluate the performance of the protocols over different buffer sizes, message TTL (Time To Leave), and simulation duration periods. We set and configure the minimum value of node buffer size (2MB), maximum simulation duration and a considerable message time to leave as default value to represents a realistic consideration and to observe the impact of multiple aspects of routing metrics constraint on forwarding performance.We ran 10 times for each experiment, and the results given represent the average values.

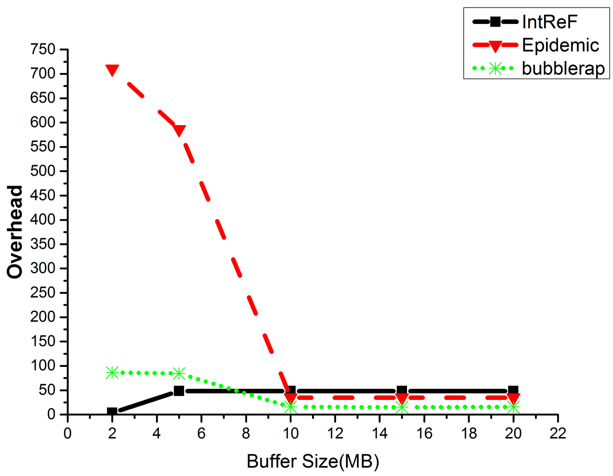
The preliminary results of simulation experiments of our work in comparison to the two schemes over different buffer sizes are shown in Figs.5.1.The figure contains three sub figures(a)(c)showing comparisons of our forwarding scheme(IntReF) in to the two schemes(Epidemic and bubblerap) for the message delivery ratio, average latency, and overhead,respectively. This performance data figure is obtained based on the real human connection trace, Cambridge data set and experimental setting, as explained in sections A and B.



(a) Delivery ratio



(b) Average Latency



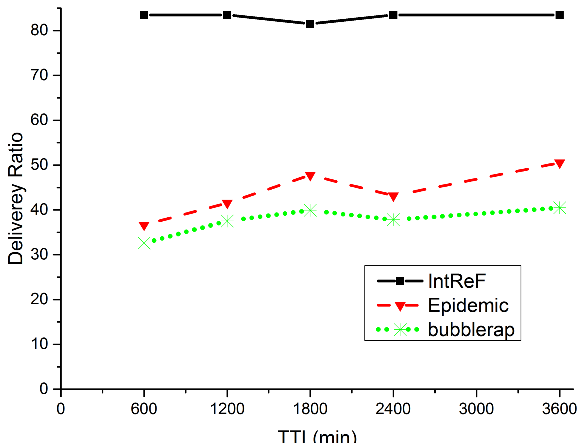
(c) Overhead Ratio

Figure 5.1: Performance with respect to buffer size

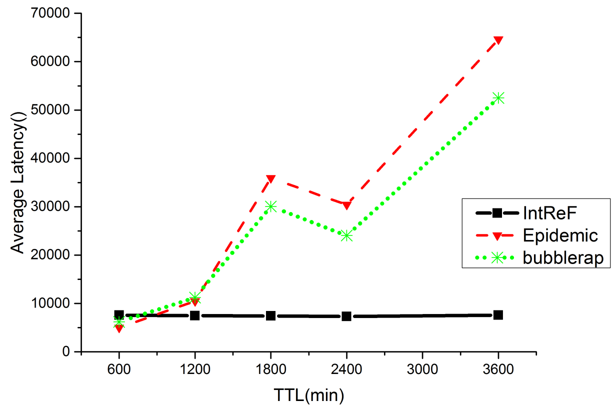
As depicted in Figs.5.1, the comparison of the forwarding performance of the three algorithms over different buffer sizes. During the size of buffer is risen, both Epidemic and bubblerap show an increase in delivery ratio, and a significant average latency increase value which is not stable, and decreasing trend for overhead. On the other hand, IntReF shows a stable record both in delivery ratio and average latency and as for the overhead ratio there is an increasing trend at the beginning and later become stable. Comparing IntReF, against both Epidemic and bubblerap in all parameters, it has a statistically significant performance improvement, except in regards to overhead with some value of buffer size. For instance, when the default value of buffer size become 2MB, as shown in fig.5.1(a), IntReF, forwards 83.72% messages where as the delivery ratio of Epidemic and bubblerap, 47.43% and 41.96% respectively. In a similar fashion, IntReF, records, a better average latency (faster, see Fig.5.1(b)) compared to Epidemic and bubblerap. But Epidemic gets faster than bubblerap, when the value of buffer size is raised particularly, at 5MB and more buffer size. In regards to the worst performance of bubblerap in average latency, since computation of community formation and detection take more time, due to the reason that dynamic nature of the network affects the most.

As depicted in fig.5.1(c), the performance of social aware schemes bubblerap and IntReF towards overhead shows a significant improvement compared to Epidemic, a flooding based scheme, due to the reason that the former replicates the message based on some kind of heuristic rather than replicates the message to every encounter nodes in the network as Epidemic did. Specifically, up to the value of buffer size reaches to 8MB, IntReF remains overhead below 50 and bubblerap remains more than 50. However, Epidemic records the worst performance at that point, more than 300. Although bubblerap and Epidemic over take IntReF after 10MB, our scheme, IntReF, take in to account, information that nodes maintains and exchange (i.e. interest information) is small enough and nodes process limited information, which saves network resources. We believe that this characteristics of our scheme helps to have a steady performance results.

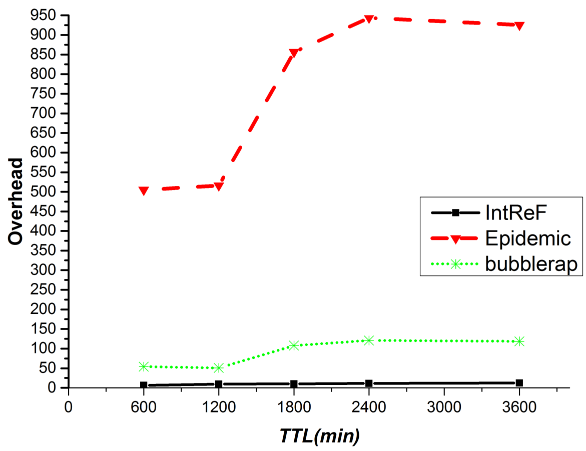
The preliminary results of simulation experiments of our work in comparison to the two schemes over different Message TTL are shown in Figs.5.2. The figure contains three sub figures (a)(c) showing comparisons of our forwarding scheme (IntReF) in comparison to the two schemes(Epidemic and bubblerap) for the message delivery ratio, average latency, and overhead, respectively. Evaluating the schemes over different values of message TTL have two fold purpose, in one hand hand, it helps to assess forwarding performance of the comparison schemes, on the other hand, finding a suitable message TTL with an acceptable performance results is necessary depending on the type of application scenario where the schemes being used.



(a) Delivery ratio



(b) Average Latency



(c) Overhead Ratio

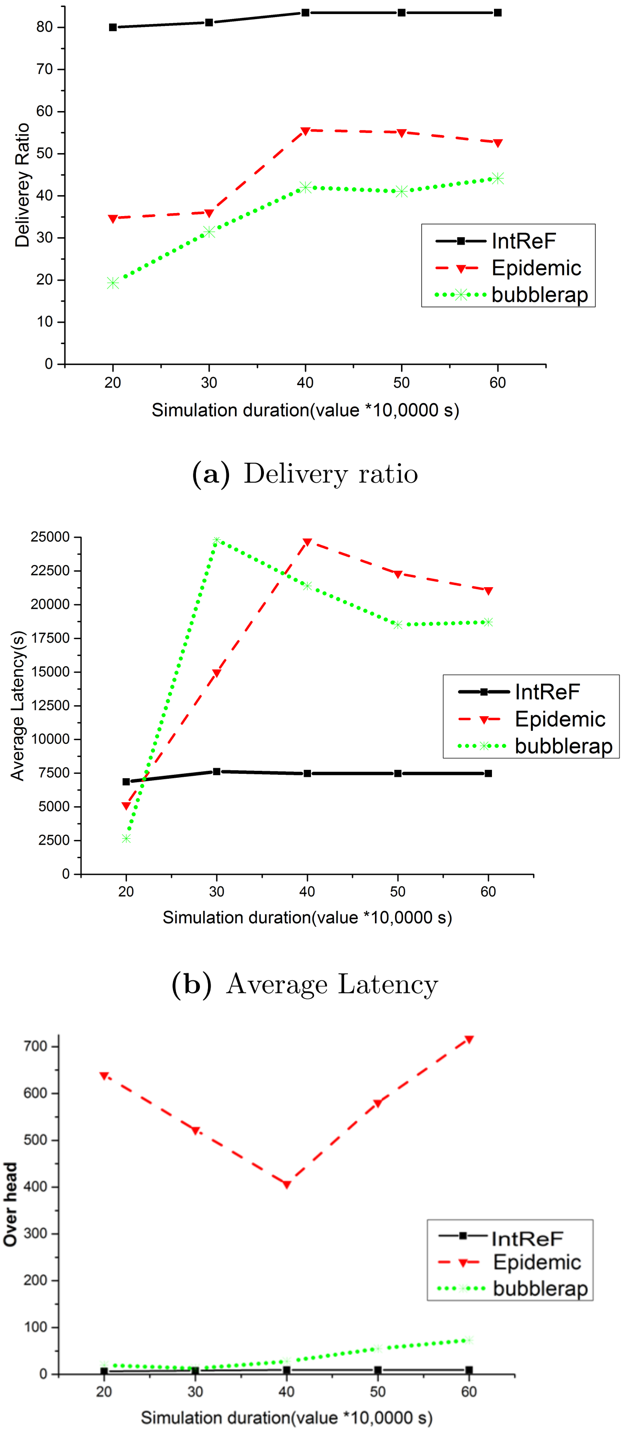
Figure 5.2: Performance with respect to message TTL

Based on the performance result shown in the figure, the performances of the three schemes have different performance characteristics over different values of message TTL. Both schemes have a stable performance trend towards delivery ratio with some Statistically significant performance variations among them. However, in regards to average latency both Epidemic and bubblerap have unstable performance. In the case of overhead ratio except Epidemic, IntReF and bubblerap have a stable performance with some Statistically significant performance variations among each other.

As the value of message TTL increases, more messages can longer stay in the network and this leads to increase the delivery ratio, however, in this experiment there is a default 2MB buffer size limitation is configured, and there is also increases message replication over the simulation duration time, therefore, this leads to buffer exhaustion, in result more message discarded, and delivery ratio affected. Due to this, there is a significant delivery ratio performance difference among IntReF against with Epidemic and bubblerap (see Fig. 5.2(a)).

When the value of message TTL less than or equals to 1200 min, both schemes have comparatively showed less sensitivity towards average latency, on the other hand, both the performance of Epidemic and bubblerap towards average latency increased as the value of message TTL grater than 1200 min (see Fig. 5.2(b)).

In the case of overhead ratio, the performance of social aware schemes bubblerap and IntReF have shown a significant improvement compared to Epidemic, a flooding based scheme, due to the reason that the former replicates the message based on some kind of heuristic rather than replicates the message to every encounter nodes in the network as Epidemic did. As the value of message TTL increases, more message stays longer time in the network in addition to the message replicated over the simulation duration, in effect, it increases the number of message copy exists in the network (see Fig. 5.2(c)). The overall overhead ratio of IntReF has a significant improvement compared to bubblerap and Epidemic, moreover, Epidemic has the worst overhead ratio performance result due to it unlimited message replication nature.



(a) Delivery ratio

(b) Average Latency

(c) Overhead Ratio

Figure 5.3: Performance with respect to simulation duration time

And similarly the performance results of the three schemes over different simulation duration time in respect to delivery ratio, average latency, and overhead are depicted in Figs. 5.3. Concerning to the value of the simulation duration time, we take in to account that more than half of the default simulation duration time (i.e., the maximum value) configuration which is used for buffer size and TTL performance evaluation setting.

According to the figures shown, as the value of the simulation duration rises, so does delivery ratio increases for all the evaluated schemes with a significant performance variation among them. However, in regards to average latency and overhead, except IntReF, Epidemic and bubblerap are more sensitive and affected most. Epidemic has the worst performance results towards overhead and bubblerap keep rising on average latency up to some value of simulation duration (see Fig. 5.3(b)), this is due to the reason that the scheme need more time to have a stable view of the community in the network.

Based on the performance result gained, IntReF achieves a better forwarding performance result compared to bubblerap and Epidemic with higher delivery ratio, less average latency. And less overhead ratio up to certain value of buffer size, and later the compared schemes outperforms with a minimal performance variations which take as a drawback.

IntReF attains this forwarding performance results, due to the reason that its design consideration, taking in to account interest relationships, various level of social strength combine with content type among nodes and forwarding capability metric to streamline forwarding decision in order to select potential forwarders during message exchange among nodes in the network. We believe that this characteristics of our scheme helps to have a steady performance results.

5.2 Summary

Due to its cost effective economic feasibility and flexibility to fine tune a range of application specific setting and parameters in model validation and performance evaluation in wireless network, simulation based techniques are a suitable approach to deploy the network and streamline performance evaluation and model validation. In line with this approach the ONE simulator tool is selected for deployment, evaluation and validation of the proposed scheme IntReF with its suitability and design considering in regards to Delay Tolerance Social Network paradigm.

The performance evaluation process follows those steps: Based on the performance evaluation metrics (delivery ratio, average latency and overhead ratio) varying different values of node buffer space, time to live, and simulation duration with a real human connection trace, IntReF performance results documented with multiple runs for statistical convergence. For comparison and evaluation of the performance result of IntReF, a selected benchmark forwarding scheme Bubblerap and Epidemic performance results documented with the same configuration and scenario as IntReF runs.

Concerning to analysis of results which depicted in data figures, IntReF performance results in relation to buffer size with deliver ratio and average latency shows a statistical significant improvement against Bubblerap and Epidemic. However, IntReF performance results towards overhead ratio cost more than Bubblerap and Epidemic. In regards to Message TTL which depicted in date figures, IntReF performance results in terms of delivery ratio, average latency and overhead ratio; it achieves a higher delivery, a reduced latency and less cost than Bubblerap and Epidemic.

In relation to simulation duration which depicted in data figures, IntReF performance results in terms of delivery ratio, average latency and overhead ratio; it achieves a higher delivery, a reduced latency and less cost than Bubblerap and Epidemic.

In conclusion based on the performance results gained, IntReF achieves a better forwarding performance in terms of delivery ratio, average latency, as result compared to bubblerap and Epidemic. The simulation results demonstrate that IntReF outperforms bubblerap and Epidemic with higher message delivery ratio and maintains a reduced average latency. With some exception in the case of buffer size parameters, the performance of IntReF towards overhead ratio shows a steady result.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this thesis, IntReF (Interest Relationship Based Forwarding) Scheme in Delay Tolerant Social Networks is proposed. The scheme has three components. Social graph component takes an input social information (interest and contact duration) to form the social graph, and formulate various social metrics and relation. In this component formation of social graph and formulation of various social strength guides by human behavior aggregation algorithm through social network analysis technique is carried out. Then interest relationship component continues to formulate information about nodes group by taking in to account nodes interest relationship dynamics which is inspired by the existence of fat-tailed probability distributions, with this each node selects and represent itself by the interest which has this features. Knowledge about this interest conveys to which group a given node belongs to. Moreover, this component takes computed social strength value as an input from the first component, and combined another social metrics interest density to formulate a forwarding capability utility function that aids message forwarding among the groups.

Using forwarding strategy component with its sub components the message forwarding operation proceeds where nodes exchanges message within the group by selecting a node which has a higher time evolving contact to interest social strength value. And message exchange between nodes in different groups by selecting nodes which has a higher forwarding capability value during nodes up on encounters.

In order to evaluate the performance of the proposed scheme, an extensive simulations experiment is conducted. Moreover, the performance is compared against two selected schemes Bubblerap and Epidemic. Based on the performance results gained, IntReF achieves a better forwarding performance in terms of delivery ratio, average latency, as result compared to bubblerap and Epidemic. The simulation results demonstrate that IntReF outperforms bubblerap and Epidemic with higher message delivery ratio and maintains a reduced average latency. Moreover, IntReF can perform around 24.66% delivery ratio in average in terms of varied value of buffer size, time to live (TTL) and simulation duration performance evaluations parameters. In regards to overhead ratio in terms of varied value of buffer size, IntReF achieves an acceptable stable performance result.

In this proposed work, we include content type towards message exchange, in addition to social metrics and relation in data forwarding operations. It is indeed, there is a significant forwarding performance improvement. So this notion of taking in to account content type during message exchange plays it part. And we showed our justification in a considered simulation experiment. However, it requires a theoretical analysis, to conclude content type is a better approach than social one in different considered scenarios.

6.2 Future Work

In regards to contact duration, one aspect we look in to further investigation is modeling the size of contact duration between encounter nodes, to better predict and streamline an effective message forwarding scheme, to aware and identify better bandwidth capacity between encounter nodes, to reduce recurring message drops, due to short contact duration or low bandwidth constraints in the heterogeneous network application scenario, where different kind of nodes characteristics included.

Added to that, there will be further investigation on the impact of selfish node on forwarding performance, Since the characteristics of some nodes in the network, non-cooperative in message exchange process (e.g., selfishness behaviors). In this work, Particularly, during message exchange among different interest social groups of nodes, there is a higher chance that nodes relay such kind of message could not interested, and unwilling to engage in the forwarding operation, compared to message exchange within group. So employing an effective incentive strategies that work with other components of forwarding scheme to stimulate this kinds of nodes to engage in and cooperative in the forwarding operation is vital, in order to have a complete forwarding scheme, which fuels the realization of DTNS in to real world application.

An other area which require further investigation is the impact of different mobility model[19] and various real human connection traces on forwarding performance, since different mobility model reflects various application scenario or use cases, and affects where nodes should move next. Therefore, it determine the underlying assumption to design an effective forwarding scheme to achieve a better network performance objectives.

Moreover, this notion of sending message M form source towards a destination, in the design consideration of forwarding scheme in delay tolerance social network, in some situations, needs to consider sending message towards group of destinations has a demand in different application scenario.

Therefore, this receiver vs. source oriented sending a message, requires further investigations, in various network applications to satisfy realistic applicability of this kind of network in real world application use cases. Moreover, it is better to consider as design paradigm integration between delay tolerance social networks, and information centric network [56][57] to benefits the opportunities of ICN in the developing region network characteristics towards message exchange and data sharing applications.

Bibliography

[1] M.Kende. "Internet Society Global Internet Report 2014". In: *Internet Society, Geneva, available at: www. internetsociety. org/sites/default/files/Global_Internet_Report_2014_0. pdf (accessed December 15, 2015)* (2014).

[2] Kara Sprague et al. "Offline and falling behind: Barriers to Internet adoption". In: *Technology, Media, and Telecom Practice Report. McKinsey & Company. Full report downloaded from: http://www. mckinsey. com/industries/high-tech/our-insights/offline-and-falling-behind-barriers-to-internet-adoption(accessed December 15, 2015)* (2014).

[3] Marco Conti, Mohan Kumar, et al. "Opportunities in opportunistic computing". In: *Computer* 43.1 (2010), pp. 42–50.

[4] Yue Cao and Zhili Sun. "Routing in Delay / Disruption Tolerant Networks : A Taxonomy , Survey and Challenges". In: 15.2 (2013), pp. 654–677.

[5] Behrouz Jedari and Feng Xia. "A survey on routing and data dissemination in opportunistic mobile social networks". In: *arXiv preprint arXiv:1311.0347* (2013).

[6] Design-related Issues et al. "A Survey of Social-Aware Routing Protocols in Delay Tolerant Networks :" in: (2013), pp. 1–23.

[7] Waldir Moreira and Paulo Mendes. "Survey on opportunistic routing for delay/disruption tolerant networks". In: (2011).

[8] Zhi-Dan Zhao et al. "Emergence of scaling in human-interest dynamics". In: *arXiv preprint arXiv:1307.7796* (2013).

[9] Feng Xia et al. "BEEINFO : Interest-Based Forwarding Using Artificial Bee Colony for Socially Aware Networking". In: 64.3 (2015), pp. 1188–1200.

[10] Feng Xia et al. "Data dissemination using interest-tree in socially aware networking". In: *Computer Networks* 91 (2015), pp. 495–507. ISSN: 1389-1286. DOI: 10.1016/j.comnet.2015.08.047. URL: http://dx.doi.org/10.1016/j.comnet.2015.08.047.

[11] Konglin Zhu, Wenzhong Li, and Xiaoming Fu. "SMART : A Social and Mobile Aware Routing Strategy for Disruption Tolerant Networks". In: (), pp. 1–12.

[12] By Dipankar Raychaudhuri and Narayan B Mandayam. "Frontiers of Wireless and Mobile Communications". In: *IEEE Communications Surveys & Tutorials* 100.4 (2012).

[13] James F Kurose and Keith W Ross. *Computer networking: a top-down approach*. Vol. 5. Addison-Wesley Reading, 2012, pp. 541–605.

[14] Miller McPherson, Lynn Smith-Lovin, and James M Cook. “Birds of a feather: Homophily in social networks”. In: *Annual review of sociology* 27.1 (2001), pp. 415–444.

[15] Linton C Freeman. “Centrality in social networks conceptual clarification”. In: *Social networks* 1.3 (1978), pp. 215–239.

[16] Elizabeth M Daly and Mads Haahr. “Social network analysis for information flow in disconnected delay-tolerant MANETs”. In: *IEEE Transactions on Mobile Computing* 8.5 (2009), pp. 606–621.

[17] Theus Hossmann, Thrasyvoulos Spyropoulos, and Franck Legendre. “A complex network analysis of human mobility”. In: *Computer communications workshops (INFOCOM WKSHPS), 2011 IEEE conference on*. IEEE. 2011, pp. 876–881.

[18] Mark EJ Newman. “Power laws, Pareto distributions and Zipf’s law”. In: *Contemporary physics* 46.5 (2005), pp. 323–351.

[19] Suvadip Batabyal and Parama Bhaumik. “Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey”. In: *IEEE Communications Surveys & Tutorials* 17.3 (2015), pp. 1679–1707.

[20] Ari Keränen, Jörg Ott, and Teemu Kärkkäinen. “The ONE Simulator for DTN Protocol Evaluation”. In: *SIMUTools ’09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*. Rome, Italy: ICST, 2009. ISBN: 978-963-9799-45-5.

[21] David B Johnson and David A Maltz. “Dynamic source routing in ad hoc wireless networks”. In: *Mobile computing* (1996), pp. 153–181.

[22] Xiaoyan Hong et al. “A group mobility model for ad hoc wireless networks”. In: *Proceedings of the 2nd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems*. ACM. 1999, pp. 53–60.

[23] Feng Xia et al. “Socially Aware Networking : A Survey”. In: 9.3 (2015), pp. 904–921.

[24] Forrest Warthman et al. “Delay-Tolerant Networks (DTNs)”. In: March (2003).

[25] K Fall. “A Delay Tolerant Network Architecture for Challenged Internets”. In: (2003). URL: `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.3.5643`.

[26] Zhensheng Zhang. “Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges”. In: *IEEE Communications Surveys & Tutorials* 8.1 (2006), pp. 24–37.

[27] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi S Raghavendra. “Single-copy routing in intermittently connected mobile networks”. In: *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*. IEEE. 2004, pp. 235–244.

[28] Sushant Jain, Kevin Fall, and Rabin Patra. *Routing in a delay tolerant network*. Vol. 34. 4. ACM, 2004.

[29] Amin Vahdat and David Becker. "Epidemic Routing for Partially-Connected Ad Hoc Networks". In: ().

[30] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi S Raghavendra. "Spray and wait: an efficient routing scheme for intermittently connected mobile networks". In: *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. ACM. 2005, pp. 252–259.

[31] Anders Lindgren and Avri Doria. "Probabilistic Routing in Intermittently Connected Networks". In: (), pp. 1–8.

[32] Pan Hui, Jon Crowcroft, and Eiko Yoneki. "in Delay Tolerant Networks BUBBLE Rap : Social-based Forwarding in Delay Tolerant Networks". In: November (2010).

[33] W.Moreira and P.Mendes. "Ad Hoc Networks Impact of human behavior on social opportunistic forwarding". In: *Ad Hoc Networks* 25 (2015), pp. 293–302. ISSN: 1570-8705. DOI: 10.1016/j.adhoc.2014.07.001. URL: http://dx.doi.org/10.1016/j.adhoc.2014.07.001.

[34] Maurice J Khabbaz, Chadi M Assi, and Wissam F Fawaz. "Disruption-tolerant networking: A comprehensive survey on recent developments and persisting challenges". In: *IEEE Communications Surveys & Tutorials* 14.2 (2012), pp. 607–640.

[35] Ying Zhu et al. "A survey of social-based routing in delay tolerant networks: Positive and negative social effects". In: *IEEE Communications Surveys & Tutorials* 15.1 (2013), pp. 387–401.

[36] Anders Lindgren and Pan Hui. "The quest for a killer app for opportunistic and delay tolerant networks". In: *Proceedings of the 4th ACM workshop on Challenged networks*. ACM. 2009, pp. 59–66.

[37] AG Voyiatzis. "The quest for a killer app for delay-tolerant networks (DTNs)". In: *Advances in Delay-tolerant Networks (DTNs): Architecture and Enhanced Performance* (2014), p. 251.

[38] Alex Pentland, Richard Fletcher, and Amir Hasson. "DakNet: Rethinking Connectivity in Developing Nations". In: *Computer* 37.1 (2004). ISSN: 00189162. DOI: 10.1109/MC.2004.1260729.

[39] S Guo et al. "Very Low-Cost Internet Access Using KioskNet". In: ().

[40] Narseo Vallina-rodriguez, Pan Hui, and Jon Crowcroft. "Has anyone seen my Goose ?" In: ().

[41] Russell J. Clark et al. "D-book: A Mobile Social Networking Application for Delay Tolerant Networks". In: *Proceedings of the Third ACM Workshop on Challenged Networks*. CHANTS '08. San Francisco, California, USA: ACM, 2008, pp. 113–116. ISBN: 978-1-60558-186-6. DOI: 10.1145/1409985.1410007. URL: http://doi.acm.org/10.1145/1409985.1410007.

[42] Philo Juang et al. "Energy-efficient computing for wildlife tracking". In: *ACM SIGOPS Operating Systems Review* 36.5 (2002), p. 96. ISSN: 01635980. DOI: `10.1145/635508.605408`.

[43] Tara Small and Zygmunt J Haas. "The Shared Wireless Infostation Model - A New Ad Hoc Networking Paradigm (or Where there is a Whale , there is a Way)". In: (2003).

[44] Piet Buys et al. "Determinants of a Digital Divide in Sub-Saharan Africa : A Spatial Econometric Analysis of Cell Phone Coverage". In: *World Development* 37.9 (2009), pp. 1494–1505. ISSN: 0305-750X. DOI: `10.1016/j.worlddev.2009.01.011`. URL: `http://dx.doi.org/10.1016/j.worlddev.2009.01.011`.

[45] Elizabeth M Daly and Mads Haahr. "Social network analysis for routing in disconnected delay-tolerant manets". In: *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing.* ACM. 2007, pp. 32–40.

[46] B Jedari et al. "Using social network analysis (SNA) to design socially aware network solutions in delay-tolerant networks (DTNs)". In: *Advances in Delay-tolerant Networks (DTNs): Architecture and Enhanced Performance* (2014), p. 205.

[47] Feng Li and Jie Wu. "LocalCom: A community-based epidemic forwarding scheme in disruption-tolerant networks". In: *2009 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, SECON 2009* (2009). ISSN: 2155-5486. DOI: `10.1109/SAHCN.2009.5168942`.

[48] Pan Hui and Jon Crowcroft. "How Small Labels Create Big Improvements". In: April (2007), pp. 65–70. DOI: `10.1109/PERCOMW.2007.55`. URL: `http://dx.doi.org/10.1109/PERCOMW.2007.55http://doi.ieeecomputersociety.org/10.1109/PERCOMW.2007.55`.

[49] Alessandro Mei et al. "Social-Aware Stateless Forwarding in Pocket Switched Networks". In: (2011), pp. 251–255.

[50] Waldir Moreira, Paulo Mendes, and Susana Sargento. "Opportunistic routing based on daily routines". In: *World of wireless, mobile and multimedia networks (WoWMoM), 2012 IEEE international symposium on a.* IEEE. 2012, pp. 1–6.

[51] Radu-ioan Ciobanu et al. "Ad Hoc Networks Interest-awareness in data dissemination for opportunistic networks". In: *Ad Hoc Networks* 25 (2015), pp. 330–345. ISSN: 1570-8705. DOI: `10.1016/j.adhoc.2014.07.004`. URL: `http://dx.doi.org/10.1016/j.adhoc.2014.07.004`.

[52] W. Moreira and P. Mendes. "Dynamics of social-aware pervasive networks". In: *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops).* 2015, pp. 463–468. DOI: `10.1109/PERCOMW.2015.7134082`.

[53] Waldir Moreira and Paulo Mendes. "Social-aware Forwarding in Opportunistic Wireless Networks : Content Awareness or Obliviousness ?" In: (). arXiv: `arXiv:1407.8342v1`.

[54] James Scott et al. *CRAWDAD dataset cambridge/haggle (v. 2006-09-15)*. Downloaded from `http://crawdad.org/cambridge/haggle/20060915/imote`. traceset: imote. 2006. DOI: `10.15783/C73S3N`.

[55] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. "Gephi: An Open Source Software for Exploring and Manipulating Networks". In: *Third International AAAI Conference on Weblogs and Social Media* (2009), pp. 361–362. ISSN: 14753898. DOI: `10.1136/qshc.2004.010033`. URL: `http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154$\backslash$npapers2://publication/uuid/CCEBC82E-0D18-4FFC-91EC-6E4A7F1A1972`.

[56] B. Ahlgren et al. "A survey of information-centric networking". In: *IEEE Communications Magazine* 50.7 (2012), pp. 26–36. ISSN: 0163-6804. DOI: `10.1109/MCOM.2012.6231276`. URL: `http://ieeexplore.ieee.org/document/6231276/`.

[57] Gareth Tyson, John Bigham, and Eliane Bodanese. "Towards an information-centric delay-tolerant network". In: *Computer Communications Workshops (INFOCOM WKSHPS), 2013 IEEE Conference on*. IEEE. 2013, pp. 387–392.

Appendix A

Appendix One

A.1 Formulation of cumulative moving average

$$CMA_n = \frac{\chi_1 + ... + \chi_n}{n}$$
$$\chi_1 + ... + \chi_n = CMA_n \times n \quad (A.1)$$

and similarly for n + 1, it is seen that

$$\chi_{n+1} = \chi_1 + ... + \chi_{n+1} - \chi_1 + ... + \chi_n \quad (A.2)$$

Solving this equation for CMA_{n+1} results in:

$$CMA_{n+1} = \frac{\chi_{n+1} + n \times CMA_n}{n + 1} \quad (A.3)$$

For n-1, the equation become:

$$CMA_n = \frac{\chi_n + n - 1 \times CMA_{n-1}}{n} \quad (A.4)$$

Tabulated performance data

Some source code section

Sample source code section of a node is created for Simulation scenario with it s properties.

```
/**
     * Creates hosts for the scenario
     */
    protected void createHosts() {
        this.hosts = new ArrayList<DTNHost>();

        for (int i=1; i<=nrofGroups; i++) {
            List<NetworkInterface> mmNetInterfaces =
                new ArrayList<NetworkInterface>();
            Settings s = new Settings(GROUP_NS+i);
            s.setSecondaryNamespace(GROUP_NS);
            String gid = s.getSetting(GROUP_ID_S);

            /** DTNHost interest and modified for FrInt  */
            int nrofInt = s.getInt(NROF_INTERESTS);
            Map<String,Double> hostInterest = new HashMap<String,Double>();
            for (int a=1; a<=nrofInt; a++){
                String interest = s.getSetting(INTEREST+a);
                is object of  Setting class
                int c=a;
                double degreeInt  = s.getDouble(DEGINTEREST+c);
                is for degree to interest
                hostInterest.put(interest,degreeInt);
            }
            /** DTNHost interest for FrInt   */

            int nrofHosts = s.getInt(NROF_HOSTS_S);
            int nrofInterfaces = s.getInt(NROF_INTERF_S);
            int appCount;

            // creates prototypes of MessageRouter and MovementModel
            MovementModel mmProto =
                (MovementModel)s.createIntializedObject(MM_PACKAGE +
                        s.getSetting(MOVEMENT_MODEL_S));
            MessageRouter mRouterProto =
                (MessageRouter)s.createIntializedObject(ROUTING_PACKAGE +
                        s.getSetting(ROUTER_S));

            // checks that these values are positive (throws Error if not)
            ensurePositiveValue(nrofHosts, NROF_HOSTS_S);
            ensurePositiveValue(nrofInterfaces, NROF_INTERF_S);

            // setup interfaces
            for (int j=1;j<=nrofInterfaces;j++) {
                String Intname = s.getSetting(INTERFACENAME_S+j);
                Settings t = new Settings(Intname);
                NetworkInterface mmInterface =
                    (NetworkInterface)t.createIntializedObject(INTTYPE_PACKAGE +
                            t.getSetting(INTTYPE_S));
                mmInterface.setClisteners(connectionListeners);
                mmNetInterfaces.add(mmInterface);
            }

            // setup applications
            if (s.contains(APPCOUNT_S)) {
                appCount = s.getInt(APPCOUNT_S);
            } else {
```

Sample source code section of a new DTNhost is created with its properties.

```
/**
     * Creates a new DTNHost  for frInt
     * @param msgLs Message listeners
     * @param movLs Movement listeners
     * @param hInt List of interests,
     * @param groupId GroupID of this host
     * @param interf List of NetworkInterfaces for the class
     * @param comBus Module communication bus object
     * @param mmProto Prototype of the movement model of this host
     * @param mRouterProto Prototype of the message router of this host
     */
    public DTNHost(List<MessageListener> msgLs,
            List<MovementListener> movLs, Map<String,Double> hInt,
            String groupId, List<NetworkInterface> interf,
            ModuleCommunicationBus comBus,
            MovementModel mmProto, MessageRouter mRouterProto) {
        this.comBus = comBus;
        this.location = new Coord(0,0);
        this.address = getNextAddress();
        this.name = groupId+address;
        this.net = new ArrayList<NetworkInterface>();

        for (NetworkInterface i : interf) {
            NetworkInterface ni = i.replicate();
            ni.setHost(this);
            net.add(ni);
        }

        // For FrInt

        this.hostInterests = new HashMap<String,Double>();
            hostInterests.putAll(hInt);

        // TODO - think about the names of the interfaces and the nodes
        //this.name = groupId + ((NetworkInterface)net.get(1)).getAddress();

        this.msgListeners = msgLs;
        this.movListeners = movLs;

        // create instances by replicating the prototypes
        this.movement = mmProto.replicate();
        this.movement.setComBus(comBus);
        setRouter(mRouterProto.replicate());

        this.location = movement.getInitialLocation();

        this.nextTimeToMove = movement.nextPathAvailable();
        this.path = null;

        if (movLs != null) { // inform movement listeners about the location
            for (MovementListener l : movLs) {
                l.initialLocation(this, this.location);
            }
        }
    }
```

-1-

Sample source code section of average density towards different interest via cumulative moving average formulation.

```
/* Called to update the average interest density of this host had
         *  with the different interests of encountered node  via via cumulative moving
         average formulation**/

        private void updateAverageDensityToInterests() {

        long currentday=SlotTimeCheck.getDay();
        int currentslot = SlotTimeCheck.getcurrentslot();
        Map<DTNHost,Double> currentAverageDensityToInterests=averageDensityToInterests.get(
        currentslot);
        Set hostsinden = InterestsDensityTo.keySet();
        Iterator<DTNHost> denTHostIterator= hostsinden.iterator();
        while(denTHostIterator.hasNext()){
            DTNHost currenthost= denTHostIterator.next();
            double oldADen=0;
            if(currentAverageDensityToInterests.get(currenthost)==null){
                oldADen=0;
            }
            else{
                oldADen=currentAverageDensityToInterests.get(currenthost);
            }
            double newADen = (InterestsDensityTo.get(currenthost)+(currentday-1)*oldADen);
            currentAverageDensityToInterests.put(currenthost,newADen);


        }
        Set<DTNHost> s = currentAverageDensityToInterests.keySet();
        Iterator<DTNHost> iter=s.iterator();
        double newvalue=0.0;   // the moving average density over a given daily sample.
        while(iter.hasNext()){
            DTNHost dtnhost = iter.next();
            if(!hostsinden.contains(dtnhost)){
                newvalue= (currentday-1)*currentAverageDensityToInterests.get(dtnhost)/
                currentday;
            }
            else{
            newvalue= currentAverageDensityToInterests.get(dtnhost)/currentday;
            }
            currentAverageDensityToInterests.put(dtnhost,newvalue);
            this.densityToDiffInterests=currentAverageDensityToInterests;  // assign in to,
            which holds average density to different interest with encounter
        }
        }
```

Portion of used Data set

Portion of used real human connection trace sample. It has the following formats, time: actionId: hostId: host2Id: and the status of connection between nodes either up or down.

time	actionId	hostId	host2Id	up or down
0.00	CONN	14	8	up
11.00	CONN	14	9	down
4.00	CONN	14	5	up
11.00	CONN	14	5	down
76.00	CONN	8	4	up
76.00	CONN	8	18	up
76.00	CONN	8	29	up
86.00	CONN	8	5	down
86.00	CONN	8	17	down
216.00	CONN	9	22	down

Sample file of message load

Sample section of file containing message load.

time	actionId	Message ID	hostId	host2Id	Size
0	C	M1	0	1	74044
0	C	M58	0	2	34196
0	C	M59	0	2	43658
0	C	M60	0	2	78824
0	C	M2	0	1	59129
0	C	M101	0	3	50247
691200	C	M1156	0	34	74044
691200	C	M1162	0	34	43034
691200	C	M1156	0	34	74044
691200	C	M1163	0	34	22936

Miscellaneous

Declaration

I, the undersigned, declare that this thesis work is my original work, has not been presented for a degree in this or any other universities, and all sources of materials used for the thesis work have been duly acknowledged.

Name: Andualem Chekol

Signature: ____________

Place: Jimma

Date: _______________

This thesis has been submitted for examination with my approval as a university advisor/co- advisor.

Name: Dr. Ahmedin Mohammed

Signature: ____________

Place: Addis Ababa

Date: _______________

Name: Mr. Kebebew Ababu

Signature: ____________

Place: Jimma

Date: _______________