Jimma University

Jimma Institute of Technology

School of Computing

Computer Networking MSc Program

Combination of Community Structure and Mechanism to Maximize Data Availability in Adhoc Social Network *(ASNET)*

By

Negasa Berhanu Fite

A Thesis Submitted to the School of Graduate Studies of Jimma University in Partial fulfilment of the Requirements for the Degree of Masters of Science in Computer Networking

May, 2017 G.C

Jimma, Ethiopia

Jimma University

Jimma Institute of Technology

School of Computing

Computer Networking MSc Program

Combination of Community Structure and Mechanism to Maximize Data Availability in Adhoc Social Network *(ASNET)*

By

Negasa Berhanu Fite

A Thesis Submitted to the School of Graduate Studies of Jimma University in Partial fulfilment of the Requirements for the Degree of Masters of Science in Computer Networking

May, 2017 G.C

Jimma, Ethiopia

Contact Information

Done By:

Mr. Negasa Berhanu
Phone Number: +251917847692
Email: negasab@gmail.com or negasaberhanu@yahoo.com

Principal Advisor
Dr. Henock Mulgeta (PhD)
Addis Abeba University
Faculty of Informatics
Addis Abeba, Ethiopia

Co-Advisor

Mr. Kebebew Abebu
Jimma University
School of Computing
Computer Networking and Operating System Chair
Jimma, Ethiopia

Declaration

I, the undersigned, declare that this thesis entitled Combination of Community Structure and Mechanism to Maximize Data Availability in Adhoc Social Network *(ASNET)* is my original work and has not been presented for a degree in this or any other universities, and all sources of references used for the thesis work have been duly acknowledged.

Name: Negasa Berhanu Fite

Signature: ______________ Date: _________________

This thesis has been submitted for examination with my approval as a University advisor.

Advisor Name: Dr. Henock Mulgeta (PhD)

Signature: ______________ Date: _________________

This thesis has been submitted for examination with my approval as a University Co-advisor.

Co-Advisor: Mr. Kebebew Abebu (MSc)

Signature: ______________ Date: _________________

DEDICATION

I dedicate this thesis for my Mother and my lover

ACKNOWLEDGMENT

First and above all, I praise the almighty God, for providing me this opportunity and granting me the capability to complete my work and who gave me patience in every endeavour of my life. Then, I would like to express my genuine gratitude to my Advisor Dr. Henock Mulgeta and my deepest appreciation goes to my Co-advisor Mr. Kebebew Abebu for their encouragement, advice, supervision, patience, and moral support throughout my thesis work. I learned a lot of bundled experience from their critical observation regarding this work and their incredible encouragement and support made me a better student in spite of myself. My sincere thank also goes to Dr. Ahmedin Ahmed for he teach me the idea of research and new innovation and to other Instructors and Classmates those share their knowledge with me and gave me comment in my work. Finally, I thank my mother, brother, my lover and other families for their unconditional support and encouragement throughout the years.

I also wish to extend my thanks to Computing School and all the staff members of this school for their cheering direction and for providing a good quality education which has enabled me to grow and develop professionally.

Table of Contents

LIST OF FIGURES

List of Tables

List of Algorithms

List of Equations

LIST OF ACRONYMS

AODV	Adhoc on Demand Distance Vector
ASNET	Adhoc Social Network
C	Community
CBR	Constant bit rate
cH	Community Head
cH-id	Community Head Identification
ComPAS	Community Partition Aware Replica Allocation
Did	Data Identification
e	Energy of node
FTP	File Transport Protocol
IEEE	Institute of Electrical Electronics Engineering
M	Mobile Node
m	Mobility
MANET	Mobile Adhoc Network
NcH	Neighbour Community Head
NS2	Network Simulation 2
RepCoH	Replica on Community Head
RPGM	Random Point Group Mobility
RREQ	Route Request (RREQ)
RT	Replica Table
RWM	Random Waypoint Mobility
s	Storage capacity of node
SAN	Social Aware Network
SEAD	Secure Efficient Adhoc Distance Vector
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

ABSTRACT

ASNET is collection of social network and MANET which is infrastructure less self configuring capability of mobile nodes connected by wireless network. In these networks due to mobility of node there is much change in network topology and this causes frequent network division. So, data sharing in network play vital role. One probable solution to avoid this problem is to take up replication techniques which increases data availability and decrease query response delay.

The replication process duplicates and preserves the consistency of multiple copies of objects in different sites so that each client node can visit a local copy of an object instead of remote ones. ASNET data replica allocation helps to avoid packet losses in case of unpredictable network partition and etc.

In this paper, the community structure based replica allocation called RepCoH (Replica on the Community Head) is proposed which address the limitation of other protocols. It collaborate mobile node fully in terms of sharing their memory space by combining the structure of the community to minimize query delay when the node move from one community to other and design the mechanism that maximize data availability and protect data loss. The protocol considers moving behaviour of nodes with properties such as storage capacity and power remain while allocating data on community head to achieve better load balancing.

The extensive simulations are performed to analyze and compare the behaviour of RepCoH with ComPAS and No Replica framework for different network scenarios generated by varying parameters such as network size and time of simulation. The performance evaluation of RepCoH through simulation indicated the proposed protocol increases data availability, and achieves better packet load balancing in comparison to ComPAS and No Replication scenarios.

Keywords: ***ASNET, Data Replication, RepCoH, Community Head***

CHAPTER ONE

INTRODUCTION

1.1 Background of the study

Wireless and Mobile network are practised an unprecedented development during the past decades. The advance of telecom, hardware's and software technologies provide wireless networking and mobile devices common place and powerful. The progress of both hardware and software technologies made possible new alternative infrastructure-less network for wireless communication in which devices form a self creating, self organizing and self administering wireless network. In the early 1990s ad hoc networks entered a new phase of growth due to popularity of notebook computers with communication equipments based on radio frequency and infrared grew [1]. An idea of the infrastructure less collection of mobile nodes was proposed, and the IEEE 802.11 subcommittee adopted the word 'ad hoc networks'. Adhoc networks are a core in the evolution of wireless networks. It is composed of mobile nodes, which communicate over wireless links without any central control. Adhoc wireless networks inherit the traditional problems of wireless and mobile communications, such as optimization of bandwidth, control of power and transmission quality enhancement[2].

The mobility patterns of mobile devices strongly depend on users movements, which are closely related to their social relationships and behaviours. Consequently, today's mobile networks are becoming increasingly human centric. This leads to the emergence of a new field we call Socially Aware Networking (SAN) which contains social properties such as social graph, community, centrality, similarities, tie strength and human mobility pattern [3]. The growing of social aware network give new social behaviour based network called Adhoc Social Network (ASNET) which work on above Mobile Adhoc Network (MANET) and greatly different from current internet based social network. Ad-hoc Social network is combination of mobile Ad-hoc network which is self-configured infrastructure less network that communicate over bandwidth constrained wireless links and social network that uses social relationships to determine the node communication [3].

ASNET is the blend of mobile ad-hoc network and social network that is a social structure made up of individuals or organizations called nodes, which are connected by one or more specific types of interdependency such as friendship, common interest, and financial exchange, relationships of beliefs, knowledge or prestige and etc. ASNET is more meaningful as users are interested to know which of other users have similar interests and are within communication range so that they can either chat or personally talk on current interests. In ASNET mobile nodes that are within the same radio range communicate directly via wireless links, while those that are far apart rely on other nodes to relay messages as routers. Since mobile hosts move freely, disconnection occur frequently and this causes frequent network division that give low data availability in ASNET than in conventional fixed networks.

Therefore increasing data availability at the point of community partitioning is extremely important issue. The solution to maximize data availability in ad-hoc social network is combining community structure and replicating data items in the community locally or that are not the owners of the original data.

The management of data in ASNET is essential due to characteristics of nodes such as mobility and resource constraints [4]. The possible solution for the data management is replication of data on different node. However, implementing replica in ASNET is required for two major issues. At the first ASNET is resource constrained environment i.e. the nodes have poor battery, CPU, storage capabilities, and the available bandwidth on the network is limited. Therefore, a replica allocation system for ASNETs should strive to consume as little resources as possible. Another major issue of implementing replica in ASNET is that the dynamicity of the environment makes making replication decisions a very strong problem. The constantly evolving topology of a ASNET makes it impossible to compute an optimal replica placement, and the unpredictable distribution of requests complicates the selection of the best data items to replicate [5].

The replica helps to avoid data loses in case of unpredictable community or network partition and reduces number of hops when data is transmitted from source to destination. The replica of data enhances the availability of data by making copies of data items and it allows better data sharing. It is used to improve data availability in distributed systems and it is applied to ASNET. The data replication in ASNET is the process of sharing information between the mobile nodes and ensures the consistency between the resources that aims to achieve improved accessibility, shorter response time, fault tolerance and to distribute the load of processing of the requests on several nodes and to avoid load of the routes of communication to the unique node. ASNET [6] nodes communicate directly with each other and relay data packets just like routers in conventional wired networks. If the distance from source to destination node is beyond the communication range of the node, data transmission can still be implemented by forwarding through other nodes which exist between them.

The paper propose combination of community structure which has significant impact on the node's mobility patterns and which cooperates the node with each other, when they share their memory space for replicating data to the neighbouring nodes and neighbouring community. This is required to access data if partition is occurred in the network, if the replica holder in the communities fails and, to improve the performance of data availability and accessibility in ad-hoc social network.

1.2 Research Motivation

The recent advances in radio communication and computer technologies led to increasing interest in Adhoc networks. Adhoc social networks are, however, far from maturity and there are tremendous difficulties in making ASNET into real working systems. Many previous research papers in Adhoc network focused on routing protocols of MANETs which allows the nodes to efficiently discover routes between communication peers. Although routing is important in ASNET, we believe providing data services between the nodes are also important thing. Unfortunately the most previous research work done on data management focus on MANET data replication without adding the social behaviour and social interest of the nodes and many of the previous papers replica systems are based on the accessibility of data and random replica without considering the nodes moving behaviour, centrality of the nodes, the remaining power and the storage capacity to hold the replicated data.

Since mobile nodes in ASNET moves freely and organize themselves without restriction of any infrastructure and every nodes play the role of router and communicates with other mobile nodes, combination of the structure of the community to manage data cammunication between the nodes are required. Therefore we develop community based data replication system which is called Replica on Community Head (RepCoH) to gets benefits of increased data availablity,decreased data accesed time and balanced workload. The RepCoh replica system address the problem of, data loss when network partition or nodes fail because of power , increase data consistency since replica doesn't distribute on each node and community head hold replica of data, decrease route discovery time because of community head manage replica.

1.3 Statement of Problems

The key issues to be considered in Adhoc social network data management are how to replicate data and manage data on community head of the nodes with maximum data accessibility. Issues related to replica of data are the following:

The first is Power; in traditional mobile networks only the power needs of clients are considered. But in ASNET every mobile devices are battery powered and the power of community head which hold and allocate replicated data is perhaps more important since it control data of every client nodes. There is a problem of power consumption since mobile nodes in ASNET are different battery power availability, therefore if the node with less power is replicated with many frequently accessed data items; it gets drained and cannot provide services any more. Thus in adhoc social network replication algorithm should replicate data in the nodes that has sufficient power by periodically checking the remaining power of the node.

The second is Resource Availability; a node in ASNET should supply mechanisms for proficient use of processing, memory and communication resources, while maintaining low power consumption. A node should bring about its basic operations without resources exhaustion. Since nodes participating in Adhoc Social Network are portable devices, their memory capacity is limited therefore the amount of replicas or replica management information that can be held by each mobile node will be limited. Hence, the replication mechanism should ensure that the replicas of different data items are distributed among the nodes those have large free storage capacity, to maximize the space utilization in nodes before sending a replica to the node, the algorithm has to check whether a node has sufficient memory capacity to hold the replica.

The third is Mobility of the nodes; due to dynamic nature of Adhoc networks and since nodes those have similar moving behaviour or nodes have strong tie moves together, it exhibits frequent and unpredictable changes of the topology. ASNETs therefore should able to adapt traffic and propagation conditions to mobility patterns of nodes. The node mobility prediction must be required in ASNET since the nodes are mobile which leads to dynamic network topology i.e. the move of the node away from the network that need replica in some other node which is expected to remain in the network for a particular period of time. For supporting mobility prediction the algorithm which combine the structure of the communities are required to minimize data loss when the node move out from the community.

The network partitioning in ASNET is also strict problem when the server or the node that contains the required data is isolated in a separate partition that reduce data accessibility to a large extent and give query delay because of disconnection between nodes.

Additional problem is in some ad-hoc social network, applications are time critical and may contain both fixed and flexible real time transactions but in some cases the transaction may miss their deadlines. If the replicas are not consistent with the updates to the data items, data accesses might not retrieve the most recent updates to the data items. This can cause problems in situations where it is critical that the most recent update to a data item is accessed. Therefore, the replication technique should be able to deliver correct information before the expiry of transaction deadlines in order to reduce the number of transactions missing their deadlines.

1.4 Contributions

In this thesis, we have proposed method to improve data accessibility and availability in adhoc social network using data replication techniques and combining the node into the community and we develop the framework of the system, community formation algorithm, community head election algorithm, RepCoH replica algorithmic for managing replica allocation among the nodes. We comprehensively evaluate data availability rate for all nodes available within the community and for the nodes in neighbouring community when network partition occur.

Subsequently, the study explores how well the declared community based replica allocation techniques work well in adhoc social network especially when network partition or nodes move from one community to other. During our experiment, we validate our desired network models by using NS2 and analyze the comparative performance of accessibility of data within the community and between the communities, evaluating the data service by comparing without replica, random replica and community head managed replica to maximize the availability and accessibility of data.

1.5 Research Questions

In our research work the following research issues considering to our system should be solved; The First one is how the node is created in the ASNET network and those nodes grouped into one cluster to form organized community, Secondly How community head is elected and by which parameter the selection procedure is preformed, At the Third How data is replicated and hold on the community heads of each community and when and how the nodes read and write there data. In general our research paper answers the question considering the structure of the nodes and data management of adhoc based social networks.

1.6 Objective of the Study

1.6.1 General Objective

The general objective of the paper is improving performance of data availability and accessibility in ad-hoc social network by combining the structure of community and designing RepCoH replica allocation mechanism.

1.6.2 Specific Objective

In this thesis we study community head based or RepCoH replica allocation mechanism for ASNET. The specific objectives of our thesis are:

- To improve the data availability, efficiency and persistent data transfer
- To Avoid data losses in case of unpredictable group mobility that causes community partition by combining the community structure
- To reduce number of hops when data transmitted from source to destination
- To reduce query delay in the network by data replication since bandwidth use of the packets are limited.
- To design community based data replication algorithm for ASNET network.
- To studying and evaluating RepCoH mechanism in adhoc social network in terms of availability of data, efficiency of data, data consistency and persistent data transfer.
- To evaluate the proposed algorithm with simulation environment and compare different replica on trace graph.

1.7 Research Methodology

The work of the paper is proceeding first by reviewing deeply of the literature on the area of the research topic. Then the method used for combining structure of the community to maximize data availability in adhoc social network is discussed by the following sequence.

i. Identifying and grouping the nodes those are closely linked to each other and those have high social relationships into clusters of entities called communities based on a variety of social connections such as social relationship, centrality and similarity, family, friends, common location, or common interest, which are decided with the social graph.

ii. Selecting community head (*cH*) by proposing election algorithm in which all nodes explain their capacity, including their current remaining energy, free storage space, and their node Identifications (*M-id*). Then the parameter information of the nodes stored into sequential table and based on the algorithm the node with top parameters is selected for the community head to handle the control of data for replica propagation.

iii. Then the selected community head flood its information to all mobile nodes within the community and to the neighbouring community head *(NcH)* including its mobile id, after that community members and the neighbouring community Head *(NcH)* register the *cH-id* on their table and request *cH* for replica by sending *M-id* and *D-id* if they need. Then *cH* respond the request of host node for holding replica of data with positive acknowledgment and nodes send their data to *cH* and *cH* allocate the data of the nodes with their *data-id* and *mobile-id* and allocate copy on *N-cH* for accessing data if partition occur.

iv. If the node requires to get their data whenever partition is occurred or when the node move from one community to other community or when node need their data, they can get the replica of data from original replica holders or copy of replica from the neighbouring community head, with or without original data holders

v. Evaluating the system by using tools such as Network Simulator version 2 or NS2 which is object oriented and discrete event driven network simulator that separates detailed protocol implementation from simulation configuration and which is the most popular network simulator for ASNET and using Microsoft Visio for designing flow of the system, structure of the framework and for designing the graphs and tables.

The methods can be done by designing the algorithm that combine the structure of neighbouring communities and by using replica allocation mechanisms that considers social communities.

1.8 Scope and Limitation

The research work includes the mechanism of replica allocation for maximizing data availability in ASNET and combination of community structure to hold replica on neighbouring communities, but the proposed method doesn't consider how to increase battery life span of the nodes those communicate and replicate the data with each other. Creating trusted relationship cryptographic keys between the nodes to communicate through medium and security of data on the replica holders are not the scope of the paper but for the sec of security we design trust chart between community heads and host nodes of different communities.

Since ad-hoc social network depends on the social relationships of the node, the probability of selfish node is low but if there is selfish node its detection is another research area. For discovering the node within and neighbouring community we use on demand reactive routing protocol since our network is within the same zone but comparison between protocol and increasing the speed of routing protocols in ASNET are another research work. The paper propose the algorithm that maximize data availability for infrastructure less ad hoc social network but combining adhoc social network structure with infrastructure based network that uses internet is not the scope of the paper.

1.9 Organization of Thesis

The organization of thesis is structured as follows. Chapter two presents and discusses the conceptual review of data replication in ASNET. The existing results and related work in replica allocation techniques over adhoc networks is addressed in chapter three. The first step to investigate the topic of this thesis is described in Chapter 4, in which the design for a combination of community structure for replica when topology changed is presented based on abstract reasoning about its fundamental building blocks. In Chapter five, the deployed community head based replica allocation techniques of proposed solution is evaluated by describing an innovative implementation design and also the performance of developed model was examined using NS2 simulation software. Some particular aspects are described in the appendices. For reasons of clarity and brevity, they are not part of the main body of the thesis.

CHAPTER TWO

LITERATURE REVIEW

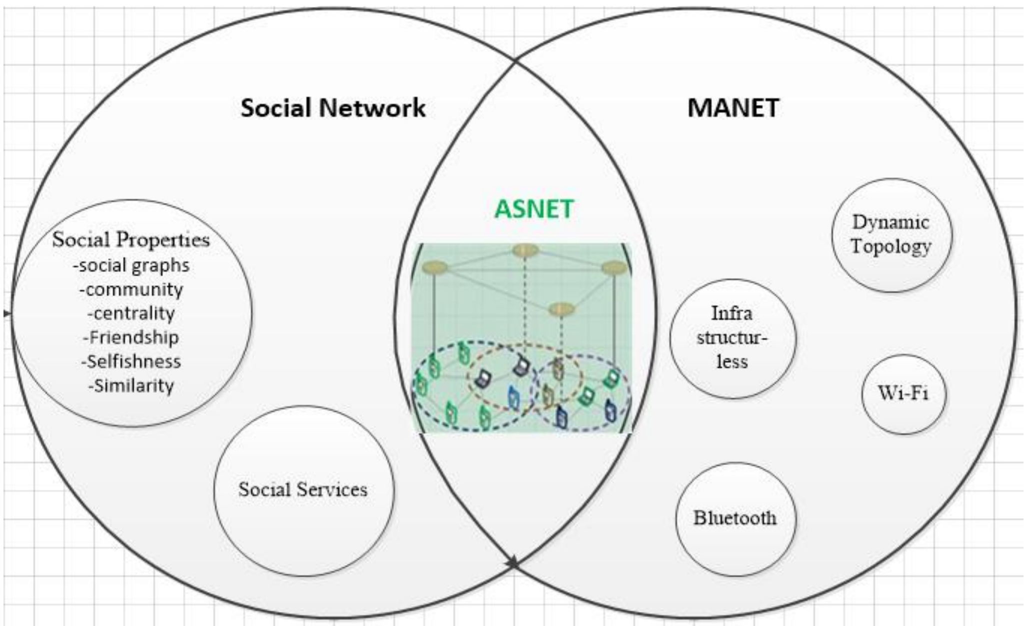
2.1 Adhoc Social Network (ASNET) Description

In today's internet technology, social networking is provided as an online service which is based on building relations among peoples. It is based on web and use centralized infrastructure. For example Facebook, LinkedIn, twitter etc provides communication among users interest to make social network. But online social network have a limitation since it depend on centralized infrastructure. For instance if the website is blocked or no internet connection on the device the social network access doesn't available. Another limitation of social network is it suggests friends based on their profile and not based on their dynamic interests i.e. present interest, context and proximity of users.

ASNET helps users to connect with other nearby users interested in similar objectives and create social network for small number of people those have similar dynamic interests or users interest in sharing common things from common source to destination and are in proximity without using infrastructure. People may be interested in spending idle time with persons having common interests while sitting on trains, participating on a conference, waiting for flights at airport or any location, to have contact information of friends on the meeting and etc. ASNET also provide information about other persons present there having similar interests in order to contact at that time and in the future.

ASNET uses routing protocol of adhoc network to facilitate the communication among interested users. Routing protocol of adhoc network allows each mobile devices as a router to communicate with each other without using any access point and router devices. The nodes in ASNET move freely and form an arbitrary topology because there is no master slave relationship for communication. The data packet sent from sender node is transmitted either directly to the replica holder if the nodes exist at the neighbour or through the intermediate nodes.

Figure 1: ASNET Structure with its contents



2.2 Applications of Adhoc Social Networks (ASNET)

Ad hoc networks are very well suited for many situations, in which an infrastructure network can't be built or it is impossible to build an infrastructure. The interest of ad hoc networks increases rapidly in recent year, because ad hoc supports mobility and freedom in the networks. Data can be exchanged without cable, access point, or portable memory space. Now days most of mobile phones are equipped with Bluetooth, Wi-Fi and Cellular radio and they can support ad-hoc communication mode. Since ASNET is infrastructure less self configuring network communication, people of related interests connected to each other using ad-hoc communication mode of mobile devices that used in different application such as information sharing in meeting, co-workers sharing calendar information, home networking, multi users games, communication between soldiers in battlefield or automated military communication system, search operations, law enforcement operations, disaster recovery , and tourists can also share their travel experiences or emergency information with other tourists and Airports can provide check in, boarding and seating services to customers as soon as they enter the airport and a lot of time can be saved by connecting to the customers' handheld device and etc.

2.3 Data Management in Wireless Adhoc Social Network

In the distributed data base management system data are replicated to improve reliability and availability. The one important issue to be considered while replicating data is the correctness of the replicated data. Replication is an important mechanism in modern wireless networks and has attracted significant efforts to improve its performance with different metrics including read cost, consistency and relocation cost [7].

Data replication is a technique stored on the multiple locations which reduce the search size. It is an appealing technique to provide services on unreliable networks caused by poor connections, host failures, and network partitions. To be most useful, data replication should be invisible to the user. The main improvement of using replication is that if you change only single copy of the data that presents in one node then data available on other nodes changes automatically. It has been studied for a long time in the areas of distributed file systems, distributed database systems, and recently in peer to peer information storage and mobile environments. It is used for improving reliability, scalability, fault-tolerance and accessibility of data services.

In distributing systems data object is retrieved from multiple locations, therefore it is beneficial to replicate data objects throughout the network[8]. Some advantages of data sharing among devices are *increased data availability and reliability* in the face of network failures, *faster query response* i.e. the queries initiated from nodes where replicas are stored can be satisfied directly without incurring network transmission delays from remote nodes and *load sharing* which means load of responding to queries can be distributed among a number of nodes rather than centralized at a single node [8].

The Important aspect in data replication of distributed systems is a mutual consistency, which is all replicas must agree on exactly one current value. When communication fails between nodes containing replicas of the same data maintaining mutual consistency between replicas becomes extremely complicated. The most disruptive of these communication failures are partition failures which separate the network into isolated sub networks which we call partitions. In general a data replication is an effective way to improve data access performance and data availability where network connectivity is not reliable and network partitioning may happen at any time. Replication therefore provides a way to support data services in ASNETs.

An Adhoc network is a collection of wireless mobile nodes forming temporary network without the need of centralized administration or support service regularly available on wide area network to which the hosts may normally be connected. In distributed system a data is usually accessed from multiple locations so it is beneficial to replicate data objects throughout the adhoc social network. The replication techniques in adhoc network increase data availability and reliability in the face of the network failures, faster query response without incurring network transmission delays from remote nodes, distributing load among the number of nodes rather than centralized at the single nodes. The other aspect of replica allocation is mutual consistency.

An ASNET is a multi hop wireless communication network which support mobile users without any existing infrastructure. To become commercially successful the technology must allow networks to support many users. Therefore to scale up the data availability among the network clustering the nodes into the community is important. Clustering provides a method to build and maintain hierarchical addresses and simplify the network routing for data transfer and allocation in ad hoc social networks. There is a lot of literature proposed for data management in ad hoc networks. We discuss the overview of the literature proposed in the following sections.

There is a number of data replication protocols have been proposed for Wireless network in the literature. Some of the proposed protocols assume that the access frequencies of devices to data items are same at every time but in real scenario access frequencies may be different at different times. Ad-Hoc Social Network (ASNET) have been proposed as anytime, anywhere social networks which use mobile ad-hoc network for its establishment and basic communication. Present social networking solutions suffers with disadvantage that they provide social networking in presence of Internet only [9]. ASNET is a solution to applications which provide anytime, anywhere social computing like sharing a cab at airport for a common destination. Mobile social networking is also an emerging area but it differs from Ad-hoc social networks in terms of optimized mobile ad hoc networks needed to support social applications for mobile devices like smart phones.

2.4 Components and Social Properties of ASNET

The main components of the ASNETs and their social properties those affect the communication patterns of the nodes are discussed as follows.

2.4.1 Mobile Devices

Mobile devices those users carry, allowing them to access the available content and form social groups. They can take any form, as long as they are mobile and might have multiple network interfaces. Mobile devices (e.g., mobile phones, PDA, or any human assisted devices) may have different wireless interfaces (e.g., Wi-Fi, cellular, and Bluetooth). Mobile users can receive data from content providers. Also, a mobile user can create and transfer the data to other users through radio link communication [10].

2.4.2 ASNET Network Connection Systems

Wireless communications is a fast-growing technology that enables people to access networks and services without cables. Deployment can be envisaged in various scenarios: different devices belonging to a single user, such as a mobile telephone, a portable computer, and a personal digital assistant (PDA), that need to interact in order to share documents; a user who receives email on a PDA; a shopping mall where customers display special offers on their PDA, car drivers loading maps and other tourist information while driving on a motorway. All of these scenarios have become reality from a technological point of view, and successful experiments are being carried out around the world.

Bluetooth and IEEE 802.11 (Wi-Fi) [11] are two communication protocol used in Ad hoc Networks and standards that define a physical layer and a MAC layer for wireless communications within a short range from a few meters up to 100 m with low power consumption from less than 1 mW up to 100 mW. Bluetooth is oriented to connecting close devices, serving as a substitute for cables, while Wi-Fi is oriented toward computer-to-computer connections, as an extension of or substitution for cabled LANs. The original idea behind Bluetooth technology was conceived in 1994, when Ericsson Mobile Communications began to study a low-power consumption system to substitute for the cables in the short-range area of its mobile phones and relevant accessories. Bluetooth [1] is a standard for wireless communications based on a radio system designed for short-range cheap communications devices suitable to substitute for cables for printers, faxes, joysticks, mice, keyboards, and so on. The devices can also be used for communications between portable computers, act as bridges between other networks, or serve as nodes of ad hoc networks. This range of applications is known as wireless personal area network (WPAN). In 1997 the IEEE approved a standard for WLAN called 802.11, which specified the characteristics of devices with a signal rate of 1 and 2 Mb/s. The standard specifies the MAC and physical layers for transmissions in the 2.4 GHz band. The aim of the IEEE 802.11 standard is to provide wireless connectivity to devices that require quick installation, such as portable computers, PDA, or generally mobile devices inside a wireless local area network (WLAN). It defines the MAC procedures for accessing the physical medium, which can be infrared or radio frequency. Mobility is handled at the MAC layer, so handoff between adjacent cells is transparent to layers built on top of an IEEE 802.11 device [11].

2.4.3 Social Network

The notion of social network and its application has recently gained much attention from the researchers in many fields. The reason is that all entities (e.g., people, devices, or systems) in this world are related to each other in one way or another. A social network not only defines the behaviour of these entities but also helps to understand different relations among them. In principle, a social network is a structure of entities (e.g., individuals, organizations, and systems) that are connected to each other through one or more interdependencies. These interdependencies could be shared values, physical contacts, financial exchanges, commodity trades, and group participations [3]. Recently, the concept of social networks has been used in the context of information and communication technologies to provide efficient data exchange, sharing, and delivery services. A social network defines the structures and ties among the users in which the users and system can use the knowledge about the relationship to improve efficiency and effectiveness of network services.

Social Networks have undergone a dramatic growth in recent years. Such networks provide an extremely suitable space to instantly share multimedia information between individuals and their neighbours in the social graph. Social networks provide a powerful refection of the structure and dynamics of the society of the 21st century and the interaction of the Internet generation with both technology and other people. Indeed, the dramatic growth of social multimedia and user generated content is revolutionising all phases of the content value chain including production, processing, distribution and consumption. It also originated and brought to the multimedia sector a new underestimated and now critical aspect of science and technology: social interaction and networking. The importance of this new rapidly evolving research field is clearly evidenced by the many associated emerging technologies and applications including online content sharing services and communities [10].

2.4.4 Social Graph

A graph is a mathematical abstraction for modelling relationships between things. A graph is constructed from nodes (the things) and edges (the relationships). Social networks exhibit the small world phenomenon that node encounters are sufficient to build a connected relationship graph. The graph is a convenient tool to represent the relational structure of social networks in a natural manner, which is generally called social graph. In a social graph, vertices (nodes) indicate human individuals, and edges (links) indicate social relationships between individuals. In some degree, the social network is equal to the social graph, so they can be used alternatively. One significant challenge in social networks is how to represent a link between two nodes [3]. According to different link meanings, several social graphs are proposed in the recent literature. The contact graph is a popular way to analyze and estimate relationships among people by observing their inter-contact time in the history. Moreover, the neighbour graph, regularity graph, and interest graph are proposed recently as well [3].

2.4.5 Community

A community is a structural subunit which can be represented as a set of individuals of a social network with high density of internal links [3] [4]. Individuals have more social connections with other individuals inside their own community than with individuals outside. The social connections may be family, friends, common location, or common interest, which are decided by the social graph. In general, individuals in the same community may meet each other more frequently. Therefore, the community structure has significant impact on people's mobility patterns and thus is beneficial for choosing the appropriate forwarding path.

In the analysis of neighbouring nodes made previously, a new term was introduced, community. According to [4] [10], community is a clustering of entities that are closely linked to each other, either by direct linkage or by some easily accessible entities that can act as intermediates. What drive individuals to form communities is a social relationship of any form and background. Common interests and family bonds constitute relationships leading to the creation of communities, such as families or art groups. The existence of communities in the social structure is undeniable. Utilizing this property in Mobile social Networking's depends on whether people clustering are taken into account to improve scalability and optimize routing. Such an attempt aiming to organize routing between social communities is presented in [10]. They argue that clustering based on social structures seems like an obvious choice if the mobile nodes are being carried by people. Routing is divided into two levels, inter-cluster and intra-cluster. In the former, messages move from cluster (community) to cluster until the destination group is found, while in the later, messages are routed inside the destination cluster until the correct destination is found. Inside each cluster, a node is assigned to be the cluster head according to a quality metric, depending on factors such as its transmission range and mobility.

2.4.6 Centrality

Aside from community, the centrality is another basic concept in social networks, which considerably affects the performance of socially aware forwarding algorithms. The experiments in [12] demonstrate that it is important to find appropriate centrality and community in the design of socially aware forwarding and data dissemination algorithms. Centrality is used to describe important and prominent nodes in a social graph. People have various roles and popularities in society. A central node has stronger capability of connection with other nodes.

2.5 Mobility Issues in ASNET

In ad hoc social network there are numerous applications in which mobile users can share information, for example sharing calendar information on the meeting. For such applications enhancing data availability is important issues and various issues are conducted with this intension. Takahiro Hara [13] quantified the influence of mobility patterns of data availability from different viewpoints. Different mobility models can be differentiated according to their spatial and temporal dependencies. A Spatial dependency is a measure of how two nodes are dependent in their motion. If two nodes are moving in same direction then they have high spatial dependency. The Temporal dependency is a measure of how magnitude and direction are related to previous velocity. Nodes having same velocity have high temporal dependency. The different mobility protocols used in MANET as well as in ASNET are as follows [14]; Random Walk (RW) which is one of the simplest mobility models and is often used in simulation experiments for MANETs. In this model, at every unit of experimental time, each mobile node randomly determines a movement direction from all directions, and randomly determines a movement speed from 0 to V [m/s]. It is known that in the long term, this model offers very low mobility similar to vibrating in the same position, because mobile nodes randomly change movement direction. The other is Random Way Point (RWP) which is one of the most popular mobility models for MANET researchers. In this model, each node remains stationary for a pause time S [s]. Then, it selects a random destination in the entire area and moves to the destination at a speed determined randomly between 0 and V [m/s]. After reaching the destination, it again pauses, and then, repeats this behaviour. It is known that in this model, mobile nodes tend to gather at the center of the area, and the movement speed tends to converge to zero (very low).The other is Manhattan Mobility (MM) [14], this model emulates the node movement on streets where nodes only travel on the pathways in the map. Manhattan grid maps of horizontal and vertical streets are used to restrict the node movement. On each street, the mobile nodes move along the lanes in both directions. At each intersection, the mobile nodes choose their directions and speed (0 to V [m/s]) randomly. Reference Point Group Mobility (RPGM) is also model of group mobility. Each group has a logical center called a reference point and group members (nodes). Each reference point moves according to the RWP model with V 0 [m/s] (maximum speed) and S0 [s] (pause time). In each group, nodes are uniformly distributed within a certain radius R from the reference point. To achieve this, we assume that each node moves according to the RW model with V [m/s] (maximum speed) within that range. Specifically, a node’s movement vector is composed by adding the movement vector based on the RW model for the node to that based on the RWP model for the reference point. The last is Random Way Point with Locality (RWP-L) model which is same as the RWP model except for the way to choose destinations. In this model, each mobile node has a home area, which is a subarea in the entire area. When determining a destination, it chooses a random destination inside the home area with high probability H and one outside the region with probability 1H. [14].

2.6 ASNET Routing Protocols

Because of the fact that it may be necessary to hop several hops or multi-hop before packet reaches the destination, routing protocol is needed. Routing protocol of Adhoc network has two functions; the first is selection of routes for various source destination pairs and delivery of messages to their correct destination. And the second is conceptually straightforward using a variety of protocols and data structure or routing tables. There are different routing protocols used in Ad hoc network basically classified as proactive, reactive and hybrid routing protocols. Mostly used routing protocols in MANET as well as ASNET are as follows; AODV(Ad Hoc On-Demand Distance Vector Routing) using distance-vector concept, but in several different ways and the above is that, AODV does not maintain a routing table, but when a node needs to communicate with another node on demand only to the approach to building routing table. When a node wants to send data to another node in the network, the first to broadcast a Route Request (RREQ) packet, RREQ where the record that this is given by which a source is to be used to find which of a destination node[15]. Secondly DSR (Dynamic Source Routing) [16] As the name suggests is the use of the concept of source routing, the routing information that is directly recorded in the inside of each packet, but to be in the MANET environment, the use of such a special, DSR is needed only when the path to find out the path, that is, on-demand. Thirdly, DSDV (Destination-Sequenced Distance-Vector Routing) [15] is based on traditional Bellman-Ford routing algorithms were developed by the improvement, and a routing table-based protocol. Each node in an operation must be stored a routing table, which records all the possible links with the nodes in the node and the distance like the number of hops, routing table within each record also contains a sequence number, which is used to determine are there any more old path in order to avoid routing table generation. Fourthly, Cluster Based Routing protocol (CBRP) [15] [16], all the nodes are separated into clusters. In order to arrange the cluster, the following algorithm is used. When a node comes up, it will go into the undecided state and broadcasts a Hello message. When a cluster-head gets this hello message it will reacts with a triggered hello message immediately. When the undecided node gets this message it changes its state to member. If the undecided node times out, then it will turn as a cluster-head. If it has a bi-directional link to some neighbour, otherwise it remains in undecided state and repeats the procedure all over again. Cluster-heads are changed as occasionally as possible. And other routing protocols used in ad hoc networks for better routing to destination nodes.

2.7 Summary

After surveying literatures corresponding to ad hoc communication we design new method for data replication of ASNET (Adhoc social network) by adding social interest and social behaviour of node in the network. Our proposed algorithm in this research combines the structure of the nodes into the community for replication of data and hold replica of data on the community head selected from inside cluster and neighbour clusters for maximizing data accessibility and availability in ASNET.

In general there are a lot of papers proposed for data replication of wireless adhoc networks. But most of the papers are designed for mobile adhoc networks without adding the social behaviour, social interest and social mobility or the same movement of the mobile nodes and their replica techniques are based on data access, type of resource they use, the properties of the nodes and etc. In the paper Duc A. Tran et al the social relationship of node is proposed by designing S-CLONE but they consider all online social networks which are based on infrastructure to design. Most of papers we reviewed show data replication techniques in MANET without adding the social services and behaviours. Its limitation are those papers doesn't show combination of the community and interest of the nodes for data replica and most of them doesn't consider the bandwidth, route time, battery power, the memory capacity, the mobility and data consistency when hold replica on the other nodes. The only and the first article proposed by Ahmedin M. et al. are based on ASNET data management by considering the social behaviour of the nodes. Ahmedin M. et al. develops the middleware which control all the system of the ASNET network and they design new protocols which is called ComPAS,[17] a community-partitioning aware replica allocation method. Its goals include integration of social relationship for placing copy of the data in the community to achieve better efficiency and consistency by keeping the replica read cost, relocation cost and traffic as low as possible. As they propose in paper [4] ComPAS is based on a perception that if we need to place a replica copy for a user i somewhere, the most desirable location should be the primary storage place for most neighbours of i; this way, most neighbours will benefit from this replica when they issue a read query. But the mobile nodes which is far from replica holder doesn't access its data easily, since limitation of bandwidth and limitation of storage and it need maximum route discovery to get its replica holders. And also the storage, power and the centrality management of data when partition for omitting data loss doesn't considered in this paper. Therefore we design the system which combines the community structure for maximizing data availability. Our proposed data management systems are based on community head which is called RepCoH, data replication on the mobile nodes which satisfy the parameters and selected for community head. Since data replica holders are community head exists in each communities and the copy of replica placed on neighbouring community heads, the accessibility and availability of data when partition occur are protected very well.

CHAPTER THREE

RELATED WORK

3.1 Overview

ASNET is a multi hop wireless communication network which support mobile users without any existing infrastructure. To become commercially successful the technology must allow networks to support many users. Therefore to scale up the data availability among the network clustering the nodes into the community is important. Clustering nodes into different communities provides a method to build and maintain hierarchical addresses and simplify the network routing for data transfer and allocation in ad hoc social networks. In this section, we discuss our work in light of the existing related papers on cluster head election techniques and different data replication techniques in social networks, MANET, social data deployment based on cloud computing and ASNET, respectively.

3.2 Cluster Head Election Techniques

To handle data transfer and storage processed under ASNET network dividing the topology into sub partition which is called the communities are required. Each community in the topologies must have community head for controlling data write and read in the group. There is different cluster head selection algorithms proposed in different articles based on their own parameters. Some of the proposed cluster head selection techniques for different wireless network are the follows.

Mainak Ch. et al. [18] proposes a weight-based distributed clustering algorithm by considering the ideal degree or cluster size, transmission power, the mobility and battery power of mobile nodes. The time required to identify the cluster heads depends on the diameter of the underlying graph. They try to keep the number of nodes in a cluster around a pre-defined threshold to facilitate the optimal operation of the medium access control (MAC) protocol. The non-periodic procedure for cluster head election is invoked on-demand, and is aimed to reduce the computation and communication costs. They assume cluster heads, operating in dual power mode connects the clusters which help in routing messages from a node to any other node.

Hao W. et al. [19] propose a novel cluster head selection and update algorithm which is called TCA (type based cluster forming algorithm) in order to reduce the update frequency of cluster heads. In their article they assume each node is aware of its exact location information which is ensured by GPS based or other positioning device, each node is equipped with an Omni-directional antenna and an idealized zero delay channel access protocol is used and also topology doesn’t change during election.

The other paper proposed on cluster head selection is by Sun B. et al. [20] which construct the stable cluster head with the help of reducing the number of cluster head reconstruction. It selects the nodes which have the stability to be the cluster heads. In this paper the authors focus on the stable availability of wireless communication connections and they propose stable cluster head selection algorithm for adhoc networks (SCSA), which discovers the mobility groups in a distributed manner and elects a cluster head based on the stability and movement vector of each mobility group.

In the paper Raihana F. et al.[21], they propose that the entire network be divided into hierarchical group of clusters. They assume that nodes are location unaware i.e. not equipped with GPS, they are left unattended at the beginning so no need for battery re-charge. In this paper, they propose Cluster head(s) selection algorithm based on an efficient trust model i.e. each group of nodes has one or more elected Cluster head, where all Cluster heads are interconnected for forming a communication backbone to transmit data. Moreover, Cluster heads should be capable of sustaining communication with limited energy sources for longer period of time. Misbehaving nodes and cluster heads can drain energy rapidly and reduce the total life span of the network. In this context, selection of best cluster heads with trusted information becomes critical for the overall performance. This algorithm aims to elect trustworthy stable cluster head(s) that can provide secure communication via cooperative nodes.

In our proposed system we add some social and community head election parameters on Raihana F. et al.[21], paper which select Cluster head(s) selection algorithm based on an efficient trust model and reduces overhead, flooding, collisions and makes the network more scalable and provide secure communication via cooperative nodes. We add social interests or relationships and parameters such as mobility, energy and storage on the paper for better clustering and data allocation.

3.3 Data Replication Techniques

Different data replication techniques are proposed by different authors for Adhoc networks, MANET and Social Networks. The authors propose the replication methods in different way such as access frequency, power consumption, memory capacity, bi-connected group of node and etc. The proposed literature discussed as follows.

3.3.1 Non-Constraints-Aware Replication Techniques (SAF, DAFN, DCG)

As T.Hara [22] proposes, data replication offers the benefit of shortening the response time for database writes or read operations and improving the data availability. Most replication strategies proposed so far address this issue. This issue includes the propagation of update operations to replicas. The latter is achieved by replicating databases and using the replicas when the site which holds the original database fails. On the other hand, since frequent division of the network is a special characteristic in adhoc networks, Their approach takes it into account and is completely different from that in distributed database systems. Shortening the response time and propagating update operations are also significant issues in ad hoc networks.

Hara [23] also proposed a series of data replication schemes in ad hoc networks. In order to guarantee data accessibility upon network partitioning, these works focus on optimizing the location of data replicas within a network and are based on the assumption that access frequencies to data items from each node are known and are fixed. His assumption limits the applicability of the schemes in practical systems. Similarly to Hara's work, Wang and Li [24], Huang et al. [25], and Derhab and Badache [26] considered the problem of replica allocation and group mobility model. Their approach takes into consideration topological information when replicating data and data replication occurs only when necessary according to certain partition detection schemes. However, in these works, partition detection depends on the mobility model. Furthermore, it is assumed that the locations and velocities of all mobile nodes are known. These assumptions are too restrictive to capture the reality of mobile ad hoc networks.

Other paper by Hara et al. in [27] for replication assume an environment where each mobile host accesses data items held by other mobile hosts in a MANET and allocates replicas of the data items on its memory space. They also assume that the area in which mobile hosts can move around is divided into several regions and the consistency of data operations on replicas is managed based on the regions. They design the protocol known as E-DCG+ [23] which creates the group of nodes that are bi-connected components, and shares replicas in larger groups of nodes to provide high stability. The work in [27] focus on classifying different replica consistency levels in a Mobile Ad hoc Network based on application requirements, and proposes protocols to realize them. In their work, each replica is valid till its original owner updates it. Hence, applying strict consistency updates may potentially degrade the system performance, given the inherently dynamic nature of the environment. Thus, the work assumes that all applications do not necessarily require such strict consistency, and it defines consistency based on group level information consistency. Particularly the protocol proposed by T.Hara et al. does not consider the social behaviour integration on the network.

In general T.Hara et al. [22] [23] [27] has proposed three data replication methods to improve data availability: Static Access Frequency (SAF), Dynamic Access Frequency and Neighbourhood (DAFN), and Dynamic Connectivity Based Grouping (DCG). All these methods assume finite node memory and predefined constant access frequency of each node towards data items. In SAF, each node creates replicas of its frequently accessed data items without considering data items replicated by its neighbours. In this methodology, [28] each mobile host allocates replicas of data items in descending order of the access frequencies. At the time of replica allocation, a mobile host may not be connected to another mobile host which has an original or a replica of a data item that this host should allocate. In this case, the memory space for this replica is retained free. The replica is created when a data access to the data item succeeds or when the mobile host connects with others. In DAFN, firstly SAF method is executed, then every pair of neighbouring mobile hosts compare their replicated data items to eliminate duplication and hence increase the number of data items that may share between two nodes. It [28] aims to eradicate the pitfall of SAF method wherein replica duplication is visualized. This method achieves replica allocation in same way as that of SAF on the basis of decreasing access frequencies. However it also further analyses the neighbouring nodes in the network to determine any duplication. If any duplication is discovered then the mobile host with low frequency of data items get its replacement. Thus the duplication problem is substantially reduced by using DAFN. The DCG method [29] first divides the nodes into stable groups by finding the bi-connected components and then shares replicas in larger groups of mobile nodes by considering group access frequency. Here group access frequency of each data item is calculated as a summation of access frequencies of nodes in the group to the data item. It is an improvement over DAFN. Here, [28] instead of allocating replicas to a pair of neighbourhood nodes, allocation is done within a larger group of nodes. The requirement of this methodology is that the group should be stable. It should sustain frequent changes in the network topology. DCG creates groups that are bi-connected components in a network. This means that each group is a maximum partial graph which is connected if an arbitrary node in the graph is deleted. This makes the group highly stable as removal of a node from the group does not divide the group. DCG method is executed in every relocation period.

Our Discussion: The three methods called SAF, DCG and DAFN suffer from several weaknesses and are not actually adapted to mobile networks. They assume that existing data items and their access frequencies are known for all nodes and they do not change. We argue that this assumption is difficult to be maintained in ASNETs. It looks like having a centralized table containing needed information that each node consults to make its decision about which data item it should replicate. In addition, these methods don't consider any of the resources constraints nor the network constraints. DCG might be the best method as it optimizes the network traffic and the allocation effort, but it is still a primitive method that can be applied only to scenarios where limited and small number of nodes participates and all nodes identify each other in advance.

3.3.2 Network Limitations Aware Replication Techniques (DAFN-S1, DAFN-S2, DCG-S1)

Extensions to the above methods have also been proposed in [29], to handle periodic and a periodic updates. Hara also proposed another extension to the above three methods by considering the link stability between connected nodes; DAFN-S1, DAFN-S2 and DCG-S1 [30]. The methods aim to eliminate duplicate replicas of data items between two nodes if the wireless link that connects them is stable. The methods assume that each node equipped with GPS device and using speed and location information are able to predict the time at which two neighbouring nodes will be disconnected. The main drawback of all these approaches is that the relocation period is constant, and hence cannot adapt to changing network topology.

Sweta J. et al. [31] propose data replication techniques called W-DCG (Weighted Dynamic Connectivity based Grouping) which aims at addressing the limitation of DCG which is proposed by T.Hara [22]. The proposed protocol considers node properties such as memory capacity and remaining battery power while replicating data on nodes to achieve better load-balancing. Moreover it uses dynamic access frequencies to replicate data items which are currently in demand. Data sharing is done in large groups, which are stable to reduce duplicity of data. This type of collaborative replication increases the availability of different data items in a group. When replication mobile host can create replicas of up to its memory size including its original data items and when a mobile host generates a request for a data item, the request is immediately satisfied if it holds the original or replica of that data item in its memory space; otherwise, it sends the request to its one-hop neighbours which again check for the requested data item else the request fails.

Our Discussion: The above three methods consider the topology changes and the network partitioning. Replication is done after each period of time considering with fixed access frequencies. The main weaknesses of the methods are calculation of the radio link stability i.e. it assumes a prior knowledge of the speed and direction of the movement of all mobile nodes which is not realistic in dynamic environments. DCG-S1 is an expensive method as it causes high traffic and needs a long execution time to allocate the replicas. Finally, the relocation period is constant; thus if the relocation period is too long, the methods cannot adapt to the topology changes, and if it is too short, the traffic increases and unnecessary replicas are diffused. So tuning this period is crucial but complex.

3.3.3 Data Aware Replication Techniques (CADRE)

As Authors propose in [32] the CADRE or is the Collaborative replica Allocation and De-allocation of Replicas with Efficiency, deploys super peer architecture called gateway node (GN). Those gateways are responsible for both the replication and the search operations. The replication is performed based on the nodes characteristics (e.g. load and energy) as well as the nodes schedules. The gateway nodes have all necessary information due to periodical messages sent by all nodes. The main aim of CADRE is to keep replicas based on the available memory size of neighbour nodes and to perform the replica de-allocation in a collaborative way. It addresses the issue of a fair replica allocation across mobile nodes. To this end, CADRE considers in addition to the access frequency of a data item DI, the origin of the requests. So the item will not be considered as ‘hot’ifonly one node issues a large number of queries asking for it. To ensure fairness, each node assigns a score to each data item. This score quantifies the importance of DI for the network as a whole. It increases as the data item is demanded from more nodes; hence, given two data items with equal access frequencies, the score of the data item that is demanded from a larger number of nodes will bi higher.

Our Discussion: CADRE is an original replication model. It addresses aspects that are rarely addressed by other models: the size of the data items, and the replication fairness. it is a centralized model, all the decisions are made by the super peer which makes the cost of the replication system very high (all nodes send their access log, load status, available memory, etc.) to the super node that makes all the operations. Finally, the client/server architecture does not suite the mobile networks.

3.3.4 Cluster Based Replication Techniques (ART)

R. Azeem et al. [33] proposed the replica allocation technique based on cluster for MANETs. In their work each mobile node is associated with a cluster and each cluster has its cluster head (CH) and each CH will maintain an Available Replica Table (ART). Whenever a node requires a new data item then node has to send the request to the CH. Then CH will check the id of required data item in its ART, if found the data id, then the request is redirected to the node with node-id pertaining to that item-id in ART. When a node receives a data item then, it will make a replica of it for future use and an update message was sent to the CH.

Our Discussion: The proposed articles by R. Azeem et al. are based on cluster and solve the problem of bandwidth usage and routing time when data is allocated from one node to another. But it doesn't consider the nodes mobility, social interaction of the node, centrality and interest of the nodes. There fore the system which considers social relationship of the nodes when community created and when data interchange is required which is proposed in our systems.

3.3.5 Resources-Limitations-Aware Techniques (CReaM)

The approach of Zeina T. [5] describe a fully decentralized replication model for MANETs which is called CReaM (Community-Centric and Resource-Aware Replication Model) that is designed to cause as little additional network traffic as possible. In CReaM data item to replicate is selected depending on the type of resource that triggered the replication process. The best data item to replicate in case of high CPU consumption is the one that can better alleviate the load of the node, i.e. a highly requested data item. Oppositely, in case of low battery, rare data items are to be replicated (a data item is considered as rare when it is tagged as a hot topic (a topic with a large community of interested users) but has not been disseminated yet to other nodes). To select the replica holder (i.e., the node that will receive a data item replica), CReaM integrates a prediction engine that estimates which location should provide the best load balancing and the best data usability.

Our Discussion: The community centric nature of the solution allows synchronized replication decisions between the nodes in the region, thus it avoids replica duplications and non useful replica creation. However, it is a costly solution; it is based on one super node per region (that is fixed or does not move outside the region); all nodes in the region communicate their information either load status, memory, Access Frequency, and etc periodically to the super node or very high network traffic, that it uses them to make the local decisions and to apply all operations gathering information of the nodes in the region, distributing replicas, and processing requests.

3.3.6 Social Network and Cloud Deployment Based Replication Techniques

In recent times, there has a recent unprecedented increase in the use of social networks and applications with the social component. Since highly scalable and elastic network resources are required to deploy new social applications, it is promising to deploy social applications in the cloud. As discussed in the F. Benevenuto et al. [34] paper, when placing social networks data among multiple clouds, it maintains the social locality that mean one-hop locality by replicating the data of a user's every neighbour to the cloud that hosts the user's own data which has been shown to be necessary for social network services due to the data access pattern that a user's most activities happen between their self and their neighbours (For example all friends must be accessed to collect their recent tweets for a Twitter user). Maintaining social locality eliminates the inter-cloud multi-get traffic and unpredictable response time, because all the data requested by a user can be retrieved from a single cloud. Compared with full replication, the storage cost is much lower because only neighbours need to be replicated across the clouds.

Our Discussion: The paper proposed by different authors on social network of cloud solve the problem of data management in infrastructure based network those use internet by holding data of online social network on different clouds. In our work, we consider offline social network that base social relationship as a social behaviour and user interactions within the community to achieve better data availability after the network division because of the dynamic nature of ASNETs.

3.3.7 Social Aware Based Replication Techniques (S-CLONE, ComPAS)

Duc A. Tran et al. [35] proposes S-CLONE; socially-aware data replication schemes which improve a social network's efficiency by taking into account social relationships of its data. S-CLONE is proposed for Online Social Network (OSN) to solve the problem of degree of data availability for every user to access data equally under any failure condition and while replicating data attempts to put those socially connected into the same server as much as possible.

Our Discussion: The Duc A. Tran et al. paper solve the online social network of data availability problem by putting socially connected nodes into the same server for data replication to improve data accessibility. But since it need infrastructure to replicate data between nodes it doesn't fully applied to offline social network which is called Adhoc social network.

Ahmedin M. et al. [6] [4] propose new data replication method by adding social behaviour on the MANET and create new network which is called ASNET (Adhoc social network). They propose new protocol called ComPAS (community-partition aware replica allocation method).The method will increase the availability of different data items in partitioned social community and it improve the performance of ASNET by exploiting social relationship while replicating in the community to achieve better efficiency and consistency while keeping the replica relocation cost as low as possible. They contribute for data replication in ASNET by underlying different modules such as consideration of social relationship as social behaviour for the exploitation of group mobility model, providing an overview and selection on existing replica allocation approaches, formulating and deriving the average read cost for the original data storage space in a community without replication and calculate the optimal Read Cost Reduction (RCR), and choosing an optimum approach and provide a comparison based on metrics such as read cost, relocation period, number of mobility group and efficiency of consistency management.

In Ahmedin M. et al. [6] work a user belongs only to one community, and created social communities should be determined not only by common interests or contacts but also by mobility related context. Their paper is the first and only paper for replica allocation of data for ad hoc social network which add social behaviour and work on the top of mobile ad hoc network. They also propose [4] ASNETs middle layer model and cross-layering approach among the layers and components which will be used as an inspiration to build an extended and full fledged data management middleware which illustrates the interaction among the four models including the application management model, reliability management model or attack model, data and resource management model or replica management and social community network model or social graph.

Our Discussion: The goal of paper proposed by Ahmedin M. et al. include integration of social relationship for placing copy of the data in the community to achieve better efficiency and consistency by keeping the replica read cost, relocation cost and traffic as low as possible. As they propose in paper [4] ComPAS is based on a perception that if we need to place a replica copy for a user i somewhere, the most desirable location should be the primary storage place for most neighbours of i; this way, most neighbours will benefit from this replica when they issue a read query. But the mobile nodes which is far from replica holder doesn’t access its data easily, since limitation of bandwidth and limitation of storage and it need maximum route discovery to get its replica holders. And also the storage, power and the centrality management of data when partition for omitting data loss doesn’t considered in this paper. Therefore we design the system which combines the community structure for maximizing data availability.

Table 1: Summary of some of Data Replication Techniques developed by different authors and our proposed work

The Authors the Papers:	The Method they Used:	Their Replica System Based On:	The Protocols Developed for	Weakness and Complexity
T.Hara et al.	SAF DCG DAFN E-DCG	- Access Frequency [22], [23] - Group level information consistency [27]	MANET	AF do not change, any node knows all other nodes, their data and their AF
Sweta J. et al.	W-DCG	- Node Properties (memory, power) [31] - Based on AF	MANET	It needs a long execution time to allocate the replicas
R. Azeem et al.	ART	- Clusters [33]	MANET	Doesn't consider mobility , social interaction of node
Duc A. Tran et al.	S-CLONE	- Social Relationship [35]	OSN	It needs infrastructure for replica
A.Mondal et al.	CADRE	- Replication based on the load status, memory, AF and the location [32]	Client/server architecture	High communication cost.
Zeina T.	CReaM	- Type of Resource [5]	MANET	Partial replication and high communication cost
Ahmedin M. et al.	ComPAS	- Social Behaviour [6] [4] - Full and Random Replication - Greedy Replica Approach	ASNET	It select random node for replica without nodes parameter consideration
Negasa et al.	-Community combination approach - RepCoH (Replication on Community Head)	- Social Interests and nodes moving behaviour - cH and neighbour community head - Mobility and partition of nodes - Parameter information of nodes such as capacity, energy , centrality etc	ASNET	Infrastructureless communication which maximize data accessibility through replica

CHAPTER FOUR

REPLICA ON THE COMMUNITY HEAD (RepCoH) MODEL

4.1 Overview of the RepCoH Model

The methodology of this thesis was presented in chapter one. These chapters describe the model proposed for our statement of problem, the topology and description for deployment of model on NS2 simulation software.

The purpose of this research work is to increase data availability and accessibility and reduce data loss in ASNET in the case of network partition and when topology is changed. In this study community combination methods and replica allocation methods are proposed. To maintain data management in ASNET, clustering the node into community is needed therefore in this model nodes form community depend on their social interests and community head is selected for handling data replication within the community and to neighbour community. The techniques are discussed in the following proposed models.

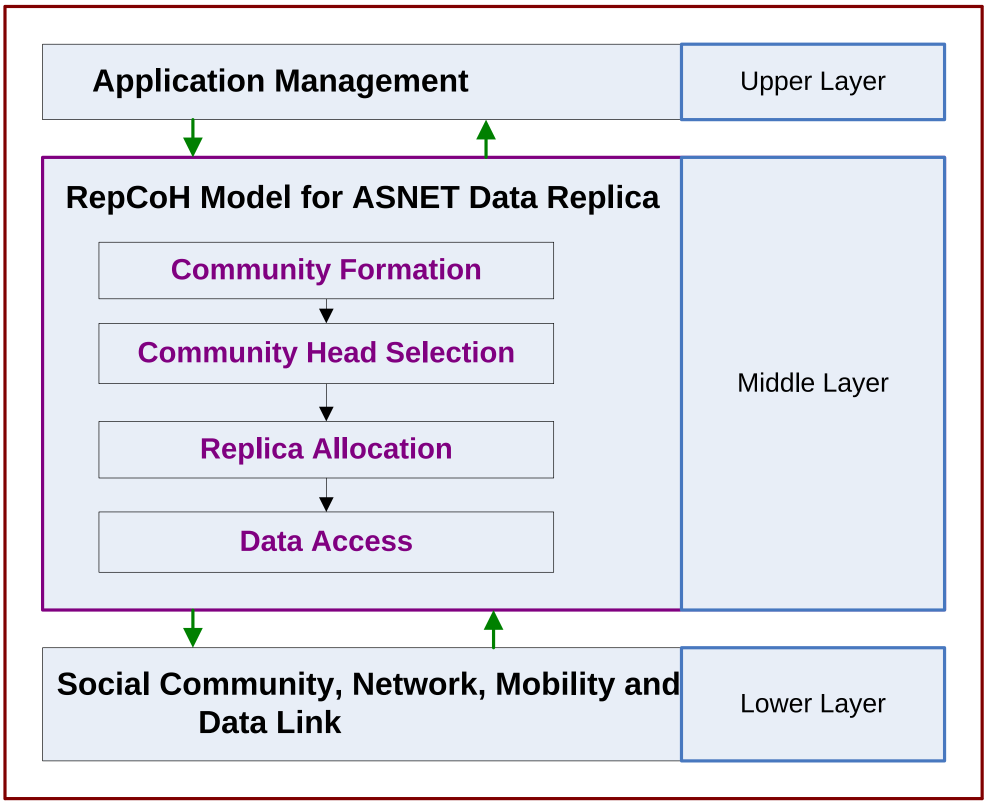


Figure 2 : *ASNET RepCoH Model for data management of ASNET*

4.2 Structure of the Community in Proposed Solution

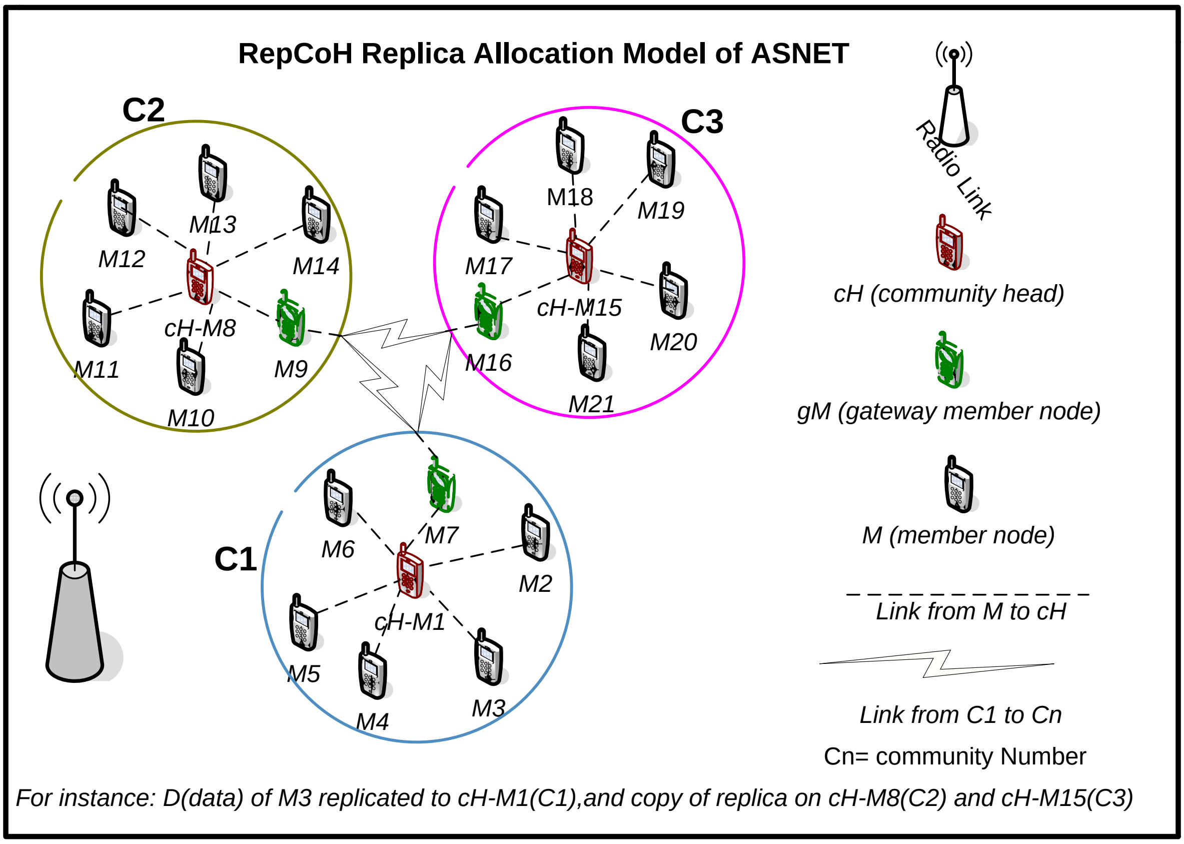


Figure 3: Scenario of proposed RepCoH Replica Model of ASNET

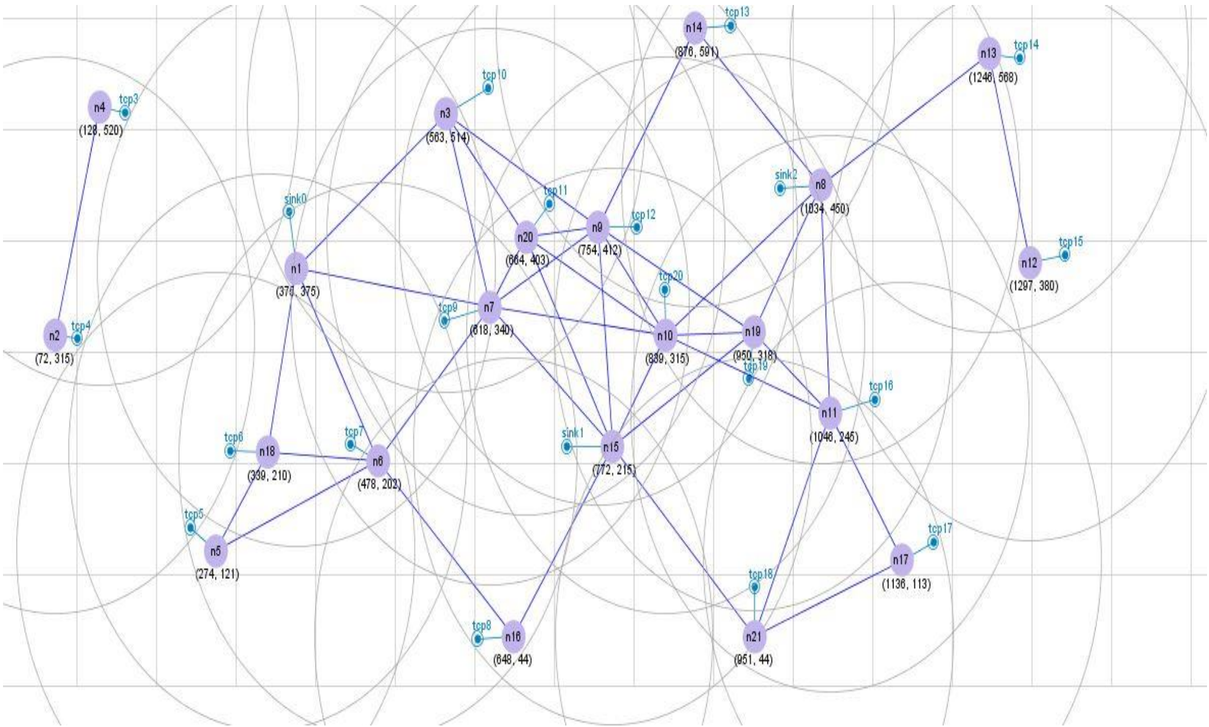


Figure 4 : The System Simulation Scenario with 21 nodes

4.3 The RepCoH System Model Description

The ASNET network system have a middleware [4] which includes Application management model which exist at the top level of the middleware and manage application exist in ASNET, Reliability management model or Attack model that handle misbehaving or selfish nodes, Data and resource management model or replication model, and Social community network model which is the social graph of the ASNET. Basically in our work we discuss the data management model that handle all process of replica of data in ASNET. Data is an important element in general wireless network, particularly in ASNETs data management turns to an essential design problem and it address challenges such as coordinating actions with other nodes, limited processing speed and storage capacity [6].

Replication is the basic techniques used in adhoc networks. Replica holding is a means of distributing consistent data at multiple servers within and at neighbouring communities. This way improve data availability and accessibility since community head of the clusters hold and manage the data within the community and its neighbours. In this section we combine the structure of the community and we define the community head based replica allocation system for the management of duplicated data items and maximizing the availability of data. Due to the mobility of node and dynamic nature of ASNET, the nodes with similar moving behaviour create a community [6]. The nodes in ASNET network are [36] directly connected or through a multi hop based on some social properties such as social graph, community, centrality, similarity, tie strength and human mobility pattern.

4.3.1 Application Management

The role of application layer in ASNET is to handle the frequent disconnection and re-connection with peer applications and to control time varying delay and packet loss statistics.

4.3.2 RepCoH Model for Data Replica

4.3.2.1 Community Formation Architecture

ASNET involves a great deal of human to human relationships, personal information, human to community relationships and human to environment relationships [1]. Individuals with related social properties such as interests, background, profession and other similarities are more likely to be friends i.e. they interact more than strangers. Specifically friends tend to encounter each other frequently and therefore have higher possibility to appear in the same position. Based on these perspectives social relationships of any form drive individuals to form communities. Common interests and family bonds cultivate relationships prominent to the creation of communities. Therefore in a social structure, the social community formation performed through different perspectives such as physical location, social relationship, interest similarity and social participation.

The main purpose of grouping nodes into community is to provide a more resourceful utilization of the network resources through spatial reuse to enhance availability, reduce overhead by providing services locally, and provide scalable architecture. The choice of a clustering algorithm affects the stability of clusters. Several algorithms have been proposed for grouping nodes into cluster in wireless networks. Since a community head (*cH*) is expected to function for longer durations once elected, we use a mobility based community formation algorithm based on social interest of nodes and nodes moving behaviour for cluster formation. In our algorithm, a node is elected as a community head (*cH*) only if the node's mobility rate is below a certain threshold. The *cH*'s are coordinators in each community and store shared information. A border node is a node that can communicate with neighbouring community or the gateway or the path for routing the next community.

Each *cH* knows how many *cH* are available in the network. Every node periodically broadcasts a *Hello* message to maintain information about the cH and its neighbours, link status, and mobility rate.

The mobility rate measures the relative mobility of a node with respect to its neighbours. We consider a relatively dense network in our study. When a new *cH* is elected, it sends a message with its own *cH-id* informing other *cH s* that a new community head has been elected and the original *cH* is replaced. After the other *cH* s receive this message, they record the new *cH* in corresponding tables, delete the original *cH*, and then send a reply message with their *cH-id*.

The following chart shows the flow when nodes those have same interest and the same moving behaviour form the community.

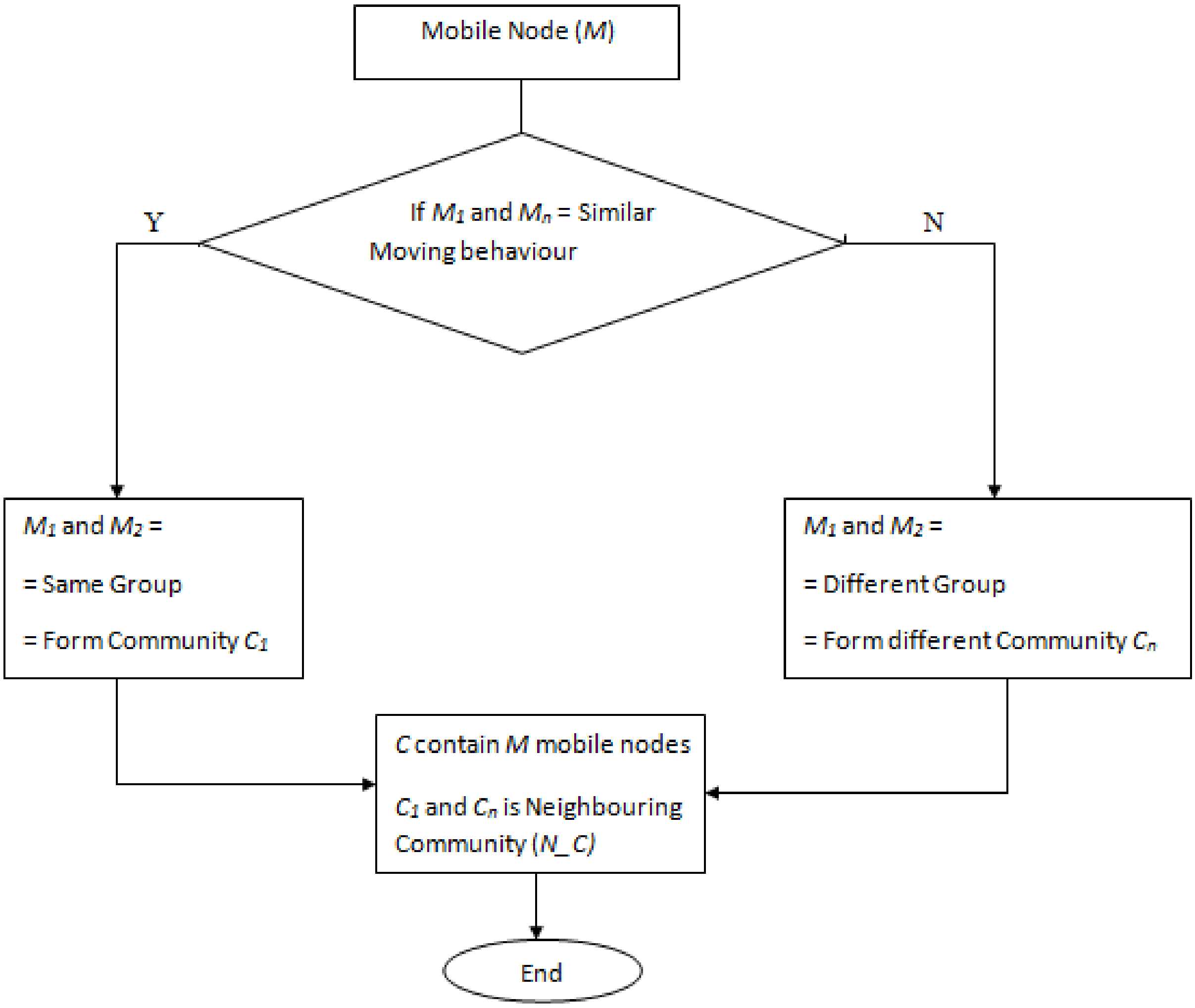


Figure 5: Flowchart of Grouping Similar Node into the Community

The following algorithm is the pseudo code of node creation and community formation in ASNET networks.

Algorithm1: Community Formation Algorithm in ASNET

01: Begin Community C =1 /* Represent community number one */

02: Load: mobile node M_1 from the network to Community C_1

03: Compute: M_1 with nodes $\sum_{i=1}^{n} M$

04: ***If*** M_1 and M_i have similar moving behaviour

05: Select: M_i for Community C_1

06: ***Repeat*** for mobile node M_n

07: Until: M_i nodes empty battery power

08: ***Else***

09: M_i doesn't enter C_1 , it compute with other M_i nodes and enter Community C_n

10: ***Repeat*** for other Community C_n

11: Output :Set of Communities

12: ***End***

Algorithm 1: Community formation Algorithm

4.3.2.2 Community Head Selection System

In ASNET since mobile nodes move freely, disconnections occur frequently and this divide the network into groups or regions. Therefore we create virtual backbone between delegate nodes in each group by forming Community. The community head is responsible for management of its groups, allows the reuse of resource which can improve the system capacity and it manage the network topology. The community head is selected from each community based on different parameters and manage data service until another community head is selected.

Community head is selected for each group to handle the control packets for replica query and update propagation. Community head or *cH* coordinates the node within the structure. All *cH* acts as a temporary base station within its community *C* and communicates with the neighbouring community heads or *NcH*. The *cH* election is the process of selecting particular node based on its parameters within community as a head node. Its responsibility is to manage data process and the nodes of its own community and to communicate with neighbouring communities. It can communicate with other communities directly through the respective *NcH* or through gateway nodes. The *cH* communicates with neighbour by sending and receiving data, compressing the data and transmitting the data to neighbouring community heads or *NcH.* Electing the specific node as *cH* is depend on the factors such as mobility of the node, centrality of the node, energy or power remaining of the node and free storage capacity of the node for data replica.

In our paper, we suppose that replicas refer to identical copies of a data and replica managers or community head are those mobile nodes that hold one or multiple replicas. Replica managers accept query requests of data replica allocation from community member mobile nodes and reply with the requested data. In community based replica allocation model, we assume each mobile node has unique address called *M-id* and is reachable through underlying adhoc routing protocols and also we assume each replica has an update count which increases when new update is committed on the replica. The update count play important role in recording the update status of each replica and helping to maintain the consistency of the replicas.

To achieve initial set of community heads, we propose election algorithm that consider combination of structure of nodes. First assume number of communities is set to *C* and all mobile nodes are aware of *C* and the total number of mobile nodes. All nodes in the community then flood their capacity including their current residual power, their free storage space and their node IDs. All nodes after receiving broadcast information, each mobile host put the *M-id* into sequential list sorted by their capacity. The M mobile nodes at the top of the list become the community head and their specific location in sequential list becomes their community head-id or *cH-id*. The number of communities remains constant during execution to maintain the stability of the replica information directory.

The following chart shows the flow when community head node which leads the network sub cluster is elected based of different parameters.

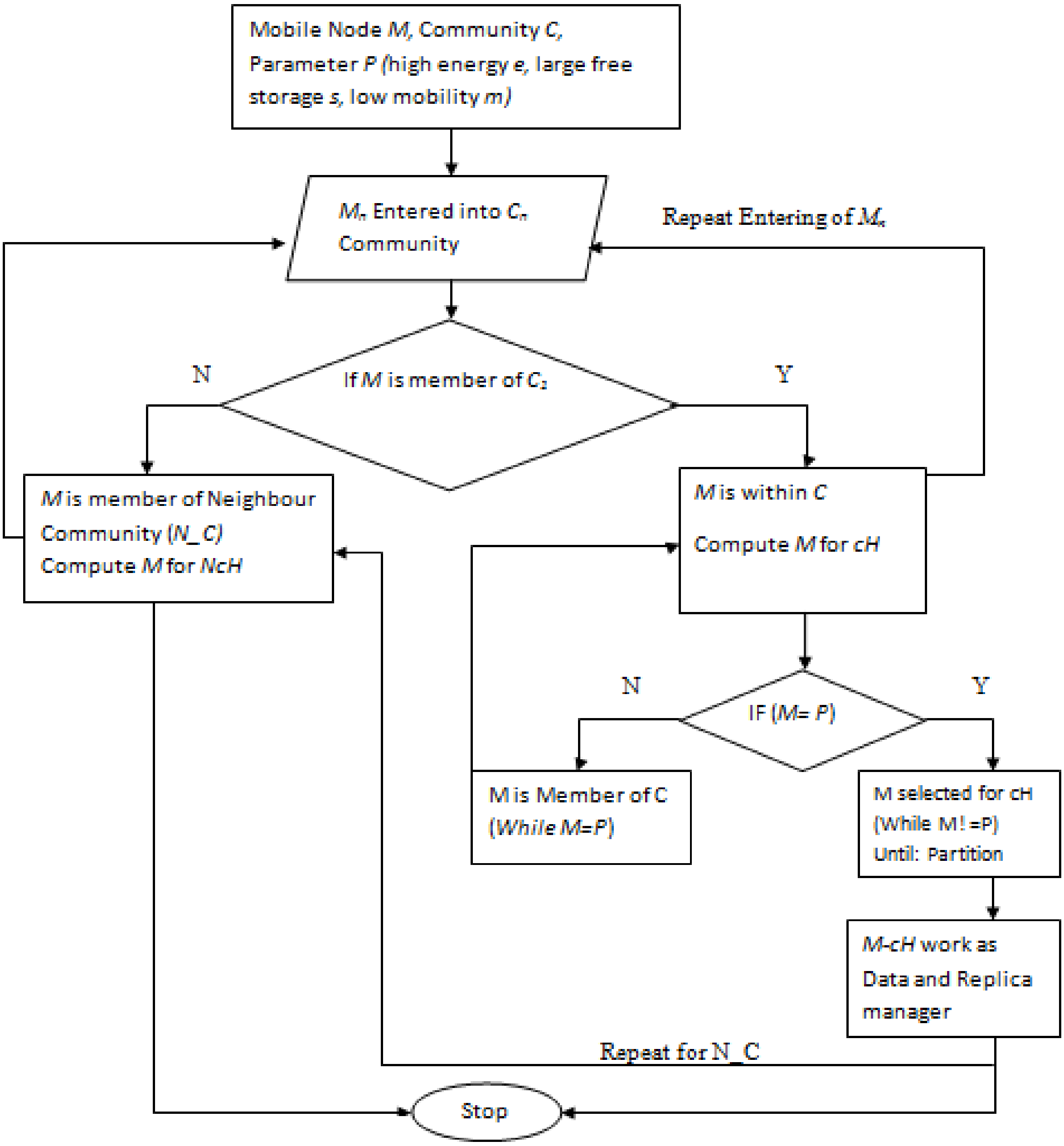


Figure 6 : Flow of Community Head Selection

Based up on Raihana F. et al.[21],and Sun B. et al [20] papers we add social relation ship and social parameters to select community head in ASNET RepCoH system. Therefore we develop new algorithm by adding parameters in the proposed algorithm to maximize data availability.

Algorithm 2: *cH* Election Algorithm

Input: M_i to M_n into C_i
Begin: $C_i = \sum_{i=1}^{n} M$
Compute: $\sum_{i=1}^{n} M$ for *cH*
do
    Broadcast_Election_Message (*M-id*, P *(e, s, m)*);
Start: *Timer_for_Recieving_Election_Messages ( ) ;*
Receive: Election_Message (*M-id*, P *(e, s, m)*);
Add: (*M-id*, P *(e, s, m)*) into Sort_List;
if (Timer_For_Election_Message_Expire ( ))
    M_i Find_Location_of_List (Own *M-id*);
if (*M-id* =Top of Sort_List)
      M_i =*cH ;*
      Flood *cH* Announce (*cH-id*) to All M_n within C_i
else
    M_i = Host Node ;
    Until
        Re-Election;
Repeat: For M_i to M_n in C_n ;
Output: *cH* Elected for data Handling ;
End;

Algorithm 2: Community Head Election Algorithm

4.3.2.2.1 Community Head Re-Selection System

Community Head is critical for implementation of replica management in RepCoH, hence we propose community head re-election algorithm for the purpose of picking new node for handling replica to replace failed community head. All mobile nodes within community are candidates to replace the head based on their parameters and they use a timer to closely monitor the status of the community head. The timer expires when *cH* doesn't hold replica and give permission of access for replica of data of the nodes and when it fail to manage data service within and neighbour community. When the timer expires, each candidate sends detection messages to the community head. If current *cH* don't respond reply message for candidates, assume *cH* is no longer available and initiate the election procedure by setting new timer and sending out an election message with its *M-id* and *P (e, s, m)* to other potential candidates. Based on Community Election Algorithm the candidate node at the top of sort list announces it self as *cH* and synchronizes with *N-cH* by broadcasting a query message. Upon receiving query messages, the *N-cH* updates their Replica Table with newly elected *cH-id* and share copy of replica of the nodes exist in newly elected *cH* Community *(C).*

Algorithm 2: *cH* Re-election Algorithm

```
Initial: M1 to Mn in Ci
Compute: Ci = ∑_{i=1}^{n} M for Re-election of cH;
Wait: cH permission for data read and holding replica;
Until : Timer Expire
if (Timer Expire)
    Send: Detection_Message (cH);
    Start: Timer_for_Detection_Message ();
if (Timer For Detection Message Expire() )
    Broadcast_Election_Message (own M-id, P (e, s, m) );
    Start: Timer_for_Recieving_Election_Messages ();
Receive: Election_Message (M-id, P (e, s, m));
    Compare: Sort_List
if ( Top Sort_List = own M-id )
    declare: New cH (own M-id);
else
    declare: New cH (other M-id at top Sort_List );
End;
```

Algorithm 3: Community Head Re-Election Algorithm

4.3.2.3 Proposed RepCoH Model

In RepCoH replica data of the member nodes in the communities are stored on the community head of the network. This is based on member nodes *cH* request for their data allocation and *cH* permission for data holding. The requirement RepCoH data replication is to minimize bandwidth and resource utilization, to minimize number of routes for data transfer i.e. if node *M1* want to send data to *M2* it send simply through *cH* since its data is exist on the *cH* and to access data if partition and also if the nodes in the community loss their battery they access their data when recharge.

Suppose adhoc social network as shown in figure below. ASNET network has no fixed infrastructure and the nodes with similar moving behaviour move together freely anywhere. Because of the mobility of nodes the network topology is dynamic and temporary or have small lifetime. Each node in this network has unique identification. All mobile nodes in the system denoted by the set $M = \{M1, M2, M3 \ldots M_n\}$, where n is number of mobile nodes in the network and $Mi\ (1 \leq i \leq n)$ is Mobile node id or *M-id*. The data is associated with mobile-id and the set of all data items is denoted by $D = \{D1, D2, D3 \ldots D_n\}$, where n is number of data items and $Di\ (1 \leq i \leq n)$ is Data id or *D-id*.

The nodes in adhoc social network those have similar interest and similar moving behaviour form the community based on community formation algorithm. The election of community head of the nodes is the basic in community head based replica allocation system. To elect community head or *cH*, we use highest degree leader election algorithm which means the nodes have high parameters elected for community head to hold and copy replica of data to other community and to control data management in the network. First we assume the number of community in the system is set to *C* where $C= \{C1, C2, C3 \ldots C_n\}$, where *Cn* is total number of community in the topology and all nodes in each community aware of *C* and the total number of nodes. Every node in the community then flood their information including their centrality within community, their free storage capacity, their current residual energy, their *Mobile-id (M-id)* and their mobility. Then after all nodes receive the broadcast information of all mobile hosts, each of them puts their parameter information with their Mobile-id into sequential list sorted by their capacity plus mobility plus battery power. Then the *C* nodes at the top of sequential lists become community head or *cH* and identified with its *cH-id.* The procedure is the same for another community *Cn.* Then other nodes remain normal nodes during execution until another re-election of community head. Each node elected for lead in each community will maintain *Replica Table or RT.* The RT will contain the information about replication that is available with different node in that community and entries for new replica for nodes exist in the community and identification of neighbour communities head or *NcH-id* for transmitting copy of replica of nodes within community to leader of neighbour community. The entries related to each node are *M-id*, *D-id,* and RT also register *NcH-id.* Then the replica propagation of data between nodes those are data owner, community head nodes and neighbouring community head will proceed.

Since the data replicated is data for communication in infrastructure-less social network replica of large capacity of data is very rare and since the existence of the topology is temporary because of the network is adhoc, the time when the memory of cH is full will be slow. But in some case if the storage and battery power of cH will be lost or if cH move to other community, re-election of cH will be preformed and until the election if the nodes within community need their data they access from the neighbouring community head. The security of data between communities head of the system are protected by encryption of data algorithm.

The following chart shows the flow when replica of the packet is allocated on community head node which manage data in the network.

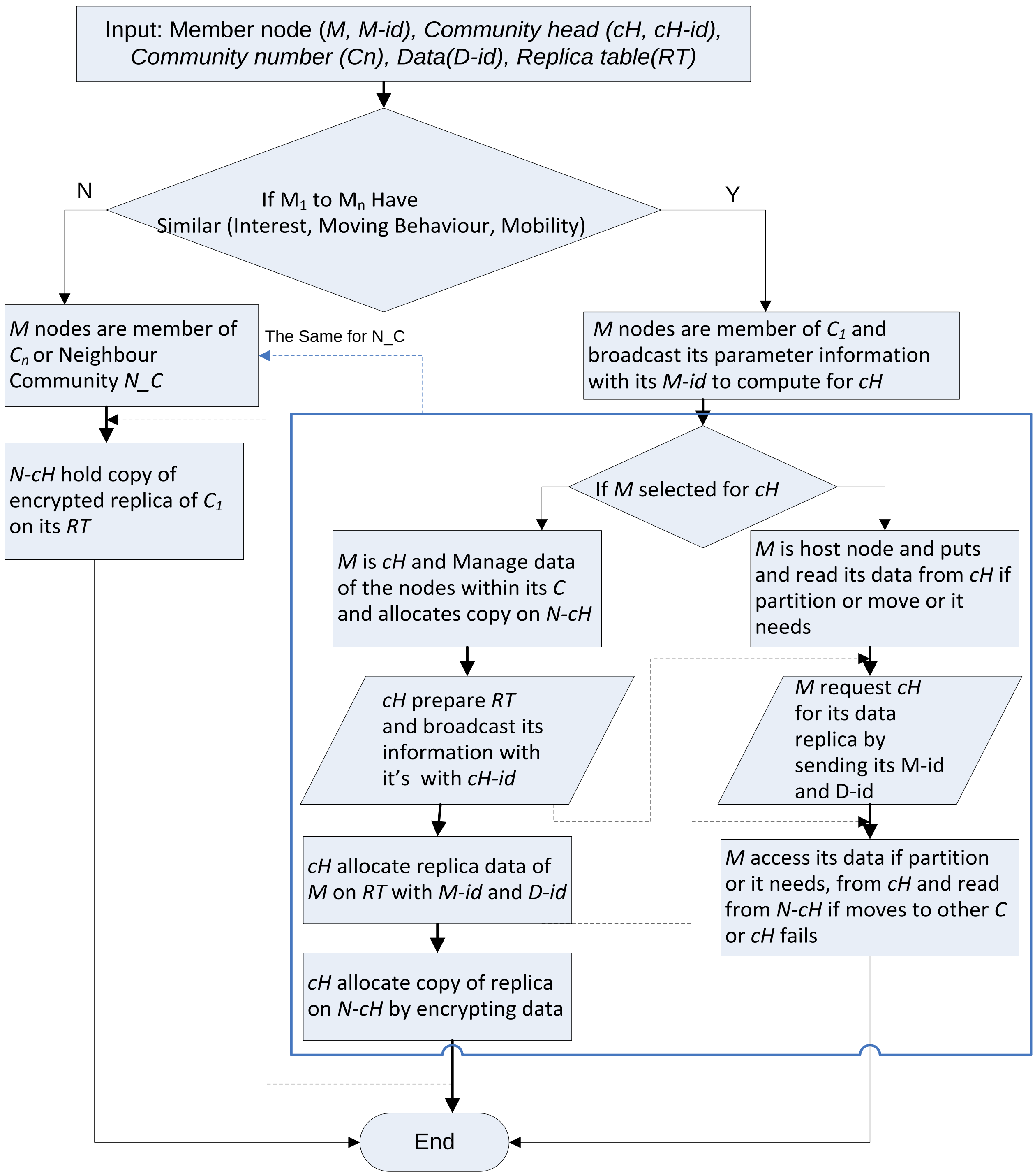


Figure 7: Flow of Replica allocation and Replica Access in RepCoH

Algorithm 2: RepCoH Data Replica Algorithm

```
Input:  Mi to Mn into Ci , cH, NcH, D;
Begin:  M-cH are cH of Ci and M-cH are NcH of Cn;
Register: D of Mi with its D-id;
do
    cH Broadcast_RT with its cH-id ;
Receive:  Mi to Mn  store cH-id() on their list ;
Response: Mi to Mn  response Positive ACK with their M-id();
then
   cH register M-id() and NcH-id() on its RT;
If (Mi request cH for Replica)
    then
       cH check its Replica table with Mi Did();
   If (available RT on cH and Did() not register before)
     then
        cH Hold replica data of Mi with its D-id();
        cH allocate copy of replica on NcH with its cH-id();
else
   Update replica of Did of same Mi;
    While (TTL for Did expires)
 Repeat: For Mi to Mn managed by cH in Cn ;
Until
   Mi partition from topology and Re-election of cH;
Output:  Allocated data on  cH and copy of replica on N-cH ;
End;
```

Algorithm 4 **:** RepCoH Data Replication Algorithm

In order to illustrate RepCoH basic operation, we presented a network composed of 33 nodes. We are given a social data graph and need to replicate its user data which have been partitioned across the storage space of the community. This case applies an ASNET in its early stage where the network size is considered small (small-scale networks). Let we assume as an example 33 nodes in our system which grouped into three communities C1, C2 and C3. i.e. nodes 1 to 11 is member of C1 ,nodes 12 to 22 is member of C2 and node 23 to 33 is member node of C2 and let M1 is cH of C1,M12 is cH of C2 and M23 is cH of C3. Assume M4 of C1 want to allocate its data, then by using M-id and other parameters of M-id and cH-id cH of C1 hold data of M4 and replicate its copy on cH of C2 and cH of C3. Then if M4 or other M want to access the data of M4 they access from cH of C1 or else while M4 partition from the community it access the copy of its data from cH of C2 and C3 communities.

4.3.2.4 Consistent Data Access

In RepCoH replica allocation mechanism data access requests are generated using AODV routing protocol. When host node request to access for its particular data item, the replica manager or *cH* check its memory and compare its replica table with requester data id and node id and then if replica of data exist in replica table, the *cH* give permission of data access for the host node otherwise the *cH* response as TTL of data replica expire for host node *M*.

In our proposed *RepCoH* algorithm we maintain the consistency of allocated data item at *cH*. When host node sends its updated data for replica to *cH*, then community head *cH* verify the update data with its data id *D-id*. The validation of data update message for data item *Di* from node having *M-id* is verified by cH based upon the ownership of the original copy of data item *Di* by node *Mi*. If *Mi* is owner of original copy then it is valid update for data item *Di* otherwise not. After verifying the validity of update message for data item *Di* Community Head broadcast the updated copy of the data item with its *D-id* to neighbour community head *N-cH* s. Now whenever a neighbour community head receive an updated copy of the data item with *D-id*, then *cH* check whether this *D-id* is available in its Replica Table *RT*. If desired *D-id* found in *RT*, then update this data item *Di*. If desired *D-id* is not found in *RT*, this means that this whole community not having a replica of *Di*. So community head create the replica of the *Di* with new *D-id* in its *RT*. In this way our system verifies the consistency of data of mobile nodes at community head and neighbouring community heads. But since our data is social network data such as calendar, meeting alert message etc which processed on adhoc network, the probability of consistency requirement is very low.

4.3.3 The Role of Social Community, Network Routing, Mobility Models and Data Link Layer in RepCoH

4.3.3.1 Social Community

The Social community can be created based on similar characteristics of individuals, including physical and social characteristics. It is the structural sub unit of a social network with high density of internal links. Individual nodes have more social connections with other individuals inside their own community than with individuals outside the community. The social connections may be family, friends, common location or common interest, which are decided by the social graph.

Social Community shows the small world phenomenon that node encounters are sufficient to build a connected relationship graph which is an abstract graph where vertices represent individual peoples and the edges describe the social ties between the individual people. With the social graph of the ASNET network, a variety of social metrics for instance social relationship, communality, centrality and similarity can be easily calculated and used by social based approaches. Therefore it is important to obtain social graphs for social based network such as adhoc social network [4], [6]. In social networks, each node represents an individual or an organization and each node is linked by some interdependency such as friendship, communication, common interest, common location, family etc. These social connections represent relationship between users. The social community of ASNET network is dynamic in nature, therefore it enables users with the common interest to communicate and share common information in the dynamic environment.

4.3.3.2 Network Routing

In wireless network packets can generally classified into data packets and routing packets. The data packet is the packet of the data file which is created by transport layer agents. The routing packet is created by routing agents. Routing packet flow mechanism begins from within a routing agent during a route discovery process. After the inception, a routing packet is configured according to the AODV protocol.

Therefore, in our system we use distance vector protocol called AODV which is very simple, efficient and effective for our replica allocation algorithm. It is reactive protocol which enables multi-hop routing between participating mobile nodes to establish and maintain adhoc network. AODV only requests route when needed and doesn't require nodes to maintain routes to destinations that are not actively used in the communication. It keep destination address, destination sequence number, hop count, next hop, life time, and active neighbor list and request buffer for each route table entry. It build routing table when node needs to communicate with another node and when it send data to other node it first broadcast Route Request (RREQ) packet to destination until it received in order to avoid routing loop generation. It is designed to solve the network loop problem, where routing packets circulate indefinitely. AODV solves this problem by discarding packets with stale sequence number. In particular, every Mobile Node maintains three sequence number counters for three types of packets: a destination counter for RREP, a broadcast counter for RREQ, and a neighbor probing counter for HELLO [16].

In AODV, each route entry contains the following fields:

- Destination address, the next hop node, and the metric (i.e., the number of hops to the destination)
- Sequence number corresponding to the destination which helps prevent a so called routing loop problem
- Active neighbors on this route entry
- Expiry time which indicates the duration where this route entry is considered fresh.

AODV defines three main types of packets. Those are *Route REQuest (RREQ)* which originated and broadcasted to every neighbor of a source node during a route discovery process. RREQ contains information such as Source address (src), Destination address (dst), Broadcast ID (bID), Source sequence number (SN s), Destination sequence number (SN d), Number of hops to destination (hop cnt).The second is Route REPly (RREP) which is a packet replied by a node. It contains routing information for the destination specified in an RREQ. RREP contains the information such as Source address (src), Number of hops to destination (hop cnt), Destination address (dst), Destination sequence number (SN d), Time where this entry is considered valid (t exp). And the third is HELLO which is a special unsolicited RREP packet. It probes neighbors within locality. HELLO contains only two pieces of information called address and sequence number of the sender [16].

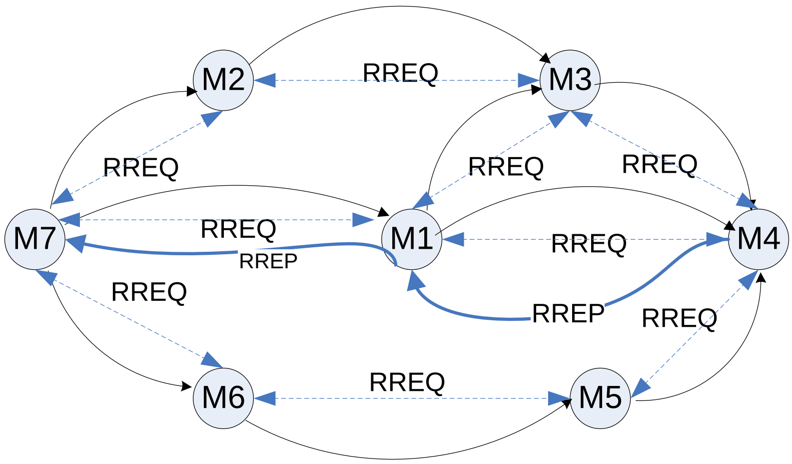


Figure 8: AODV route discovery of mobile nodes within the communities [16]

4.3.3.3 Mobility Model

In wireless network, especially in adhoc social network the big problem come up the replica is the mobility of node, since the node are mobile all the time. The bandwidth reservation made or control information exchange may end up being of no use if node mobility is very high. The MAC protocol has no role to play in influencing the mobility. The MAC protocol design must take this mobility factor into consideration such that the performance of the system is not significantly affected due to node mobility. Mobility allows ASNET created on the fly in any situation where there are multiple wireless devices. This gives flexible ad hoc can be temporarily setup at anytime in any place, lower getting started costs due to decentralized administration and the nodes in ad hoc network need not rely on any hardware and software. So, it can be connected and communicated quickly.

On our proposed system each node moves according to the Random Waypoint mobility model when they are created [37]. This model is widely used in ASNET simulations as it is considered to be the closest to typical movement patterns of real mobile nodes. Random Waypoint breaks down the movements of a node into two types of phases, pauses and motion periods. In a pause period, a node stays at its current location for a certain time. In a motion period, a node moves toward a new random destination at a speed selected in a predefined interval. The destination, speed and direction for each node are all chosen randomly and independently of other nodes. In our simulation, the moving speed of each node was chosen randomly in the range 0 typical range of walking speed of human beings; pause times were set at 3 seconds. The initial position of each host was set randomly. And we use Random point group mobility (RPGM) when community is formed and during system work .Here each group has a logical centre or group leader that determines the group's motion behaviour. Initially each member of the group is uniformly distributed in the neighbourhood of the group leader. Subsequently, at each instant, every node has speed and direction that is derived by randomly deviating from that of the group leader.

4.3.3.4 Data Link and Physical layer

The data link layer exist at lower layer of our system which is responsible for establishing reliable link, error control and security can model a link bandwidth, propagation delay and packet framing for instance sequence number and acknowledge number. It bridges the routing layer i.e. routing agent to the MAC layer. It has two sub layers called Medium Access Control (MAC) and Logical Link Control (LLC). A MAC protocol defines communications rules to which Mobile Nodes fulfill with to access a shared medium. It is responsible for channel sharing, allotting time slots, frequency or code space among them. The role of data link layer is important in multiplexing, frame detection, MAC and error control to ensure point to point and point to multipoint connections for reliable communication. Link layer also handles rate, power, and source/channel coding. The main objective in MAC protocols is to optimize spectral reuse in order to maximize the channel utilization.

Other layer at the bottom contains physical network interfaces such as hardware (e.g., radio modem, antenna) which creates and sends out data bits and channels which model the medium shared by all Mobile Nodes. On the sender side, a higher-layer Mobile Node component (a Mac object in most cases) sends a packet to the physical network interface. The network interface sets up physical layer parameters and forwards the packet to the channel. On the receiver side, the process starts when the channel receives a packet via its function and the direction of incoming packet is upward. In wireless physical layer effects include interference, noise, signal reception, fading, and path loss, which usually have an impact on routing protocols as well as on the performance of other higher layers. In ad hoc network physical layer handles data reception, modulation, data encryption, power management. Since there is no centralized mechanism to govern the ad hoc network, nodes are solely responsible to use their resources in the most efficient manner, therefore, power management is of pivotal importance.

4.3.3.5 Security

Security is very important in wireless communication. In adhoc networks the problem of security is complicated because of absence of infrastructure and dynamic network topology. The lack of centralized control makes the system more vulnerable to attacks and threats. Since the nodes can enter and leave the network structure at any time, it makes difficulty to identify malicious node which re-joining the system and launch attacks from different locations repeatedly. Although the wireless channel, dynamic network topology, the mobility of nodes and the dependency on routing protocols all makes the ASNET data management more prone to attacks. Therefore concerning security protection such protecting from malicious users, protecting the availability, authentication, confidentiality, integrity, non-repudiation and routing protocol protection is important for securing our system.

The issue of adhoc social network security should be solved by protecting the MAC layer, encrypting data, verification of authenticity, making sure distributed public and private keys are secure and different security levels at every layers and also nodes builds its trust chart based on the nodes it trusts when communication with the node, assists in solving security issues. Another security measures to protect ASNET network secure is by using secure routing protocol which is called Secure Efficient Adhoc Distance Vector (SEAD) where source selects a random seed and sets the maximum hop count, and since security of adhoc depends on number of neighbours in the community and maximum key length, frequently changing the key also helps in improving security by making it less vulnerable.

In general when we secure ASNET data replication with public key security management, we ensure each node keeps exchanging information with other nodes on which group it belongs to, each node is able to monitor the behaviour of its group mates and obtain their public keys and each node keeps a trust table for storing trust values of other nodes.

CHAPTER FIVE

SIMULATION RESULTS AND EVALUATION

5.1 Simulation Environment

This section discusses the simulation and the performance of our algorithm. Various simulation environments were used for evaluation and implementation of our model. The network simulation is the most useful and common methodology used to evaluate different network topologies without real world implementation [38]. There are different wireless simulators such as QUALNET, OMNET++, NS-2, NS-3, OPNET, J-SIM and etc. The simulators are selected based on its language support, platform support, its licensing, GUI support, animation support and etc.

The model of our simulation is constructed on the basis of open source event-driven simulator called NS2 (version2). It is an open source network simulation tool. NS2 is a simulation tool designed specifically for communication networks. The main functionalities of NS2 are to set up a network of connecting nodes and to pass packets from one node to another. A network object is one of the main NS2 components which are responsible for packet forwarding.

Network Simulator version 2 widely known as NS-2 is a discrete event driven network simulation tool for simulation of network to study the dynamic nature of network [39]. It is an open source solution implemented in C++ and Otcl programming languages at UCB (University Of Carolina Berkley).NS-2 provides highly modular platform for simulation of wired as well as wireless networks. NS-2 supports different network components, protocols like TCP, UDP, FTP and traffic sources like CBR etc. Results of the simulations are provided with in a trace files that contains the entire occurred event during the simulation. The output of simulation is visualized through Nam (Network Animator). Otcl/Tcl is a scripting language with very simple syntaxes which allows fast development. Otcl in NS-2 is used to specify the parameters for simulated networks like node, links, protocols, topography and so on. C++ objects oriented language is used for byte manipulation, packet processing and implementation of new algorithm or protocol in NS-2. Otcl and C++ interact with each other through TCL/C++ interface called as TclCl [39].

It is an object oriented, discrete event driven simulator written in C++ (i.e., a backend) and Otcl (i.e., a frontend). The primary use of NS is in network researches to simulate various types of wired/wireless local and wide area networks; to implement network protocols such as TCP and UPD, traffic source behavior such as FTP, Telnet, Web, CBR and VBR, router queue management mechanism such as Drop Tail, RED and CBQ, routing algorithms such as Dijkstra, and many more.

Ns2 is written in C++ and Otcl to separate the control and data path implementations. The simulator supports a class hierarchy in C++ (the compiled hierarchy) and a corresponding hierarchy within the Otcl interpreter (interpreted hierarchy). The reason why ns2 uses two languages is that different tasks have different requirements: For example simulation of protocols requires efficient manipulation of bytes and packet headers making the run-time speed very important. On the other hand, in network studies where the aim is to vary some parameters and to quickly examine a number of scenarios the time to change the model and run it again is more important.

In ns2, C++ is used for detailed protocol implementation and in general for such cases where every packet of a flow has to be processed. For instance, if you want to implement a new queuing discipline, then C++ is the language of choice. Otcl, on the other hand, is suitable for configuration and setup. Otcl runs quite slowly, but it can be changed very quickly making the construction of simulations easier. In ns2, the compiled C++ objects can be made available to the Otcl interpreter. In this way, the ready-made C++ objects can be controlled from the OTcl level [39].

After simulation, NS2 outputs either text-based or animation-based simulation results. To interpret these results graphically and interactively, tools such as NAM (Network Animator) and Xgraph are used. The result of the simulations is an output trace file that can be used to do data processing (calculate delay, throughput etc) and to visualize the simulation with a program called Network Animator (NAM). NAM is a very good visualization tool that visualizes the packets as they propagate through the network. An overview of how our proposed solution is implemented in NS2 is shown in the figure below.

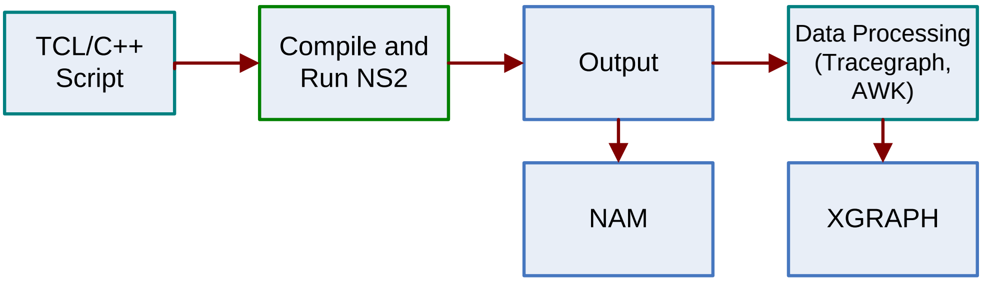


Figure 9: The Implementation view of our Proposed Model on NS2 [39]

5.2 Simulation Model

The simulation experiment is carried out in LINUX (UBUNTU distribution). The detailed simulation model based on network simulator-2 (ver-2.35) is used in the evaluation. The NS2 instructions are used to define the topology structure of the network, the motion mode of the nodes, to configure the service source and the receiver, to create the statistical data track file and so on. The studied scenario consists of 20 to 100 mobile nodes. In our simulation setting, a group of mobile nodes spread randomly in an area of 2000*2000 square meters. Each node is equipped with a radio transceiver capable of transmitting signals up to approximately 230 meters over a 2 Mbps wireless channel, using the two-ray path loss model without fading. All simulations are run for 10 to 300 seconds of simulated time. All mobile nodes are constant bit rate traffic sources. They are distributed randomly within the mobile ad hoc network. The sources continue sending data until one second before the end of the simulation. The radius of a *cH* is set to one community.

The Ad-hoc On Demand Distance Vector (AODV) routing protocol is employed to carry out the unicast routing operation. A node moves according to the random waypoint model , in which each node selects a random destination in the specified area and moves toward the destination with a speed selected randomly from the range varies from 2 to 20 m/s, with the default 1 m/s.

We implement our proposed model by using NS2 simulation software i.e. first we install the software package on Ubuntu operating system then install and configure packages such as NAM, Trace Graph, Xgraph and different protocols used in our model. Then we write different source code to implement our proposed model by using language TCL and C++ and also the AWK script for performance evaluation of our system. At the last we generate the graph from the simulation result which we call tracegraph by using Xgraph and we convert it into line chart for better description and neatly explanation of the model.

5.2.1 Simulation Parameters

We simulate the ad hoc social network using NS2 simulator version 2.35 having following parameters. Simulation process and results analysis first of all, to set the topology and the configuration of nodes properties, and also properties of MAC layer for some address type, protocol type, channel type, simulation time and transmission way of wireless. In our simulation, we used topography size 2000 m x 2000 m, number of wireless nodes 20 to 100 nodes with maximum moving speed 20 m/s. We did the Simulation for 10 to 300 sec. with maximum 6 connections at a time allowing UDP traffic. The network simulation parameters we have used for our simulation purpose shown in the following table 2.

Parameters	Values
Channel Type	Wireless Channel
Radio Propagation Model	Two Way Ground
MAC Type	802.11
Antenna Model	Omni Antenna
Maximum Packet	500 bytes
Number Of Mobile Nodes	20 to 100
Routing Protocols	AODV
Dimension of Topography	2000*2000 meters
Time Of Simulation End	10 sec to 300sec
Transmission range	250 m
Traffic type	CBR(Constant Bit Rate)
Number of Replica	1 to 25
Maximum speed	20 m/s
Maximum Replica Time to Live	1min

Table 2: Simulation Environment Parameters with their Values

5.2.1.1 Simulation Parameter Description

In our simulation we use wireless channel since our network is ad hoc network and Two Way ground radio propagation model that predicts path loss when the signal received consists of the line of sight component and multi path component formed predominately by a single ground reflected wave. We also used IEEE 802.11 specification for implementing our network and Omni directional antenna which radiates radio wave power uniformly in all directions in one plane, with the radiated power decreasing with elevation angle above or below the plane. Another parameter we used in our simulation are constant bit rate for traffic generation and different protocols such as on demand routing protocol called AODV, group mobility model called RPGM and we simulate our scenario within 2000 m^2 and compare with different simulation time and number of nodes to get optimum results.

5.2.2 Prototype Implementation

In Adhoc Social Networks the movement of the network nodes may quickly change the topology and results link failure and partition between the nodes. In data management of ASNET keeping the nodes data at node partition or when nodes lost its battery power is great issue. Therefore to control data management applying community head based replica allocation or RepCoH is important to maximize data accessibility between nodes. Figure 10 shows the simulation of 21 mobile nodes using NS2 with their transmission range. Nodes in our simulation scenario can communicate with each other by forming the community based on their mobility group and data between nodes managed by community heads elected from the community members. Then Random traffic connections of UDP and CBR are setup between mobile nodes using a traffic-scenario generator script. The CBR and UDP traffics connections are created between wireless mobile nodes. To create a traffic-connection file, the type of traffic connection, the number of nodes and maximum number of connections to be setup between them, a random seed and in case of CBR connections, a rate whose inverse value is used to compute the interval time between the CBR packets are defined.

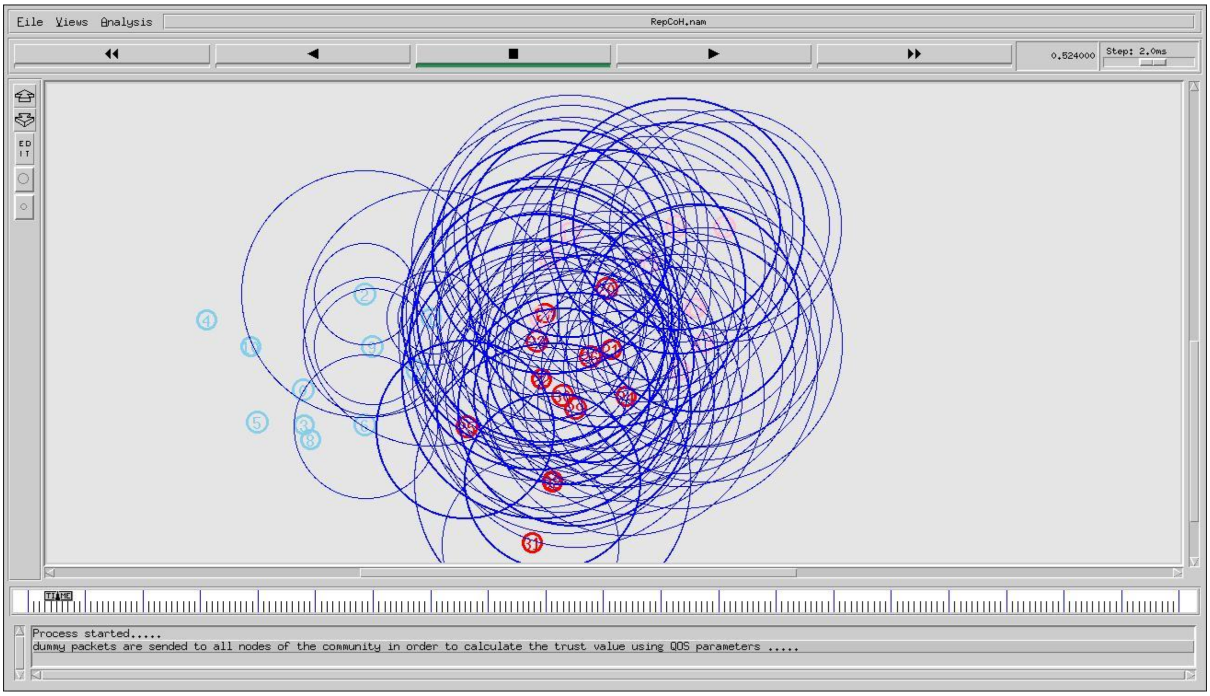


Figure 10: Simulation Screenshot of when Community is formed based on mobility

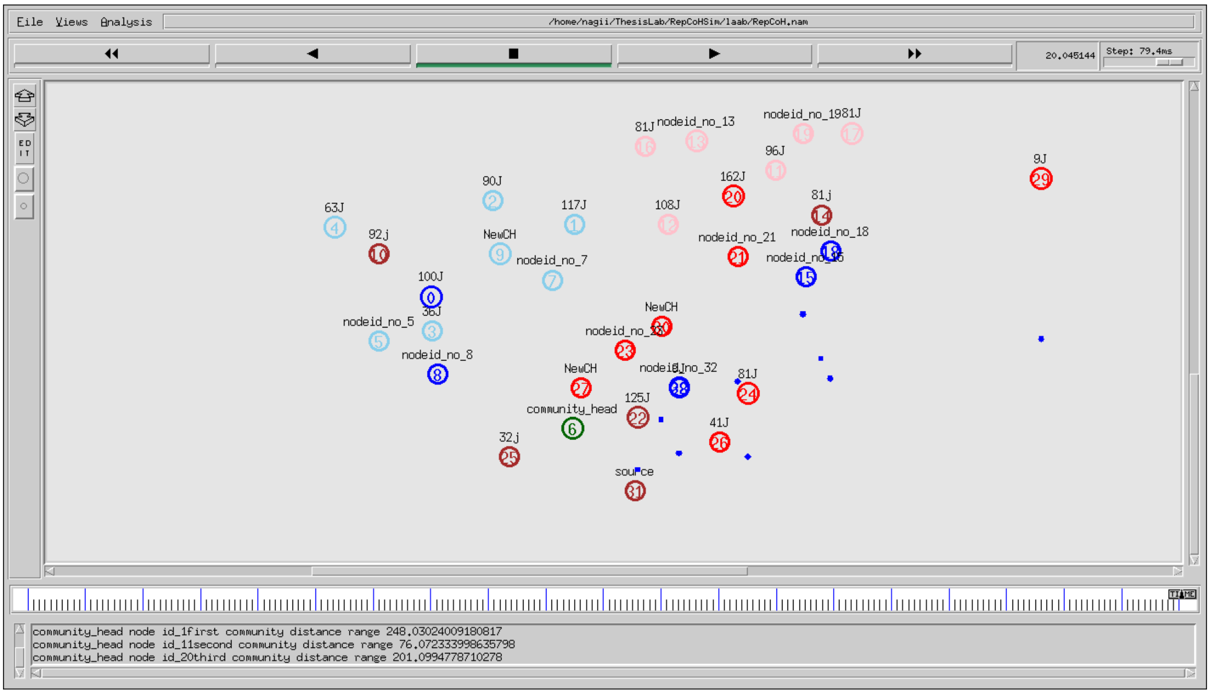


Figure 11: Simulation Screenshot of when Community Head is selected based on parameters

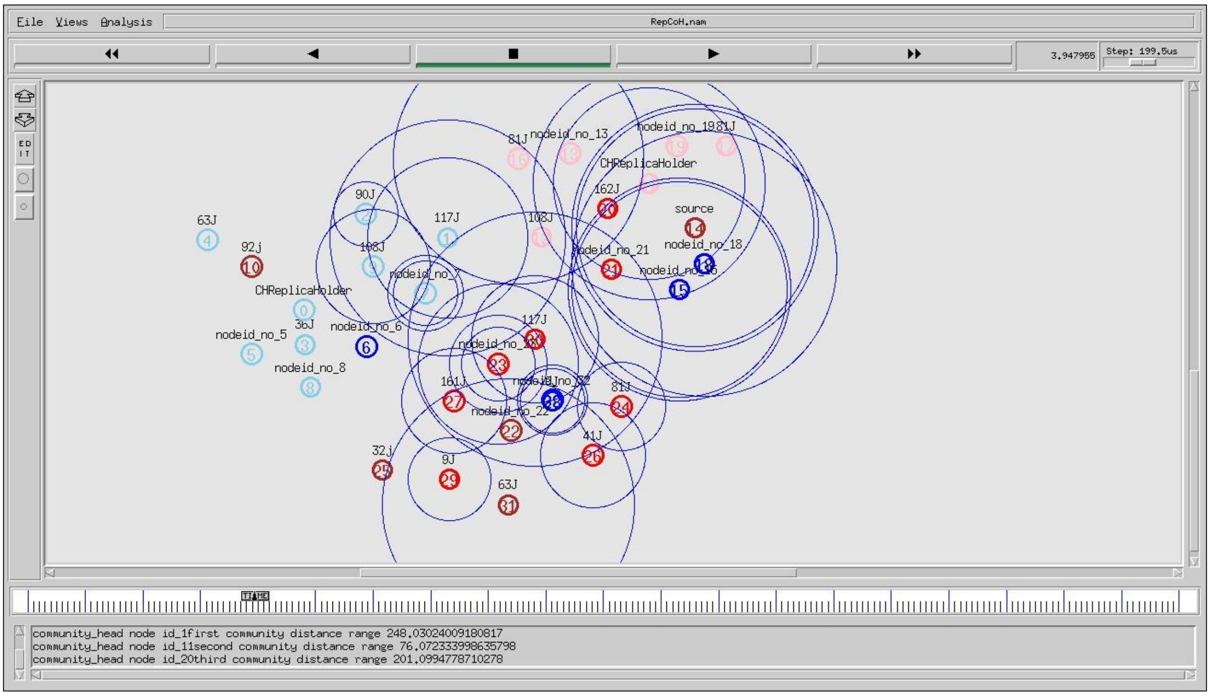


Figure 12: Simulation Screenshot of when RepCoH replica allocation works

5.3 Simulation Results and Analysis

The simulation results are obtained under several experiments. The result for proposed work has been compared with the results of ComPAS protocol and NoReplica using Xgraph. Xgraph has been used to conduct qualitative analysis. The parameters under consideration are Average Packet delivery ratio, End-to-End Delay, Throughput and Packet Loss. The following graphical analysis shows the performance results of NoReplica, ComPAS and RepCoH protocols.

5.3.1 Performance Metrics

These metrics are interesting because it can be used to point out what really happened during the simulation and provide valuable information about proposed system. In our system we propose different performance evaluation metrics. In the following sections the metrics are described. We have used packet delivery and data accessibility ratio, data end to end delay, data loss ratio, and throughput for evaluating the performance of RepCoH and other algorithm.

5.3.1.1 Average Packet Delivery Ratio

The Average packet delivery ratio presents the ratio between number of packet sent from application layer and the number of packets actually received at the destination nodes. It is defined as the ratio of total number of packets received by the destination which means the community heads and the total number of packets sent by the source which means the member nodes. The delivery of packets was investigated as the function of the speed of the mobile nodes, mobility of nodes and the size of the community.

In our simulation work, data accessibility is defined as the ratio of successful data requests to the total data requests made by a node. If the node receives its data item requested from community head or vice versa within a specified time threshold the request is deemed successful otherwise the request is considered a failure. The data accessibility was investigated as the function of the number of communities, community head changes, the transmission range, update interval of data and the size of networks.

In our scenario it is the ratio of number of data packets delivered to community head to the number of data packets supposed to be delivered to the receivers. This ratio represents the effectiveness of data accessibility.

$$\boldsymbol{Average\ PDR} = \sum_{k=0}^{n} \left(\frac{Total\ Packet\ Delivered}{Total\ Packet\ Sent}\right)$$

(Equation 1: Average Packet Delivery Ratio)

Figure 13 shows the Xgraph for RepCoH and ComPAS with pause time set to 25ms. The X-axis of the graph indicates the time and Y-axis shows the Packet Delivery Ratio. Here the PDR of ComPAS will decrease and PDR of RepCoH will increase when simulation time increases.

The following graph shows the average packet delivery performance of RepCoH comparing with ComPAS and NoReplica. We get maximum packet deliverence from RepCoH protocol than ComPAS and No Replica sinc commuity head manage the data. We compare this with the maximum bandwidth used and diffirent number of nodes.

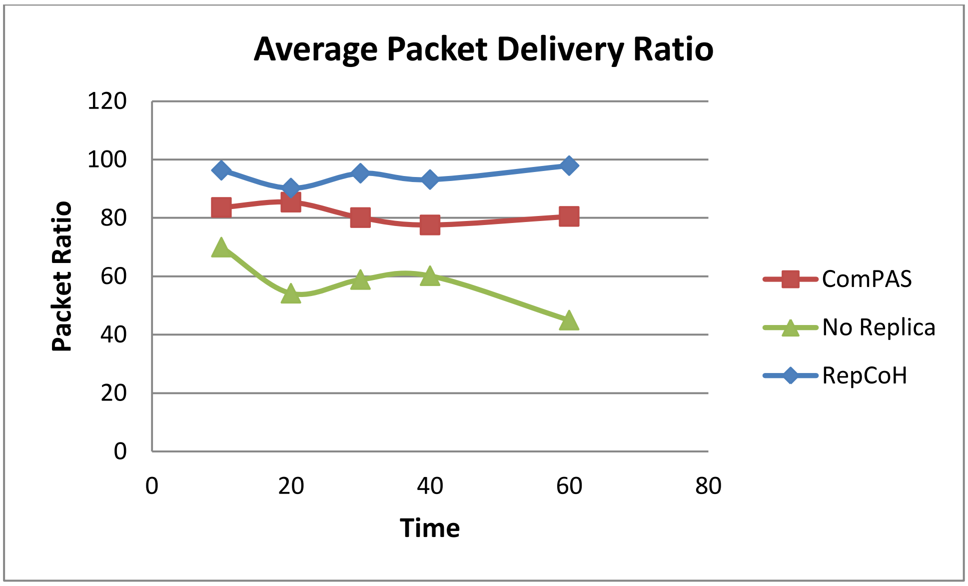


Figure 13: Average Packet Delivery Performance

5.3.1.2 Average End To End Delay

The term and to end is used to measure an average of performance between nodes in an ad hoc social network. It involves source and destination and it is the total delay that data packets experiences as it are travelling through the ASNET network. The end-to-end delay is the time taken for a packet to be transmitted from the source to destination in the network i.e. the average response time for successful requests, from sending a request until receiving the response. This metric reflects the response latency of a caching system. It is measured in second. In the discussion that follows, the node local statistics for each designated workstations have been collected in our study in order to analyze the End-to-End delay performance. In NS2 terminology, the local statistics generally at the node level, act as a source of statistic wires. Such statistics are scoped to the process module in which a process operates. In other words, each separate instance of the process model can individually maintain a local statistic with the same name. These statistics vary independently over time, can be separately probed, and can be used to feed statistic wires. The delay is built up by several smaller delays in the network that adds together. These delays might be time spent in packet queues, forwarding delays, propagation delay or the time it takes for the packet to travel through the medium and the time needed to make retransmission if a packet got lost.

The average delay is calculated by taking time of last packet received or last packet time attribute of Loss Monitor Agent, subtracting the initial time of transmitting that packet and then dividing by the total number of packets sent. We achieve average end to end delay for packet sent from member node to replica holder or community head.

$$\boldsymbol{Average\ Delay} = \sum_{k=0}^{n} \left(\frac{Time\ of\ last\ packet\ Recieved - Time\ of\ Packet\ Sent}{n} \right)$$

$$*\ n\ is\ the\ total\ packet\ received$$

(Equation 2: Average Delay)

Figure 14 shows that the delay performance for ComPAS and RepCoH. The X-axis of the graph represents the time of simulation end and the Y-Axis represents the end to end delay. While comparing the RepCoH and ComPAS the ComPAS and NoReplica gives more delay result than RepCoH. The result depicts that delay for data transfer gets reduced by using RepCoH scheme instead of using ComPAS when the simulation time increased.

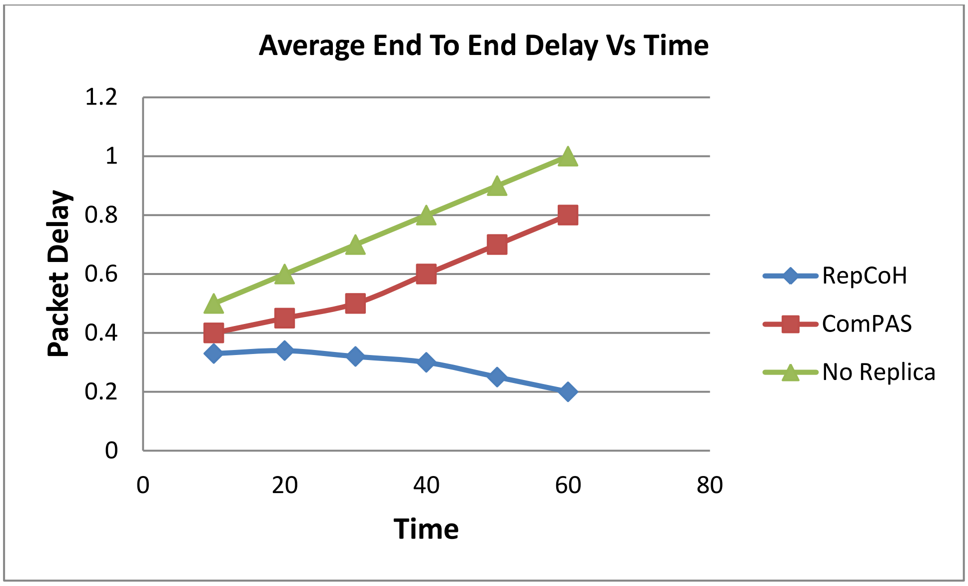


Figure 14: The Average Delay Performance

Figure 15 shows that the delay performance for ComPAS and RepCoH. The X-axis of the graph represents the effect of different number of nodes and the Y-Axis represents the end to end delay. While comparing the RepCoH and ComPAS the ComPAS and NoReplica gives more delay result than RepCoH. The result depicts that delay for data transfer gets reduced by using RepCoH scheme instead of ComPAS and NoReplica when the number of nodes increases.

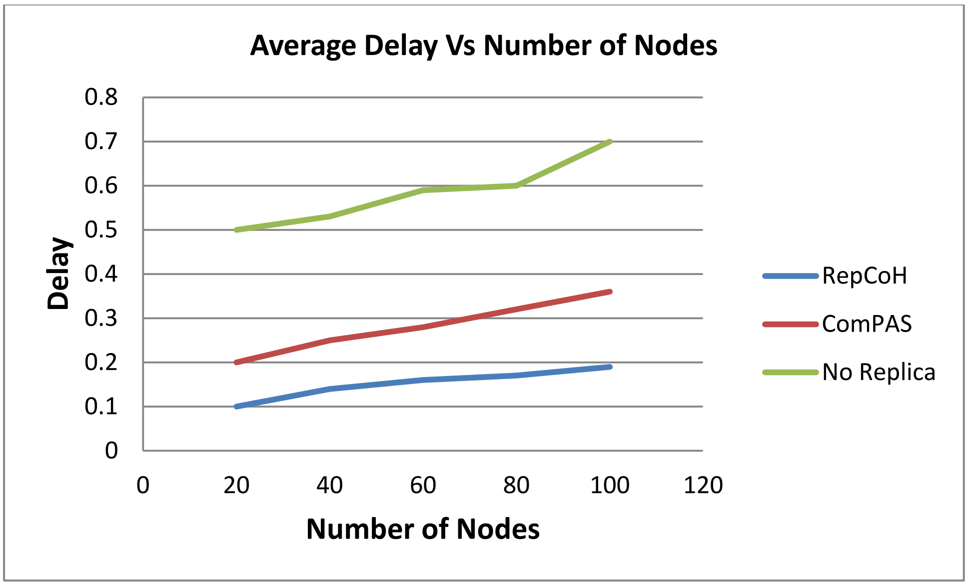


Figure 15: Delay Vs Number of Nodes

Figure 16 shows the results of Mobility Vs Delay. From the results, we can see that RepCoH scheme has slightly lower delay than the ComPAS and NoReplica because in ComPAS when the most neighbour node which hold the replica move to other location there is occurrence of packet delivery delay.

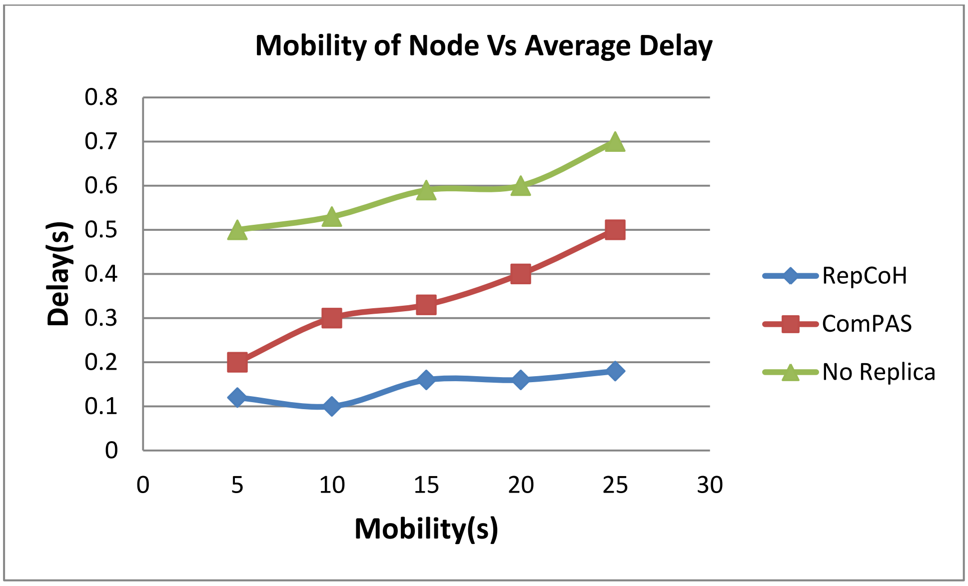


Figure 16: The Effect of Mobility of Node on Delay of Packet

5.3.1.3 Average Packet Loss

The term packet loss occurs when one or more packets of data travelling across a network fail to reach their destination. Packet loss is typically caused by network congestion. Packet loss is measured as a percentage of packets lost with respect to packets sent. The packet drop rate is calculated first by storing the number of packets lost by the sink. The number of lost attribute of the Loss Monitor agent provides the packets lost. We then take this value and divide by the sample time to achieve our packets loss per unit time or packet drop rate.

$$\textbf{\textit{Average Packet Loss}} = \sum_{k=0}^{n} \left(\frac{Number\ of\ packet\ lost}{t}\right) * Time\ of\ Simulation$$

(Equation 3: Average Packet Loss)

The Average Packet drop or loss rate is the average number of data packets that are not successfully sent to the destination nodes. In ComPAS when the time of simulation increases the number of packets dropped also increases i.e. number of packets not successfully reaching the destination is high. The RepCoH performs consistently well with increase in time of simulation. The number of packets dropped is negligible which means that almost many packets reach the destination successfully. The packets dropped are much less compared to performance of ComPAS.

Figure 17 and 18 displays the Xgraph plot of our network's packet drop rate. The X-axis represents simulation time and the Y-axis represents packet drop rate (packets/s). Since a lot of replica allocation and replica read performed when time of simulation increases there is occurrence of packet loss because of bandwidth limitation and routing discovery.

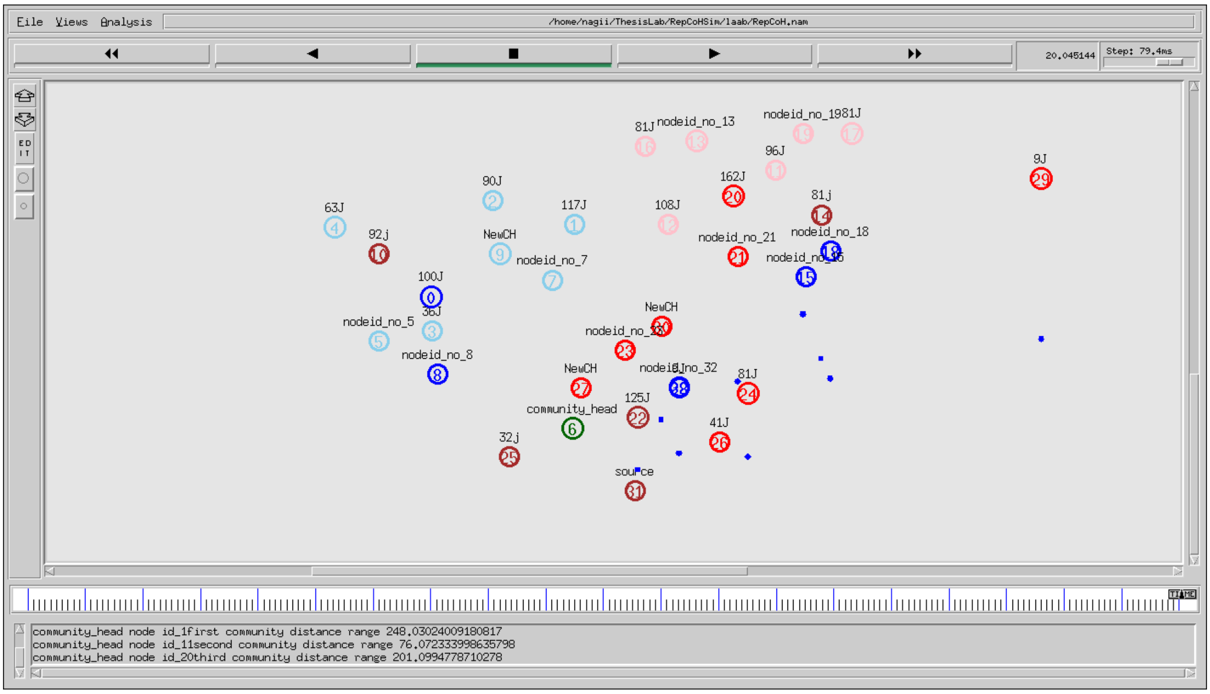


Figure 17: NS2 result shows that when packet is dropped

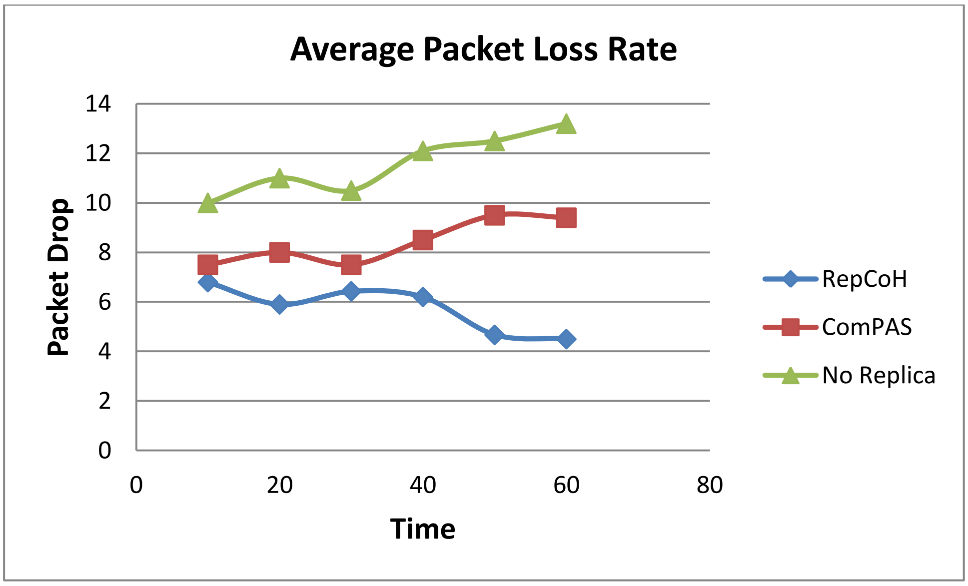


Figure 18: Average Packet Loss Rate Performance

5.3.1.4 Throughput Performance

Throughput is a total number of delivered data packets divided by the total duration of the simulation time. In our case we evaluate throughput of replica allocation in terms of packets delivered per second and in terms of number of nodes in the communities.

The term throughput is also defined as the total useful data received per unit of time. We calculating the throughput, by storing the number of bytes received by the sink or replica holder node. We use the bytes attribute of the Loss Monitor agent to access the bytes received, then convert this number to mega bits per second as seen in Appendix code. The number of bytes received is multiplied by 8, divided by sample time and then divided by 1000000 which achieve the mega bits per second.

$$\boldsymbol{Throughput} = \sum_{k=0}^{n} \left(\frac{Number\ of\ packets\ in\ bytes\ recieved * 8}{time * 100000}\right) * Time\ of\ Simulation$$

(Equation 4: Throughput)

Figure 19 illustrate the comparison of throughput for ComPAS and RepCoH. In this metrics the throughput of the protocol in terms of the number of message delivered per one second or Mega bits per second is analyzed. As the simulation result shows the average throughput performance of RepCoH is 93.7 % which is more than ComPAS and NoReplica when we compare with their result.

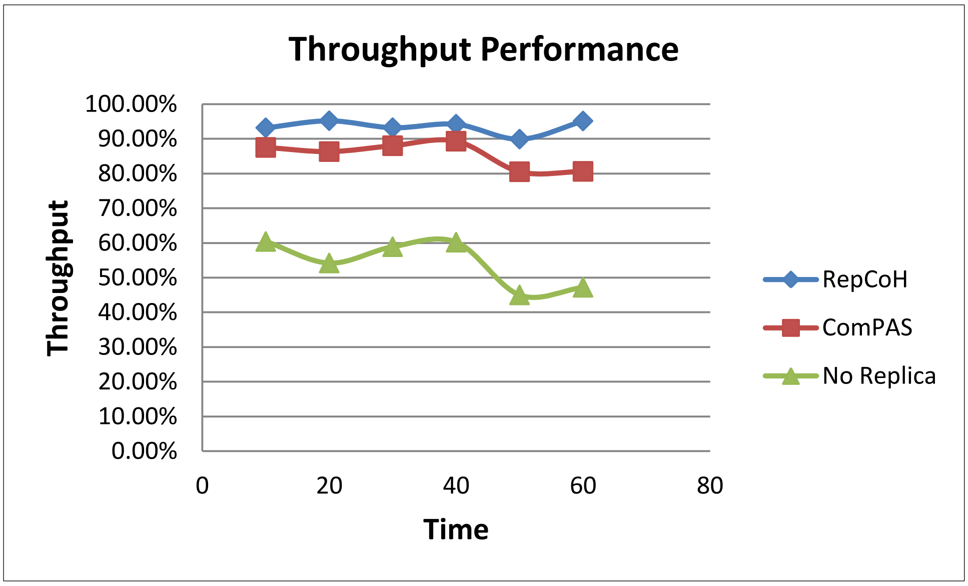


Figure 19: The Throughput Performance

As the figure above indicates the average throughput of RepCoH which is 0.93 is higher than ComPAS which is 0.9 and NoReplica which is 0.6 when we compare it. Therefore our community head consideration replica performs better in throughput than others.

We illustrate also the throughput performance with different number of nodes by comparing with our RepCoH protocol with ComPAS and No Replica schemes. And the throughput result of RepCoH are superior than other as the simulation results shows. The following Figure 20 shows the percentage of throughput performance with different number of nodes.

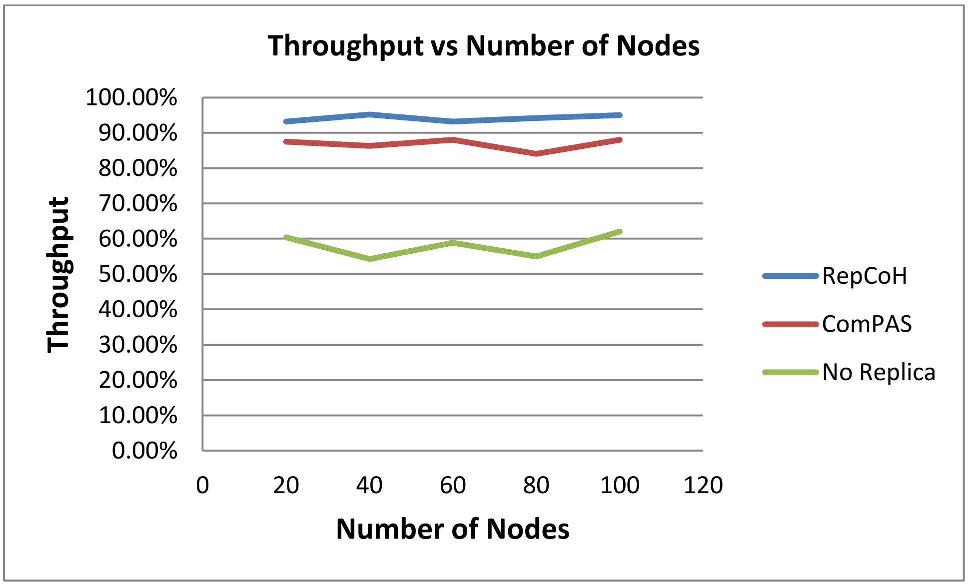


Figure 20: Throughput Vs Number of Nodes

The effect of number of nodes also shows high throughput as the comparison of others. As the figure above indicates the average throughput of RepCoH which is 0.93 is higher than ComPAS which is 0.9 and NoReplica which is 0.6 when we compare it. Therefore our community head consideration replica performs better in throughput than others.

5.3.1.5 The Effects of Number of Nodes on Data Availablity

Basically, the data accessibility is the ratio of the number of successful access requests to the number of all access requests issued. The ASNET replication protocol aims to increase the accessibility of data items in the network. In mobile network environments achieving 100% data accessibility is nearly impossible, due to mobility of nodes and changing network topology.

Figure 21 shows the variation of data accessibility with varying number of nodes in the networks in all the three replication methods. As the number of nodes increases, the number of data items to be replicated in the network also increases, hence the data accessibility reduces. The graph in figure 19 illustrates that RepCoH provides slightly higher data accessibility than ComPAS essentially due to the three reasons; first in case of RepCoH, replica allocation of data items is based community head, while for ComPAS replica allocation is based on mostly neighbour nodes. Second, RepCoH allocate hot data items to nodes having high parameters for replica allocation, thereby ensuring short waiting times for queries at the job queues of these nodes and consequently reduced query response times. Third, RepCoH also considers the node battery power while replicating data item, instead ComPAS doesn't consider the node battery power which may results the allocation hot data item on a node having less battery power or which are going to exhaust quickly.

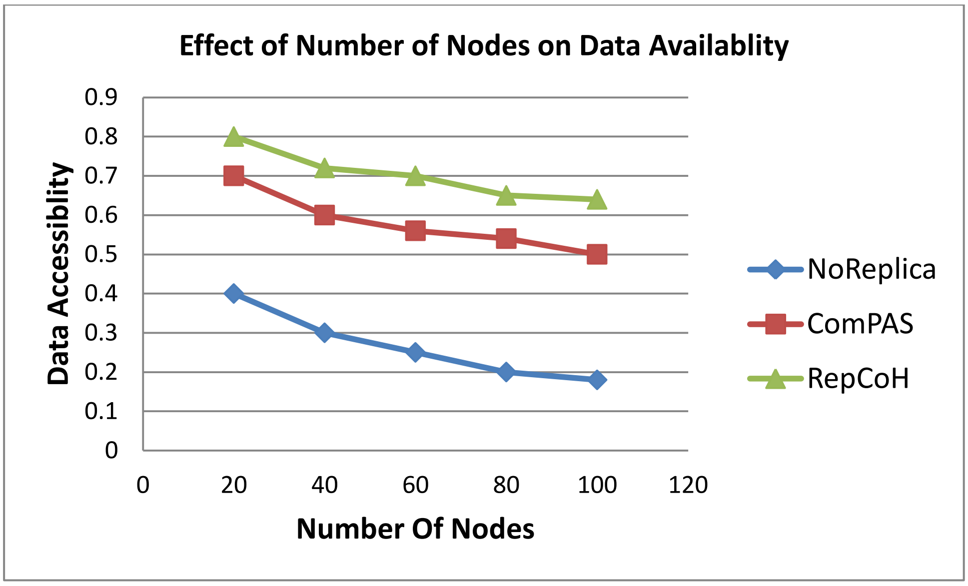


Figure 21: Data Accessibility Ratio on the Effect of Number of Nodes

5.3.1.6 The Effects of Number of Nodes on Packet Loss

Since the wireless communication is prone to packet errors, packet loss is common in ad hoc networks. We investigate in this section the effect of packet loss rate on different number of nodes. Packet loss rate is intuitive that data accessibilities of both schemes degrade as packet loss rate increases. In addition, as shown in Figure below, the superiority of scheme RepCoH over scheme ComPAS and No Replica diminishes when the pack loss rate is high. Since scheme RepCoH relies on the exchange of message among nodes and community heads in the case of high packet loss rate, the replicated packet is lost therefore, recognize replica on community heads minimize packet loss when comparing to replica on most neighbour nodes and no replication. When number of nodes within each communities increases the average packet loss will be changed because of maximum number of hops occur between community heads and between member nodes and community head of the community.

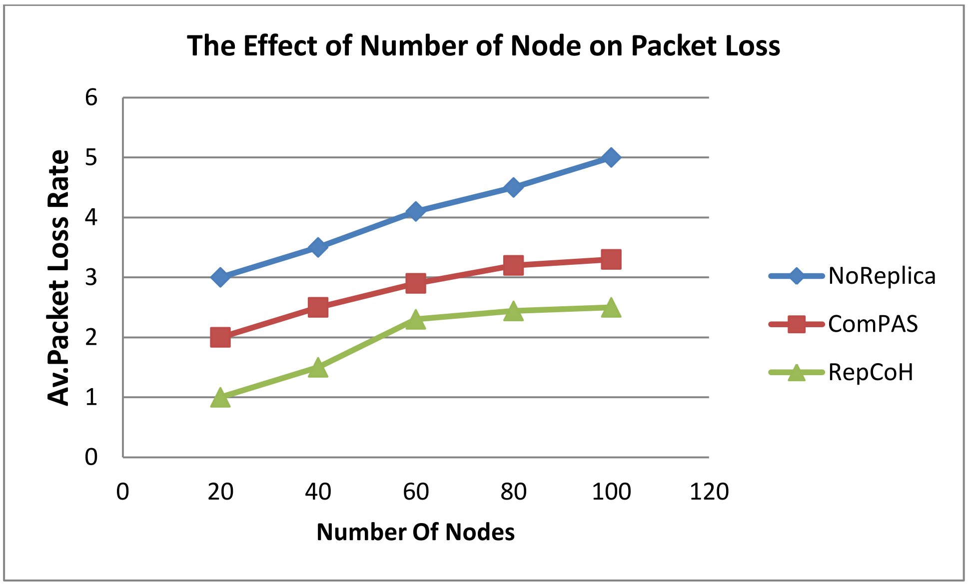


Figure 22: The Effect of Number of Node on Packet Loss

5.3.1.7 Effects of Number of Nodes on Energy Consumption

In ASNET environment, there are no base stations and hence the member nodes cannot predict when they will receive traffic. Energy Consumption is defined as the sum of units required to the key transmission throughout the duration of simulation. When a node sends or receives a packet decrements the available energy according to size of the packets and used bandwidth. Although actual equipment consume energy not only when sending and receiving but also while replicating data and data read. The following equations represent the energy used (in Joules) when a packet is transmitted (Eq. a) or received (Eq. b) and packet size is represented in bits.

$$Energy(Tx) = \frac{(330*5*PacketSize)}{2*10^6} \dots Eq(a) \quad Energy(Rx) = \frac{(230*5*PacketSize)}{2*10^6} \dots Eq(b)$$

... (Equation 5: Energy Consumption when packet transmitted (Tx) and Received (Rx))

We illustrate the energy consumption of community head nodes of RepCoH and replica holder of ComPAS i.e. replica holder are most neighbour node which make occurrence of maximum energy consumed on every nodes and energy used to transmit packet in NoReplica scheme as follows with comparing with different number of nodes and time of simulation.

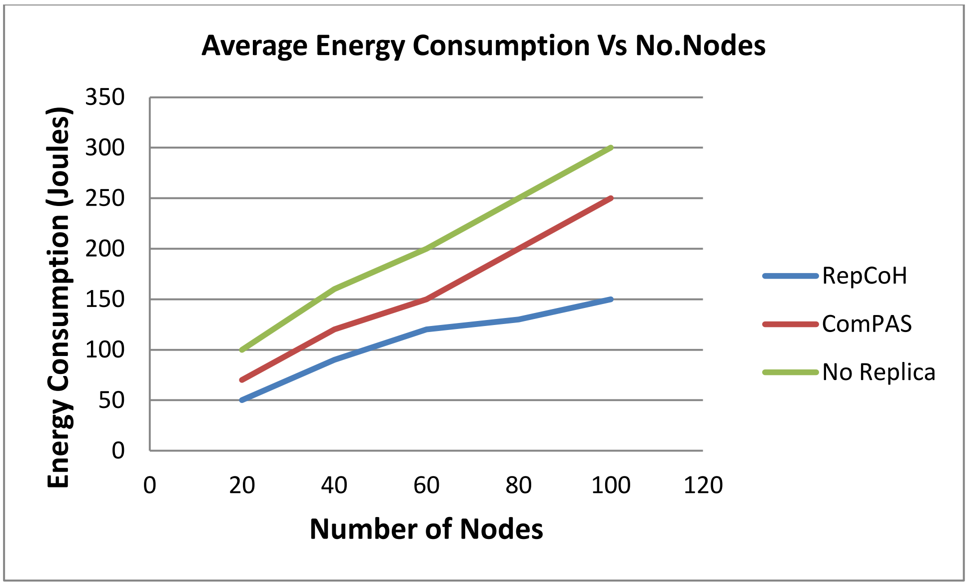


Figure 23: Energy Consumption when Data allocated with different number of nodes

In Figure below we also illustrate the energy consumption of community head nodes of RepCoH and replica Holder of ComPAS and energy used to transmit packet in NoReplica scheme with comparing energy consumed and time of simulation. Since many replica hold and replica read proceed as time of simulation increases, there is occurrence of energy consumption. Therefore we evaluate our model to get optimum results.

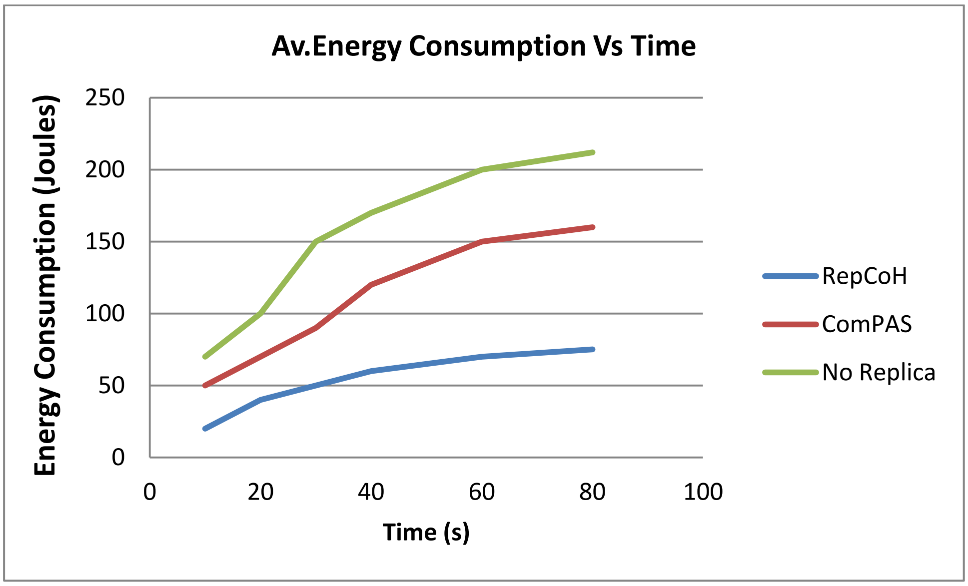


Figure 24: Energy Consumption Vs Time

5.3.1.8 The Efficiency of RepCoH Consistency

To validate the performance of our proposed RepCoH scheme, we have also performed an evaluation to measure the consistency of RepCoH superiority to ComPAS and No Replication techniques in terms of replica allocation efficiency. We compute a measure called community head replication to improve the performance of data and reliable data existence. This implies that is RepCoH constantly superior to ComPAS and No Replication. Schemes Besides, increase in the number of simulation time and even high number of nodes can't have negative impact on the effectiveness and consistency of RepCoH. Finally, the results confirm the efficiency and consistency of the replica allocation method that we have proposed over ComPAS and Without Replication techniques.

5.4 Validation of the Simulation Results

The correctness of simulation model is a primary concern of the developers and users of the model, because we obtain credible simulation results. Verification and validation are activities conducted during the development of a simulation model with the ultimate goal of producing an accurate and credible model [40]. There is couple of techniques in relation to the validation, verification and testing. In order to verify accuracy of the network models and interpreted simulation results we have considered the works of the others person, which is described in literature review and different portions of this course. In the course of simulation results verification, packet delivery performance obtained from the performance analysis. Therefore it can be ensured that network models and the simulation results are correct. With the aim of validating of our simulation results, the techniques which have been used in the analysis are statistical outcomes. Furthermore, the course of action in validating the simulation results has also been followed in this work. The results from this study have favoured packet delivery ratio of RepCoH based data allocation system. Verification of a model is the process of confirming that it is correctly implemented with respect to the conceptual model that is, it matches specifications and assumptions deemed acceptable for the given purpose of application. During verification, our model is tested to find and fix errors in the implementation of the model. Validation checks the accuracy of the model's representation of the real system. Model validation is defined to mean "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model". A model should be built for a specific purpose or set of objectives and its validity determined for that purpose. We validate our model with simulation environment.

Testing techniques can be of significant help during model verification. Testing can also help to ensure that a simulation model that once passed validation and verification will also remain correct for an extended period. Software testing is an art on its own, with several techniques and methodologies.

There is couple of techniques in relation to the validation, verification and testing. In order to verify accuracy of our models and interpreted simulation results we have considered the works of the others person, which is described in literature review and different portions of this course. In the course of simulation results verification, different performance obtained from the performance analysis and the simulation is found to be almost similar. Therefore it can be ensured that our proposed models and the simulation results are correct. With the aim of validating of our simulation results, the techniques which have been used in the analysis are statistical outcomes. Furthermore, the course of action in validating the simulation results has also been followed in this work. The results from this study have favoured the performance rate where RepCoH maintains higher the performance results than others.

CHAPTER SIX

CONCLUSION AND FUTURE WORK

6.1 Conclusion

This paper has proposed a scheme that is used for combining nodes in ad hoc network and replicating data's on various selected nodes in a shapeless, self configuring and self-motivated network, which replicate the objects by determining community head of the node.

In this paper, we evaluated data accessibility of ASNET in terms of various performance metrics such as average data delivery, average data loss, and average delay of data, average replica allocation and average data accessibility of nodes in the above chapter. And we compare different replica techniques in NS2 simulation. Based on this thesis the research question presented in chapter one were set out to examine the data availability performance of ASNET applications.

We have established the importance of considering social relationships in the replica allocation of a user data for Adhoc Social Networks. We have particularly proposed RepCoH, a system based on community head of social community combined with social relationship for data allocation so that data availability for all users is guaranteed. The system gives a number of replicas required on community head of each community for each user that results in an efficient replication solution. The efficiency of the proposed method is better compared to ComPAS, random replication scheme, and no replica scheme as shown in our evaluation results and analysis. We showed that RepCoH offers significant gains in data availability. However, our end result remains at the initial stage as it doesn't address all the data replication protocol design issues and, optimizing all performance metrics is hard to achieve.

6.2 Future Work

As the future works, we will expand the proposed method to work with different cases in social relationship of a community. Following the designed system model as an extension plan for the proposed scheme is the first priority. Since there are a lot of design issues that need to be considered i.e., Security of data, more consistency, more reliability, scalability, better cooperation should be found between these categories to achieve most out of it. We also intend to test RepCoH in real ASNETs scenarios. Finally, based on this work, it is visible to maintain that there is still a lot that needs to be investigated, researched, and formulated before designing an overall efficient data management. To this end, we go one step further to envision more novel solutions being introduced for this important area of research in the future.

APPENDIX

NS2 TCL FOR COMMUNITY FORMATION, COMMUNITY HEAD SELECTION AND RepCoH DATA REPLICA

 # Sample Code for providing initial location of mobile nodes

```
set node_(0) [$ns node]

$node_(0) set X_ 299

$node_(0) set Y_ 248

$node_(0) set Z_ 0.0

$ns initial_node_pos $node_(0) 35

set node_(1) [$ns node]

$node_(1) set X_ 297

$node_(1) set Y_ 496

$node_(1) set Z_ 0.0

$ns initial_node_pos $node_(1) 35
```

 # Sample code of Community Forming Based on Mobility of Nodes

```
$ns at 0.1 "$node_(0) setdest 279.0 369.0 500.0"

$ns at 0.1 "$node_(1) setdest 537.0 498.0 500.0"

$ns at 0.1 "$node_(2) setdest 390.0 541.0 500.0"

$ns at 0.1 "$node_(3) setdest 280.0 305.0 500.0"

$ns at 0.1 "$node_(4) setdest 105.0 494.0 500.0"

$ns at 0.1 "$node_(5) setdest 184.0 288.0 500.0"

$ns at 0.1 "$node_(6) setdest 392.0 302.0 500.0"

$ns at 0.1 "$node_(7) setdest 498.0 398.0 500.0"

$ns at 0.1 "$node_(8) setdest 290.0 228.0 500.0"

$ns at 0.1 "$node_(9) setdest 404.0 445.0 500.0"

$ns at 0.1 "$node_(10) setdest 184.0 446.0 500.0"
```

#Community Head Selection NS2 Source Code

```
proc community {} {
global val ns_ node_ topo contador_nodos rng
Phy/WirelessPhy set Pt_ $val(pt_CommunityHead)
set node_($contador_nodos) [$ns_ node]
$node_($contador_nodos) random-motion 0
set x [$rng uniform 0.0 $val(x)]
set y [$rng uniform 0.0 $val(y)]
$node_($contador_nodos) set X_ $x
$node_($contador_nodos) set Y_ $y
$node_($contador_nodos) set Z_ 0.0
set interval [$rng uniform 0.0 1.0]
Node/MobileNode/SensorNode set processingPower 0.36
Node/MobileNode/SensorNode set instructionsPerSecond_ 150000000
Phy/WirelessPhy set  bandwidth_ 1000000.0
set udp_($contador_nodos) [new Agent/UDP]
set app_($contador_nodos) [create_CommunityHead_app [$node_(1) node-addr]
$val(disseminating_type) $val(CommunityHead_disseminating_interval)]
$node_($contador_nodos) attach $udp_($contador_nodos) $val(port)
$node_($contador_nodos) add-app $app_($contador_nodos)
set processing_($contador_nodos) [new Processing/AggregateProcessing]
$app_($contador_nodos) node $node_($contador_nodos)
$app_($contador_nodos) attach-agent $udp_($contador_nodos)
$app_($contador_nodos) attach-processing $processing_($contador_nodos)
$processing_($contador_nodos) node $node_($contador_nodos)
$ns_ at [expr $val(start) + 1 + $interval] "$app_($contador_nodos) start"
$ns_ at $val(stop) "$app_($contador_nodos) stop"
incr contador_nodos
```

```
if (CommunityHead())
    {
        set $nsmsg*data=(msg*)pkt.data udp;
        set $nsmember_node = node_rep;
        set $community1.node_rep[11]=[0,1,2,3,4,5,6,7,8,9,10]
        set $community2.node_rep[11]=[11,12,13,14,15,16,17,18,19,29,30]
        set $community3.node_rep[11]=[20,21,22,23,24,25,26,27,28,31,32]
        set int $energy_level;
        set int $energy=10joules;
        set $communityhead1=max_energy-level.community1.node_rep;
        set $communityhead2=max_energy-level.community2.node_rep;
        set $communityhead3=max_energy-level.community3.node_rep;
        set threshold_value=5.5
        config_.rp_dsr;
        set rt_upd=routing table_updation
        energy_level=(msg*data)*energy-(msg*data);
}
else
{
        begin ()
      {
  int energy_level;
        community1.node_rep[11]=energy_level;
  if (community1.node_rep[11]=max_energy_level)
  {
    select communityhead1=community1.node_rep->max_energy_level;
            community1.node_rep++;
  }
      else
          {
              communityhead();
          }
```

```
            if (community2.node_rep[11]=max_energy_level)
        {
           select communityhead2=community2.node_rep->max_energy_level;
                    community2.node_rep++;
        }
             else
                 {
                       communityhead();
                   }
            if (community3.node_rep[11]=max_energy_level)
        {
           select communityhead3=community3.node_rep->max_energy_level;
           community3.node_rep++;
        }
               else
                 {
                        communityhead();
                   }
}
}
```

 # Sample Code for setting energy parameters of Node

```
   set node0 [expr ($bw0*$energy)-$bw0]
   set node1 [expr ($bw1*$energy)-$bw1]
   set node2 [expr ($bw2*$energy)-$bw2]
   set node3 [expr ($bw3*$energy)-$bw3]
   set node4 [expr ($bw4*$energy)-$bw4]
$ns at 1.85 "$node_(0)  label  45J"
$ns at 1.86 "$node_(1)  label  117J"
$ns at 1.87 "$node_(2)  label  90J"
$ns at 1.88 "$node_(3)  label  36J"
$ns at 1.89 "$node_(4)  label  63J"
```

 # Packet Replication NS2 Source code

```
set tcp(1) [$ns create-connection TCP $node_(31) TCPSink $node_(20) 2]
set ftp(1) [$tcp(1) attach-app FTP]
set tcp(2) [$ns create-connection TCP $node_(20) TCPSink $node_(32) 2]
set ftp(2) [$tcp(2) attach-app FTP]
set tcp(3) [$ns create-connection TCP $node_(32) TCPSink $node_(18) 2]
set ftp(3) [$tcp(3) attach-app FTP]
set tcp(4) [$ns create-connection TCP $node_(18) TCPSink $node_(11) 2]
set ftp(4) [$tcp(4) attach-app FTP]
set tcp(5) [$ns create-connection TCP $node_(10) TCPSink $node_(1) 2]
set ftp(5) [$tcp(5) attach-app FTP]
set tcp(6) [$ns create-connection TCP $node_(1) TCPSink $node_(9) 2]
set ftp(6) [$tcp(6) attach-app FTP]
set tcp(7) [$ns create-connection TCP $node_(9) TCPSink $node_(6) 2]
set ftp(7) [$tcp(7) attach-app FTP]
set tcp(8) [$ns create-connection TCP $node_(6) TCPSink $node_(27) 2]
```

REFERENCE

[1] I. Chlamtac, M. Conti, and J. J. Liu, “Mobile ad hoc networking : imperatives and challenges,” vol. 1, pp. 13–64, 2003.

[2] S. Giordano, “Mobile Ad-Hoc Networks,” 2000.

[3] F. Xia, S. Member, L. Liu, J. Li, J. Ma, A. V Vasilakos, and S. Member, “Socially Aware Networking : A Survey,” pp. 1–18, 2013.

[4] A. M. Ahmed, F. Xia, N. Y. Asabere, H. Bin Liaqat, and J. Li, “Social Community-Partition Aware Replica Allocation in Ad-hoc Social Networks,” vol. 1, 2013.

[5] Z. T. Takkouz, “Increasing data availability in mobile ad-hoc networks : A community-centric and resource-aware replication approach Increasing Data Availability in Mobile Ad-hoc Networks : Zeina Torbey Takkouz,” 2013.

[6] F. Xia, S. Member, A. M. Ahmed, S. Member, L. T. Yang, J. Ma, J. J. P. C. Rodrigues, and S. Member, “Exploiting Social Relationship to Enable Efficient Replica Allocation in Ad-hoc Social Networks,” vol. XX, no. X, 2013.

[7] P. Padmanabhan, “MANAGING DATA REPLICATION IN MOBILE AD- HOC NETWORK DATABASES (Invited Paper) *,” 2006.

[8] W. Jiang, “A Performance Evaluation of Dynamic replication in MANET.” 2006.

[9] S. Gambhir and N. Aneja, “Ad-hoc Social Network: A Comprehensive Survey,” vol. 4, no. 8, pp. 156–160, 2013.

[10] N. Vastardis and K. Yang, “Mobile Social Networks : Architectures , Social Properties and Key Research Challenges,” pp. 1–17, 2012.

[11] E. F. and F. Potorti, “BLUETOOTH AND WI-FI WIRELESS P ROTOCOLS :,” *IEEE Wirel. Commun.*, no. February, pp. 12–26, 2005.

[12] F. Xia, S. Member, L. Liu, J. Li, J. Ma, and A. V Vasilakos, “Socially-Aware Networking : A Survey,”pp. 1–18,2014.

[13] S. N. , T. Hara, Y.-H. Loh, “Data replication methods based on the stability of radio links in ad hoc networks.,” p. 969, 2003.

[14] T. Hara and S. Member, “Quantifying Impact of Mobility on Data Availability in Mobile Ad Hoc Networks,” vol. 9, no. 2, pp. 241–258, 2010.

[15] P. S. Saleh Ali K.Al-Omari, “An Overview of Mobile Ad Hoc Networks For The Existing Protocols and Applications,” *Graph-Hoc*, vol. vol.2, pp. 87–110, 2010.

[16] A. S. Study, “MASTER ’ S THESIS Routing Protocols in Wireless Ad-hoc Networks - ,” 1998.

[17] A. M. Ahmed, Q. Yang, N. Y. Asabere, T. Qiu, and F. Xia, “ComPAS : Maximizing Data Availability with Replication in Ad-hoc Social Networks,” pp. 215–216, 2014.

[18] A. H. Networks, M. Chatterjee, S. K. Das, and D. Turgut, “WCA : A Weighted Clustering Algorithm for Mobile,” pp. 193–204, 2002.

[19] H. Wu and Z. Zhong, “A Cluster-head Selection and Update Algorithm for Ad Hoc Networks,” 2010.

[20] S. Baolin, C. Gui, Y. Song, and C. Hu, “Stable Clusterhead Selection Algorithm for Ad Hoc Networks,” vol. 6, no. 3, pp. 95–106, 2013.

[21] R. Ferdous, V. Muthukkumarasamy, and E. Sithirasenan, “Trust-based Cluster head Selection Algorithm for Mobile Ad hoc Networks.”,2012.

[22] T. Hara, “Effective Replica Allocation in Ad Hoc Networks for Improving Data Accessibility,” pp. 1568–1576, 2001.

[23] S. K. Madria, T. Hara, S. K. Madria, and S. Member, “Data replication for improving data accessibility in ad hoc networks Data Replication for Improving Data Accessibility in Ad Hoc Networks,” 2006.

[24] K. H. W. and B. L, “Efficient and guaranteed service coverage in partitionable mobile ad-hoc networks,” 2002, no. the IEEE 21st Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 1089–1098.

[25] J. Huang and M. Chen, “On the Effect of Group Mobility to Data Replication in Ad Hoc Networks,” vol. 5, no. 5, pp. 492–507, 2006.

[26] A. D. and N. Badache, “Data replication protocols for mobile ad-hoc networks: a survey and taxonomy,” vol. 11, no. IEEE Communications Surveys and Tutorials, pp. 31–55, 2009.

[27] T. Hara, T. Hara, S. Member, S. K. Madria, and S. Member, “Consistency management strategies for data replication in mobile ad hoc networks Consistency Management Strategies for Data Replication in Mobile Ad Hoc Networks,” 2009.

[28] R. Ann, S. Saritha, and V. Kuttamath, “CRM : Cluster Based Replica Allocation over MANET,” vol. 3, no. 2, pp. 699–703, 2014.

[29] T. Hara, “Replica Allocation Methods in Ad Hoc Networks with Data,” pp. 343–354, 2003.

[30] T. Hara, Y. Loh, and S. Nishio, “Data Replication Methods Based on the Stability of Radio Links in Ad Hoc Networks,” pp. 0–4, 2003.

[31] S. Jain, M. Chawla, and N. Gupta, “Weighted Dynamic Bi-connected Group based Replica Allocation method for Mobile Ad Hoc Networks,” vol. 2013, no. Cac2s, pp. 189–197, 2013.

[32] and M. K. A. Mondal, S.K. Madria, “CADRE: A Collaborative Replica Allocation and Deallocation Approach for Mobile-P2P Networks.,” *10th Int. Conf. Database Eng. Appl. Symp.*, no. Conference, p. pages 21–28, 2006.

[33] R. Azeem, M. Ishfaq, and A. Khan, “Techniques about Data Replication for Mobile Ad-hoc Network Databases,” pp. 53–57, 2012.

[34] Z. Wang, B. Li, L. Sun, and S. Yang, “Cloud-based Social Application Deployment,” pp. 301–312, 2012.

[35] D. A. Tran, K. Nguyen, and C. Pham, “S-CLONE : Socially-aware data replication for social networks,” *Comput. Networks*, vol. 56, no. 7, pp. 2001–2013, 2013.

[36] H. Bin Liaqat, F. Xia, S. Member, J. Ma, L. T. Yang, A. M. Ahmed, S. Member, and N. Y. Asabere, “Social-Similarity-Aware TCP With Collision Avoidance in Ad Hoc Social Networks,” pp. 1–12, 2014.

[37] et al. C. Bettstetter, C. Wagner, “The Spatial Node Distribution of the Random Waypoint Mobility Model.,” *Workshop*, no. In German Workshop on Mobile Ad Hoc Networks (WMAN), p. pages 41–58, 2002.

[38] " Mishra, Vinita, and Smita Jangale, “Analysis and Comparison of different wireless network simulators,” 2014, vol. Vol. 4.

[39] E. H. Teerawat Issariyakul, *Introduction to Network Simulator NS2*, Second edi. Springer Science+Business Media, LLC 2012.

[40] András Varga and OpenSim, “OMNeT++ Simulation Manual,” vol. version 5., 2016.