OPTIMIZATION OF ROUTING PROTOCOLS AND PROFILE MATCHING ALGORITHMS IN AD-HOC SOCIAL NETWORK

THESIS

submitted in fulfillment of the requirement of the degree of

DOCTOR OF PHILOSOPHY

J.C. BOSE UNIVERSITY OF SCIENCE & TECHNOLOGY YMCA, FARIDABAD

by

Nagender Aneja

Registration No.: YMCAUST/Ph55/2011

Under the Supervision of

Dr. Sapna Gambhir

Assistant Professor

Department of Computer Engineering

Faculty of Engineering and Technology

J.C. Bose University of Science & Technology YMCA, Faridabad

Sector-6, Mathura Road, Faridabad, Haryana, India

December, 2018

DEDICATION

to

My Parents

DECLARATION

I hereby declare that this thesis entitled "**Optimization of Routing Protocols and Profile Matching Algorithms in Ad-hoc Social Network**" *by* **Nagender Aneja**, being submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy in **Department of Computer Engineering** under Faculty of Engineering and Technology of J.C. Bose University of Science & Technology YMCA, Faridabad, during the academic year March 2012 - December 2018, is a bona fide record of my original work carried out under the guidance and supervision of **Dr. Sapna Gambhir, Assistant Professor, Department of Computer Engineering, Faculty of Informatics & Computing** and has not been presented elsewhere.

I further declare that the thesis does not contain any part of any work which has been submitted for the award of any degree either in this university or in any other university.

Nagender Aneja
Registration No.: YMCAUST/Ph55/2011

CERTIFICATE

This is to certify that this Thesis entitled "**Optimization of Routing Protocols and Profile Matching Algorithms in Ad-hoc Social Network**" by **Nagender Aneja**, submitted in fulfillment of the requirement for the Degree of Doctor of Philosophy in **Department of Computer Engineering** under Faculty of Engineering and Technology of J.C. Bose University of Science & Technology YMCA, Faridabad, during the academic year March 2012 - December 2018, is a bonafide record of work carried out under my guidance and supervision.

I further declare that to the best of my knowledge, the thesis does not contain any part of any work which has been submitted for the award of any degree either in this university or in any other university.

(Signature of Supervisor)

Dr. Sapna Gambhir

Assistant Professor

Department of Computer Engineering

Faculty of Informatics & Computing

J.C. Bose University of Science & Technology YMCA, Faridabad

Dated:

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my Supervisor **Dr. Sapna Gambhir** for giving me the opportunity to work in this area. It would never be possible for me to take this thesis to this level without her innovative ideas and her relentless support and encouragement.

Besides my advisor, I would like to thank **all faculty members** for attending progress presentations and for their insightful comments and encouragement that incented to widen the research from various perspectives.

I also wish to express my gratitude to **Prof. Komal Bhatia** and **Prof. Atul Mishra** for motivating me and continuously persuading the right way to pursue the research work. I appreciate him for all his guidance and instructions.

I am grateful to **my parents**, who have provided me with moral and emotional support in my life.

My word of thanks cannot compensate the contribution of my wife **Dr. Sandhya Aneja**, who has been extremely supportive throughout this entire process and has made countless sacrifices to help me get to this point.

I am thankful to my son **Kartikeya** for giving me happiness, love, and affection.

I also express my sense of gratitude to one and all, who directly or indirectly, helped in this research work.

Nagender Aneja

Registration No.: YMCAUST/Ph55/2011

Abstract

Ad-hoc social networks aim to manage user social connections in a novel way by founding short social relationships among users with shared interests or requirements and provide an innovative mechanism to enhance existing as well as beginning new social relations. However, the existing approaches lack location-based profile matching that utilizes dynamic profile, while using traditional similarity metrics to score profile matching. Further, existing approaches used Mobile Ad-hoc Network (MANET) based routing protocol without considering the fact that not all MANET nodes will create and maintain Ad-hoc Social Network (ASN). Researchers have used WiFi Direct Peer-to-Peer networking (P2P) technology to create and facilitate communication among ASN nodes. However, no attempt has been made to match dynamic profiles over P2P without any central entity. This research solved the above challenges.

This research work has contributed in *providing middleware architecture, users' preferences that can be used to propose new application, optimizing profile matching, profile similarity metric, routing algorithms, and android based mobile application* for ASN.

The middleware architecture proposed in the research work can be used to create ASN applications to provide location-based social networking. The architecture comprises four layers Application Layer, Transport Layer, Ad-hoc Social Layer, and Ad-hoc Communication Layer. The Ad-hoc Social Layer provides necessary functions like profile management to broadcast and receive profiles and perform the perform profile matching.

The research work also conducted a survey of users to understand users' preferences better. Results from 108 participants comprising 82% from academic field and 18% from industry field shows 91% users actively accessing social networking and while 47% users have the opinion that the current social networking sites don't provide enough opportunities to know people of similar interest. One of the objective of the survey was to know how much similarity users' prefer when a person nearby with similar interest available, and results indicates the average similarity users' prefer is 75%.

Prior studies have described profile as a set of keywords describing attributes about user and matching users based on these attributes. While most of the studies have considered static interests as attributes or have used dynamic interests in some cases changed by users manually.

In few cases profile has been built by extracting keywords from browsed URLs. The research analyzed the prior mechanisms and propose a novel location-based profile matching algorithm, called Geo-Social Profile Matching Algorithm. Another variant of the proposed algorithms called Semantic Geo-Social Profile Matching Algorithms is also presented. Results indicate Geo-Social Profile Matching Algorithm performs better when the required profile similarity threshold is higher or equal to 70%. Since as per survey users' preferred profile similarity is 75%, the geo-social profile matching will perform better in real-life scenario. The results are further improved by semantic geo-social profile matching algorithm.

Profile similarity scoring is the most fundamental and important operation that computes similarity score such that the value is directly proportional to the interests similarity. Cosine Similarity has been used in the research studies, while the analysis indicates the cosine similarity needs improvement for weighted interests. The proposed metric Piecewise Maximal Similarity has been proposed that computed attribute based similarity and analysis over real-data indicates that Piecewise Maximal Similarity (PMS) performs better than Cosine Similarity. The performance of PMS over Cosine Similarity is 4% when compared with Facebook and 3% when compared with Bluetooth.

The analysis of PMS and Bluetooth indicates that people with similar profile have tendency to have higher Bluetooth encounters. Thus, a routing algorithm that selects intermediate nodes with similar profile will be stable and improve the network efficiency. The proposed Social Profile Aware Ad-hoc On-demand Distance Vector (SPA-AODV) is a modification of Ad-hoc On-demand Distance Vector (AODV) routing protocol that uses contextual information i.e. social profiles of the neighboring nodes for routing decisions. Simulation results indicate the SPA-AODV performs better when the number of nodes are high and also even when the nodes are moving with walking speed of 1 m/sec - 2 m/sec.

WiFi Direct Peer-to-Peer networking (P2P) has been preferred by researchers to create ASN for Android based mobile devices since P2P doesn't use any infrastructure. This research not only used P2P for network communication but also shared the limited interests over P2P without any central resource. The proposed android based mobile application, OffAT-Chat in Airplane Mode, is available at Google Play Store to create ASN with nearby users based on similar profiles.

Thus, this research achieved its objectives of optimizing profile matching algorithms and routing protocols by proposing geo-social profile matching algorithm, semantic geo-social profile matching algorithm, profile similarity metric, and social profile-aware ad-hoc on demand distance vector routing protocol. The proposed middleware architecture and the android based mobile application demonstrates the ASN from practical feasibility of ASN.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES/GRAPHS

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

AES Advanced Encryption Standard

AODV Ad-hoc On-demand Distance Vector

ASN Ad-hoc Social Network

CS Cosine Similarity

FAR False Alarm Ratio

GO P2P Group Owner

GPS Global Position System

MAC Media Access Control

MANET Mobile Ad-hoc Network

ns2 The Network Simulator - ns-2

OffAT Offline Chat

PC Percent Correct

PMS Piecewise Maximal Similarity

POD Probability of Detection

P2P WiFi Direct Peer-to-Peer networking

RREQ Route Request

RREP Route Reply

RERR Route Error

SPA-AODV Social Profile Aware Ad-hoc On-demand Distance Vector

TS Threat Score

TTL Time-To-Live

WPA2 Wi-Fi Protected Access II

Chapter I

INTRODUCTION

Chapter I

INTRODUCTION

This chapter describes introduction to Ad-hoc Social Network, related research issues, problem definition, objectives of the research, thesis contribution, and organization of the thesis.

1.1 Ad-hoc Social Network (ASN)

The outgrowth of wireless and mobile communication technologies have directed to new networking concept called Mobile Ad-hoc Network (MANET). A MANET comprises independent and self-governing mobile nodes that are not dependent on fixed infrastructure for communication with nearby nodes. However, mobile nodes rely on intermediate nodes to route packets from a source to a destination. There are many applications including unplanned meetings, costly access to fixed infrastructure, disaster relief, emergencies, military battlefield communications, remote areas with no connectivity, foreigners visiting different countries without local SIM, researchers attending international conferences, persons in the airplane, etc. where MANET is highly useful. MANET is also being explored by researchers to provide on-demand social networking that doesn't use the Internet.

Ad-hoc Social Network is defined as a Mobile Ad-hoc Network consisting of wireless devices also called nodes equipped with Wi-Fi and Bluetooth, which are connected to each other via some social patterns e.g. interests and are free to move [1]. The ASN implemented on MANET is useful due to various reasons including easy deployment without additional cost. The importance of ASN was recognized by Google with implementation of P2P networking in Android 4.0+ devices. The P2P is a Wi-Fi communication that connects devices without the need for infrastructure hardware or fixed router. The P2P enabled devices have a built-in soft access point that helps devices to act both as router and client.

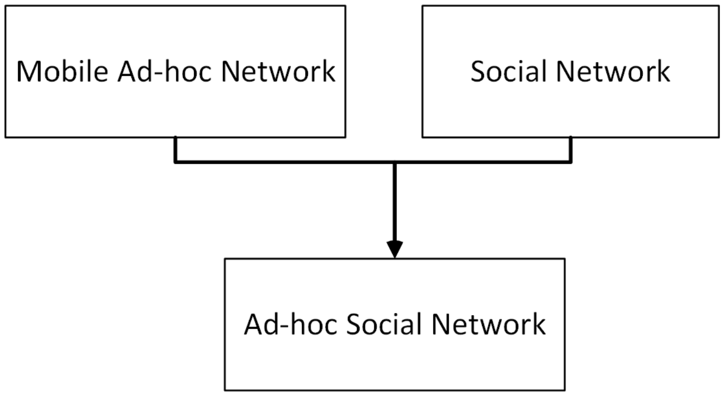


Figure 1.1: Ad-hoc Social Network (ASN)

Thus, ASN, as shown in Figure 1.1 is a combination of MANET and Social Network. ASN can connect users who are in a proximate location to each other and have common interests. ASN uses routing protocols of the MANET to facilitate communication among interested users.

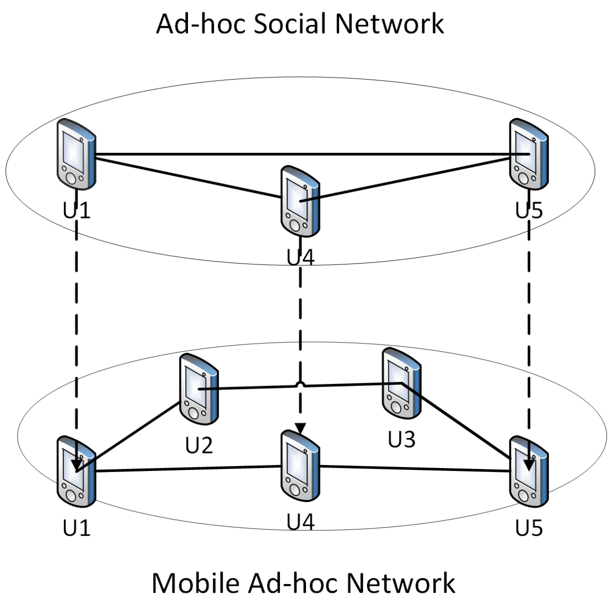


Figure 1.2: Nodes U_1, U_4, and U_5 participating to create ASN

Out of many nodes that form the MANET like $\{U_1, U_2, U_3, U_4, U_5\}$, some nodes, e.g., $\{U_1, U_4, U_5\}$ can create ASN as shown in Figure 1.2. ASN also helps users to communicate and share content with nearby users who have similar interests. The content can be shared without the need of Internet or other Infrastructure including hot-spot and is useful in applications like location-based photo sharing especially in a group with International Roaming, playing networking games in Aeroplane or Train, chatting, broadcasting location-based advertisements in a shopping mall, communication in emergency situations, etc.

The current technologies that support ASN in the mobile devices include Bluetooth, Wi-Fi, and Wi-Fi Direct. Bluetooth and Wi-Fi have been used traditionally and have certain limitations e.g. range in case of Bluetooth is limited, and Wi-Fi needs central infrastructure to facilitate communication. Recently in late 2011, Google provided support of Wi-Fi Direct in the Android 4.0+ devices that provides peer-to-peer networking and group communication. Figure 1.3 shows an example scenario where devices can be connected using WiFi Direct Peer-to-Peer networking (P2P).

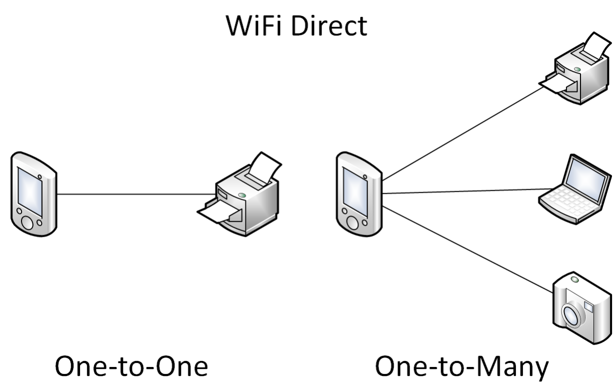


Figure 1.3: WiFi-Direct for one-to-one and one-to-many devices

1.2 ASN related Research Issues

Researchers have recommended extending cellular infrastructure via ad-hoc communication technology like P2P. Several studies of research have shown that ad-hoc communication has potential to optimize resources like network bandwidth and battery lifetime. The multi-hop wireless network or MANET has also been proposed to provide networking applications when an infrastructure-based network is hard to deploy or is not possible in some emergencies. MANET allows intermediate nodes to route data for connectivity, and these applications are highly useful for military and emergencies. The rapid advancement of wireless technology has made possible to establish social network services over MANET. The current challenges in ASN can be classified into following three categories:

1. *Profile Matching:* Profile Matching refers to determining interests of a particular user and suggesting other nearby users who have interests similar to the user. The role of matching interests is important since the probability of two users being connected is higher when they have similar interests. Profile matching is a challenge in ASN since it doesn't have any centralized infrastructure to maintain users' profiles and further nodes are resource constraint and highly mobile.

2. *Profile Similarity Metric:* Profile Similarity Metric refers to a similarity measure between nodes. Since the number of users available for ASN may not be large and thus, as a result, the probability of users getting matched based on static interests will also be small. The static interests may include current city, education, job, etc. Thus, the need of computing the similarity of users instantly without accessing centralized server and preferably using dynamic location-based interests instead of static interests is

advantageous for ASN.

3. *Routing:* Routing refers to selecting a path to send data packets. The routing protocol provides necessary functions and services to nodes to help them decide and select intermediate nodes to route packets from a source to destination. The nodes are not aware of the existing topology, and further the topology is also dynamic due to nodes mobility. There exists number of MANET routing protocols handle these issues, however, the protocols don't consider the similarity metric as the topology. The probability of nodes remain connected is higher when the interests similarity between nodes is higher. Thus, incorporating the profile similarity in routing decisions can help provide a stable path and increase network throughput.

1.3 Problem Definition

The implementation and acceptability of ASN rely on the simple and efficient algorithm that can search and match people having similar interests by providing a stable and reliable path between connected and similar nodes. This includes improving the profile similarity and matching algorithms along with improving the routing protocol.

Profile similarity and matching have been addressed by some researchers, however, the current approaches include synchronizing the online profile to determine location-based nearby users, computing interests after being connected, using published information at a centralized resource, or using interests irrespective of the current location. Moreover, current approaches used the traditional similarity metrics to score the matched profiles. Thus, there is a need to improve the profile management by using location-based interests and scoring the matched profiles that provide a high score to two profiles which best match based on location-based interests.

Some studies have been published in creating a multi-hop wireless ad-hoc network to be used for ASN, however, using the traditional routing protocols like AODV may not be advantageous since not all MANET nodes participate in ASN. Thus, MANET routing protocol that incorporates social networking characteristics into routing decisions shall be advantageous for ASN.

1.4 Objectives of the Research

The prime objective of ASN is to search nearby persons for a given user whose interests best match with the user. The algorithm to search nearby friends should be intelligent so that it helps users to extend their existing social network. There are several methodologies to discover friends over MANET. However, these have either high complexity or low success ratio. Further, the routing algorithms that have been used in ASN are traditional MANET routing algorithms that need to be customized ASN.

The objectives of the research work titled “Optimization of Routing Protocols and Profile Matching Algorithms in Ad-hoc Social Network” were to optimize *Profile Matching Algorithm, Profile Similarity Metric, Routing Algorithm*.

The objective of optimizing profile matching algorithm is to include location-based features in determining interests. Using location information is not just check-ins since the fact that a user checked-in is the first connection but the most valuable information is the data that includes information and behavior, which may come from search and browsing history around that location. Thus, there is a need to look the location broadly than just a particular geo-location.

The profile similarity metric is used to match profiles of two users. A typical user profile is characterized by some words comprising hometown, interests, professional associations, location, etc. The profiles are matched and scored based on the intersection on these parameters. However, it becomes a challenge when the number of users is less since scores based on these keywords may not be sufficient to be used as ranking similarity. Thus, profile similarity metric needs to be optimized for ASN so that the metric provides weight to the number of times words appear in the search and browsing history in addition to static information like hometown and professional information.

The traditional routing protocol e.g. AODV, DSDV, DSR can work for ASN; however, the performance of the ASN can be improved by adding social characteristics in the routing decisions. The reason to add social characteristics is that ASN nodes are a subset of MANET nodes since not all nodes will remain connected and the probability of nodes being connected is high when the profiles are similar.

1.5 Thesis Contribution

This research work has contributed in the area of ASN to optimize profile matching and routing algorithms. Specifically, the contribution of the research work is as follows.

A survey of ASN and it's major challenges was published. This research also conducted a users' survey by an online questionnaire to know users' preferences for ASN. A Middleware architecture has also been proposed as a framework to create ASN. The middleware introduces Ad-hoc Social Layer and Ad-hoc Communication Layer along with Application and Transport Layers for profile management functions.

A novel Location-based Profile Matching Algorithm is proposed that is suitable for ASN. The core to create a social network is the capability of locating people that have similar interests by using their profile information such as family names, locations, etc. The proposed implementation of location-based profile improved the profile matching algorithm. The proposed algorithm creates a location-based profile based on prior user actions and browsing history and is more propitious than static or global profile. The location-based profile supports managing different location-based interests in a forest data structure with the root node as a location node. A global profile of the user is a union of all local profiles. In case of local profiles, the interests are computed from browsing history and prior user interaction on a mobile device at that location. Simulation results show the performance of local-profile matching is better than profile matching based on a global profile especially when the similarity threshold is higher.

A novel Piecewise Maximal Similarity (PMS) is proposed that performs better than Cosine Similarity. Computing profile similarity is an essential requirement in the area of social networking to recommend similar new social connections that have a high likelihood of being affirmed as an actual connection. Cosine Similarity has been widely used to match users interests. This research analyzed cosine similarity for social profiles and observed that cosine similarity is not the best profile similarity metric especially in case of weighted interests for location-based ASN. The PMS has been proposed [2] that computes maximal profile similarity based on a minimum weight for each interest attribute. The PMS was analyzed by considering a dataset that provides interests of a set of users collected through SocialBlueConn application [3] at the University of Calabria. The dataset also includes Facebook friendship connections among participants and traces of Bluetooth encounters among participants during the experiment. The performance of the PMS was computed based on a contingency table with four parameters, i.e., Percent Correct, Probability of Detection, False Alarm Ratio, Threat Score. The Piecewise Maximal Similarity was compared with the Cosine Similarity for real data of Facebook friends data in a first scenario and actual encounters of users detected by Bluetooth in a second scenario. Results indicate that PMS performs better than Cosine Similarity (4% better correct results when compared with Facebook and 3% better correct results when compared with Bluetooth) in both scenarios and is more consistent overall with Facebook and Bluetooth Contacts.

This research also proposed modifications in AODV Routing protocol now called Social Profile Aware Ad-hoc On-demand Distance Vector (SPA-AODV). The modified protocol incorporates social profile metric in the routing decisions and performs better than AODV. Routing Protocol based on contextual information has been proposed to transfer data packets for the reason that contextual routing is more advantageous for the ASN [4]. This is because nodes are carried by humans that stay together when interests are matched. The need for contextual aware routing protocol can be explained by considering a scenario in which there exists two paths between a source and a destination, wherein one of the paths is shortest, while the other path is longer but intermediate nodes on the longer path have higher profile similarity. Since the nodes are carried by human beings who tend to stay together if the profile similarity is higher, thus, the longer path tends to be more stable and will improve network efficiency. The protocol has been simulated on The Network Simulator - ns-2 (ns2) and simulation results indicate that control packets increase with increasing number of nodes in the proposed protocol SPA-AODV, however, overhead is less than or equal to AODV for different scenarios comprising number of nodes ranging from 4, 8, 16, 32, and 64. Further, SPA-AODV performs better when the number of nodes increases, thus favorable to large ASN. Packet Delivery Ratio was also computed using different speed ranging from 1 m/sec to 3 m/sec. Results indicate that the packet delivery ratio ratio decreases when speed is increased, however, SPA-AODV performs better than AODV for a higher number of nodes. Reduction of packet delivery ratio at higher speed is due to the reason that nodes may become unreachable as they travel from one location to another or packets were dropped till new path was established.

An android based mobile app for ASN to test the feasibility study for matching location-based interests without a central server was developed. The android application scan interests of nearby users and provide profile similarity without using any centralized infrastructure. The application uses three-layer ASN architecture comprising: Geographical Layer, Social Layer, and Content Layer. The process of creating ASN is initiated by selecting a user from a list of nearby users based on maximal profile similarity or otherwise. The first step is to send a connection request using P2P. The waiting time to accept the request by the other user is configured to 30 seconds. The users may share content or add other users to the group once a request is accepted. The geographical layer allows seeing nearby devices that are interested in ASN. A user can also access the nearby available users along with profile similarity in an Airplane Mode. The request to create the connection using P2P is sent when the icon in front of a user is clicked, while the content that includes file and text can be shared after a connection request is accepted. The current version supports exchanging typed and hand-drawn text messages. Once two users are connected, other users can join the group and share text and files in the group. The application does not need any centralized infrastructure and has been tested with a combination of devices that were in Airplane Mode, but WiFi switched on without being connected to an access point, connected with an access point, without SIM, but WiFi switched on. The delivery of messages was tested with a group of students in a class, and all messages were delivered successfully.

1.6 Organization of Thesis

The thesis is mainly divided in six chapters as listed below:

Chapter 2 discusses the applications, characteristics, and related approaches for ASN. The related approaches have been classified into four categories *Architecture, Framework, and Mobile Application Implementation; Profile Management; Similarity Metric; Routing Protocols*. The chapter also provides a comparison analysis of current state-of-the-art.

Chapter 3 presents proposed middleware architecture and finding from users' preferences. The middleware architecture comprising four layers provide a framework for ASN. The various parameter can be set by an application developer or users for ASN using the proposed middleware, however, it is important to know preferences of the users also. This chapter also presents findings from a users' survey for ASN.

Chapter 4 presents the proposed Geo-Social Profile Matching Algorithm that is a location-based profile matching algorithm. The chapter also describes Piecewise Maximal Similarity, which scores profile similarity to help decide a user about potential nearby connections.

Chapter 5 describes Social Profile Aware Ad-hoc On-demand Distance Vector Routing protocol, which includes social profile similarity in the routing decisions. A path with similar intermediate nodes will be more stable than the others. This chapter also describes the android based mobile application developed to demonstrate ASN. The mobile application can broadcast limited interests in the nearby region using P2P and provides information about profile similarity with neighboring nodes.

Chapter 6 presents Conclusion and Future Work. It summarizes the significant accomplishments of the research work and illustrates the scope for future work in this technology area.

Chapter II

LITERATURE SURVEY

Chapter II

LITERATURE SURVEY

This chapter presents the uses, properties, and relevant strategies, organized in various categories *Architecture, Framework, and Mobile Application Implementation; Profile Management; Similarity Metric; Routing Protocols*. The chapter also contributes a comparative critique of principal work.

2.1 Introduction

ASN consists of a highly dynamic, decentralized and self-organizing network of autonomous, mobile devices that interact with peers. ASN is a social network over MANET that extends online social network. Despite the extensive progress of social networks, there are still particular restrictions like exploitation by central server on the privacy and thus, need new design solutions. People often ask ASN to strengthen local interaction, e.g., in scenarios like conferences and exhibition. Following are some applications of ASN in the shopping mall, classroom, office, public events or other public places:

2.2 ASN Applications

ASN implemented on MANET are useful due to the reasons of easy and inexpensive deployment instantly. There are a number of scenarios where ASN can be useful.

Extending Cellular Network: The cellular network service providers can provide customers a mobile application that uses ASN to extend the networking capability in the remote areas where there is no connectivity e.g. during International Roaming.

Academic Conferences: Most of the conferences being organized have participants from international institutes who may not prefer to buy local SIM. Some conference organizers don't share WI-Fi due to security concerns. However, even if the connectivity is available, it is difficult to extend social networking with strangers. The ASN applications can help to not just connect people nearby but also match interests and notify.

Airplane: ASN can be used to provide social networking applications and multi-player games inside the airplane. There is no connectivity available, and many flights are around 12 hours that creates boredom especially for children. People can use ASN to search for other people with similar interests, children can play multi-player games, and airlines can also provide passenger to cabin crew communication using ASN. ASN applications can also provide in-flight entertainment and shopping experience for low-budget airlines.

Public Places: ASN can also be used in restaurants, malls to provide customized offers to nearby potential customers and customers can create groups to get bulk discounts from shops. People can play networking games or chat inside metro station, airports, trains, etc.

Private Places: In residential societies or organization, many times a consensus is required for certain decisions. ASN can help people to connect and vote for different options.

Emergency Situation: ASN can be used in Natural Disasters even to locate human presence in emergency situations.

Military: Soldiers can't use public network or GPS since there is fear of being tracked by the enemy. In these situations, soldiers can use ASN to transmit information like messages or pictures of their location to team members.

2.3 ASN Characteristics

There is no straightforward way in current social networks to facilitate local social communication. Although face-to-face conversation is possible, it has limited use for file sharing and group discussions without disclosing the mobile number or email address. Further, the existing social network services assume Internet by cellular or WiFi is always available, which is not true in emergencies, due to cost in International Conference in foreign countries, or due to security aspects. Recently, the ASN has become widespread for universal usage of mobile devices. ASN sets up local communication via mobile devices without using a fixed infrastructure. Typically, ASN has following characteristics, also known as *SPOT*:

Spontaneous: Spontaneous refers to networking when the nodes discover each other within a short period without any prior planning.

Proximity-based: The objective of ASN is to bring together people who are nearby and have similar interests; thus Physical Proximity of two or more individuals is utmost importance. Further, nodes are also required to establish and maintain route between source and destination.

Opportunistic: Opportunistic networks are special wireless networks that exploit the human social characteristics, e.g. similarities, mobility patterns, and interests to perform routing and data sharing. ASN is opportunistic since these are created on-the-fly to connect people nearby.

Transient: ASN are set-up for temporary purpose and are thus short lived.

2.4 Related Work

ASN relaxes the condition that network infrastructure is necessary and establishes local community by using short-range communication techniques, e.g., Bluetooth, ZigBee, or P2P. There have been various attempts to allow direct communications within IEEE 802.11 radio devices [5], e.g., IEEE 802.11 DCF, 802.11s, and 802.11z but their diffusion is very restricted. However, the new protocol named WiFi Direct, based on IEEE 802.11, addressed the weaknesses of IEEE 802.11 DCF. The WiFi Direct is now natively available on most mobile devices and is getting exceptional attention from the research society. The WiFi Direct endeavors to advance WiFi based ad-hoc communications with energy preservation features.

The prior studies related to ASN can be classified into following categories: (i) *Architecture, Framework, and Mobile Application Implementation*, (ii) *Profile Management*, (iii) *Similarity Metric*, and (iv) *Routing Protocols*.

2.4.1 Architecture, Framework, and Mobile Application Implementation

Seada and Perkins [6] proposed MANET architecture for proximate social networks using existing wireless interfaces. The authors introduced many research challenges in areas of, e.g., high-density scenario, management of friend lists, and localized search.

Zhang et al. [7] proposed MoNet, which is a Wi-Fi-based multihop networking system. The authors proposed WiFace on the top of MoNet to share content over MANET. MoNet uses virtual NIC with 48-bit virtual Ethernet address to support MANET routing protocol. The infrastructure used many mobile nodes deployed in a distributed environment connected with wired Internet and used VPN to connect clients. The content in the WiFace is broadcasted and received and doesn't provide a mechanism for locating people with similar interests on-the-fly.

Zhang et al. [8] proposed BASA - Building Mobile Ad-hoc Social Networks on top of Android. Network Layer of the BASA allows proximate devices to communicate. A device that initiates a network connection subscribe to the published services and create contact after determining the connection details. BASA also conducts network operations to discover nodes, services, and connection changes. The limitation of the BASA is that it needs rooted Android Phone and further there is no interest matching of users before a connection is established.

Hoang and Ogawa [9] developed MANET using P2P over Android. Simulation results indicate advantageous of P2P over Bluetooth. The authors observed stability in the discovery time of P2P even when the devices are increased.

Chung et al. [10] and Joy et al. [11] proposed DiscoverFriends application that creates android multi-hop P2P using IPv6. The communication is facilitated by a confidential ID that is known only to user's friends. It doesn't allow to search new friends with active interests.

Ramos et al. [12] presented a case study, FrameGeoSocial, to create a social network among friends using MANET and Global Position System (GPS).

Rahman and Hossain [13] described a framework to create ASN of millions of people by offering context-aware serious-game services as an incentive. The authors introduced a framework that connects various portable devices using the cloud. It promotes heterogeneity amongst people relating to diverse cultures by grouping the devices to form the interest-based community. The framework uses a requester as a client, and a crowdsource as a server.

Table 2.1 summarizes the important features and provide remarks for different prior studies in the area of Architecture, Framework, and Mobile App Implementation of ASN. The comparison of categories shows need to work in the area of providing social applications over P2P among similar users.

Table 2.1: Comparison of related work for Architecture, Framework, and Mobile App

Paper	Features	Remark
Seada and Perkins [6], 2006	Architecture for proximate social networks	Proposed research challenges
Zhang et al. [7], 2010	WiFace over MoNet using virtual NIC with 48-bit virtual Ethernet address	No interest matching nearby
Zhang et al. [8], 2014	BASA creates ASN over Android	Needs rooting devices and no Interest Matching
Hoang and Ogawa [9], 2014	P2P over MANET	Shows P2P advantageous over Bluetooth
Ramos et al. [12], 2015	FrameGeoSocial is a social network over MANET and GPS	Profile managed manually
Chung et al. [10], 2015 and Joy et al. [11], 2016	Discover Friends that creates P2P using confidential ID	Doesn't allow to search new friends with similar interests
Rahman and Hossain [13], 2017	ASN provides game services and incentives	No location-based dynamic profile matching

2.4.2 Profile Management

Bottazzi et al. [14] suggested a middleware for ASN comprising two components named Dependent Social Network Manager (PSNM) and Global Network Manager (GSNM) for building a user profile. PSNM announces user profile depending on its diverse interests, and GSNM consolidates PSNM profile with its location id. However, the profiles are not refreshed dynamically.

Campbell et al. [15] suggested adding sensing ability into social networking applications. The authors introduced a system named Cence Me that receives information about neighboring users and brief facts which can be applied in many applications. The authors also proposed a buddy locate service so that a user can receive instant notification if a CenceMe user has a similar profile. The system interprets activities of a user by sensing various sensors of a mobile device, however, in addition to requiring Internet and servers it gives no weight to extract interests from users search and browsing history.

Sarigöl et al. [16] presented Ad Social that fosters social network applications in an ad-hoc network and exhibited ASN on 10 - 15 Nokia N810 handhelds besides economic overhead. In a standard online social network (e.g., LinkedIn, Facebook) a user's list of friends includes friends believed by the user, while in ASN, peers are nearby users whose nearness has been identified by Ad Social. Users can locate profile of nearby fellow quickly and start a chat session. Alternatively, they can also explore peers resembling some specific interest. However, Ad Social matches interests applying a string matching algorithm and asks the user to record his/her profile manually.

Yiu et al. [17] exhibited an application that discovers mates in proximity based on provided threshold Euclidean distance. The authors stated that the application tunes its proximity distance according to communication cost. The application can be utilized in multiplayer gaming to find players that have similar interests.

Sarigöl et al. [18] manifested a tuple space that abstracts underlying network as a shared memory space wherein nodes can save and view key-value pairs (i.e., tuples). However, each application institutes its own "shared memory" rather than all tuples residing in particular shared memory. The authors applied tuple space to execute a buddy presence service that permits users to see all buddies in their vicinity as well as search for buddies with particular interests. Every user builds a profile that covers a list of interests. The profile is interchanged among users to sift friends. However, this method does not address learning interests automatically to generate a profile of the user.

Lee and Hong [19] proposed a mechanism to extract changing interests and uses cosine similarity. The method discloses using meaningful keywords from browsed URLs. However, it doesn't include prior user action or location-based interests. Further, there is a need to improve cosine similarity so that the profiling metric is suitable to ASN.

Trieu and Pham [20] introduced a system named STARS, which is an ad-hoc network of smartphones. It is a data sharing paradigm wherein users wish to share information with different people inside a small group for a reduced amount of time. The system implements features to create a social network and share interests like text, comments, pictures, etc. User records an identifier on the timeline, and the application advertises the id in the network. A decentralized application operating on user's device forms an Interest-based network and implement security and privacy requirement.

Li et al. [21] proposed location-based social network over MANET. The authors defined profile as a vector of keywords with weights. The profile similarity is computed based on if the keyword is present in both profiles or not irrespective of the weight of the keyword. This may have disadvantageous since weight is an important factor to determine as which pair of users are highly similar.

Zhang et al. [22] suggested a privacy-preserving approach to matching the profile of users in a decentralized manner for multi-hop ASN. The method protects privacy by not exposing profile of participants and the submitted preference profile due to a reliable communication channel connecting an initiator and matching users.

Zhang et al. [23] recommended Proximity-based mobile social networking (PMSN) that permits two users to complete profile matching without revealing any data about their profiles exceeding comparison result. However, the user has to pick interest level manually and to set values of different attributes that are numerous is a very tedious process.

Wang et al. [24] proposed a network of nodes connected using P2P. However, it doesn't offer any profile matching before being connected.

Wang et al. [25] proposed the idea of G-friends that stands for geographical location-based friends. The authors computed life-style vectors of users to used cosine similarity for the recommendation. However, cosine similarity as discussed later has disadvantages. Further, the need for a centralized server to process the lifestyle is a limitation for ASN.

Khan et al. [26] proposed status and challenges for P2P. The authors shows that P2P provides the data rate of upto 250 Mbps in range of around 200 meters and can be used for online gaming, streaming media, sharing content over P2P. The authors listed challenges of P2P to provide multi-hop routing. However, the mechanism doesn't allow using P2P without profile matching.

Table 2.2 summarizes the important features and provide remarks for different prior studies in the area of Profile Management of ASN. The comparison shows that there is need to work in the area of creating location-profile and optimizing profile matching approach.

Table 2.2: Comparison of related work for Profile Management

Paper	Features	Remark
Bottazzi et al. [14], 2007	Middleware for ASN to build and manage profile	Profile not refreshed dynamically
Campbell et al. [15], 2008	CenceMe	No weight to user search and browsing history and considers a global profile
Sarigöl et al. [16], 2009	AdSocial	No dynamic interests
Yiu et al. [17], 2010	Proximate Finds a group of users so that each pair of users have similarity within threshold	Uses Internet and cosine similarity and needs improvement in computing similarity and to implement for P2P
Sarigöl et al. [18], 2010	Provides buddy presence service	Profile keywords entered manually by user and doesn't represent location-based interests
Lee and Hong [19], 2011	Profile Management	No location-based interest and profile matching
Trieu and Pham [20], 2012	Supports interest-based network	Interest shared after connection
Li et al. [21], 2013	Location-based social network	No weightage to keywords
Zhang et al. [22], 2013	Match profiles of a user in a decentralized manner to protect privacy	Uses a centralized trusted server

Continued on next page

Table 2.2 – continued from previous page

Paper	Features	Remark
Zhang et al. [23], 2013	Matches profile of a user without revealing data	Profile interest lever set by user and no weight given to current location
Wang et al. [24], 2014	P2P	No profile matching before connected
Wang et al. [25], 2015	G-friends	Used cosine similarity that has limitations in weighted interest vector
Khan et al. [26], 2017	Shows online gaming, streaming media, sharing content over P2P	Allows using P2P without profile matching

2.4.3 Similarity Metric

Spertus et al. [27] analyzed several similarity functions, e.g., L1-Norm, L2-Norm (cosine similarity), Pointwise Mutual Information: positive correlations (MI1), Pointwise Mutual Information: positive and negative correlations (MI2), Salton (IDF), and Log Odds. The authors inferred that L2-norm or the cosine similarity is the proper similarity standard.

Li and Khan [28] proposes to shift existing social networking archetype towards MANET based social network. The authors used Ontological profile for semantic similarity matching and semantic-based distance vector routing protocol. The routing protocol helps to find users with similar interests. However, the routing protocol has limited use when the profile is not similar, or the threshold is high.

Anderson et al. [29] investigated as to how similarity in the properties of two users can influence the evaluation that one user gives to another.

The evaluation allows one user to say about another whether he or she admires the content of another or trusts another user or not. The authors found that analysis of user-to-user evaluations can be enhanced by taking account of similarity in users. Cosine similarity was also applied to examine that how user similarity influences evaluations.

Sanguankotchakorn et al. [1] posed a problem of finding nodes of similar interests in the social network over MANET.

Symeonidis et al. [30] proposed recommender system that uses location history to determine users' similarity.

Li et al. [31] proposed a design of location-based social network over MANET. The authors used similarity of users regarding common interests as the social relation. Each node only connects to nodes that are nearby and have common interests; therefore, search success is increased and overhead decreased.

Liaqat et al. [32] exploited similarity-matching social properties of intermediate nodes to maximize bandwidth utilization in ASN.

Han et al. [33] measured interest similarity depending on the probability of sharing interest and degree of interest similarity utilizing weighted cosine similarity.

Mizzaro et al. [34] applied cosine similarity on Twitter to interpret semantic relations linking the words occurring in the same tweet and the similar topics.

Zhang et al. [35] recognized users from different social networks to build a profile by uniting the various profiles over the networks. The profile is represented as a bag-of-words vector, where the words are weighted by term frequency-inverse document frequency (TF-IDF). The authors also calculated profile content similarity by computing inner product and cosine distance.

Kraus et al. [36] suggested Locality Sensitive Hashing (LSH) that restricts the search to collections of objects, called buckets, that have a large probability to be similar to the query and further, used cosine similarity function.

Yu et al. [37] proposed a method to recommend friend suggestion based on point-of-interest and check-in data as who can provide more information.

Several other researchers used cosine similarity in the area of social networks and addressed the intricacy to analyze and improve profile similarity algorithm.

Table 2.3 summarizes the important features and provide remarks for different prior studies in the area of Similarity Metric of ASN.

Table 2.3: Comparison of related work for Similarity Metric

Paper	**Features**	**Remark**
Spertus et al. [27], 2005	Analyzed various similarity functions and recommended cosine similarity	Cosine similarity has limitation for ASN interest vector with weights assigned to different interests
Li and Khan [28], 2009	Ontological semantic profile matching	Limited use when profile not similar or the threshold is high
Anderson et al. [29], 2012	Studies effect of similairty of two users in evaluation of another user	Used cosine similarity that can be improved for ASN
Sanguankotchakorn et al. [1], 2012	Proposed challenges in searching similar nodes over MANET	Nodes similarity is a challenge
Symeonidis et al. [30], 2014	Location history to determine users' similarty	The locations previously visited may not reprresent current-location interests

Continued on next page

Table 2.3 – continued from previous page

Paper	Features	Remark
Li et al. [31], 2014	Used similarity of users as the social relation	Static profile without considering different weight to different keywords
Liaqat et al. [32], 2015	Used social properties of intermediate nodes for bandwidth utilization	Controls data rate based on similarity
Han et al. [33], 2015	Utilized weighted cosine similarity based on probability of sharing interest. Assigns less weight to popular interests	No weight to the frequency of a keyword
Mizzaro et al. [34], 2015	Created profile from keywords used in twitter and used cosine similarity	Needs improvement to implement it for ASN to extract text
Zhang et al. [35], 2015	Created a profile by uniting various profile over different social networks	Used cosine similarity that needs improvement for ASN
Kraus et al. [36], 2016	Used cosine similarity over collection of objects that have large probability to be similar	Needs improvement in profile similarity for ASN
Yu et al. [37], 2017	Recommend friend suggestion based on check-in data	Considers historical data without giving weight to location-based interests

2.4.4 Routing Protocols

Li and Khan [28] proposed MobiSN for self-configured mobile ASN implemented in Java. MobiSN is a semantics-based framework and presents functions and services such as friend matchmaking, generation of ontology-based profiles, and automatic forming of groups, etc.

However, MobiSN considers shared ancestor and root notion in determining the similarity. Further, it is challenging for developers to build applications and services for ASN to meet complex requirements of diverse ASN users.

Kayastha et al. [38] studied data delivery services concerning the social relationship among mobile users and proffered network architecture and protocol design issues. One of the open research problems and future directions that authors suggested is using context-aware data distribution in Mobile peer-to-peer network.

Ahmed et al. [39] proposed community-partition aware replica allocation for ASN to increase the availability of data in partition social community.

Liaqat et al. [40] presented challenges in scheduling algorithms in ASN and proposed Pop-aware scheduling algorithm that computes traffic at the intermediate node and assigns priority to incoming traffic based on centrality, which is social property.

Marinho et al. [41] expanded Wi-Fi Direct technology to transmit the information over multi-hop and measured the number of exchanged messages using routing protocols: (i) flooding, (ii) Ad hoc On-demand Distance Vector (AODV), (iii) AODV-Backup Route (AODV-BR), and (iv) Location-Aided Routing (LAR).

Xia et al. [42] presented a review of socially aware networking to utilize social characteristics of nodes. The base of socially aware networking is that people usually carry the mobile devices. The mobility of devices or nodes is truly due to the movement of users and by examining social relationships, getting the mobility regularities of mobile devices as well as predicting the connection opportunities can be applied to devise routing protocols. Creating routing protocols based on social patterns is beneficial but is not favorable if nodes diverge from the social pattern. Accordingly, there is a requirement for routing protocol that attains the best path based on user-defined parameters.

Palani et al. [43] recommended utilizing interest and contact occurrence for effective file distribution. The authors stated applying interest mining that is the identification of nodes interests, area structure, constructs familiar-interest nodes with shared contacts into communities.

Gupta et al. [44] proposed social-tie-strengths-based routing in MANET, where nodes travel in group and mobility is planned. The method is particularly useful in military and emergency situations where soldiers or unmanned vehicles travel as per defined plan. The social tie based strength considers the frequency of encounters. The routing algorithm has two activities (i) table exchanges and (ii) route selection. Table exchange activity provides sufficient information to nodes to select reliable path during route selection. The protocol uses social networking for routing decision. However, it is not a general protocol that can be used to deploy ASN where the mobility is not planned, and the nodes are not aware of each other.

Table 2.4 summarizes the important features and provide remarks for different prior studies in the area of Routing Protocols of ASN.

Table 2.4: Comparison of related work for Routing Protocols

Paper	**Features**	**Remark**
Li and Khan [28], 2009	Proposed semantics-based distance-vector query routing that allows a node to make routing decision by knowing its immediate neighbors and limited resource information	Overhead in case of high mobility

Continued on next page

Table 2.4 – continued from previous page

Paper	**Features**	**Remark**
Kayastha et al. [38], 2011	Presented survey of a mobile social network for its applications, network architectures, and protocol design issues	Challenges in community detection, content distribution, mobility, and privacy
Ahmed et al. [39], 2013	Proposed data replication to avoid data losses in case of unpredictable community or network partition in ASN	Used social context in the community-partition aware replica allocation method
Liaqat et al. [40], 2014	proposed scheduling algorithm for ASN that computed traffic load and prioritize incoming flow using degree centrality	Used social property in scheduling packets
Marinho et al. [41], 2015	Tested routing protocols e.g. AODV over P2P and found routing load ok for two or three devices	Context aware routing can help to reduce the routing load
Xia et al. [42], 2015	Reviewed social aware networking and presented open challenges in mobile social sensing, privacy, selfishness, and scalability	The survey indicates that social property of a node is a powerful source for designing network and routing and forwarding protocols
Palani et al. [43], 2015	Proposed a social network-based peer-to-peer substance file distribution system in disjointed mobile ad hoc Networks	Used interests of a node and contact occurrences for file distribution

Continued on next page

Table 2.4 – continued from previous page

Paper	Features	Remark
Gupta et al. [44], 2017	Proposed routing algorithm of nodes for which the planned mobility of the nodes are partially known and the nodes travel in groups	Use social tie as strength as a measure

2.5 Comparison Analysis of Related Work over ASN Features

Table 2.5 compares related work for different ASN features including *Architecture and Implementation; Profile Management; Similarity Metric; Routing*. The comparison analysis shows the need of location-based profile management over P2P.

Table 2.5: Comparison of related work for ASN parameters Protocols

Paper	Architecture / Implementation	Profile Management	Similarity Metric	Routing
Zhang et al. [7], 2010	Virtual NIC	Content shared to all	Not Applicable	DSR
Zhang et al. [8], 2014	Modified Android	Aggregates profiles and stores in a repository and grants access to other users	Not Applicable	Uses WiFi, SIP, Bluetooth, and TCP/UDP
Kraus et al. [36], 2016	P2P Architecture	Profile managed by users	Cosine Similarity	P2P

Continued on next page

Table 2.5 – continued from previous page

Paper	Architecture / Implementation	Profile Management	Similarity Metric	Routing
Funai et al. [45], 2017	P2P over modified Android	Not Applicable	Not Applicable	Gateway node switches between different P2P goups
Khan et al. [26], 2017	Shows online gaming, streaming media, sharing content over P2P	Not Applicable	Not Applicable	Proposed multi-hop P2P as a challenge

2.6 Summary

This chapter explained the ASN importance, characteristics, and the recent work done by various researchers in the area of ASN that provides new challenges due to local communication as indicated by a survey of related studies. The survey indicates lack of interest matching with people nearby, which is important for location-based social networking or ASN. The survey for profile management indicates the need to improve profile management in the area of creating a location-based profile based on users' prior actions without getting connected to a central server. The prior study also used the traditional similarity metrics and routing protocol, and there is a need to improve or customize using social features of ASN.

Chapter III

PROPOSED WORK: ARCHITECTURE OF ASN AND USERS' PREFERENCES

Chapter III

PROPOSED WORK: ARCHITECTURE OF ASN AND USERS' PREFERENCES

Review of prior work indicates need to have a middleware architecture in accordance to users' preferences. The middleware architecture as proposed in this chapter describes functions and methodology required to establish ASN. The results from users' survey for their preferences are also presented in this chapter.

3.1 Introduction

ASN introduces a layer of social network on the top of MANET. Figure 1.1 shows a sample scenario where the mobile nodes U_1, U_2, U_3, U_4, U_5 forms MANET. In a network of five nodes, three nodes U_1, U_4, U_5 forms a social network. While previous works have introduced ASN architecture, this research proposed a general middleware architecture that can be adapted to create ASN with distinct features and at the same time provide extensibility to fix specific parameters. One of the reasons that middleware is beneficial is that it simplifies the profile management, peer-discovery, and peer-communication process.

This chapter, instead of solving specific issues, proposes a design of middleware architecture that identifies and discusses the essential components to establish ASN. The middleware components include exchanging users' profiles and providing interaction among ASN users. The suggested architecture comprises overlay modules for distributing information and maintaining the ASN, thereby presenting a comprehensive platform-independent solution for application developers to build ASN and other applications.

This chapter also discusses users' perception on ASN and the need to establish the social network. Users' perception is computed using online questionnaire survey.

3.2 Middleware Architecture

Middleware is a software that connects different parts of an application or various applications and presents services to the applications beyond those accessible from the operating system. It provides application programming interface (API) for utilizing operating system features.

The proposed architecture [46] comprises four layers *Application Layer, Transport Layer, Ad-hoc Social Layer, Ad-hoc Communication Layer*.

The application layer assists users to use the ASN by receiving social links and also to configure the parameters of other layers or modules.

The transport layer results in end-to-end communication for the ASN and ensures the transcendent throughput per connection by fixing congestion due to gradual changes in the topology.

The ad-hoc social layer handles user profiles and props to build a social network over MANET.

The ad-hoc communication layer addresses concerns related to the administration of the network amongst nodes by providing essential communication services among interested nodes.

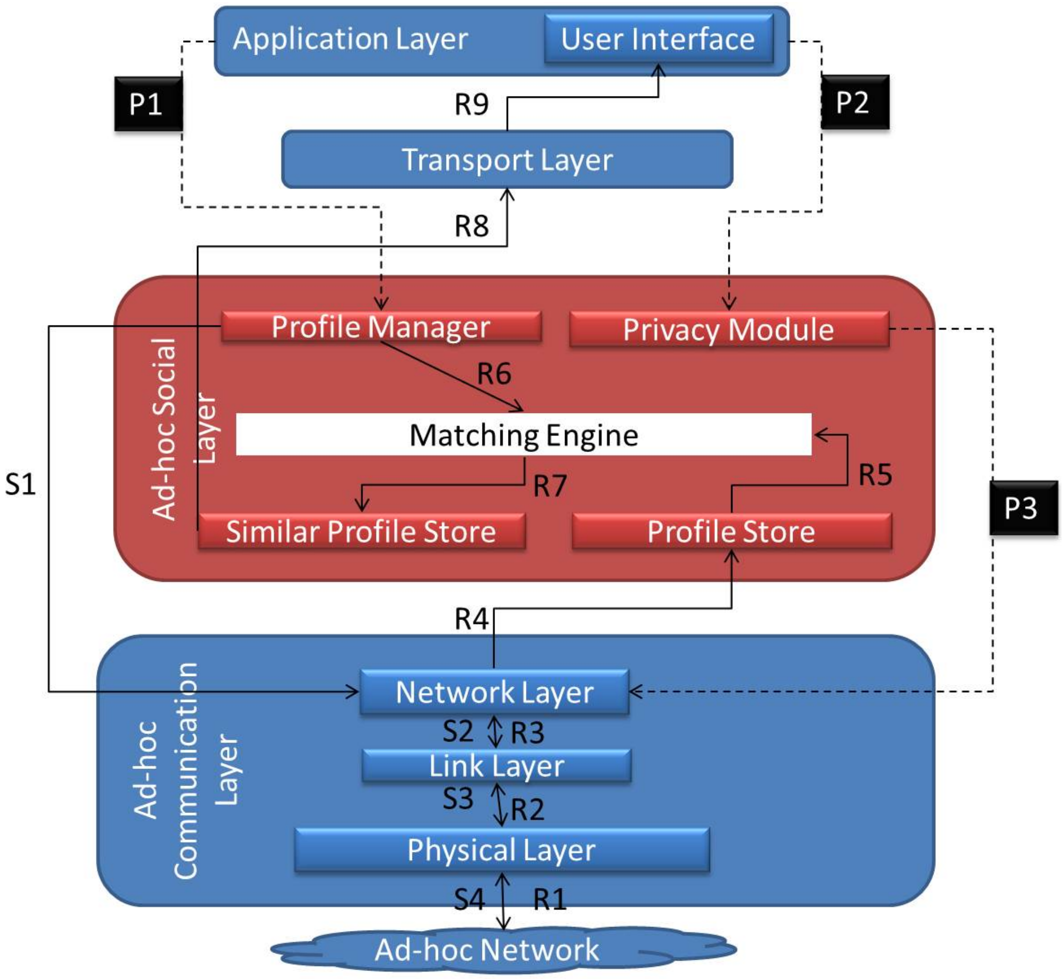


Figure 3.1: Middleware Architecture for Ad-hoc Social Network (ASN)

Figure 3.1 demonstrates the proposed architecture for end-to-end communication in an ASN. The labels R_n, and S_n denote the progress of accepting and transmitting profile data packets from one node to another and P_n depicts processing step, where n is the step number. The profile packets of other nearby users are received at physical layer (step $R1$), link layer (step $R2$), network layer (step $R3$), profile store (step $R4$), matching engine (others' profile from profile store in step $R5$ and from self-profile from profile manager in step $R6$), similar profile store (step $R7$), transport layer (step $R8$), and user interface (step $R9$).

Similarly, the profile of a user is broadcasted to nearby users from profile manager to network layer (step $S1$), the link layer (step

$S2$), physical layer (step $S3$), and to other nearby users via wireless communication (step $S4$). The processing steps labeled as $P1, P2$, and $P3$ allows to pass configured parameters to lower layers as selected by a user. For example, a user may select a location-based profile, global profile, or customized profile and this value is passed to Profile Manager (step $P1$). Similarly, a user may want to limit broadcasting its profile to a fixed maximum number of hops, which can be set to configure time to leave or Time-To-Live (TTL) field in the privacy module in the ad-hoc social layer (step $P2$) and further in the lower layer (step $P3$).

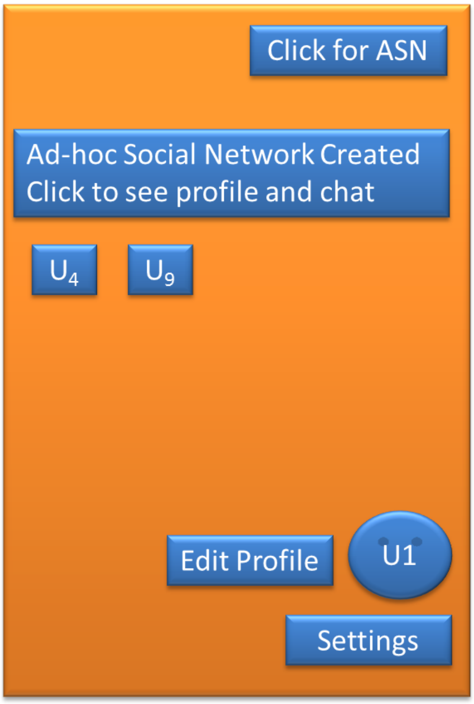


Figure 3.2: User Interface to start ASN

Figure 3.2 demonstrates a high-level user interface to set different values as per user preferences for ASN. The button labeled "Click for ASN" broadcasts the user profile to nearby users (steps $S1 - S4$). Once a user device receives the profiles of other nearby users (steps $R1 - R9$), it computes the profile similarity and displays the similar profile (e.g., $U4$ and $U9$), that can be clicked to chat and exchange the messages, files, images, etc. A user may edit or update its profile by clicking $U1$ (step $P1$) or configure other privacy settings (step $P2$) by clicking the button labeled "Settings".

3.2.1 Application Layer

The application layer presents a user interface to manage user profile and to track other nearby users with similar interests for exchanging text, pictures, video, etc. The user interface also incorporates settings to fix parameters of the ad-hoc social layer and ad-hoc communication layer. User U_1 may utilize the user-interface to control its profile, e.g., the algorithm may initiate the location-based profile, but the user can override the default option. User U_1 may also utilize the interface as displayed in Figure 3.2 by clicking a button named "Click for ASN" to start the ASN followed by chatting or sharing images. The profiles of similar users are received by data packets via communication path $R_1 - R_9$. The application layer further facilitates the user to enter the TTL value (step $P2$). The value of TTL is used by lower layers to constrain the visibility of profile over the number of hops. For example, if a user $U1$ is not interested in establishing ASN with users higher than three hops, then $U1$ may initiate TTL for three hops. The user may also fix similarity as a medium, high, or very high in the settings.

3.2.2 Transport Layer

The transport layer offers end-to-end communication for the ASN and assures the highest throughput per connection. The transport layer protocol has a mechanism to regulate congestion following dynamic changes in the topology and possible bandwidth. The transport layer receives the same profiles as discussed after this.

3.2.3 Ad-hoc social layer

The ad-hoc social layer is an essential layer from the research perspective because not enough work has been proposed regarding components or services for this layer. This study introduces an ad-hoc social layer that encompasses the following elements:

3.2.3.1 Profile manager

The profile manager maintains the user profile including features, e.g., name, birth date, gender, education, profession, buddies, internet search record, browsing history, GPS of places visited and core characteristics of profiles of persons with prior chat records. The profile can be a static profile, dynamic profile, or customized as per user's preferences. The profile may be set based on the current GPS position. The profile manager transmits a relevant profile of the user to the matching engine(step $R6$) and initiates broadcasting the profile (step $S1$). The profile vector of a user can be defined as:

$User_A = \{\{\text{keyword 1, keyword 2, keyword 3}\}; \{n_1, n_2, n_3\}\}$

Keywords $keyword_i$ represents user interests and n_i represent the weight of the corresponding keyword. A higher value of the weight indicates higher interest level for that particular keyword.

3.2.3.2 Profile store

The profile store saves the profile of every user obtained by step $R4$. The collected profiles are given to matching element. The profile store is refreshed dynamically based on the data obtained from neighbors as fresh members enter or depart ASN.

3.2.3.3 Matching engine

The matching engine compares the user's profile with profiles acquired from the profile repository called profile store. Similar profiles that are alike with more than the threshold limit are shifted to "Similar Profile Store" as illustrated by $R7$ in Figure 3.1. The matching operation can be based on syntax or semantics. The matching engine compares the profile of $U1$ with saved profiles based on the similarity selections established by $U1$. For clearness, let us consider that the matching engine agrees on profiles of $U1$ with $U4$ and $U5$ while discarding others as shown in Figure 1.2.

3.2.3.4 Similar profile store

The similar profile store keeps the profiles that satisfy the least threshold of similarity. Hence, the similar profile store will save profiles and connection features of $U4$ and $U5$ because of higher similarity. If $U1$ and $U5$ participate in ASN and exchange messages, the stored profile of $U5$ will be considered as a preferred profile for $U1$.

3.2.3.5 Privacy module

The privacy module enables users to regulate settings (step $P2$) regarding who can get the profile concerning the number of hops, thus, controlling the user's visibility regarding hops via TTL. The TTL is introduced as the number of hops a packet is allowed to traverse before being dropped. Each node will reduce the profile TTL range by one unit, and the packet is rejected when TTL field equals zero (step $P3$).

3.2.4 Ad-hoc Communication Layer

The ad-hoc communication layer is intended to be adaptable in choosing the optimal data link, routing and application protocols. This layer encompasses a network layer, link layer, and physical layer. The ad-hoc communication layer serves to connected devices to flow messages in the network without employing a base station. Each node of ASN aids to route data packets between different nodes.

3.2.4.1 Network layer

The network layer incorporates routing algorithms utilized to transmit packets in the ad-hoc network. A routing algorithm is selected in accordance to the ASN if it is dense or sparse. For example, flooding may be employed if the network is sparse. The network is classified as sparse or dense based on edge density, which is determined as the ratio of existing links (m) to the total number of possible links as shown in equation 1. The highest link density D of a fully connected network is one.

$$D = \frac{m}{0.5 * N * (N-1)} \tag{1}$$

3.2.4.2 Data link layer

The data link layer manages data link connections linking nodes. The data link component keeps records of neighboring nodes and manages the connections. It supports transmitting and receiving routing data as per the routing table. Data is forwarded straight to the destination if devices are within reach of one hop; otherwise, to an intermediary device if source and destination are at longer than one-hop distance. This component handles the connection between nodes, e.g., $U1, U2, U3, ...Un$. The communication includes transmitting hello packets comprising profile information, maintaining the network and controlling failures in the packets.

3.2.4.3 Physical Layer

The physical layer comprises a transceiver to transfer and receive data packets. Consideration is required concerning signal reception, interference and noise and preamble length.

3.3 Middleware Architecture Applications

The proposed middleware architecture can be used to develop various ASN applications that including sharing content e.g. text, images, audio, video with similar users. The need of profile matching is important since it will help users to have sufficient information about other nearby users before accepting connection request.

3.4 User Preferences for ASN

This section of the chapter presents results from a user survey. This research surveyed users to understand users' preferences and perception about ASN.

3.4.1 Methodology

The intentions of this study [47] were to know users preferences for ASN and in particular the preferred percentage value of profile similarity. A survey questionnaire [48] was prepared using Google Docs and made available online to potential participants. The link to the questionnaire and request to complete was shared personally by email and social networking sites including Facebook and LinkedIn. The social networking sites Facebook and LinkedIn were selected since users of these sites are already aware of the social networking applications and become relevant participants for the study.

The questionnaire asked specific questions with multiple-choice answers. Questions 1, 2, and 3 form a group to determine relevancy of a respondent for the study, e.g., Question 1 was to know if a respondent is using any social networking application on a mobile device or not, Question 2 asked how many times generally a user uses social application on a mobile device, Question 3 asked if a user uses mobile browser to search information on a mobile device.

Although Google Trends [49], reporting top search terms for different regions, indicates that search and browsing pattern and thus interests are different in various geographical regions, Question 4 was specific to know from respondents if they also feel this. Thus, Question 4 asked if the search query of the users on a mobile device varies with their location whether they are at home, office, or at some other location. Questions 5, 6, and 7 were asked to know their opinion if the users are satisfied with online social applications or they prefer location-based ASN. Question 5 asked if the current social networking applications that they are using also help them to expand their social network, Question 6 asked if a respondent is interested in getting notification of another user who is nearby and have similar interests, Question 7 particularly asked about ASN if the respondent is interested in social networking using MANET. Question 8 was the final question to know how much similarity a user prefers when looking for a new social connection.

3.4.2 Results and Analysis

The questionnaire was divided into four sections comprising:

Section 1: Trend of social networking on mobile device

Q1: Do you use Social Networking sites, e.g., Facebook, LinkedIn, Google+ on Mobile Phone or Tablet?

Q2: What is the frequency of visiting these online social networking sites on your mobile phone or tablet?

Q3: Do you search information on Google, Bing, or any other search engine using a mobile device?

Section 2: Importance of location-based interests for social networking

Q4: To what extent your search query vary with respect to your location. e.g. you search different information at your home (like movies, recipes etc) and different information outside your home (like sharing a cab, restaurants etc

Section 3: Satisfaction level with current social networking applications

Q5: Does these social networking sites provides enough opportunities to know people of your interest anytime and anywhere?

Q6: Are you interested in getting information on your mobile device if some person with interests similar to you is present nearby you?

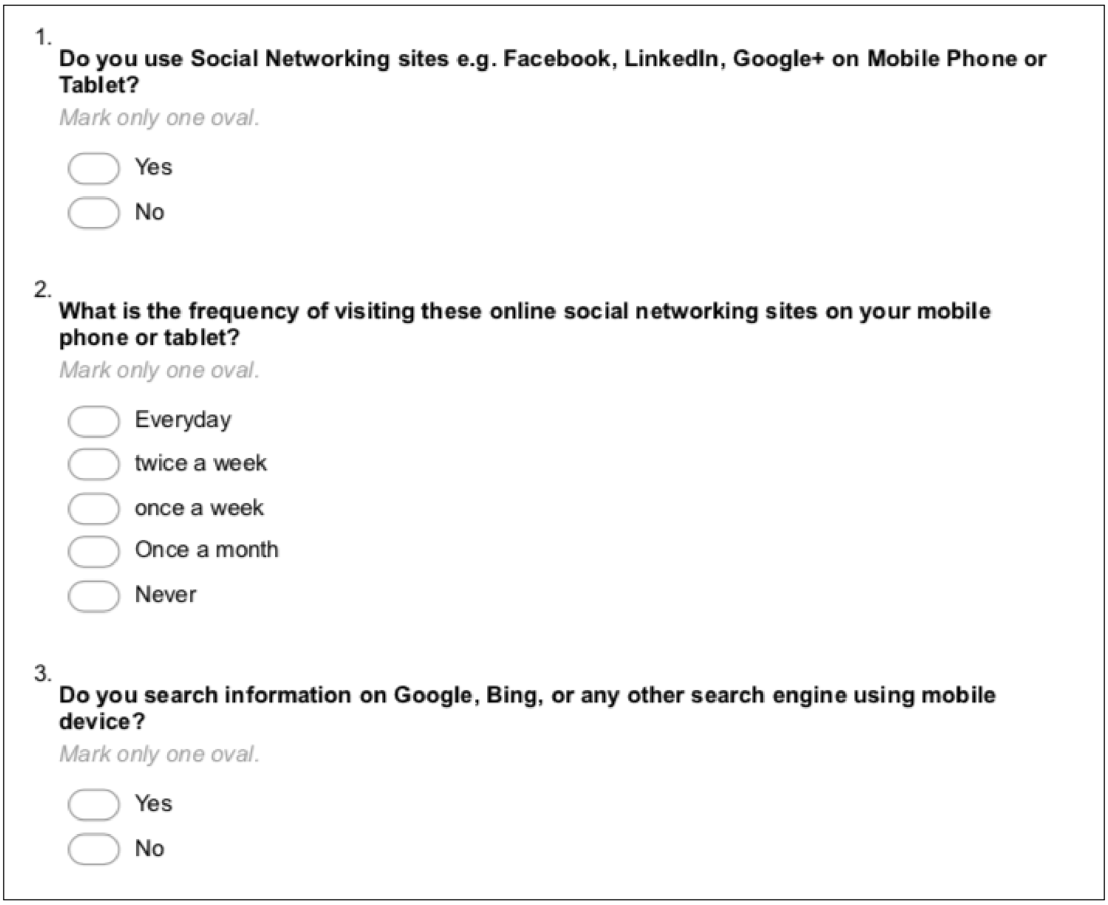
Q7: Would you like to access a location-based social networking site which can suggest you friends based on your present interest, context and GPS location that doesn't use your mobile internet connection?

Section 4: Preferred value of profile similarity

Q8: How much interest similarity you would like to have with your friend on such social networking site?

The total number of 102 respondents completed the questionnaire that includes online and offline submissions. The respondents are classified as 82% from the academic field include engineering students and faculty members and 18% from industry including professionals using online social networking.

3.4.2.1 Trend of Social Networking on the Mobile Device (Q1-Q3)



1.
Do you use Social Networking sites e.g. Facebook, LinkedIn, Google+ on Mobile Phone or Tablet?
Mark only one oval.

- Yes
- No

2.
What is the frequency of visiting these online social networking sites on your mobile phone or tablet?
Mark only one oval.

- Everyday
- twice a week
- once a week
- Once a month
- Never

3.
Do you search information on Google, Bing, or any other search engine using mobile device?
Mark only one oval.

- Yes
- No

Figure 3.3: Q1-Q3: Usage of Social Networking on the Mobile device

Figure 3.3, displays Question 1, Question 2, and Question 3 asks to know current device trends in accessing social networking applications. The questions are (i) Do you use Social Networking sites, e.g., Facebook, LinkedIn, Google+ on Mobile Phone or Tablet?, (ii) What is the frequency of visiting these online social networking sites on your mobile phone or tablet?, (iii) Do you search information on Google, Bing, or any other search engine using a mobile device?

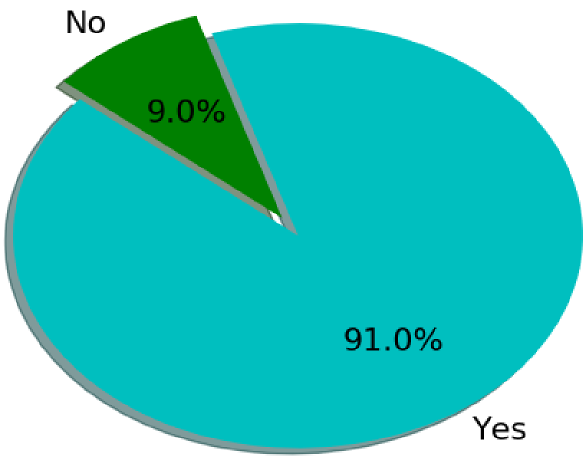


Figure 3.4: Results (Q1): Do you use Social Networking sites, e.g. Facebook, LinkedIn, Google+ on Mobile Phone or Tablet?

Figure 3.4 shows results for Question 1, illustrating that 91% of respondents agree using social networking applications on the mobile device.

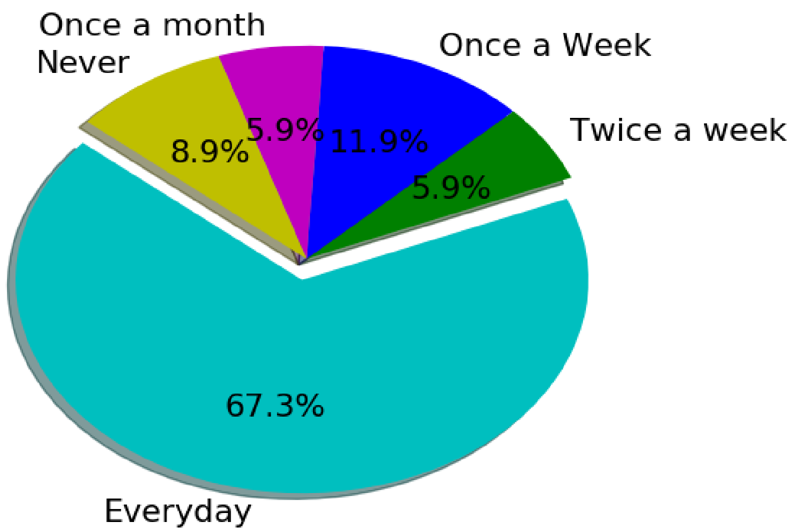


Figure 3.5: Results (Q2): What is the frequency of visiting these online social networking sites on your mobile phone or tablet?

Figure 3.5 shows results for Question 2, illustrating 68% of respondents agree that they are using social networking application every day at their mobile device indicating a growing trend of the mobile device for social applications (Figure 3.5). 18% of users indicated usage of social networking application on a weekly basis (either once a week or twice a week). There was a low number of users, e.g., 6% users using social networking application on a monthly basis and 9% of users not using the applications on the mobile device.

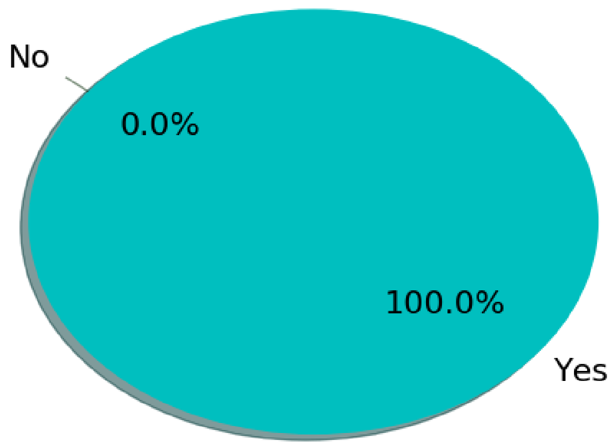
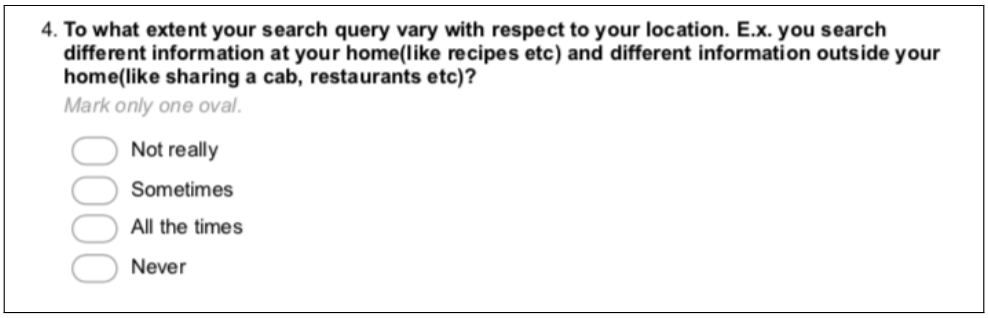


Figure 3.6: Results (Q3): Do you search information on Google, Bing, or any other search engine using mobile device?

Figure 3.6 shows results for Question 3, illustrating that people are indeed using a mobile device for search and browsing information and Internet since 100% users said usage of the mobile device for searching information using Google, Bing, or other browsers.

3.4.2.2 Variation of Search and Browsing history with geographical location (Q4)



4. To what extent your search query vary with respect to your location. E.x. you search different information at your home(like recipes etc) and different information outside your home(like sharing a cab, restaurants etc)?

Mark only one oval.

- Not really
- Sometimes
- All the times
- Never

Figure 3.7: Q4: To what extent your search query vary concerning your location. e.g., you search different information at your home (like movies, recipes, etc.) and different information outside your home (like sharing a cab, restaurants, etc.)

Figure 3.7, displaying Question 4, asks to know if the search query varies with location. This is to know if the profile matching that builds profile utilizing local search queries is beneficial to the user or not.

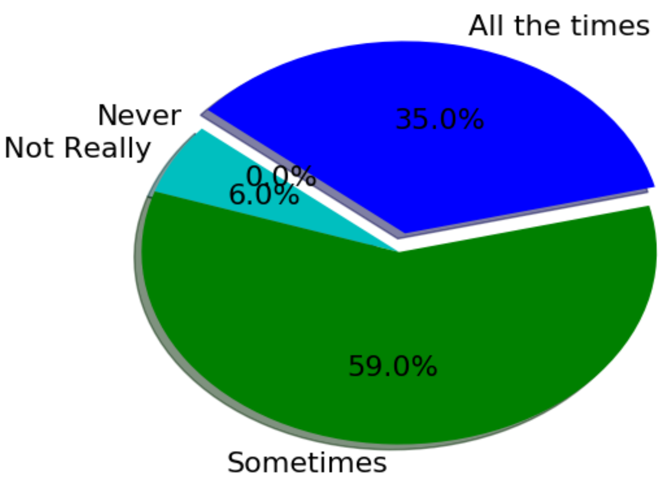
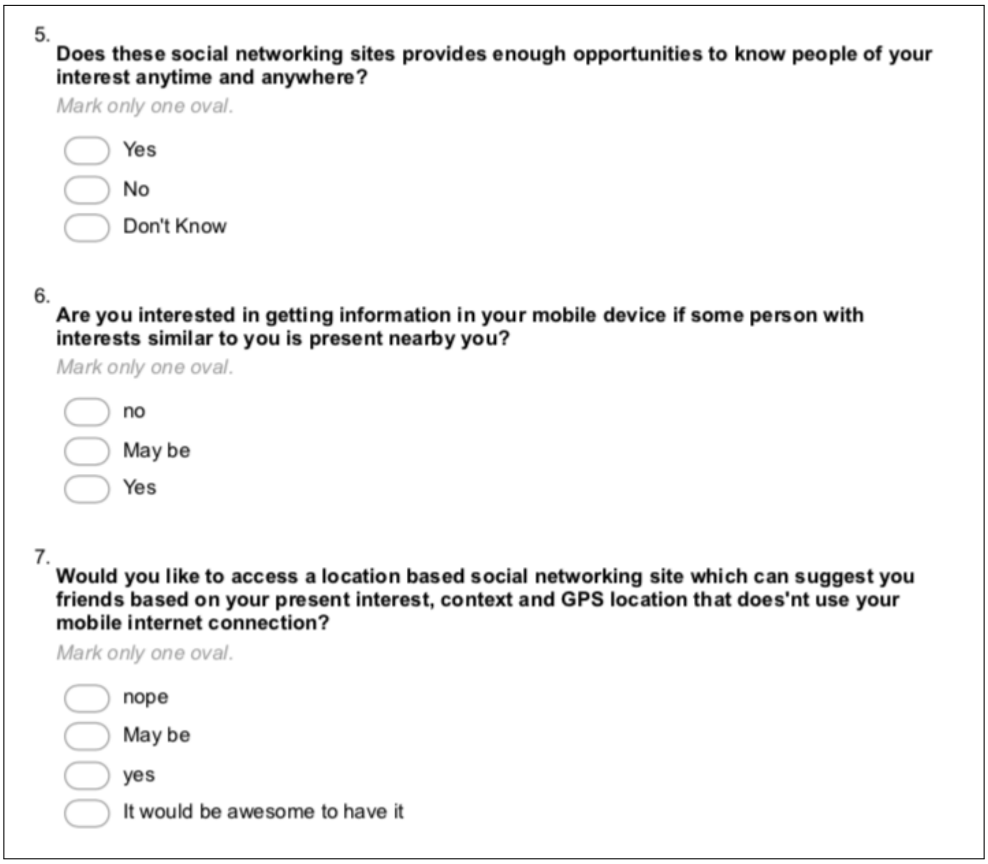


Figure 3.8: Results(Q4):Location-dependent interests

Figure 3.8 shows results for Question 4, illustrating increasing trend of geographical location-based interests since 94% of users said that their queries either always (35%) or sometimes (59%) are dependent on their current location. This also shows the need of using dynamic interests for social networking rather than using static interests currently being used by online social network applications.

3.4.2.3 Satisfaction level for existing social networking application (Q5-Q7)


5.
Does these social networking sites provides enough opportunities to know people of your interest anytime and anywhere?
Mark only one oval.

- Yes
- No
- Don't Know

6.
Are you interested in getting information in your mobile device if some person with interests similar to you is present nearby you?
Mark only one oval.

- no
- May be
- Yes

7.
Would you like to access a location based social networking site which can suggest you friends based on your present interest, context and GPS location that does'nt use your mobile internet connection?
Mark only one oval.

- nope
- May be
- yes
- It would be awesome to have it

Figure 3.9: Q5-Q7: Satisfaction level for existing social networking application

Figure 3.9 displays Question 5, Question 6, and Question 7 asks to know the satisfaction level of existing social network applications.

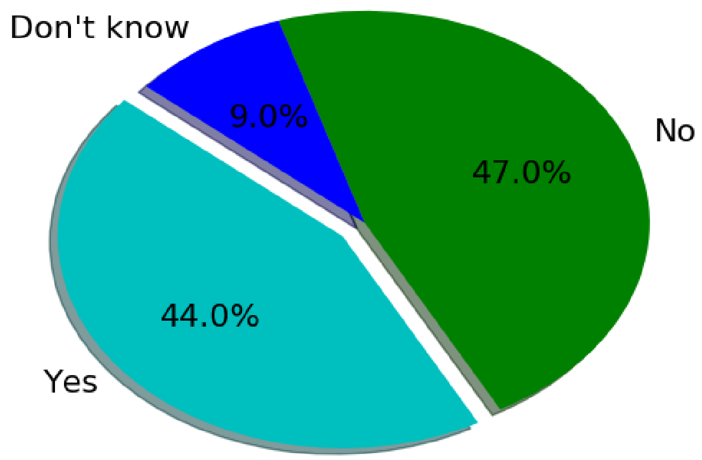


Figure 3.10: Results(Q5): Does these social networking sites provides enough opportunities to know people of your interest anytime and anywhere?

Figure 3.10 shows results for Question 5, illustrating whether current social networking applications provide sufficient opportunities to extend the social network, the opinion was equally divided, 47% of users said that current social networking applications don't provide sufficient opportunities and 44% of users are satisfied with the existing applications.

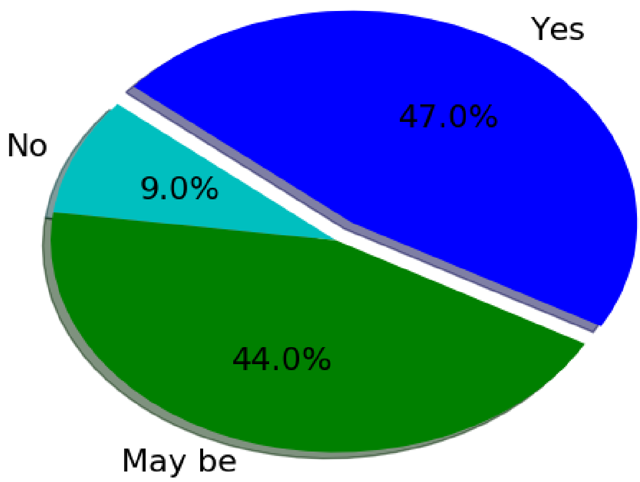


Figure 3.11: Results(Q6): Are you interested in getting information in your mobile device if some person with interests similar to you is present nearby you?

Figure 3.11 shows results for Question 6, illustrating 47% of users are interested in receiving notification about a nearby person if the interest matches, while 44% of users were not sure about receiving notification.

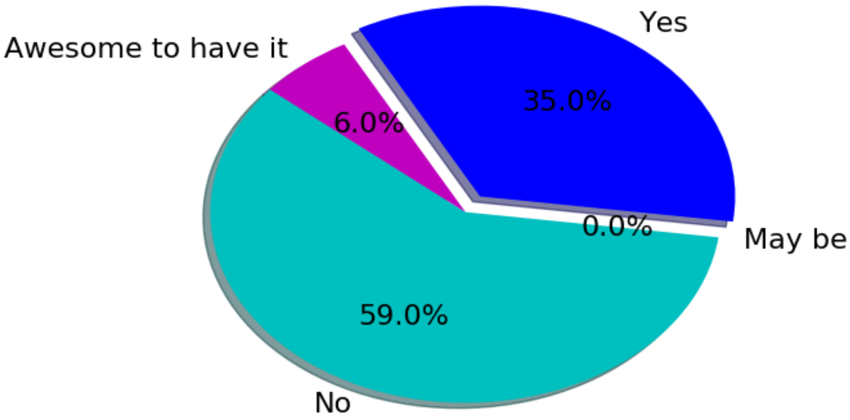
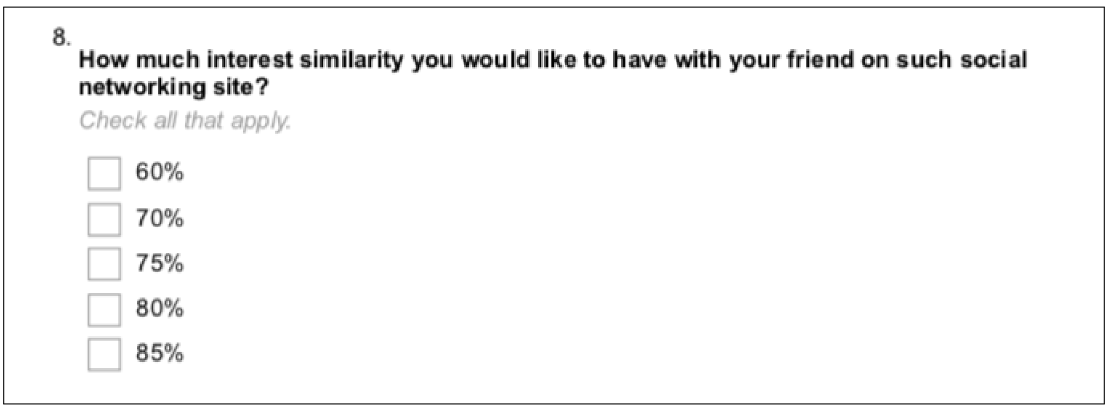


Figure 3.12: Results(Q7): Would you like to access a location based social networking site which can suggest you friends based on your present interest, context and GPS location that doesn't use your mobile internet connection?

Figure 3.12 shows results for Question 7, illustrating 43% of users were interested in ASN (35% responded yes and 6% responded awesome to have it).

3.4.2.4 Preferred value of Profile Similarity



8.
How much interest similarity you would like to have with your friend on such social networking site?
Check all that apply.

- [ ] 60%
- [ ] 70%
- [ ] 75%
- [ ] 80%
- [ ] 85%

Figure 3.13: Q8:How much interest similarity you would like to have with your friend on such social networking site?

Figure 3.13, displaying Question 8, asks to know the users' preferred value of profiles similarity.

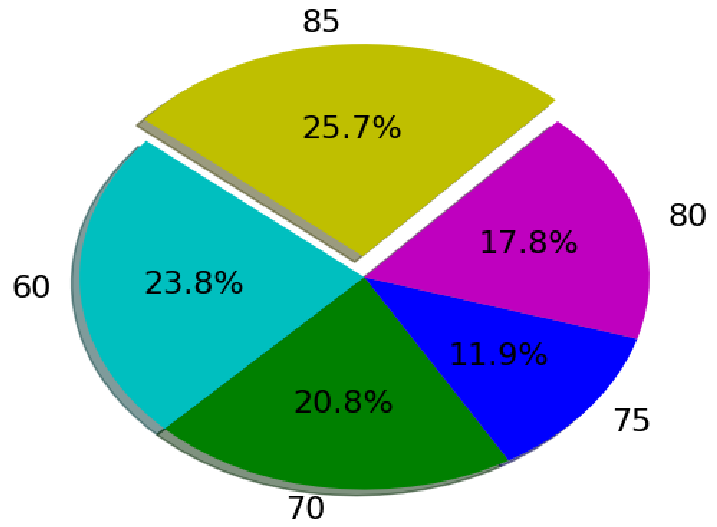


Figure 3.14: Results(Q8): Preferred value of Profile Similarity

Figure 3.14 shows results for Question 8, illustrate the preferences for a value of Profile Similarity when users want to receive notification about another user for possible new social connection. 24% of users preferred 60% similarity, 21% of users preferred 70% similarity, 12% of users preferred 75% similarity, 18% of users preferred 80% similarity, and 26% of users preferred 85% similarity. Based on the results, it can be assumed that users on the average prefer 75% similarity.

3.5 Summary

This chapter presented layered-architecture of ASN and results of users' survey to determine the users' preferences for ASN. The system architecture can be used to build and maintain ASN to exploit contact opportunities between friends and other people sharing similar interests. The layered-architecture comprising layers Application Layer, Transport Layer, Ad-hoc Social Layer, and Ad-hoc Communication Layer can also be used to provide an API to develop a third-party application. The ASN social functions e.g. profile manager, profile store, matching engine, similarity profile store, and privacy module under the ad-hoc social layer.

This chapter also provides results from a user survey conducted to know users' preferences for ASN. The results indicate growing interests towards ASN and users' need features per the proposed layer-architecture. 91% users are accessing social networking applications from a mobile device, while 68% are using the social networking applications every day. All users indicated usage of the search engine, e.g., Google or Bing to search for the desired information indicating that users' interests can be extracted from search and browsing history. Regarding location-based interests, 94% of respondents acknowledged that their interests change with the change of location always (35%) or sometimes (59%). Further, users' preferred similarity score for a profile with nearby users is 75%.

Chapter IV

PROPOSED WORK: OPTIMIZATION OF PROFILE MATCHING AND SIMILARITY METRIC

Chapter IV

PROPOSED WORK: OPTIMIZATION OF PROFILE MATCHING AND SIMILARITY METRIC

Results from users' preferences indicate that there is a growing trend of dynamic interests where the interests are location-dependent. One of the possible ways to determine the location-based interests is to use the users' prior interaction e.g. search and browsing history at the mobile device for a particular location. This chapter presents a location-based profile matching algorithm and further improved profile similarity metric.

4.1 Introduction

Profile of a user is an important parameter to recommend friends in a social network since connected users represent the success of a social network. Thus, the profile must be meaningful and well-organized. One of the objectives of ASN is to connect people with similar interests when the people are nearby. Thus, the strength of ASN lies over strong profile management that includes features like building the profile, organizing profiles, and matching profiles. This chapter discusses proposed profile Matching algorithm, profile Similarity Metric, and social-aware routing protocol.

4.2 Geo-Social Profile Matching Algorithm

Lee and Hong [19] proposed a profile as a set of keywords extracted from human-readable keywords from URLs browsed by a user. The profile built using this mechanism is dynamic. However, the assumption that URL has meaningful words is not always true. Further, considering previously browsed URLs or browsed URLs at different locations to determine recent dynamic interests is also not suitable for Location-based social networks. According to a survey by Gambhir et al. [47], 35% users feel that most times their search queries varies over the geographic location and 59% users feel that the variation is sometimes. Thus, 94% of users responded variation of queries or indirectly variation of interests depending on the location. Thus, there is a need for managing location-based dynamic profile.

This chapter proposes a hierarchical profile model [50] with a root node as a location. All browsed URLs can be divided into some locations with a predetermined radius, and the current interests are defined as interests of the cluster that is near to current location.

4.2.1 Algorithm

The algorithm creates multiple location-based profiles using prior user actions including keywords used in search history, keywords appearing in URLs, metadata of browsed web pages or mobile applications.

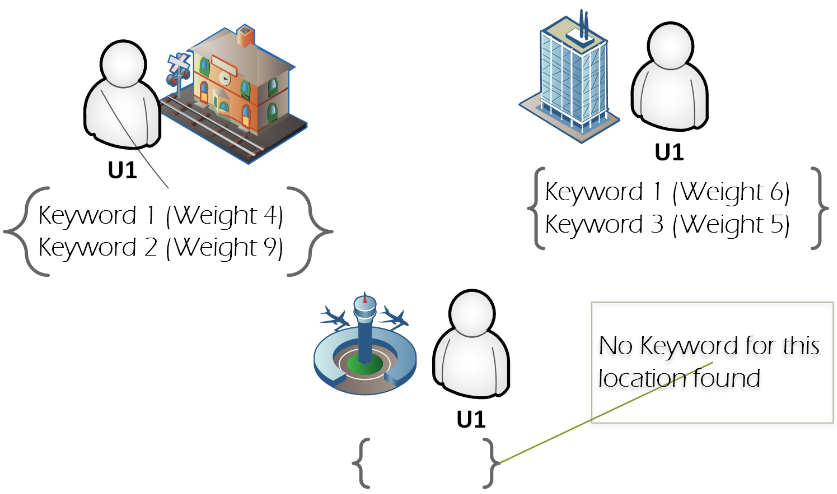


Figure 4.1: Geo-Social Profile for Three Locations

The algorithm uses a forest structure to manage multiple location-based profiles. The user may also prefer to combine all existing profiles to create a single Global Profile. Figure 4.1 shows an example of the proposed location-based profile for three Locations. *Location* 1 has *Keyword* 1 with *weight* 4 and *Keyword* 2 with *weight* 9. The weight of keyword represents the number of times a keyword appears in search and browsing history. Similarly, *Location* 2 has *Keyword* 1 with *weight* 6 and *Keyword* 3 with *weight* 5. *Location* 3 has not been initialized yet, however, a global profile that is a summation of *Location* 1 and *Location* 2 is available at *Location* 3.

Algorithm 4.1 Creating Location Based Profile

$Node_A$ *is a node in* ASN

$F_A \leftarrow Forest\ of\ Node_A$

$L_A \leftarrow GPS\ Locations\ of\ Node_A$

$L_A(t) \leftarrow Location\ of\ Node\ A\ at\ time\ t$

$List_A(L_A(t)) \leftarrow Hierarchical\ data\ structure\ of\ keywords\ for\ location\ L_A(t)$

$K_i \leftarrow a\ Keyword\ that\ represents\ interest\ of\ a\ user$

$I(K_i) \leftarrow Interest\ Level\ of\ Keyword\ K_i$

Require: $Node_A\ receives\ K_i\ from\ user\ action\ at\ L_A(t)$

if $L_A(t) \in F_A$ **then**

  $updateInterests(List_A(L_A(t)),\ K_i)$

else

  $Create\ an\ empty\ list\ List_A(L_A(t))\ and\ add\ to\ F_A$

  $updateInterests(List_A(L_A(t)),\ K_i)$

end if

Function updateInterests($List_A,\ K_i$)

if $K_i \in List_A$ **then**

  $I(K_i) = I(K_i) + 1$

else

  Insert K_i into $List_A$

  $I(K_i) = 0$

end if

EndFunction

Algorithm 4.1 demonstrates a method used to update profile based on location-based keywords. The algorithm can be explained with an example scenario for $Node_A$ in the ASN. The other notations to explain the algorithm are F_A as forest of $Node_A$ that is similar to Figure 4.1; L_A as array of visited GPS locations for $Node_A$ e.g. $\{$Location 1, Location 2$\}$; $L_A(t)$ as the current location of a node at time t; $List_A(L_A(t))$ as the root node for a list of keywords that interests for location $L_A(t)$ at time t; K_i represents a keyword that is interest of a user; and $I(K_i)$ shows the weight or level of the keyword K_i that is the number of times the keyword K_i appears in the browsing history.

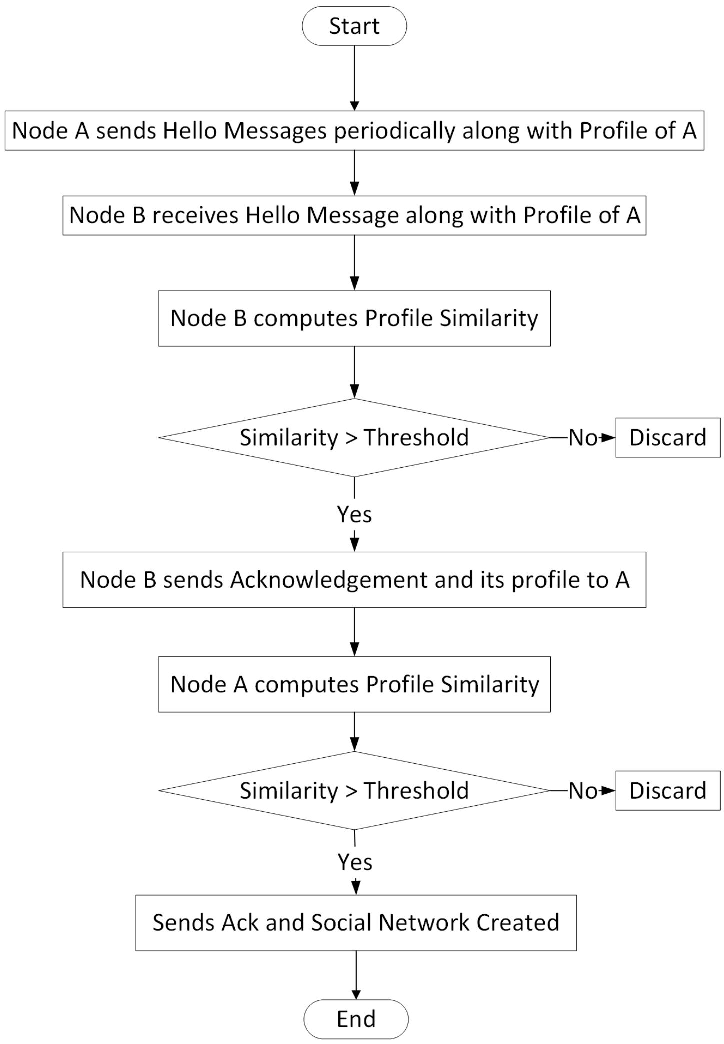


Figure 4.2: Method of Creating ASN

The process of creating or updating location-based profile is triggered in the background at the appropriate user action at a location L_A. The algorithm first checks F_A to verify if the L_A is already existing in the list, if it exists meaning the location is the previous location visited by the user, the keyword is added if it is a new keyword else interest level of the keyword is increased. In case, the location is a new location that is not existing in the F_A, the location is added to F_A, and the keyword is also added to the location with interests level initialized to one. In this way, the algorithm helps to build and maintains location-based user profile.

Once the profile has been created and set as per current location and user's preferences, the process of Profile Matching starts as explained in Figure 4.2. The process of creating ASN includes sending $Hello$ messages that include a profile of a user. The profile is a list of keywords and the interest level (number of times a keyword appears in the search and browsing history). A node computes profile similarity when it receives the profile of another user. The profile similarity can be computed by determining cosine similarity or other similarity algorithms. The profile is discarded if the computed similarity value is less than a predetermined threshold value of the user. However, if the similarity is greater than the threshold, an acknowledgment is sent. A node that receives the acknowledgment also computes profile similarity and create ASN if the similarity value is greater than its threshold.

Algorithm 4.2 Creating ASN

```
  Trigger: Users Join Ad-hoc Network
Require: Users' exchange profile and compute profile similarity
  if Similarity ≥ Threshold then
    Suggestion to create ASN
  else
    Profile Discarded
  end if
```

Algorithm 4.2 explains the process demonstrated in the Figure 4.2 of creating ASN when the computed profile similarity is higher than the threshold. The Profile Similarity is computed using Cosine Similarity (CS) that is defined as:

$$Cosine\ Similarity(\mathrm{CS}) = \frac{\sum_{i=1}^{n}(A_i * B_i)}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{2}$$

where Profile(A) = $\{A_1, A_2, ..., A_n\}$ and Profile(B) = $\{B_1, B_2, ..., B_n\}$

4.2.2 Example

The process of creating a location-based ASN can be understood by the following example.

Suppose Profile of a user A at a location L_1 is represented by $P_{L_1}(A)$ and Interests Level as $W_{L_1}(A)$. Interest Level represents a frequency of keywords appearing in user's prior actions and browsing history. Consider two users A and B that have similar keywords $\{sports$, $enternatinment\}$ in their browsing history at Location L_1 and $\{politics$, $national\}$ at Location L_2. Thus, their profile may vary w.r.t. Interest Level, i.e., the frequency of these keywords as appeared in the browsing history and prior user actions.

At Location L_1: $P_{L_1}(A) = P_{L_1}(B) = \{sports, enternatinment\}$
$W_{L_1}(A) = \{2,\ 4\}$
$W_{L_1}(B) = \{24,\ 22\}$

At Location L_2: $P_{L2}(A) = P_{L2}(B) = \{politics, national\}$
$W_{L2}(A) = \{13,\ 16\}$
$W_{L2}(B) = \{12,\ 12\}$

Global Profile: $P(A) = P(B) = \{sports, enternatinment, politics, national\}$
$W(A) = \{2,\ 4,\ 13,\ 16\}$
$W(B) = \{24,\ 22,\ 12,\ 12\}$

Now, cosine similarity for Global Profile, Location L_1, and Location

L_2 is computed as

$$
\begin{aligned}
CS_{Global}(A,B) &= \frac{2.24+4.22+13.12+16.12}{\sqrt{2^2+4^2+13^2+16^2}\sqrt{24^2+22^2+12^2+12^2}} \\
&= 0.62 \\
CS_{L2}(A,B) &= \frac{13.12+16.12}{\sqrt{13^2+16^2}\sqrt{12^2+12^2}} \\
&= 0.99 \\
CS_{L1}(A,B) &= \frac{2.24+4.22}{\sqrt{2^2+4^2}\sqrt{24^2+22^2}} \\
&= 0.93
\end{aligned}
$$

The above example considers a scenario when two users have similar interests at two different locations while different weights at different locations. Cosine similarity based on the global profile is 0.62 and based on local profile at location $L1$ is 0.93 and at $L2$ is 0.99 as shown in the example. Therefore, cosine similarity does not perform well based on global profiles and a social network may not be recommended by a recommender system due to the low value of profile similarity. Moreover, in this case, the profile similarity based on Global Profile is less than users' preferred similarity that is 0.75 as determined by Gambhir et al. [47].

4.2.3 Implementation

The algorithm was implemented using Python Programming Language. A simple approach to simulate Geo-Social Profile Matching Algorithm consists of using Google Trends [49] that reports keywords from historical search volume. The keywords reported by Google Trends represent interests of a large number of users in different geographical locations. The following Table 4.1 includes top ten keywords in four countries India, USA, UK, and Australia.

Table 4.1: List of Top 10 keywords for different regions

India	Blockchain, Infosys, Finacle, India, Bitcoin, EdgeVerve, Bitcoin, Cryptocurrency, OPEC, Petroleum
USA	OPEC, Petroleum, Juniper Networks, New York Stock Exchange, T. Boone Pickens, Texas, TE Connectivity Ltd., New York Stock Exchange, Amazon.com, NASDAQ
UK	China, Purchasing Managers Index, NASDAQ, Dow Jones Industrial Average, Uber, William Haskell Alsup, Waymo, McAfee, Anthony Levandowski, Ransomware
Australia	Oroton, Australia, Administration, Australian Securities Exchange, Victoria, Weather, Tabcorp Holdings, Tatts Group Limited, Crown Resorts, Air New Zealand

The four countries were considered as different geographical regions. The simulation was implemented considering some users ranging from $10-100$. The threshold for profile similarity was also considered in the range of $0.4-0.9$.

Thus, a user can have five profile at a time meaning four local profiles respectively for different locations and a global profile that is the union of all local profiles. The interest level of each keyword was assigned randomly that varies from 0 to 20. Thus, interest level assigned to keyword *Blockchain* for a $User$ 1 in India as $five$ indicates that *Blockchain* appears in Search and Browsing History as $five$ times. The research simulated local profiles of all users in different geographical locations and then created a Global Profile by taking a union of all local profiles.

Thus, the dataset includes $four$ local profiles and *one* global profile for some users ranging from $10-100$. The profile similarity was computed using cosine similarity for all users in the respective location using local profile of the respective location and then using a global profile. In case of 100 users, the total comparison for *each* location were 9900 since each profile of each user is compared with all other users except self. Local Match is counted if the computed profile similarity is greater than the threshold. Once some matched users were computed for different locations, the average Local Matched Ratio was computed to compare with Global Matched Ratio.

Global Matched Ratio is defined as:

$$GlobalRatio = \frac{gMatched}{(nUser * nUser) - nUser} \tag{3}$$

gMatched are cases where cosine similarity $\geq$ threshold

Local Matched Ratio is defined as:

$$LocalRatio = \frac{lMatched}{((nUser * nUser) - nUser) * nLocations} \tag{4}$$

lMatched are cases where cosine similarity $\geq$ threshold for all locations

The experiment didn't compute profile similarity of a user at one location with other users at the different location(s) since it is not a practical assumption. All users were compared with each other within their respective cluster using local profile and finally based on global profile ignoring cluster.

4.2.4 Results and Analysis

This section discusses the performance of profile matching based on respective Local Profiles and Global Profiles for different values of threshold ranging from 0.4 to 0.9. The simulation results indicate that performance of local profile is better when the required threshold is high.

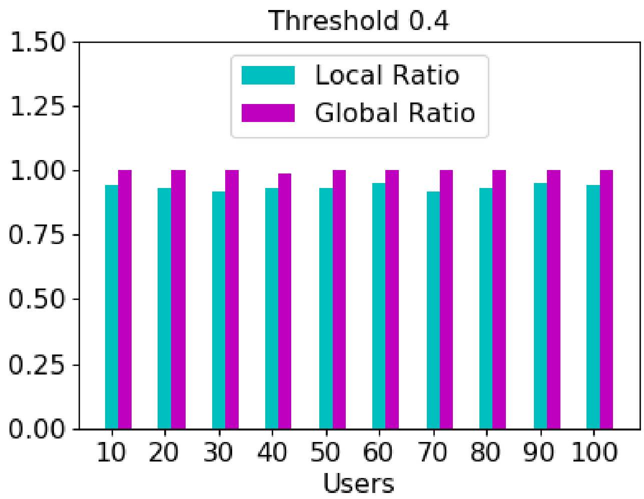


Figure 4.3: Geo-Social Profile Matching Algorithm with threshold value 0.4

Figure 4.3 shows comparison results of Profile Matched Ratio using local profile (Local Ratio) and global profile (Global Ratio) for threshold value of 0.4. The results show slightly better performance of global ratio.

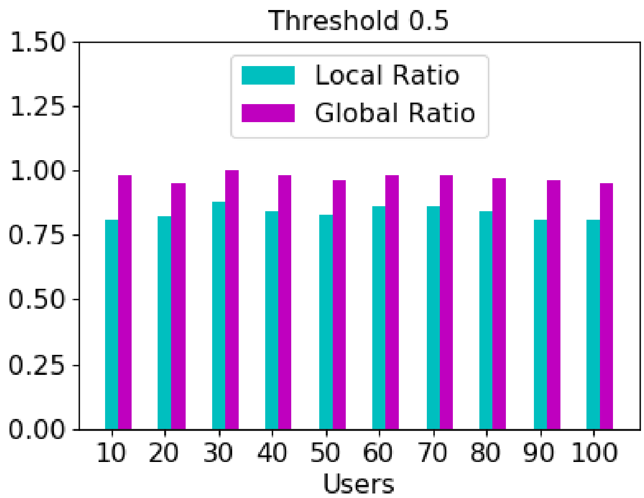


Figure 4.4: Geo-Social Profile Matching Algorithm with threshold value 0.5

Figure 4.4 shows comparison results of Local Ratio and Global Ratio for threshold value of 0.5. The results show slightly better performance of Global Ratio than Local Ratio .

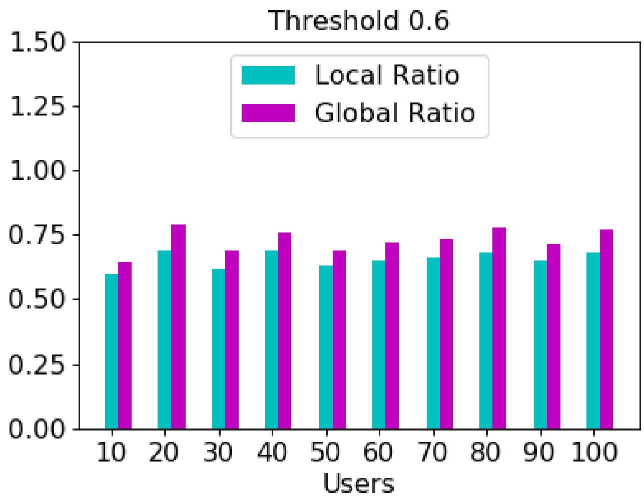


Figure 4.5: Geo-Social Profile Matching Algorithm with threshold value 0.6

Figure 4.5 shows comparison results of Local Ratio and Global Ratio for threshold value of 0.6. The results show better performance of Global Ratio than Local Ratio .

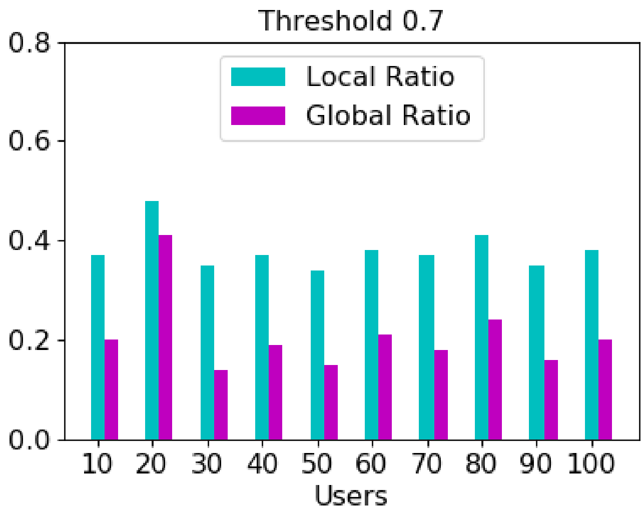


Figure 4.6: Geo-Social Profile Matching Algorithm with threshold value 0.7

Figure 4.6 shows comparison results of Local Ratio and Global Ratio for threshold value of 0.7. The results show Local Ratio started giving better performance than Global Ratio.

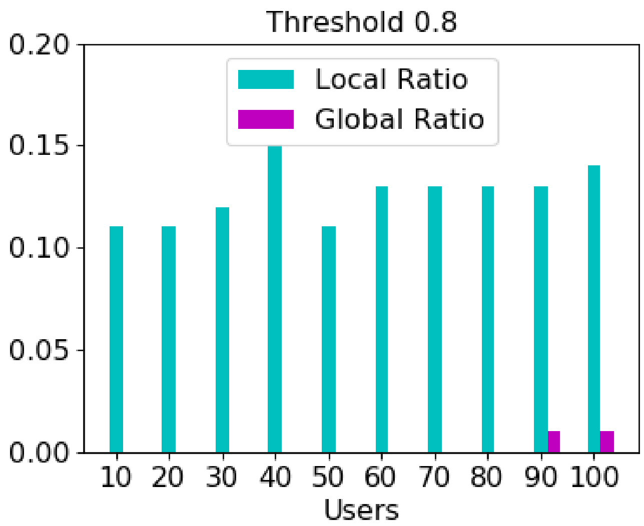


Figure 4.7: Geo-Social Profile Matching Algorithm with threshold value 0.8

Figure 4.7 shows comparison results of Local Ratio and Global Ratio for threshold value of 0.8. The results show Local Ratio provides better performance than Global Ratio.

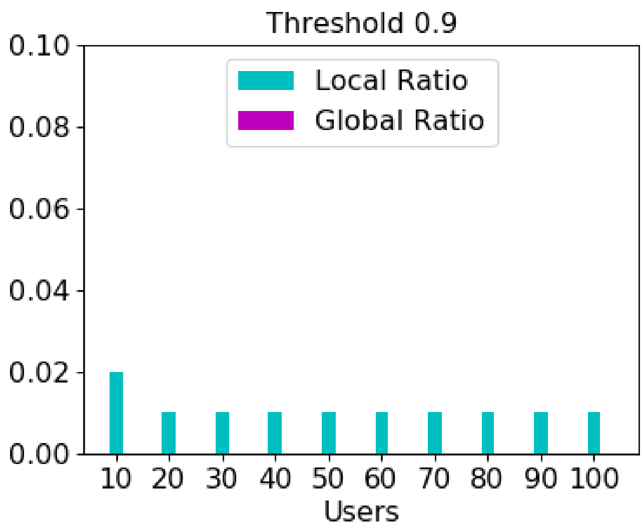


Figure 4.8: Geo-Social Profile Matching Algorithm with threshold value 0.9

Figure 4.8 shows comparison results of Local Ratio and Global Ratio for threshold value of 0.9. The results show Local Ratio provides better performance than Global Ratio. In fact, the value of global ratio is null.

Thus, results show better performance of local profile similarity for the higher value of the threshold. The performance of the location-based matching algorithm is advantageous not only for better social networking but also to target advertisements to highly relevant users.

4.3 Semantic Geo-Social Profile Matching Algorithm

The Profile Matching algorithm explained earlier discusses location-based profile matching algorithm. The location-based profile matching can be further improved by integrating semantic similarity [51].

Algorithm 4.3 Creating Location Based Semantic Profile

$Node_A$ *is a node in* ASN

$F_A \leftarrow Forest\ of\ Node_A$

$L_A \leftarrow GPS\ Locations\ of\ Node_A$

$L_A(t) \leftarrow Location\ of\ Node\ A\ at\ time\ t$

$List_A(L_A(t)) \leftarrow Hierarchical\ data\ structure\ of\ keywords\ for\ location\ L_A(t)$

$K_i \leftarrow a\ Keyword\ that\ represents\ interest\ of\ a\ user$

$I(K_i) \leftarrow Interest\ Level\ of\ Keyword\ K_i$

Require: $Node_A\ receives\ K_i\ from\ user\ action\ at\ L_A(t)$

if $L_A(t) \in F_A$ **then**

  $updateSemanticInterests(List_A(L_A(t)),\ K_i)$

else

  $Create\ an\ empty\ list\ List_A(L_A(t))\ and\ add\ to\ F_A$

  $updateSemanticInterests(List_A(L_A(t)),\ K_i)$

end if

Function updateSemanticInterests($List_A,\ K_i$)

if $K_i \in (semantic)List_A$ **then**

  $I(K_i) = I(K_i) + 1$

else

  Insert K_i into $List_A$

  $I(K_i) = 0$

end if

EndFunction

Semantic matching is a matching operation that identifies if two keywords of the profiles correspond to one another, e.g., vehicle and car are semantic equivalents to automobile since these are synonyms in the English Language. The semantic similarity can be calculated with WordNet that is a lexical database to find if keywords are synonymous. Algorithm 4.3 is a modification of algorithm 4.1 concerning module that updates interest level when processing keywords from user interests. The module *updateSemanticInterests* of the Algorithm 4.3 increases the interest level of existing keyword if the new keyword is semantically equivalent.

4.3.1 Example

The process of creating location-based semantic ASN can be explained with the following example. In a location-based ASN, there can be two situations, first situation (*Situation 1*) in which there is no similar keyword intra-location but similar available in inter-location. The second situation (*Situation 2*) in which related keywords have been used in intra-location also. The study of these scenarios will describe advantageous of semantic profile matching.

This research computed three types of similarities namely Global Similarity, Similarity at Location X, and Similarity at Location Y. Global Similarity means that user profile is not separate for different locations and thus is not a forest instead it is a tree that is the summation or union of all local profiles. Interest levels are added in case keywords are semantically similar at the different locations. Similarity at Location X and Similarity at Location Y considers location-based profile for their respective locations to compute semantic similarity.

Situation 1

In the first situation, assume that both users *A* and *B* searched or browsed URLs comprising keywords (football, badminton, soccer and tennis) at two geo-locations $Location_X$ and $Location_Y$ at time $t1$ and $t2$ respectively. Both users searched for {football, badminton} at $Location_X$ at time t_1 and {soccer, tennis} at $Location_Y$ at time t_2. The number of times the keywords {football, badminton} appear are {2, 4} for $User_A$ and {24, 22} for $User_B$ at $Location_X$. Similarly, the number of times the keywords {soccer, tennis} appear are {13, 16} for $User_A$ and {12, 12} for $User_B$ at location Y.

Thus, profiles at $Location_X$:

$User_A$ = {{football, badminton}; {2, 4}}
$User_B$ = {{football, badminton}; {24, 22}} and

$$
\begin{aligned}
Cosine\ Similarity(User_A, User_B) &= \frac{User_A \cdot User_B}{\|User_A\|_2 \cdot \|W_B\|_2} \\
&= \frac{(2,4)\cdot(24,22)}{\|(2,4)\|_2 \cdot \|(24,22)\|_2} \\
&= \frac{(2\cdot 24)+(4\cdot 22)}{\sqrt{2^2+4^2}\cdot\sqrt{24^2+22^2}} \\
&= \frac{48+88}{4.47\cdot 32.56} = \frac{136}{145.54} \\
&= 0.93
\end{aligned}
$$

Similarly, profiles at $Location_Y$:

$User_A$ = {{soccer, tennis}; {13, 16}}
$User_B$ = {{soccer, tennis}; {12, 12}} and

$$
\begin{aligned}
Cosine\ Similarity(User_A, User_B) &= \frac{User_A \cdot User_B}{\|User_A\|_2 \cdot \|W_B\|_2} \\
&= \frac{(13,\ 16)\cdot(12,12)}{\|(13,\ 16)\|_2 \cdot \|(12,12)\|_2} \\
&= 0.99
\end{aligned}
$$

Since the Global profile is a union of all local profiles, global profiles will be

$$\begin{aligned}
User_A &= \{\{football,\ badminton\};\{2,\ 4\}\} \cup \{\{soccer,\ tennis\};\{13,\ 16\}\}\\
&= \{\{football,\ badminton, soccer,\ tennis\};\{2,\ 4,\ 13,\ 16\}\}\\
User_B &= \{\{football,\ badminton\};\{24,\ 22\}\} \cup \{\{soccer,\ tennis\};\{12,\ 12\}\}\\
&= \{\{football,\ badminton, soccer,\ tennis\};\{24,\ 22,\ 12,\ 12\}\}
\end{aligned}$$

The Global semantic profile will be different from the above Global profile since two keywords, namely, $football$ and $soccer$, can be considered semantically similar. Thus, Global Semantic Profile will be:

$$\begin{aligned}
User_A &= \{\{football,\ badminton\};\{2,\ 4\}\} \cup \{\{soccer,\ tennis\};\{13,\ 16\}\}\\
&= \{\{football,\ badminton,\ tennis\};\{15,\ 4,\ 16\}\}\\
User_B &= \{\{football,\ badminton\};\{24,\ 22\}\} \cup \{\{soccer,\ tennis\};\{12,\ 12\}\}\\
&= \{\{football,\ badminton,\ tennis\};\{36,\ 22,\ 12\}\}
\end{aligned}$$

As a result, Semantic Cosine Similarity for Global Profile may be different than Cosine Similarity for Global Profile, which is computed as below:

For Global Profile:

$$\begin{aligned}
Cosine\ Similarity &= \frac{User_A \cdot User_B}{\|User_A\|_2 \cdot \|W_B\|_2}\\
&= \frac{(2,4,13,16)\cdot(24,22,12,12)}{\|(2,4,13,16)\|_2 \cdot \|(24,22,12,12)\|_2}\\
&= \frac{(2\cdot 24)+(4\cdot 22)+(13\cdot 12)+(16\cdot 12)}{\sqrt{2^2+4^2+13^2+16^2}\cdot\sqrt{24^2+22^2+12^2+12^2}}\\
&= \frac{48+88+156+192}{21.1\cdot 36.72} = \frac{484}{774.79}\\
&= 0.62
\end{aligned}$$

$$
\begin{aligned}
Semantic\ Cosine\ Similarity &= \frac{User_A \cdot User_B}{\|User_A\|_2 \cdot \|W_B\|_2} \\
&= \frac{(15,4,16)\cdot(36,22,12)}{\|(15,4,16)\|_2 \cdot \|(36,22,12)\|_2} \\
&= 0.84
\end{aligned}
$$

Thus, the example in situation 1 indicates improvement in cosine similarity when the profile is semantically matched as shown in the Table 4.2.

Table 4.2: Inter-Location Semantic Cosine Similarity

Profile	User	Interests $Location_X$		Interests $Location_Y$		Similarity Score	Semantic Similarity Score
		Football	Badminton	Soccer	Tennis		
$Location_X$	A	2	4			0.93	0.93
	B	24	22				
$Location_Y$	A			13	16	0.99	0.99
	B			12	12		
Global	A	2	4	13	16	0.62	
	B	24	22	12	12		
Semantic Global	A	15	4	16			0.84
	B	36	22	12			

Situation 2

In the second situation, the keywords are changed so that similar semantic keywords are used intra-location. Thus, local profiles for two users will be:

Thus, profiles at $Location_X$ will be:

$User_A = \{\{\text{football, tennis}\}; \{2,\ 16\}\}$

$User_B = \{\{\text{soccer, tennis}\}; \{12,\ 12\}\}$

For cosine similarity, when a user receives a profile of another user that has different keywords, it calculates the union of keywords and places null values for the words that don't exist. Thus, cosine similarity will use following profiles for computation:

$User_A = \{\{$football, tennis, soccer$\}; \{2,\ 16,\ 0\}\}$
$User_B = \{\{$football, tennis, soccer$\}; \{0,\ 12,\ 12\}\}$.

Now, cosine similarity and semantic cosine similarity is computed as follows:

$$
\begin{aligned}
Cosine\ Similarity &= \frac{User_A \cdot User_B}{\|User_A\|_2 \cdot \|W_B\|_2} \\
&= \frac{(2,16,0)\cdot(0,12,12)}{\|(2,16,0)\|_2 \cdot \|(0,12,12)\|_2} \\
&= \frac{(2\cdot 0)+(16\cdot 12)+(0\cdot 12)}{\sqrt{2^2+16^2+0^2}\cdot\sqrt{0^2+12^2+12^2}} \\
&= \frac{192}{16.12\cdot 16.97} = \frac{192}{273.56} \\
&= 0.70
\end{aligned}
$$

However, in case of Semantic Cosine Similarity, since the keywords football and soccer are considered semantically similar, the profile will be:

$User_A = \{\{$football/soccer, tennis$\}; \{2,\ 16\}\}$
$User_B = \{\{$soccer/football, tennis$\}; \{12,\ 12\}\}$

$$
\begin{aligned}
Semantic\ Cosine\ Similarity &= \frac{User_A \cdot User_B}{\|User_A\|_2 \cdot \|W_B\|_2} \\
&= \frac{(2,16)\cdot(12,12)}{\|(2,16)\|_2 \cdot \|(12,12)\|_2} \\
&= \frac{(2\cdot 12)+(16\cdot 12)}{\sqrt{2^2+16^2}\cdot\sqrt{12^2+12^2}} \\
&= \frac{24+192}{16.12\cdot 16.97} = \frac{216}{273.56} \\
&= 0.79
\end{aligned}
$$

Thus, example in situation 2 also indicates improvement in cosine similarity when profile are matched semantically as shown in Table 4.3.

Table 4.3: Intra-Location Semantic Cosine Similarity

Profile	User	Interests $Location_X$			Similarity Score	Semantic Similarity Score
		Football	Tennis	Soccer		
$Location_X$	A	2	16		0.70	
	B		12	12		
Semantic $Location_X$	A	2	16			0.79
	B	12	12			

4.3.2 Implementation

This research is simulated using a network of 20 users with the weights or interest level for the keywords generated randomly for two locations. Once the local profiles are available, global profile and global semantic profiles were computed. Cosine similarity and Semantic cosine similarity were calculated using global profile and global semantic profile respectively. Thus, using above example, global profile matrix comprises the global profile of 20 users in a matrix of size $(20,4)$, where 20 represents the rows or users and 4 represents the columns or interest level for different keywords {football, badminton, soccer, tennis}. Similarly, the semantic profile matrix comprises a global semantic profile of 20 users in a matrix of size $(20, 3)$, where 20 represents the rows or users and 3 represents the columns or interest level for different keywords {football/soccer, badminton, tennis}. The semantic profile is created by adding interest level corresponding to $football$ and *soccer* since these words are considered semantically similar.

The matrices, global profile matrix (20x4) and semantic profile matrix (20x3), were used to compute cosine similarity (20x20) and semantic cosine similarity (20x20) respectively. The similarity matrices comprise similarity of each user with all other users, and thus the size is 20x20. The cosine similarity and semantic cosine semantic have been computed using *sklearn* library of *Python*. Based on cosine similarity matrix (20x20) and semantic cosine similarity matrix (20x20), this research computed friend matrix (20x20) and semantic friend matrix (20x20) for different similarity thresholds e.g. 0.6, 0.7, 0.75, 0.80. The friend matrix is essentially a binary matrix that shows a friend with a value 1 if the similarity is higher than the threshold. The below graphs (polynomial fit of three degrees) shows the number of friends for different users for various thresholds.

Figures 4.9 - 4.12 shows analysis of friends identified based on similarity threshold values of 0.6, 0.7, 0.75, and 0.8 for cosine similarity and semantic cosine similarity respectively. The graph shows better performance of semantic cosine similarity than cosine similarity.

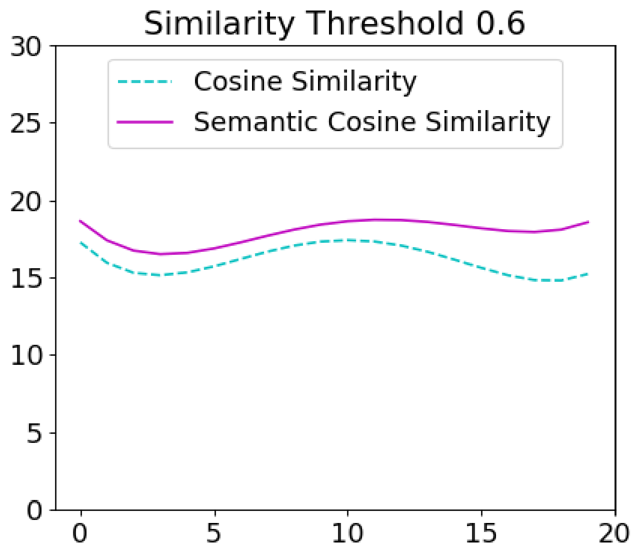


Figure 4.9: Friendship data of 20 users based on threshold value 0.6

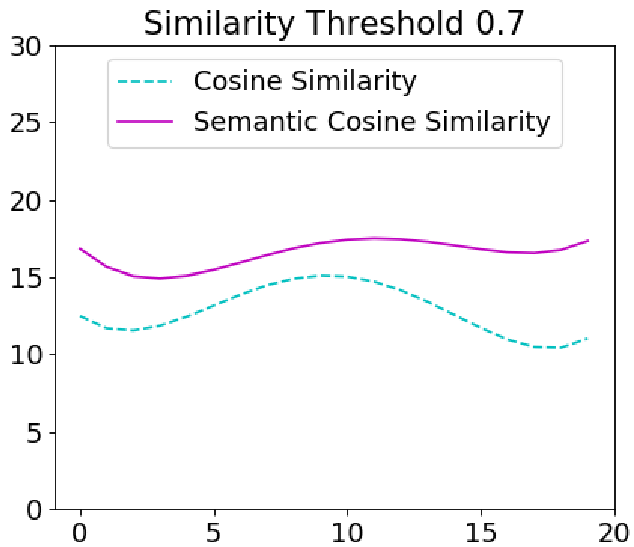


Figure 4.10: Friendship data of 20 users based on threshold value 0.7

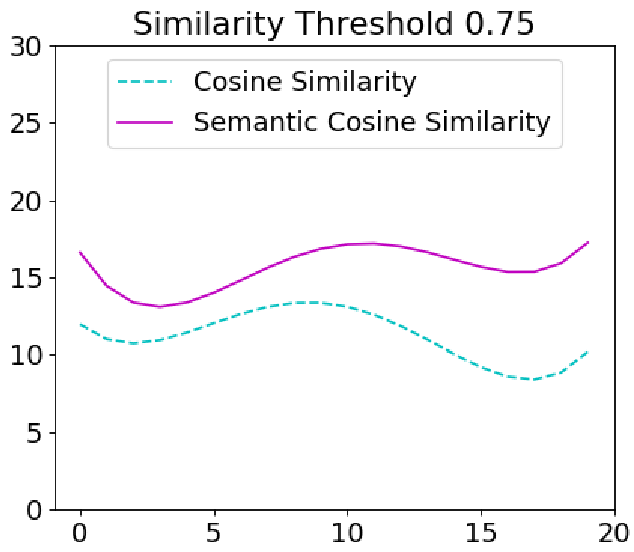


Figure 4.11: Friendship data of 20 users based on threshold value 0.75

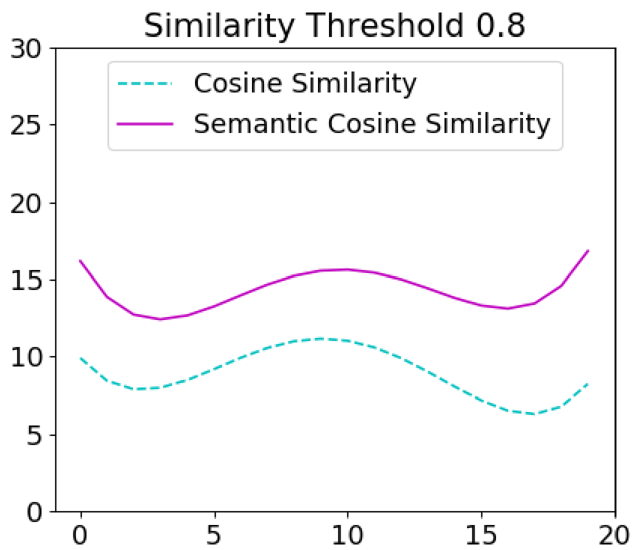


Figure 4.12: Friendship data of 20 users based on threshold value 0.8

4.4 Piecewise Maximal Similarity (PMS)

One of the most common operations in social networking is a similarity assessment. The similarity is computed based on a user profile that comprises different attributes. Centralized access to attributes is not feasible due to lack of infrastructure in ASN. Thus, the profile similarity is largely based on static or dynamic keywords or interests available on a mobile device.

A profile of a user can be defined as:

Definition 1 ***(User Profile)*** *Profile of a user is based on a vector space model defined by set of keywords* $P = \{k_1,\ k_2,\ k_3,...,k_n\}$ *where* k_i *represents users interests in terms of different keywords and Interest Level* $W = \{w_1,\ w_2,\ w_3,...,w_n\}$ *where* w_i *represents weight of keyword* k_i*. The weight of a keyword is the number of times that keyword appears in prior user action or search and browsing history. The keywords can be provided by a user or can be dynamically extracted from the user behavior.*

This research analyzed profile similarity computed by $L1 - Norm$ and $L2 - Norm$ and proposed a new profile similarity metric, Piecewise Maximal Similarity (PMS), for ASN [2].

$L1 - Norm$ is defined as:

Definition 2 ***(L1-Norm)*** *L1-Norm or L1-Similarity [27] of two users* U_1 *and* U_2 *with profiles* P_1*,* P_2 *and Interests Levels as* W_1 *and* W_2 *is:*

$$\begin{aligned} L1(P_A, P_B) &= \frac{W_A \cdot W_B}{||W_A||_1 \cdot ||W_B||_1} \\ &= \frac{\Sigma(wA_i \cdot wB_i)}{\Sigma|wA_i| \cdot \Sigma|wB_i|} \end{aligned} \tag{5}$$

$L2 - Norm$ is defined as:

Definition 3 *($L2 - Norm$ **or cosine similarity**) $L2 - Norm$ or cosine similarity [27] measures cosine of the angle between two vectors. Since, the cosine of* 0° *is* 1 *and less than* 1 *for other angles, thus, the value of cosine similarity is always between* 0 *and* 1. *Cosine similarity of two users with profiles P_A and P_B with interest weight vectors W_A and W_B respectively is defined as*

$$\begin{aligned} L2(P_A, P_B) &= \frac{W_A \cdot W_B}{\|W_A\|_2 \cdot \|W_B\|_2} \\ &= \frac{\Sigma(wA_i \cdot wB_i)}{\sqrt{(}\Sigma(wA_i)^2) \cdot \sqrt{(}\Sigma(wB_i)^2)} \end{aligned} \tag{6}$$

4.4.1 Algorithm

This research computed L1-norm and cosine similarity for some random profile vectors and examined the relationship between profile vectors and computed similarity values manually. This was to know the reliability of computed profile similarity for ASN and human interpretation of profile vectors. The below example provides a brief analysis of L1-Norm and cosine similarity:

Example 1 *(Example demonstrating that cosine similarity is better than $L1 - Norm$).*

Table 4.4: Comparison of cosine similarity and $L1 - Norm$

W_A	W_B	L1	cosine similarity
(1, 0, 0)	(1, 0, 0)	1	1
(1, 1, 0)	(1, 1, 0)	0.5	1
(1, 1, 1)	(1, 1, 1)	0.33	1

Table 4.4 provides L1-norm and cosine similarity for selected three sets of profile vectors that are similar. Analysis of L1-Norm indicates that the value of the L1-norm ranges from 0.33 to 1 in cases where profile vector is similar. However, value of cosine similarity provides

the correct measure and performs better than $L1-Norm$.

Table 4.5: Cosine similarity for random weight vectors

Sr No	Profile User A	Profile User B	cosine similarity
1	(1,0,0)	(1,0,0)	1
2	(2,0,0)	(2,0,0)	1
3	(3,0,0)	(3,0,0)	1
4	(1,1,0)	(1,1,0)	1
5	(1,1,1)	(1,1,1)	1
6	(2,2,2)	(2,2,2)	1
7	(3,3,3)	(3,3,3)	1
8	(1,2,1)	(2,1,1)	0.83
9	(1,3,1)	(3,1,1)	0.63
10	(2,3,1)	(1,3,2)	0.92
11	(3,2,2)	(2,2,3)	0.94
12	(3,0,3)	(2,2,2)	0.81

Detailed analysis of cosine similarity for additional profile vectors with random weights as shown in Table 4.5, demonstrates that two sets are inconsistent with profile similarity. For example, if we compare case 9 and case 12 of Table 4.5, the profile vectors [(1,3,1) and (3,1,1)] of case 9 are similar higher than the profile vectors [(3,0,3) and (2,2,2)] of case 12 because of the null value of one attribute in case 12. However, cosine similarity is high for case 12 than case 9. Thus, cosine similarity may provide inconsistent results for many cases like this.

Table 4.6: Cosine similarity for different user profiles

Sr No.	W_A	W_B	cosine similarity
1	(1, 1, 1)	(10, 10, 10)	1
2	(10, 1, 1)	(10, 10, 10)	0.69

Further, Table 4.6 shows another example demonstrating that cosine similarity may not provide better user experience in computing profile similarity since case 2 should have higher value of profile similarity than case 1.

This research proposes a new metric, called Piecewise Maximal Similarity (PMS), to compute profile similarity that measures maximal similarity on each attribute of the profile vector. The maximum similarity on each attribute is the minimum value of the attribute in corresponding vectors as shown in Equation 7 For example, if a user has a weight of 5 and another user has a weight of 10 for the same attribute, it means two users can be assumed to have the maximum similarity of weight 5 for this attribute.

$$\mathrm{PMS}(P_A, P_B) = \sum_{i=1}^{n} min(wA_i, wB_i) \tag{7}$$

The final value is normalized by dividing with average of total weights:

$$\mathrm{PMS}(P_A, P_B) = \frac{\sum_{i=1}^{n} min(wA_i, wB_i)}{(\sum wA_i + \sum wB_i)/2} \tag{8}$$

Definition 4 *(Piecewise Maximal Similarity (PMS)). PMS of two users with profile vectors P_A and P_B with weighted interest vectors W_A and W_B respectively is defined as*

$$PMS(P_A, P_B) = \frac{\Sigma min(wA_i, wB_i)}{(\Sigma wA_i + \Sigma wB_i)/2} \tag{9}$$

PMS is computed on the example of Table 4.5 and results are shown in Table 4.7. PMS performs better in particular for scenarios e.g. cases 9 and 12 and equally for all other cases.

Table 4.7: Cosine similarity and PMS for random weight vectors

Sr.No.	Profile User A	Profile User B	cosine similarity	PMS
1	(1,0,0)	(1,0,0)	1	1
2	(2,0,0)	(2,0,0)	1	1
3	(3,0,0)	(3,0,0)	1	1
4	(1,1,0)	(1,1,0)	1	1
5	(1,1,1)	(1,1,1)	1	1
6	(2,2,2)	(2,2,2)	1	1
7	(3,3,3)	(3,3,3)	1	1
8	(1,2,1)	(2,1,1)	0.83	0.75
9	(1,3,1)	(3,1,1)	0.63	0.6
10	(2,3,1)	(1,3,2)	0.92	0.83
11	(3,2,2)	(2,2,3)	0.94	0.86
12	(3,0,3)	(2,2,2)	0.81	0.67

4.4.2 Implementation

PMS was implemented on dataset [3] using Python. The dataset includes evidence of Bluetooth encounters, Facebook friendships, and interests of a set of users obtained through SocialBlueConn application at the University of Calabria. The dataset contains Bluetooth device vicinity data obtained by an ad-hoc Android application named SocialBlueConn [3]. 15 scholars used the application at University of Calabria campus in Rende, Cosenza (Italy) during working days from Jan 28, 2014 to Feb 5, 2014. The dataset also includes the social profiles (Facebook friends and self-declared interests) of the users.

Each device performed a periodic Bluetooth device discovery after 180 seconds to locate nearby devices. Different android smartphones were used but with Bluetooth range of about 10 meters. Each participant was told to keep the device with himself or herself and powered on from 12:00 AM to 08:00 PM, and to use the application for social networking during the experiment. Each device recorded the results of the periodic device discovery. Participants logged into Facebook using an ad-hoc website accessing Facebook API to provide Facebook friendship data. Participants' interests were acquired at the start of the experiment through an offline questionnaire. The questionnaire comprised a list of questions concerning participants' preferences for following nine macro-categories: (A) mobility, (B) sport, (C) music, (D) cinema, (E) literature, (F) multimedia entertainment, (G) politics, (H) other hobbies and (I) social networks.

The interest of the users was based on the following categories:

Table 4.8: Participant's Interests

Category	Category Name	Details	Number of categories
A	Mobility	A1 = Public transport; A2 = Own vehicle; A3 = Bicycle; A4 = Walking	4
B	Sport	B1 = Athletics; B2 = Soccer; B3 = Swimming; B4 = Volley; B5 = Martial arts; B6 = Tennis; B7 = Basket	7
C	Music	C1 = Rock; C2 = Pop; C3 = Ethnic/popular; C4 = Classic; C5 = Jazz; C6 = Rap	6
D	Cinema	D1 = Adventure; D2 = Comedy; D3 = Action; D4 = Biographic; D5 = Sci-fi; D6 = Horror	6
E	Literature	E1 = Detective stories; E2 = Biographical stories; E3 = Fantasy stories; E4 = Novel stories	4
F	Multimedia Entertainmnet	F1 = Terrestial TV; F2 = satellite TV; F3 = Web TV; F4 = Console/PC	4
G	Politics	G1 = Rigth; G2 = Center; G3 = Left; G4 = M5S party; G5 = Others; G6 = Not responding	6
H	Other hobbies	H1 = Fitness &Wellness; H2 = Cinema; H3 = Technology; H4 = Arts; H5 = Motors; H6 = Theater	6
I	Social Networks	I1 = Facebook; I2 = Twitter; I3 = GooglePlus; I4 = LinkedIn; I5 = Badoo	5

For each category, the participants could choose one or more sub-categories. Thus, there were 48 categories of interests, i.e., each profile vector included 48 attributes. Cosine similarity and PMS were computed for each user with all of the other users. Once cosine similarity and PMS are available, a friend matrix for both cosine similarity and PMS was constructed based on a threshold value. This research used thresholds of 0.50 and 0.75 and computed following four matrices. The rationale for selecting 0.75 was the users' preferences following the finding by Gambhir et al. [47]. A lower threshold was added 0.50 to verify the results for lower threshold value. However, the other similarity thresholds ranging from $0 - 1$ were considered to compare cosine similarity and PMS for the evaluation metrics comprising PC, POD, FAR, and TS.

1. *Cosine Similarity-Friend Matrix*: Friends based on cosine similarity
2. *PMS-Friend Matrix*: Friends based on PMS
3. *BT-Friend Matrix*: Friends based on Bluetooth
4. *FB-Friend Matrix*: Friends based on Facebook

The criteria for marking two users as a friend in the matrices is explained in the Table 4.9.

Table 4.9: Criteria that two users are Friends

Cosine Similarity-Friend Matrix	If the cosine similarity between the two users was equal or higher than the threshold (two cases when the threshold value is 0.50 and 0.75)
PMS-Friend Matrix	If the PMS between the two users was equal or higher than the threshold (two cases when the threshold value is 0.50 and 0.75)
BT-Friend Matrix	If the number of encounters between two users were 150 or higher during experiment days
FB-Friend Matrix	If the two users were friends on the Facebook

4.4.3 Results and Analysis

Friendship data computed using cosine similarity and PMS are being called Predicted friendship in this research. The observed Friendship data includes from Bluetooth and Facebook. The results of cosine similarity and PMS were compared to see which friendship predicted data (whether PMS or cosine similarity) is more consistent with the observed data (Facebook and Bluetooth).

Table 4.10: Contingency Table to compare predicted value (cosine similarity and Piecewise Maximal Similarity (PMS)) with observed value (Facebook and Bluetooth)

	Observed-Yes	Observed-No	Total
Predicted-Yes	a	b	a+b
Predicted-No	c	d	c+d
Total	a+c	b+d	a+b+c+d

For comparison, the Contingency Table as shown in Table 4.10 was used. The table tabulated the number of instances when the prediction was right/wrong when compared with the observed data. The *"a"* table entry represents *"Hits"* i.e. the number of predictions that correspond to observations; entry *"b"* represents *"False Alarms"* i.e. the number of predictions that do not correspond to observations; entry *"c"* represents *"Misses"* i.e. the number of no-predictions that correspond to observations; and entry *"d"* represents *"Correct Rejections"* i.e. the number of no-predictions that correspond to no-observations.

The following comparison were achieved:

(i) Cosine Similarity (predicted) vs Facebook (observed)

(ii) Cosine Similarity (predicted) vs Bluetooth (observed)

(iii) PMS (predicted) vs Facebook (observed)

(iv) PMS (predicted) vs Bluetooth (observed)

Each comparison was measured with the following performance metrics:

1. *Percent Correct (PC):* PC is the percent of predictions that are correct and is computed as

$$\begin{aligned} \text{PC} &= \frac{\text{number of correct predictions}}{\text{total number of predictions}} \\ &= \frac{Hits + Correct\ Rejections}{Hits + False\ Alarms + Misses + Correct\ Rejections} \\ &= \frac{a+d}{a+b+c+d} \end{aligned} \tag{10}$$

2. *Probability of Detection (POD):* POD is the fraction of observed events that is predicted correctly as:

$$\begin{aligned} \text{POD} &= \frac{\text{number of correct friend predictions}}{\text{total number of observed friends}} \\ &= \frac{Hits}{Hits + Misses} \\ &= \frac{a}{a+c} \end{aligned} \tag{11}$$

3. *False Alarm Ratio (FAR):* FAR is the fraction of ”yes” predictions that were wrong, and is:

$$\begin{aligned} \text{FAR} &= \frac{\text{number of incorrect friend predictions}}{\text{total number of predictions}} \\ &= \frac{False\ Alarmss}{Hits + False\ Alarms} \\ &= \frac{b}{a+b} \end{aligned} \tag{12}$$

4. *Threat Score (TS):* TS, also known as Critical Success Index, combines Probability of Detection and False Alarm Ratio into one

score for low-frequency events and is computed as:

$$\begin{aligned}
\text{TS} &= \frac{\text{number of correct friend predictions}}{\text{total friend predictions+total observed but missed}} \\
&= \frac{Hits}{Hits + False\ Alarms + Misses} \\
&= \frac{a}{a+b+c}
\end{aligned} \tag{13}$$

Table 4.11: Comparison of cosine similarity and PMS with Facebook (threshold 0.50)

	Observed-Yes	Observed-No	Total
Predicted-Yes	46	47	93
Predicted-No	50	82	132
Total	96	129	225

(a) Predicted by cosine similarity

	Observed-Yes	Observed-No	Total
Predicted-Yes	46	39	85
Predicted-No	50	90	140
Total	96	129	225

(b) Predicted by PMS

Table 4.11 shows comparison of observed and predicted data where Observed data is Facebook friendship data while Predicted data is from Cosine similarity (Table 4.11a) and PMS (Table 4.11b). The results of Table 4.11 are based on the comparison of 15 users when compared with each other after assuming similarity threshold value 0.50. The data for threshold 0.75 is not shown since it was similar for both cosine similarity and PMS

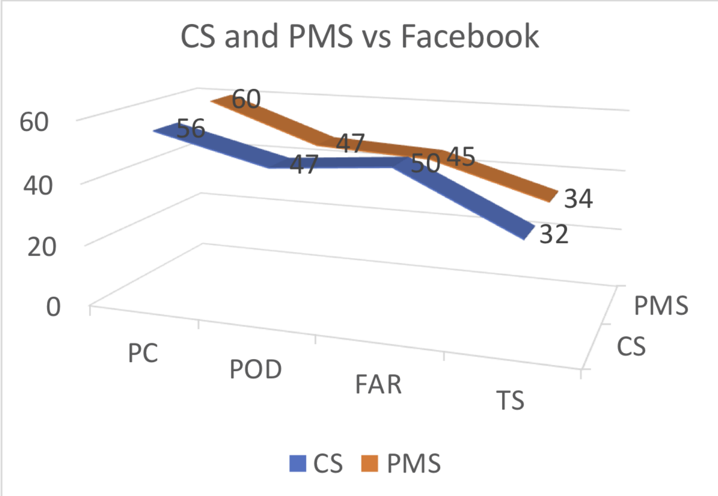


Figure 4.13: PC, POD, TS, and FAR for cosine similarity and PMS vs Facebook

The values of PC, POD, TS, and FAR were computed to analyze as which metric is better to predict friendship and is consistent with Facebook. Figure 4.13 shows that PMS performs better since PC was found 4% higher and FAR 5% lower in comparison of cosine similarity.

Table 4.12: Comparison of cosine similarity and PMS with Bluetooth

(a) Predicted by cosine similarity

	Observed-Yes	Observed-No	Total
Predicted-Yes	16	77	93
Predicted-No	19	113	132
Total	35	190	225

(b) Predicted by PMS

	Observed-Yes	Observed-No	Total
Predicted-Yes	15	70	85
Predicted-No	20	120	140
Total	35	190	225

Table 4.12 shows comparison of observed and predicted data where the Observed data is from Bluetooth encounters while Predicted data is from cosine similarity (Table 4.12a) and PMS (Table 4.12b). The results of Table 4.12 are based on 15 users when compared with each other assuming similarity threshold 0.50 and Bluetooth encounters

threshold value 150. The data for similarity threshold 0.75 is not shown since it was similar for both cosine similarity and PMS

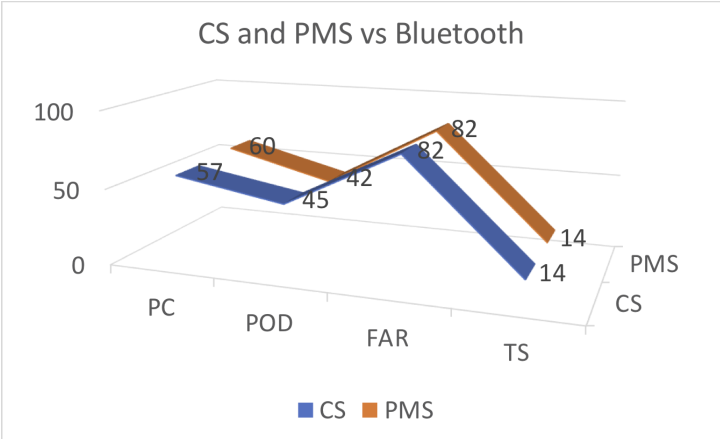


Figure 4.14: PC, POD, TS, and FAR for cosine similarity and PMS vs Bluetooth

The values of PC, POD, TS, and FAR were computed to analyze which metric is the better metric to predict friendship and is consistent with Bluetooth. Figure 4.14 shows that PMS performs reasonably well since PC was found 3% higher and FAR equal in comparison of cosine similarity. However, POD was found lower.

Thus, in both cases when compared with Facebook or Bluetooth, results indicate that PMS works better than cosine similarity in both situations and is more consistent with Facebook or Bluetooth Contacts. The metrics cosine similarity and PMS in comparison to Facebook and Bluetooth were also analyzed for different values of similarity threshold, and results as shown in Figures 4.15 - 4.22.

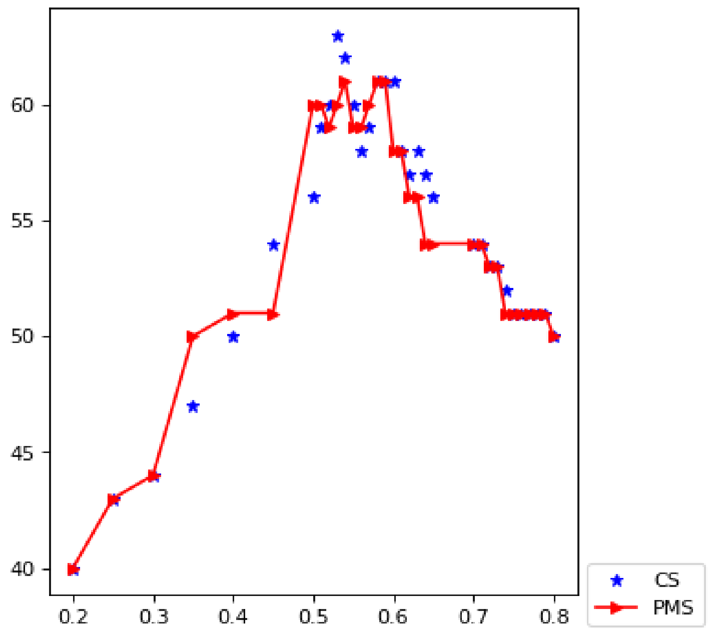


Figure 4.15: PC for Cosine Similarity and PMS vs Facebook for different values of similarity threshold

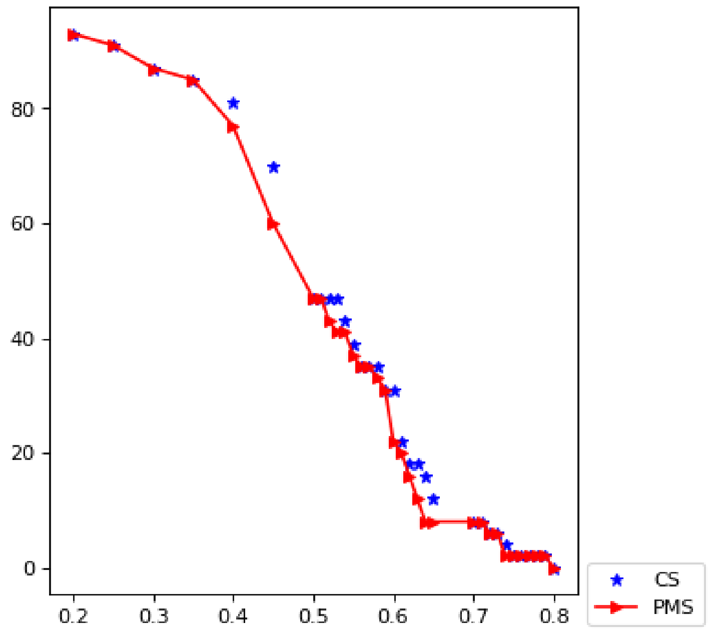


Figure 4.16: POD for Cosine Similarity and PMS vs Facebook for different values of similarity threshold

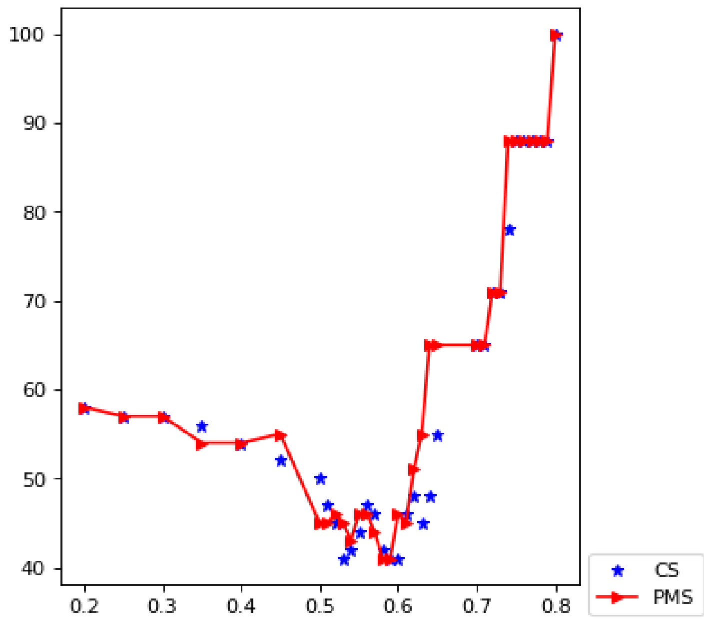


Figure 4.17: FAR for Cosine Similarity and PMS vs Facebook for different values of similarity threshold

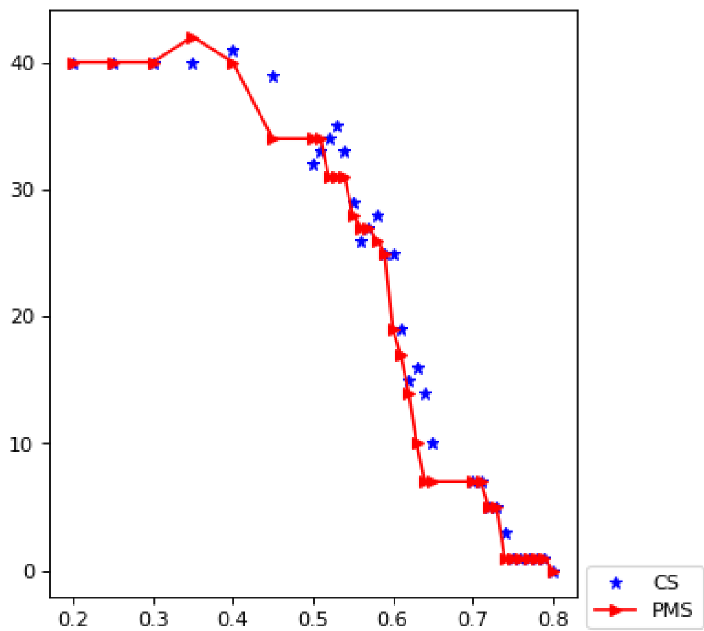


Figure 4.18: TS for Cosine Similarity and PMS vs Facebook for different values of similarity threshold

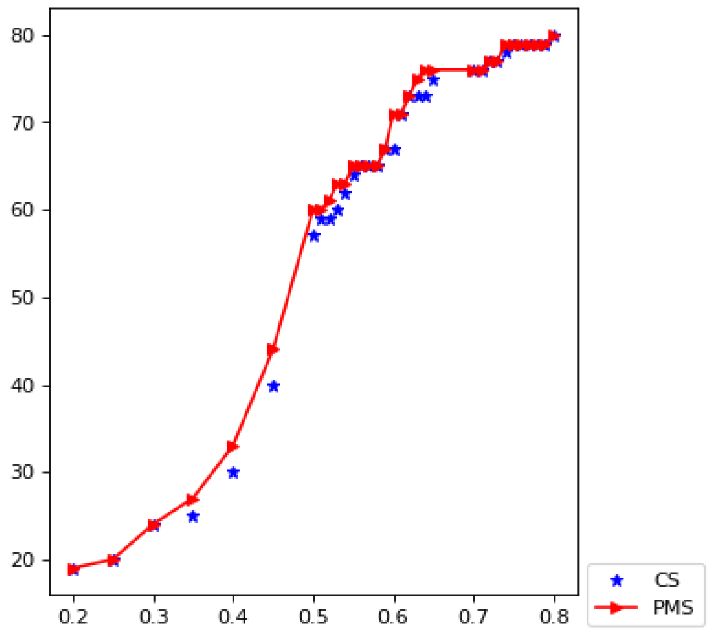


Figure 4.19: PC for cosine similarity and PMS vs Bluetooth for different values of similarity threshold

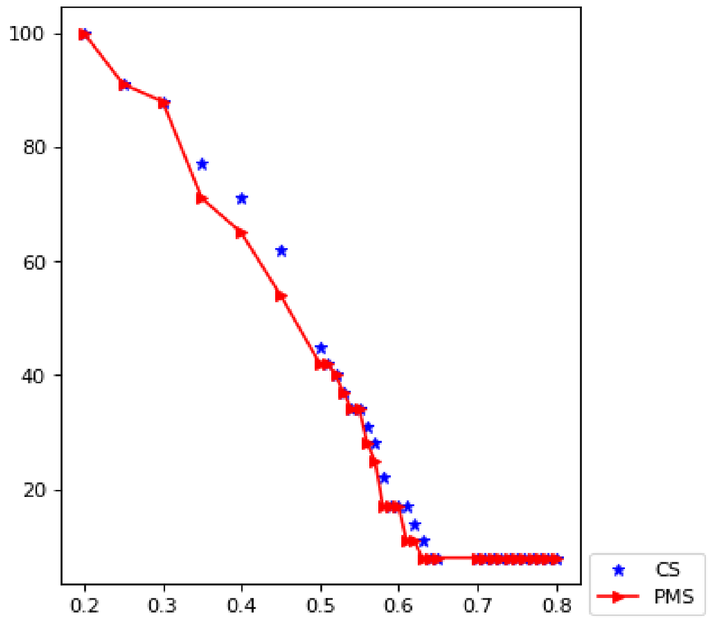


Figure 4.20: POD for cosine similarity and PMS vs Bluetooth for different values of similarity threshold

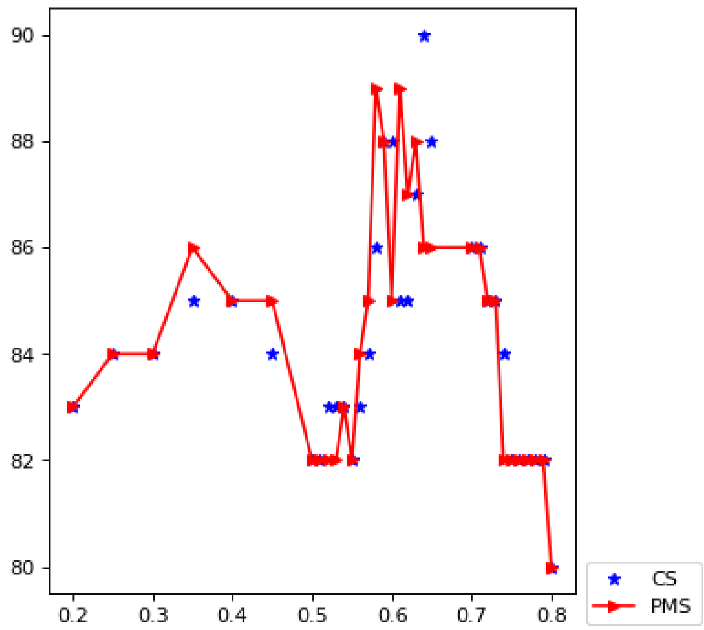


Figure 4.21: FAR for cosine similarity and PMS vs Bluetooth for different values of similarity threshold

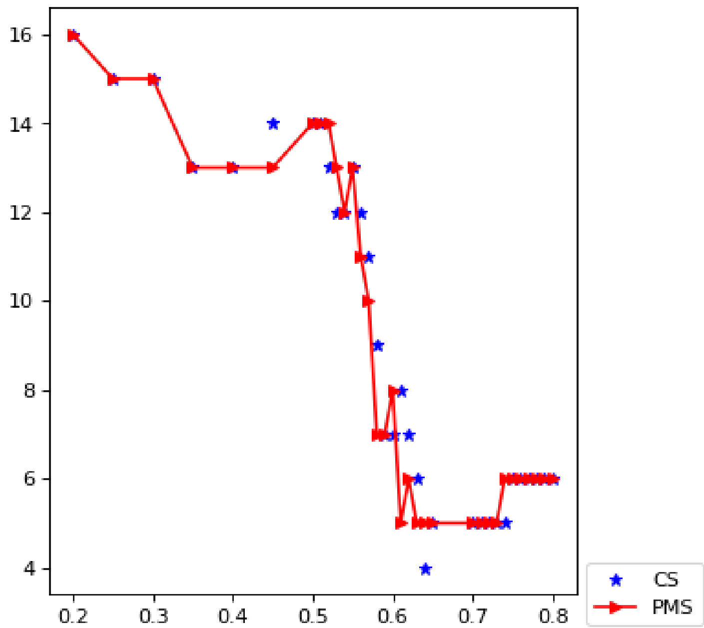


Figure 4.22: TS for cosine similarity and PMS vs Bluetooth for different values of similarity threshold

Thus, results indicate that values of both cosine similarity and PMS are close together indicating almost similar performance for this data set. However, PMS will perform better in general where the cases like profile vectors as mentioned in Table 4.6 are high.

4.5 Summary

This chapter explained geo-social profile matching algorithm that creates multiple location-based profiles by extracting keywords from users' location-based search and browsing history. The algorithm compares the location-based local profile and location-independent global profile by computing profile similarity for various thresholds 0.4 to 0.9. Results indicate improved performance of local profile matching than global profile matching when the similarity threshold is high.

This chapter also described the effect of semantic matching in addition to location-based profile matching. Results indicate that semantic cosine similarity performs better than cosine similarity.

Additionally, the chapter explains new way known as Piecewise Maximal Similarity (PMS) to compute profile similarity metric. PMS computes profile similarity based on the maximal profile similarity on each attribute. PMS was compared with cosine similarity using the real-data set that has interests categories and actual value whether two users are friends as indicated by Facebook or meeting frequency as shown by Bluetooth. Results indicate 4% improvement in PMS accurately predicting friendship when compared with Facebook and 3% improvement when compared with Bluetooth.

Chapter V

PROPOSED WORK: OPTIMIZATION OF ROUTING PROTOCOL AND DEVELOPMENT OF ANDROID BASED MOBILE APPLICATION

Chapter V

PROPOSED WORK: OPTIMIZATION OF ROUTING PROTOCOL AND DEVELOPMENT OF ANDROID BASED MOBILE APPLICATION

The performance of local profile matching is higher than using global profile matching when compared to Facebook and Bluetooth. Since there is a correlation between profile similarity matching and Bluetooth meeting patterns, this fact can be exploited to improve routing protocol to give stable path. A path that has highly similar nodes as intermediate nodes will be a stable path than with random or dissimilar nodes. This chapter presents an improvement in the routing protocol and also introduces an android based mobile app that creates a digital aura and broadcast users interest in a nearby region and computer profile similarity using P2P in the airplane mode.

5.1 Introduction

SPA-AODV routing protocol [4] has been proposed in this research for MANET-based social networking applications by using contextual aware information [4]. The need of contextual aware routing protocol can be explained with the help of Figure 5.1 that shows a network topology comprising of five nodes N_0, N_1, N_2, N_3, N_4.

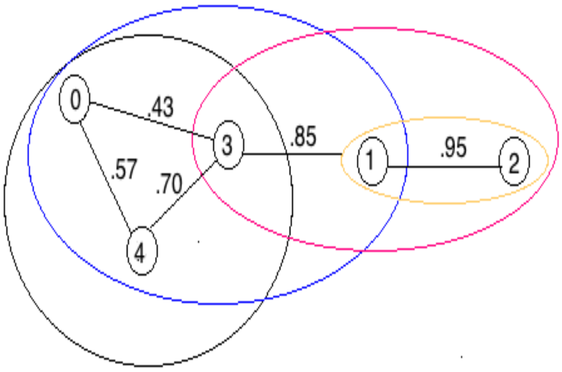


Figure 5.1: ASN of five nodes with edges represented by similarity measures

Nodes N_0, N_3, N_4 are in range of each other and node N_3 is in range of N_1 but not in range of N_2. Edges have been represented by values of similarity metric for profiles of connecting nodes. For demonstration by an example, the profile similarity between different nodes are assumed as below:

$PS(N_0,N_4) = 0.57, \qquad PS(N_0,N_3) = 0.43$

$PS(N_3,N_1) = 0.85, \qquad PS(N_1,N_2) = 0.95$

Neighbors of Node 3 = $Nb(N_3)$ =$\{$ N_0, N_1, and $N_4$$\}$.

Figure 5.1 shows two options for communication between a source node N_0 and a destination node N_2. The first option is that the routing algorithm finds a path that is shortest between N_0 and N_2. The other option is to find a path that may be longer, but all nodes on the path have a similarity greater than required by the source node N_0. The motivation to create such a path is due to the reason that ASN is based on cooperation and nodes having similar interests are motivated to maintain the ad-hoc network. Additionally, there is a pattern among nodes as mobile nodes are carried by human beings who tend to stay together if interests are similar.

The proposed routing protocol covers both situations. The protocol allows setting Profile Similarity value by a source node. If the value of the profile similarity threshold is set by source, the routing algorithm selects a path so that the nodes on the path are similar to the source node. However, if the profile similarity is not set or set to null, the routing algorithm automatically selects the shortest path following the base protocol which in this case is AODV.

5.1.1 Routing Protocol

SPA-AODV has two phases called (i) Pre-processing phase and (ii) Route establishment phase.

5.1.1.1 Phase 1 - Preprocessing phase

The preprocessing phase provides sufficient time for nodes to create search and browsing history and thus profile, as nodes move from one zone to another.

Algorithm 5.1 Preprocessing Phase for SPA-AODV

Notations:

Profile Table of $Node_i$ = PT_i

Profile Entry of $Node_j$ in table of $Node_i$ (PT_i) = $PT_{i.j}$

Local Profile of $Node_i$ = $LProfile_i$

Global Profile of $Node_i$ = $GProfile_i$

Local Profile Similarity computed by $Node_i$ and $Node_j$ = $LProfile_{i.j}$

Global Profile Similarity computed by $Node_i$ and $Node_j$ = $GProfile_{i.j}$

for (for all nodes $Node_i$ receiving hello packets periodically from $Node_j$) **do**
  Calculate $LProfile_{i.j}$ and $GProfile_{i.j}$
  if $LProfile_{i.j} \geq threshold$ and $PT_{i.j} \notin PT_i$ **then**
    Write $LProfile_j$ and $GProfile_j$ in PT_i
    Increment Hop Count and broadcast by updating sequence number
  end if
  if $LProfile_{i.j} \geq threshold$ and $PT_{i.j} \in PT_i$ **then**
    Update $LProfile_j$ and $GProfile_j$ in PT_i
    Increment Hop Count and broadcast by updating sequence number
  end if
end for

The profile is exchanged among nodes using $Profile\ Unit$. Each node stores the received profiles in respective $Profile\ Store$. Algorithm 5.1 explains the method of exchanging profiles to store in Profile Tables when the profile similarity metric is higher than the threshold. Thus, node N_3 can exchange profile with nearby nodes Nb(N_3)=$\{N_0, N_1, N_4\}$ using Profile Units (as shown in Figure 5.1). The profile store of the nodes is also updated as profile units received and similarity computed.

Table 5.1: Profile Store of Node N_3

Node	Profile Vector	Similarity Metric	Hop Count
N_0	PV_{N_0}	0.43	1
N_4	PV_{N_4}	0.70	1
N_1	PV_{N_1}	0.85	1
N_2	PV_{N_2}	0.95	2

Table 5.1 shows profile store of N_3 after preprocessing phase. A node that wants to establish ASN can scan profile table to determine a destination and can establish a relatively stable path so that intermediate nodes have profile similarity higher than the threshold.

5.1.1.2 Phase 2 - Route Establishment Phase

Nodes connect each other to create ASN in the route establishment phase by sending RREQ and RREP packets. Referring to Figure 5.1 in which there are following two paths from N_0 to N_2:

Path 1: $N_0 \xrightarrow{0.43} N_3 \xrightarrow{0.85} N_1 \xrightarrow{0.95} N_2$

Path 2: $N_0 \xrightarrow{0.57} N_4 \xrightarrow{0.70} N_3 \xrightarrow{0.85} N_1 \xrightarrow{0.95} N_2$

Path 2 is preferred when N_0 needs a path to N_2 with profile similarity higher than 0.44, else Path 1 is the best path.

Algorithm 5.2 RREQ in Route Establishment Phase

Trigger:

$Node_i$ receives route request from $Node_j$, wherein source is $Node_A$ and destination is $Node_D$

Notations:

Routing table of $Node_i$ = RT_i

Routing table entry of $Node_j$ in RT_i = $RT_{i.j}$

Threshold requested by $Node_A$ = TH_A

if $Node_i \neq D$ **then**
  if $Node_i$ received first RREQ **then**
    Create entry in RT_i for $Node_j$
    Broadcast RREQ in its neighborhood
  else
    if $SequenceNumber$ is newer OR Profile Similarity is higher than TH_A **then**
      update RT_i with entry $RT_{i.j}$
      Broadcast RREQ in its neighborhood
    end if
  end if
end if

Algorithm 5.3 RREP in Route Establishment Phase

Trigger:

$Node_i$ receives RREP from $Node_j$, wherein source is $Node_A$ and destination is $Node_D$, the RREP sent by $Node_D$

Notations:

Routing table of node $Node_i$ = RT_i

Route Entry of $node_j$ in RT_i = $RT_{i.j}$

Threshold requested by $Node_A$ = TH_A

if $Node_i \neq A$ **then**
  if $Node_i$ receives first $RREP$ **then**
    Create entry in RT_i for $Node_j$
  else
    if $SequenceNumber$ is newer OR Profile Similarity is higher than $Threshold_A$ **then**
      Update RT_i with entry $RT_{i.j}$
    end if
  end if
end if

5.1.2 Implementation

SPA-AODV has been implemented using ns2 by modifying existing AODV routing protocol. The AODV protocol is an on-demand IP routing protocol that uses route discovery process to find a route to a destination. It maintains the active routes only and the routes not in use get expired. It determines multiple routes from source to destination but keeps only a single route. Each node maintains a sequence number that increases when the node notices a change in the neighborhood topology. The sequence number helps to prevent loops and provide route freshness. The protocol uses routing tables to store routing information. The routing table stores (i) Destination Address, (ii) Next-Hop Address, (iii) Destination Sequence Number, and (iv) Life-Time. A node maintains a list of precursor nodes for each destination for route maintenance. Life-time of the route is updated when used else it gets expired. A node that wants to send a data packet to a destination checks its routing table if it has a current route to the destination and forwards the packet to next-hop if the route is available. A route discovery process is initiated by generating RREQ when the route is not available. RREQ contains IP address and the sequence number of the source node and the destination. The packet also includes broadcast ID that gets an increase when source node sends RREQ. A RREQ is uniquely identified using the source IP address and broadcast ID. A route to the destination is discovered by flooding control packets, and the discovered routes are later used to send data packets.

An intermediate node that receives RREQ sets up a reverse route entry for the source node. The reverse route entry includes IP address and the sequence number of source address, number of hops for the source node, next-hop address, a life-time. A node can respond to RREQ by unicasting RREP if it has an unexpired route to destination and sequence number of the destination is higher than in RREQ. RREP contains IP address of the source and the destination, destination sequence number, hop-count, and lifetime. An intermediate node receiving RREP sets up a forward path entry to the destination. The entry includes IP address of the destination and next-hop, hop-count for the destination, and lifetime. In case a node receives multiple RREPs, it forwards the first RREP and may forward another RREP if the destination sequence number is higher or has lower hop-count. The reverse path entry in the routing table entry is purged after a time-out interval that is sufficient to allow RREP to comeback. The forward path entry is purged when it is not used for an active route timeout interval. If an intermediate node or destination node moves, all active neighbors are informed by Route Error (RERR). A node receiving an RERR, marks route to the destination as invalid by setting the distance to destination as infinity. Link failures are detected by *Hello* messages. Neighboring nodes periodically exchange *Hello* messages and absence of *Hello* message indicate link failure.

AODV is implemented in ns2 with help of files namely *aodv.h*, *aodv.cc*, *aodv_logs.cc*, *aodv_packet.h*, *aodv_rqueue.h*, *aodv_rqueue.cc*, *aodv_rtable.h*, *aodv_rtable.cc*.

The routing table entry has below fields:

Listing 5.1: Routing Table Entry

```
class aodv_rt_entry {
    nsaddr_t rt_dst;                  // Destination Address
    nsaddr_t rt_nexthop;              // Next-hop
    u_int16_t rt_hops;                // Hop count - Metric
    u_int32_t rt_seqno;               // Sequence Number
    aodv_precursors rt_pclist;        // Precursors
    aodv_ncache rt_nblist;            // Active Neighbors
    double rt_req_timeout;            // when I can send another req
    u_int8_t rt_req_cnt;              // Number of broadcasted RREQ
    double rt_expire;                 // when entry expires
    u_int8_t rt_flags;                // can be RTF_DOWN, RTF_UP, RTF_IN_REPAIR
};
```

The RREQ has following fields:

Listing 5.2: Route Request (RREQ) Format (aodv_packet.h)

```
struct  hdr_aodv_request  {
    u_int8_t          rq_type ;          //  Packet  Type
    u_int8_t          reserved [2];
    u_int8_t          rq_hop_count;      //  Hop Count
    u_int32_t         rq_bcast_id ;      //  Broadcast  ID
    nsaddr_t          rq_dst ;           //  Destination  IP  Address
    u_int32_t         rq_dst_seqno ;     //  Destination  Sequence Number
    nsaddr_t          rq_src ;           //  Source  IP  Address
    u_int32_t         rq_src_seqno ;     //  Source  Sequence Number
    double            rq_timestamp;      //  when REQUEST sent;
};
```

The RREP has following fields:

Listing 5.3: Route Reply (RREP) Format (aodv_packet.h)

```
struct  hdr_aodv_reply  {
    u_int8_t          rp_type ;               //  Packet  Type
    u_int8_t          reserved [2];
    u_int8_t          rp_hop_count;           //  Hop Count
    nsaddr_t          rp_dst ;                //  Destination  IP  Address
    u_int32_t         rp_dst_seqno ;          //  Destination  Sequence Number
    nsaddr_t          rp_src ;                //  Source  IP  Address
    double            rp_lifetime ;           //  Lifetime
    double            rp_timestamp;           //  when corresponding  REQ sent;
};
```

The RERR has following fields:

Listing 5.4: Route Error (RERR) Format (aodv_packet.h)

```
struct  hdr_aodv_error  {
    u_int8_t          re_type ;          //  Type
    u_int8_t          reserved [2];      //  Reserved
    u_int8_t          DestCount;         //  DestCount
    //  List  of  Unreachable  destination  IP  addresses  and sequence  numbers
    nsaddr_t          unreachable_dst [AODV_MAX_ERRORS];
    u_int32_t         unreachable_dst_seqno [AODV_MAX_ERRORS];
```

The *Hello* Message has following fields:

Listing 5.5: Hello Message Format (aodv_packet.h)

```
struct  hdr_aodv_reply  {
    u_int8_t          rp_type ;               //  AODVTYPE_HELLO
```

```
    u_int8_t        reserved [2];
    u_int8_t        rp_hop_count;         // Hop Count as 1
    nsaddr_t        rp_dst ;              // Destination IP Address
    u_int32_t       rp_dst_seqno ;        // Destination Sequence Number
    nsaddr_t        rp_src ;              // src_ from hdr_ip (ip.h)
    double rp_lifetime ; // (1 + ALLOWED_HELLO_LOSS) * HELLO_INTERVAL
};
```

The routing protocol SPA-AODV was registered as ASN in ns2 to set the protocol in the tcl script. This research created a copy of AODV protocol folder of the ns2 and first renamed all class files to ASN and then modified ASN to simulate the SPA-AODV. The newly added protocol ASN was registered in *ns − packet.tcl*, *ns − mobilenode.tcl*, *ns − lib.tcl*, *Makefile* of following files in a similar way as AODV is mentioned:

- /home/parallels/ns2/ns-allinone-2.35/ns-2.35/tcl/lib/ns-packet.tcl
- /home/parallels/ns2/ns-allinone-2.35/ns-2.35/tcl/lib/ns-lib.tcl
- /home/parallels/ns2/ns-allinone-2.35/ns-2.35/tcl/lib/ns-mobilenode.tcl
- /home/parallels/ns2/ns-allinone-2.35/Makefile
- /home/parallels/ns2/ns-allinone-2.35/ns-2.35/Makefile

The routing table entry as explained in Source Code Listing 5.1 was amended to include fields specific to ASN as below:

Listing 5.6: Additional fields of a Routing Table Entry for ASN

```
class asn_rt_entry {
    u_int8_t    rt_lprofile [ZONES][LDIC_SIZE];
    u_int8_t    rt_gprofile [GDIC_SIZE];
};
```

The *RREQ* as explained in Source Code Listing 5.2 was amended to include fields specific to ASN as below:

Listing 5.7: Additional Fields for RREQ (asn_packet.h) for ASN

```
struct hdr_asn_request {
```

```
    u_int8_t     rq_lprofile [ZONES][LDIC_SIZE];   // Local Profile of source
    u_int8_t     rq_gprofile [GDIC_SIZE];          // Global Profile of source
    double       rq_threshold ;                    // Threshold for Profile Match
}
```

The *RREP* as explained in Source Code Listing 5.3 above was amended to include fields specific to ASN as below:

Listing 5.8: Additional Fields for RREP (asn_packet.h) for ASN

```
struct hdr_asn_reply {
    u_int8_t   rp_lprofile [ZONES][LDIC_SIZE]; // Local Profile of node sending reply
    u_int8_t   rp_gprofile [GDIC_SIZE];        // Global Profile of node sending reply
}
// updated sendReply by including profile information
sendReply(nsaddr_t ipdst , u_int32_t hop_count,
nsaddr_t rpdst , u_int8_t lprofile [ZONES][LDIC_SIZE],
u_int8_t *gprofile , u_int32_t rpseq ,
u_int32_t lifetime , double timestamp);
```

$./ns-2.35/common/mobilenode.h$ was updated to set profile, retrieve a profile, and to retrieve the current zone. The profile of a node is dynamically set during preprocessing phase based on random keywords selected from a dictionary specific to the current zone. The following constants and functions were declared in the $mobilenode.h$.

Listing 5.9: Common Profile Parameters

```
#define ZONES          4
#define LDIC_SIZE      10
#define GDIC_SIZE      (LDIC_SIZE * ZONES)
void   set_profile () ;    // Sets profile based on the current zone of the node.
int    get_Zone() ;        // retreives the zone of a moving node
// retreives the profile of a moving node
void  getProfile ( u_int8_t mlweight[ZONES][LDIC_SIZE], u_int8_t mgweight[GDIC_SIZE])
```

$./ns-2.35/common/mobilenode.cc$ bound local profiles and number of zones to allow setting the variables from TCL files. The function to set local and global profile of a node was also defined (Source Code Listing 5.9).

5.1.3 Results and Analysis

Table 5.2: Simulation Parameters

Parameter Name	Value
Number of Nodes	4, 8, 16, 32, 64
Speed	1, 2, 3
Area	1000*1000, 1500*1500, and 2000*2000
Number of Zones	8, 16
Packets Sent	1000 packets with 512 bytes payload sent in 100 seconds interval

Table 5.2 provides list of parameters used for simulation. A randomly selected node was configured to connect other nodes that have profile similarity higher than a preconfigured threshold. Awk scripts were used on trace files to compute Overhead and Packet Delivery Ratio.

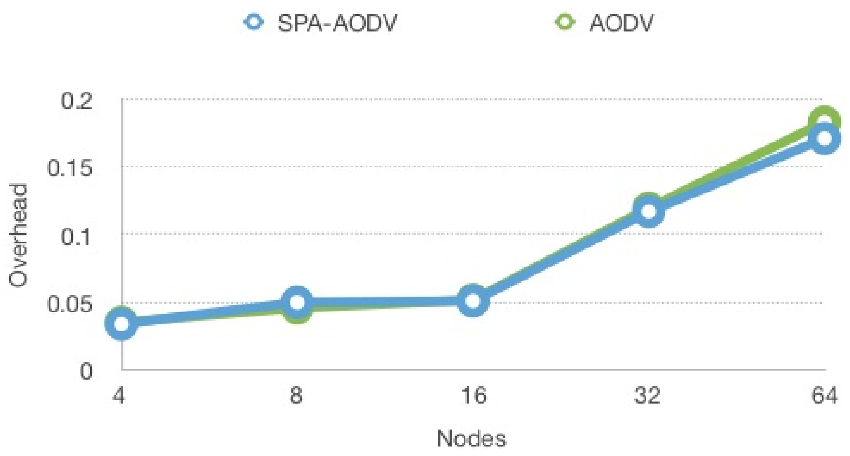


Figure 5.2: Overhead

Figure 5.2 shows overhead comparison of SPA-AODV and AODV. Overhead is the number of routing packets or control packets for network communication. Results indicate that the overhead in SPA-AODV is slightly less in case of 64 nodes and in all other cases it is almost equal.

Packet Delivery Ratio (PDR) is computed using generated and received packets as logged in the trace file. PDR is defined as the ratio of received packets and the generated packets. The below figures displays packet delivery ratio when nodes are moving with speed of 1 m/sec, 2 m/sec, and 3 m/sec.

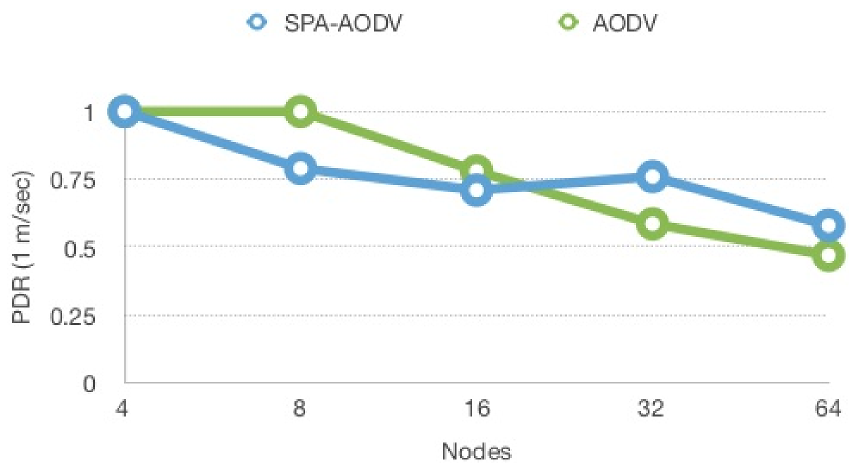


Figure 5.3: Packet Delivery Ratio when nodes are moving with speed of 1 m/sec

Figure 5.3 displays packet delivery ratio when nodes are moving with speed of 1 m/sec.

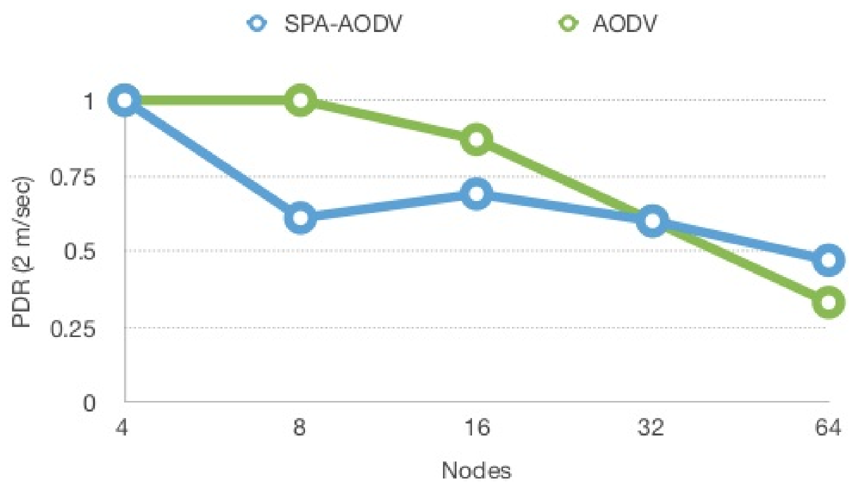


Figure 5.4: Packet Delivery Ratio when nodes are moving with speed of 2 m/sec

Figure 5.4 displays packet delivery ratio when nodes are moving with speed of 2 m/sec

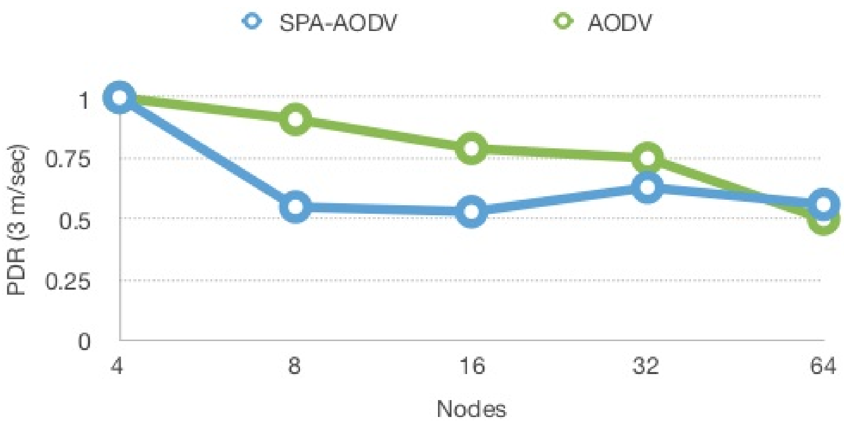


Figure 5.5: Packet Delivery Ratio when nodes are moving with speed of 3 m/sec

Figure 5.5 displays packet delivery ratio when nodes are moving with speed of 3 m/sec.

Results indicate that SPA-AODV performs better when number of nodes are high, for example, PDR is good when number of nodes are greater than 16 in case of speed 1 m/sec (Figure 5.3) and when greater than 32 in case of speed of 2 m/sec (Figure 5.4).

One of the reasons that SPA-AODV performs better in case of higher nodes is the increased chances of profile similarity among nodes. As discussed above, SPA-AODV outperforms in overhead also with the higher number of nodes.

5.2 ASN using WiFi Direct Peer-to-Peer networking (P2P) Technology on the top of Android

This section presents the implementation of ASN at one-hop for Android Mobile Phones. The one-hop MANET is created using P2P that is supported by Android Phones higher than API 14 (Android 4.0 Ice Cream Sandwich or higher version). Mobile devices can be connected with each other to form groups by P2P technology without any centralized infrastructure or base station. The participating devices negotiate roles, and one of the devices become P2P Group Owner (GO), and the other device becomes a client. The devices use Media Access Control (MAC) address as the device ID for discovery and communication. P2P is based on IEEE 802.11 infrastructure technology, but a soft access point is created after the negotiation [45]. This research proposes architecture and implementation of Offline Chat (OffAT), which is an android based mobile application to create ASN using P2P.

5.2.1 Architecture

The OffAT [52] broadcasts users interests provide Profile Similarity and group communication without using any centralized infrastructure. The main components of the OffAT include Device Discovery, Interest-based Social network, and Inter-group communication.

Devices discover neighboring devices by selecting a listen channel and alternate between search-state and listen-state. The time allocated to each state can vary from 100 ms to 300 ms. Literature survey indicate that researchers have either used MAC address or some other confidential information to discover and being discovered [11]. Some researchers have also proposed using Bluetooth to discover and a centralized server to compute profile similarity. However, this research proposes a new method of sharing limited profile and username information without sharing any confidential information and compute profile similarity without any centralized server. The device id is modified with username and profile information so that the information is available to other nearby users without any other communication and without the need for any centralized server. The username and interests are broadcasted during device discovery phase. The devices participate to create a ASN if the profile similarity is higher than a threshold value. The interest-based ASN can be created among nearby users with similar profiles.

Once the devices are connected, the devices participate in sharing messages and files using P2P. Security of network is a prime concern nowadays since wireless networks are vulnerable to active and passive attacks. However, communication security and privacy are available in the proposed architecture since ASN uses P2P that further uses Wi-Fi Protected Access II (WPA2) implementing Advanced Encryption Standard (AES), which is a stronger encryption algorithm. Further, the proposed architecture doesn't require sharing the mobile number or email for the purpose of creating ASN. The connection request can also be declined or ignored and be disconnected anytime after being connected.

5.2.2 Implementation

The architecture has been implemented using Android Studio.

Listing 5.10: Receiving User Name and Interests

```
// MainActivity.java
public void startOffatAddListener (){
    startOffat.setOnClickListener(new View.OnClickListener() {
        @Override
        public void onClick(View v) {
            WifiManager wifi = (WifiManager)
             getApplicationContext().getSystemService(Context.WIFI_SERVICE);
             wifi.reconnect();
```

```
            if  (!e1.getText().toString().equals("") &&
            !e2.getText().toString().equals("")) {
                GlobalData.name = e1.getText().toString();
                GlobalData.interests =
                e2.getText().toString().replaceAll("\\s+",",").replaceAll(",+",",");
                Intent intent = new Intent(v.getContext(), WiFiDirectActivity.class);
                startActivity(intent);
            }
        }
    });
}
```

The code Listing 5.10 (*MainActivity.java*) explains extracting user id and interests from a user interface.

Listing 5.11: Setting Name and Interests

```
// WiFiDirectActivity.java
public void setNameInterests(){
    try {
        Method method = manager.getClass().getMethod("setDeviceName",
        WifiP2pManager.Channel.class, String.class, WifiP2pManager.ActionListener.class);
        method.invoke( manager, channel,
        GlobalData.name+"#"+GlobalData.interests, new WifiP2pManager.ActionListener() {
            public void onSuccess() {}
            public void onFailure(int reason) {}
        });
    } catch (Exception e){
    Toast.makeText(getApplicationContext(),
    "Unable send User name and Interests ",Toast.LENGTH_SHORT).show();
    }
}
```

The code Listing 5.11 (*WiFiDirectActivity.java*) explains setting the device id as user name#interests to participate in the discovery process. Setting name and interests as device id help other devices to receive profile information without using any centralized server and without being connected with irrelevant users.

Listing 5.12: Computing Profile Similarity and displaying nearby users

```
// DeviceListFragment.java
// Array adapter for ListFragment that maintains WifiP2pDevice list.
private class WiFiPeerListAdapter extends ArrayAdapter<WifiP2pDevice> {
    @Override
```

```
    public View getView(final int position, View convertView, ViewGroup parent) {
        View v = convertView;
        if (v == null) {
            LayoutInflater vi = (LayoutInflater)
            getActivity().getSystemService(Context.LAYOUT_INFLATER_SERVICE);
            v = vi.inflate(R.layout.row_devices, null);
        }
        device = items.get(position);
        String interestArray = "";
        if (device != null) {
            TextView device_name = (TextView) v.findViewById(R.id.device_name);
            TextView device_details = (TextView) v.findViewById(R.id.device_details);
            if (device_name != null) {
                String[] yourInterestWords = device.deviceName.split("#")[1].split(",");
                String[] myInterestWords = myDeviceName.split("#")[1].split(",");
                List<String> yourInterestWordsList = Arrays.asList(yourInterestWords);
                List<String> myInterestWordsList = Arrays.asList(myInterestWords);
                List<String> matchedWords = matched(myInterestWordsList, yourInterestWordsList);
                String str1 = "";
                for (String s : matchedWords)
                {
                    str1 += s.trim() + " ";
                }
                int similarityPercentage = matchedWords.size()*100/myInterestWordsList.size();

                str = device.deviceName;
                if (similarityPercentage >0){
                    str += "\nMatched: "+str1+"("+ similarityPercentage +"%)";
                }else{
                    str += "\nMatched: none (0%)";
                }
                device_name.setText(str);
                str = "";
            }
        }
    }
}
```

The code Listing 5.12 (*DeviceListFragment.java*) explains computing profile similarity percentage based on common keywords.

5.2.3 Results and Analysis

In the existing implementation, the interests are entered manually, but in future work, the interests can be computed based on prior user action and browsing history. Top three keywords along with their frequencies appearing in browsing history and prior user actions can be considered to automatically set as location-based interests.

Figure 5.6: User Interface to configure Device ID

Figure 5.6 displays a user interface to enter username and interests.

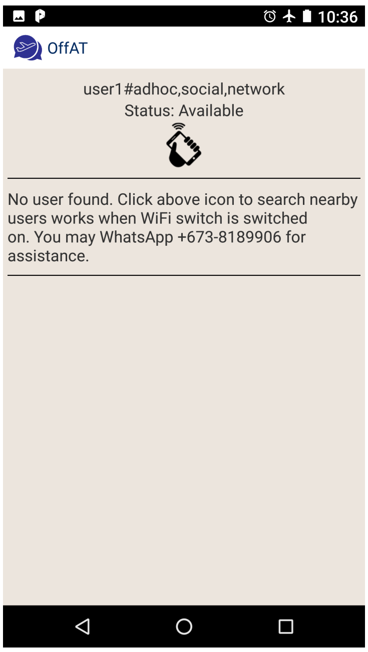


Figure 5.7: Device ID (displaying user name#user interests) in discovery phase with status as available

The Figure 5.7 displays user's status as available along with user name#interests as visible to other nearby users. A device can have two status as Available or Connected. A connected device can either be group owner of an existing group member or a client. Figure 5.7 shows device discovery phase and displays the device status as available.

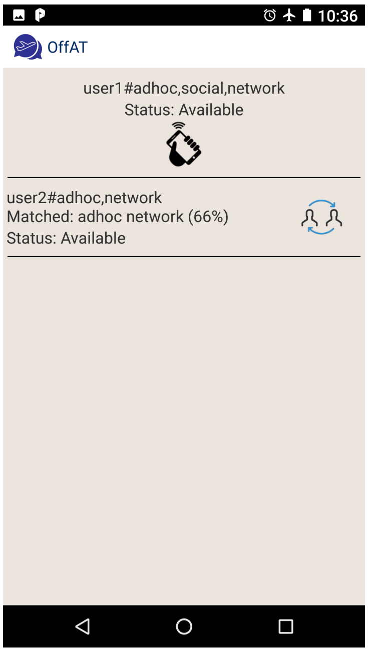


Figure 5.8: Nearby available devices with their interests and profile similarity

Figure 5.8 shows two devices with status as available. Figure 5.8 is the interface of *user*1 that can see another nearby user *user*2, with profile keywords ad-hoc and network, is available with profile similarity as 66%. The figure also shows an icon along with the nearby user to start connection request. If the connection request from *user*1 is accepted by the user *user*2, the status gets converted to connected as shown in Figure 5.9 with icons to start a chat or to disconnect.

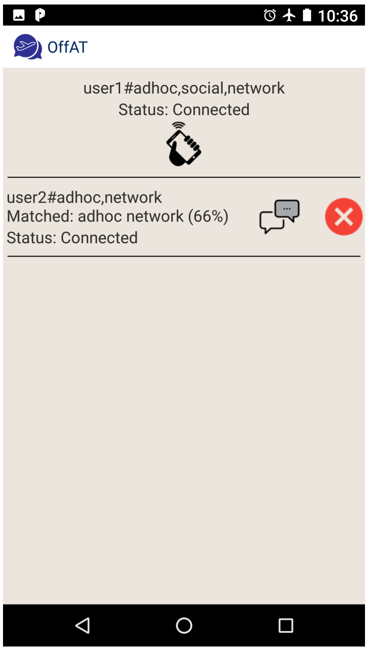


Figure 5.9: Nearby connected device with option to chat/disconnect along with profile and similarity

Figure 5.9 shows two devices with status as connected with icons to start a chat or to disconnect.

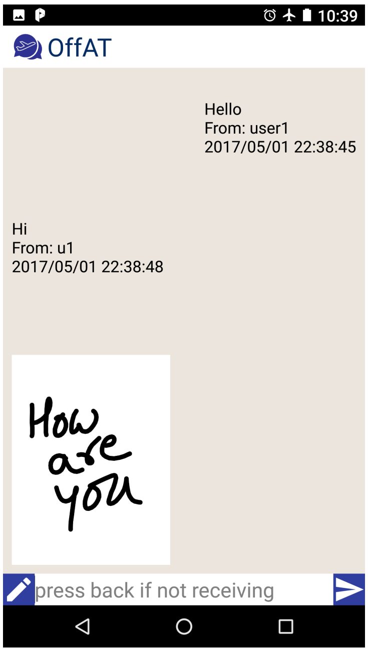


Figure 5.10: Text and handwritten messages being exchanged using P2P in Airplane Mode

The Figure 5.10 shows the real-time chat communication including text and handwritten messages in an Airplane Mode.

The Android-based mobile application is available at Google Play Store [52]. The application has been developed using Android Studio with minimum SDK as API 14 (Android 4.0 Ice Cream Sandwich version). The application was tested by connecting two devices in the following combinations:

Table 5.3: Status of Devices when connected with P2P

Device 1	Device 2
Airplane Mode	Airplane Mode
Airplane Mode	WiFi
Airplane Mode	4G
Airplane Mode	3G
Airplane Mode	2G

The devices were able to connect with each other and able to communicate successfully in all the combinations mentioned in the Table 5.3.

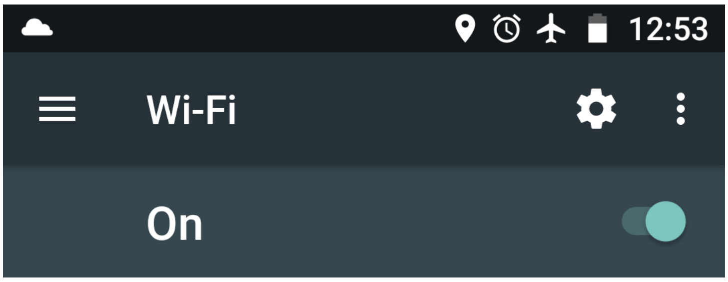


Figure 5.11: WiFi Status when using P2P in Airplane Mode

The only concern was when trying in Airplane Mode since the WiFi switched should still be switched on. There is no need to connect to WiFi Hotspot, but WiFi should be switched on as shown in Figure 5.11

This research tried to programmatically switch on WiFi with permission from users. However, it shows conflict when changing device id immediately after switching to WiFi. Thus, future work should handle to broadcast username and interests in Airplane mode without manually switching on WiFi after activating airplane node.

Implementation results and feedback posted at Google Play Store (4.5 rating and 100-500 installs) indicate user preferences towards location-based social networking. Users have suggested storing username and interests, and it can be configured by Shared Preferences in the Android.

5.3 Summary

This chapter discussed context-aware routing protocol and an android based mobile application that creates ASN over android.

The social profile aware ad-hoc on-demand distance vector routing protocol that is context-aware routing protocol has been proposed in this chapter. The protocol selects intermediate nodes based on profile similarity threshold value for a path between a source and a destination. The protocol was simulated in two phases comprising pre-processing phase to set-up interest value for different zones and route establishment phase to select a path between source and destination. The proposed protocol was compared with ad-hoc on-demand distance vector routing protocol for parameters overhead and packet delivery ratio. Results indicate that the proposed protocol performs better when the number of nodes is greater.

An android based mobile app, called OffAT-Chat in Airplane Mode, was developed for Android using WiFi peer-to-peer networking. The application allows users to set dynamic interests anonymously and access the interests of nearby users without getting connected. The mechanism helps users to know other users interests and profile similarity value before sending or accepting the connection request. The application was tested for group communication among students and works well with a group of students.

Chapter VI

CONCLUSION AND FUTURE WORK

Chapter VI

CONCLUSION AND FUTURE WORK

This chapter provides the major accomplishments of this research work and provides directions to further research in the area of ASN.

6.1 Conclusion

The significant achievements of this research work are proposed middleware architecture, findings from users survey that can be used to customize ASN applications, geo-social profile matching algorithm that can be used to provide location-based social networking applications, semantic geo-social profile matching algorithm for location-based profile matching, piecewise maximal similarity to evaluate and score profile matching, social profile aware routing protocol to provide contextual aware routing, and an Android-based mobile application to connect nearby users based on their interests.

The suggested middleware architecture is a layered-architecture for ASN containing four layers Application Layer, Transport Layer, Ad-hoc Social Layer, and Ad-hoc Communication Layer. The architecture can be used to build and maintain ASN and present functions like profile manager, profile store, matching engine, similarity profile store, and privacy module under the ad-hoc social layer.

The findings from a users' survey reveal an escalating trend for ASN. The results from users survey show the mobile device as the platform to access the current social networking applications. 91% users conceded using a mobile device to access social networking. Further, the usage of the device and social networking applications is common since 68% users access these applications every day. All users indicated utilizing a search engine, and further 94% users agreed that the interests change with the location always or sometimes. Thus, ASN can exploit the users' preferences to provide new profile managing algorithm. Regarding favored similarity score, users average choice was 75%.

The geo-social profile algorithm is a location-aware profile algorithm that provides location-based social networking applications based on local-profiles. A local profile is created based on location-based prior actions including search and browsing history. Implementation of geo-social profile algorithm using Google Trends that provides top keywords at the different location indicates better performance of profile matching using location-based interests when the similarity threshold is high.

The semantic geo-social profile algorithm is a modified version of the geo-social algorithm. The modification includes in creating a profile and matching profile wherein the interest level of a keyword is increased when the keywords are semantically equal. The implementation of the algorithm for 20 users with random interest level for a particular example of keywords with semantically equal keywords shows better performance of semantically match when the required similarity threshold value is greater than 60%.

The piecewise maximal similarity is an improvement over cosine similarity for weighted interests for ASN. The piecewise maximal similarity is the maximum similarity based one each interest attribute that is computed by taking summation of the minimum of the interest level of two users for a particular keyword. The minimum value of the two interest level indicates the scale that these two users are maximal similar. Results and implementation of the piecewise similarity for random users and real-time data set indicate improved performance. The shown results compare friendship predicted by cosine similarity and piecewise maximal similarity with that of observed by facebook and Bluetooth for different values of the threshold. There is 4% performance gain over cosine similarity in accuracy (percent correct) and 5% in false alarm ratio (predicting friend when the users were not friends as observed) when compared with Facebook data and 3% performance gain over cosine similarity in accuracy when compared with Bluetooth.

A context-aware social routing protocol, social profile aware ad-hoc on demand distance vector routing protocol, proposed in this research work is capable of selecting intermediate nodes so that the profile similarity of the nodes on the path is high. Results indicate the proposed protocol performs better than AODV when the number of nodes is high. The overhead in case of the proposed protocol is almost same as AODV for up to 32 nodes and marginally higher for 64 nodes. The packet delivery ratio in the proposed protocol is better for nodes higher than 16 nodes when nodes are moving with a speed of 1m/sec and better when nodes are higher than 32 nodes when nodes are moving with speed of 2 m/sec. In a normal real-life situation, the speed of 1 m/sec or 2 m/sec is a realistic walking speed of users, and thus the proposed protocol will perform better in real-life scenarios.

A mobile application was developed using the stock android version that can create ASN and match users interests using P2P. The app, OffAT-Chat in Airplane Mode, is available at Google Play Store and can be used to create ASN on Android 4.0+ devices. The application is capable to broadcast selected username and interests in a nearby region and shows a simple keyword-based profile similarity score with the other received profiles. The application also allows exchanging text or handwritten messages to the other connected users.

6.2 Future Work

Profile Management and Matching profile is not just an issue in ASN but also in an online social network like Facebook, Twitter, and LinkedIn etc. The future work can include:

- Usage of geo-social profile algorithm and PMS Metric for the application of online social network.
- Usage of SPA-AODV for the MANET communication in general

Currently, most of the flights ask to stop using Mobile phones and wireless network, but no technical study has been conducted whether the usage of mobile phones with P2P interferes with flight communication. A technical study should be conducted to verify if P2P or MANET communication is flight safe or not. The study can be conducted by measuring signal strength when the number of users using P2P is very high to know if it interferes with flight communication. The current practice of keeping mobile in airplane mode is because the mobile phone is customized to increase the signal strength when it couldn't find communication cellular tower. However, Wi-Fi or P2P doesn't work like that since the mobile phone doesn't increase signal strength if it is unable to connect to Wi-Fi hotspot or it P2P peers. This will open new applications for ASN. Thus, the future work may include:

- Measuring signal strength of P2P when the number of mobile devices are high to verify if the signal is sufficiently high to interfere with airplane communication
- Developing ASN applications for airplane passengers to play online games and passenger to airplane communication
- Developing ASN applications to push notifications to nearby users by businesses

REFERENCES

[1] T. Sanguankotchakorn, S. Shrestha, and N. Sugino, "Effective ad hoc social networking on OLSR MANET using similarity of interest approach," *Internet and Distributed Computing Systems*, pp. 15–28, 2012.

[2] S. Gambhir, N. Aneja, and L. C. De Silva, "Piecewise Maximal Similarity for Ad-hoc Social Networks," *Wireless Personal Communications*, vol. 97, no. 3, pp. 3519–3529, 2017.

[3] A. Caputo, A. Socievole, and F. D. Rango, "CRAWDAD dataset unical/socialblueconn (v. 2015-02-08)," Downloaded from http://crawdad.org/unical/socialblueconn/20150208, Feb. 2015.

[4] N. Aneja and S. Gambhir, "Social Profile Aware AODV Protocol for Ad-hoc Social Networks," *Wireless Personal Communications*, vol. 97, no. 3, pp. 4161–4182, 2017.

[5] C. Funai, C. Tapparello, and W. Heinzelman, "Supporting multi-hop device-to-device networks through wifi direct multi-group networking," *arXiv preprint arXiv:1601.00028*, 2015.

[6] K. Seada and C. Perkins, "Social networks: the killer app for wireless ad hoc networks?" *Nokia Research Centre NRC-TR-2006-010*, 2006.

[7] L. Zhang, X. Ding, Z. Wan, M. Gu, and X.-Y. Li, "Wiface: a secure geosocial networking system using wifi-based multi-hop manet," in *Proceedings of the 1st ACM Workshop on Mobile Cloud Computing & Services: Social Networks and Beyond*. ACM, 2010, p. 3.

[8] D. Zhang, D. Zhang, H. Xiong, C.-H. Hsu, and A. V. Vasilakos, "BASA: Building mobile Ad-Hoc social networks on top of android," *IEEE Network*, vol. 28, no. 1, pp. 4–9, 2014.

[9] L. V. Hoang and H. Ogawa, "A platform for building ad hoc social networks based on Wi-Fi direct," in *Global Conference on Consumer Electronics (GCCE)*. IEEE, 2014, pp. 626–629.

[10] E. Chung, J. Joy, and M. Gerla, "DiscoverFriends: Secure social network communication in mobile ad hoc networks," in *International Wireless Communications and Mobile Computing Conference (IWCMC)*. IEEE, 2015, pp. 7–12.

[11] J. Joy, E. Chung, Z. Yuan, J. Li, L. Zou, and M. Gerla, "DiscoverFriends: secure social network communication in mobile ad hoc networks," *Wireless Communications and Mobile Computing*, vol. 16, no. 11, pp. 1401–1413, 2016.

[12] T. M. Ramos, B. R. Moura, M. B. Damasceno, M. Holanda, and A. Araujo, "FrameGeoSocial mobile Ad-hoc social network based on geolocation," in *Iberian Conference on Information Systems and Technologies (CISTI)*. IEEE, 2015, pp. 1–5.

[13] M. A. Rahman and M. S. Hossain, "A Location-Based Mobile Crowdsensing Framework Supporting a Massive Ad Hoc Social Network Environment," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 76–85, 2017.

[14] D. Bottazzi, R. Montanari, and A. Toninelli, "Context-aware middleware for anytime, anywhere social networks," *IEEE Intelligent Systems*, vol. 22, no. 5, 2007.

[15] A. Campbell, S. Eisenman, K. Fodor, N. Lane, H. Lu, E. Miluzzo, M. Musolesi, R. Peterson, and X. Zheng, "Cenceme: Injecting sensing presence into social network applications using mobile

phones (demo abstract)," in *Proc. of Ninth ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'08), Hong Kong*, 2008.

[16] E. Sarigöl, O. Riva, P. Stuedi, and G. Alonso, "Enabling social networking in ad hoc networks of mobile phones," *Proceedings of the VLDB Endowment*, vol. 2, no. 2, pp. 1634–1637, 2009.

[17] M. L. Yiu, S. Šaltenis, K. Tzoumas *et al.*, "Efficient proximity detection among mobile users via self-tuning policies," *Proceedings of the VLDB Endowment*, vol. 3, no. 1-2, pp. 985–996, 2010.

[18] E. Sarigöl, O. Riva, and G. Alonso, "A tuple space for social networking on mobile phones," in *Data Engineering (ICDE), 2010 IEEE 26th International Conference on*. IEEE, 2010, pp. 988–991.

[19] J. Lee and C. S. Hong, "A mechanism for building ad-hoc social network based on user's interest," in *Asia-Pacific Network Operations and Management Symposium*. IEEE, 2011, pp. 1–4.

[20] Q. L. Trieu and T. V. Pham, "STARS: Ad-hoc peer-to-peer online social network," in *International Conference on Computational Collective Intelligence*. Springer, 2012, pp. 385–394.

[21] H. Li, K. Bok, and J. Yoo, "Mobile P2P social network using location and profile," in *Ubiquitous Information Technologies and Applications*. Springer, 2013, pp. 333–339.

[22] L. Zhang, X.-Y. Li, and Y. Liu, "Message in a sealed bottle: Privacy preserving friending in social networks," in *Distributed Computing Systems (ICDCS), 2013 IEEE 33rd International Conference on*. IEEE, 2013, pp. 327–336.

[23] R. Zhang, J. Zhang, Y. Zhang, J. Sun, and G. Yan, "Privacy-preserving profile matching for proximity-based mobile social networking," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 656–668, 2013.

[24] Y. Wang, A. V. Vasilakos, Q. Jin, and J. Ma, "A Wi-Fi Direct Based P2P Application Prototype for Mobile Social Networking in Proximity (MSNP)," in *International Conference on Dependable, Autonomic and Secure Computing*. IEEE, 2014, pp. 283–288.

[25] Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Friendbook: a semantic-based friend recommendation system for social networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 3, pp. 538–551, 2015.

[26] M. A. Khan, W. Cherif, F. Filali, and R. Hamila, "Wi-fi direct research-current status and future perspectives," *Journal of Network and Computer Applications*, vol. 93, pp. 245–258, 2017.

[27] E. Spertus, M. Sahami, and O. Buyukkokten, "Evaluating similarity measures: a large-scale study in the orkut social network," in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 2005, pp. 678–684.

[28] J. Li and S. U. Khan, "MobiSN: Semantics-based mobile ad hoc social network framework," in *Global Telecommunications Conference (GLOBECOM)*. IEEE, 2009, pp. 1–6.

[29] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec, "Effects of user similarity in social media," in *Proceedings of the fifth ACM international conference on Web search and data mining*. ACM, 2012, pp. 703–712.

[30] P. Symeonidis, D. Ntempos, and Y. Manolopoulos, “Location-based social networks,” in *Recommender systems for location-based social networks*. Springer, 2014, pp. 35–48.

[31] H. Li, K. S. Bok, and J. S. Yoo, “P2P Based Social Network over Mobile Ad-Hoc Networks,” *IEICE Transactions on Information and Systems*, vol. 97, no. 3, pp. 597–600, 2014.

[32] H. B. Liaqat, F. Xia, J. Ma, L. T. Yang, A. M. Ahmed, and N. Y. Asabere, “Social-similarity-aware TCP with collision avoidance in ad hoc social networks,” *IEEE Systems Journal*, vol. 9, no. 4, pp. 1273–1284, 2015.

[33] X. Han, L. Wang, N. Crespi, S. Park, and Á. Cuevas, “Alike people, alike interests? Inferring interest similarity in online social networks,” *Decision Support Systems*, vol. 69, pp. 92–106, 2015.

[34] S. Mizzaro, M. Pavan, and I. Scagnetto, “Content-based similarity of twitter users,” in *European Conference on Information Retrieval*. Springer, 2015, pp. 507–512.

[35] Y. Zhang, J. Tang, Z. Yang, J. Pei, and P. S. Yu, “Cosnet: Connecting heterogeneous social networks with local and global consistency,” in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015, pp. 1485–1494.

[36] N. Kraus, D. Carmel, I. Keidar, and M. Orenbach, “NearBucket-LSH: Efficient similarity search in P2P networks,” in *International Conference on Similarity Search and Applications*. Springer, 2016, pp. 236–249.

[37] F. Yu, N. Che, Z. Li, K. Li, and S. Jiang, “Friend Recommendation Considering Preference Coverage in Location-Based Social

Networks," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer, 2017, pp. 91–105.

[38] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, "Applications, architectures, and protocol design issues for mobile social networks: A survey," *Proceedings of the IEEE*, vol. 99, no. 12, pp. 2130–2158, 2011.

[39] A. M. Ahmed, F. Xia, N. Y. Asabere, H. B. Liaqat, and J. Li, "Social community-partition aware replica allocation in ad-hoc social networks," in *IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*. IEEE, 2013, pp. 834–841.

[40] H. B. Liaqat, Q. Yang, A. M. Ahmed, Z. Xu, T. Qiu, and F. Xia, "A social popularity aware scheduling algorithm for ad-hoc social networks," in *International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, 2014, pp. 28–33.

[41] R. P. Marinho, U. B. Menegato, and R. A. R. de Oliveira, "Mobile devices routing using wi-fi direct technology," in *Advanced International Conference on Telecomunications*, 2015, pp. 83–89.

[42] F. Xia, L. Liu, J. Li, J. Ma, and A. V. Vasilakos, "Socially aware networking: A survey," *IEEE Systems Journal*, vol. 9, no. 3, pp. 904–921, 2015.

[43] S. Palani, S. Kumar, and P. Madhuri, "Influencing social networks for P2P substance-based file distribution in disjointed MANETs," *International Journal of Computer Applications*, vol. 122, no. 2, 2015.

[44] R. Gupta, N. Krishnamurthi, U.-T. Wang, T. Tamminedi, and M. Gerla, "Routing in Mobile Ad-Hoc Networks using Social Tie

Strengths and Mobility Plans," in *Wireless Communications and Networking Conference (WCNC), 2017 IEEE*. IEEE, 2017, pp. 1–6.

[45] C. Funai, C. Tapparello, and W. Heinzelman, "Enabling multi-hop ad hoc networks through WiFi Direct multi-group networking," in *International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2017, pp. 491–497.

[46] N. Aneja and S. Gambhir, "Middleware Architecture for Ad-hoc Social Network ," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 13, no. 9, pp. 690–695, 2016.

[47] S. Gambhir, N. Aneja, and S. Mangla, "Need of Ad-hoc Social Network based on Users' Dynamic Interests," in *International Conference on Soft Computing Techniques and Implementations (ICSCTI)*. IEEE, Oct 2015, pp. 52–56.

[48] S. Gambhir, N. Aneja, and S. Mangla, "Survey about Social Networking Groups and Making Friends Anytime Anywhere," Google Form, 2015.

[49] Google, "Google Trends," Extracted from https://trends.google.com/trends/home/b/IN, 2017.

[50] N. Aneja and S. Gambhir, "Geo-Social Profile Matching Algorithm for Dynamic Interests in Ad-Hoc Social Network," *Social Networking*, vol. 3, no. 5, pp. 240–247, 2014.

[51] N. Aneja and S. Gambhir, "Geo-Social Semantic Profile Matching Algorithm for Dynamic Interests in Ad-hoc Social Network," in *International Conference on Computational Intelligence Communication Technology (ICCICT)*. IEEE, Feb 2015, pp. 354–358.

[52] N. Aneja and S. Gambhir, “OffAT: Ad-hoc Social Network on the top of Android,” Google Play Store. [Online]. Available: https://play.google.com/store/apps/details?id=com.naneja.offat1&hl=en

BRIEF PROFILE OF THE RESEARCH SCHOLAR

Name:	Nagender Aneja
Designation:	Researcher, Institute of Applied Data Analytics and Senior Manager, Office of Innovation Enterprise, Universiti Brunei Darussalam, Brunei Darussalam
Qualification:	Ph.D. (2011 to current) M.E. Computer Technology and Applications (2002), Delhi College of Engineering
Research Interests:	Mobile Ad-hoc Network, Social Network, Data Analytics, Deep Learning, Natural Language Processing
Work Experience:	Researcher, Institute of Applied Data Analytics and Senior Manager, Office of Innovation Enterprise, Universiti Brunei Darussalam, (2013 - to current) Associate IP Lead, Microsoft Patent Research Services, CPA Global Noida (2007-2012) Assistant Professor, Tecnia Institute of Advanced Studies Delhi, IP University (2005-2007) Assistant Professor, ABES Engineering College Ghaziabad, UP Technical University (2004-2005) Senior Lecturer, ITM Gurgaon, MD University (1999-2004) Lecturer, Lala Lajpat Rai Engineering College Moga Punjab, Punjan Technical University (1998-1999)

LIST OF PUBLICATIONS OUT OF THE THESIS

List of Published Papers

Sr. No.	Title of Paper	Name of Journal/Conference where published	No	Volume and Issue	Year	Pages
1	Profile-Based Ad Hoc Social Networking Using Wi-Fi Direct on the Top of Android	Mobile Information Systems, Hindawi			2018	Article ID 9469536
2	Social Profile Aware AODV Protocol for Ad-hoc Social Networks	Wireless Personal Communications, Springer	3	97	2017	4161-4182
3	Piecewise Maximal Similarity for Ad-hoc Social Networks	Wireless Personal Communications, Springer	3	97	2017	3519-3529
4	Middleware Architecture for Ad-hoc Social Network	Research Journal of Applied Sciences, Engineering and Technology	9	13	2016	690-695
5	Geo-Social Profile Matching Algorithm for Dynamic Interests in Ad-Hoc Social Network	Social Networking	5	3	2014	240-247
6	Ad-hoc Social Network - A Comprehensive Survey	International Journal of Scientific and Engineering Research	8	4	2013	156-160

Continued on next page

– continued from previous page

Sr. No.	Title of Paper	Name of Journal/Conference where published	No	Volume and Issue	Year	Pages
7	Geo-Social Semantic Profile Matching Algorithm for Dynamic Interests in Ad-hoc Social Network	IEEE International Conference on Computational Intelligence and Communication Technology, Ghaziabad			2015	354-358
8	Need of ad-hoc social network based on users' dynamic interests	IEEE International Conference on Soft Computing Techniques and Implementations (ICSCTI), Faridabad			2015	52-56
9	Software design for social profile matching algorithm to create ad-hoc social network on top of Android	International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi			2016	3450-3454
10	Various Issues in Ad-hoc Social Networks	National Conference on Recent Trends in Computer Science and Information Technology			2012	6-9

List of Patents

Sr. No.	Title	Country	App No	Pub No	Filing Date	Status
1	Ad-Hoc Social Networking and Profile Matching System	USA	US 14/981,146	US 2016/0330773	Dec 28, 2015	Pending
2	Ad-Hoc Social Networking and Profile Matching System	India	3105/ DEL/ 2015	Official Journal of The Patent Office, Issue No. 13/2017, page 527	Sept 29, 2015	Pending

List of Mobile Applications

Sr. No.	Title	OS	Status
1	OffAT: Ad-hoc Social Network on the top of Android	Android 4.0+	Available at Google Play Store