Enhancing Recommendations in Mobile Social Network

Hadeer Hassan Ibrahim,
Faculty of Computer and Information Sciences,
Ain Shams University, Cairo, Egypt,
Hadeer.h.shahin@gmail.com

Tamer Abdelkader,
Faculty of Computer and Information Sciences,
Ain Shams University, Cairo, Egypt,
tammabde@cis.asu.edu.eg

Rania El Gohary
Faculty of Computer and Information Sciences,
Ain Shams University, Cairo, Egypt,
Rania.Elgohary@cis.asu.edu.

Abstract— In Mobile Social Networks (MSNs), people contact each other through mobile devices, such as smartphones and tablets, while they move freely. The communication takes place on-the-fly by the opportunistic contacts between mobile users via local wireless bandwidth, such as Bluetooth or WiFi without a network infrastructure. Social Multicast is an important routing service in MSNs where data transmission is addressed to a group of users according to their social features. The aim of this paper is to find and recommend mobile nodes that can efficiently relay and consume messages based on their social features. Efficiency in this context is to achieve high delivery ratio while reducing considering resources constraints and limitations such as power and space. The proposed algorithm, TESS, measures social similarity based on **T**ime-based **E**ncounter of **S**ocially **S**imilar nodes. We compare the proposed algorithm with the known social multicast algorithms: Multi-CSDO, EncoCent and Epidemic. Simulations results show that the proposed algorithm outperforms others in terms of delivery ratio and network overhead.

I. Introduction

In Mobile Social Networks (MSNs) individuals with similar interests are likely to meet and contact each other through their mobile devices, such as smartphones and tablets. This process is similar to web-based social networking, except that in MSNs there is no infrastructure.

Many web-based social networks, such as Facebook and Twitter, have created mobile applications to give their users instant real-time access as long as they are connected to the internet.

Moreover, there is an inherent connection between mobile and web networks, as mobile apps use existing web social networks to create native communities and promote discovery. The web-based social networks take advantage of mobility features and accessibility, such as in location based apps.

On MSN, the connection between nodes is usually short-term, time-dependent, and unstable as nodes move freely.

Most of the previous multicast routing protocols do not consider the social features of the mobile nodes [1, 2]. Recently, social aware protocols have been proposed that use the users' social interaction as well as their interests to find and recommend target destinations [3] and [4]. Other social aware protocols propose community detection, where the mobile nodes are classified into groups, the central node is determined to be the most connected node, and achieve the shortest path to destinations [5, 6, and 7]. The authors in [8] propose a social profile-based multicast algorithm (SPM) based on static social features in user profiles. The work in [9], [10] and [11] use assigned profiles to nodes and take advantage of the fact that people having more similar social features in common tend to meet and contact more often in MSNs. One of the old is usually used as a which family known protocol-well benchmark to compare with ،is Epidemic [12]. In this protocol, all nodes are used as relays to find the target destinations .Epidemic usually achieves a high delivery ratio, given unlimited resources. However, its overhead on the bandwidth and power consumption is considerably high ،which renders the protocol non-efficient.

In MSNs, each node in the network can meet other nodes and can send messages to them, and all nodes are discriminated by their social feature values extracted from a web social network.

In this paper, the social features of mobile nodes will be exploited to efficiently recommend destinations and find candidate relays to hold messages. The proposed algorithm, **T**ime-based **E**ncounter of **S**ocially **S**imilar nodes, TESS, measures the similarity between encountered nodes and use it to determine if the message should be relayed or not. Similarity includes the values of social features, count of the number of meetings and the age of each meeting (how old when they met). To test the performance of the proposed algorithm, simulation experiments have been conducted to measure several performance metrics: hop count, network overhead, and the number of dropped messages. Results show that the proposed algorithm outperforms previous ones in the previously mentioned metrics.

The rest of the paper is organized as follows: Section II references the related work. Section III presents our proposed algorithm. Section IV shows and discussessimulation results; Conclusions are drawn in Section V.

II. Related Work

Routing in MSNs should be capable of handling opportunistic contacts, intermittent connectivity, highly mobile nodes, power and storage constrained devices, and the possible nonexistence of end-to-end paths.

In [3], the authors recommend friends based on building social communities. These communities depend on multiple keywords, such as interests, history and current locations of different users.

In [6], the authors use the concept of home-aware community in MSN. They propose community aware opportunistic routing (CAOR) algorithm, and the opportunistic routing depends only on a few nodes in the network.

In [7], the work detects communities of friends based on the complexity and type of community detected.

The work in [4] studies interests and daily life of mobile nodes. The authors proposed SCORP, which is based on a utility function. This function reflects the probability of encountering nodes with a certain interest among the ones that have similar daily social habits. It considers one interest for each node so it doesn't reflect all interests of the node.

In [5], a community-based multicast routing scheme was proposed by exploiting node centrality and social community structures. This approach exploits the awareness of individual nodes to their data forwarding probabilities to destinations.

In [8], a social-profile-based multicast (SPM) algorithm is proposed, which uses static social features in user profiles. The aim of this algorithm is to select the most relevant relays in the multicast routing in MSNs. This algorithm discovers the most important and representative social features affiliation distance and common language from the Infocom06 trace [14] is collected from an experiment in which people carried Bluetooth devices during the Infocom 2006 conference. It contains the contact records as well as reports the social profile information of the participants. The selection of candidate relays is calculated with a small average affiliation distance or high common language ratio to the destinations. Therefore, the algorithm does not need to record node contact history, but the people's static behavior may not reflect the actual behavior of people.

In [11], the static social feature was extended to dynamic social feature. The authors propose a social-similarity-based multicast (Multi-Sosim) routing algorithm. In dynamic social features, what is recorded is not only if a node has the same social feature values with the destination, but also the frequency this node has met other nodes. These nodes have the same social features values during the recent time interval. This definition of dynamic social features is based on frequency, which cannot distinguish the cases of two nodes have the same frequencies but one of them is closer to a destination more than the other one.

The multicast algorithm in MSNs can be implemented using simple approaches, such as Epidemic routing [12]. Epidemic forwards the message to all nodes in the network without any condition, which enhances the delivery ratio if resources are unlimited, but it has high forwarding cost.

In [9], the relationships between destinations and the relationships between each relay and community of destinations are considered. Two Community and Social feature based multicast algorithms are proposed, named Multi-CSDO and Multi-CSDR. Multi-CSDO involves destination nodes only in community detection, while Multi-CSDR involves both the destination nodes and the relay candidates in community detection. The social similarity between a node and a community should include all of the social features values of the nodes involved. The social similarity between two destination nodes is calculated using static social features. The social similarity between a relay candidate and a destination is calculated using enhanced dynamic social features. In every connection between two nodes, they start to combine all destinations held by two nodes. In CSDO, they force two communities of these destinations to make the destinations that have higher similarity near to each other. Then, in CSDR they measure the similarity between the two nodes and these communities, to make the node hold the community of destinations that have higher similarity to it, to decide which is better to deliver the message to the destination.

In [10], the node's social context information is used to calculate the social similarity utility between a node and destination. The priority of these social features is statically set by adding weight to each social feature. The social connection of network nodes is used to calculate the betweenness centrality utility of a node. The authors propose that each node has a set of evidence and each evidence has a weight (the importance of the evidence in the network).

We can categorize the work done into three categories: creating communities, exploiting social profiles, and social independence. In creating communities, a central node is selected that has the most connected number of nodes. This enables the central node to reach most of the other nodes in the network. In exploiting social profiles, weights are assigned to nodes based on their social similarity. These works are based on the assumption that nodes with similar social profile meet more than the less similar nodes. The third category does not depend on the social interaction between nodes, which is considered the traditional multicast protocols for Mobile adhoc networks. Our proposed protocol belongs to the second category, which is to exploit the social profile. We enhance the weights calculated based on social features to efficiently deliver messages and reduce cost.

III. PROPOSED ALGORITHM: TIME-BASED ENCOUNTER OF SOCIALLY SIMILAR NODES (TESS)

The proposed algorithm aims to enhance the process of finding destinations known by their profiles. To achieve this goal, two enhancing features are introduced: Two-message based scheme, and time-based encounter weight calculation. In the next subsections, we are going to explain each enhancing feature.

A. Two-message based scheme

The source nodes hold two types of messages. The first one is a small message that contains the social features values describing the destinations, where destinations are determined by a set of social features values. The small message will be sent to all encountered nodes in an epidemic style routing. The second one is the main data message, which contains the actual data required to reach the target destinations. The data message will be sent to the found destinations and the relay candidates that have a high opportunity to meet these destinations.

The procedure of the two-message based scheme between two connected nodes i and$j$, can be detailed as follows:

Time	At node	Action
	i	Send the small information message to j, containing the social features of the required destination
	j	1. Receive the message 2. Calculate the similarity weight between itself and the target destinations 3. Reply back to i with an acknowledgement, Ack, containing the similarity weight
	i	1. Receive the Ack message 2. Determine if j could be a destination and relay node or not 3. If yes, send a copy of the main data message to j
	j	If the data message arrived: 1. Receive the Data message 2. Read the list of target destinations 3. If it is one of them, forward the message to the application layer. 4. Whether it is a destination or not, repeat the procedure with other contacting nodes.

B. Time-based Encounter Similarity Weight Calculation

Measuring the weight of encounter between nodes was based only on its count [9]. For example, Assume the target destination profile consists of three social features: Location, Language, and Age. Assume a source node is multicasting a message to all nodes with social profile: <Location=”Egypt”, Language=”Arabic”, Age=20>. The source node meets two candidate relay nodes. One of them has met the same social profile one day ago, and the other has met the same social profile one hour ago. In previous encounter-based algorithms, both nodes have the same weight towards this social profile, which is not true from our perspective.

We develop with the following equations to calculate the weight between a node and target destinations,

In Eq. (1), we calculate the weight of the common social features values between an encountered node and a destination. Let f_i be the value of the i’s feature of the common social features vector $f = \{f_i, \forall i = 1..k\}$, then the weight W_{f_i} equals

$$W_f = {\left(\sum_{j=1}^{n} t_A - t_j\right)}\Big/{n} \qquad (1)$$

Where,

t_A: the time at which the two nodes appear in the communication range of each other (System time),

t_j: the element j in the vector of encounter times with nodes having the same value of the social feature,

n: length of the vector of encounter times (count of nodes whose feature values equal to this node), and

In TESS, the node has a high opportunity to meet the destination, if it has common social feature values with destination as well as it is a central node that it is trading with many others have similar social profile.

After calculating the weight of each social feature, we can calculate the similarity weight between the encountered node and the target destination using Eq. (2).

$$W_{node-D} = {\left(\sum_{i=1}^{k} W_{f_i}\right)}\Big/{m} \qquad (2)$$

Where,

W_{node-D}: is the similarity weight of the node with the destination at a specific time,

k: number of the common social features values between the node and the destination, and

m: is the count of connections made between this node and encountered nodes in the past, whether it has common social features values with these encountered nodes or not.

If m is increasing and summation of common social features values for this node is decreasing, therefore this node is not similar to the nodes surrounding it, and hence, having less opportunity to meet the destination.

If the similarity is high between a node and the total nodes that it came across, this indicates that this is a central node.

The pseudo-code for the proposed protocol is presented in Algorithm 1.

C. **TESS Example**

Assume the Actual time, t_A=18:00 .Table 1 and Table 2 are profiles of two nodes in the network that have a connection in this actual time.

Each node has three social features (Location, Language and Age) with different values, node1’s social features vector is <Location=USA, Language=English, Age=20>.

Let propose that each node has made 20 connections in the history that has common social features values or not (m=20).

Each feature is associated with a vector of times. In our example node1 has met other nodes with the social feature value (Location=USA) in 12:00, 13:00, 14:00 and 17;00.

Because node1 and node2 have a current connection, the current time will be added to the vector of times of the common social feature values.

The common social feature value is Age, so 18:00 is added to the vector of encounter times of this social feature value in the two nodes. The updated profiles are shownin Table 3 and Table 4.

Profile of the target destination, D, is <Location=USA, Language =Arabic, Age=20>. For node1, the common social feature values with D is < Location =USA, Age=20 >. For node2, the result is < Language =Arabic, Age=20 >, using Eq. (1), the weight of these social feature values as shown in Table. 5 and Table 6.

Using Eq. (2), the similarity weight of node1 with D,

$$W_{1-D} = (4+5)/20 = 0.45,$$

The similarity weight of node2 with D.

$$W_{2\text{-}D} = (5+0)/20 = 0.25,$$

As common languages between two nodes with D is equal, so the calculation showed that the weight between node1 and D is higher than that between node2 and D. Therefore, node1 is preferred to be a relay than node2 to reach destination D. These steps are repeated to all destinations gathered from node1 and node2.

Algorithm 1: TESS

Require: The source node x and its destination set D_s = {d1, d2, · · ,dn};
1: While not all of the destinations receive the message do,
2: On contact between a message holder x and node y (x, y profiles contain social features values and vector of times with other nodes)
3: if $y \in D_x$ then
4: /* Found destination y */
5: y gets a copy of the original message
6: endif
7: /* Compare node social similarity and split the destinations
*/
8: Calculate common social features values between each node and D_x
9: Calculate weight to common social features values of node x (W_x) and common social features values of node y (W_y) according to D_x
9: if $W_x < W_y$
10: if common social features values of x, D_x = common social features values of y, D_y
11: send a copy of the original message to D_y
12: endif
13: else if common social features values of x, D_x < common social features values of y, D_y
14: send a copy of the original message to D_y
15: endif

Table 1: Profile of node1.

	Feature name	Vector Time
1	Location=USA	<12:00, 13:00, 14:00, 17:00>
2	Language=English	<9:00, 10:00>
3	Age=20	<10:00, 11:00>

Table 2: Profile of node2

	Feature name	Vector Time
1	Location=Egypt	<12:00, 13:00>
2	Language=Arabic	<11:00, 15:00>
3	Age=20	<>//this node doesn't meet nodes with Age=20

Table 3: Updated profile of node1

	Feature name	Vector Time
1	Location=USA	<12:00, 13:00, 14:00, 17:00>
2	Language=English	<9:00, 10:00>
	Age=20	<10:00, 11:00, 18:00>

Table 4: Updated profile of node2

	Feature name	Vector Time
1	Location=Egypt	<12:00, 13:00>
2	Language=Arabic	<11:00, 15:00>
3	Age=20	<18:00>

Table 5: Calculated weights of features of node1

	Feature name	Vector Time	Weight
1	Location =USA	<12:00, 13:00, 14:00, 17:00>	$\frac{\left(\begin{array}{c}(18{:}00-17{:}00)+(18{:}00-14{:}00)\\+(18{:}00-13{:}00)+(18{:}00-12{:}00)\end{array}\right)}{4} = 4$
2	Age=20	<10:00, 11:00, 18:00>	$\frac{\left(\begin{array}{c}(18-11)+(18-10)\\+(18-18)\end{array}\right)}{3}=5$

Table 6: Calculated weights of features of node2

	Feature name	Vector Time	Weight
1	*Language =Arabic*	*<11:00, 15:00>*	$\frac{\left(\begin{array}{c}((18{:}00-11{:}00)\\+(18{:}00-15{:}00))\end{array}\right)}{2}=5$
2	*Age=20*	*<18:00>*	$\frac{(18{:}00-18{:}00)}{1}=0$

IV. SIMULATION RESULTS

We compare the proposed algorithm, TESS, with Multi-CSDO [9], EncoCent [10] and Epidemic [12]. All algorithms use the context information about networks and forwards multiple message copies. Simulations were conducted using the Opportunistic Network Environment (ONE) simulator [13].

In the simulation script, we set the mobility area of nodes that have common social features values to be near to each other.

Simulations were done with a different number of nodes in the network. Message inter-arrival (creation) time is a uniform random number between 25 and 35 seconds. Each message has a set of destination social features to reach. We run several simulations to test the impact of changing buffer capacity and messages TTL. Simulation run time is set to 23200 seconds (around 6.44 hours).

The measured performance metrics are:

- Relay count: shows the number of nodes holding the message to be delivered to the destinations,
- Delivery ratio: the ratio of successful messages delivered to its destination to the number of all the messages that are created in the system,

- Overhead ratio: the ratio of the number of redundant packets to the total packets in the network,
- Hop count average: shows the average of the number of relay nodes that the delivered messages have passed through,
- Drop packets ratio: The number of deleted packets on the network.

There are two sets of experiments with different network sizes, (N=20), (N=100), (N=150) and (N=200), where N is the total number of nodes in the network. The two sets are detailed in the following.

I) Impact of Varying the Buffer Capacity

Increasing the buffer capacity of a node enables it to store more packets, and hence increases the probability of the packet to be delivered. As shown in Fig.1 increasing buffer capacity increases delivery ratio for all protocols. This is justified by the reduction of dropped packets inFig.2 and Fig. 3.

II) Impact of Varying the Packet Lifetime (TTL)

As shown in Fig. 2 and Fig. 3, increasing packets TTL decreases the dropped packets as it gives the packet more time in the network to reach the destination. As shown in Fig.4 and Fig. 5 changing the packet TTL doesn't affect the hop count.

As shown in Fig. 6, Fig. 7, Fig. 8 and Fig. 9 the proposed algorithm outperforms other algorithms in relay count and overhead ratio. This shows that the proposed algorithm is efficiently selecting the next hop to destination, which minimizes the message routing cost.

It can be found also, that the proposed algorithm achieves a delivery ratio results near to the other proposals. This means that reducing cost did not negatively impact the main objective which is the delivery ratio.

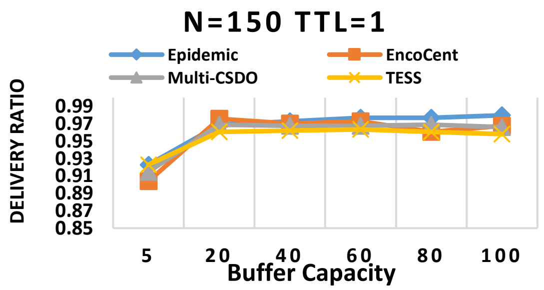

Fig.1. Impact of changing buffer capacities on Delivery Ratio in the network Size=150.

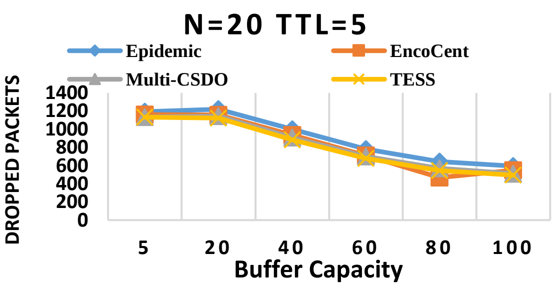

Fig.2. Impact of Buffer capacity on Dropped packets in the network size=20.

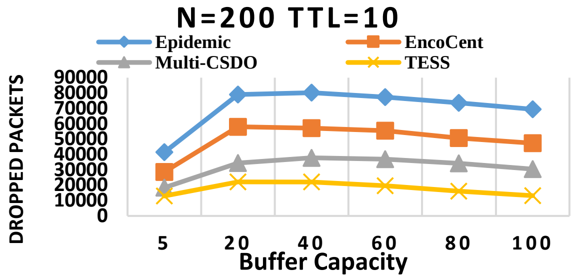

Fig.3. Impact of changing Buffer capacity on Dropped packets in the network size=200.

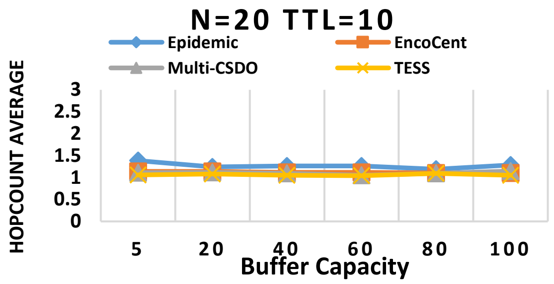

Fig. 4. Impact of changing Buffer capacity on Hop count in the network size=20.

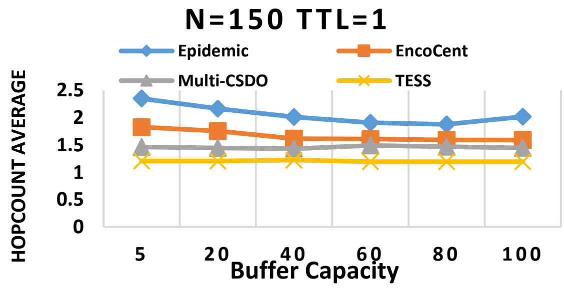

Fig.5. Impact of changing Buffer capacity on Hop count in the network size=150.

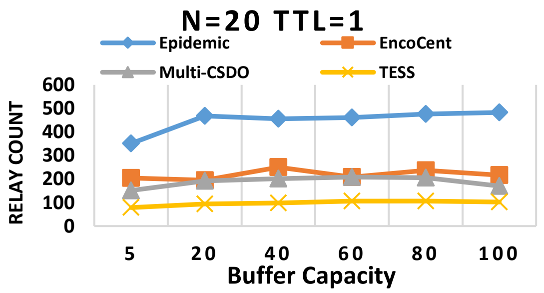

Fig.6. Impact of changing Buffer capacity on Relay count in the network size=20.

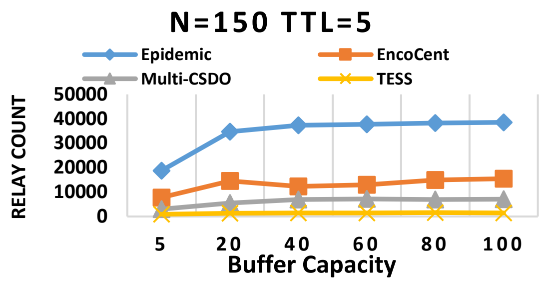

Fig.7. Impact of changing Buffer capacity on Relay count in the network size=150.

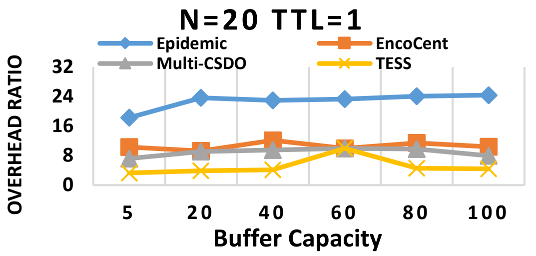

Fig.8. Impact of changing Buffer capacity on Overhead ratio in the network size=20.

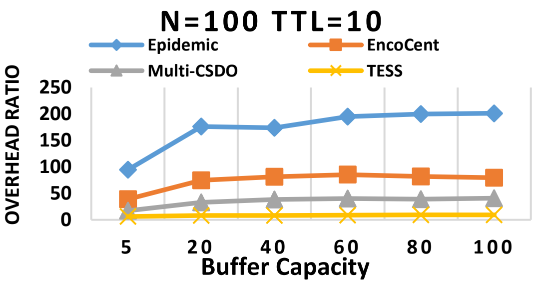

Fig.9. Impact of changing Buffer capacity on Overhead ratio in the network size=100.

V. CONCLUSION

Mobile Social Networks (MSN) are gaining attention due to the spread of mobile devices, in addition to the popularity of social networks. In this paper, we exploited the mobility of mobile devices in addition to their engagement with the web social networks to enable efficient recommendation of target destinations in Mobile Social Networks (MSNs). We proposed and implemented TESS, a time-based encounter of socially similar nodes algorithm. The proposed algorithm uses a two-message based scheme, in addition to a time-based count of encountered nodes, to determine the weight of social similarity. Using the calculated weight, relays and destinations are determined. Results show that the proposed algorithm outperforms the previous ones: Multi-CSDO, EncoCent, and Epidemicin number of relays, hop count, number of dropped messages, and overhead ratio, and has convergent results in the delivery ratios. As an extension to this work, we will consider dynamic social features, where the features and their values change with time.

REFERENCES

[1]Luo Junhai; Ye Danxia; Xue Liu; Fan Mingyu. A survey of multicast routing protocols for mobile Ad-Hoc networks. IEEE Communications Surveys & Tutorials.

[2] Lavanya, P.; Reddy, V. Siva Kumar; Prasad, A. Mallikarjuna. Research and survey on multicast routing protocols for MANETs. IEEE 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT) - Coimbatore, Tamil Nadu, India (2017.2.22-2017.2.24)] 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT) .

[3] Li, He; Bok, Kyoungsoo; Yoo, Jaesoo. An Efficient Mobile Social Network for Enhancing Contents Sharing over Mobile Ad-hoc Networkss.IEEE 2012 13th International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT) - Beijing, China (2012.12.14-2012.12.16)

[4] Moreira W., Mendes P., Sargento S. (2014) Social-Aware OpportuNistic Routing Protocol Based on User's Interactions and Interests. In: Sherif M., Mellouk A., Li J., Bellavista P. (eds) Ad Hoc Networks. ADHOCNETS 2013. Lecture Notes of the Institute for Computer Sciences, Social Informatics and TelecommuNications Engineering, vol 129. Springer, Cham

[5] W. Gao, Q. Li, B. Zhao, and G. Cao. Multicasting in delay tolerant networks: a social network perspective. In Proceedings of ACM MobiHoc, 2009.

[6] M. Xiao, J. Wu, and L. Huang. CommuNity-Aware OpportuNistic Routing in Mobile Social Networks. IEEE Trans. on Computers 63(7):1682–1695, 2014.

[7] Didwania, Ankit; Narmawala, Zunnun. A comparative study of various community detection algorithms in the mobile social network. [IEEE 2015 5th Nirma University International Conference on Engineering (NUiCONE) - Ahmedabad, India (2015.11.26-2015.11.28)] 2015 5th Nirma University International Conference on Engineering (NUiCONE).

[8] X. Deng, L. Chang, J. Tao, J. Pan, and J. Wang. Social profile-based multicast routing scheme for delay-tolerant networks. In Proceedings of IEEE ICC, pages 1857–1861, 2013.

[9] Charles Shang, Britney Wong, Xiao Chen, Wenzhong Li, Suho Oh, "CommuNity and Social Feature-Based Multicast in OpportuNistic Mobile Social Networks", Computer CommuNication and Networks (ICCCN) 2015 24th International Conference on, pp. 1-8, 2015.

[10] Xu F, Zhang H, Deng M, Xu N, Wang Z. Social-aware data forwarding in smartphone-based Delay-Tolerant Networks. InComputational Electromagnetics (ICCEM), 2016 IEEE International Conference on 2016 Feb 23 (pp. 84-86). IEEE.

[11] Y. Xu and X Chen. Social-similarity-based multicast algorithm in impromptu mobile social networks. In IEEE Globecom, 2014.

[12]A. Vahdat and D. Becker. Epidemic routing for partially connected ad hoc networks. TechNical report, Dept. of Comp. Sci., Duke UNiv., 2000.

[13] Ari Ker¨anen and r J¨org Ott. Opportunistic Network Environment simulator. Helsinki University of Technology Department of Communications and Networking Special assignment.

[14]J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD trace: cambridge/haggle/imote/infocom2006 (v.2009-05-29)," http://crawdad.cs.dartmouth.edu/cambridge/haggle/imote/infocom2006, May 2009.