A Novel Social Context-Aware Data Forwarding Scheme in Mobile Social Networks

Fang Xu^{1,2}, Huyin Zhang², Min Deng¹*, Ning Xu², Zhiyong Wang², Zenggang Xiong² and Conghuan Ye²

 ¹ School of Computer and Information Science, Hubei Engineering University, Xiaogan, Hubei, China
 ² School of Computer, Wuhan University, Wuhan, Hubei, China * dengmin@hbeu.edu.cn

Abstract

Routing in disconnected delay-tolerant mobile ad hoc networks (MANETs) continues to be a challenging issue. Several works have been done to address the routing issues using the social behaviors of each node. Mobile Social Networks (MSNs) are increasingly popular type of Delay Tolerant Networks (DTNs). The routing performance improves when knowledge regarding the expected topology and the social context information of the networks. In this paper, we introduce a new metric for data forwarding based on social context information, in which node's social context information is used to calculate the encounter utility between a node and destination, and the social relationship of network nodes is used to calculate the betweenness centrality utility of a node. We combine two utility functions to derive the social strength among users and their importance. We also present social context-based data forwarding algorithm for routing decision. Extensive simulations on real traces show that the introduced algorithm is more efficient than the existing algorithms.

Keywords: social networks; forwarding algorithms; delay-tolerant networks; encounter probability; centrality

1. Introduction

With wireless technologies, such as IEEE 802.11, WiMAX, Bluetooth, and other short range radio solutions (sensor devices), devices equipped with wireless networking capabilities are pervasively used in daily life. The ubiquity of these networked devices can facilitate communication even in the worst situations where there is no networking infrastructure, such as in earthquake or tsunami areas [1]. The pervasive deployment of wireless personal devices is creating the opportunity for the development of novel applications. The exploitation of such applications with a good performance-cost tradeoff is possible by allowing devices to use free spectrum to exchange data whenever they are within wireless range. Lately, the fast growing number of devices carried by people make also possible to form Mobile Social Networks (MSNs) for local networking services as an inexpensive alternative [2]. These networks are usually sparse and the connection between their nodes changes frequently. Many MANETs and some delay tolerant networks (DTNs) routing algorithms [3] provide forwarding by building and updating routing tables whenever mobility occurs. We believe this approach is not appropriate for MSNs, since mobility is often unpredictable and topology structure is highly dynamic.

Among many real life examples of DTNs, MSNs are of growing significance as a result of the rapid and wide spread use of various personal wireless devices (*e.g.*, cell phones, GPS devices) among people and their surroundings [4]. MSNs are a special kind

Min Deng is the corresponding author.

of DTNs, in which mobile users move around and communicate with each other via their carried short-distance wireless communication devices. In mobile social networks, there is a potential of collaborative data gathering via already deployed and human maintained devices. As more users use portable devices to contact each other, MSNs attract more attention, and opportunistic routing of messages in these networks has been studied by many researchers [5, 6]. However, due to the challenging network environment (intermittent connectivity causing lack of stable end-to-end path between nodes) in these networks, efficient routing of messages is not an easy task. To ease these difficulties and enable nodes to make right forwarding decisions while routing messages, inherent social network properties of these networks need to be utilized [7, 8].

In this paper, we propose a social context-aware data forwarding scheme, called EncoCent. In this scheme, node's social context information is used to calculate the encounter utility between a node and destination, and social relationship of network nodes is used to calculate the betweenness centrality utility of a node. As such, if the destination node is unknown to the sending node or its contacts, the message is routed to a structurally more central node where the potential of finding a suitable carrier is dramatically increased. As we know, up to now, this is the first time the two social context-based utilities are used to make data forwarding decision in mobile social networks. EncoCent combines two utility functions to derive the social strength among users and their importance, so that the source devices try to find the best carrier for data forwarding.

This paper is structured as follows. Section 2 provides a brief overview of related work on DTNs routing and MSNs routing. The system model, the data forwarding utility function and our design of EncoCent algorithm are presented in Section 3. The performance of EncoCent by comparing with existing famous DTN routing protocols is evaluated in Section 4. Finally, Section 5 concludes our work.

2. Related Work

In this section, we first overview the state-of-the-art in general DTN routing algorithms and then specifically look at the social-based routing that utilize the social network features in their designs and are proposed mainly for MSNs.

Research on data forwarding in DTNs originates from Epidemic routing [9] which floods the network. Later studies develop forwarding strategies to approach the performance of Epidemic routing with lower cost, which is measured by the number of data copies created in the network. As the result, the fastest spread of copies is achieved yielding the shortest delivery time and the minimum delay. The major drawback of this approach is excessive usage of bandwidth, buffer space and energy due to the greedy spreading of copies. Therefore, several algorithms were proposed (*e.g.*, Controlled Flooding [10], Spray and Wait [11], Multiperiod Spray and Wait [12] to limit the distribution of the message copies while still achieving high delivery rates. While the most conservative approach always keeps a single data copy and one of the first studies in the category of single-copy based algorithms is Prophet [13]. The approach is based on the observation that the movement of nodes in a typical mobile ad hoc network might be predictable based on repeating behavioral patterns.

Recently, some social-aware routing algorithms that are based on social network analysis have been proposed, such as Bubble Rap [14], HiBOp [7], Fresh [15] and dLife [16], *etc.* These studies have focused on MSNs and analyzed the social network properties of these networks to assist the design of efficient routing algorithms. When social contact patterns are exploited, data forwarding metric is centrality based, and is generally destination-independent. Various metrics have been proposed to calculate node centrality. Two key concepts in social network analysis are: (1) community, which is a group of people with social relations; (2) centrality, which indicates the social relations between a

node and other nodes in a community. Based on the two concepts, these algorithms detect the communities and compute the centrality value for each node. Messages are delivered via the nodes with good centralities. Such prior-art shows that social-based solutions are more stable than those which only consider node mobility.

However, they only consider encounter probability between nodes or centrality metrics of nodes in data forwarding, which may create bottlenecks in the network. In this paper, we introduce a new data forwarding scheme (EncoCent) different from all above studies. The EncoCent routing metrics is comprised of both a node's encounter probability calculated by social context information and its betweenness centrality. If the destination node is unknown to the sending node or its contacts, the message is routed to a structurally more central node where the potential of finding a suitable carrier is dramatically increased.

3. The EncoCent Scheme

3.1. Social Context Modeling

In order to compute the delivery probability at each node, we first define a common set of evidences that is the list of evidences used in the network and the related weight (importance of this evidence in the network). Each node, will build its node profile, which is an instantiation of the common set of evidences.

3.1.1. Encounter Probability: In the mobile social network, a node profile contains the node's social context information about the device holder (name, nationality, workplace, residence address, hobbies, *etc.*). The node profile is represented by the set of evidence/value couples. For security reasons, each evidence/value couple in the node profile is hashed by the hash function H(e, v), and the results (hashed values) are added to the node profile.

The profile structures are the same for every node. More specifically, the node profiles are composed by the same set of evidences and these evidences are in the same order. The only difference can be the values of the evidences. One of the most innovative points of our social context based routing approaches in comparison with the other context-based routing techniques is that we identify the importance of context information. In this node profile, these evidences must be in different importance priorities, and thus can be presented in different ways by using weights. We reasonably assume that, when the source node wants to send a message, it must have some information (attributes) about the destination node. Starting from the common set of evidences, each node can create its own node profile.

To classify which evidence is important to successfully deliver the message to the destination, assume that each evidence has a weight, let N(e, v) be the set of the evidence/value pairs in the profile of node N, D(e, v) is the set of the evidence/value pairs in the profile of destination D.

The set of the matched evidence/value pairs between node N and destination D is denoted by M(e, v):

$$M(e,v) = N(e,v) \cap D(e,v)$$

(1)

Definition 1. *Encounter probability*. Encounter probability is the probability that the node N meet with destination node D, as $P_{[N, D]}$.

Each node has a node profile that is composed by evidence/value pairs. The encounter probability $P_{[N, D]}$ between node N and destination D is computed by matching the profile of node N with the profile of destination D. The matching between the two profiles is done by comparing the hashed value of each evidence/value pair in the profile of N with the hashed values of the evidence/value pairs in the destination profile. It is worth

recalling that the outputs of the hash function always have a fixed size length and that the hashed values are compared in order.

Evidence weight shows that the importance of evidence in information delivery, and each evidence is assigned a weight. If *W* is the set of the weights W_m of the evidences in M(e, v); W_D is the set of the weights of the evidences in D(e, v). The encounter probability $P_{[N, D]}$ between node *N* and destination *D* is computed as follows:

$$P_{[N,D]} = \frac{\sum_{m \in W} W_m}{\sum_{d \in W_D} W_d} \tag{2}$$

Definition 2. Encounter utility within period *i*. The encounter utility that a node will encounter destination node within period i, as $EncoUtil_N(D)$.

Suppose that the set X_i contains the nodes that have the encounter probability to D higher than the encounter probability of N to $D(P_{[S, D]})$ and that these nodes are encountered by N within period i. $|X_i|$ represents the cardinality of X_i , the encounter utility between N and D within period i is computed as formula (3): with *EncoUtil_N* (D) is the encounter utility within period i, $P_{[A, D]}$ is the encounter probability between a node and the destination.

$$EncoUtil_{N}(D) = \frac{\sum_{A \in X_{i}} P_{[A,D]}}{|X_{i}|}$$
(3)

3.1.2. Betweenness Centrality

We estimate a node's centrality in the network in order to identify bridges. Centrality have proved of great value in the analysis and understanding of the roles played by actors in social networks, as well as by vertices in networks of other types, including citation networks, computer networks, and biological networks. The centrality of a node in a network is a measure of the structural importance of the node. A central node, typically, has a stronger capability of connecting other network members. There are several ways to measure centrality. The one of most widely used centrality measures is betweenness measures [17].

Determining betweenness is simple and straightforward when only one geodesic connects each pair of points. There, the central point can more or less completely control communication between pairs of others. But when there are several geodesics connecting a pair of points, the situation becomes more complicated. A point that falls on some but not all of the geodesics connecting a pair of others has a more limited potential for control. A node with a high betweenness centrality has a capacity to facilitate interactions between the nodes that it links. In our case it can be regarded as how well a node can facilitate communication to other nodes in the network. Betweenness centrality is calculated as:

$$BetCent(P_i) = \sum_{j=1}^{n} \sum_{k=1}^{j-1} \frac{g_{jk}(p_i)}{g_{jk}}$$
(4)

where g_{jk} is the total number of geodesic paths linking p_j and p_k , and $g_{jk}(p_i)$ is the number of those geodesic paths that include p_i . Betweenness centrality is a measure of network centrality that counts the paths between vertex pairs on a network that pass through a given vertex. Vertices with high betweenness lie on paths between many others and may thus have some influence over the spread of information across the network.

Freeman's centrality metrics are based on analysis of a complete and bounded network which is sometimes referred to as a sociocentric network. These metrics become difficult to evaluate in networks with a large node population because they require complete knowledge of the network topology. For this reason the concept of ego networks has been introduced. Ego network analysis can be performed locally by individual nodes without complete knowledge of the entire network. Marsden introduces centrality measures calculated using ego networks and compares these to Freeman's centrality measures of a sociocentric network, and he calculated the egocentric and the sociocentric betweenness centrality measure for the network in [18].

The betweenness centrality $BetCent(P_i)$ based on the egocentric measures does not correspond perfectly to the sociocentric measures. However, it can be seen that the ranking of nodes based on the two measures of betweenness are identical in this network. This means that two nodes may compare their own locally calculated betweenness value, and the node with the higher betweenness value may be determined.

3.2. EncoCent Utility Calculation

The utility cost of The EncoCent is a normalized value in the [0, 1] range influenced by two factors: encounter utility and betweenness centrality utility. Selecting which node represents the best carrier for the message becomes a multiple attribute decision problem, where we wish to select the node that provides the maximum utility for carrying the message. This is achieved using a pairwise comparison matrix on the normalized relative weights of the attributes. The encounter utility (*EncoUtil_N*) of node N for delivering a message to destination node D is calculated as formula (3). Here, *Cent_N* represents node the sociocentric betweenness centrality of node N which calculated by Marsden, and we reasonably assume that *Cent_N* is approximately equal to *BetCent(P_i)*. The betweenness centrality utility *CentUtil_N* of node N for delivering a message to destination node compared to node V is given as formula (5).

$$CentUtil_{N} = \frac{Cent_{N}}{Cent_{N} + Cent_{V}}$$
(5)

The *EncoCent*(*N*) is given by combining the normalised relative weights of the attributes given as formula (6), where α and β are tunable parameters and $\alpha+\beta=1$. Consequently these parameters allow for the adjustment of the relative importance of the two utility values.

$$EncoCent(N) = \alpha EncoUtil_{N}(D) + \beta CentUtil_{N}$$
(6)

3.3. Data Forwarding Algorithm

In this section, we will present our proposed data forwarding algorithm to solve the routing problem in MSNs. Based on the above Analysis, we develop a lightweight socialaware algorithm called EncoCent. The forwarding procedures when nodes N and V meet are summarized in Algorithm 1. When node N received a Hello Message from node V, node N verifies that node V is a new neighbor. If this is the case, node V send an encounter request and delivered the messages destined for node V. A list of nodes it has encountered is replied from node V, and then the list is used to update the encounter utility value on node N and the betweenness centrality utility value as described in sections 3.1.1 and 3.1.2 respectively. The two nodes then exchange a summary vector containing a list of destination nodes. This vector contains currently carrying messages and destination node profile for along with their own locally determined betweenness value and the encounter utility value for each destination. The EncoCent utility of node N is calculated for each destination in summary vector which described in section 3.2. If the EncoCent utility of node N is higher than the EncoCent utility of node V for a given destination, the destination is added to a vector of destinations for which messages are requested. When all destinations in the summary vector have been compared, the message request list is sent to node V from node N. Any messages destined for the destination node is removed from the queue of node V, and then the messages is forwarded to node N. Upon receiving a transfer message from node V the message is added to the message queue of node N. Based on the above algorithm design principles, our objective is basically trying to find the best forwarding node, which contributes to the routing in distributed approach.



4. Performance Evaluation

In this section, we conduct extensive simulations to evaluate the performance of the EncoCent algorithm. Then, our considerations about the results obtained when comparing EncoCent with dLife, Fresh and Prophet are presented. Simulation results demonstrate that the EncoCent algorithm can significantly improve MSNs routing performance.

4.1. Experiment Setup

Performance analysis is done on the Opportunistic Network Environment (ONE) [19]. In order to evaluate the premise of routing we utilized a trace of node contacts from MIT reality data set [20]. This data set consists of the traces of 97 Nokia 6600 smart phones which were carried by students and staff at MIT over nine months. These traces record contacts among users that mainly use the Bluetooth. In our simulations, we used the contacts logged during a three month period from the beginning of February to the end of April. This is the time of the second academic semester where human relations are relatively stable and participants are active on campus.

In order to simplify the analysis of the problem, we assume that there is a node profile and a set of weight of evidences in each node. In this paper, we do not consider how to obtain social context information and weights of evidences. Because of EncoUtilN is the same important as CentUtilN in this context, both α and β values are 0.5. The other simulation parameters are summarized in Table 1.

Parameters	values	Unit
Duration period	400000	S
Time window	30	s
Warm up	5000	s
Number of simulation	20	-
Area	4500×3400	m ²
Message TTL	600	min
Message size	256~512	KB

Table 1. Simulation Parameters

We ran each simulation 20 times, and we collected statistics by running each algorithm on the same set of packets. For the evaluation metrics, we use the following metrics, namely, the average delivery ratio, the delivery cost, and the routing efficiency. The average delivery ratio represents the ratio between the numbers of messages that are successfully delivered to their destination nodes within their lifetime to the total number of messages generated. The delivery cost represents the average number of forwarding used for the successfully delivered messages. On the other hand, the routing efficiency represents the ratio of the delivery probability to the delivery cost.

4.2. Simulation Results

Using simulations, we compare the proposed routing algorithm with three other benchmark algorithms described above. Each algorithm (including ours) uses the context information of nodes and deals with the forwarding of a single copy of the message to make the comparison fair.



Figure 1. Delivery Ratio vs. Time

For main simulations, we assume that the nodes have sufficiently large buffer space to store every message they receive and the bandwidth is high enough to allow the exchange of all messages between nodes at encounter times. Any change in the current assumptions is expected to affect the performance of the compared algorithms in the same way, since they all use one copy of the message. In Figure 1, it shows comparison of all algorithms including our EncoCent algorithm in terms of packet delivery ratio. Our algorithm achieves the higher delivery ratio (75 percent, close to Fresh which has the highest delivery ratio). Figure 2 shows that our EncoCent algorithm has the minimum average delivery cost, compared with the dLife who decrease by 17% with near minimum average

delivery cost. Although Fresh can achieve good delivery ratio, the average cost per packet is almost the highest among all other algorithms. Thus, the efficiency of our EncoCent illustrated in Figure 3. Consequently, its efficiency is the best among all algorithms with a 10, 61, and 102 percent improvement over dLife, Fresh, and Prophet, respectively. Based on the simulation results, the average delay that our EncoCent can achieve 75 percent of packet delivery is approximately 20 days, which outperforms all existing algorithms.



Figure 2. Delivery Cost vs. Time



Figure 3. Routing Efficiency vs. Time

According to these simulation results, we can observe that the routing strategy can be improved when more information about the social behaviors of each mobile node is available. To recall, our rationale behind is to search for a forwarding node that has a higher chance to meet with the destination node based on encounter utility and betweenness centrality utility. By finding a node that shares a similar attribute as the destination node using social profile matching coupled with the social interaction matching, the probability of delivery the packet successfully increases significantly. By these simulation results, we have verified the validity of our algorithm outperforming the others.

5. Conclusion

Since social information is quite useful to aid data forwarding in MANETs, we introduce EncoCent, which combines the $EncoUtil_N$ and $CentUtil_N$ utility functions to derive the social strength among users and their importance. The EncoCent routing metric

is comprised of both a node's encounter probability and a node's social betweenness centrality. As such, if the destination node is unknown to the sending node or its contacts, the message is routed to a structurally more central node where the potential of finding a suitable carrier is dramatically increased. We extensively evaluate our proposed algorithm using ONE simulator and compare it with the benchmark algorithms. Simulation results show that EncoCent outperform the current available methods and validate the applicability of our proposed scheme in MSNs. Finally, we are setting up real experiments to see the performance in realistic scenarios.

Acknowledgments

This research was supported in part by the National Natural Science Foundation of China (61272454, 61370092); in part by Specialized Research Fund for the Doctoral Program of Higher Education (20130141110022); in part by the NSF of Hubei Province in China (2014CFB188); in part by the Science Research Fund of Hubei provincial education office (Q20152703); in part by the Science Research Fund of Hubei Engineering University (201518); in part by Hubei Provincial Department of Education Outstanding Youth Scientific Innovation Team Support Foundation (T201410).

References

- [1] W. Kaimin, X. Liang, and K. Xu, "A survey of social-aware routing protocols in delay tolerant networks: applications, taxonmy and design-related issues", IEEE Communications Surveys & Tutorials, vol. 16, no. 1, (2014), pp. 556-578.
- [2] M. Waldir, P. Mendes, and S. Sargento, "Social-aware opportunistic routing protocol based on user's interactions and interests", Ad Hoc Networks, Springer International Publishing, (2014), pp. 100-115.
- [3] L., Anders, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks", Service Assurance with Partial and Intermittent Resources. Springer Berlin Heidelberg, (2004), pp. 239-254.
- [4] B. Eyuphan, and B. K. Szymanski, "Exploiting friendship relations for efficient routing in mobile social networks", IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 12, (2012), pp. 2254-2265.
- [5] X. Mingjun, J. Wu, and L. Huang, "Community-aware opportunistic routing in mobile social networks", IEEE Transactions on Computers, vol. 63,no. 7, (2014), pp. 1682-1695.
- [6] M. Alessandro, *et al*, "Social-Aware Stateless Routing in Pocket Switched Networks", IEEE Transactions on Parallel & Distributed Systems, vol. 26, no. 1, (2015), pp. 252-261.
- [7] C. Boldrini, M. Conti, and A. Passarella, "Exploiting users' social relations to forward data in opportunistic networks: The hibop solution", Pervasive and Mobile Computing, vol. 4, no.5, (2008), pp. 633–657.
- [8] M. Waldir, P. Mendes, and S. Sargento, "Structure and Dynamics of Social-aware Pervasive Networks", ACM SIGCOMM (2014).
- [9] V., Amin, and D. Becker. "Epidemic routing for partially connected ad hoc networks." Technical Report CS-200006, Duke University, (2000).
- [10] H. Khaled A., K. C. Almeroth, and E. M. Belding-Royer, "Delay tolerant mobile networks (dtmns): Controlled flooding in sparse mobile networks." NETWORKING 2005. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems. Springer Berlin Heidelberg, (2005), pp. 1180-1192.
- [11] Spyropoulos, Thrasyvoulos, Konstantinos Psounis, and Cauligi S. Raghavendra, "Efficient routing in intermittently connected mobile networks: The single-copy case", IEEE/ACM Transactions on Networking (TON), vol. 16, no. 1 (2008), pp. 63-76.
- [12] B. Eyuphan, Z. Wang, and B. Karol Szymanski, "Cost-effective multiperiod spraying for routing in delay-tolerant networks", IEEE/ACM Transactions on Networking (TON), vol.18, no. 5, (2010), pp. 1530-1543.
- [13] L., Anders, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks." Service Assurance with Partial and Intermittent Resources. Springer Berlin Heidelberg, (2004), pp. 239-254.
- [14] H. Pan, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay-tolerant networks", IEEE Transactions on Mobile Computing, vol. 10, no. 11, (2011), pp. 1576-1589.
- [15] D-F, Henri, M. Grossglauser, and M. Vetterli, "Age matters: efficient route discovery in mobile ad hoc networks using encounter ages", Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing, New York, USA, (2003), June 257-266.

- [16] M. Waldir, P.Mendes, and S. Sargento, "Opportunistic routing based on daily routines", IEEE international symposium on world of wireless, mobile and multimedia networks (WoWMoM), San Francisco. USA, (2012), June 1-6.
- [17] F., Linton C, "Centrality in social networks conceptual clarification", Social networks, vol.1, no. 3, (1979), pp. 215-239.
- [18] M. Peter V, "Egocentric and sociocentric measures of network centrality", Social networks, vol. 24, no. 4, (2002), pp. 407-422.
- [19] K. Ari, Jörg Ott, and T. Kärkkäinen, "The ONE simulator for DTN protocol evaluation", Proceedings of the 2nd international conference on simulation tools and techniques. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Rome, Italy, (2009).
- [20] E. Nathan, A. Sandy Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data", Proceedings of the National Academy of Sciences 106.36 (2009), pp. 15274-15278.

Authors



Fang Xu, was born in 1981. He is currently working toward the Ph.D degree with the school of computer, Wuhan University, Wuhan, China. His research interests include context aware computing, mobile social networks, and ad hoc networks.



Huyin ZHANG, was born in 1962. He is currently a Professor and a Ph.D Supervisor with the school of computer, Wuhan University, Wuhan, China. His research interests include wireless mobile networks, ad hoc networks, high-performance computing.