Community-based Message Forwarding in Mobile Social Networks

Zhiming Chen[1], Yang Xiang[2]
[1]*School of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210023, China*
[2]*Jiangsu Key Laboratory of Big Data Security & Intelligent Processing, Nanjing 210023, China*

Keywords: Delay tolerant networks, Community, Forwarding, Centrality, Movement direction.

Abstract: With the popularity of various smart devices and the application of sensor network technology, message transmission using mobile devices is becoming widespread. This paper focuses on the forwarding in mobile social network (MSN). The MSN is a special Delay Tolerant Network (DTN) consisting of mobile nodes. In MSN, nodes move and share information with each other through carried short-range wireless communication devices. Mobile nodes in the MSN typically access some building areas more frequently, such as schools, companies, or apartments, while visiting other areas, such as the roads between buildings, less frequently. The building areas that nodes frequently visit are called communities. To increase delivery ratio and reduce transmission time in MSN, this paper proposes a novel zero-knowledge multi-copy routing algorithm, Mixed Message Forwarding (MMF) which exploits and improves the metric, namely centrality. Centrality reflects the importance of a node in the network. MMF improves copy diffusion by using different directions of node movement as well. Special facilities called boundary boxes are added to the network scenario. Boundary boxes are special throw boxes. Throw boxes are relays with large storage space and fixed position. MMF is designed and evaluated, which utilizes the aforementioned boundary boxes to reduce transmission delay. The simulation results show that the MMF can improve the delivery ratio and reduce the transmission delay, compared with other algorithms.

1 INTRODUCTION

Delay tolerant networks (DTN) are a type of challenged networks, wherein the contacts between the communicating devices are intermittent. Consequently, a contemporaneous end-to-end path between the source and destination rarely exists. In DTNs, the node is usually highly mobile and often moves out of the ranges of other nodes, causing only periodic connectivity throughout the network (Hom J et al, 2017).

Mobile social networks (MSNs) are composed of mobile users that move around and use their carried wireless communication devices to share information via online social network services, such as Facebook, Twitter, etc (Xiao M et al, 2013). Recently, the short-distance communication model has also been adopted by encountered mobile users in MSNs to share information, such as multimedia, large-size files, etc., at a low cost. Such MSNs can be seen as a special kind of delay tolerant network (DTN).

Message forwarding is one of the most challenging aspects of this network because of the inherent intermittent connectivity. In this paper, we seek to address this particular problem by employing the theory of node centrality and movement direction. We propose a novel forwarding strategy, Mixed Message Forwarding (MMF), which exploits special facilities such as delay to decrease transmission delay.

The rest of this article is organized as follows: Section 2 discusses related work. Section 3 is a brief description of MMF. The forwarding process of the MMF algorithm is presented in Section 4. Section 5 shows the performance of MMF through a number of simulation experiments. We make a conclusion in Section 6.

2 RELATED WORK

Epidemic routing (Vahdate A et al, 2000), which indiscriminately floods the network with messages, has the highest delivery ratio and delivery time but

Chen, Z. and Xiang, Y.
Community-based Message Forwarding in Mobile Social Networks.
DOI: 10.5220/0008097401970202
In *Proceedings of the International Conference on Advances in Computer Technology, Information Science and Communications (CTISC 2019)*, pages 197-202
ISBN: 978-989-758-357-5

also the highest delivery cost. In order to reduce this cost, researchers interested in social network dynamics have utilized various social metrics to select the relay node. Three influential social-based protocols are SimBet (Daly E et al, 2007), Bubble RAP (Pan H et al, 2011), and Friendship routing (Bulut E et al, 2014). SimBet uses similarity and betweenness centrality metrics to determine relay nodes with higher probabilities of delivering the message. Similarly, Bubble RAP uses centrality and community to make forwarding decisions, and friendship routing considers the relationships between nodes by introducing a metric that measures quality of friendship between nodes.

The protocol proposed in Huijuan and Kai (Huijuan Z et al, 2015) expands on the work in Kim et al. (Kim CM et al, 2014) by additionally considering endpoint-biased expanded ego betweenness centrality. Contact Frequency Based Approach (CFBA) and Contact Duration Based Approach (CDBA) are two very similar routing protocols proposed in a single paper (Gondaliya N et al, 2016). They both separate the nodes into communities using the k-clique method based on contact duration, and use centrality, a metric that represents the connectivity of the network, to select the relay node. Social-Based Single Copy Routing, or SBSCR, (Gondaliya N et al, 2016) is a community-based routing mechanism in which routing decisions are made based on a calculated social based utility (SBU) that considers similarity and friendship values. The two protocols proposed in Chen and Lou (Chen H et al, 2015), Expected Encounter Routing (EER) and Community Aware Routing (CAR), are based on metrics determined through history of interaction between nodes. IRS (Singh AK et al, 2018) is an incentive based routing strategy. In this approach intermediate nodes can participate and earn incentives for sacrificing their selfishness. Choksatid et al. propose the protocol SED (Choksatid T et al, 2016), which is the renovation of Epidemic Routing scheme. (Igarashi Y et al, 2018) proposed by Igarashi et al. controls message forwarding in each terminal using parameters named “Community” and “Centrality.”

3 MMF ALGORITHM OVERVIEW

In order to solve these problems, this paper proposes a community-based opportunity network algorithm MMF algorithm. The MMF algorithm is divided into four stages: internal forwarding, external forwarding, roaming and acquisition. To more intuitively see the entire delivery process, Fig. 1 is presented a static scene. Certainly, the nodes in our simulated scene are mobile.

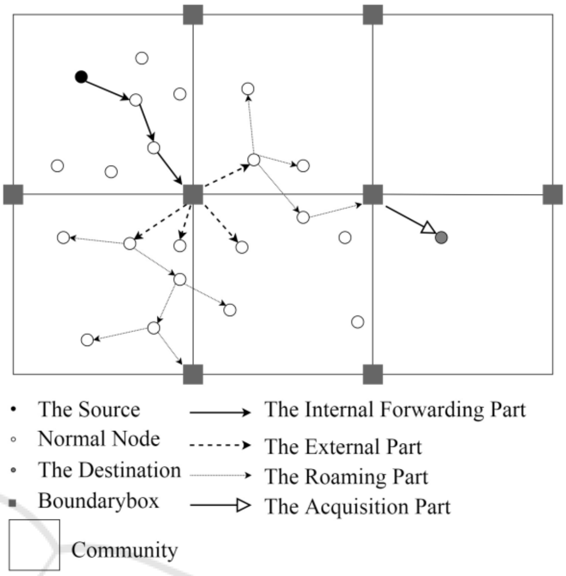


Figure 1: Network model.

(1) In the internal forwarding stage, nodes transfer the message in the community to which the source belongs until a node encounters a boundary box. Nodes are more inclined to send the message to boundary boxes.

(2) In the external forwarding stage, when a boundary box receives the copy of the messages, it will spread the messages to the neighboring communities’ nodes within its transmission area.

(3) In the roaming stage, the node forwards the messages priority to boundary boxes. We use a multi-copy approach based on node movement direction to spread information to other communities.

(4) In the acquisition stage, the destination will extract message from any first-time carrier. The device carrying the message can be a node or a boundary box.

In the above scheme, the four phases do not necessarily follow the order, which is determined by the location of the source and the destination. The source and the destination are in one community, so it is possible to go directly to the fourth phase. Or, if the destination is near the source node’s community, then the third phase will not be executed.

4 MMF ALGORITHM

In this section, we presents the MMF algorithm in detail. This article focuses on only message transfer, and it can send messages as long as each carrier has enough cache space and the link has enough bandwidth. In addition, the communication time between any two nodes and between the node and the boundary box is independent.

4.1 Internal Forwarding Stage

The central utility value can be used to measure the importance of a node in message transmission. The central utility value defined in this paper consists of two parameters, the number of node forwardings and the number of neighbor node changes. The central utility value is calculated as follows:

u_j^t is the number of node j forwardings in a certain time slot $t. u_j^t$ reflects the number of the node j acts as a relay in a certain time slot. The higher u_j^t, the higher the degree to which a node is available in that time slot.

The standard calculation method for the u_j^t is as follows:

$$u = \frac{u_j^t - \mu}{\sigma} \quad (1)$$

where $\mu = \frac{\sum_{i=1}^{t} u_j^t}{t}$ is the average of the u_j^t from the beginning to the time slot t, and $\sigma = \sqrt{\sum_{i=1}^{t}\left(u_j^t - \mu\right)^2}$ is the dispersion.

v_j^t is the number of neighbor node changes for node j at a certain time slot $t. v_j^t$ reflects the number of the node j encounters other nodes and the number of changes in neighbor nodes in the certain time slot t. The higher v_j^t means that the node meets more new nodes in the time slot t. In a time slot t, v_j^t is defined as:

$$v_j^t = \left|v_j(t) \cup v_j(t_{old})\right| - \left|v_j(t) \cap v_j(t_{old})\right| \quad (2)$$

where $v_j(t)$ is the neighbor nodes of the current time slot t, and $v_j(t_{old})$ is the neighbor nodes of the previous time slot t_{old}.

The standardized calculation method of the v_j^t is as follows:

$$v = \frac{v_j^t - \eta}{\varepsilon} \quad (3)$$

where $\eta = \frac{\sum_{i=1}^{t} v_j^t}{t}$ is the average of the v_j^t from the beginning to the time slot t, $\varepsilon = \sqrt{\sum_{i=1}^{t}\left(v_j^t - \eta\right)^2}$ is the dispersion.

The node have highter u and v values and more likely to encounter the destination node. So we combine these two parameters by the following formula to define the central utility value U_j of node j:

$$U_j = \alpha v + \beta u, \quad (4)$$

where α, β are the set weights.

When the node receives the central utility U_j of the neighbor nodes, this node compares its value to the maximum one. If this node's value is lower, then it will forward a message to the node with the largest utility value.

4.2 Spread Stage

In the spread phase, when a boundary box receives a copy of the message, the boundary box spreads message to nodes with no message outside of the community to which this copy sender belongs.

4.3 Roaming Stage

In the roaming stage, in order to speed up transmission, we use a multi-copy method. The node j carries the number of the belonging community, the current position coordinates (x_j, y_j) and the previous time slot position coordinates, such as (x_j', y_j'). These two position coordinates can be used to calculate the current slot movement direction of the node j.

Suppose that in the scene, there is a node a with copy, and there is a node b without copy in the transmission range of a. If the cosine of the current slot movement direction of the two nodes satisfies within a certain threshold range, a sends a copy to b. The range of the cosine value indicates that the arc of the angle θ between the two moving directions of a and b is around $\frac{\pi}{2}$. That is to say, in the transmission range of the node a having message, there is a node b having no message, and if the angle of movement direction of a and b is in the arc of $\frac{\pi}{2} + \triangle\theta, \triangle\theta \epsilon \left(0, \theta - \frac{\pi}{2}\right)$, the node a transmits a copy to b. The cosine value is calculated as follows:

$$|\cos\theta| = \left|\frac{(x_a - x_a', y_a - y_a') \times (x_b - x_b', y_b - y_b')}{\sqrt{(x_a - x_a')^2 + (y_a - y_a')^2} \times \sqrt{(x_b - x_b')^2 + (y_b - y_b')^2}}\right| \quad (5)$$

where $k \leq 1$.

4.4 Acquisition Stage

In the extraction phase, the destination gets the message when it encounters the message carrier. This message carrier may be in the internal forwarding stage, external forwarding stage or roaming stage. In the worst case, the message is spread to each community, then the destination just gets the message.

5 SIMULATION

In this section, we conducted a number of simulations to evaluate the performance of MMF. The comparison of the algorithms, evaluation methods, parameter settings and results are shown below.

In this paper, we only focus on zero-knowledge multi-copy routing algorithm in MSN. In order to obtain a fair performance comparison, we only compare our algorithm with the existing zero-knowledge multi-copy routing algorithm: Epidemic algorithm, Prophet algorithm and HS algorithm.

Epidemic, Prophet and HS algorithms all transfer messages through replication. The Epidemic algorithm can achieve the best delivery delay in all routing algorithms. The Prophet algorithm is a utility-based multi-probability routing algorithm. In the HS algorithm, the message carrier uses a binary method to send a message to the encountered node or throwbox.

5.1 Parameter Settings

The scenes simulated in this paper are relatively large, the nodes are assumed to move randomly, and the initial positions of the nodes are also randomly generated. The boundary box is generated at initialization time and has a fixed location. In addition, we can alter the parameter value as needed, so that we can observe the impact of each parameter value on the result and bring out the optimal results, which is beneficial to compare with other algorithms and evaluate the pros and cons of the algorithm.

In the simulation of this paper, the scene of the model is a large rectangle. To simplify the simulation, we set the length and width of the scene to be the same. The node's transmission radius is set to 15, and the number of communities is fixed at a value of 9. The experiment is roughly divided into three parts in terms of the average delivery rate, the average transmission time and the influence of the angle cosine value on the average transmission time.

In the first experiment, the community length and width D of the four algorithms were adjusted to 50, 60 and 70, and then we compare the experimental results. In the second experiment, the nodes' number N of the four algorithms was adjusted to 1000, 2000, and 3000, and then we compare the experimental results. The evaluation parameters are shown in Table 1.

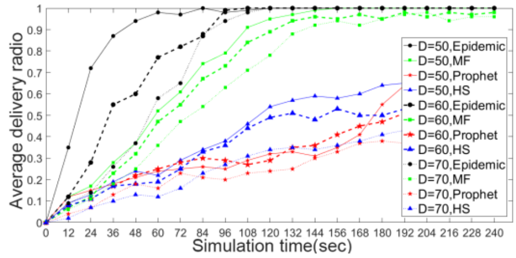
Table 1: Parameter Settings.

Parameter Name	Range
Experimental Area $(L*W)$	150*150/180*180/210*210
Number of Nodes (N)	1000/2000/3000
Node Transmission Radius	15
Number of Communities $(DMAX)$	9
Community Length And Width (D)	50/60/70
Duration (t/s)	240/600/6000
Angle cosine (K)	0. 1/0. 5/0. 8/1

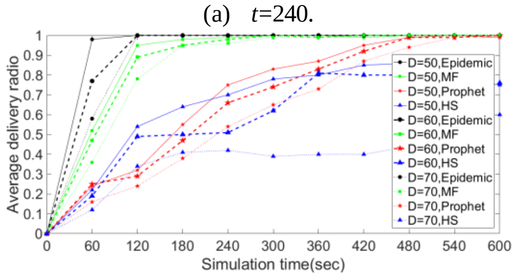
The metrics evaluated in this simulation are the average delivery rate and the average transmission time. The average delivery rate is the ratio of the number of successful deliveries to the total number of messages. The average transmission time is the delivery time of the first copy to its destination.

5.2 Simulation Results

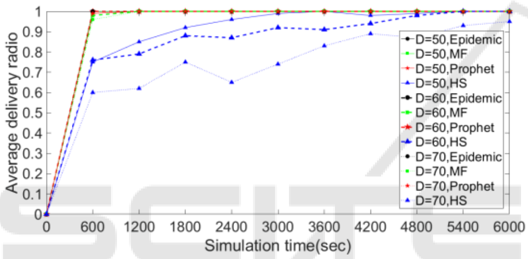
We conduct three sets of simulation experiments to evaluate the impact of various parameters on performance. In the first set of simulations, we adjusted the community length and width D of the four algorithms to 50, 60, and 70, and set N=2000, θ=0. 5. In the second set of simulations, we adjust the number of nodes N of the four algorithms to 1000, 2000, and 3000, and set D=50, θ=0. 5. In order to observe the difference between the algorithms more specifically, in these two sets of experiments, we will observe the trend of each algorithm under the condition of t=240, t=600, t=6000, as shown in Figure 2 and Figure 3.



(a) *t*=240.



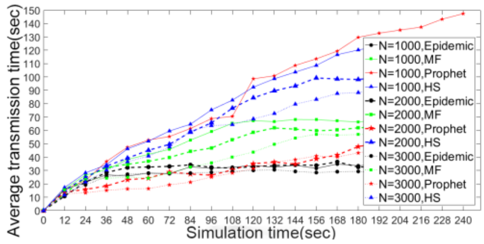
(b) *t*=600.



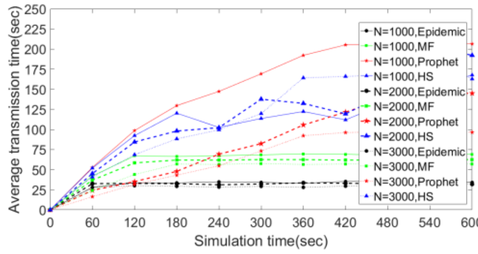
(c) *t*=6000.

Figure 2: Comparison of average delivery ratio.

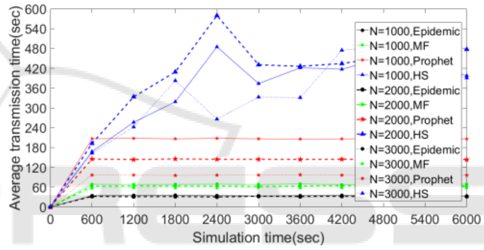
Fig. 2 shows that the average delivery rate of the four algorithms decreases as the community length and width increase. In contrast, the HS algorithm has the worst average delivery rate at D=70. Prophet's probabilistic selection causes the message to spread continuously close to the destination, but since Prophet only relies on nodes to forward messages, it has moderately lower performance. MMF is mainly due to the external forwarding function of boundary boxes, which enables the message to spread rapidly to other communities, so it performs better than the two algorithms above. Epidemic's message number and number of forwarding nodes are not limited, so the diffusion speed is very fast in a small scenario, but as the entire scene becomes larger, spreading messages requires more transmission time.



(a) *t*=240.



(b) *t*=600.



(c) *t*=6000.

Figure 3: Comparison of average transmission time.

Next, we also perform three sets of simulations to evaluate the performance of the above algorithm in terms of average transmission time. The results in Fig. 3 show that as the number of nodes increases, the average transmission time all decreases. The results show that at t=240 and t=600, the Prophet has the longest transmission time at n=1000. When t=6000, the transmission time of HS is more than the other three algorithms. Overall, the longer t is, the more unstable HS is. Because the role of the boundary box in the HS is only to spread the messages in the community, the spread of the messages between the communities only depends on the single copy of the node, so the number of nodes has a great impact on it and the transmission time is long. MMF uses the external forwarding function of boundary boxes and relatively few nodes, so that the messages can quickly spread to the other communities, therefore it has a shorter average transmission time among the three algorithms except Epidemic.

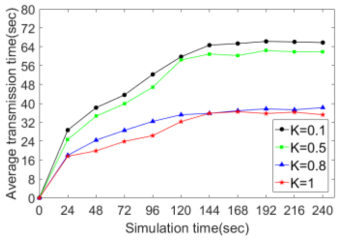


Figure 4: Effect of MMF angle cosine on average transmission time.

Next, we modify the angle cosine value K used by MMF in the roaming phase for four times. We set N=2000, D=50, t=240. As shown in Fig. 4, we compared the average transmission time at K=0.1, K=0.5, K=0.8, and K=1. The result shows that the closer K is to 1, the smaller the average transmission time is, the more similar it is to flooding, but this will cause massive copies, unnecessary energy consumption and useless bandwidth occupation.

6 CONCLUSIONS

In this paper, we study a special mobile social network, in which the running scenario includes some nodes, communities and boundary boxes, and propose a zero-knowledge multi-copy routing algorithm called MMF. MMF set a higher priority for boundary boxes to help spread information quickly. Theoretical analysis and simulation results show that boundary boxes play an important role in the message dissemination process. By using boundary boxes, MMF achieves better performance than several existing zero-knowledge MSN routing algorithms.

ACKNOWLEDGMENTS

This research is supported by National Natural Science Foundation of China under Grant Nos. 61872191, 41571389, 61872193.

REFERENCES

Hom J, Good L, Yang S. A survey of social-based routing protocols in Delay Tolerant Networks. *International Conference on Computing. Santa Clara: IEEE*, 2017: 788-792.

Xiao M, Wu J, Huang L. Homing spread: Community-home-based Multi-copy Routing in Mobile Social Networks. *IEEE Infocom.Turin: IEEE*, 2013: 2319-2327.

Vahdate A, Becker D. Epidemic routing for partially connected ad hoc networks. *Technique Report, CS-2000-06, Department of Computer Science, Duke University, Durham, NC*, 2000.

Daly E, Haahr M. Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs. *Acm Interational Symposium on Mobile Ad Hoc*, 2007.

Pan H, Crowcroft J, Yoneki E. Bubble rap: social-based forwarding in delay tolerant networks. *IEEE Transactions on Mobile Computing*, 2011, 10(11): 1576-1589.

Bulut E, Szymanski B. Exploiting Friendship Relations for Efficient Routing in Mobile Social Networks. *IEEE Transactions on Parallel and Distributed Systems*, 2014, 23(12), 2254-2265.

Huijuan Z, Kai L. A Routing Mechanism Based on Social Networks and Betweenness Centrality in Delay Tolerant Networks. *International Journal of Computer Science & Information Technology*, 2015, 7(6), 107-116.

Kim CM, Han YH, Youn JS, Jeong YS. A Socially Aware Routing Based on Local Contact Information in Delay-Tolerant Networks. *The Scientific World Journal*, 2014.

Gondaliya N, Kathiriya D. Community detection using inter contact time and social characteristics based single copy routing in delay tolerant networks. *International Journal of Ad hoc, Sensor & Ubiquitous Computing*, 2016, 7(1), 21-35.

Gondaliya N, Kathiriya D. Community detection using inter contact time and social characteristics based single copy routing in delay tolerant networks. *International Journal of Ad hoc, Sensor & Ubiquitous Computing*, 2016, 7(1), 21-35.

Chen H, Lou W. Contact expectation based routing for delay tolerant networks. *Ad Hoc Networks*, 2015, 36(1), 244-257.

Singh AK, Pamula R. IRS: Incentive Based Routing Strategy for Socially Aware Delay Tolerant Networks. *2018 5th International Conference on Signal Processing and Integrated Networks*, 2018, 343-347.

Choksatid T, Narongkhachavana W, Sumet Prabhavat. An efficient spreading Epidemic Routing for Delay-Tolerant Network. *2016 13th IEEE Annual Consumer Communications & Networking Conference*, 2016, 473 – 476.

Igarashi Y, Miyazaki T. A DTN routing algorithm adopting the “Community” and “Centrality” parameters used in social networks. 2018 International Conference on Information Networking, 2018, 211 – 216