Accepted Manuscript

Intelligent distributed routing scheme based on social similarity for mobile social networks

Fang Xu, Qiong Xu, Zenggang Xiong, Nan Xiao, Yong Xie, Min Deng, Huibing Hao

PII: S0167-739X(19)30125-6
DOI: https://doi.org/10.1016/j.future.2019.02.033
Reference: FUTURE 4786

To appear in: *Future Generation Computer Systems*

Received date : 16 January 2019
Revised date : 9 February 2019
Accepted date : 18 February 2019

Please cite this article as: F. Xu, Q. Xu, Z. Xiong et al., Intelligent distributed routing scheme based on social similarity for mobile social networks, *Future Generation Computer Systems* (2019), https://doi.org/10.1016/j.future.2019.02.033

Intelligent Distributed Routing Scheme Based on Social Similarity for Mobile Social Networks

Fang Xu[a,c,*], Qiong Xu[a], Zenggang Xiong[a,*], Nan Xiao[a], Yong Xie[b], Min Dong[a], Huibing hao[a]

[a] School of Computer and Information Science, Hubei Engineering University, Hubei, China

[b] Department of Computer Technology and Application, Qinghai University, Qinghai, China

[c] School of Computer, Wuhan University, Wuhan, China

*Corresponding author: Fang Xu (e-mail: xf2012@whu.edu.cn).
Zenggang Xiong (e-mail: xzg@hbeu.edu.cn)

Abstract: In mobile social networks (MSNs), the social attributes of nodes are important factors driving the mobility of nodes. By studying the mobility of the daily activities of node carriers, an intelligent distributed routing algorithm based on social context information prediction was proposed. First, we study the data forwarding problem of mobile social networks from two aspects, the daily behavior of mobile nodes and the similarity of social attributes respectively. Then, our algorithm uses BP neural network to predict the encounter regularity of mobile nodes in terms of time and space dimensions. This information can provide a basis for routing decisions. Finally, a routing algorithm with predictive capability is designed in combination with synchronous delivery and asynchronous delivery. Simulation analysis and experimental results show that the proposed routing algorithm can effectively improve the message delivery ratio and reduce the network overhead.

Keywords: mobile social network; social similarity; prediction model; routing; social context

1. Introduction

Mobile social networks (MSNs) are one of the important development directions of mobile ad hoc networks. In the real world, there are a lot of phenomena such as node movement, frequent link interruption and network partition in MSNs. Therefore, there is often no end-to-end path between two nodes when they need to communicate. Especially the distributed MSNs systems, where no centralized infrastructure is available, can greatly benefit from efficient opportunistic networking [1]. The emergence of opportunistic routing is precisely to solve this problem [2]. It allows nodes deliver messages as the way of storing, carrying and forwarding. When there is no forwarding path to the destination node, the relay node stores the information in its own cache, and then waits for the opportunity to contact other nodes in order to send the message towards the destination node. Based on this forwarding mechanism, in the past few years, researchers have put a lot of effort into optimizing routing protocols and forwarding strategies [3-5].

There have been numerous approaches attempting to devise new socially-aware metrics, in order to increase the effectiveness and efficiency of the opportunistic routing algorithms [6]. Research shows that, in the routing protocol design, the context information of the network can effectively improve the performance of the routing protocol. The context information in the network includes the current network operating environment, the behavior of the device carrier and the information about the device carrier, such as the location of work, the time of visiting a location or seeing a user, and so on. Any information that helps to facilitate network routing decisions is considered context information. Using this context information, the routing protocol can easily find a suitable forwarding node and improve the forwarding efficiency of data packets. At present, some researchers studied the mobile social network composed of smart devices [7] carried by people. They believe that the mobile social network is a collection of social relationships, and social relations significantly affect the encounter mode between network nodes. Therefore, it is very important to collect and utilize context information in MSNs to optimize routing strategy [8-10].

In this paper, we propose an intelligent distributed routing scheme based on social similarity for MSNs, called Similarity-Aware Intelligent Routing (SAIR). The routing algorithm uses BP neural network (Back-Propagation Neural Networks, BNN) model to predict contact information between nodes, and then use prediction information to make routing decisions. According to the current context information of the node and the destination node, the forwarding node starts the message forwarding process at an appropriate time and place. This method can reduce the delivery delay and network overhead.

The structure of this paper is organized as follows. Section 2 introduces the related research work. The system model and assumptions is presented in Section 3, including the definition of encounter probability and neural network prediction model. Section 4 introduces the Routing algorithm based on social similarity. We experimentally compare the performance of the routing proposed by us with other state-of-art routing protocols in Section 5. Finally, Section 6 concludes our work.

2. Related Work

The original opportunistic routing algorithm used a flood-based approach to forwarding data. It went from the initial blind flooding to control flooding later. The more famous algorithms in this type of algorithm are Epidemic routing [11] and network coding routing [12]. However, social relations significantly affect the encounter pattern between nodes in MSNs. Therefore, researchers are actively exploring the use of social context information in the network to optimize routing strategies.

The researchers proposed some community-based routing protocols. People with social ties, common interests, and similarities usually form a group. Compared with the social interaction between groups, the social interaction among the members in the same group will be more frequent. In mobile social networks, we call these groups communities. First, the community detection algorithms are used to assign nodes to different communities. The social graph structure is then designed based on the detected community later. In order to forward data between different communities, active nodes between communities can be used as a bridge for data delivery. The premise of this approach is that the community must be established firstly. However, the construction of the community will incur some costs. These are some of the more well-known community-based routing protocols, such as LABEL [13], Bubble Rap [14], CAOR [15], SGBR [16], etc.

If a network does not have obvious community characteristics, community detection not only consumes valuable network resources but also cannot achieve the desired effect. As a result, researchers have proposed routing protocols that do not require community support, such as SimBet [17], PRoPHET [18], FairRoute [19], PeopleRank [20] and LASS [21].

The social context-aware algorithm not only uses the context information related to node mobility, but also considers the social aspects of the node as an important parameter. In fact, in most cases, the moving characteristics of a node are determined by the behavior of the carrier which may be a person, an animal or a vehicle. Therefore, the social relationship of the carrier greatly affects the encounter of nodes in the network. The advantage of this method is that it is more adaptable to the real world. There are three typical social context-based routing algorithms, HiBOp [22] [51], Rubble Rap [14] and dLife [23] respectively.

Although researchers have proposed some context-based routing algorithms, these algorithms rarely take advantage of both time and space factors. Even if there is no neighbor node with a high probability of delivery around the sending node, the sending node still starts the process of data forwarding. If we solve this problem, it is possible to further improve the forwarding efficiency and reduce the network overhead.

3. System Model

Observing the mobile behavior in MSNs, it is found that the mobile mode of the node carrier is usually repetitive. Therefore, we can obtain the contact information between nodes by utilizing the social context information of carriers. It provides a new research idea for improving the performance of routing algorithm. The relay node carrying the information attempts to find the best relay node among the neighbor nodes in order to efficiently forward the data to the destination. It is more flexible when forwarding messages by predicting the time and space information when nodes meet.

3.1. Description of the problem

In a mobile social network, the success rate of data forwarding is not high enough, and the delivery overhead is large. Based on the current international research progress, we summarize the main problems faced by mobile social networks.

1. It is common to design routing protocols based on community characteristics. However, community detection and maintenance is a difficult problem, and requires the consumption of certain network and computing resources. Routing design without considering community information can avoid the cost of community detection and maintenance. This is especially suitable for network environments where the community characteristic isn't obvious.

2. When there is no end-to-end path in the network, how does the relay node effectively select the optimal carrier and the forwarding time?

3. The resources (energy, buffer space, and bandwidth) of the nodes in the mobile social networks are very limited. Although the multi-copy forwarding mode can improve the forwarding efficiency of data, it is a huge waste of valuable network resources by using the method of flooding or similar algorithms. Therefore, it is great significant to study the efficient data forwarding algorithm based on resource-saving single copy mode.

4. The accuracy of node movement prediction is a key problem in routing design. We can analyze the social attributes and mobile model characteristics of mobile nodes, and then make full use of the social context information of the carriers. In this way, the accuracy of encounter prediction can be improved and the forwarding efficiency of data can be further improved.

3.2. Models and Assumptions

To formalize the routing problem in mobile social networks, we propose some assumptions for the models we designed.

1. Mobile social networks are described as G (V, E). Nodes in a network are carried by people. G is an undirected unconnected graph. In the graph, V is the set of nodes and E is the set of links between nodes.

2. When nodes x and y are in communication range with each other, the communication links

between them are bidirectional. The connection state of a link varies with time. In other words, at different times the links may be connected or disconnected.

3. In the graph G, there is at least one cut point and one cut edge. If any cut edge or the cut point is deleted, then the network G will evolve into a disconnected graph composed of several subgraphs. We think of the single subgraph as a connected domain.

4. Each node in the network has a limited cache space, and it is the same size. Each node does not refuse to forward data for other nodes.

5. The nodes in the network are all in a dynamic moving state. During the movement, the nodes continuously meet other nodes, and the movement of the nodes shows a certain movement law for a long time.

6. If the carriers of the two nodes have similar social context attributes, there is a high probability of encounter between them.

3.3. Social similarity and encounter probability

The movement of nodes is closely related to the social activities of the node carriers. First we create a social attribute list (SAL) for each node to store relevant social context information. The appropriate weight is then set according to the importance of the context information. Finally, the prediction model of BP neural network is used to calculate the delivery probability.

The social attribute list contains information about the device carrier, such as name, address, work unit, hobbies, etc. The node's social attribute list consists of the social evidence and the corresponding value. The social attribute list plays an important role in calculating the social similarity between two nodes and predicting the movement of nodes. The nodes in the mobile social network are not completely random moving. They have repeated movements at different times, so the movements are predictable. If a node has visited a place several times before, it will probably also visit this place next time. Social network theory holds that the higher the similarity of social attributes between people, the greater the encounter probability. Therefore, the social similarity of carriers can be used to represent the encounter probability between nodes. The set of evidence and corresponding values in the social attribute list is denoted as $N(e,v)$. The social attribute list of destination node D is denoted as $D(e,v)$. The intersection of social attribute list between the node N and destination D is formalized as $M(e,v)$, which is expressed as formula (1).

$$M(e,v) = N(e,v) \bigcap D(e,v) \tag{1}$$

The encounter probability between network node N and network node D can be expressed by the matching degree of SAL between N and D. We calculate the matching degree by comparing the hash values of the corresponding attributes in the two SALs. The calculation method of the encounter probability between the two nodes is shown in formula (2). Where, W is the set of attribute weight W_m in $M(e,v)$ and W_D is the set of attribute weight W_d in $D(e,v)$.

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

The encounter probability of period refers to the probability that node S meets the destination D during the time period i, denoted as P_i. Assume Xi is a set of nodes encountered by the relay node S in the period i, and the probability of nodes in Xi meeting the destination D are higher than the probability of node S meeting the destination D. $|X_i|$ is the absolute value of X_i. The calculation method of the encounter probability between network node S and destination node D in period i is shown in formula (3).

$$P_i = \frac{\sum_{A \in X_i} F_{[A,D]}}{|X_i|} \quad (3)$$

In daily life, people's activities have periodic trend. In order to analyze this rule, the activities of nodes are divided into cycles (historical parameters) and periods (current parameters), as shown in Fig. 1. Device carriers are likely to repeat their activities day after day. When the source node needs to send messages, it first predicts the encounter probability in the next period, and then use this information to send the messages to the receiving node at the appropriate time. Therefore, based on the predicted information, the relay node knows when and where to send the message with a high success ratio.

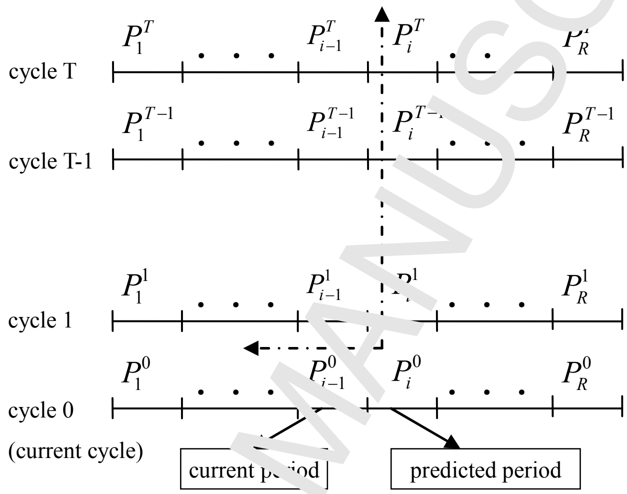


Fig. 1. Current and historical parameters

3.4. Prediction model

In order to calculate the encounter probability of the sending node with other nodes in the next period, it is necessary to know the contact history of the sending node with other nodes. As long as the sender obtains the historical encounter information of other nodes, it will be able to predict contact information in the next period.

In this paper, BP neural network is used to design the prediction model. In order to model, we need to introduce the following two parameters, the encounter probability in the current period (current input) and the historical encounter probability of the same period in the previous cycle (historical input). The node of MSNs calculates the encounter probability in each period and stores these values for the prediction of the next cycle.

3.4.1. Prediction process

As shown in Fig. 1, $P_i^j (j \in ([1, ..., T]), (i \in [1, ..., R])$ is the probability that the current node meets the destination node at period i in cycle j. T is the number of historical periods. R is the number of periods divided into each period. For example, the encounter probability at the period (i-1) of the cycle 0 can be expressed as P_{i-1}^0. Because mobile nodes have limited resources, simple solutions should be used to solve problems. Considering the practicality, the calculation of the historical input in time period i is the cumulative probability of two encounters, as shown in formula (4).

$$Pc_i^1 = \frac{P_i^1 + Pc_i^2}{2} \tag{4}$$

In the equation (4), Pc_i^2 is the cumulative value of the encounter probability at period i in the cycle 2. Pc_i^1 is the cumulative value of the encounter probability at period i in the cycle 1.

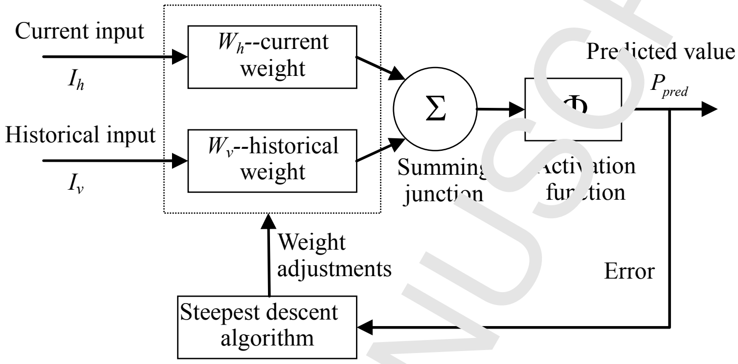


Fig. 2. Prediction process of BP neural network model

Fig. 2 shows the basic model based on BP neural network. There are two input parameters, the current input I_h and the historical input I_v. The current input I_h is formalized as $I_h = P_{i-1}^0$. The historical input I_v is the cumulative value of $P_i^j (j \in ([1, ..., T])$, formalized as $I_v = Pc_i^1$. Its calculation method is shown in formula (4). The synaptic weights of current input and historical input correspond to W_h and W_v, respectively. The summing junction function is described in formula (5).

$$F = I_v W_v + I_h W_h \tag{5}$$

The activation function is a normalized function whose output is between 0 and 1. The output value of the activation function is the encounter probability at the next period, and the encounter probability P_{pred} in the next period is as shown in formula (6).

$$P_{pred} = \Phi = \frac{F}{W_v + W_h} = \frac{I_v W_v + I_h W_h}{W_v + W_h} \tag{6}$$

3.4.2. Error calculation

The BP neural network algorithm uses the gradient descent method to find the minimum value of the error function. The error function is used to calculate the mean square error between the true value and the expected value for a given sample.

In this paper, the expected output is the predicted output value. The error is calculated in formula (7).

$$E = \frac{(P_{pred} - P_{actual})^2}{2} \tag{7}$$

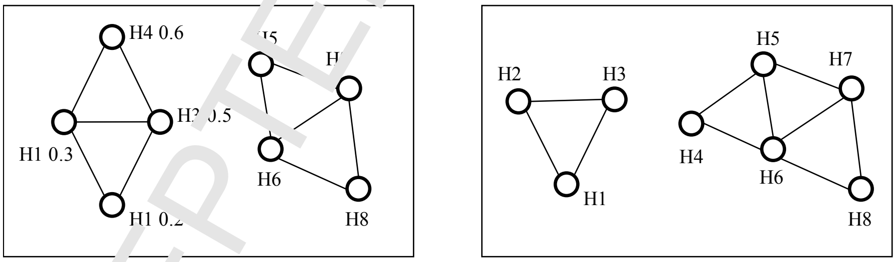
In above equation, P_{pred} is the predicted value of the period, and P_{actual} refers to the true value of the period.

4. Routing algorithm

Mobile social networks are often divided into interconnected domains. As the nodes are in a moving state, the nodes in the connected domain are constantly changing. Our proposed routing algorithm adopts two methods for data forwarding, namely synchronous forwarding and asynchronous forwarding. When the receiving node and the sending node are in the same connected domain, there is an end-to-end path between the two nodes. At this point, the sending node forwards the data packet directly to the receiving node in a synchronous manner. On the contrary, when two nodes are not in the same connected domain, the asynchronous forwarding such as the opportunistic forwarding is adopted. In asynchronous forwarding, firstly, the node with the highest delivery probability to the destination is searched in the connected domain. This node is selected as the relay node. The sending node then sends the message to the relay node by synchronous forwarding. The relay node looks for the appropriate opportunity to send the message to the next relay node by asynchronous forwarding until the message is received by the destination node. The delivery probability mentioned in this paper refers to the probability that a node meets the destination node in a certain period.

4.1. Algorithm summary

The schematic diagram of message forwarding in the MSNs is shown in Fig. 3. There are two connected domains. In Fig. 3 (a), the node H1 needs to send a message *M* to the node H8. However, at this time, there is no end-to-end path between two nodes, so synchronous forwarding cannot be adopted. Fig. 3 shows the delivery probability of each node to H8. At this time, in the connected domain where H1 is located, the node with the highest delivery probability to H8 is H4. Therefore, H1 uses the synchronous forwarding to send *M* directly to node H4. H4 adds *M* to its cache. After some time, H4 moves to another connected domain, as shown in Fig. 3 (b). At this time, H4 and H8 are in the same connected domain, and then the synchronous forwarding is used. H4 immediately sends the message to H8, and the final message *M* is delivered to the destination.



(a) time t-1 (b) time t

Fig. 3. Process of message forwarding

In this paper, an improved DSDV [24] routing algorithm (I-DSDV) is adopted for synchronous forwarding in the connected domain. The original DSDV routing table item is extended to include the best carrier ID (bestCarrier) and corresponding delivery probability (deliveryProb), as shown in Fig. 4. The prediction module in each node periodically calculates the delivery probability of that node to other nodes, and then stores it in the delivery probability table. Each node periodically broadcasts routing information and delivery probability tables to its neighbors.

targetId	nextId	dist	bestCarrier	deliveryProb

Fig. 4. Key fields of routing table

This algorithm makes full use of node mobility, network topology changes and social attributes of nodes. According to whether two nodes need to communicate are in the same connected domain, different data forwarding mode is selected to improve the efficiency of forwarding.

4.2. Route Table Building

I-DSDV algorithm is needed for synchronous forwarding in the connected domain. Therefore, the routing table will be established. To facilitate the description of the algorithm, we have defined some symbols. The key symbols used in the algorithms are described as Table 1.

Table 1

Key symbols used in the algorithms

Symbol	Description
BPPM	BP neural network prediction model
$P_{[N,D]}$	The probability of encounter between node N and node D
L_m	Neighbor set of m
RA	Router Advertisements
T_{period}	The time period of prediction model
ΔT	The interval between the current time and time of the latest prediction (or initialize routing table)
d	destination node of MSG
MSG	The message waiting to be sent
CD	The same connected domain in MSNs

Algorithm 1 describes the pseudo code of the routing table building. The prediction module BPPM is deployed on any node *m* in the network. The input parameters of the algorithm are: $P_{[N,D]}$, L_m, and *RA*. First, the routing table is initialized. Then, the algorithm enters the loop. Node m gets the neighborhood set L_m at this time. In each period T_{period}, the BPPM module runs once and stores the calculation results in the delivery probability table (DPT). At the same time, node *m* updates the corresponding items in the routing table and broadcasts *RA* to the neighbors. When *m* receives the *RA* from a neighbor, it updates its routing table and related parameters. If the time interval exceeds the period T_{period}, node *m* still does not receive *RA* from neighbor *n*, then node *n* is considered unreachable.

Algorithm 1 Route Table Building
(Node *m* is a node where BPPM are deployed.)

Input: $P_{[N,D]}$, L_m, *RA*
initialize the routing table
Repeat
  Obtain the neighborhood set L_m of *m*
  If $\Delta T = T_{period}$ then
    Execute BPPM($P_{[N,D]}$)
    Return *delivery_probability_table=DPT*
    Update *routing_table=R*
    Broadcast *RA* to *n* ($n \in L_m$)
  End if
  While *m* receive *RA* of *n* ($n \in L_m$) do
      Update *R* of *m*
      Update *bestCarrier* and *deliveryProb*
  End while
  If $\Delta T > T_{period}$ *m* don't receive the *RA* from *n* ($n \in L_m$)
    then
      nextId←null
      dist←16

```
  End if
Until {m shutdown or routing disable}
```

4.3. Message fowarding

Algorithm 2 describes the process of node *m* forwarding messages in the mobile social network. This algorithm explains in detail how to choose synchronous and asynchronous forwarding under different circumstances. This algorithm will be started when there are messages needing to be forwarded. The input parameters of the algorithm are: *d*, *MSG*. If both source and destination are in the same connected domain, node *m* adopts the method of synchronous forwarding. Otherwise, the message is sent to the relay node with the highest delivery probability in the same connected domain, and then the asynchronous forwarding is started. The relay node stores and carries the messages, sending it to the next relay node at the appropriate time and place until the message finally reaches its destination.

Algorithm 2 Message Fowarding
(Node *m* is a node in MSNs.)

```
Input: d, MSG
  If d=m then
    Node m receive and process MSG
    Else if m∈CD and d∈CD then
      Deliver MSG using I_DSDV synchronously
      Else
        Relay_node={m'∈CD | max(deliveryProb}
        Deliver MSG to relay_node
        relay_node forward MSG asynchronously at right
time and place
      End if
  End if
```

5. Performance evaluation

In order to evaluate the performance of the SAIR routing, we adopted the widely used network simulation platform ONE [25]. We compare the SAIR route with the other three classic routes, namely Epidemic [11], PRoPHET [18] and dLife [23]. Then the experimental results are analyzed in detail. The simulation experiment uses the following four statistics: delivery ratio, overhead ratio, average delay, and average hop count.

5.1. Simulation environment

In this paper, we abstract the nodes' mobile behavior at the network layer, mainly focusing on the encounter information and connectivity between nodes. We ignored issues with radio reception and the MAC layer, such as packet loss due to information interference.

The simulation experiment is based on the following assumptions. When two nodes are within the communication range of each other, the message can be forwarded. The constraints of battery power and storage space in nodes are not considered. When a node sends messages to another node, it can receive it correctly without losing packet. The network uses a social-based mobile model (CMM) [26], and the experiment simulates the social contact situation for 10 days. The initial values of current weight and historical weight are set to 0.5 respectively. The learning rate is set to 0.5 and the number of iterations is 550. In the simulation process, the nodes sending message were randomly selects, and the final simulation result is the average of 20 simulation results. The other parameters in simulation processes are shown in Table 2.

Table 2
Detailed simulation parameters

parameter	value	unit
Topological area	3000 × 2500	m^2
Mobility model	CMM	-
Number of nodes	50～100	-
Number of connected domain	15	-
Communication mode	WiFi	-
Transmission radius	40	m
Cache space	6～22	MB
Message size	512	Byte
Traffic model	Random pairs	-
Packet transmission rate	4	packet/s

5.2. Prediction accuracy

The BP neural network model has the self-learning ability. The process of information forward propagation and error back propagation can adjust the weights and thresholds of the network continuously. After several cycles, the predicted value is close to the expected value. Fig. 5 shows the error ratio for prediction using BP neural network model. In the first four cycles, there was a large gap between the predicted value and the expected value due to insufficient data samples. With the increase of cycle number, BP neural network model gradually adjusts the weight through self-learning and back propagation of error, so that the error of prediction gradually decreases. After the fourth cycle, the prediction error stabilized at about 5%. Given the inherent nature of mobile social networks, this level of error is acceptable.

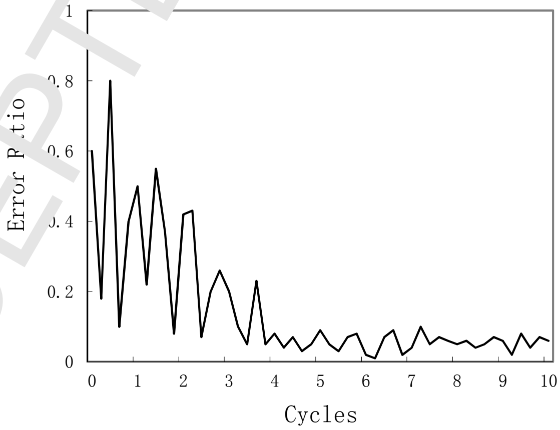


Fig. 5. Error ratio of the prediction

5.3. Impact of the node number

Obviously, after giving network area and speed of node movement, the total number of nodes' encountering is directly related to the number of nodes in the network. This section mainly verifies the impact of changes in the number of nodes on message forwarding. The evaluation is mainly based on the following four aspects: delivery ratio, overhead ratio, average delay, and average hop count. Here the cache for each node is set to 12MB.

The delivery ratio of the four routing at different node numbers is shown in Fig. 6. The results show that the network delivery ratio increases as the number of nodes increases. The delivery ratio of SAIR is on average 14% higher than that of Epidemic, 12% higher than that of ProPHET and 5% higher than that of dLife.

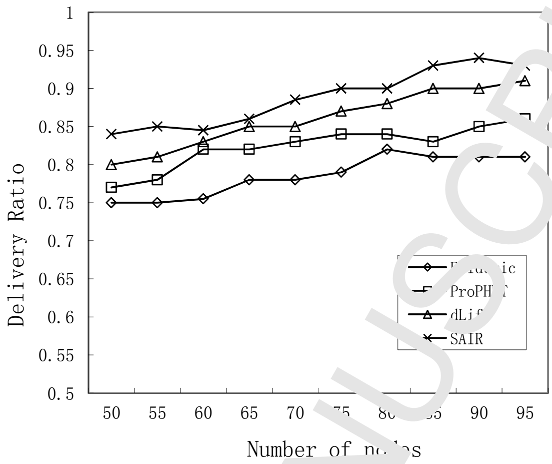


Fig. 6. Delivery ratios vs. node number

The effect of the node numbers on the network overhead ratio is shown in Fig. 7. As the number of nodes increases, the overhead ratio of the four types of routing all appear the increasing tendency. Epidemic's overhead ratio is much higher than other routing algorithms, and dLife and SAIR have similar overhead ratio. The difference in overhead ratio between dLife and SAIR is small. Compared with Epidemic, ProPHET, and dLife, the overhead ratio of SAIR was reduced respectively by an average of 71%, 39%, and 16%. It is shown in Fig. 7 that SAIR has the lowest network overhead in the same network scenario.

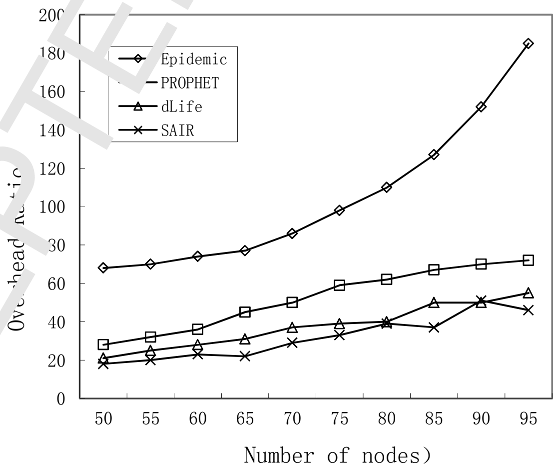


Fig. 7. Overhead ratios vs. node number

The influence of the number of nodes on the average delay for data forwarding is described in Fig. 8. The average delay of the four algorithms decreases rapidly with the increase of the number of nodes. Concerning the Epidemic algorithm, the average delay is the smallest, while the ProPHET algorithm has the largest. DLife and SAIR have similar average delay. The average delay of SAIR was 22%

higher than the Epidemic. However, the average delay of SAIR was respectively 16% and 5% lower than the ProPHET and dLife.

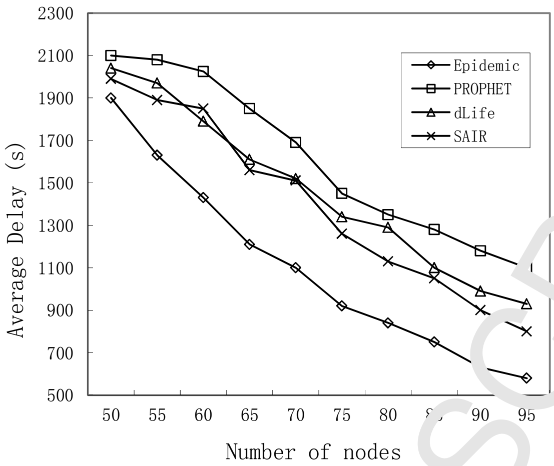


Fig. 8. Average delays vs. node number

The average hop count of each algorithm varies with the number of nodes as shown in Fig. 9. According to the figure, the Epidemic algorithm has the highest average hop count because it employs a flooding method. The average hop count of ProPHET is also at a high level. When the number of nodes is 95, SAIR is 0.1 higher than dLife. When the number of nodes is 95, SAIR is 0.1 higher than dLife, while it is 4.5 and 1.7 lower than Epidemic and ProPHET respectively.

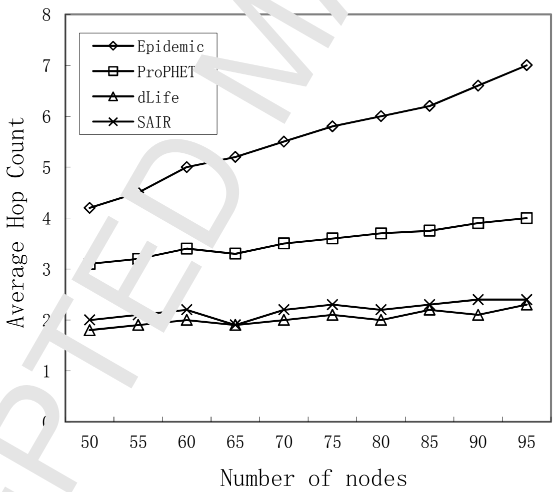


Fig. 9. Average hop count vs. node number

Next, the simulation results are analyzed. Epidemic forward messages using flood methods without taking advantage of any context information. Although the minimum average delay can be obtained, it is a great waste of limited network resources. ProPHET fails to consider the social mobility mode of the node carrier, and overestimates the link between nodes, resulting in the increase of data forwarding times. This increases overhead ratio of the network and affects the delivery ratio. Like SAIR, dLife makes use of the laws of carriers' daily activities and social relations to help improve routing efficiency. However, by using the BP neural network model to predict the future movement of nodes, SAIR has achieved better performance. In particular, when the number of nodes increases, SAIR uses the prediction information to start the forwarding process at an appropriate time and place, which can reduce the network overhead and improve the delivery ratio.

5.4. Impact of the memory capacity

The nodes of the mobile social network are generally resource-constrained nodes, and the cache capacity of the nodes is also very limited. The cache capacity of a node determines the amount of messages it can carry. Therefore, the size of cache space has a great impact on the performance of routing algorithm. Then, through simulation experiments, we analyze the impact of cache capacity on the performance of routing algorithms. The number of nodes in the MSNs is set as 60 here.

Based on different cache Spaces, the delivery ratio of the four routing is shown in Fig. 10. With the increase of node cache space, the delivery ratio of the four routing algorithms increases. When the node cache space is small, dLife and SAIR have a big advantage. However, the performance growth rate decreases as the cache space increases. When the size of cache space is 18MB, Epidemic has a higher delivery ratio than ProPHET.

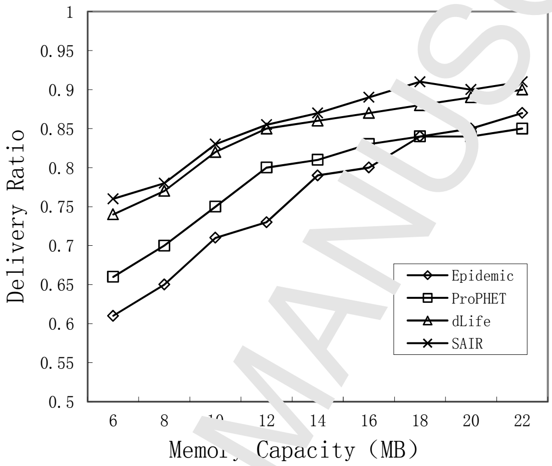


Fig. 10. Delivery ratio vs. memory capacity

The effect of the size of the cache space on the network overhead ratio is shown in Fig. 11. With an increase in the size of the cache space, the Epidemic overhead ratio exhibits a rapidly decreasing trend. The change trend of the other three algorithms is relatively gentle. SAIR has the best overhead ratio.

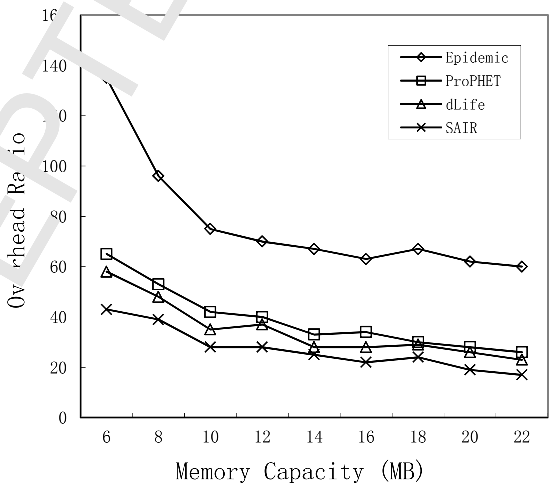


Fig. 11. Overhead ratio vs. memory capacity

Fig. 12 shows the impact of cache space changes on the average latency of message delivery. According to the comparison results, with the increase of node cache space, the average delay of Epidemic falls rapidly, while the average delay of SAIR is stable. When the cache space is less than 8MB, the average delay of Epidemic is larger than that of the other three algorithms. With an increase in the cache space, when it is larger than 8MB, the average delay of Epidemic rapidly becomes smaller than the other three algorithms.

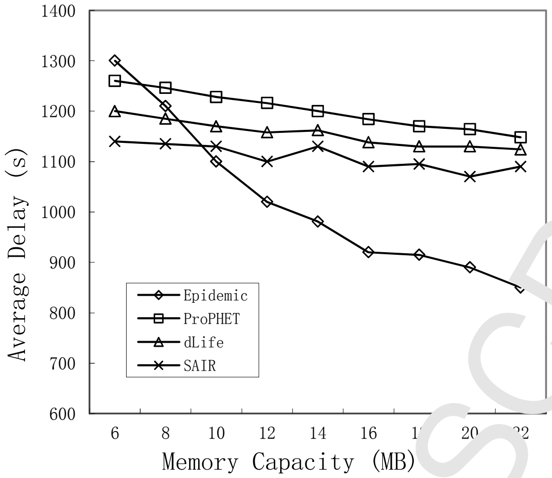


Fig. 12. Average delay vs. memory capacity

The average hop count of each algorithm vary with the cache space, as shown in Fig. 13. According to the figure, the Epidemic has the highest average hop count because of flooding method. In general, regarding the Epidemic and ProPHET, the average hop count are decreasing with the increase of the cache space. This phenomenon indicates that the two algorithms have a greater dependence on the cache space of the nodes, and that they consume a larger cache space. When the cache space changes, the average hop count of dLife and SAIR change slightly. This phenomenon indicates that the two algorithms mainly use the encounter prediction information to forward data. The number of replicas of messages in the network is relatively small, thus saving the resources of cache space in node. As show in the Fig. 13, the average hop count of dLife is the best. Our proposed SAIR is slightly worse than dLife, while the performance of Epidemic and ProPHET is poor.

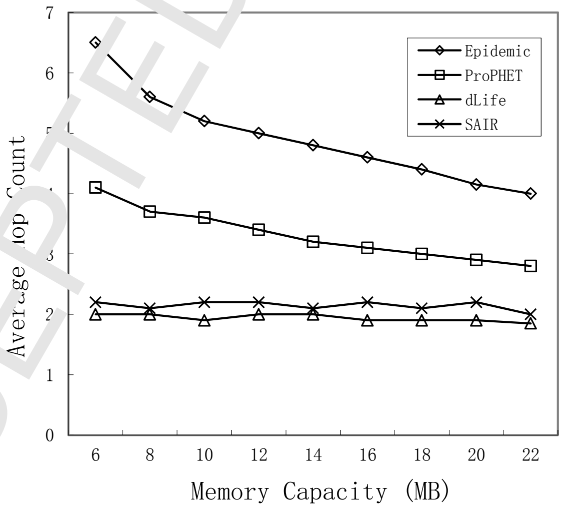


Fig. 13. Average hop count vs. memory capacity

From the above analysis, it can be seen that when the cache space increases, more messages can be carried by nodes, and more messages can be exchanged when nodes meet. Therefore, it is beneficial to improve the performance of message forwarding, such as delivery ratio and average delay. SAIR considers not only the context information, but also the synchronous forwarding within the connected domain and the asynchronous forwarding between the connected domains. In general, the performance of SAIR algorithm achieves the desired effect under the single copy mode with limited network resources.

6. Conclusion

We propose an intelligent distributed routing algorithm based on social similarity by studying the influence of social activities of node carriers on the encounter pattern between nodes. This algorithm can use social context information in the network to predict the mobile behavior of network nodes through the BP neural network. The routing decision process takes full account of the time and space attributes of mobile nodes. When the receiving node and the sending node are in the same connected domain at the same time, the message forwarding adopts synchronous mode, otherwise adopts asynchronous mode. Finally, through the simulation experiment, we compare and analyze the routing algorithm with the existing famous algorithm. Our algorithm can improve the ability of network to adapt to topology change. It has the characteristics of distribution, adaptive and intelligent optimization. Next, we will study the incentives of selfish nodes and routing security based on this algorithm.

Acknowledgments

This research was supported by the MOE (Ministry of Education of China) Project of Humanities and Social Sciences (17YJCZH203), the Science Research Fund of Hubei provincial education office (D20182702), the National Natural Science Foundation of China (No.61502154, 61370092), Hubei Provincial Department of Education Outstanding Youth Scientific Innovation Team Support Foundation (T201410).

References

[1] Vastardis N, Yang K, Leng S. Socially aware multi-phase opportunistic routing for distributed mobile social networks[J]. Wireless personal communications, 2014, 79(2): 1343-1368.

[2] Jing Chen, Kun He, Ruiying Du, Yang Xiang. Dominating Set and Network Coding-based Routing in Wireless Mesh Networks. IEEE Transactions on Parallel and Distributed Systems. 2015,26(2):423-433

[3] Araniti G, Orsino A, Militano L, et al. Context-aware Information Diffusion for Alerting Messages in 5G Mobile Social Networks[J]. IEEE Internet of Things Journal, 2017, 4(2): 427-436.

[4] Wu D, Zhang F, Wang H, et al. Security-oriented opportunistic data forwarding in mobile social networks[J]. Future Generation Computer Systems, 2018, 87: 803-815.

[5] Tao J, Wu H, Shi S, et al. Contacts-aware opportunistic forwarding in mobile social networks: A community perspective[C]//Wireless Communications and Networking Conference (WCNC), 2018 IEEE. IEEE, 2018: 1-6.

[6] Lu Z, Wen Y, Zhang W, et al. Towards information diffusion in mobile social networks[J]. IEEE Transactions on Mobile Computing, 2016, 15(5): 1292-1304.

[7] Jing Chen, Chiheng Wang, Ziming Zhao, Kai Chen, Ruiying Du, Gail-Joon Ahn. Uncovering the Face of Android Ransomware: Characterization and Real-time Detection. IEEE Transactions on Information Forensic & Security. 2018, 13(5): 1286- 1300

[8] Rahim A, Kong X, Xia F, et al. Vehicular social networks: A survey[J]. Pervasive and Mobile Computing, 2018, 43: 96-113.

[9] Ahmed M, Li Y, Waqas M, et al. A Survey on Socially-Aware Device-to-Device Communications[J]. IEEE Communications Surveys & Tutorials, 2018.

[10] Hu X, Chu T H S, Leung V C M, et al. A survey on mobile social networks: Applications, platforms, system architectures, and future research directions[J]. IEEE Communications Surveys & Tutorials, 2015, 17(3): 1557-1581.

[11] A. Vahdat and D. Becker. Epidemic routing for partially connected ad hoc networks[R]. Technical

Report CS-200006, Duke University, 2000.
[12] Wu Y, Chou P A, Kung S Y. Minimum-energy multicast in mobile ad hoc networks using network coding[J]. IEEE Transactions on communications, 2005, 53(11): 1906-1918.
[13] Hui P, Crowcroft J. How small labels create big improvements[C]//Pervasive Computing and Communications Workshops, 2007. PerCom Workshops' 07. Fifth Annual IEEE International Conference on. IEEE, 2007: 65-70.
[14] Hui P, Crowcroft J, Yoneki E. Bubble rap: Social-based forwarding in delay-tolerant networks[J]. Mobile Computing, IEEE Transactions on, 2011, 10(11): 1576-1589.
[15] Xiao M, Wu J, Huang L. Community-aware opportunistic routing in mobile social networks[J]. Computers, IEEE Transactions on, 2014, 63(7): 1682-1695.
[16] Abdelkader T, Naik K, Nayak A, et al. SGBR: A routing protocol for delay tolerant networks using social grouping[J]. Parallel and Distributed Systems, IEEE Transactions on, 2013, 24(12): 2472-2481.
[17] Daly E M, Haahr M. Social network analysis for routing in disconnected delay-tolerant manets[C]//Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing. ACM, 2007: 32-40.
[18] Anshul Verma, K.K.Pattanaik,et al. Routing in Opportunistic Networks[M]. New York: Springer, 2013: 69-97.
[19] Pujol J M, Toledo A L, Rodriguez P. Fair routing in delay tolerant networks[C]//INFOCOM 2009, IEEE. IEEE, 2009: 837-845.
[20] Mitibaa A, May M, Ammar M, et al. Peoplerank, combining social and contact information for opportunistc forwarding[C]//World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE International Symposium, San Francisco CA, 2012: 1-5.
[21] Li Z, Wang C, Yang S, et al. Lass: Local-activity and social-similarity based data forwarding in mobile social networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(1): 174-184.
[22] Boldrini C, Conti M, Jacopini J, et al. Hibop: a history based routing protocol for opportunistic networks[C]//World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007. IEEE International Symposium on a. IEEE, 2007: 1-12.
[23] Moreira W, Mendes P, Sargento S. Opportunistic routing based on daily routines[C]//World of wireless, mobile and multimedia networks (WoWMoM), 2012 IEEE international symposium on a. IEEE, 2012: 1-6.
[24] Singh K, Verma A K. Experimental analysis of AODV, DSDV and OLSR routing protocol for flying adhoc networks (FANETs)[C]//Electrical, Computer and Communication Technologies (ICECCT), 2015 IEEE International Conference on. IEEE, 2015: 1-4.
[25] Keranen A. Opportunistic network environment simulator[J]. Special Assignment report, Helsinki University of Technology, Department of Communications and Networking, 2008.
[26] Musolesi M, Mascolo C. Designing mobility models based on social network theory[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2007, 11(3): 59-70.

This research was supported by the MOE (Ministry of Education of China) Project of Humanities and Social Sciences (17YJCZH203), the Science Research Fund of Hubei provincial education office (D20182702), the National Natural Science Foundation of China (No.61502154, 61370092), Hubei Provincial Department of Education Outstanding Youth Scientific Innovation Team Support Foundation (T201410).

Fang Xu received the M.S. and Ph.D. degrees from Wuhan University, Wuhan, China, in 2009 and 2016, respectively. He is an Associate Professor in the School of Computer and Information Science, Hubei Engineering University, Hubei, China. His research interests include social computing, wireless mobile networks, and context aware computing. He is a member of the IEEE Computer

Qiong Xu was born in 1991. She received the B.S. degree in computer science and technology from Hubei University of China，in 2013，and she is currently pursuing the M.S. degree in computer technology with Hubei University of China. She is involved in research on routing algorithms for mobile social networks

Zenggang Xiong received the M.S degree in computer application from Hubei University in 2005 and the Ph.D. degree in computer application from University of Science and Technology Beijng in 2009 respectively. He is a professor in computer science at Hubei Engineering University. His research interest includes cloud computing and big data.

Nan Xiao was born in 1995. She received the B.S. degree in computer science and technology from Hubei Engineeing University of China，in 2018，and she is currently pursuing the M.S. degree in computer technology with Hubei University of China. She is involved in research on incentives for selfish nodes in Mobile social network.

Yong Xie received Ph.D. degree incomputer science from Wuhan University, Wuhan, China, in 2016. He is currentlya Associate Professor with the Department of Computer Technology andApplication, Qinghai University. His current research interests include nextgeneration Internet, network protocol and protocol security.

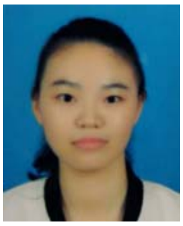
Min Deng received the M.S. degrees from Wuhan University, Wuhan, China, in 2010. She is a lecturer in the School of Computer and Information Science, Hubei Engineering University, Xiaogan, China. Her research interests include wireless mobile networks, and context aware computing.

Huibing Hao obtained his PhD degree in Mechanical Engineering from Southeast University in 2015. His research interests include reliability evaluation and maintenance decision. He is currently an assistant professor at the Department of Mathematics at Hubei Engineering University, Xiaogan, China. His research interests include big data, reliability prediction, quality management, fault diagnosis.

Fang Xu

Qiong Xu

Zenggang Xiong



Nan Xiao

Yong Xie

Min Deng


Huibing Hao

Highlights

- Social Similarity is used to predict the probability of encounter.

- A hybrid message forwarding is adopted combining synchronous forwarding and asynchronous forwarding.

- An Intelligent Distributed Routing is proposed to improve message delivery ratio.