Activity-based routing algorithm in opportunistic mobile social networks

Sheng Zhang[1] ⓘ**, Houzhong Liu[1], Caisen Chen[2], Zhaojun Shi[1] and William Wei Song[3]**

Abstract

In opportunistic mobile social networks, nodes are clustered according to their interests or hobbies and take part in different activities regularly. We delve into the temporal and spatial mobility characteristics of network nodes and put forward an activity-based message opportunistic forwarding algorithm. The main idea of the algorithm is that we choose different message forwarding methods according to the situation of nodes participating in activities. If the source node and the destination node are both attend in the same activities, we select the best relay node which has the biggest delivery probability. While the source node and the destination node are not in the same activities at the same time, we need to find the optimal path which owns highest indirect delivery probability, and messages will be transmitted through the optimal path. The simulation results show that the proposed routing algorithm can not only improve the successful delivery ratio of messages but also reduce the network delay and the network overhead obviously, in comparison with the classical opportunistic routing algorithms, such as community-aware message opportunistic transmission algorithm, community-based message transmission scheme algorithm, PRoPHET, Epidemic algorithm, and interest characteristic probability prediction algorithm.

Keywords

Opportunistic mobile social network, temporal and spatial characteristics, activity-based, message forwarding, delivery probability, indirect delivery probability

Introduction

The opportunistic mobile social networks (OMSNs) is a form of mobile self-organizing networks that exploit the human social characteristics, such as similarities, daily routines, mobility patterns, and interests to perform the message routing and data sharing.[1] In such networks, the users with mobile devices are able to form on-the-fly social networks to communicate with each other via Bluetooth or WiFi while within each other's transmission range and share data objects.[2]

The OMSNs occur not only in real communities but also in virtual communities. There are many OMSNs in real world, such as the opportunistic computing,[3] the vehicular ad hoc network,[4,5] the pocket switched network,[6] and the mobile wireless sensor networks.[7] The OMSNs in virtual communities include Facebook, Twitter, WeChat, QQ, and Instagram.

[1]School of Information Engineering, Nanchang Hangkong University, Nanchang, China
[2]Military Exercise and Training Center, Academy of Army Armored Force, Beijing, China
[3]Business Intelligence and Informatics, Dalarna University, Borlänge, Sweden

Corresponding author:
Sheng Zhang, School of Information Engineering, Nanchang Hangkong University, 696 Fenghe South Avenue, Nanchang 330063, Jiangxi, China.
Email: Zwxzs168@126.com

In the era of fifth-generation (5G) communication, the existence of massive communication data puts forward higher requirements for the rapid and accurate transmission of messages. In recent years, the OMSNs emerged as a new mechanism of communications in wireless networks. Unlike mobile ad hoc networks that require end-to-end communication paths for message exchange, the communication in OMSNs takes place on the establishment of opportunistic contacts among mobile nodes, without the availability of end-to-end message routing paths. As the mobile devices can make contact only when humans come into contact, such networks are tightly coupled with human social networks. Therefore, the OMSNs exploit the human behaviors and social relationships to build more efficient and trustworthy message dissemination schemes.[8] Therefore, it is not possible to run the traditional routing protocols in the OMSNs.

The OMSNs only depends on the encounter opportunity among nodes to forward messages. Consequently, the "Storage-Carry-Forward" strategy is usually used to deliver messages in the OMSNs. As the nodes generally tend to a group according to social relations, and the nodes are much active in the same group, while they hardly move to other group, we call it with the "spatial characteristics" of the OMSNs. Based on the same interest in the OMSNs, some nodes usually visit the specific group regularly. We call it with the "temporal characteristics" of the OMSNs. There are just a few nodes which can visit other groups according to their preference, they are likely to set up ties between different groups. In this case, how to forward messages efficiently from the source node to destination node is challenging. So we try to propose a new strategy to forward messages in the OMSNs, using fully the temporal and spatial characteristics of nodes mobility. In some special occasions, such as lack of communication infrastructure, bad network environment, and emergency response, the emergence of new message forwarding strategy can better deal with the problems that existing wireless network technology is difficult to solve, such as network splitting and delay.

Literature survey

The OMSNs is a special kind of opportunistic networks. According to the forwarding strategy and node characteristics, opportunistic network routing algorithms can be divided into the following three categories.

Routing algorithm based on opportunistic forwarding

The routing algorithm based on opportunistic forwarding is a common routing algorithm in opportunistic networks. When two nodes meet, the node X carrying the data packet forwards the data packet to another node Y, and the node Y may only be an intermediate node. Moreover, node X does not make a decision on whether it is the best path to forward the data packet through the Y forwarding node. Typical algorithms include Epidemic,[9] Spray and Wait,[10] and Spray and Focus.[11] Based on the Epidemic algorithm, Chancay-García et al.[12] considered that the amount of transmitted information is limited by the transmission speed and contact time and proposed an epidemic extension protocol called Xpread. Similarly, on the basis of the Epidemic algorithm, Chunyue et al.[13] proposed an asynchronous sleep mechanism that allows nodes to determine whether to enter the sleep state based on their own energy status, real-time status, and other information. It dynamically controls the sleep time and wake-up time to ensure that nodes will not miss possible communication opportunities. In order to improve the delivery rate of the spray and wait algorithm, Derakhshanfard[14] proposed an Erlang-based buffer management and routing algorithm, which considered the probability of message delivery and buffer when selecting relay nodes and allocating tokens sent to nodes status and message delivery time. Also, taking into account the difference in node forwarding capabilities, Wu et al.[15] proposed a spray strategy based on social circles. Under the premise of not fixing the initial number of copies, they selectively spray messages multiple times during the waiting phase to determine the appropriate number of redundant message copies. These methods all establish routing paths through encounters between nodes. They adopt a "storage-carry-forward" strategy to forward messages in OMSNs. In order to increase the successful delivery ratio of messages and reduce network delays, researchers often increase the number of copies of messages, meanwhile it brings huge communication burden in OMSNs.

Prediction-based routing algorithm

The prediction-based routing algorithm attempts to predict the best forwarding node through historical information, thereby increasing the successful delivery ratio of data packet transmission, thereby improving forwarding efficiency and reducing resource consumption. Typical routing algorithms based on predictive information include PRoPHET[16,17] and MaxProp.[18] Some scholars have improved on these typical methods. Wu et al.[19] established an interest characteristic probability prediction (ICPP) method based on the user's selection or promotion of feature interest with neighbors, selecting the node's interest neighbors and receiving or sending information. In order to improve the quality of message delivery in sparse opportunistic networks, based on the fusion of the historical encounter probability and the time series, Xu et al.[20] proposed a

routing algorithm based on node intimacy. The node intimacy describes the effective encounters of nodes, can predict the encounter probability of nodes in future time slices. Srinidhi et al.[21] proposed an optimized multi-copy routing algorithm based on an improved PRoPHET-random forest. The algorithm uses past and current context information to classify intermediate forwarding nodes as reliable forwarders and unreliable forwarders, so as to achieve efficient multi-copy routing in communication. In order to reduce the cost of the regular update mechanism of node location information, Wang et al.[22] proposed a location-aware opportunity routing algorithm. First, a location-aware prediction model based on Markov process was established to predict the next location of the node. Then a message forwarding decision-making scheme is designed, and the required relay node is selected according to the sending probability of the node.

Routing algorithm considering community characteristics or node periodicity

On the basis of previous studies, researchers have integrated the community characteristics of nodes and the periodicity of node movement into the new routing algorithms in OMSNs. Niu et al.[23] proposed a community-based message transmission scheme (CMTS). In dealing with intra-community communication, CMTS determines the number of message copies in the network and chooses more active nodes as the relay nodes. For inter-community communication, messages are transmitted to target communities mainly relying on the nodes which have larger social degrees. Liu et al.[24] proposed an interest community routing (ICR) scheme. The network is divided into different communities based on the interest and the encounter number among nodes. Similar to the CMTS, the different message routing methods are applied for intra-community and inter-community. Our research team also proposed a community-aware message opportunistic transmission (CMOT) algorithm.[25] The traditional methods often rely on one or two nodes to transmit information when taking advantage of opportunities to mobile social networks. The consequence of this is that data transmission requires a long wait. Based on this, Wu et al.[26] reorganized the community, and then filtered duplicate nodes in the proposed method. By evaluating the contribution of nodes and reorganizing the community, the delivery ratio and cost in the opportunity mobile social network are improved. In order to ensure the safe and effective transmission of medical data, Chen et al.[27] proposed an adaptive routing optimization algorithm. The routing scheme defines the average message forwarding delay as a new indicator for selectively forwarding messages, and then designs a

local community detection algorithm based on this indicator to adapt to the characteristics of the opportunistic network, and finally relies on a certain super node to transfer messages between different nodes. Aiming at intermittently connected flying ad hoc networks, Li et al.[28] proposed a mobile-assisted adaptive routing algorithm, which combines routing strategies with location services, and uses different adaptive strategies to update the information of neighboring nodes and other nodes in the network.

The above community-based message forwarding algorithm considers the spatial characteristics of the network and improves the success rate but does not consider the time characteristics of the network. Considering the node movement is periodic, Wu et al.[29] proposed a periodic intermittently connected–based data (PICD) delivery algorithm. First step, for a certain node, based on its periodic intermittently connected paths to the sink and its data delay tolerance, PICD establishes the delay probability model by means of random dynamic programming. Second step, PICD will calculate the node's periodic delay probability distribution matrix by the function space iteration. Third step, according to the node's distribution matrix, the current data delivery probability can be calculated by specifying period and tolerable delay, and this will become the key-player of the next-hop choosing. Although the PICD can reduce network delay, it does not consider the spatial characteristics of OMSN. Meanwhile, the PICD is too complex. But the PICD is too complex to reduce the computation overhead.

The traditional opportunistic network routing algorithms are simply designed based on their physical properties of nodes. However, in the opportunistic network application scenario with human activity participation, the movement mode of the node is reflected the social behavior of the device holder or the social relationship in reality, so it has certain social attributes. Compared with physical attributes, the behavior of nodes has obvious paroxysmal, memory, and heterogeneity. Among the aforementioned existing studies, some studies did not consider the social characteristics of OMSNs, other studies used the spatial characteristics of the network when forwarding messages, and forwarded messages through intermittent connections through nodes periodically, but this kind of method is computationally expensive. Only by explaining the social connections between nodes from multiple dimensions, such as time dimension and space dimension, we can fully understand the characteristics of the OMSNs.

In this article, we set up a spatiotemporal constrained network model and propose an activity-based message opportunistic forwarding (AMOF) algorithm, using fully the temporal and spatial characteristics of mobile nodes in the OMSNs. The nodes fall into two categories in OMSNs: the free nodes and the

constrained nodes. For the constrained nodes, if they attend the same activity, we try to find out the node with the biggest delivery probability as relay node to forward messages. If they are not present the same activity, we try to find the optimal path using multi-hops to forward messages. Meanwhile, we consider the encounter cycles of nodes when the delivery probability is calculated. Our major contributions are summarized as following:

- We define a spatiotemporal constrained network model that can be suited to describe the spatio-temporal characteristics of OMSNs. The network model includes a series of activities, the cycles of nodes taking part in activities, and the encounter probability of nodes in different activity.
- We use different message forwarding strategies for different cases. If nodes attend the same activity, we pick out the best relay node which has the biggest delivery probability to forward messages. If nodes are not present the same activity, we try to find out the optimal path with multi-hops to forward messages.
- When we calculate the delivery probability, we not only consider the spatial characteristics in space dimension, but also consider the temporal characteristics in time dimension of nodes in OMSNs.

Modeling and assumptions for OMSNs

Activity-based mobility model

In this section, we define a spatiotemporal constrained network model, which we call it the activity-based mobility model. The description of proposed model is shown as follows:

Definition 1

   Network model. Assume that a network has n nodes which attend m activities, each activity takes place in itself cycle, respectively, and the nodes encounter with a certain probability in an activity. The network model is expressed with $G = (A, T, P)$ where $A = \{a_1, a_2, ..., a_m\}$ is the set of activities, $T = \{T_{a_1}, T_{a_2}, ..., T_{a_m}\}$ is the set of activity cycles that a node attends, and $P = \{P_{a_1}, P_{a_2}, ..., P_{a_m}\}$ is the set of node encounter probability in each activity. The T_{a_x} indicates a cycle table of $n \times n$, and P_{a_x} indicates an encounter probability table of $n \times n$.

Definition 2

   Activities. An activity is periodical behavior of nodes that occur in a specific time and place. Any node in the network can participate in several different activities or

not participate in any activities, but the node will not participate in 2 or more activities at the same time.

In the proposed network model, nodes fall into two categories: (1) the constrained nodes that need to attend a certain activity periodically and (2) the free nodes that can move freely, they are not part of any activity.

Assumptions of the activity-based mobility model

Each node attends respective activities for the different purpose. Nodes will encounter each other with a higher probability in the same activities than that of the nodes in different activities or out of any activity. Based on the spatiotemporal characteristics of the network and the mobility characteristics of nodes, we assume the activity-based mobility model as follows:

1. All nodes in the network are homogeneous and each node has the limited storage space.
2. The occurrence of activities in the network has a strict periodicity, and nodes may not participate in activities on time, or even "absent" an activity.
3. The node i may attend several activities to form the activity set A_i.
4. When nodes attend an activity, the nodes mobility within an activity follows random way point mobility model (RWP). If the node does not attend any activity, the node mobility within simulation area follows random walk mobility model (RW).

For the convenience of narration, Table 1 summarizes the main variable symbols in this article.

The proposed algorithm

If the source node and destination node include the same activities, that is, $A_i \cap A_d \neq \emptyset$, we will find out the best relay node by the biggest encounter probability in the same activity. It can avoid the messages spreading to other activities, reduces the network resource consumption.

If the source node and destination node does not exist the same activities, that is, $A_i \cap A_d = \emptyset$, we try to find the optimal path from the source node activity to the destination node activity. Therefore, AMOF routing algorithm includes two parts: (1) the message forwarding strategy with the same activities and (2) the message forwarding strategy without the same activities.

Message forwarding strategy with same activities

If the source node and destination node have the same activities, we can calculate the delivery probability by

Table 1. The detailed description of variable symbols.

| Symbol	Meaning of the symbol
$P_{a_x}(i,j)$	The encounter probability of the node i and node j in the activity a_x
$T_{a_x}(i,j)$	The encounter cycle of the node i and node j in the activity a_x
$TTLr$	The remaining lifetime of the message to be forwarded
a^P_x	The activity with the highest probability of meeting the target node
a^T_x	The activity with the shortest cycle of meeting the target node
a^E_x	The activity with the highest probability and shortest cycle of meeting the target node
N^a_E	The frequency of a^E_x in $TTLr$
N^a_P	The frequency of a^P_x in $TTLr$
N^a_T	The frequency of a^T_x in $TTLr$
N^a_{PT}	The frequency of the a^T_x when the a^P_x does not occur
$P_{rm}(i,j)$	The delivery probability of message
$P^a_m(i,j)$	The approximate expression of message delivery probability
$P_l(i,j)$	The encounter probability of node i & j in the path
$T_l(i,j)$	The encounter cycle of node i & j in the path
l^P	The path that there is biggest encounter probability to the target node
l^T	The path that there is smallest encounter cycle to the target node
l^E	The path that there is biggest encounter probability and smallest encounter cycle to the target node
N^l_E	The frequency of the l^E in the $TTLr$
N^l_P	The frequency of the l^P in the $TTLr$
N^l_T	The frequency of the l^T in the $TTLr$
N^l_{PT}	The frequency of the l^P when the l^T does not occur
$P^l_m(i,j)$	The indirect delivery probability of message

the encounter probability, the encounter cycle, and the remainder of time to live (TTL). According to the delivery probability, we estimate whether the message is forwarded. So we build the encounter probability table for each node to store the encounter probability when it encounters certain node. Then, we build the encounter cycle table for each node to store the encounter cycle when it encounters certain node.

How to get the encounter probability. When the node i encounters the node j in the activities of a_x, the encounter probability can be updated according to the formula (1). When node i and j often encounter, the formula ensures that the nodes have high delivery probability, where $P^{old}_{a_x}(i,j)$ is the encounter probability of node i and j before update, P_{init} is the initialization constant, and the initial value is 0.75.

$$P_{a_x}(i,j) = P^{old}_{a_x}(i,j) + \left(1 - P^{old}_{a_x}(i,j)\right) \times P_{init} \quad (1)$$

If the node i does not encounters the node j in the activities of a_x during a time interval, they are less likely to become good message forwarders to each other. So the encounter probability of the node i and j must gradually age. The aging equation is shown in formula (2), where $\gamma \in [0, 1]$ is the aging constant, and the k is the number of the time interval. The initial value of γ is 0.99 in our simulation experiments.

$$P_{a_x}(i,j) = P^{old}_{a_x}(i,j) \times \gamma^k \quad (2)$$
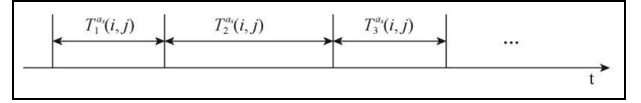
Figure 1. The schematic diagram of the encounter cycle.

How to get the encounter cycle. The node i and j took part in the activity a_x for several times in the history. The schematic diagram of encounter cycle for the node i and j is shown in Figure 1.

The encounter cycle is calculated with formula (3), it is a weighted average of the historical multiple encounter cycles, where $T^{a_x}_f(i,j)$ is the encounter cycle that the node i and j attend the activity at the fth times, $w^{a_x}_f$ is the weight coefficient that the node i and j attend the activity at the fth times, and g is the total number of recent encounters.

$$T_{a_x}(i,j) = \sum_{f=1}^{g} w^{a_x}_f T^{a_x}_f(i,j), f \leqslant g \quad (3)$$

We expect that messages are forwarded to the node which encountered with the highest frequency recently. So we set different weights for each encounter cycle according to the order of occurrence, the weight calculation formula is shown as formula (4). The nearest encounter cycle has the highest weight, while the farthest encounter cycle has lowest weight.

$$w_f^{a_x} = \begin{cases} \dfrac{1}{2^{g-1}}, & f = 1 \\ \dfrac{2^{f-2}}{2^{g-1}}, & f \geqslant 2 \end{cases} \tag{4}$$

How to get the delivery probability. Considering the spatio-temporal characteristics of OMSNs and the messages

TTL restrictions, the delivery probability can be calculated by formula (5), which measures the node ability for forwarding messages. In the formula, S is an activity set that the node i and j both attend. a_x^S indicates an activity in S. S' is an activity subset of S before the activity a_x^S appears. $a_x^{S'}$ indicates an activity in S'. $P_{a_x^S}(i,j)$ indicates the encounter probability in activity a_x^S, while $P_{a_x^{S'}}(i,j)$ indicates the encounter probability in activity $a_x^{S'}$.

$$P_{rm}(i,j) = \sum_{a_x^S \in S} \left[P_{a_x^S}(i,j) \times \prod_{a_x^{S'} \in S'} \left(1 - P_{a_x^{S'}}(i,j)\right) \right] \tag{5}$$

We can accurately calculate the delivery probability of message by formula (5), but it takes a huge amount of computing overhead. So we try to find out the key activities, which have a greater impact on message forwarding, using the encounter probability table and encounter cycle table based on the activities. The key activities have two kinds: one is the activity that has shortest encounter cycle to the destination node, called a_x^T, the other is the activity that has biggest encounter probability to the destination node, called a_x^P. We can get the approximate value of $P_{rm}(i,j)$, called $P_m^a(i,j)$.

If a_x^T and a_x^P are not the same activity, we assume that the M represents the number of a_x^T and the N represents the number of a_x^P. M can be calculated approximately by expression $N_{PT}^a = [T_{a_x^P}(i,j) / T_{a_x^T}(i,j)]$, which represents the number of a_x^T when a_x^P does not occur. N can be calculated approximately by expression $N_P^a = [TTL_r / T_{a_x^P}(i,j)]$, which represents the number of a_x^P in TTL_r. Where TTL_r indicates the remaining TTL time of message.

The approximate delivery probability $P_m^a(i,j)$ can be calculated, $P_m^a(i,j) = 1 - \left[\left(1 - P_{a_x^T}(i,j)\right)^M \left(1 - P_{a_x^P}(i,j)\right) \right]^N$.

We also assume that N_T^a represents the number of a_x^T in TTL_r. It can be calculated by $N_T^a = [TTL_r / T_{a_x^T}(i,j)]$. If $N_P^a = 0$ and $N_T^a \neq 0$, then $P_m^a(i,j) = 1 - \left(1 - P_{a_x^T}(i,j)\right)^{N_T^a}$. If $N_P^a = 0$ and $N_T^a = 0$, then $P_m^a(i,j) = P_{a_x^P}(i,j)$. So the delivery probability at any time is shown in formula (6).

$$P_m^a(i,j) = \begin{cases} 1 - \left[\left(1 - P_{a_x^T}(i,j)\right)^{N_{PT}^a} \left(1 - P_{a_x^P}(i,j)\right) \right]^{N_P^a}, & N_P^a \neq 0 \\ 1 - \left(1 - P_{a_x^T}(i,j)\right)^{N_T^a}, & N_T^a \neq 0, \; N_P^a = 0 \\ P_{a_x^P}(i,j), & N_T^a = 0 \end{cases} \tag{6}$$

Due to the reasonable assumption of the model, the maximum difference between $P_m^a(i,j)$ and $P_{rm}(i,j)$ is $M + 1$ activities error. We can calculate the convergence of power series by formula (7)

$$\lim_{N_P^a \to \infty} \frac{1 - \left[\left(1 - P_{a_x^T}(i,j)\right)^{N_{PT}^a} \left(1 - P_{a_x^P}(i,j)\right) \right]^{N_P^a + 1}}{1 - \left[\left(1 - P_{a_x^T}(i,j)\right)^{N_{PT}^a} \left(1 - P_{a_x^P}(i,j)\right) \right]^{N_P^a}} = 1 \tag{7}$$

When messages be produced, $TTL_r \approx TTL(TTL \gg T_{a_x^P}(i,j) > T_{a_x^T}(i,j))$, N_P^a is large numbers, and $1 - \left[\left(1 - P_{a_x^T}(i,j)\right)^{N_{PT}^a} \left(1 - P_{a_x^P}(i,j)\right) \right]^{N_P^a} \leqslant P_{rm}(i,j) \leqslant 1 - \left[\left(1 - P_{a_x^T}(i,j)\right)^{N_{PT}^a} \left(1 - P_{a_x^P}(i,j)\right) \right]^{N_P^a + 1}$, so $P_m^a(i,j) \approx P_{rm}(i,j)$. But the error of $P_m^a(i,j)$ becomes larger when TTL is going to run out.

If a_x^T and a_x^P are the same activity, that is, the activity has not only shortest encounter cycle, but also has biggest encounter probability to the destination node, we indicate it as a_x^E. So the delivery probability can be calculated by formula (8)

$$\begin{aligned} P_m^a(i,j) &= P_{a_x^E}(i,j) + \left(1 - P_{a_x^E}(i,j)\right) P_{a_x^E}(i,j) \\ &+ \left(1 - P_{a_x^E}(i,j)\right)^2 P_{a_x^E}(i,j) + \cdots \\ &+ \left(1 - P_{a_x^E}(i,j)\right)^{N-1} P_{a_x^E}(i,j) \\ &= 1 - \left(1 - P_{a_x^E}(i,j)\right)^N \end{aligned} \tag{8}$$

In the formula, N can be calculated approximately by expression $N_E^a = [TTL_r / T_{a_x^E}(i,j)]$, and N_E^a represents the number of a_x^E in TTL_r. If $N_E^a = 0$, then $P_m^a(i,j) = P_{a_x^E}(i,j)$. Therefore, the delivery probability at any time is shown in formula (9)

$$P_m^a(i,j) = \begin{cases} 1 - \left(1 - P_{a_x^E}(i,j)\right)^{N_E^a}, & N_E^a \neq 0 \\ P_{a_x^E}(i,j), & N_E^a = 0 \end{cases} \quad (9)$$

Message forwarding strategy without same activities

Because the source node and the destination node does not exist at the same activity, that is, $A_i \cap A_d = \emptyset$, messages cannot be transferred directly from the source node to the destination node. Messages need relay nodes in the relay activities to forward. There could be multiple paths to forward messages from the source node to the destination node through relay nodes. So we try to find the optimal path from the source node to the destination node.

Transmission probability based on link. If there is a node h that shares the same activities with the message carried node and attends the same activities with the destination node, we can build a link of message forwarding by the relay node h, that is, $(i \rightarrow h \rightarrow j)$. So the message transmission probability based on the link can be described in formula (10), where $P_l(i,j)$ represents the transmission probability of the node i and j by the link l, a_x is the same activity of node i and h, while a_y is the same activity of node h and j. $P_{a_x}(i,h)$ is the encounter probability of node i and h, and $P_{a_y}(h,j)$ is the encounter probability of node h and j.

$$P_l(i,j) = P_{a_x}(i,h) \times P_{a_y}(h,j) \quad (10)$$

We can get the encounter cycle of two nodes in the link, shown in formula (11), where $T_{a_y}(h,j)$ indicates the encounter cycle of the node h and j in the activity a_y, $z = \min\{N^*\}$.

$$T_l(i,j) = \begin{cases} \max\{T_{a_x}(i,h), T_{a_y}(h,j)\}, & T_{a_x}(i,h) \leqslant T_{a_y}(h,j) \\ z \times T_{a_y}(h,j) \geqslant T_{a_x}(i,h), & T_{a_x}(i,h) > T_{a_y}(h,j) \end{cases}$$
$$(11)$$

How to get the indirect delivery probability. There could be multiple links between the source node and the destination node. We get the optimal path by the transmission probability and the encounter cycle in the link. In order to find the best link, we assume l^T as the link that has the shortest encounter cycle, and assume l^P as the link that has biggest encounter probability. Similar to the solving process of the delivery probability in section "Message forwarding strategy with same activities," we can get the solution of the indirect delivery probability, as shown in formulas (12) and (13).

If l^T and l^P are not the same link, the indirect delivery probability can be calculated as formula (12), where $N_{PT}^l = [T_{l^P}(i,j)/T_{l^T}(i,j)]$ represents the number of l^T

when l^P does not occur, $N_T^l = [TTL_r/T_{l^T}(i,j)]$ represents the number of l^T in TTL_r, and $N_P^l = [TTL_r/T_{l^P}(i,j)]$ represents the number of l^P in TTL_r.

$$P_m^l(i,j) =$$
$$\begin{cases} 1 - \left[(1 - P_{l^T}(i,j))^{N_{PT}^l}(1 - P_{l^P}(i,j))\right]^{N_P^l}, & N_P^l \neq 0 \\ 1 - (1 - P_{l^T}(i,j))^{N_T^l}, & N_P^l = 0, N_T^l \neq 0 \\ P_{l^P}(i,j), & N_T^l = 0 \end{cases}$$
$$(12)$$

If l^T and l^P are the same link, noted l^E, then the indirect delivery probability can be calculated by formula (13), where $N_E^l = [TTL_r/T_{l^E}(i,j)]$ represents the number of l^E in TTL_r.

$$P_m^l(i,j) = \begin{cases} 1 - (1 - P_{l^E}(i,j))^{N_E^l}, & N_E^l \neq 0 \\ P_{l^E}(i,j), & N_E^l = 0 \end{cases} \quad (13)$$

Message forwarding procedure

If $A_i \cap A_d \neq \emptyset$, we will find out the best relay node by the biggest encounter probability in the same activity. If $A_i \cap A_d = \emptyset$, we try to find the optimal path from the source node activities to the destination node activities. The *Procedure A* and *Procedure B* will finish the different task respectively as following.

Message forwarding procedure with same activities. The node, which carries the messages, will send messages to the encounter node, if the delivery probability of the encounter node is bigger than that of current node. Otherwise, messages cannot be forwarded to the encounter node.

> *Procedure A.*
> Step 1. Node i encounters node j.
> Step 2. If node j is the destination node, then send messages to node j, go to Step 6.
> Step 3. If there are no same activities between node i and destination node, then turn to the message forwarding strategy without the same activities.
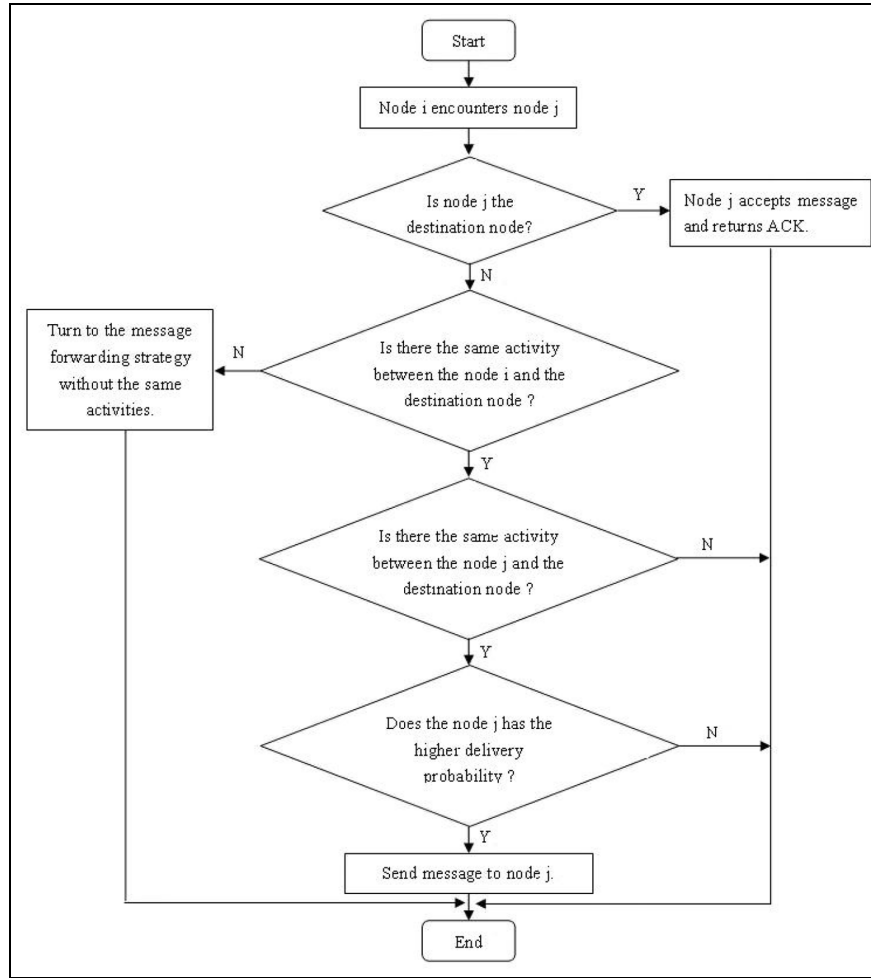> Step 4. If there are no same activities between the node j and destination node then go to Step 6.
> Step 5. If node j has higher delivery probability, then node i sends messages to node j.
> Step 6. End of the algorithm.

The block diagram of message forwarding procedure with same activities is shown in Figure 2.

Message forwarding procedure without same activities. The main idea of messages forwarding without the same

Figure 2. Messages forwarding flow chart with the same activities.

activities is shown as follow. The node which carries the messages sends messages to the encounter node, if the indirect delivery probability of the encounter node is bigger than that of current node. Otherwise, messages are not be sent to the encounter node.

Procedure B.
Step 1. Node i encounters node j.
Step 2. If node j is the destination node, then send messages to node j, go to Step 6.
Step 3. If there are the same activities between node i and destination node then turn to the message forwarding strategy with the same activities.
Step 4. If there are the same activities between the node j and destination node, then send messages to node j, and go to Step 6.
Step 5. If node j has higher delivery probability, then node i sends messages to node j.
Step 6. End of the algorithm.

The block diagram of message forwarding procedure without same activities is shown in Figure 3.

Simulation scenario and results analysis

Simulation scenario and parameters setting

In this article, we use the ONE (the opportunistic network environment) to simulate, compare with the typical algorithms, such as PRoPHET, Epidemic, CMTS, CMOT, and ICPP. Before the simulation begins, the AMOF algorithm runs 1000 s to complete the preprocessing for building activities. The specific simulation parameters are set in Table 2.

In the above table, the value of the radius of activity region is set to 100/150/200/250/300/350, respectively, so that the influence of the range of activity occurrence on message forwarding can be investigated. Similarly, setting different parameter values for cache size, number of node, message TTL, the interval of message generation, and so on, can comprehensively consider the effectiveness of the method proposed in this article.

Experiment results and analysis

Based on the simulation scenario in section "Simulation scenario and parameters setting," we compare and
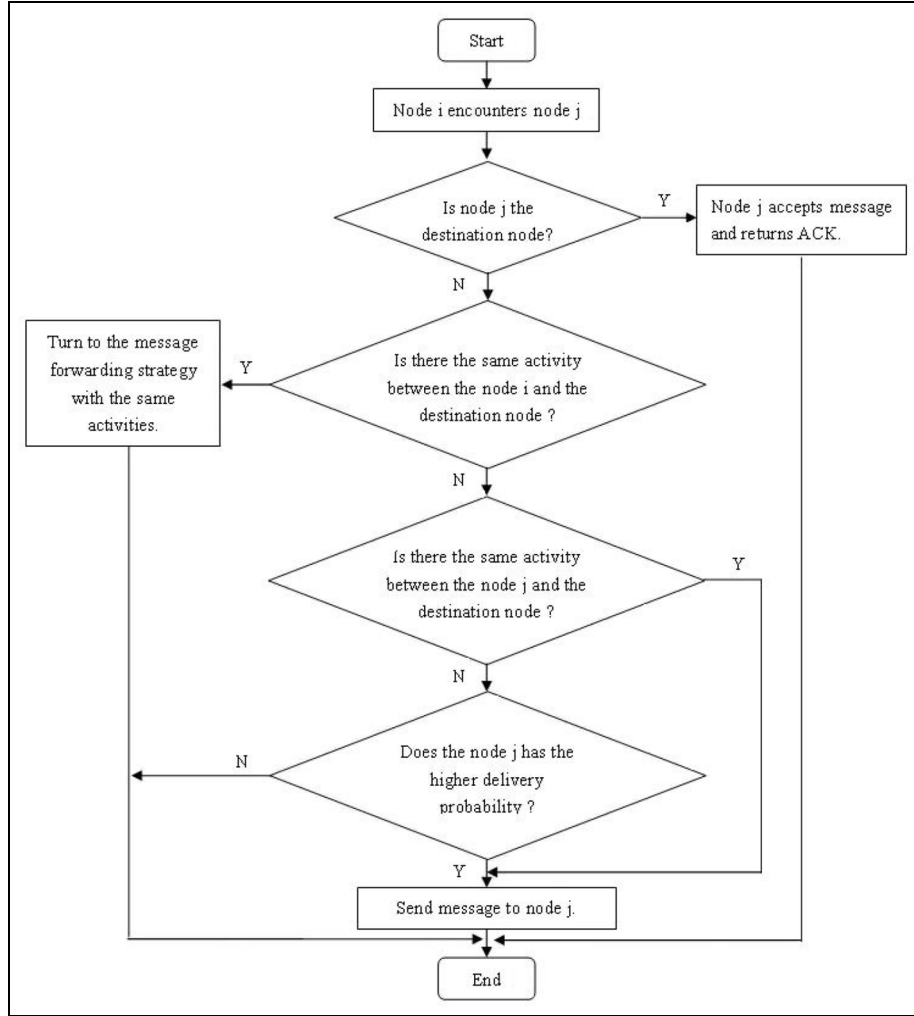
Figure 3. Message forwarding flow chart without the same activities.

Table 2. The parameters of simulation scenario.

Parameters (unit)	Values
Simulation region (m²)	7500 × 7500
Simulation time (h)	12
The radius of activity region (m)	100/150/200/250/300/350
The number of node	64/112/160/208/256
The movement velocity of node (m/s)	3~9
The size of message (MB)	0.5~1
Cache size (MB)	5/7.5/10/12.5/15/17.5/20
The interval of message generation (s)	25~35
TTL (min)	180/240/300/360/420/480

analyze the performance of six algorithms with different parameters for activity radii, cache size, number of nodes, TTL, and message interval. The metrics include (1) the successful delivery ratio, (2) the network overhead ratio, and (3) the average network delay.

The impact of the activity radii. In the current experiment, we set that the number of nodes is 160, the node cache size is 20 MB, and the message TTL is 360 min. We survey the impact for the aforementioned different routing algorithms by changing the radius of activities. The simulation results are shown in Figures 4–6.

In Figure 4, we can observe that the AMOF, CMOT, CMTS, and ICPP routing algorithms have higher successful delivery ratio than that of the PRoPHET and Epidemic. The reason is that the former four routing algorithms have considered the social characteristics (i.e. the space characteristic of clustering) of the OMSNs, while the latter two routing algorithms have not considered the network social characteristics. The AMOF shows the highest successful delivery ratio among them, because it not only considers space characteristic, but also considers the time cycle characteristic.

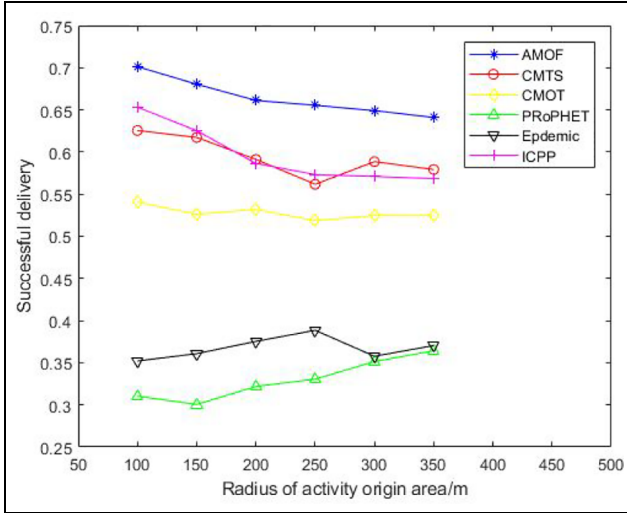Figure 5 shows that the overhead of the CMTS, CMOT, ICPP, and AMOF remains unchanged, while

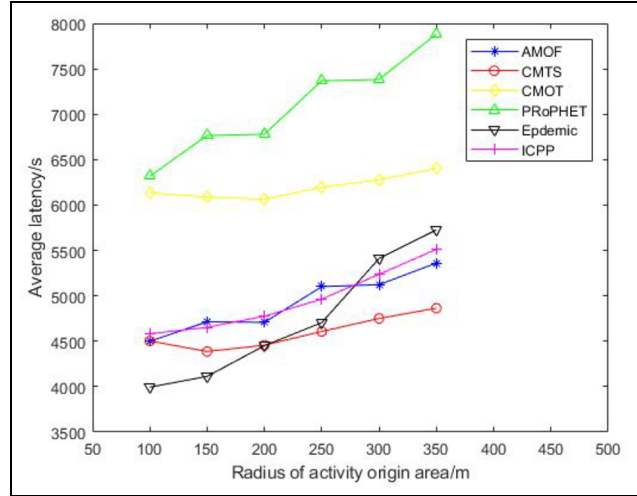**Figure 4.** Comparison of the delivery ratio in different radii.

Figure 6. Comparison of the average delay in different radii.
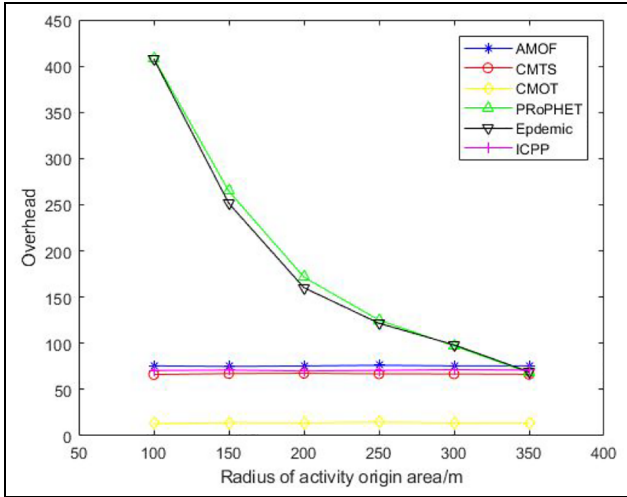

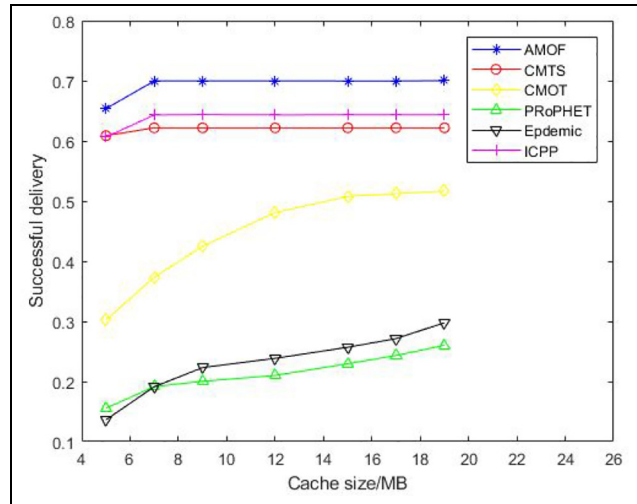**Figure 5.** Comparison of the overhead ratio in different radii.

Figure 7. Comparison of delivery ratio in different cache sizes.

the overhead of the PRoPHET and Epidemic decreases when the radius of activity increases. The former four algorithms have lower overhead than the latter two algorithms, and the CMOT has the lowest overhead among them. The reason is that the former four algorithms consider the clustering characteristics and use different message forwarding strategies for intra-cluster and inter-cluster. The CMOT shows the best performance because it uses "one-copy and one-hop for forwarding" strategy in intra-cluster.

We can see that all algorithms increase the average delay with increasing the radius of activity, and the average delay of the Epidemic, CMTS, ICPP, and AMOF is lower than that of the CMOT and PRoPHET.

Synthesizing Figures 4–6, the AMOF increases the successful delivery ratio, respectively, by 15.77%,

16.12%, and 31.36% comparing with the CMTS, ICPP, and CMOT, when they have the similar overhead and network delay.

The impact of cache size. In the current experiment, we set that the number of nodes is 160, the activity radius are 100 m, and the message TTL is 360 min. We survey the impact for the aforementioned different routing algorithms by changing the cache size. The simulation results are shown in Figures 7–9.

In Figures 7 and 8, we can observe that the AMOF, CMOT, ICPP, and CMTS routing algorithms have higher successful delivery ratio and lower network overhead ratio than that of the PRoPHET and Epidemic. The AMOF shows the highest successful delivery ratio among them, and the CMOT shows the lowest
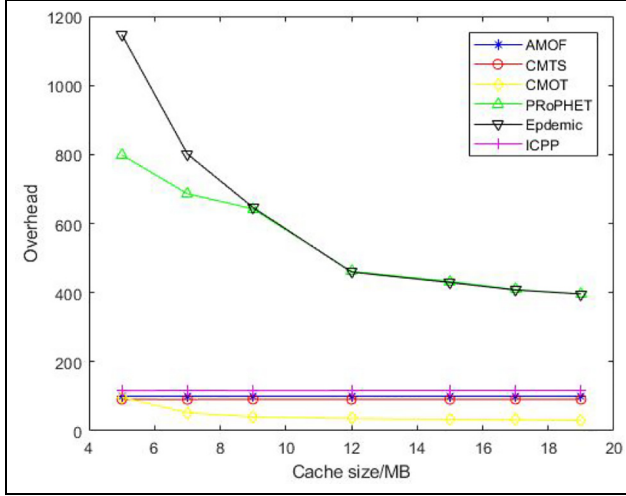
Figure 8. Comparison of overhead ratio in different cache sizes.
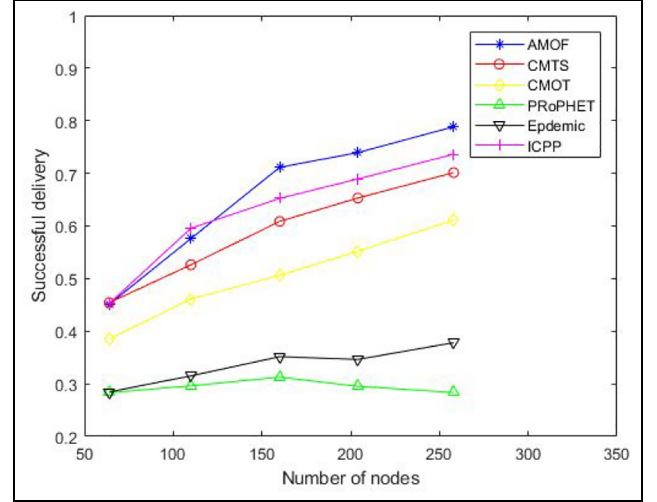
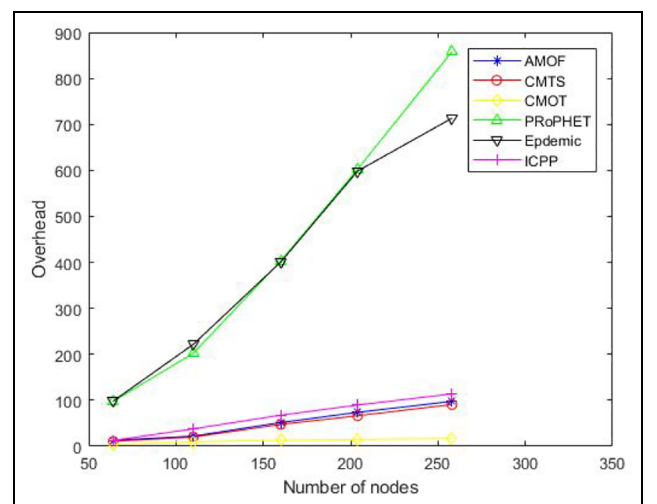**Figure 10.** Comparison of delivery ratio in different node densities.
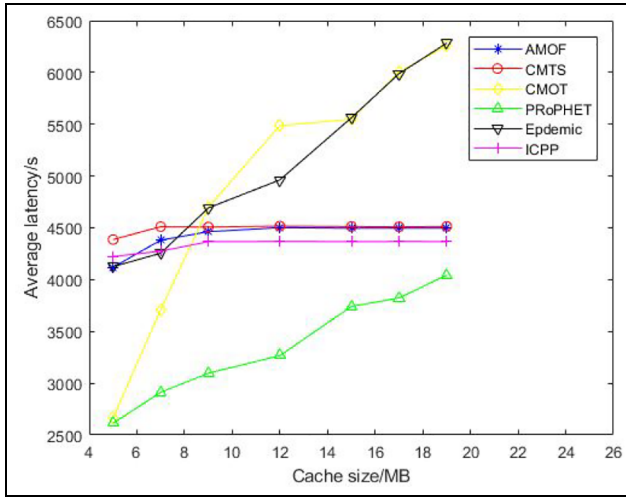
Figure 9. Comparison of average delay in different cache sizes.

Figure 11. Comparison of overhead ratio in different node densities.

overhead ratio. The Figure 9 shows that the average delay of the AMOF, ICPP, and CMTS are not affected by changing of cache size, while the average delay of the CMOT, PRoPHET, and Epidemic increase with increasing the cache size. The PRoPHET has the lowest average delay among all algorithms.

The reason is that the AMOF, CMOT, CMTS, and ICPP have considered the social characteristics of the OMSNs, while the PRoPHET and Epidemic have not considered the network social characteristics. The PRoPHET uses "multi-copy and multi-hop forwarding" strategy to win the lowest average delay.

In summary, the overhead ratio and average network delay of the AMOF algorithm almost are not affected, the successful delivery ratio is highest value among all compared algorithms, when the cache size increases step by step.

The impact of the node number. In the current experiment, we set that the cache size is 20 MB, the activity radius are 100 m, and the message TTL is 360 min. We survey the impact for the aforementioned different routing algorithms by changing the node density. The simulation results are shown in Figures 10–12.
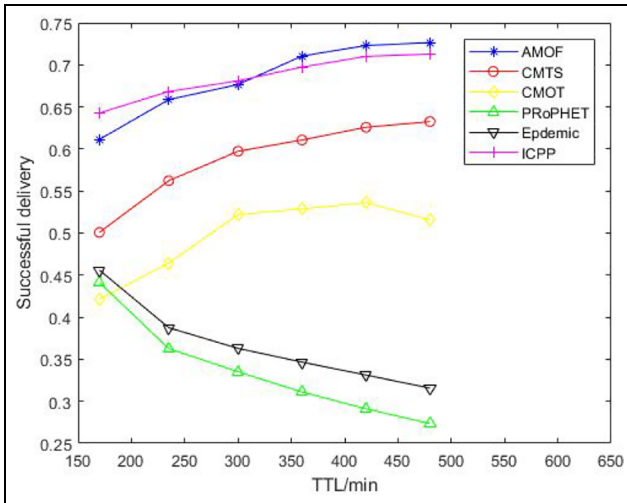
In Figure 10, we can observe that all routing algorithms increase the successful delivery ratio with increasing the number of nodes. The successful delivery ratio of the AMOF, CMOT, ICPP, and CMTS which have considered the social characteristics of network increase faster than that of PRoPHET and Epidemic. The successful delivery ratio of the Epidemic increase slowly before the number of nodes is 160, then it decrease slowly after the number of nodes is 160.
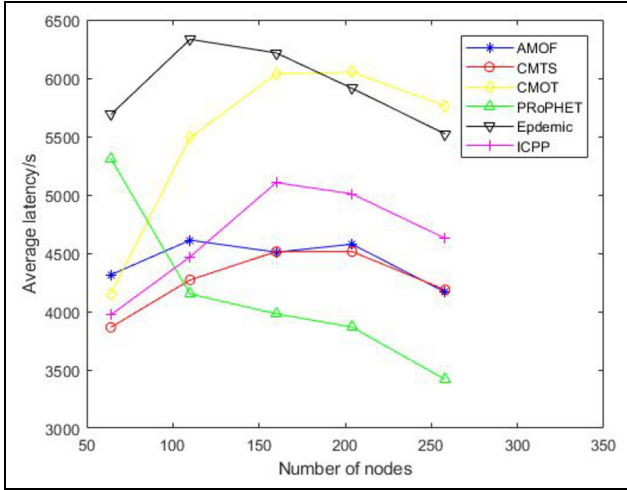
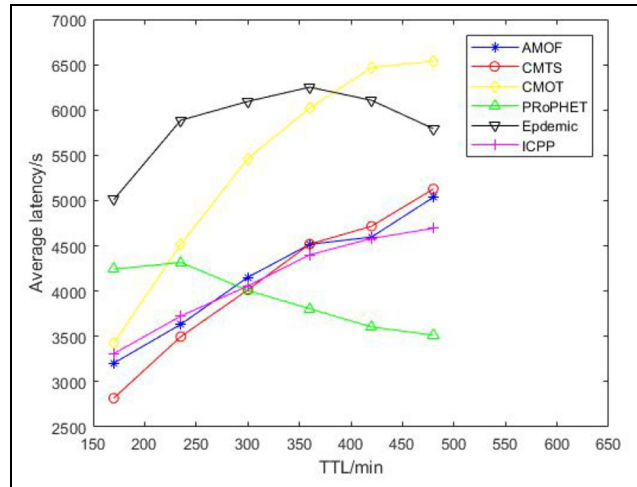**Figure 12.** Comparison of average delay in different node densities.
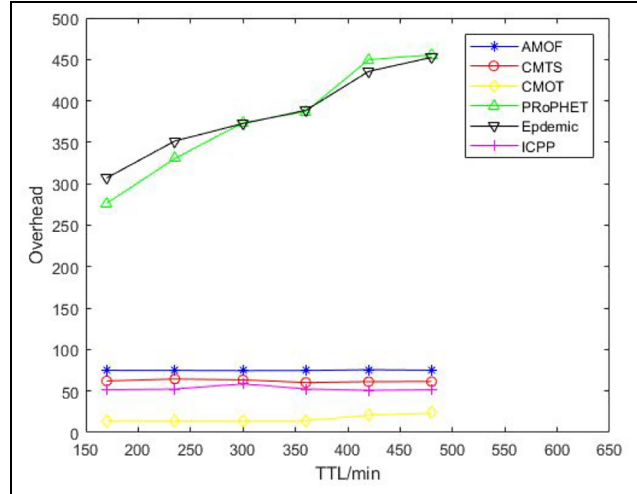
Figure 14. Comparison of overhead ratio in different TTLs.

Figure 13. Comparison of delivery ratio in different TTLs.

Figure 15. Comparison of average delay in different TTLs.

Among all compared routing algorithm, the AMOF gets the highest successful delivery ratio, its successful delivery ratio increases 7.37%, 13.58%, and 30.63% than that of the ICPP, CMTS, and CMOT, respectively.

Figure 11 shows that the AMOF, CMOT, ICPP, and CMTS routing algorithms have lower network overhead ratio than that of the PRoPHET and Epidemic, so they are not sensitive to the changing of the node number. Meanwhile the network overhead of the PRoPHET and Epidemic increases quickly with increasing the number of nodes. The reason is that the node density becomes high with increasing the number of nodes when the region of network is not change. There is more message invalid forwarding in the routing algorithm of the PRoPHET and Epidemic.
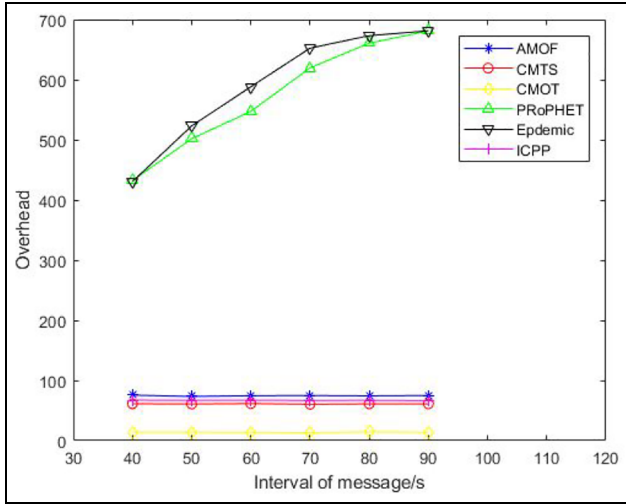
Figure 12 shows that the AMOF and CMTS are not sensitive to the changing of the node number, and have lower average network delay than that of the CMOT, ICPP, and Epidemic. The PRoPHET uses "multi-copy and multi-hop forwarding" strategy to get the lowest average delay, but its performance is not stable. When the number of nodes is less than about 80, the average network delay of the PRoPHET is higher than that of the AMOF, CMOT, ICPP, and CMTS and is only lower than that of the Epdemic.

The impact of the message TTL. In the current experiment, we set that the number of nodes is 160, the cache size is 20 MB, and the radius of activity is 100 m. We survey the impact for the aforementioned different routing algorithms by changing the TTL of messages. The simulation results are shown in Figures 13–15.
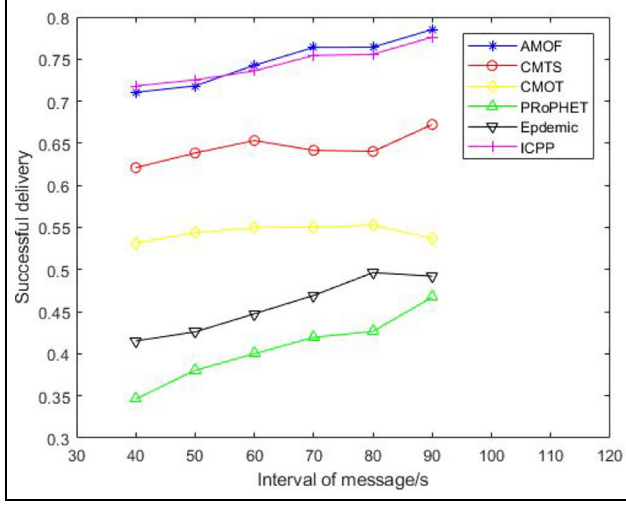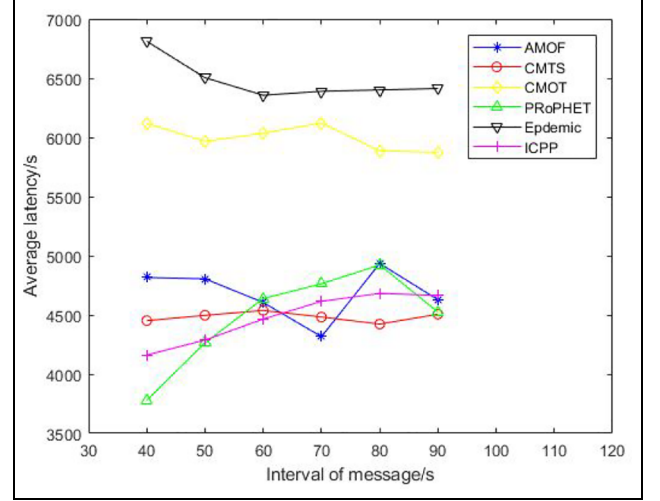
Figure 16. Comparison of delivery ratio in different message intervals.

Figure 18. Comparison of average delay in different message intervals.

Figure 17. Comparison of overhead ratio in different message intervals.

Figures 13 and 14 show that with the increasing of message TTL, the successful delivery ratio of the Epidemic and PRoPHET decreased, and the network overhead of the Epidemic and PRoPHET increased. The reason is that the Epidemic and PRoPHET have not the strategies to control redundant messages effectively. While the AMOF, CMTS, ICPP, and CMOT routing algorithms consider fully the constraints of space and time, and use different message forwarding strategies for intra-cluster and inter-cluster. So they get higher successful delivery probability and lower overhead ratio than that of Epidemic and PRoPHET.

In Figure 15, we can observe that the TTL of message has a great influence on the average network delay. When the TTL becomes long, the time for waiting and picking the relay nodes also becomes correspondingly

long in routing algorithms of the AMOF, CMTS ICPP, and CMOT. The CMOT is most significantly affected by TTL. However, the Epidemic algorithm does not wait and pick the relay node. The message is forwarded once the node is encountered. When the TTL becomes long, the Epidemic has more time to encounter more relay nodes to forward messages.

The impact of the message interval. In the current experiment, we set that the number of nodes is 160, the radius of activity is 100 m, the node cache size is 20 MB, and the message TTL is 360 min. We survey the impact for the aforementioned different routing algorithms by changing the message interval. The simulation results are shown in Figures 16–18.

In Figure 16, we can observe that all routing algorithms increase slowly the successful delivery ratio with increasing the message interval. The reason is that the number of messages decreased when the message interval increased, so the forwarding probability of each message becomes high, the successful delivery ratio increased. The AMOF is best algorithm among them.

Figure 17 shows that the AMOF, CMOT, ICPP, and CMTS routing algorithms have lower network overhead ratio than that of the PRoPHET and Epidemic. It shows that AMOF, CMOT, ICPP, and CMTS routing algorithms are not sensitive to the changing of the node number. Meanwhile the network overhead of the PRoPHET and Epidemic increases slowly with increasing the number of nodes. The reason is that the AMOF, CMOT, ICPP, and CMTS control the message copies, while the PRoPHET and Epidemic do not limit the message copies. With increasing the message interval, there is more message invalid forwarding in the routing algorithm of the PRoPHET and Epidemic.

Figure 18 shows that the AMOF, ICPP, CMTS, and PRoPHET have lower average network delay than that of the CMOT and Epidemic. The AMOF, ICPP, and CMTS are not sensitive to the changing of the message interval. The PRoPHET uses "multi-copy and multi-hop forwarding" strategy to get the lowest average delay when the message interval is less than 60 s. But the CMTS has most stable performance of the network average delay in these algorithms.

Conclusion

Based on spatial and temporal characteristics of nodes in OMSNs, we proposed an AMOF algorithm. The main idea is that we defined the activity which a group of nodes attend it periodically and modeled a time-varying activity-based network, denoted as $G = (A, T, P)$. We select the relay node which has the biggest delivery probability to forward messages when the source node and destination node include the same activity. Otherwise, we get the best path according to the indirection delivery probability to transfer messages.

In the simulation experiments, compared with previous classical opportunistic routing algorithms, the AMOF algorithm achieved the optimal performance in terms of message successful delivery ratio and has lower network overhead and delay than other algorithms.

In the follow-up work, we will further study the energy consumption of nodes in 5G network environment. On the premise of ensuring the security of data transmission, the network structure and various attributes of nodes are analyzed to improve the transmission performance and reduce energy consumption as much as possible.

ORCID iD

Sheng Zhang https://orcid.org/0000-0002-5303-236X

References

1. Sachdeva R and Dev A. Review of opportunistic network: assessing past, present, and future. *Int J Commun Syst* 2021; 34: e4860.
2. Wu J, Chen Z and Zhao M. Information cache management and data transmission algorithm in opportunistic social networks. *Wirel Netw* 2019; 25(6): 2977–2988.
3. Singh AK and Pamula R. An efficient and intelligent routing strategy for vehicular delay tolerant networks. *Wirel Netw* 2021; 27(1): 383–400.
4. Zhou H, Wang H, Chen X, et al. Data offloading techniques through vehicular ad hoc networks: a survey. *IEEE Access* 2018; 6: 65250–65259.
5. Huang CM, Chen YF, Xu S, et al. The vehicular social network (VSN)-based sharing of downloaded geo data using the credit-based clustering scheme. *IEEE Access* 2018; 6: 58254–58271.
6. Zhang C, Du Z, Parmar MD, et al. Pocket-switch-network based services optimization in crowdsourced delivery systems. *Comput Electr Eng* 2017; 62: 53–63.
7. Zhang H, Chen Z, Wu J, et al. FRRF: a fuzzy reasoning routing-forwarding algorithm using mobile device similarity in mobile edge computing-based opportunistic mobile social networks. *IEEE Access* 2019; 7: 35874–35889.
8. Abualsaud K, Elfouly TM, Khattab T, et al. A survey on mobile crowd-sensing and its applications in the IoT era. *IEEE Access* 2018; 7: 3855–3881.
9. Lenando H and Alrfaay M. EpSoc: social-based epidemic-based routing protocol in opportunistic mobile social network. *Mob Inform Syst* 2018; 2018: 6462826.
10. Derakhshanfard N, Sabaei M and Rahmani AM. Sharing spray and wait routing algorithm in opportunistic networks. *Wirel Netw* 2016; 22(7): 2403–2414.
11. Burgess J, Gallagher B, Jensen D, et al. MaxProp: routing for vehicle-based disruption-tolerant networks. In: *Proceedings of the 25th IEEE international conference on computer communications (INFOCOM 2006)*, Barcelona, 23–29 April 2006, pp.1–11. New York: IEEE.
12. Chancay-García L, Hernández-Orallo E, Manzoni P, et al. Optimising message broadcasting in opportunistic networks. *Comput Commun* 2020; 157: 162–178.
13. Chunyue Z, Tian H, Dong Y, et al. An energy-saving routing algorithm for opportunity networks based on asynchronous sleeping mode. *Comput Electr Eng* 2021; 92: 107088.
14. Derakhshanfard N. Erlang based buffer management and routing in opportunistic networks. *Wirel Pers Commun* 2020; 110(2): 2165–2177.
15. Wu L, Cao S, Chen Y, et al. An adaptive multiple spray-and-wait routing algorithm based on social circles in delay tolerant networks. *Comput Netw* 2021; 189(12): 107901.
16. Lindgren A, Doria A and Schelén O. Probabilistic routing in intermittently connected networks. In: *Proceedings of the international workshop on service assurance with partial and intermittent resources*, Fortaleza, Brazil, 1–6 August 2004, pp.239–254. Berlin: Springer.
17. Ramesh S. An efficient secure routing for intermittently connected mobile networks. *Wirel Pers Commun* 2017; 94(4): 2705–2718.

18. Abdelkader T, Naik K, Nayak A, et al. A performance comparison of delay-tolerant network routing protocols. *IEEE Netw* 2016; 30(2): 46–53.

19. Wu J, Yu G and Guan P. Interest characteristic probability predicted method in social opportunistic networks. *IEEE Access* 2019; 7(99): 59002–59012.

20. Xu G, Wang X, Zhang N, et al. A routing algorithm for the sparse opportunistic networks based on node intimacy. *Wirel Commun Mob Comput* 2021; 2021(3): 1–9.

21. Srinidhi NN, Sagar CS, Deepak Chethan S, et al. An improved PRoPHET —random forest based optimized multi-copy routing for opportunistic IoT networks. *Intern Things J* 2020; 11: 100203.

22. Wang X, Wang X, Wang L, et al. An efficient location-aware routing approach in opportunistic networks. *IEEJ Trans Electr Electron Eng* 2020; 15(5): 704–713.

23. Niu J, Zhou X, Liu Y, et al. A message transmission scheme for community-based opportunistic network. *J Comput Res Dev* 2009; 46(12): 2068–2075.

24. Liu QL, Hu CF, Zhu DL, et al. Interest community detecting method and routing scheme in opportunistic networks. *J Beijing Univ Post Telecom* 2014; 37(3): 62–66.

25. Zhang S, Liu XD, Bao XL, et al. Community-based message opportunistic transmission scheme in mobile social networks. *J Softw* 2015; 26: 29–38.

26. Wu J, Chen Z and Zhao M. Community recombination and duplication node traverse algorithm in opportunistic social networks. *Peer-to-Peer Netw Appl* 2020; 13(13): 940–947.

27. Chen W, Chen Z and Cui F. Adaptive routing optimization algorithm in community-oriented opportunistic networks for mobile health. *Sensors* 2019; 19(8): 1876.

28. Li X, Deng F and Yan J. Mobility-assisted adaptive routing for intermittently connected FANETs. *IOP Conf Ser Mater Sci Eng* 2020; 715(1): 012028.

29. Wu L, Wu DA, Liu M, et al. Periodic intermittently connected-based data delivery in opportunistic networks. *J Softw* 2013; 24(3): 507–525.