ARTICLE

Enhancing Data Forwarding Efficiency in SIoT with Multidimensional Social Relations

Fang Xu[1,2,3], Songhao Jiang[1,2], Yi Ma[1,2,3,*], Manzoor Ahmed[1,3,*], Zenggang Xiong[1,2,3] and Yuanlin Lyu[1,2]

[1]School of Computer and Information Science, Hubei Engineering University, Xiaogan, 432000, China

[2]School of Computer Science and Information Engineering, Hubei University, Wuhan, 430062, China

[3]Artificial Intelligence Industrial Technology Research Institute, Hubei Engineering University, Xiaogan, 432000, China

*Corresponding Authors: Yi Ma. Email: mayi@whu.edu.cn; Manzoor Ahmed. Email: manzoor.achakzai@gmail.com

ABSTRACT

Effective data communication is a crucial aspect of the Social Internet of Things (SIoT) and continues to be a significant research focus. This paper proposes a data forwarding algorithm based on Multidimensional Social Relations (MSRR) in SIoT to solve this problem. The proposed algorithm separates message forwarding into intra- and cross-community forwarding by analyzing interest traits and social connections among nodes. Three new metrics are defined: the intensity of node social relationships, node activity, and community connectivity. Within the community, messages are sent by determining which node is most similar to the sender by weighing the strength of social connections and node activity. When a node performs cross-community forwarding, the message is forwarded to the most reasonable relay community by measuring the node activity and the connection between communities. The proposed algorithm was compared to three existing routing algorithms in simulation experiments. Results indicate that the proposed algorithm substantially improves message delivery efficiency while lessening network overhead and enhancing connectivity and coordination in the SIoT context.

KEYWORDS

SIoT; data forwarding; social attributes; social relations; community

1 Introduction

With the rapid development of SIot since the 21st century, the performance of various mobile devices has been significantly improved. Smartphones have also become an irreplaceable part of people's daily life. Its usage rate and coverage rate have comprehensively surpassed that of traditional computers, according to China Internet Network Information Center (CINIC). By December 2022, the Internet penetration rate reached 75.6%. With the popularity of mobile devices, various network communication theories are constantly improving. Social Internet of Things (SIoT) [1] emerged at times and required diversified applications and increasing network and product scale. The processing of wireless network connection, forwarding, and interruption becomes increasingly essential.

Traditional mobile networks are mainly cellular networks. A cellular network consists of three components: the mobile station, the base station subsystem, and the network subsystem. The emergence of SIoT networks has been fuelled by developments in wireless network technology, sensing technology and sociology. It is developed from the Delay Tolerant Network (DTN) [2], which combines the connectivity of IoT devices with the interactive nature of social networks. Its goal is to improve the quality of service and user experience by enabling IoT devices to connect to each other and share information. They do not have a fixed transmission path. The transmission of information between network nodes is obtained through encounters between nodes. The transmission is usually carried out in a multi-hop mode. In the SIoT, different types of wireless communication devices can act as nodes that transmit information in the network, which is more in line with the needs of self-organizing networks. Because SIoT can effectively improve the lack of communication infrastructure in wireless networks, the network environment could be better, and the connectivity and interruption between nodes are frequently changed. Therefore, there are more and more research and development related to SIoT, and its research and application value is increasingly shown in the research of wireless network, which can be widely used in the fields of network transmission in remote areas, vehicle-connected network [3], post-disaster rescue, medical health, smart home city management, supply chain management, agriculture and other fields. Meanwhile, SIoT can effectively increase the overall efficiency of the system, improve the quality of service, enhance flexibility, promote innovation, and improve security. Overall, SIoT brings significant benefits to a wide range of application areas by combining the connectivity capabilities of IoT devices with the interactive nature of social networks.

SIoT is a new field that combines IoT with social networking, utilising the principles of social networking to enhance interoperability and collaboration between devices. In this environment, network imbalance may manifest itself in the form of uneven resource allocation, inconsistent quality of service, and network congestion. The network performance characteristics of imbalance mainly include uneven resource allocation, inconsistent quality of service, network congestion, and inconsistent response time. Therefore we can solve the problem of network imbalance by resource allocation strategy, load balancing, optimising network topology, performance monitoring and management, dynamic adjustment strategy, and social network analysis.

Under certain conditions, SIoT nodes will move randomly; they work in a "store-carry-forward" mode. Furthermore, nodes in the network may be objects with sensing devices. Smart devices play an indispensable role in people's daily life. The concept of the SIoT is also gaining popularity. SIoT has the characteristics of the Internet of Things (IoT) and human society [4]. As a social network, the aim is to turn the connection of various devices in the Internet of Things into something like how humans interact in society. With the advancement of AI technology, various smart devices are expected to act and think like humans [5].

As the main participants in social activities, human beings' social attributes and relationships are indispensable to interpersonal communication. For example, communication between two people who share the same interests is often closer than communication between two people who do not share the same interests. Nodes with common interests and social connections will gather to form new community structures [6]. Most of the time, nodes will be active in several communities they are interested in, and a small part will go to other communities to enhance communication. Nodes belonging to the same interest community interact frequently and are more likely to meet; however, there are fewer interactions from different communities. Based on the information transmission mode of the above nodes, a better data forwarding algorithm can be designed to improve the information delivery rate.

Research shows that dividing nodes into different community structures can effectively increase information delivery rates [7]. Relationships and attributes between nodes greatly influence the probability of nodes meeting each other. Nodes with high social relationships also have a higher probability of meeting other nodes, and the same holds between communities. Therefore, message-forwarding efficiency can be effectively improved when nodes or communities with strong social relationships act as relay roles. This paper defines three metrics: node social relationship strength, node activity, and community connection. A linear combination of three indexes designs a new data-forwarding algorithm. Firstly, the nodes are divided into communities based on their social attributes. The source node carries the message through the network and determines whether the source and target nodes are in the same community. If yes, the source node will select the node in that community with the most similar interest attributes and the closest social ties to it as the relay node. If not, the community with the highest social connection will be selected as the relay community. And relatively similar and active nodes in that community will be selected as relay nodes for message forwarding.

1.1 Related Work

The routing strategy design in the SIoT is a very active research field, and many scholars have classified the routing algorithms according to different perspectives and starting points. In SIoT, the path between nodes is still being determined, greatly limiting the forwarding of messages. The original Direct Delivery Routing (DDR) algorithm [8] allows the node to carry the message and only forwards them when they encounter the destination node. Therefore, it produces only some copies of the message. It has the advantage of low network overhead and low average latency. However, its message delivery rate is low. In 2000, researchers proposed the Epidemic Routing (ER) [9] algorithm based on flooding ideas to address its low delivery rate. It adopts the transmission mechanism of copying when meeting and indiscriminate direct diffusion of information to the network, similar to the "contagion" model. Undeniably, it has the highest transmission rates but also pays the highest transmission costs. Researchers also proposed the Prophet Routing (PR) [10] algorithm to reduce the routing overhead based on historical encounters and information transfer probability. The algorithm uses historical encounter information to predict the possibility that nodes will meet soon. Copy data packets to nodes that are more likely to complete data transfer. In this routing protocol, the node does not blindly forward the information to the neighboring node. However, it calculates the prediction value of the neighboring node to the destination node in advance. The predicted probability determines whether the saved information should be forwarded. If the predicted value of the adjacent node is greater than that node, the message will be forwarded.

On this basis, researchers have begun to investigate the factors affecting node attributes. Popular social attributes currently include community structure, similarity, centrality, personnel flow pattern [11], social graph, and similarity. For example, the authors in [12] used similarity and centrality to improve the message delivery probability. While the work in [13] introduced the concept of community and defined the centrality between the same community and different communities to indicate the importance of nodes forwarding messages within and across communities. In [14], the authors measured social similarity by the activity of nodes, with nodes with higher activity being more likely to receive messages, thus increasing the message delivery rate. In another work, Deng et al. [15] measured social similarity in terms of relationships and attributes between nodes and the centrality of nodes in terms of their social connectivity. By combining these two functions, the optimal relay node is determined. The authors in [16] first explored the impact of three social features, namely physical proximity, user interests, and social relationship on users' daily routines. Then, they proposed a multi-dimensional routing protocol called Proximity-Interest-Social (PIS) protocol in which the three

different social dimensions are integrated into a unified distance function in order to select optimal intermediate data carriers. Similarly, authors in [17] proposed a mode in which users select or push patterns with similar hobbies to predict similar nodes and deliver messages. The work in [18] fully considered the sociability and mobility of nodes. Routing uses node activity as a measurement index for message forwarding in the same community. It selects a high contact frequency node with the destination node as a relay node for message forwarding. In [19], the authors sought to improve our understanding of human mobility in terms of social structures, and to use these structures in the design of forwarding algorithms for Pocket Switched Networks (PSNs). Taking human mobility traces from the real world, they discover that human interaction is heterogeneous both in terms of hubs (popular individuals) and groups or communities. The authors in [20] divided communities according to the social characteristics of nodes. Nodes are given priority in forwarding messages to nodes in the same community. For nodes in different communities, the intimacy of their nodes would be considered, and the social correlation would be used as a measurement index, effectively improving the message delivery rate. In [21], the authors combined the similarity of nodes in social networks, communication, and propagation factors and proposed a routing algorithm based on user-adaptive data transmission. It effectively utilized the edge nodes, divided them into several communities, reconstructed the community structure, and improved the information transmission rate. The work in [22] explored community and friendship as social features to effectively route messages. The nodes develop their community in a distributed way based on the concept of a familiar set. Merging of communities of nodes is governed by certain rules. Friendship in this paper evolves based on various factors like inter-contact duration, frequency of encounters and successful delivery of messages forwarded in a few past time intervals between the nodes. Routing of messages is accomplished by utilizing community and friendship as social characteristics. Moreover, the work in [23] proposed a mobile network routing query algorithm based on time-varying relation groups. It solves the problems of unstable node communication and dependence on erroneous relay nodes in the network. The authors introduced selfish nodes in [24,25] with improvements in detection and motivation, effectively improving the efficiency of node forwarding. Another work [26] proposed an encryption algorithm that effectively ensures data security when transmitting messages between nodes.

However, many studies have not yet fully considered the social attributes of the nodes and are more or less deficient in measuring the relevant metrics, which do not correctly reflect the relationships between the nodes. The routing performance will be significantly improved if we can fully use nodes' and communities' social attributes and relations and formulate reasonable metrics.

1.2 Motivation and Contribution

With the rapid development of social networks and the widespread use of mobile devices, especially smartphones, have made effective data communication a crucial component of the SIoT. As Internet consumers and demand for various applications continue to rise, the importance of efficient network connection, forwarding, and interruption increases. The SIoT, which derives from the Delay Tolerant Network (DTN), addresses the dearth of communication infrastructure in wireless networks and improves network environments with frequently changing connectivity and interruption between nodes. However, data forwarding remains a significant challenge in the IoT, necessitating the development of novel approaches to enhance the efficiency of information transmission.

Our proposed article presents a novel data forwarding scheme, the algorithm based on Multidimensional Social Relations (MSRR) in SIoT. The algorithm optimizes the message-forwarding process by utilizing the social attributes and node connections. The proposed algorithm introduces three critical metrics: the intensity of node social relationships, node activity, and community connection.

The algorithm divides nodes into communities based on their social attributes and identifies the most efficient relay nodes for message forwarding by analyzing these metrics.

The following are the main contributions of this article:

- We analysed the diverse applications of SIoT in today's society as well as future trends and summarise existing work.
- We outline a novel SIoT data forwarding algorithm based on multidimensional social relationships. The algorithm takes into account the multidimensional information such as the node's encounter time, encounter frequency, interest characteristics, euclidean distance and so on. And it combines them linearly to select more suitable relay nodes for information forwarding, which is proved from the theory that it can effectively improve the efficiency of information transmission.
- We classify message forwarding into intra-community forwarding and inter-community forwarding. And the connectivity between communities is analysed. It ensures communication relaying through nodes with stronger social connections and higher activity levels.
- We compare the proposed algorithm's performance to three existing routing algorithms through simulation. Results indicate that the proposed algorithm significantly improves message delivery efficiency, reduces network overhead, and enhances connectivity in the SIoT context.

This article thoroughly analyzes the concept of SIoT and its various applications. Section 1 summarizes related works on data-forwarding algorithms to set the stage for the proposed algorithm. Section 2 provides a clear comprehension of the underlying framework by introducing the system model and assumptions that form the basis of the study. Moving on to Section 3, the article provides a detailed explanation of the algorithm's implementation process. This section explores the algorithm's complexities, emphasizing its key characteristics and describing how it addresses the challenges of data forwarding in SIoT. In addition, the article compares the Multidimensional Social Relationships algorithm to other frequently used algorithms in Section 4. In this evaluation, the performance of various algorithms is compared, and experimental results are presented to demonstrate that the proposed algorithm is superior. The paper's conclusion is Section 5, which summarizes the study's findings and discusses their implications.

2 System Model and Assumptions

In the SIoT, a network node with a community structure tends to select objects for message transmission according to their social attributes. The community division according to the nodes' social attributes and interests will make the nodes' movement more regular. Nodes with the same properties communicate more frequently. Moreover, they are more likely to cluster in the same area. Their transmission efficiency will be higher.

This community structure in SIoT divides the source and destination nodes into two cases, within and outside the community, according to their location. Therefore, the message forwarding of nodes can be divided into two stages: forwarding within the community and forwarding across the community. Fig. 1 illustrates how a node forwards a message.

Therefore, we define three measurement indicators: "node social relationship strength," "node activity," and "community connection degree." When a node forwards a message in a community, the node's social attributes and the node's activity in the community are considered as the criteria to measure the relay capability of the node. When a node forwards messages across communities, the

connectivity between the source and destination communities is considered the standard to measure the relay capability of nodes.
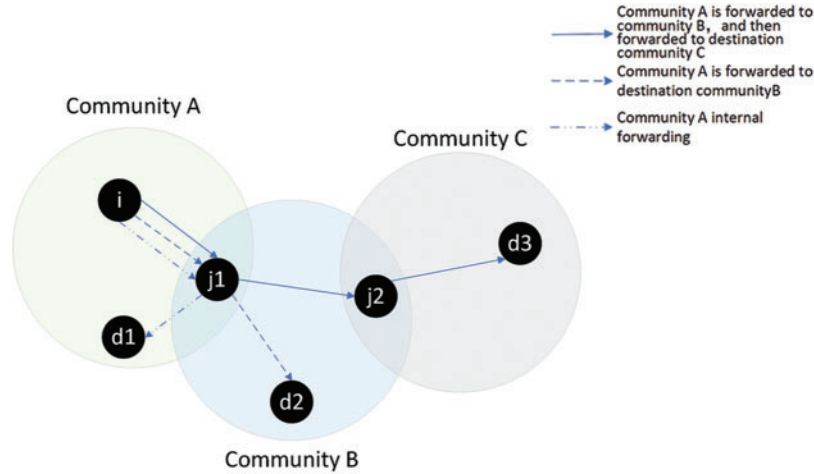

Figure 1: Message forwarding process

2.1 Definitions and Assumptions

For the convenience of description, we use the following symbols to describe this (as shown in Table 1). And make the following assumptions according to the need:

1. We represent the relationship between nodes in SIoT as an undirected graph $G = (V, E)$, where N is the total number of nodes, E is the set of edges, and V is the set of nodes. When two nodes meet and perform message forwarding, an edge is formed between them. $E(i, j)$ is denoted as the number of connections between nodes, $i \neq j$, but i and j can swap positions.
2. We abstract the relationship between communities as an undirected graph $G'(V', E')$, where E' is the collection of community edges and V' is the collection of communities. We refer to the community that contains node i as C_i and the community that contains node j as C_j. The number of connected nodes in C_i and C_j is denoted as $E'(C_i, C_j)$, and each connected node is counted only once. If the node is located in the overlap region of C_i and C_j, then $i = j$, but C_i and C_j are represented as two different communities and $E'(C_i, C_j) \neq 0$.
3. Anyone node in the SAN network belongs to at least one community. A community has multiple nodes. There may be intersections between communities, so the muster of communities containing node i is defined as $C(i)$. K_i is denoted as the sequence number of the community and $C(i) = \{C_{Ki}(i) \mid K_i \in \{1,2,3\ldots\}\}$. We assume each node knows its community and whether the destination node is in the same community when delivering the message.
4. We assume that nodes are always in motion in the network and unconditionally help to forward messages to other nodes.

Table 1: Main symbols

Symbol	Description
$G(V, E)$	An abstract graph of network nodes

(Continued)

Table 1 (continued)

Symbol	Description
$G'(V', E')$	An abstract graph of community structure
t_n	The duration of the node's nth encounter
I	Node's interest features
N	Total number of nodes in the network
N^i	The total number of nodes in the community containing node i
$i \& j$	The node i and j in the network
$i' \& j' \& i'' \& j''$	Any node in a network that satisfies the condition of the formula they are in
$E(i, j)$	Connection times between node i and node j
$E'(C_i, C_j)$	Connection times between two communities
K_i	The serial number of the community where node i is located
$C(i)$	All communities containing node i
C_i	Community containing node i

2.2 Social Relationship Model

Social relationships between nodes in the SIoT can be expressed in various ways, such as node interest attribute, node encounter frequency, node social context information, node encounter duration, etc. If two nodes have the same interest attribute and the encounter lasts for a long time, their communication is closer, and the social relationship is better. Therefore, through much research, this paper combines the interest attribute between nodes and the connection strength proportionally to measure node social relationships.

Definition 1: T (i, j). In one interaction cycle, the total encounter duration of node i and node j. Fig. 2 shows the time series of historical encounters of node i and node j. t_n represents the duration of node i and node j at the nth encounter, i.e.,

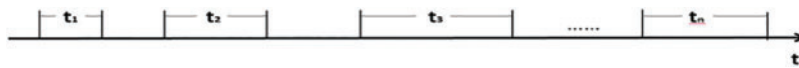$$T(i,j) = \sum_{k=1}^{n} t_n. \tag{1}$$

Figure 2: Mode encounter history timing table

Definition 2: $F_{(i)}$. Represents the number of connections between i and other nodes in an interaction cycle, i.e.,

$$F_{(i)} = \sum_{j \in N(i)} E(i,j), \tag{2}$$

where $E(i, j)$ denotes the number of connections between node i and node j. $N(i)$ represents the collection of other nodes encountered by node i in the current cycle.

Definition 3: $L_{(i,j)}$. According to formulas (1) and (2), the connection strength between nodes, i.e.,

$$L_{(i,j)} = \frac{E(i,j) \, T_{(i,j)}}{F_{(i)} \sum_{x \in N(i)} T_{(i,x)}}. \tag{3}$$

Definition 4: Define an n-dimensional hobby attribute vector to represent the hobby characteristics of node i, i.e.,

$$IA_i = \{I_1, I_2, I_3, \ldots, I_N\}, \tag{4}$$

where IA_i indicates the historical preference attribute of node i. Since the value range of each dimension attribute may differ, we normalize the data here so that the value range of each dimension attribute is controlled between [0,1].

Definition 5: According to the interest attribute vector of node i and node j. We can figure out the Euclidean distance between two nodes, i.e.,

$$D_{(i,j)=} \sqrt{\sum_{k=1}^{n} (I_k - J_k)^2}. \tag{5}$$

Definition 6: The more similar the interests of two nodes are, the more similar the two nodes are, and the smaller the Euclidean distance between them. The similarity of the two nodes is inversely proportional to the Euclidean distance. The interest similarity of node i and node j is expressed as

$$S_{(i,j)} = \frac{1}{1 + D_{(i,j)}}. \tag{6}$$

Definition 7: The social relationship strength between nodes is a linear combination of node connection strength and node similarity, i.e.,

$$SRS_{(i,j)} = \alpha L_{(i,j)} + \beta S_{(i,j)}, \tag{7}$$

where α and β are the weights, which represent the importance of node connection strength and similarity, respectively. They take values dynamically according to the actual situation of simulation parameter settings. In this paper, all the weights are adopted (considering that the implementation environment of the algorithm and the focus of the study may be different, in order to achieve the expected results, the researcher needs to set the weights flexibly according to the actual situation).

2.3 Node Activity Model

A node's activity reflects a node's ability to forward messages. The more times a node meets other nodes in a period, the higher its activity and the stronger its relay ability. When two nodes belong to the same community, the frequency of their encounters will increase. However, when two nodes are not in the same community, the frequency of encounters will be lower. Therefore, we will choose different metrics to calculate depending on whether the source node is in the same community or not.

Definition 8: The activity in the community is the ratio of the total number of encounters between node i and the local community nodes to the total number of encounters between all nodes in the local community, i.e.,

$$SI_{(i)} = \sum_{k \in K_i} \frac{\sum_{(i,i') \in C_k(i)} E(i,i')}{\sum_{(i',j') \in C_k(i)} E(i',j')}. \tag{8}$$

Definition 9: Out-of-community activity refers to the ratio of the sum of encounters of node i with nodes outside its local community to the sum of encounters of all nodes in other communities (communities adjacent to node i), i.e.,

$$SO_{(i)} = \sum_{k \in K_i, k' \in K_i'} \frac{\sum_{i \in C_k, i'' \in C_{k'}} E(i, i'')}{\sum_{(i'', B'') \in C_{k'}} E(i'', j'')}. \tag{9}$$

2.4 Community Connection Model

In the SIoT, nodes are not unlimited in activity. When nodes are not in the same communication range as each other, messages must be forwarded to the destination node via a relay node. This is when multiple communities are involved. To enable messages to be forwarded across communities, community connectivity is defined as a measure of the association between communities.

Communities can be divided into overlapping and independent relationships, depending on their different locations in relation to each other. As shown in Fig. 1. There are nodes in an overlapping community that belong to two or more different communities. Nodes in this zone are more likely to be relay nodes for cross-community messaging. The more nodes in common between two communities, the more connected the two communities are. In a separate community, there is no common node. Therefore, selecting a more suitable relay node can improve the delivery rate of nodes across community transitions, with due consideration of community overlap relationships.

Definition 10: Community connectivity. It consists of the relationship between familiar and unfamiliar communities. The relationship between communities is expressed as the ratio of the number of nodes intersecting between communities to the number of nodes merged between communities. The relationship between unfamiliar communities is expressed as the ratio of the number of nodes connected between communities to the sum of the number of nodes between communities, i.e.,

$$CC_{(C_i, C_j)} = \sum_{i \in K_i, j \in K_j} \left(\gamma \left| \frac{C_i \cap C_j}{c_i \cup C_j} \right| + \delta \left| \frac{E'(C_i, C_j)}{N^i + N^j} \right| \right). \tag{10}$$

3 Data Forwarding Algorithm

Message delivery is the most critical part of the SIoT. Traditional networks have stable end-to-end connectivity. However, the topology in SIoT is unstable; the messages are transmitted in the "storage-carrier-forward" mode. In this transmission mode, the source node first carries the message through the network in motion. The message is then forwarded to a relay node. Finally, the relay node passes the message to the destination node. Therefore, choosing a suitable relay node is the first task of passing messages in SIoT.

Through the comprehensive evaluation of the above "node social relationship strength," "node activity," and "community connection degree," the multidimensional social relation data forwarding algorithm is designed. Where the strength of node social relationship takes into account information such as encounter duration, number of encounters, interests and Euclidean distance between nodes and is determined by a linear combination. The larger the value, the stronger the social relationship between nodes and the more efficient the information transfer. Node activity is determined by considering how often the node meets with other nodes inside and outside the community. The higher the node's activity, the more frequently it communicates with other nodes, the stronger the relay capability, and the more efficient the information transmission. Community connectivity reflects whether a node is in multiple

communities. If a node is in more than one community, the node is more efficient in transmitting information across communities. By comparing the forwarding efficiency of each node, better relay nodes are selected to improve the delivery rate of the node.

3.1 Forwarding Strategy

Definition 11: Intra-community forwarding *CIR*. In intra-community forwarding, the social relations and community nodes' activities should be considered. *CIR* is a linear combination of the node's social relationship strength and the node's activity in the community, i.e.,

$$CIR_{(i)} = \varepsilon SRS_{(i,j)} + \zeta SI_{(i)}. \tag{11}$$

Definition 12: Cross-community forwarding needs to consider the node activity outside the community and community connectivity. *COR* is a linear combination of the node's community connectivity and the node's activity outside the community, i.e.,

$$COR_{(i)} = \eta CC_{(c_i, c_j)} + \theta SO_{(i)}. \tag{12}$$

where ε, ζ, η and θ represent the weight coefficient, respectively.

3.2 Algorithm Description

Based on the above analysis, we propose a new model. The model is optimised for analysing the attributes and social behaviour of nodes in SIoT. It makes it more suitable for analysing SIoT.

The encounter duration and the number of encounters of nodes reflect whether the communication between nodes is close or not, and nodes with closer communication have better social relationship and more efficient message delivery. The Euclidean distance obtained through the defined attributes of nodes' interests and hobbies reacts to whether two nodes are similar or not, the smaller the Euclidean distance, the more similar the interests between the nodes, the higher the rate of contact, and the higher the message delivery efficiency. Combining them linearly further filters out the nodes with high transmission efficiency in the network. In particular, since most of the devices in SIoT are carried by users. Therefore the social properties of nodes are consistent with those of humans. Social competence and social diversity can reflect the social attributes of nodes.

The MSRR algorithm is thus proposed based on the above assumptions and model description. First, the nodes are divided into different communities. Then determine whether the source and target nodes are in the same community. If yes, it will select the node within the community that is most similar to it and has a high activity level as the relay node. Otherwise, it selects the community with the highest connectivity as the relay community. Finally, it continues to select the most appropriate relay node within the relay community. Regardless of whether the relay node is in the community or not, as long as it is a suitable next-hop node, the node will forward the message in the shortest time, avoiding the problem that the TTL of the message expires in advance due to too long a stay time, which significantly improves the improved message forwarding efficiency.

Fig. 3 shows the message-forwarding process of the algorithm, which proceeds in the following steps:

(1) Firstly nodes A and B meet in the network. And determine whether node B is the destination node or not. If B is the destination node then A forwards the message to B.

(2) If B is not the destination node, it will be determined whether node A and the destination node are in the same community.

(3) If A and the destination node are in the same community, will determine if B is in the same community as the destination node.

If B and the destination node are in the same community, will judge the size of CIR(A) and CIR(B). If CIR(A)$ < $CIR(B), then A will forward the message to B. Conversely A will continue to carry the message for movement.

If B and the destination node are not in the same community, A will carry the message to continue the movement.

(4) If A and the destination node are not in the same community. It will be determined if B and the destination node are in the same community.

If B and the destination node are in the same community, A forwards the message to B. If B and the destination node are not in the same community, will determine the size of COR(A) and COR(B).

If COR(A)$ < $COR(B), A forwards the message to B. Conversely A will carry the message and continue the movement.
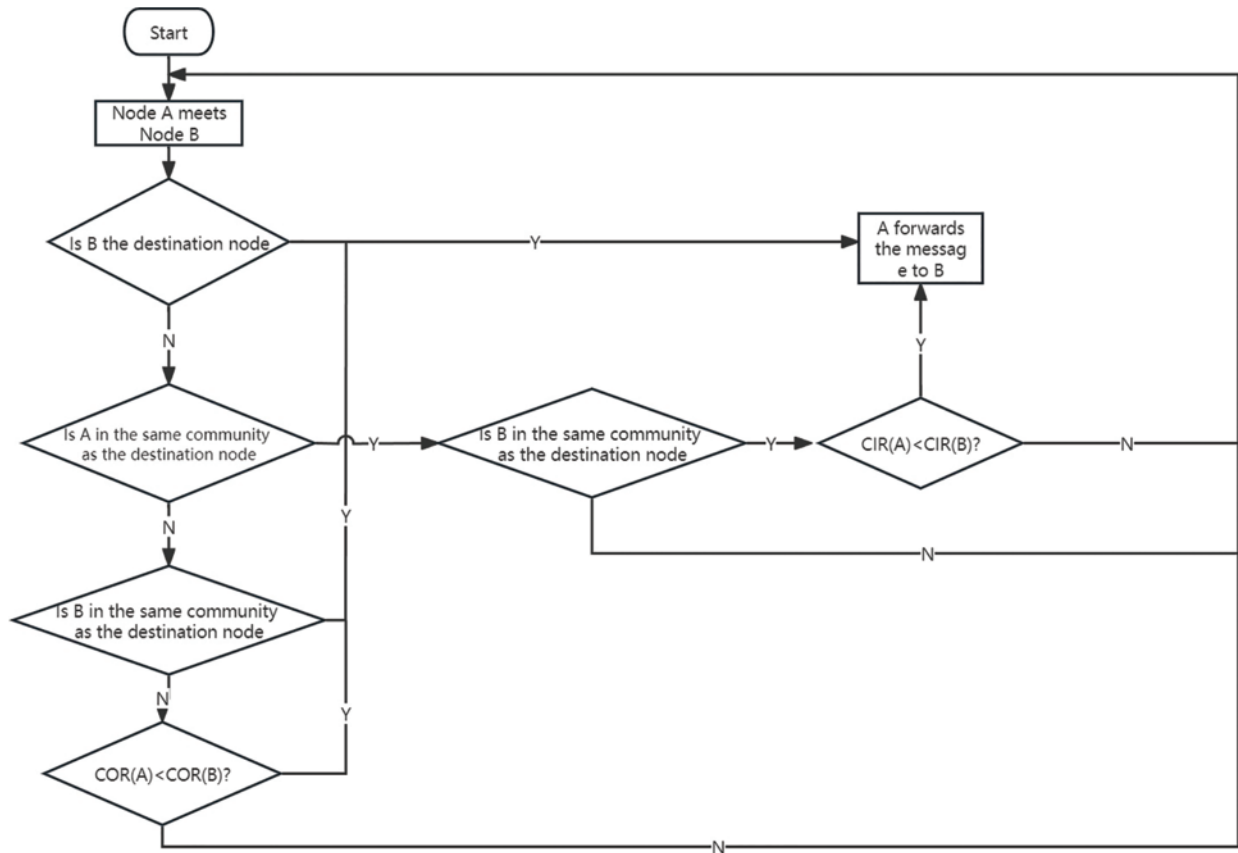
Figure 3: Algorithm flow chat

4 Simulation Experiment and Analysis

To compare and evaluate the performance of the MSRR algorithm, simulation experiments have been conducted in this paper using the ONE [27] simulator, which is designed for the DTN

environment. It takes advantage of the differences between various routing policies. The DTN packet forwarding process is reproduced. It can more realistically track and analyze the operation of complex routed networks. The hardware environment is as follows: CPU model is i7-10870H (2.2 GHz), disk size is 1 TB, GPU is NVIDIA GeForce RTX 3060 6 GB, etc. The software environment is as follows: The IntelliJ IDEA supports one. The JDK version is 1.8. The ONE version is 1.5.1. Our research combines nodes' direct and latent social relationships or attributes to improve routing performance. Therefore, we will ignore the detailed effects of weights; all weights are set to 1 in this paper.

This paper compares the MSRR algorithm with three other routing algorithms: DDR, ER, and PR. These four algorithms' delivery rates, network overhead rates, and average latencies are compared for different simulation durations and node buffer sizes.

4.1 Performance Evaluation Index

The optimization goal of the SIoT routing algorithm is to improve the message transmission rate, reduce transmission delay and network overhead. Therefore, three metrics will be used to evaluate the algorithm's performance in this paper.

Delivery rate: The ratio of successfully forwarded messages to the total messages forwarded. The formula is calculated as follows:

$$deliver_{ratio} = \frac{delivery}{all\ messages}. \tag{13}$$

Network overhead: The network overhead is the ratio of the difference between the total number of messages forwarded and the number of messages successfully delivered to the total number of messages successfully delivered. The formula is calculated as follows:

$$overhead_{ratio} = \frac{relayed - delivered}{delivered}. \tag{14}$$

Average latency: Average latency represents the total time it takes for a message to be delivered to the destination node from when it is sent. The formula is calculated as follows:

$$latency_{average} = \frac{\sum each\ Message_{delivered_{times}}}{delivered}. \tag{15}$$

4.2 Simulation Environment and Parameter Settings

In order to make the experimental data more realistic and convincing as well as to better simulate the real and natural network environment. We refer to a large number of literatures and synthesise the information of their experimental parameters. The various parameters of the experiment are finally determined as shown in Table 2.

4.3 Experimental Results and Analysis

To show the efficiency of our proposed data forwarding algorithm, we compare the performance of our algorithm, labeled as MSRR, with the following algorithms:

1. Direct Delivery Routing (DDR): It allows the node to carry the message and will not forward it until it meets the destination node. Its network overhead and delay are very low, but the delivery rate is also very low.

2. Epidemic Routing (ER): It indiscriminately transmits messages when nodes meet and directly diffuses messages to the network. It has a high transmission path but pays the highest transmission cost.
3. Prophet Routing (PR): It analyzes nodes' historical encounter information, predicts the next hop node, and transmits the data to nodes more likely to meet.

Table 2: Experimental parameter

Parameter	Parameter value
Network topology area	4500 m $*$ 3400 m
Mobile model	Shortest Path Map Based Movement
Node communication method	Bluetooth
Signal Transmission Radius	10 m
Message generation interval	25 to 35 s
Message size	500 KB to 1000 KB
Number of nodes	126
Number of communities	6
Transfer speed	250 KB/s
Cache size	50 M
Message lifetime	300 min
Simulation time	8 h

The parameter settings of DDR, ER, and PR involved in the comparative experiment are consistent to make the experiment more realistic. Each simulation was run 10 times, and the data was then analyzed separately for statistics. We prove the algorithm's performance by changing the simulation time and node cache size to make the results more convincing.

4.3.1 Simulation Effect of Simulation Time on Algorithm Performance

In this experiment, the node cache size is set to 50 M, and the experiment time is set to 2, 4, 6, 8, and 10 h, respectively.

As shown in Fig. 4, as time increases, the message delivery rate of DDR is always at the lowest level. The source node only forwards the message once it encounters the destination node. However, many nodes are in a complex opportunistic network. It is difficult for the source node to forward the message directly to the destination node. Therefore, the single-process working mode of DDR leads to low delivery efficiency. The message delivery rate of PR shows an overall upward trend with the increase in time, and the increase is most evident at 4 h, which is at a medium level among the four algorithms. The message delivery rate of ER is at the highest level at 2 h, and it shows an upward trend with the increase of time. The principle of the algorithm is similar to the way epidemics spread. The message will be forwarded as long as the nodes meet each other. The unlimited forwarding of the message dramatically increases the possibility of the message reaching the destination node. The MSRR algorithm is moderate at 2 h but increases significantly at 4 h. As time increases, the algorithm has been in a steady growth state, even surpassing the ER. From the simulation, the MSRR algorithm is the best in message delivery rate. The algorithm fully considers the preference attributes of nodes and

the social relations between nodes. Moreover, it can better select relay nodes, significantly improving the message delivery rate.
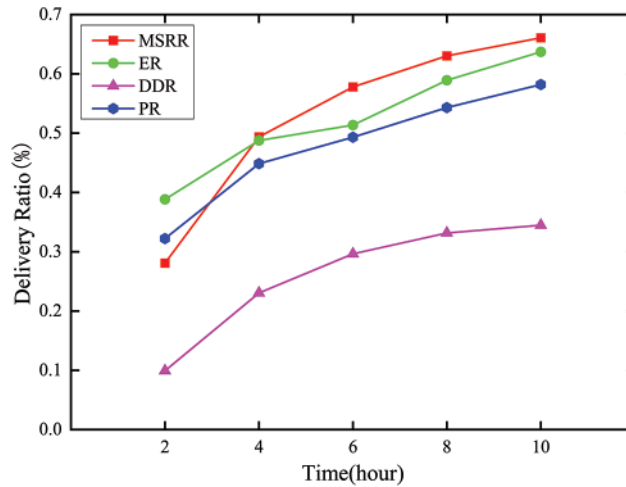

Figure 4: Delivery rate under different simulation times

As shown in Fig. 5, the DDR algorithm does not generate message copies as it does not perform message replication and hence its network overhead is always 0. Whereas the PR and ER algorithms replicate a large number of messages in the network and hence the network overhead becomes very high and the performance becomes unstable. The MSRR algorithm selects more appropriate relay nodes while forwarding messages and hence reduces the number of forwarded messages and the generation of message copies. Its network overhead ratio is always low and stable at 14%–15%.
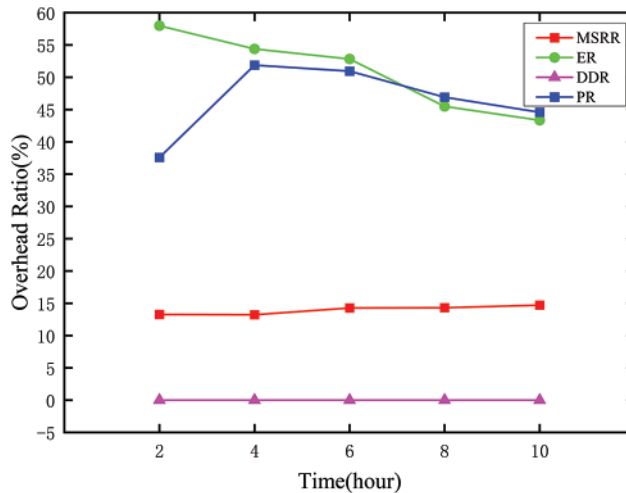
Figure 5: Network overhead under different simulation times

According to Fig. 6, the average latency of direct DDR is relatively low during 2–4 h as time increases. However, it has been at the highest level after 6 h because the source node will only forward the message when it encounters the destination node. This makes the message forwarding time very long. The average latency of ER and PR is at a high level of 2 h but has been consistently at a

low level from 4 to 8 h. The delay of ER is close to DDR in 10 h. And PR has always been at its lowest level. The average latency of the MSRR algorithm is at the lowest level at 2 h, and the delay is slightly higher at 4–8 h. However, from the perspective of the whole test cycle, the delay level has been at a moderately low level after 8 h. The average latency of the PR is the best, lower than the other three algorithms. Because the algorithm will predict the probability based on the historical encounter information of the node, which effectively reduces the waiting time and calculation amount and improves the forwarding efficiency. The average latency of the MSRR algorithm has a relatively stable upward trend. It rationally uses the social attributes of nodes to select the optimal relay node, significantly improving forwarding efficiency, and its average latency is slightly higher than the PR.
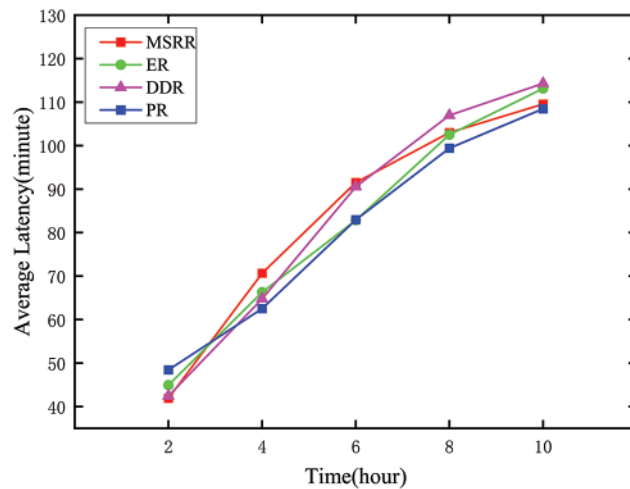

Figure 6: Average latency under different simulation times

4.3.2 Effect of Node Buffer Size on Algorithm Performance

In this experiment, we set the node buffer sizes to 5 M, 10 M, 15 M, 20 M, and 25 M, and the experiment time is 8 h. To better prove the algorithm's performance, we will combine the experimental results and data for comparative analysis.

As shown in Fig. 7, the delivery rate of DDR has been stable at around 33%. Its delivery rate does not change with the cache size. Because the algorithm does not duplicate messages, the delivery rate will not be affected regardless of changes in its cache size. As the cache size increases, PR and ER delivery rates increase steadily. Nevertheless, the delivery rate of the ER is slightly higher than that of the PR. The ER and the PR will generate many message copies during the message delivery process, so the impact of the cache size on its delivery rate is very obvious. The MSRR algorithm has the highest message delivery rate; even when the cache is small, it can have a high delivery rate. As the buffer size increases, the delivery rate increases significantly and is at the highest level.

Fig. 8 shows a comparison of the network overhead of the four algorithms for different cache sizes. In the case of smaller buffers, DDR incurs no network overhead since there is no relaying process. Whereas multi-replica routing such as PR and ER replicates a large number of message copies when the cache is small, resulting in network congestion and higher packet loss rate. Therefore the network overhead of these two algorithms is always at a high level. As the buffer area increases, their network overheads are mitigated. The MSRR algorithm fully considers the multidimensional social information of the nodes. It will select more suitable relay nodes for message forwarding, so that the

network overhead can be effectively controlled even with a small cache. The network overhead of the MSRR algorithm is significantly lower than that of the PR and ER algorithms, which has been stable at about 15%, but slightly higher than that of the DDR algorithm.
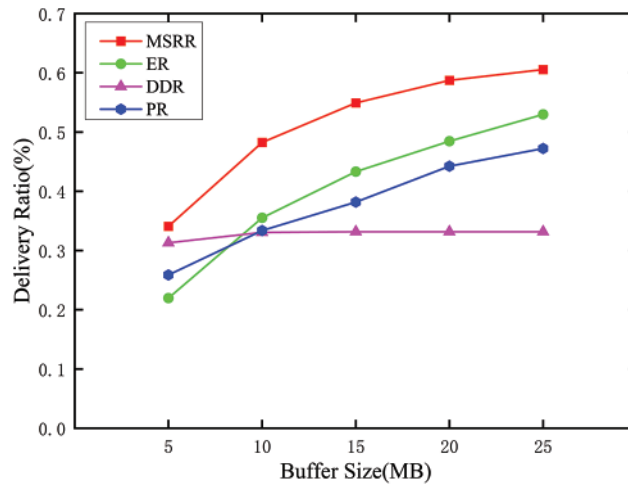

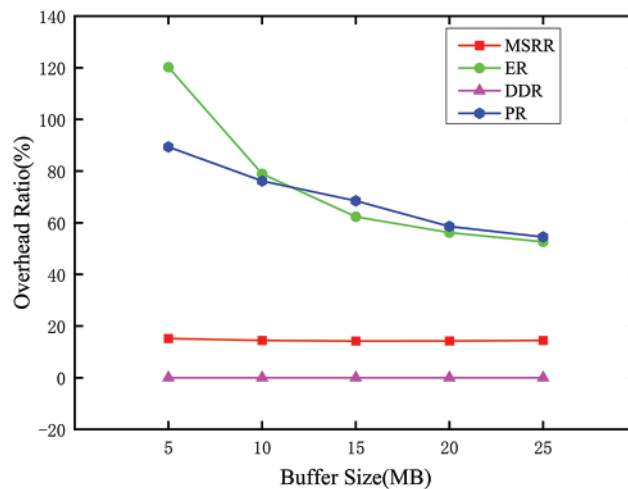**Figure 7:** Delivery rate under different cache sizes

Figure 8: Network overhead at different cache sizes

As shown in Fig. 9, the average delay of DDR has always been at the highest level because the source node of this algorithm will only forward the message when it encounters the destination node, so the waiting time is the longest. Its average delay is always the highest and varies small. The average delay of PR and ER increases steadily as the cache size increases. However, the average delay of the MSRR algorithm is higher than that of PR and ER because each node in the ER participates in the delivery of messages, significantly reducing the waiting time. PR uses node history information to predict node prediction probability, saving much time. The MSRR algorithm comprehensively considers factors such as the node's hobby attribute and the node's social relationship so that the relay node's selection is more targeted and the calculation is more complicated, which is why the average delay of the MSRR algorithm is higher.
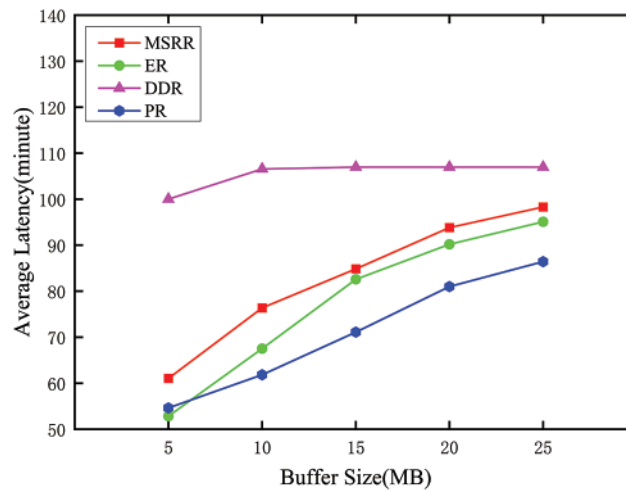
Figure 9: Average latency at different cache sizes

Through the analysis of the above simulation experiment data, the overall performance of the MSRR algorithm proposed in this paper is better than that of the other three algorithms. While maintaining a high message delivery rate, it reduces network overhead and average latency. The MSRR algorithm comprehensively considers the node's hobby attribute and social relationship in selecting relay nodes. At the same time, it considers the relationship between different communities when forwarding messages across communities. The other three algorithms consider too single factors and only perform better in a specific data item.

5 Conclusions

This paper concludes by presenting a novel data forwarding algorithm for the SIoT that leverages node attributes and community relations. By analyzing nodes' interest attributes and social relationships, the algorithm effectively considers various factors to select relay nodes within and between communities. The algorithm considers the node's interest attributes and encounter time within the community to identify appropriate relay nodes for message forwarding, thereby promoting efficient communication. Similarly, when forwarding messages between various communities, external node activity and inter-community connectivity are considered to select relay nodes that facilitate collaboration. The experimental results show that the MSRR algorithm has the highest message transfer rate under almost all conditions. Because it makes reasonable use of the social attributes and social relationship information of nodes, only the average delay is slightly higher than ER and PR, and the network overhead is slightly higher than DR. The forwarding pattern of the ER algorithm is similar to the message forwarding pattern of "contagion", which greatly reduces the waiting time. The PR algorithm predicts the encounter probability based on the node's history information, which saves a lot of time. The DR algorithm will forward the message only when it encounters the destination node. So its network overhead is always 0. The MSRR algorithm needs to analyse multi-dimensional information such as the node's own attributes as well as the complex social relationship between the nodes to choose a more suitable relay node. This increases the average delay and network overhead to some extent. However, it also results in a higher message delivery rate. Overall, the MSRR algorithm only sacrifices a little average delay and network overhead, but gains a huge information transfer rate. Therefore, it can provide the global optimal solution. Future efforts will concentrate on further

optimizing the algorithm to reduce the average delay while maintaining a high delivery rate. These optimization efforts will improve the algorithm's efficacy and the SIoT's overall message-forwarding efficiency.

Author Contributions: Conceptualization, Fang Xu, Songhao Jiang, Yi Ma, Manzoor Ahmed; Methodology, Fang Xu, Songhao Jiang; Simulations & Validation, Fang Xu, Songhao Jiang; Formal analysis, Zenggang Xiong, Yuanlin Lyu; Writing-review & editing, Fang Xu, Songhao Jiang, Yi Ma, Manzoor Ahmed; Funding acquisition, Fang Xu. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Materials: No new data and materials were created or analyzed in this study. Data and material sharing are not applicable for this article.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Z. Cai and J. Wu, "Efficient content transmission algorithm based on multi-community and edge-caching in ICN-SIoT," *Peer-to-Peer Networking and Applications*, vol. 73, no. 1, pp. 277–294, 2023.

[2] Q. Y. Tan, D. D. Liu and J. Zhang, "Research on adaptive congestion control mechanism in delay tolerance networks," *Computer Engineering and Applications*, vol. 54, no. 11, pp. 109–115, 2018.

[3] H. T. Zhao, H. Cheng, Y. Ding, H. Zhang and H. Zhu, "Research on traffic accident risk prediction algorithm for vehicle connected edge network based on deep learning," *Journal of Electronics and Information Technology*, vol. 42, no. 1, pp. 50–57, 2020.

[4] S. Alam, S. Zardari, S. Noor, S. Ahmed and H. Mouratidis, "Trust management in Social Internet of Things (SIoT): A survey," *IEEE Access*, vol. 10, pp. 108924–108954, 2020. https://doi.org/10.1109/ACCESS.2022.3213699

[5] M. C. Zhou, "Evolution from AI, IoT and big data analytics to metaverse," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 12, pp. 2041–2042, 2022.

[6] M. A. Javed, M. S. Younis, S. Latif, J. Qadir and A. Baig, "Community detection in networks: A multidisciplinary review," *Journal of Network and Computer Applications*, vol. 108, pp. 87–111, 2018. https://doi.org/10.1016/j.jnca.2018.02.011

[7] Q. Zou, Y. Li, X. Yang and Z. Zhou, "Identification of key nodes in directed network based on overlapping community structure," *Automatic Control and Computer Sciences*, vol. 55, no. 2, pp. 167–176, 2021.

[8] M. Abdelmoumen, E. Dhib, M. Frikha and T. Chahed, "How to improve the performance in delay tolerant networks under Manhattan mobility model," in *21st Annual IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications*, Istanbul, Turkey, IEEE, pp. 2008–2013, 2010.

[9] G. Liu and Y. M. Li, "Research on spread routing protocol for delay tolerance networks," *Measurement and Control Technology*, vol. 37, no. 12, pp. 62–65, 2018.

[10] R. Zhu, D. Gu, S. Li, S. Wei and Y. Huang, "Optimization and simulation of prophet routing algorithm for DTN networks," *Computer and Digital Engineering*, vol. 49, no. 11, pp. 2237–2241+2286, 2021.

[11] S. Tsugawa, "A survey of social network analysis techniques and their applications to socially aware networking," *IEICE Transactions on Communications*, vol. 102, no. 1, pp. 17–39, 2019.

[12] J. W. Chang and C. Chen, "CROP: Community-relevance-based opportunistic routing in delay tolerant networks," *IEICE Transactions on Communications*, vol. 97, no. 9, pp. 1875–1888, 2014.

[13] J. Niu, X. Zhou, Y. Liu, L. Sun and J. Ma, "A message transmission scheme for community-based opportunistic network," *Journal of Computer Research and Development*, vol. 46, no. 12, pp. 2068–2075, 2009.

[14] Z. Li, C. Wang, S. Yang, C. jiang and X. Li, "Lass: Local-activity and social-similarity based data forwarding in mobile social networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 1, pp. 174–184, 2014.

[15] M. Deng, F. Xu, Z. Xiong, C. Ye and H. Xia, "Social-aware data forwarding algorithm in mobile social networks," *Computer Engineering and Design*, vol. 39, no. 7, pp. 1835–1840, 2018.

[16] F. Xia, L. Liu, B. Jedari and S. K. Das, "PIS: A multi-dimensional routing protocol for socially-aware networking," *IEEE Transactions on Mobile Computing*, vol. 15, no. 11, pp. 2825–2836, 2016.

[17] J. Wu, G. Yu and P. Guan, "Interest characteristic probability predicted method in social opportunistic networks," *IEEE Access*, vol. 7, pp. 59002–59012, 2019.

[18] W. Wang, "Research on message transmission strategy and caching mechanism in mobile social networks," M.S. dissertation, Central China Normal University, China, 2018.

[19] P. Hui, J. Crowcroft and E. Yoneki, "Bubble rap: Social-based forwarding in delay tolerant networks," in *Proc. of the 9th ACM Int. Symp. on Mobile Ad Hoc Networking and Computing*, Hong Kong, China, pp. 241–250, 2008.

[20] Y. Xie, "Research on high performance message transmission algorithm and cache management mechanism in mobile social networks," M.S. dissertation, Central China Normal University, China, 2019.

[21] Y. Lu, C. Liu, J. Luo and J. Wu, "Routing algorithm based on user adaptive data transmission scheme in opportunistic social networks," *Electronics*, vol. 10, no. 10, pp. 1138, 2021.

[22] M. R. Penurkar and U. A. Deshpande, "Social characteristics-based routing algorithm for a mobile social network," *Computing*, vol. 103, no. 1, pp. 133–153, 2021.

[23] Y. Dong, C. Liu, J. Luo and J. Wu, "A routing query algorithm based on time-varying relationship group in opportunistic social networks," *Electronics*, vol. 10, no. 13, pp. 1595, 2021.

[24] Z. Xiong, N. Xiao, F. Xu, X. Zhang, Q. Xu et al., "An equivalent exchange based data forwarding incentive scheme for socially aware networks," *Journal of Signal Processing Systems*, vol. 93, pp. 249–263, 2021.

[25] N. Xiao, Z. Xiong, F. Xu, Q. Xu, X. Zhang et al., "A diversity-based selfish node detection algorithm for socially aware networking," *Journal of Signal Processing Systems*, vol. 93, no. 7, pp. 811–825, 2021.

[26] M. Revanesh, J. M. Acken and V. Sridhar, "DAG block: Trust aware load balanced routing and lightweight authentication encryption in WSN," *Future Generation Computer Systems*, vol. 140, pp. 402–421, 2022. https://doi.org/10.1016/j.future.2022.10.011

[27] Z. Wang, X. H. Wang and J. Q. Sui, "Opportunity network simulator ONE and its extension research," *Computer Application Research*, vol. 29, no. 1, pp. 272–277, 2012.