IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 29, 2019, accepted April 29, 2019, date of publication May 2, 2019, date of current version May 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2914505

An Effective Transmission Strategy Exploiting Node Preference and Social Relations in Opportunistic Social Networks

YEQING YAN, ZHIGANG CHEN, (Member, IEEE), JIA WU, (Member, IEEE), LEILEI WANG, KANGHUAI LIU, AND PENG ZHENG

China Mobile Joint Laboratory, School of Computer Science and Engineering, Central South University, Changsha 410075, China

Corresponding authors: Zhigang Chen (czg@csu.edu.cn) and Jia Wu (jiawu5110@163.com)

This work was supported in part by the Major Program of the National Natural Science Foundation of China under Grant 61672540, in part by the National Natural Science Foundation China under Grant 71633006, in part by the China Postdoctoral Science Foundation under Project 2017M612586, in part by the Postdoctoral Science Foundation, Central South University, under Grant 185684, and in part by China Mobile Joint Laboratory, Ministry of Education, through Mobile Health.

ABSTRACT With the development of network technology and the advent of 5G communication era, high-frequency radio waves with high bandwidth will become the main choice of signal sources. While promoting the development of a high-speed communication network, a high-frequency radio wave has the problem of limited coverage, which makes opportunistic social network become the mainstream communication method. Since the transmission of a large amount of data in a short time will cause the problem of data redundancy, opportunistic social networks suggest that the most appropriate next hop should be selected to achieve efficient data transmission. At present, there are several routing algorithms based on social relations, which attempt to select the most suitable next-hop node among neighbor nodes by making use of relevant context information and historical interaction between nodes. However, existing data transmission methods in opportunistic social networks mainly focus on the influence of a few social attributes on the similarity between nodes but ignore the transmission preference caused by individual characteristics of nodes. To improve the transmission efficiency, this paper establishes an effective data transmission strategy (ENPSR) exploiting node preference and social relations in opportunistic social networks. In our scheme, individual transmission preferences are obtained by measuring the social attributes and historical information of nodes in the transmission process. The appropriate message delivery decision is determined by the prediction scheme, and the continuous and stable data transmission are realized through the recommendation mechanism. According to the simulation experiments, the average delivery ratio of ENPSR algorithm is 0.85, which is 20% higher than that of the epidemic algorithm.

INDEX TERMS Opportunistic social networks, node preference, routing, social relations, data transmission.

I. INTRODUCTION

As the development of web technology and people's demand, surfing online has become an indispensable part of people's life. The internet is an important tool for delivering messages and communication in people's daily life, which brings people great convenience [1], [2]. It has become an important part of people's lives to share video and photos via social platforms to get enough attention from groups. In this case, people's demand of network bandwidth and information storage requirements have been greatly improved, which gave birth to the arrival of the era of 5G [3].

Instead of low-frequency radio waves with small bandwidth and small amount of information, 5G networks use high-frequency radio waves [4]. Using high-frequency radio as signal source can greatly improve the speed of information transmission, which means that a large amount of data can be transmitted in a short time. High-frequency radio waves have large bandwidth and storage room for information, but are vulnerable to obstacles and have limited coverage. To solve this problem, 5G network proposes the concept of micro base stations, which are different from the existing large

The associate editor coordinating the review of this manuscript and approving it for publication was Nan Wu.

base stations. They are small and can be installed anywhere, alleviating the problem of limited high-frequency radio coverage [5]. With the help of micro-base station, device-to-device communication will become the main form of network communication.

5G communication can directly transmit data packets between devices, only need to inform base stations, greatly improving the efficiency of information transmission [6]. But compared with traditional D2D communication, the communication range of this new way is very limited, so the transmission of messages needs to be achieved through the movement of mobile devices [7]. The "storage-carry-forward" mechanism in opportunistic network is a perfect work structure for this new way of transmission. Due to the 5G communication is closely related to social life, we define this kind of network structure as the opportunistic social networks [8]–[11]. On account of the development of high-speed communication mode brought by 5G network, the transmission of a large amount of data in a short time will also bring data redundancy, and the problems brought by it need to be solved by appropriate opportunistic social network routing scheme [12].

However, current researches on opportunistic social network algorithms do not take into account the redundant impact of high-speed communication networks on data transmission, which lead to problems such as reduced data transmission efficiency, greatly increased network overhead and time delay [13]. The solution of these problems is to find a way to reduce the impact of data redundancy on message transmission in 5G environment. An effective way is to filter the nodes in the network and only use nodes that have a greater possibility of successful transmission to achieve forwarding. There are many schemes about how to select relay nodes in existing researches. As opportunistic social networks are closely related to social life, another problem is how to bring complex social relations into consideration of routing algorithms and propose the most effective scheme for selecting relay nodes. Due to human participation in network activities, the behavior of nodes has a large number of attributes, which represent the relationship between different users, making the indicators related to social forwarding multi-layer and multi-metrics [14]–[16]. Therefore, different algorithms based on social relations consider attribute similarity factors between nodes, which can play different roles in data transmission between nodes [17], [18]. In real social scenarios, data transmission between nodes is similar to commercial cooperation between humans, and the ability of partners should be evaluated before cooperation [19]. When a node evaluates its partner, different nodes have individual preferences for the partner's attributes in the transmission process due to their different social status and living environment. The transmission preference of nodes has a significant impact on data transmission strategy, which may affect the data transmission speed and efficiency of routing algorithm.

To solve these problems, an effective transmission strategy exploiting node preference and social relations is proposed in this paper. By measuring complex social attributes and historical information of nodes in data transmission process, this strategy proposes three attribute indexes that affect the data transmission ability of nodes. Through the measurement of attribute characteristics, the degree of individual preference in data transmission can be determined, and the concept of preference similarity is proposed. The determination of node preference similarity can make the transmission close to the real 5G communication environment in the maximum limit. Meanwhile, in order to solve the problem of data redundancy caused by mass data transmission in a short time under the 5G environment, this paper also proposes a relay node prediction and recommendation scheme based on matrix decomposition. This scheme not only creatively adds the concept of node preference, but also solves the problem of incomplete evaluation information of other nodes in the network, thus improving the recommendation accuracy of the algorithm. In conclusion, ENPSR is a novel routing-forwarding method, which takes into account the individual attribute preference degree of nodes in the transmission process.

The contributions of this paper are listed as follows:

1) In order to estimate node attributes during information transmission in 5G environment, three indexes of communication quality, residual cache and active level are proposed. These three indicators can be used to estimate the availability of nodes, which is very important for the effective information transmission process.
2) By effectively analyzing the attribute preference of nodes in the transmission process, we can define the individual preference of nodes as preference similarity, so as to measure their impact on information transmission.
3) To achieve efficient and reliable message transmission, we add neighbor information and preference similarity to the evaluation and prediction method of node scoring matrix. In this way, the effective transmission strategy proposed in this paper can be effectively transplanted into real 5G communication environment.
4) In accordance with the simulation results in the Opportunistic Networking Environment (ONE), we analyze the performance of this novel algorithm (ENPSR) and compared it with some other algorithms. Our algorithm shows enhanced performances in increasing the delivery ratio and reducing transmission delay and routing overhead.

The rest of this paper is organized as follows. In Section 2, we will give a brief introduction of existing works related. The system model will be proposed and analyzed in Section 3. In Section 4, we evaluate the performance of our proposal via extensive simulations. The conclusion of the paper is shown in the last section.

II. RELATED WORK

Due to the novelty of exploiting the broadcast and mobile characteristics of wireless communication and improving the transmission success ratio and channel utilization, opportunistic routings receive the favor of many researches, for different application scenarios. Consequently, in this section, we will give a brief introduction to the state-of-the-art of both social-aware and social-ignored routing algorithms in opportunistic networks.

A. EXISTING SOCIAL-IGNORANT ALGORITHMS

Social-ignorant algorithms mean that social information related to nodes is not used to make appropriate messaging decisions during data transmission. The Epidemic algorithm presented in literature [20] is a traditional opportunistic algorithm, and its data-transfer process is similar to the propagation process of infectious diseases. This algorithm has high packet delivery ratio, but its transmission delay and routing overhead is large because source node send messages to all nodes it encountered. Borah *et al.* [21] proposed an algorithm which calculates the forwarding probability and thus obtain the possible path of optimal routing. This algorithm has good performance in terms of packet delivery ratio, but its routing overhead is also high. Literature [22] proposed a flooding routing algorithm, the Spray and Wait algorithm, which divides the forwarding process for information into two steps. The first step is to copy messages and the transfer process is in the second phrase. This algorithm has high delivery ratio, but it can easily cause excessive transmission delay and data redundancy.

Wang *et al.* [23] proposed a routing algorithm based on node locations to predict the probability of meeting. This algorithm predicts the distance between nodes adopts markov model and then determines the possible positions of nodes, so as to obtain the optimal path planning. Tang *et al.* [24] proposes a scheme based on reinforcement learning (RL), which can be applied to opportunistic routing transmission requiring high reliability and low latency. This work designed an expected arbitrary path delay (EAD) for opportunistic networks as an innovative way to measure path costs and estimate possible end-to-end delays between the current node and the target node. As EAD is dynamically measured and updated, an appropriate trade-off between transport reliability and latency can be achieved. However, this opportunistic routing scheme can only be applied to specific scenarios and is not suitable for opportunistic social networks. Zhou *et al.* [25] proposed a new method to predict the social pattern of nodes from the perspective of time, so as to improve the performance of data forwarding in opportunistic mobile networks. The core idea is to capture and use temporal correlation to infer possible temporal social contact patterns of data in the remaining effective time. This work has a similar delivery rate to epidemic when it reduces its cost, but our algorithm performs better than epidemic.

B. EXISTING SOCIAL-AWARE ALGORITHMS

In the social-aware algorithms of opportunistic social networks, the social relationship between nodes is usually used to evaluate the transmission relationship between nodes, such as the possible meeting probability between nodes, the mobile habit of nodes or the situation of the community in which the nodes are located. Bubble Rap algorithm proposed by Jain and Yadav [26] can rank nodes according to their activity level and historical transmission records. In data forwarding process, the node passes the message to nodes in its communication domain in order according to the ranking situation until the destination node receives the information. Ciobanu and Dobre [27] proposed a viewpoint that the movement and interaction behavior of nodes in the future time is predictable, and the contact trajectory of nodes is regular even if the time series is not in strict and controlled situations. Zhang *et al.* [28] proposed a community detection label propagation algorithm based on node importance and label influence for large networks. This work can obviously improve the quality of community detection result and shorten the iteration time. In literature [18], Floriano *et al.* proposed a data forwarding strategy of opportunistic network based on social information, which can make use of offline and online user information of social networks. This method can improve the success rate of information transmission and reduce the number of message copies in the network by extracting the central index of the node interaction history in the online network. Li *et al.* [29] proposed a community detection algorithm based on user profile by studying the potential attributes of user profiles. Combined with the similarity of users' complex attributes, users with similar interests, action purposes and daily behaviors could be can be distinguished, so as to realize effective community detection.

Literature [17] explores a framework of relational, social, and personal contexts as predictors of matching opportunities. This algorithm can map the predictive impact of personal context and the social ability of neighbors on the optimal routing of opportunistic networks. ML-SOR algorithm proposed by [14] combines the characteristics of social networks, from node centrality, community structure, bond strength and link prediction and other characteristics. This measurement model is used to measure the forwarding ability of a node relative to the node encountered so as to conduct data forwarding. Yan *et al.* [30] proposed an algorithm to divide nodes in the social network into several communities and adopt a reduction strategy based on multiple attributes to obtain several efficient communities, so as to make efficient routing and transmission through the community. In literature [31], Kanghuai *et al.* proposed a concept of mobile similarity, which combined the social relationships and node transmission features. The FCNS algorithm has good performance on packet delivery ratio and routing overhead, but not good at transmission delay. Socievole *et al.* [16] also proposed a way to predict node behavior based on node characteristics, which discussed multi-layer and multiple metrics related to the social forwarding. Jia *et al.* [32] proposed a routing algorithm to establish the weight distribution and community reconstruction among nodes, which can solve the problem of large transmission delay in social opportunistic networks due to insufficient cache. This method not only has low energy consumption, but also can improve the transmission ratio, overhead and end-to-end delay of social opportunistic network. By analyzing people's sociality and subjectivity, Wu *et al.* [33] proposed an optimal routing scheme for cooperative nodes based on the characteristics of opportunistic networks. In this scheme, reliability, availability and weight factors are taken as the weights of human activities to obtain the optimal cooperation nodes, thus has good performance in real social life.

The routing algorithm based on sociality generally considers the social relationship between nodes and puts forward a lot of indicators to measure the characteristics of nodes. However, with the gradual popularization of 5G technology, the indicators should fit into the application scenario. In order to make the opportunistic social network routings can be good transplanted into real 5G scene, we proposed an effective transmission strategy exploiting node preference and social relations. The main idea of this paper is to propose a concept of preference similarity by analyzing the individual preference of nodes in the transmission process. Then the similarity and neighbor information of nodes are added when the next hop node is measured so as to obtain the prediction and recommendation scheme of optimal relay node and transmission path. The algorithm proposed in this paper combines the advantages of existing algorithms and makes it applicable to the real 5G scene.

III. SYSTEM MODEL DESIGN

We have witnessed the fruitful and exciting advancement of opportunistic network routing algorithms based on node attributes and complex connectivity patterns. These natural and artificial social connection attributes are caused by the social, economic and financial status of people at different levels. By quantifying the attributes of human society and analyzing the historical interaction information, this work propose an optimal next hop prediction algorithm based on social attributes in 5G environment, which can improve routing performance by obtaining the transmission preferences of nodes.

A. FEATURE ANALYSIS

Each node has its own feature preference during information transmission process, which has a variety of performance presentations due to individual differences. Considering the conditions that transmission process needs to meet is very important for the node to select relay nodes with good performance. Based on such criteria, we define the following features to quantify feature attributes in 5G environment.

1) COMMUNICATION QUALITY

The new 5G multiple access technology adopts the end-to-end data transmission mode, which can greatly improve the access amount of mobile network. In this case, whether the equipment can meet the communication needs has become a major issue to be considered for efficient transmission in the 5G environment. Communication quality is a concept refers to the effectiveness degree of the network in meeting the needs of users for communication. In opportunistic social networks, since nodes are in a constant moving state, the connection relations established with relay nodes is intermittent and discontinuous. If we want to measure the ability of a node to transmit message, we need to evaluate the total amount of connected time that the node has successfully communicated over a period of time. In a time interval, if the total maintenance time of the stable connection state between node V_a and node V_b is long, it can be considered that the communication capability of node V_a is strong during this time.

As shown in the figure below, we divide the communication states of nodes in time interval t into two types, connection and disconnection state. If a node establishes a connection with another node at time t_1 and disconnect at time t_2, the state of the node within the time interval $[t_1, t_2]$ is referred to as a stable connection state, and this period of time is referred to as connection time. If the node is disconnected at time t_2, and it enters the connected state again at time t_3, then the state of the node within the time interval $[t_2, t_3]$ is called the disconnected state, and this period of time is referred to as disconnection time. The transition between the connected state and the disconnected state may be due to node movements, which is caused by a node exceeding the communication domain of another node. For example, node a is located in the communication domain of node b at t_1, and is not within the communication domain of node b at time t_2.

The total lifetime of the node during the message passing process is defined as communication time, denoted as T_i. The connection time is defined as TC_{re}, which represents the total connection duration between node V_r and the encounter node V_e. The disconnection time is denoted by TD_{re}, representing the total disconnection duration between node V_r and the encounter node V_e. According to the connection characteristics of nodes in its lifetime, the calculation methods of the connection time and the disconnection time can be obtained, which are expressed as:

$$TC_{re} = \sum_{i=1}^{m} TC_{re}(i) = \sum_{i=1}^{m} [TC_{re\_end}(i) - TC_{re\_start}(i)] \tag{1}$$

$$TD_{re} = \sum_{i=1}^{m} TD_{re}(i) = \sum_{i=1}^{m} [TC_{re\_start}(i+1) - TC_{re\_end}(i)] \tag{2}$$

where m represents the number of successful connections of node V_r during its lifetime. $TC_{re\_end}(i)$ and $TC_{re\_start}(i)$ respectively represent the end time and start time of the $i-th$ connection of node V_r and the encountered node V_e, and V_e represents all nodes that node V_r meets and successfully connects. By quantifying and analyzing the connection time and communication time of node, communication quality

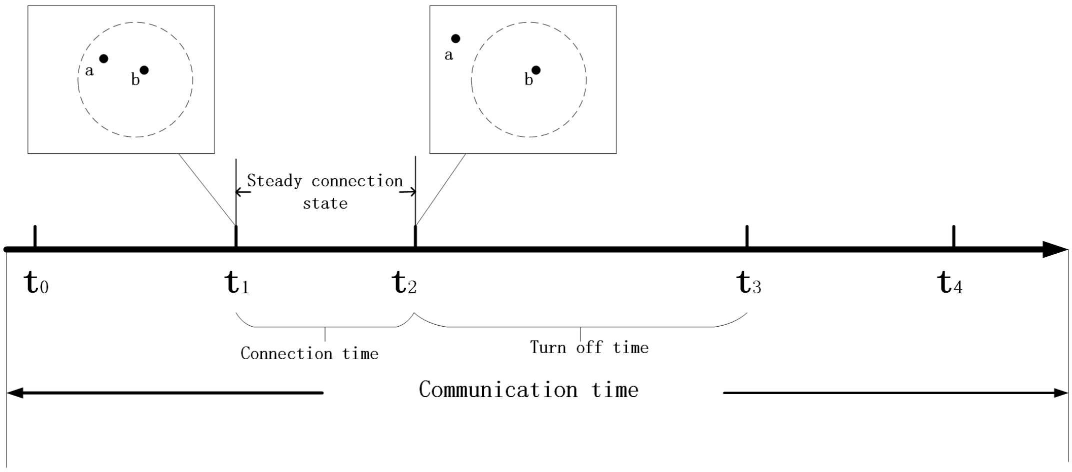


FIGURE 1. This diagram describes two cases of communication time, connected time and disconnected time.

$CQ(i)$ of node V_i is defined as:

$$CQ(i) = \sum_{i=1}^{m} \left[\begin{array}{l} (T_i - TC_{re}(i)) - \\ \frac{1}{m}\left(T_i - \sum_{i=1}^{m}\left[TC_{re\_end}(i) - TC_{re\_start}(i)\right]\right) \end{array} \right]^2 \tag{3}$$

2) REMAINING CACHE OF NODES

The development of 5G technology has brought a boom in the development of high-speed communication network. Under this situation, the transmission of massive data in a short time has certain requirements for the storage space of equipments. In the information transmission and data forwarding process, it is necessary to consider whether the remaining buffer space of the next hop node is sufficient to support information transmit, thus the remaining cache of node is a very important and critical feature when selecting relay nodes.

We define the time interval as t, and the amount of data received by node n at time t is expressed as $C_n(t)$. The required buffer space is linearly related to the amount of data $R_n(t)$ received during time interval t, and can be expressed as $C_S * R_n(t)$, where C_S is the buffer room occupied by node when collecting unit amount data. When the information is transmitted, it needs to consume a certain amount of buffer space. If the sink node allocates channels to node n, the node n transmits the message at the unit cache consumption rate C_T.

Therefore, the total cache room C_n^{total} of node n in time interval t can be expressed as the cache occupied by the amount of data received in the time interval plus the buffer consumed when transmitting messages, as shown in the formula [4].

$$C_n^{total}(t) = C_S * R_n(t) + \sum_{k\in\kappa} J_{n,k}(t)C_T, \quad \forall n \in N \tag{4}$$

Since the amount of data received by the node in the time gap has an upper bound, $R_n(t) \le R_{\max}$, and the number of allocated channels $\sum_{k\in\kappa} J_{n,k}(t)\, C_T \le 1$, So the maximum cache value of each node within a time interval t can be expressed as $C_{\max} = C_s r_{\max} + C_T$. So the remaining cache of node n at the current time t_0 can be expressed as:

$$\begin{aligned} C_n^{remain}(t_0) &= C_{\max} - C_n^{total}(t_0) \\ &= C_{\max} - C_S r_n(t_0) - \sum_{k\in\kappa} J_{n,k}(t_0)C_T, \quad \forall n \in N \end{aligned} \tag{5}$$

3) ACTIVE LEVEL OF NODES

In this section, we examine another important factor affecting data transmission in 5G environment, node active level, which is not considered by existing routing algorithms. 5G network adopts multiple access, which greatly improves the number of links between devices and the efficiency of data transmission. Therefore, light and portable mobile communication equipment will replace base station as the main carrier of data transmission and realize super-reliable large-scale mechanical communication. Human beings are carriers of wireless communication devices, so the regularity in human life has a critical impact on the performance of routing algorithms. In the normal daily life of human beings, the time period during which information can be transmitted is generally continuous, and the active level is relatively low at night and midnight. If we want routing algorithms to be more efficient, the impact of user activity changes on information transmission should be eliminated, so we propose the concept of node active level and use the mapping time instead of the wall time to measure active level of nodes in the opportunistic social networks.

The units of mapping time $\widehat{H}$ and wall time H are defined as $\tilde{\omega}$ and ω respectively, representing every second in a day, and their value range both is [0, 86400). The biggest difference between this two time is that H can only be a positive integer but $\widehat{H}$ can be a decimal. By analyzing the transmitted data, we can get the number of messages transmitted per second (E_ω), and obtain the average number of messages transmitted per second (E^*). Obviously, the value of E_ω will continue to change as ω changes, and the value of E^* will

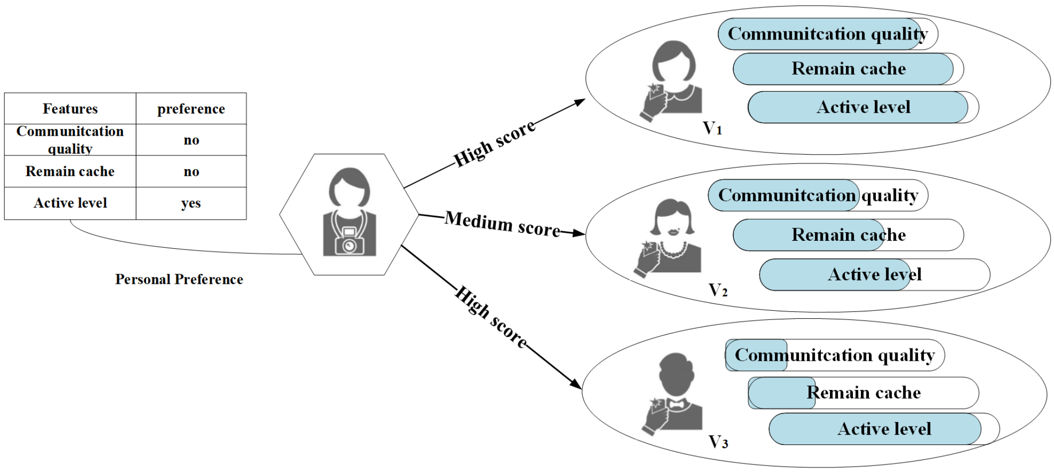


FIGURE 2. In this figure, the individual preferences of the nodes are shown. Due to the different preference level of nodes for transmission attributes, the result of inter-node scoring has individual preference obviously.

not change. It is easy to find that E_ω in the working time period is significantly higher than it in the night rest time, so E_ω can be used to describe the continuously change related to the node active level.

The conversion relationship between wall time and mapping time is related to the number of messages transmitted in unit time. We define the mapping function g here as equation [6]:

$$\tilde{\omega}= g(\omega) = \sum_{j=0}^{\omega} E_j \big/ E^* \tag{6}$$

For the original wall timestamp H of the dataset, we define its corresponding mapping timestamp as $\widehat{H}$. In order to convert the wall timestamp into its map timestamp, we first need to convert H to the minimum unit ω_* of day by the following formula:

$$\omega_* = \mathrm{mod}(H + T_{zone} * H_{\sec onds}, D_{\sec onds}) \tag{7}$$

where mod (a, b) returns the modulus after division of a by b. T_{zone} represents the time zone in which the node is located, for example, if the node is in China, and China belongs to the East Eight District, then T_{zone} takes 8 as its value. $H_{\sec onds}$ is the number of seconds in an hour. $D_{\sec onds}$ represents the total number of seconds in a day, which is the maximum value of $\tilde{\omega}$ and ω. According to the expressions of equations [6] and [8], the mapping time can be calculated as:

$$\widehat{H}_i = H_i - \omega_* + \tilde{\omega}_* = H_i - \omega_* + g(\omega_*) \tag{8}$$

The communication chance of nodes in the opportunistic social network are closely related to the frequency of node movements in the time interval. Thus, the active level of a node can be judged according to the total distance of its movements in the mapping time period, expressed as equation [9], in which v represents the average speed of the node in the history records:

$$AL_i = v * \widehat{H}_i = v * [H_i - \omega_* + g(\omega_*)] \tag{9}$$

B. PREFERENCE SIMILARITY CALCULATION

In the process of node interactions and information transmission in opportunistic social networks, we will take the user as the node of the sender. Each user is required to give an evaluation score for the nodes they interact with after the successful transmission. However, through the analysis of the transmission history, we can find that the users usually have feature preferences in the information transmission process. This kind of feature preference is similar to the personal interests of human beings, which may affect the fairness and accuracy of their judgments.

Figure 2 is a simplified diagram depicting node n scoring three interactive nodes. As shown in the diagram, node n has its feature preference, and when it scores the interactive nodes V_1, V_2, and V_3, the result of the evaluation is obviously subjective. It is easy to find that the score of node V_1 is fair because all of its feature levels are high. However, the integrated transmission capability of node V_3 is lower than that of node V_2, but since the feature C_i of node V_3 satisfies the preference of node n, the node n scores higher on node V_3 than node V_2.

Due to this kind of transmission preference of nodes, we propose a concept of preference similarity. The preference similarity can be composed of two parts: the score similarity of interactive nodes and the similarity of the user's attention to the different attributes of the interactive node. The score similarity can express the user's preference for a single node, and the attribute focused similarity can reflect the user's preference for various features of the node. When selecting the optimal next hop, the user may have a low evaluation of node V_i but is very concerned about a certain attribute of it. It is reasonable to recommend such a node V_i to this user. And vice versa. Therefore, these two similarities should be considered comprehensively.

Set user set $V = \{v_1, v_2, \ldots, v_n\}$, the set of interactive nodes $I = \{i_1, i_2, \ldots, i_m\}$. I_c represents a collection of interactive nodes that user v_h and user v_l have jointly evaluated.

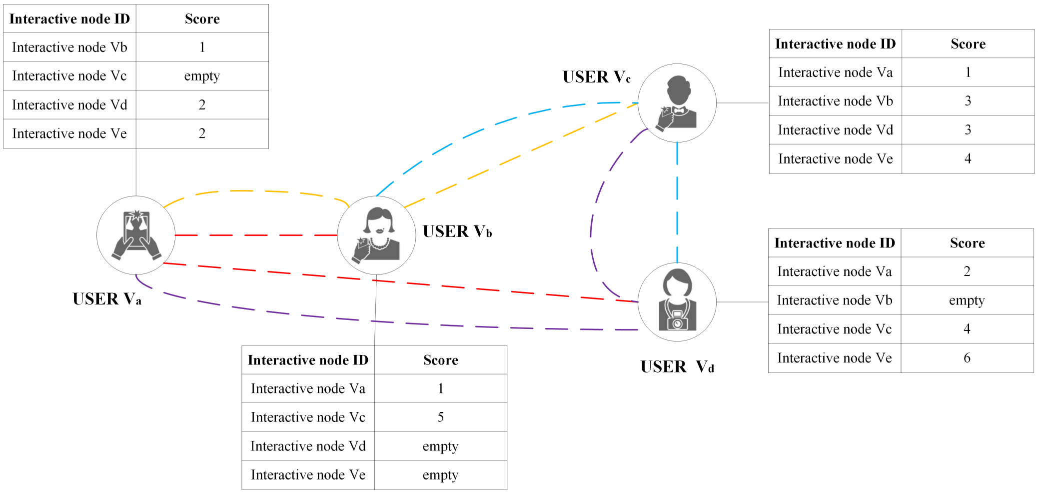


FIGURE 3. **This figure describes the score between nodes. If there is a historical interaction between node i and node j, node i will give a score to node j by the situation of communication.**

r_{vi} represents the rating of user v to interactive node i, and $\bar{r}_v$ represents the average rating of user v for set I_c. In this paper, pearson correlation similarity is used to calculate the similarity of users' scores to interactive nodes, which can represent as:

$$Sim(v_h, v_l)_a = \frac{\sum\limits_{i \in I_c} \left(r_{vi} - \bar{r}_{v_h}\right)\left(r_{vi} - \bar{r}_{v_l}\right)}{\sqrt{\sum\limits_{i \in I_c} \left(r_{vi} - \bar{r}_{v_h}\right)^2 \sum\limits_{i \in I_c} \left(r_{vi} - \bar{r}_{v_l}\right)^2}} \tag{10}$$

The cosine similarity is used to calculate the similarity of the user's attention to different attributes of interactive nodes, denoted as:

$$Sim(v_h,v_l)_b = \frac{\sum\limits_{c_i \in C} \left(n_{v_h c_i} * n_{v_l c_i}\right)}{\sqrt{\sum\limits_{c_i \in C} n_{v_h c_i}^2} * \sqrt{\sum\limits_{c_i \in C} n_{v_l c_i}^2}} \tag{11}$$

where $C = \{c_1, c_2, \ldots, c_i\}$ represents the attributes of interactive nodes. $n_{v_h c_i}$ indicates the rating numbers of user v_h under the interactive node's attribute c_i. According to the above two formulas, the user's preference similarity to the interactive node is expressed as equation [12], where α is the balance coefficient (in this paper, $\alpha = 0.5$).

$$Sim(v_h, v_l) = \alpha Sim(v_h, v_l)_a + (1 - \alpha)\, Sim(v_h, v_l)_b \tag{12}$$

C. USING THE MATRIX DECOMPOSITION METHOD TO PREDICT OPTIMAL RELAY NODES

In the process of selecting the next hop in opportunistic social networks, the user will score its interactive nodes, as shown in figure 3. We define the sender nodes as user nodes, nodes that are in the receiver and have interacted with the user nodes is defined as corresponding interaction nodes. Each user node scores its interactive nodes after they achieve a successful transmission.

From the figure we can see that the node V_a evaluates node V_b, node V_d and node V_e after the process of information transmission, but does not evaluate node V_c. This may be because node V_a has no interaction with node V_c, or it may be because of other objective reasons that node V_a has not yet scored node V_c but it will in the future. Node V_c may interact with node V_a as a user node and evaluate it, but the evaluation between them is not symmetric in value. That is to say, when scoring, the score of node V_b to node V_c is different from that of node V_c to node V_b, which is caused by the attribute preference of each node. To calculate scores between user nodes and interaction nodes in the network, we can get the score matrix of this network.

Corresponding calculation based on the scoring matrix of the user nodes-interaction nodes can be used to predict the score for the unrated node. Therefore, this method proposed in this section can be used to predict the optimal relay nodes, thereby improving the performance of routing algorithm. Specifically, it is assumed that there are N user nodes and M nodes that have interacted with them in the network, and the user node it constitutes is $V = \{v_1, v_2, \ldots, v_n\}$. The score matrix of user nodes and interactive nodes is express as $R = [r_{vi}]_{N \times M}$. In this score matrix, r_{vi} represents the score of user node v on interactive node i. The matrix decomposition method which considered preference similarity can analyze the potential feature vectors of user nodes (interactive nodes) and then predict unknown scores. Suppose $U \in R^{K \times M}$ and $V \in R^{K \times M}$ respectively represent the potential eigenvectors of user nodes and interaction nodes, where U_v and V_i represent the K-dimensional eigenvectors of the user nodes and the interaction nodes, respectively.

The matrix decomposition model gives the user nodes and the interactive nodes an implicit feature vector with a gaussian prior which average value is 0, express as equation [13]. σ_U^2 and σ_V^2 represent the vector variances of the user node and the interaction node respectively, and E is the identity matrix.

$$p\left(U\left|\sigma_U^2\right.\right)=\prod_{v=1}^{N}N\left(U_v\left|0,\sigma_U^2E\right.\right)$$

$$p\left(V\left|\sigma_V^2\right.\right)=\prod_{i=1}^{N}N\left(V_i\left|0,\sigma_V^2E\right.\right) \tag{13}$$

Assume that the conditional probability of the score data that has been obtained is as follows:

$$p\left(R\left|U,V,\sigma_R^2\right.\right)=\prod_{v=1}^{N}\prod_{i=1}^{M}\left[N\left(R_{vi}\left|U_v^TV_i,\sigma_R^2\right.\right)\right]^{I_{vi}^R} \tag{14}$$

wherein, I_{vi}^R takes 1 when user node v has already scored the interactive node i, otherwise, it takes 0. After Bayesian derivation, the posteriori probability of the implicit feature of the user node and the interactive node can be obtained to satisfy the formula [15]:

$$\begin{aligned}&p\left(U,V\left|R,\sigma_U^2,\sigma_V^2,\sigma_R^2\right.\right)\\&=\propto p\left(R\left|U,V,\sigma_R^2\right.\right)\times p\left(U\left|\sigma_U^2\right.\right)\times p\left(V\left|\sigma_V^2\right.\right)\end{aligned} \tag{15}$$

The purpose of setting the score of interactive nodes like this is to maximize the posterior probability of implicit features as large as possible. In order to simplify the subsequent derivation process, equation [16] takes a natural logarithm operation on the posterior probability of the implicit feature.

$$\begin{aligned}&\ln p\left(U,V\left|R,\sigma_U^2,\sigma_V^2,\sigma_R^2\right.\right)\\&=-\frac{1}{2\sigma_R^2}\sum_{v=1}^{N}\sum_{i=1}^{M}I_{vi}^R\left(R_{vi}-U_v^TV_i\right)\\&\quad-\frac{1}{2\sigma_U^2}\sum_{v=1}^{N}U_v^TU_M-\frac{1}{2\sigma_V^2}\sum_{i=1}^{M}U_i^TV_i\\&\quad-\frac{1}{2}\left[\left(\sum_{v=1}^{N}\sum_{i=1}^{M}I_{vi}^R\right)+N\times K\ln\sigma_U^2+M\times K\ln\sigma_V^2\right]+C\end{aligned} \tag{16}$$

To maximize the formula [16], you need to minimize the following objective function, express as equation [17].

$$\begin{aligned}F&=\frac{1}{2}\sum_{v=1}^{N}\sum_{i=1}^{M}I_{vi}^R\left(R_{vi}-U_v^TV_i\right)^2+\frac{\sigma_R^2}{2\sigma_U^2}\sum_{v=1}^{N}\|U_v\|^2\\&\quad+\frac{\sigma_R^2}{2\sigma_V^2}\sum_{i=1}^{M}\|V_i\|^2\end{aligned} \tag{17}$$

Since neighbor nodes affect the judgments of the node on the interactive node when scoring, that is, the score is also closely related to neighbor nodes. In order to make the scoring situation as close as possible to the real social life in 5G environment, our probabilistic decomposition model incorporates the user's neighbor information. The feature vector of user v_h can be represented as a weighted sum of its neighbor vectors. N_v represents the neighbor nodes of user node v, and S_{vj} represents the preference similarity degree between user node v and user node j. Next, we normalize S_{vj} according to these two equation below:

$$\sum_{j\in N_v}S_{vj}=1 \tag{18}$$

$$U_v=\sum_{j\in N_v}U_jS_{vj} \tag{19}$$

In order to make the matrix decomposition method more suitable for the optimal relay node selection scheme, this paper adds a conditional probability to this method, which expressed as equation [20]:

$$p\left(U\left|\sigma_U^2,\sigma_S^2\right.\right)=p\left(U\left|\sigma_U^2\right.\right)\times p\left(U\left|S,\sigma_S^2\right.\right) \tag{20}$$

where $p\left(U\left|S,\sigma_S^2\right.\right)$ represents the probability of the eigenvector of user nodes under the condition of the known neighbor relationships, and the probability is consistent with the gaussian distribution below:

$$p\left(U\left|S,\sigma_S^2\right.\right)=\prod_{v=1}^{N}N\left(U_v\left|\sum_{j\in N_v}U_jS_{vj},\sigma_S^2E\right.\right) \tag{21}$$

Based on these improvements, the new posterior probability of feature vectors of user nodes and interaction nodes can be obtained, expressed as:

$$\begin{aligned}&p\left(U,V\left|R,\sigma_U^2,\sigma_V^2,\sigma_R^2\right.\right)\\&=\propto p\left(R\left|U,V,\sigma_R^2\right.\right)\times p\left(U\left|\sigma_U^2\right.\right)\\&\quad\times p\left(U\left|S,\sigma_S^2\right.\right)\times p\left(V\left|\sigma_V^2\right.\right)\end{aligned} \tag{22}$$

After these steps, we can get the improved objective function:

$$\begin{aligned}F&=\frac{1}{2}\sum_{v=1}^{N}\sum_{i=1}^{M}I_{vi}^R\left(R_{vi}-U_v^TV_i\right)^2+\frac{\sigma_R^2}{2\sigma_U^2}\sum_{v=1}^{N}\|U_v\|^2\\&\quad+\frac{\sigma_R^2}{2\sigma_V^2}\sum_{i=1}^{M}\|V_i\|^2+\frac{\sigma_R^2}{2\sigma_S^2}\sum_{v=1}^{M}\begin{bmatrix}\left(U_v-\sum_{p\in N_v}U_pS_{vp}\right)^T\\\left(U_v-\sum_{p\in N_v}U_pS_{vp}\right)\end{bmatrix}\end{aligned} \tag{23}$$

The above objective function can be solved by the stochastic gradient descent method, and the partial derivative formulas of the characteristics of the user node and the interactive node can be respectively obtained, which is expressed as:

$$\begin{aligned}\frac{\partial F}{\partial U_v} &= \sum_{i=1}^{M} I_{vi}^{R} V_i \left(U_v^T V_i - R_{vi}\right) + \lambda_U U_v \\ &\quad + \lambda_S \left(U_v - \sum_{p \in N_v} U_p S_{vp}\right) \\ &\quad - \lambda_S \sum_{p | v \in N_p} \left(U_p - \sum_{x \in N_p} U_p S_{px}\right) \\ \frac{\partial F}{\partial V_i} &= \sum_{v=1}^{N} I_{vi}^{R} U_v \left(U_v^T V_i - R_{vi}\right) + \lambda_V V_i \qquad (24)\end{aligned}$$

Among them, $\lambda_U = \frac{\sigma_R^2}{\sigma_U^2}$, $\lambda_S = \frac{\sigma_R^2}{\sigma_S^2}$, and $\lambda_V = \frac{\sigma_R^2}{\sigma_V^2}$ are penalty coefficients, which can prevent the score matrix of user nodes and interactive nodes from over fitting. Therefore, the matrix decomposition model converts the original problem into an optimization problem among neighbor nodes, for which there must be an optimal solution. Through these mathematical calculations, we can get the feature vector of user nodes and interaction nodes, so that we can predict all the score of interaction nodes which are evaluated by user nodes. The prediction scheme proposed in this section can obtain the optimal next-hop node recommendation scheme in the current communication domain. In each forwarding decision, the prediction of next hop node can improve the selection efficiency of relay nodes and the performance of routing algorithm.

IV. COMPLEXITY ANALYSIS OF SYSTEM MODEL

In summary, an effective data transmission strategy exploiting node preference and social relations is proposed in this paper. Meanwhile, to improve the understanding of the whole algorithm, specific steps of the message transmission process are listed as follows:

STEP 1: In order to make the routing algorithm can be well applied in real social life, the social attributes of nodes are considered in the data transmission strategy. The concepts of communication quality, remain cache, and active level are presented in this paper to quantify the social attributes of nodes in 5G environment.

STEP 2: Due to the social characteristics of nodes, individual bias exists in the evaluation of interactive nodes. Therefore, preference similarity is proposed to measure the preference value of nodes for attributes, so that the preference of nodes can be well considered in data transmission process.

STEP 3: The evaluation value of nodes for interactive nodes can be used to obtain the score matrix. By using matrix decomposition method and bayesian derivation, we can predict the score of nodes for non-interactive nodes and recommend the best next-hop node in the current communication domain of nodes.

STEP 4: Selecting the most suitable relay node in the current communication domain in each data transmission process can obtain the final routing propagation path.

Algorithm 1 An Data Transmission Strategy Exploiting Node Preference and Social Relations

Input: A graph $G(V, E)$, a source node S, a destination node D;

Output: optimal path;

Begin

//Social feature analysis

Calculate communication quality $CQ(i)$, remain cache $C_n^{remain}(t_0)$ and active level AL_i

//Preference similarity quantification

For each node v_h

Calculate preference similarity $Sim(v_h, v_l)$

End for

//Utilize the Probabilistic decomposition model

For each node

  Give a score to nodes it interacted with;

End for

The objective function is obtained by bayesian derivation $p\left(U \left| \sigma_U^2\right.\right) = \prod_{v=1}^{N} N\left(U_v \left| 0, \sigma_U^2 E\right.\right)$ and $p\left(V \left| \sigma_V^2\right.\right) = \prod_{i=1}^{N} N\left(V_i \left| 0, \sigma_V^2 E\right.\right)$

The new objective function is obtained by introducing preference similarity $Sim(v_h, v_l)$ into the objective function

For each process

  The optimal relay node in the current communication domain is obtained by optimizing the objective function

End For

End

Algorithm 1 is constructed to introduce the ENPSR algorithm in details for better readability. Specifically, the social attributes of each node are considered during the feature quantification process, so the time complexity of this phase is $O(n)$. In the phase of preference similarity measurement, the score similarity of nodes and the attribute similarity of nodes are calculated, thus, the time complexity of this process is $O(n)$. Based on the theory of matrix decomposition, the score of the nodes for the non-interactive nodes can be predicted and the optimal next hop of the nodes in the current communication domain can be obtained. The time complexity of this process is $O(\log_2 n)$. Therefore, the total computational complexity of the ENPSR algorithm can be computed as $O(n + n + \log_2 n) = O(n)$ through rigorous mathematical analysis. In retrospect, the time complexity of SCR is $O(n^2)$ and the time complexity in the Epidemic routing algorithm is $O(n)$.

V. SIMULATION AND ANALYSIS

In the simulations, we adopt the open source simulation tool, opportunistic network environment (ONE), to evaluate

TABLE 1. Characteristics of the four experimental data sets.

Dataset	Infocom5	Infocom6	Cambridge	Intel
Device	iMote	iMote	iMote	iMote
Duration(days)	3.5	4	11.5	4
Number of experimental devices	41	98	52	9
Number of internal contacts iMote	22459	170601	10873	1364

TABLE 2. Simulation parameters of four experimental datasets in ONE.

Dataset	Infocom5	Infocom6	Cambridge	Intel
Number of nodes	41	98	52	9
Buffer size	5M	5M	5M	5M
TTL	60min	60min	2 days	0.5 days

ENPSR by performance comparison with FCNS [31] (fuzzy routing-forwarding algorithm), SCR [34] (effective social relationship measurement and cluster based routing), SCANE [32] (sensor communication area and node extend routing algorithm) and epidemic algorithm [20]. The latest routing algorithm FCNS, SCR and SCANE are published in recent years, and epidemic algorithm is a classic method of opportunistic networks. This work adopts the real datasets download from CRAWDAD to drive node activity. By considering the data information required by the proposed algorithm in 5G environment, dataset Infocom 5, Infocom 6, Cambridge and Intel are selected for our simulations, and the detailed information is shown in table 1. In the evaluations, the cache size of node is set to 5M, and the message size is 1K. The number of nodes and TTL sets in this paper are different to original datasets, so the changes are shown in table 2.

The simulation scheme of data transmission in 5G environment is designed in this paper, which follows the characteristics of mass data transmission within a short time in 5G environment. In such an environment, data transfer efficiency can be described by packet delivery ratio. The transmission delay caused by transmission is the end-to-end delay between devices. Routing overhead is also an important aspect that must be measured, as the transmission of large amounts of data is bound to cause large routing overhead. The ENPSR algorithm and other four algorithms were run in the same simulation environment to compare their performance. As a reference, we set a list of evaluation indicators [35], which is used for comparison, as follows:

1) *Delivery Ratio:* This parameter refers to the ratio of the number of messages passed to the total number of messages generated within a certain time interval. This time interval is very short in a 5G environment.
2) *Average End-to-End Delay:* This parameter comprehensively measures the delay of transmission in 5G environment, which can be divided into three types: delay caused by routing selection, delay caused by relay node waiting for message and delay caused by message forwarding.
3) *Routing Overhead:* This parameter measures the overhead between node pairs during information transfer.

A. DELIVERY RATIO

In the simulation, ENPSR, FCNS, SCR, SCANE and Epidemic algorithm run in the four datasets respectively, and the parameter settings of simulation environment are shown in table 1. Figures 4 and 5 shows the change in the packet delivery ratio when the five algorithms above run in four different data sets. We use three dimensional broken line diagram to analyse the experimental results in figures 4. In figure 4, three dimensional broken line diagram can express the distribution of delivery ratio intuitively on the vision. Experimental results show that the performance of ENPSR is much higher than the other algorithms in terms of delivery ratio.

With the simulation time varying, figure 5 shows the packet delivery ratio of the ENPSR, FCNS, SCR, SCANE and epidemic algorithm. When the simulation time is short, the performance advantages of ENPSR algorithm is less obvious than that of the other four algorithms, but with the increase of simulation time, the success rate of ENPSR algorithm is significantly higher than that of the other four algorithms. This is because our solution quantifies the social attribute in 5G environment of nodes and fully considers the influence of node's attribute preference on message transmission strategy by measuring individual preference characteristics of nodes. Other schemes fail to take full account of the transmission bias of nodes, which may easily result in poor data transmission effect. Therefore, our algorithm can combine the social attribute and transmission preference of nodes when predict and recommend the next hop node, thus improving the selection accuracy of optimal node and path in the process of message transmission. Moreover, the SCR and SCANE algorithms transmit messages through the cooperation of neighbor nodes, but this is not an efficient transmission scheme when the cache room of nodes is limited. As for FCNS algorithm, the remaining cache and active level of nodes are not taken into account when calculating node similarity, thus, selected relay nodes may not be able to effectively participate in the process of message transmission. On the whole, in the ENPSR algorithm, the delivery ratio is 0.8447 on average, is 20% higher than that of epidemic algorithm.

B. AVERAGE END-TO-END DELAY

Figure 6 exhibits the comparison results of average end-to-end delay among the five different algorithms. The ENPSR algorithm has the smallest average end-to-end delay in the comparison scheme because it optimizes the routing strategy of data transmission more effectively than other schemes. The ENPSR algorithm quantifies and utilizes the message transmission preference of nodes, which can reduce a large

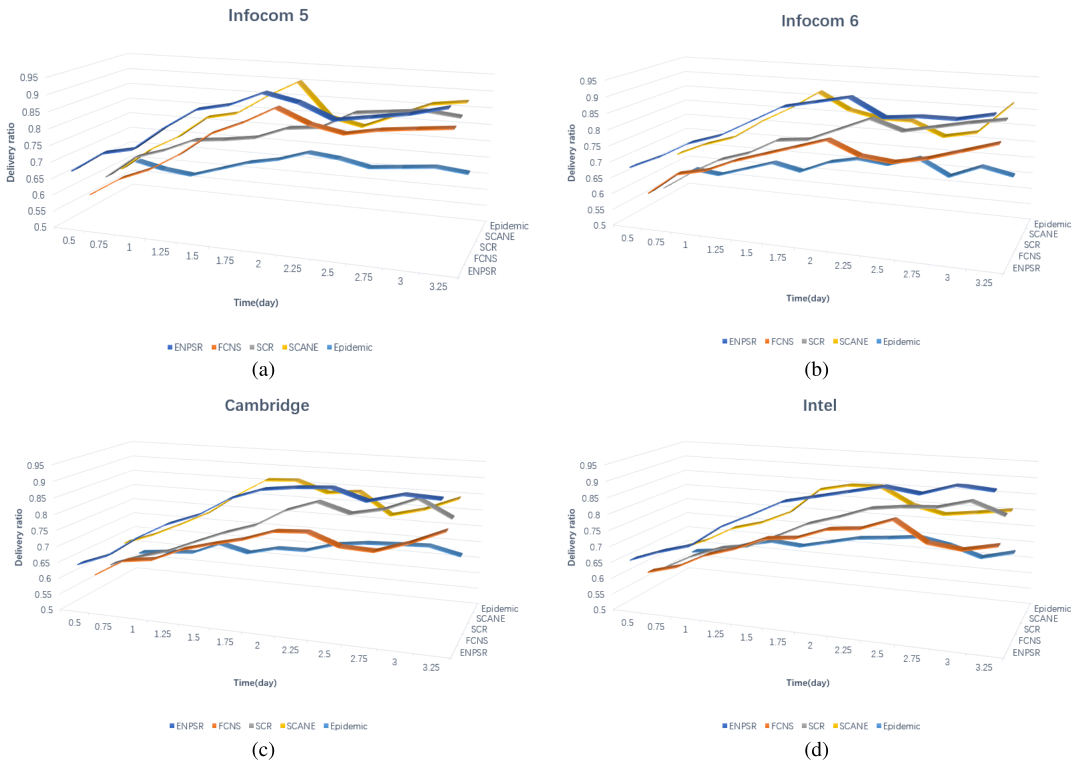


FIGURE 4. **Three dimensional broken line diagram of packet delivery ratio.**

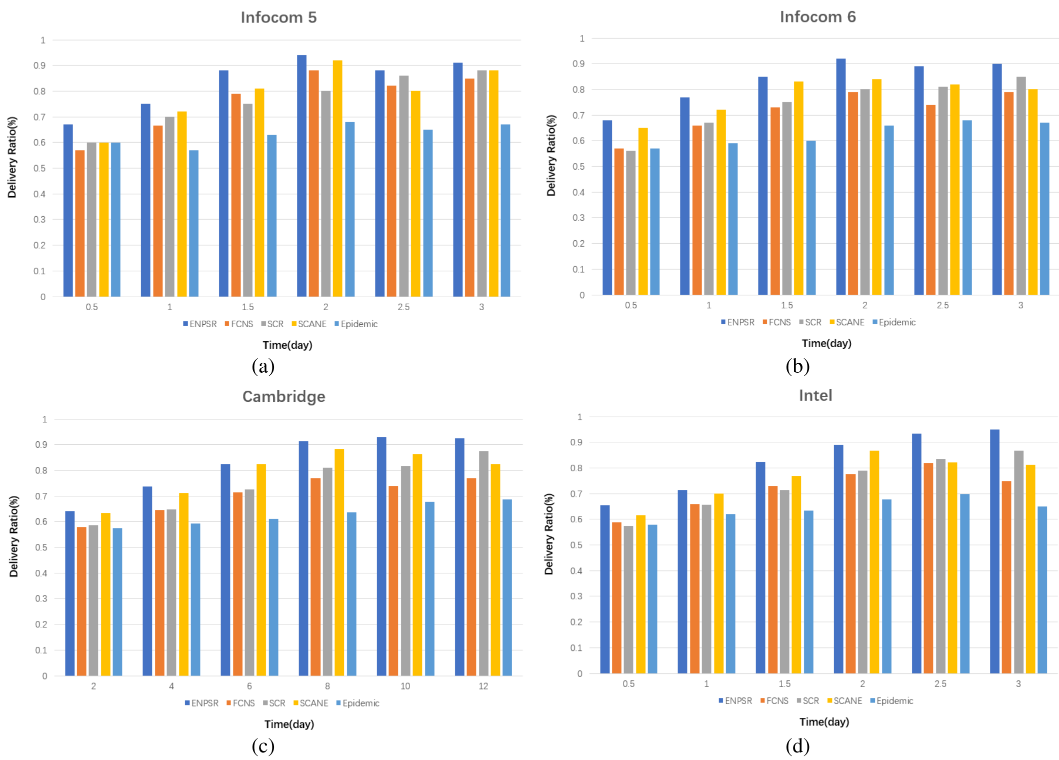


FIGURE 5. **Packet delivery ratio comparisons.**

number of waiting delays. The prediction and recommendation method proposed by ENPSR method also can reduce the delay of routing selection. In comparison, regarding the epidemic algorithm, a large number of copies of information are generated during data transmission, which will lead to an increase in forwarding delays. In addition, SCR and SCANE

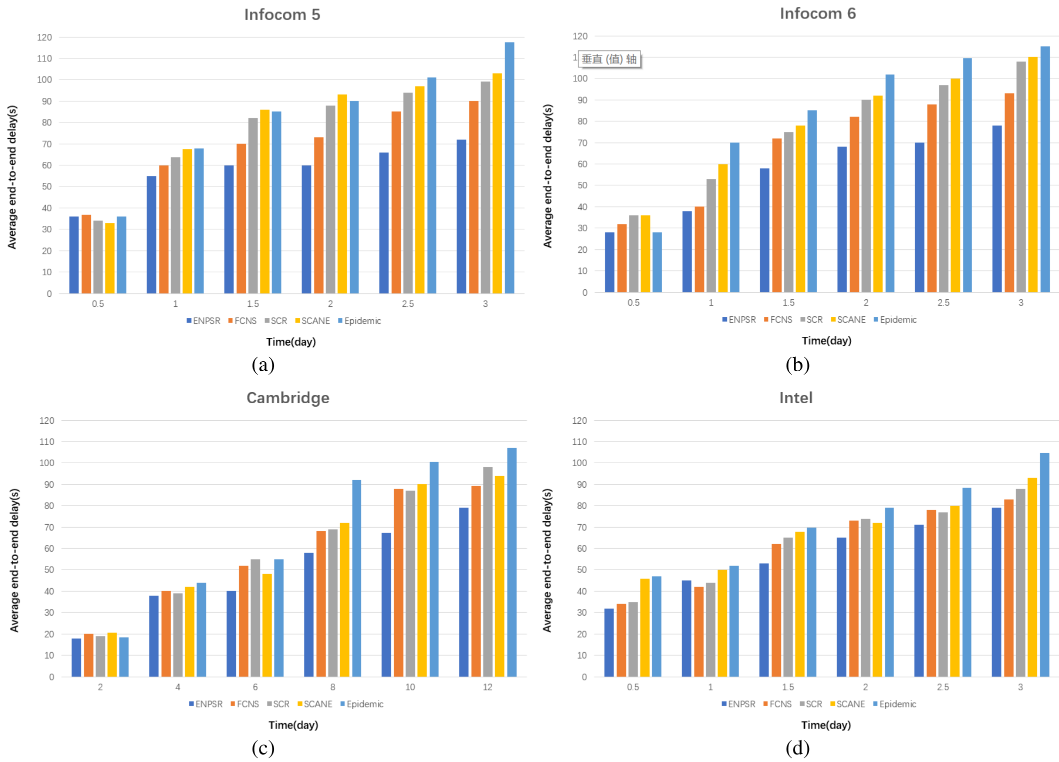


FIGURE 6. Average end-to-end delay comparisons.

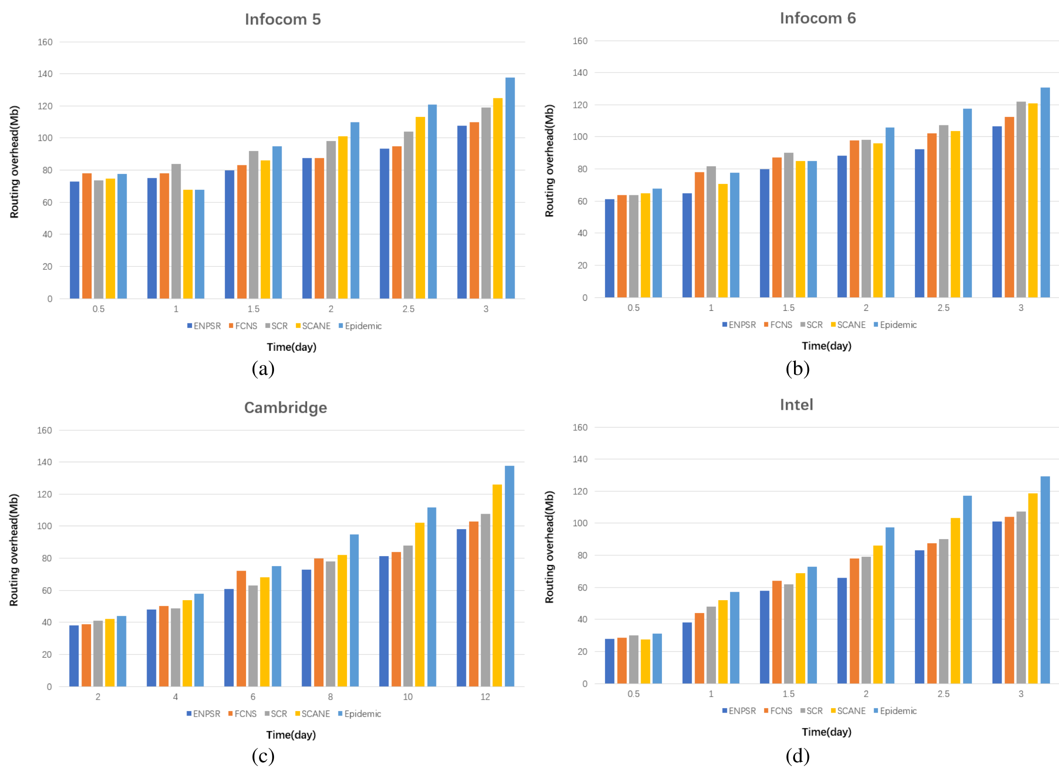


FIGURE 7. Routing overhead comparisons.

algorithms adopt the strategy of cooperative transmission by neighbor nodes, which reduces the impact of node cache on message transmission, but still has a high delay when the experiment time is long. FCNS algorithm is affected by the node cache, and the performance is poor when the node cache space is small. To sum up, the ENPSR algorithm is the best method to improve the performance of average end-to-end delays compared with other algorithms in high speed communication scenario.

C. ROUTING OVERHEAD

Eventually, the comparison results of routing overhead among these five different algorithms are demonstrated in figure 7. Compared with other models, ENPSR can predict the next-hop nodes better. The routing cost of sending messages to other non-cooperative nodes can be effectively reduced by sending messages to nodes that satisfy the transmission preference in the communication domain. In addition, the nodes do not need to use the computational model for continuous calculation and decision-making in the process of information transmission, which can reduce the cost of time and routing resources. Regarding the epidemic algorithm, a large number of redundant information copies require time and computing resources, and the routing overhead is significantly higher than that of other algorithms. For SCR and SCANE algorithms, the cooperation mechanism is conducive to reasonable allocation of computing resources, so the cost of these two algorithms is in the middle level. FCNS algorithm takes the mobile similarity of nodes into account, but it does not fully consider the transmission preference of nodes, so its performance is worse than that of ENPSR algorithm. Compared with the result, ENPSR has better performance than other four models in terms of routing overhead.

VI. CONCLUSION

This work proposed an effective data transmission strategy exploiting node preference and social relations (ENPSR) for opportunistic social networks. The algorithm combines multiple social attributes of nodes in 5G environment, quantifies the communication quality, remain cache and active level of nodes in the process of data transmission to measure the social relations of nodes. Unlike other routing algorithms based on node features, the ENPSR synthetically considers the individual preference of nodes and proposes an concept of preference, which is a combination of score and feature similarities. Then, based on the theory of matrix decomposition, the score of the nodes for the non-interactive nodes can be predicted. The optimal next hop of the nodes in the current communication domain and optimal recommended route can be obtained. Simulations have been done, and the results show that the proposed protocol performs better than the FCNS, SCR, SCANE and epidemic algorithms in terms of delivery ratio, average end-to-end delays and routing overheads.

In the future, with the improvements of computing capacity of mobile devices in opportunistic social networks, the ENPSR algorithm proposed in this paper can be applied to the transmission environment of 5G and big data networks. We will collect bigger real data sets in social scenarios and explore methods to improve the performance of information transmission.

REFERENCES

[1] S. Liu, L. Zhang, and Z. Yan, "Predict pairwise trust based on machine learning in online social networks: A survey," *IEEE Access*, vol. 6, pp. 51297–51318, 2018. doi: 10.1109/ACCESS.2018.2869699.

[2] P. Giridhar, S. Wang, T. Abdelzaher, R. Ganti, L. Kaplan, and J. George, "On localizing urban events with Instagram," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, May 2017, pp. 1–9.

[3] M. Patzold, "5G is coming around the corner [mobile radio]," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 4–10, Mar. 2019.

[4] A. Awada and D. S. Michalopoulos, "A radio resource-efficient paging method in high frequency 5G networks," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Singapore, Dec. 2017, pp. 1–6.

[5] Z. Ning, X. Wang, and J. Huang, "Mobile edge computing-enabled 5G vehicular networks: Toward the integration of communication and computing," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 54–61, Mar. 2019.

[6] K. Poularakis, G. Iosifidis, V. Sourlas, and L. Tassiulas, "Exploiting caching and multicast for 5G wireless networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 4, pp. 2995–3007, Apr. 2016.

[7] S. Li, L. Da Xu, and S. Zhao, "5G Internet of Things: A survey," *J. Ind. Inf. Integr.*, vol. 10, pp. 1–9, Jun. 2018.

[8] J. Wu, Z. Chen, And M. Zhao, "Information cache management and data transmission algorithm in opportunistic social networks," in *Wireless Networks*, 2018, pp. 1–12.

[9] N. Wang and J. Wu, "Optimal data partitioning and forwarding in opportunistic mobile networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–6.

[10] H. Zhou, J. Wu, H. Zhao, S. Tang, C. Chen, and J. Chen, "Incentive-driven and freshness-aware content dissemination in selfish opportunistic mobile networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 9, pp. 2493–2505, Sep. 2015.

[11] Y. Yan, Z. Chen, J. Wu, and L. Wang, "An effective data transmission algorithm based on social relationships in opportunistic mobile social networks," *Algorithms*, vol. 11, no. 8, p. 125, 2018.

[12] L. Chancay-García, E. Hernández-Orallo, P. Manzoni, C. T. Calafate, and J. Cano, "Evaluating and enhancing information dissemination in urban areas of interest using opportunistic networks," *IEEE Access*, vol. 6, pp. 32514–32531, 2018.

[13] J. Wu, Z. Chen, and M. Zhao, "Information transmission probability and cache management method in opportunistic networks," *Wireless Commun. Mobile Comput.*, vol. 2018, Mar. 2018, Art. no. 1571974.

[14] A. Socievole, E. Yoneki, F. De Rango, and J. Crowcroft, "ML-SOR: Message routing using multi-layer social networks in opportunistic communications," *Comput. Netw.*, vol. 81, pp. 201–219, Apr. 2015.

[15] S. J. Borah, S. K. Dhurandher, I. Woungang, and V. Kumar, "A game theoretic context-based routing protocol for opportunistic networks in an IoT scenario," *Comput. Netw.*, vol. 129, pp. 572–584, Dec. 2017.

[16] A. Socievole, E. Yoneki, F. De Rango, and J. Crowcroft, "Opportunistic message routing using multi-layer social networks," in *Proc. 2nd ACM Workshop High Perform. Mobile Opportunistic Syst.*, 2013, pp. 39–46.

[17] J. M. Mayer, S. R. Hiltz, L. Barkhuus, K. Väänänen, and Q. Jones, "Supporting opportunities for context-aware social matching: An experience sampling study," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, San Jose, CA, USA, May 2016, pp. 2430–2441.

[18] F. De Rango, A. Socievole, and S. Marano, "Exploiting online and offline activity-based metrics for opportunistic forwarding," *Wireless Netw.*, vol. 21, no. 4, pp. 1163–1179, 2015.

[19] K. Jang, J. Lee, S.-K. Kim, J.-H. Yoon, and S.-B. Yang, "An adaptive routing algorithm considering position and social similarities in an opportunistic network," *Wireless Networks*, vol. 22, no. 5, pp. 1537–1551, 2016.

[20] M. Chitra and S. S. Sathya, "Selective epidemic broadcast algorithm to suppress broadcast storm in vehicular ad hoc networks," *Egyptian Inform. J.*, vol. 19, pp. 1–9, Mar. 2017.

[21] S. J. Borah, S. K. Dhurandher, S. Tibarewala, I. Woungang, and M. S. Obaidat, "Energy-efficient prophet-PRoWait-EDR protocols for opportunistic networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

[22] S. Sisodiya, P. Sharma, and S. K. Tiwari, "A new modified spray and wait routing algorithm for heterogeneous delay tolerant network," in *Proc. Int. Conf. I-SMAC (IoT in Social, Mobile, Anal. Cloud) (I-SMAC)*, Coimbatore, India, Feb. 2017, pp. 843–848.

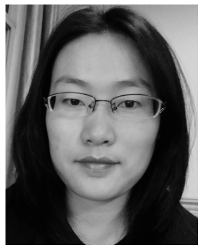
[23] X. Wang, X. Wang, L. Zhang, Y. Lin, R. Zhao, *A Routing Algorithm Based on the Prediction of Node Meeting Location in Opportunistic Networks*. Singapore: Springer, 2017.

[24] K. Tang, C. Li, H. Xiong, J. Zou, and P. Frossard, "Reinforcement learning-based opportunistic routing for live video streaming over multi-hop wireless networks," in *Proc. IEEE 19th Int. Workshop Multimedia Signal Process.*, Oct. 2017, pp. 1–6.
[25] H. Zhou, V. C. M. Leung, C. Zhu, S. Xu, and J. Fan, "Predicting temporal social contact patterns for data forwarding in opportunistic mobile networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10372–10383, Nov. 2017.
[26] S. Jain and P. Yadav, "Controlled replication based bubble rap routing algorithm in delay tolerant network," in *Proc. Int. Conf. Next Gener. Comput. Technol.* Singapore: Springer, 2017, pp. 70–87.
[27] R. I. Ciobanu and C. Dobre, "Predicting encounters in opportunistic networks," in *Proc. 1st ACM Workshop High Perform. Mobile Opportunistic Syst.*, 2012, pp. 9–14.
[28] X.-K. Zhang, J. Ren, C. Song, J. Jia, and Q. Zhang "Label propagation algorithm for community detection based on node importance and label influence," *Phys. Lett. A*, vol. 381, pp. 2691–2698, Sep. 2017.
[29] D. Li, L. Ma, Q. Yu, Y. Zhao, and Y. Mao, "An community detection algorithm based on the multi-attribute similarity," in *Proc. Int. Conf. Commun.*, Nov. 2013, pp. 118–121.
[30] Y. Yan, Z. Chen, J. Wu, and L. Wang, "An effective data transmission algorithm based on social relationships in opportunistic mobile social networks," *Algorithms*, vol. 11, no. 8, p. 125, 2018.
[31] K. Liu, Z. Chen, J. Wu, and L. Wang, "FCNS: A fuzzy routing-forwarding algorithm exploiting comprehensive node similarity in opportunistic social networks," *Symmetry*, vol. 10, no. 8, p. 338, Aug. 2018. doi: 10.3390/sym10080338.
[32] J. Wu, Z. Chen, And M. Zhao, "Weight distribution and community reconstitution based on communities communications in social opportunistic networks," *Peer-to-Peer Netw. Appl.*., vol. 12, no. 1, pp. 158–166, 2018. doi: 10.1007/s12083-018-0649-x.
[33] J. Wu and Z. Chen, "Human activity optimal cooperation objects selection routing scheme in opportunistic networks communication," *Wireless Pers. Commun.*, vol. 95, pp. 3357–3375, Aug. 2017.
[34] F. Zeng, N. Zhao, and W. Li, "Effective social relationship measurement and cluster based routing in mobile opportunistic networks," *Sensors*, vol. 17, no. 5, p. 1109, 2017.
[35] J. Wu And Z. Chen, "Sensor communication area and node extend routing algorithm in opportunistic networks," *Peer-to-Peer Netw. Appl.*, vol. 11, no. 1, pp. 90–100, Jan. 2018. doi: 10.1007/s12083-016-0526-4.

YEQING YAN is currently pursuing the master's degree with the School of Computer Science and Engineering, Central South University, where she is also a Researcher in mobile health with the China Mobile Joint Laboratory, Ministry of Education. Her main research interests include communications and networking, complex networks, and opportunistic networks.

ZHIGANG CHEN was born in 1964. He received the B.E., M.S., and Ph.D. degrees from Central South University, China, in 1984, 1987, and 1998, where he is currently a Professor, a Supervisor of the Ph.D. degree, and the Dean of the School of Computer Science and Engineering. He is also the Director and an Advanced Member of the China Computer Federation (CCF) and a member of the Pervasive Computing Committee, CCF. His research interests include the general areas of cluster computing, parallel and distributed systems, computer security, and wireless networks.

JIA WU received the Ph.D. degree in software engineering from Central South University, Changsha, Hunan, China, in 2016, where he is currently an Engineer in mobile health with the China Mobile Joint Laboratory, Ministry of Education, and a Distinguished Associate Professor with the School of Computer Science and Engineering. He is a Senior Member of the CCF and a member of the ACM. His research interests include wireless communications and networking, wireless networks, and mobile health in network communication.



LEILEI WANG is currently pursuing the master's degree with the School of Computer Science and Engineering, Central South University. She is also a Researcher in mobile health, China Mobile Joint Laboratory, Ministry of Education. Her research interests include opportunistic networks and vehicular networks.

KANGHUAI LIU is currently pursuing the master's degree with the School of Computer Science and Engineering, Central South University. He is also a Researcher in mobile health with the China Mobile Joint Laboratory, Ministry of Education. His research interests include wireless communications and networking, wireless networks, opportunistic networks, medical decision-making, big data, machine learning, deep learning, and data mining.

PENG ZHENG is currently pursuing the master's degree with the School of Computer Science and Engineering, Central South University. His main research interests include wireless communications, complex networks, and opportunistic networks.