




## Article

# FCNS: A Fuzzy Routing-Forwarding Algorithm Exploiting Comprehensive Node Similarity in Opportunistic Social Networks

Kanghuai Liu <sup>1,2</sup> , Zhigang Chen <sup>1,2,\*</sup>, Jia Wu <sup>1,2,\*</sup>  and Leilei Wang <sup>1,2</sup> 

<sup>1</sup> School of Software, Central South University, Changsha 410075, China; liukanghuai@yeah.net (K.L.); wll\_1234@163.com (L.W.)

<sup>2</sup> “Mobile Health” Ministry of Education-China Mobile Joint Laboratory, Changsha 410083, China

\* Correspondence: czg@csu.edu.cn (Z.C.); jiauwu0510@csu.edu.cn (J.W.); Tel.: +86-133-8748-0797 (Z.C.)

Received: 20 July 2018; Accepted: 10 August 2018; Published: 13 August 2018



**Abstract:** At the dawn of big data and 5G networks, end-to-end communication with large amounts of data between mobile devices is difficult to be implemented through the traditional face-to-face transmission mechanism in social networks. Consequently, opportunistic social networks proposed that message applications should choose proper relay nodes to perform effective data transmission processes. At present, several routing algorithms, based on node similarity, attempt to use the contextual information related to nodes and the special relationships between them to select a suitable relay node among neighbors. However, when evaluating the similarity degree between a pair of nodes, most existing algorithms in opportunistic social networks pay attention to only a few similar factors, and even ignore the importance of mobile similarity in the data transmission process. To improve the transmission environment, this study establishes a fuzzy routing-forwarding algorithm (FCNS) exploiting comprehensive node similarity (the mobile and social similarities) in opportunistic social networks. In our proposed scheme, the transmission preference of the node is determined through the fuzzy evaluation of mobile and social similarities. The suitable message delivery decision is made by collecting and comparing the transmission preference of nodes, and the sustainable and stable data transmission process is performed through the feedback mechanism. Through simulations and the comparison of social network algorithms, the delivery ratio in the proposed algorithm is 0.85 on average, and the routing delay and network overhead of this algorithm are always the lowest.

**Keywords:** opportunistic social networks; routing algorithm; node similarity; fuzzy inference system; data transmission

## 1. Introduction

In recent years, opportunistic networking (OPPNET) [1] has emerged from both mobile ad hoc networks (MANETs) [2] and delay-tolerant networks (DTNs) [3]. It is a type of intermittent connected network (ICN) [4] where the source node communicates with the destination without a complete and reliable end-to-end connectivity. Consequently, as a sort of multi-hop wireless network, opportunistic networks proposed that end-to-end data transmission could be implemented by ‘opportunistic communication’, which is created by the movement of nodes. In opportunistic networks, nodes in unpredictable locations attempt to communicate with each other at unpredictable intervals. Nowadays, the application areas of opportunistic networks mainly include wildlife monitoring networks [5], interplanetary networks [6,7], vehicular networks [8], networks in underdeveloped areas [9,10], and social opportunistic networks [11–14].

At the dawn of 5G networks and big data, technological development is leading to the wide distribution of mobile devices, such as smartphones, Bluetooth, Wi-Fi technologies, or notebooks. Thus, it can be said that opportunistic social networks [10,13] have originated from the social scene in which mobile devices carried by humans have adopted the meeting opportunities to implement peer-to-peer communication between users. A simple example of opportunistic social networks could be how mobile phone users move as they wish and communicate with each other via WiFi, Hotspot, or Bluetooth within each other's communication area. While users are out of reach of each other's transmission domain, the store-carry-forward strategy [4] is commonly applied in the data transmission process of opportunistic social networks, and nodes carrying data packets adopt the encounter opportunities created by movements to forward messages in order to relay nodes or destinations.

Without an instantaneous end-to-end path, message or data transmission between source nodes and destinations may go through a long distance and cannot always be in a connected state. Furthermore, traditional face-to-face transmission methods in opportunistic social networks are difficult to implement complete big data communication without continuous connectivity. Theoretical analysis [4] points out that relay nodes can distribute parts of cache spaces for messages and properly wait for the destination. Therefore, it is particularly important for message applications to choose suitable relay nodes to perform an efficient data transmission process. Additionally, how reliable relay nodes are selected to form the surrounding neighbors is a momentous issue for routing algorithm design in opportunistic social networks.

The social attribute of humans always reflects some special connections between users, which can be a strong basis for choosing between relay nodes. Research [5] has presented a comprehensive survey of human mobility of opportunistic social networks in three main aspects: the attribute characteristics of users, mobility patterns and traces, and mobility prediction methods. The research shows that nodes with a high social similarity are divided into the same community, and the frequent encounter between them indicates that more communication opportunities will exist so as to transmit messages to each other. Consequently, in opportunistic social networks, context-aware routing algorithms based on node similarity attempt to adopt the contextual information related to nodes, as well as the social relationships among nodes, to choose a proper relay node from the surrounding neighbors.

In general, each node in an opportunistic social network may possess a great number of social attributes, and different social attributes represent different relationships between users. Therefore, different routing algorithms based on node similarity consider different similarity factors of nodes in the network, such as interests, locations, workplaces, professions, physical characteristics, or social relations. It is clear that each similarity factor is more likely to play a different role in successful data transmission. This means that transmission mechanisms based on node similarity are supposed to take into consideration a large number of similarity factors, rather than a single one. Moreover, the temporal and spatial attributes of nodes [15] represent the geographical positions of users at different times (the mobile trajectories) in opportunistic social networks. In social scenarios, humans are more likely to repeat the same mobile route every day, such as from places of residence to workplaces, meaning there is a high probability of meeting the same person each day [16]. In other words, two users may encounter each other frequently during an activity cycle if there is a relatively high degree of mobile similarity between them, which indicates message carriers can decide when and where to forward messages to the suitable relay nodes. Consequently, it is able to obtain an enhanced transmission performance to exploit the mobile similarity between users in opportunistic social networks.

However, when evaluating the similarity degree between a pair of nodes, few algorithms [10,12] synthetically consider multiple similarity factors, and mobile similarity is even disregarded. While transmission mechanisms accurately assess the similarity degree between nodes, the biggest difficulty is how to comprehensively evaluate the different effects of each similarity factor and highlight the importance of moving user trajectories in data transmission. On one hand, humans taking along

mobile devices usually move as they like within a certain range [5], and the distance between them is constantly changing. Therefore, only when users are in each other's transmission range, messages can be successfully forwarded to relay nodes. Meanwhile, due to the changing location information, mobile trajectories of users in the networks are difficult to measure. On the other hand, the social attributes of users (except for mobility information) in opportunistic social networks are relatively stable [15], meaning the collection and updating of effective information also becomes a significant basis for defining node similarity. For different application scenarios, each social attribute plays a different role in the transmission environment—thus, it is hard for source nodes or message carriers to comprehensively weigh the influence of a larger number of similarity factors between users.

To solve these urgent issues, this work proposes a fuzzy routing-forwarding algorithm (FCNS) exploiting comprehensive node similarity (the social and mobile similarities) in opportunistic social networks. This strategy employs the social attributes and historical mobile information of nodes to determine the similarity degrees between them. Through reasonable weight adjustments and the fuzzy inference system [17], the transmission preference values are adopted to comprehensively evaluate the social and mobile similarities between nodes and the destination. Meanwhile, according to the feedback mechanism [18], the scheme creatively proposes that messages will be sent to the future two-hop relay nodes, which have a relatively higher transmission preference value than message carriers. In conclusion, the FCNS is a novel routing-forwarding method, which considers the multiple and comprehensive similarity factors of nodes—especially mobile similarity. The contributions of this paper are listed as follows:

- By effectively collecting and updating user information, message applications are able to accurately assess the degree of comprehensive similarity between users (the social and mobile similarities) and determine the special transmission relationships between them.
- According to the information entropy and fuzzy inference system, we synthetically consider the social and mobile similarities of nodes and obtain the level of transmission preferences of each node. Additionally, message carriers can select suitable relay nodes by collecting and comparing the transmission preference values of nodes in opportunistic social networks.
- To achieve an efficient and reliable data transmission process, this scheme uses the feedback mechanism to ensure that the future two-hop relay nodes have a higher transmission preference value than message carriers, which makes the routing and forwarding of messages more continuous and stable.
- In accordance with the simulation results in the Opportunistic Networking Environment (ONE), this novel algorithm (FCNS) enhances performances on the delivery ratio, network overhead, and end-to-end data transmission delay.

The rest of the paper is organized as follows: In Section 2, we will give a brief introduction to the related works; the FCNS routing-forwarding algorithm will be proposed and analyzed in Section 3; in Section 4, the simulation results and performance analysis are provided; the discussion of the paper will be shown in Section 5, and; finally, the conclusion of the paper is shown in the last section.

## 2. Related Works

In recent years, research around the routing algorithm [8,11–14] has been a hot issue in opportunistic social networks, where different methods suitable for different application scenarios have been proposed. In opportunistic social networks, routing algorithms can be roughly divided into two categories: context-aware algorithms, and context-ignorant algorithms. Context-aware routing algorithms, based on node similarity [8,12,15,18], attempt to adopt the contextual information related to nodes and the social relationships between nodes to choose a proper relay node among the surrounding neighbors for the data transmission process. Moreover, in order to improve the transmission environment, the context-aware algorithm needs to process heavy computing tasks and manage large amounts of information. On the contrary, context-ignorant routing algorithms [2,4,6,11,13,14] that

don't utilize the contextual information of nodes are more likely to execute the flooding transmission in opportunistic social networks. Because of the existence of superfluous message group copies in the network, the context-ignorant routing algorithm commonly has very high end-to-end delay and network overhead. Consequently, in this section, we will give a detailed introduction to the state-of-the-art versions of the context-aware and context-ignorant algorithms in opportunistic social networks.

### 2.1. The Proposed Context-Aware Routing Algorithms

In the proposed context-aware algorithms of opportunistic social networks, the similarity degree among nodes is usually used to evaluate transmission relationships between users, such as the future meeting probability between, the mobility patterns of, or the community division of users. Mayer et al. [15] presented a context-based social matching system based on the contextual information of nodes. This algorithm uses the relationships between and the social and personal context of users to predict the matching opportunities of data transmission between nodes. Meanwhile, the matching opportunity between users can be adopted to implement end-to-end communication in the network. To evaluate the impact of the position information of nodes, Kiyoungh et al. [12] proposed a routing algorithm (PaSS) based on the social and positional similarities of nodes in opportunistic social networks. The PaSS scheme adopts the position and contact information of nodes to judge the degree of similarity between them. Moreover, in this algorithm, a node with a relatively high similarity degree to the destination may be a reliable relay node in the data transmission process. Nevertheless, the PaSS routing algorithm doesn't consider the time factor when determining the degree of similarity between a pair of nodes. In order to effectively evaluate the influence of the time factor, Kumar et al. [19] presented a new routing protocol (GD-CAR) based on the varying contextual information of each node. The GD-CAR routing algorithm collects and updates the constantly changing contextual information of nodes, determines the properties of the neighbors by executing the genetic search operation, and eventually forwards messages to the destination. In the GD-CAR scheme, the genetic search operation finds and collects only a small amount of social information of users in the networks, so the transmission relationships between users may not be accurately assessed.

In context-aware algorithms, mathematical methods and models are often used to optimize the routing and forwarding of messages, such as the graph theory, set theory, or the Markov model. Talipov et al. [20] recommended a message delivery algorithm based on the graph theory and the context data of users in opportunistic smartphone networks. This algorithm uses a certain number of message replications to improve the delivery ratio of message forwarding, but only pays attention to the data information transmitted by nodes, which implies that the decision of message delivery is not always accurate in the data transmission process. Satya et al. [21] proposed a context-based routing protocol based on the game theory in opportunistic networks. In this algorithm, the contextual information of nodes is used to establish a stable combination of contexts through the analysis and transformation of the game theory. Deepak, Nguyen, and Nahrstedt [22] recommended a context-based strategy based on human-centric information and crowd-sensing assignment in opportunistic social networks. This method only uses human-centric information of nodes to implement the individual coverage of a social vertex, and the data transmission process can be performed by using the graph theory.

### 2.2. The Proposed Context-Ignorant Routing Algorithms

In opportunistic social networks, context-ignorant algorithms do not use the contextual information related to nodes to make suitable message delivery decisions in the data transmission process. Additionally, according to the context-ignorant algorithm, a mass of message copies will be diffused all over the network and the behavior of relay nodes can be reasonably used to transmit messages to the destination. Dingwei et al. [23] proposed a community division algorithm based on multiple attribute similarities of nodes in opportunistic social networks. This strategy proposed

that each attribute of a node contains a series of keywords, and each keyword corresponds to a feature vector. Each user in opportunistic social networks may possess a great number of social attributes, and each social attribute of a node plays a different role in the data transmission process. Therefore, this algorithm attempts to give different weights to each keyword of node attributes and calculates the social similarity between nodes. Nguyen and Giordano [16] recommended a socially-based routing algorithm using the location prediction of users, where the data transmission in this strategy is divided into two situations—i.e., the destination is an occasional contact node or a frequent contact node. This algorithm adopts a backpropagation neural network (BNN) model [24] to predict the future meeting probability of nodes, and the message carrier makes a suitable decision on when and where to transmit messages to the relay node. In this algorithm, the mobile periodicity of nodes is creatively adopted to predict the location and time of node encounters.

Most context-ignorant algorithms may also focus on the interests of, the trust relationships between, or the influence of users in the community. Xueyang et al. [25] recommended a routing algorithm based on the interest and trust models of users. This algorithm employs the interest and trust relationships of nodes to determine the similarity degrees between them. Meanwhile, the interest similarity between nodes can be obtained through the calculation and updating of users' attributes. In addition, the trust relationships between nodes can be employed to accurately assess the reliability of data transmission in opportunistic social networks. Allen et al. [26] presented a routing-forwarding algorithm based on exploiting user interest similarity and social links for micro-blogging in opportunistic social networks. This scheme pays attention to the activities and interests of nodes, uses the push-based protocols of micro-blogs to forward utterance information, and adopts different methods to evaluate the friendship or interest similarity between nodes in real scenarios. Additionally, interest and trust management is also commonly used to ensure the secure mechanism of opportunistic social networks, Yao et al. [27] proposed a secure routing algorithm (TRSS) based on the social similarity between nodes in opportunistic social networks. The TRSS scheme adopts a novel trust management mechanism to distinguish between malicious, selfish, misbehaving, and normal nodes in the networks, and attempts to determine the node's trustworthiness through special relationships between users.

### 3. System Model Design

In opportunistic social networks, when measuring the node similarity between a pair of nodes, most traditional routing algorithms based on node similarity mainly focus only on a single factor, such as the number of common neighbors, the position similarity, or the number of common social attributes. Additionally, none of them takes the mobility similarity between nodes into consideration. It should be noted that mobility similarity needs to consider both the time and location information of the node in the network, while position similarity just takes the location information into account. Consequently, we present a fuzzy routing-forwarding algorithm exploiting a comprehensive node similarity, which is the combination of the mobility and social similarities of nodes. Next, we will give a detailed introduction to the proposed algorithm.

#### 3.1. Collecting Enough Information about the State of the Network

In our proposed strategy, there exists a special warming-up period in which each node gains and updates adequate and accurate information about the state of the network. The information about the state of the network is collected and updated only at the warming-up phase, meaning this phase is actually a period for data collection and updating. The length of the warming-up phase  $\Delta T$  is generally set according to the user's activity cycle [15]. In the real social scene, users are more likely to repeat the same mobile route every day, such as from places of residence to workplaces, and there is a high probability of meeting the same person each day [16]. Therefore, the optimal value for the warming-up phase in the real social scene should be set to 24 h. After the warming-up phase, message carriers begin to make suitable message delivery decisions by the fuzzy evaluation of social and mobile similarities of nodes in the network. Moreover, in order to quantify the whole process of information collection



and updating in the warming-up phase, we define the following Equations (1) and (2) as the state sequence of node  $u$ .

$$List(u) = \langle GP_u, \gamma_u, *List_m \rangle \quad (1)$$

$$*List_m = \langle List(b), List(c) \dots List(n) \rangle \quad (2)$$

where  $GP_u$  represents the mobile distribution of the node  $u$ , and  $\gamma_u$  represents a feature vector of the node  $u$  that includes different social attributes of this node.  $*List_m$  is a state sequence set of the other nodes encountered by the node  $u$  in a given warming-up period  $\Delta T$ . Meanwhile, if a node  $u$  meets another node  $v$ , we adopt the following Equation (3) to construct a table item  $List(1)$  for the meeting information between nodes  $u$  and  $v$ . Moreover, the encounter status matrix  $S(n)$  for all table items can be established by the following Equation (4).

$$List(1) = List(u) \ll = \gg List(v) = \langle GP_u, \gamma_u, GP_v, \gamma_v, *List_m \rangle \quad (3)$$

$$S(n) = (List(1), List(2), List(3) \dots List(n)) \quad (4)$$

Consequently, during the warm-up period, each node adopts its own mobile information, social attributes, and historical meeting information to build a state sequence  $List$ . As there are movements of nodes in the network, when a node encounters another node, the two nodes update their own  $*List_m$  through the state sequence information obtained from each other. Furthermore, the information exchanging process can also be completed through the cooperation of multiple nodes in the network. For each encounter between nodes, this scheme builds a table item that contains all information about the meeting of the two nodes. Moreover, in order to establish a unified data set about the node similarity, the encounter status matrix  $S(n)$  will be built in a given warming-up period  $\Delta T$ . Unlike the social attribute of nodes in the network, the mobile information changes over time, so promptly updating the mobile information of users can facilitate the timeliness, uniformity, and accuracy of the statistics collected.

More specifically, as shown in Figure 1, when node  $u$  encounters node  $v$ , the node  $u$  sends  $List(u)$  to the node  $v$ , and the node  $v$  also transmits its own state sequence  $List(v)$  to the node  $u$ . Moreover, through the state sequence obtained from each other, the two nodes update their own encounter status matrix  $S(n)$  by adding this meeting information  $List(1)$  to their own cache spaces. As there is movement of nodes in the network, the node  $u$  may encounter another node  $w$ , and the two nodes will continue to update the encounter status matrix  $S(n)$ , which effectively establishes a unified dataset about the state of the network. After the warming-up period  $\Delta T$ , message carriers start to make suitable message delivery decisions by the fuzzy evaluating of social and mobile similarities of nodes in the network. When the whole process of data transmission is complete, each node in the network clears the encounter status matrix  $S(n)$  from its cache spaces and enters into the next warming-up phase, which indicates that the FCNS algorithm is a scheme based on periodic updates.

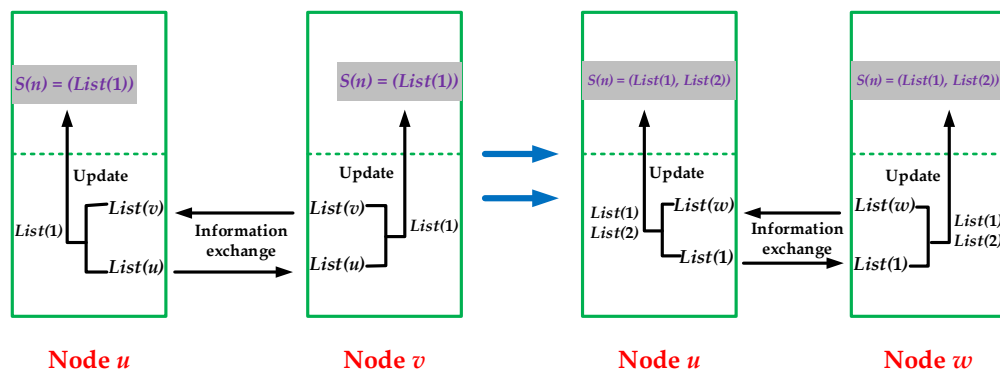


Figure 1. Information collection and updating in the warming-up phase.

### 3.2. Calculation of Node Similarity

In general, the mobile information of nodes is also a type of social attribute. However, due to its specialty in the data transmission process, the proposed algorithm attempts to distinguish the mobile information from other social attributes and analyzes it separately. In the following section, we first calculate the mobile similarity between nodes.

#### 3.2.1. Mobile Similarity Calculation

Similar to the research on social relationships between nodes in social networks, mobile similarity, based on both spatial and temporal information, can also describe the closeness degree of node relationships. Generally, the higher the similarity of two nodes on the geographic trajectory, the higher the probability of successful message forwarding between them will be. According to this theory, the proposed algorithm means that the moving trajectory of the node in an opportunistic social network can be represented by the combination of the time stamp and communication area such as  $\langle t_1, l_1 \rangle, \langle t_2, l_2 \rangle, \dots, \langle t_n, l_n \rangle$ , where  $t$  represents the time stamp and  $l$  is the communication area. In addition, we define  $u$  as a node's label, and  $n$  represents the number of voices or data services between the node  $u$  and others. In previous sections we have already defined  $GP(u, r)$  as the mobile distribution of the node  $u$ , so it can be computed by the following formula:

$$GP(u, r) = \sum_{i=1}^n \frac{\delta(r, l_i(u))}{n(u)} \quad (5)$$

where  $r$  is the location set of the communication area in an opportunistic social network. In the above Equation (5), when  $r = l_i(u)$ ,  $\delta(r, l_i(u)) = 1$ ; otherwise, it is equal to 0. This means that only when the node  $u$  contacts with others in the same communication area  $r$ , the value of  $\delta(r, l_i(u))$  is 1.

$$MS(u, v) = \frac{\sum_{i=1}^{n(u)} \sum_{j=1}^{n(v)} \theta(\Delta t - |t_i(u) - t_j(v)|) (\delta(l_i(u), l_j(v)))}{\sum_{i=1}^{n(u)} \sum_{j=1}^{n(v)} \theta(\Delta t - |t_i(u) - t_j(v)|)} \quad (6)$$

Meanwhile, considering the time factor, we strictly define  $\Delta t$  as the time precision, which reflects the proportion of all nodes in the same geographical location at the time. Moreover, taking into consideration the impact of working time or nonworking time, we set different weight values  $\theta$  for different time periods. During working hours, humans in real-life situations may repeat a similar moving trajectory every day [16], such as from where they live to where they work, so the weight value  $\theta$  for this period may be a bit bigger. On the contrary, because human activities are unlimited during non-working hours [15], they are more likely to move randomly in the communication area. To reduce the impact of non-working time, we should give a small value for the weight  $\theta$  of this period. Consequently, the mobile similarity  $MS(u, v)$  between nodes  $u$  and  $v$  can be calculated by the above Equation (6), where  $t_i(u)$  represents the  $i$ -th communication time of the node  $u$ ,  $l_j(v)$  is the geographical location of the node  $u$  on the  $j$ -th communication,  $|t_i(u) - t_j(v)|$  represents the time interval between nodes  $u$  and  $v$ , and  $(\delta(l_i(u), l_j(v)))$  is the position similarity between nodes  $u$  and  $v$ .

#### 3.2.2. Social Similarity Calculation

As described in the previous sections, we have already analyzed the degree of mobile similarity between nodes. In this section, we will assess the social similarity between nodes by exploring the correlations of social attributes of nodes.

Firstly, we define the following Equation (7) as a feature vector of the node  $u$ , where  $A_u, B_u, C_u, D_u \dots N_u$  represents different values of social attributes of the node  $u$ . In fact,  $A_u, B_u, C_u, D_u \dots N_u$  are different sub-vectors corresponding to different social attributes of the node  $u$ . To obtain detailed comparisons of the social similarity, we also define different feature words for each

sub-vector in the feature vector, such as the interest sub-vector  $A_u = (a_1, a_2, a_3, a_4)$ , where  $a_1, a_2, a_3, a_4$  respectively represent different feature words (such as basketball, game, music or film). If the node  $u$  possesses feature words in the sub-vector, the corresponding value is set to 1; otherwise it is set to 0. For instance, if the node  $u$  likes basketball and film, its sub-vector of interest would be  $A_u = (1, 0, 0, 1)$ . Moreover, for the same sub-vector between a pair of nodes  $u$  and  $v$ , we use the following formula (8) to compute the sub-vector similarity between them:

$$\gamma_u = (A_u, B_u, C_u, D_u \dots N_u) \quad (7)$$

$$\text{sim}(\beta_u, \beta_v) = \frac{\beta_u \cdot \beta_v + \varepsilon}{\max(\|\beta_u\|^2, \|\beta_v\|^2) + \varepsilon} + \frac{\min(\|\beta_u\|^2, \|\beta_v\|^2) - \beta_u \cdot \beta_v}{\|\beta_{\max}\|^2} (u \neq v) \quad (8)$$

where  $\beta_{\max}$  is a full sub-vector and all elements in it are 1,  $\varepsilon$  is the minimum sub-vector and  $\|\beta_u\|$  represents the sub-vector model of the feature vector of the node  $u$ . Therefore, we can easily calculate that the range of  $\text{sim}(\beta_u, \beta_v)$  is from 0 to 1. Additionally, in order to evaluate the similarity degree of multiple social attributes between a pair of nodes, we employ the following Equation (9) to calculate the social similarity between nodes  $u$  and  $v$ :

$$SS(\gamma_u, \gamma_v) = w_1 \text{sim}(A_u, A_v) + w_2 \text{sim}(B_u, B_v) + w_3 \text{sim}(C_u, C_v) + \dots + w_n \text{sim}(N_u, N_v) (u \neq v) \quad (9)$$

where  $w_1, w_2, w_3 \dots w_n$  represent different weights of the sub-vector similarities between nodes  $u$  and  $v$ . In order to determine the weights of the sub-vector similarity between two nodes, the algorithm uses an improved entropy evaluation method to compute the weights  $w_1, w_2, w_3 \dots w_n$  and the weight decision matrix  $W$  can be shown as:

$$W = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \quad (10)$$

where  $x_{ij}$  represents the module of the  $j$ -th sub-vector of the  $i$ -th node. Moreover, the contribution degree of the  $j$ -th attribute value of the  $i$ -th node can be shown as:

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (11)$$

To evaluate the impact of all nodes in the network on social similarity, we define  $E_j$  as the total contribution degree of the  $j$ -th attribute value of all nodes, and  $E_j$  can be computed by the following Equation (12), where the range of  $E_j$  is from 0 to 1. When the contribution degree of a certain attributional value of each node tends to be equal,  $E_j$  is equal to 1.

$$E_j = -\frac{1}{\ln(m)} \sum_{i=1}^m p_{ij} \ln(p_{ij}) \quad (12)$$

Therefore, differences in the contributory degree can be used to determine the weight of the sub-vector. We define  $d_j$  as the consistent degree of the contribution of each node at the  $j$ -th social attribute, and it can be shown as  $d_j = 1 - E_j$ . Additionally, the weights of each sub-vector are shown as the following Equation (13). If the message sender can give the subjective estimate weight  $\lambda_j$



based on its historical experience, the corrected weight value  $w_j^*$  can be computed by the following Equation (14).

$$w_j = \frac{d_j}{\sum_{j=1}^n d_j} \quad (13)$$

$$w_j^* = \frac{\lambda_j w_j}{\sum_{j=1}^n \lambda_j w_j} \quad (14)$$

In conclusion, the purpose of exploring the social similarity is to divide nodes in the network into different communities, and the nodes belonging to the same community may have more opportunities to transmit data packages to each other, which signifies that the efficiency of data transmission between these nodes can be improved greatly.

### 3.3. Using the Fuzzy Inference System to Compute the Transmission Preference

Node similarity measures the special relationships between nodes that are conducive to the data transmission process in opportunistic social networks, such as the distance between nodes, the probability of the future meeting between nodes, or the community division of nodes. However, the unstable factors between nodes may cause changes in their similarity. In addition, if the priority of data transmission between nodes is measured through the variational information of node similarity, the node in the network may obtain inaccurate statistics about the state of the network. Therefore, we have tried using the node similarity to obtain a transmission metric representing the fuzzy level of message delivery between nodes, which effectively avoids the influence of uncertain node similarity. We creatively employ the fuzzy inference system to determine the grade of mobile and social similarities and to implement the calculation of transmission preferences between nodes in our proposed algorithm. Moreover, we adopt the Mamdani fuzzy system [17] as our fuzzy inference system because of its extensive applicability. The fuzzy system consists of three components: the fuzzifier, fuzzy inference, and defuzzifier. In the following sections, we will describe the implementation of each component of the fuzzy inference system in detail.

#### 3.3.1. The Fuzzifier Component

The fuzzifier is a common component that is used to evaluate the membership degree of fuzzy sets in the fuzzy inference system. In our fuzzifier component, there are two different input variables (the social and mobile similarities) and we define three different fuzzy sets (low, medium, and high) for each variable. Moreover, each fuzzy set in the fuzzifier can be computed by a membership function. Consequently, we need to define three different membership functions for the fuzzy set in the fuzzy inference. In the traditional membership function, there exist many different membership function paradigms that can be applied to different scenarios, such as rectangular, triangular, trapezoidal, or inverse. Due to the normal distribution [28] of the movement of nodes in opportunistic social networks, the normal membership function (15) [29] will be defined to assess the membership degree of each input in our inference system.

$$f_{\mu}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\left(\frac{x-\mu}{2\sigma^2}\right)^2} \quad (15)$$

$$\begin{cases} f_{low}(x) = \frac{1}{\sqrt{2\pi} \times 0.3} e^{-\left(\frac{(x-0.2)^2}{2 \times (0.3)^2}\right)} \\ f_{medium}(x) = \frac{1}{\sqrt{2\pi} \times 0.3} e^{-\left(\frac{(x-0.5)^2}{2 \times (0.3)^2}\right)} \\ f_{high}(x) = \frac{1}{\sqrt{2\pi} \times 0.3} e^{-\left(\frac{(x-0.8)^2}{2 \times (0.3)^2}\right)} \end{cases} \quad (16)$$

Additionally, for the sake of the requirement of low computation in mobile devices, the three different membership functions (low, medium, and high) for the node similarity can be defined as the above system of Equations (16), and their geometric graph is shown in Figure 2. In mathematical expressions of the three membership functions, different values of the node similarity correspond to three different levels in the above system of Equations (16). For instance, when  $MS(u, v) = 1/2$ , it corresponds to  $f_{low}(1/2) = 0.8066$ ,  $f_{medium}(1/2) = 1.33$ , and  $f_{high}(1/2) = 0.8066$  for the three different fuzzy sets of mobile similarity.

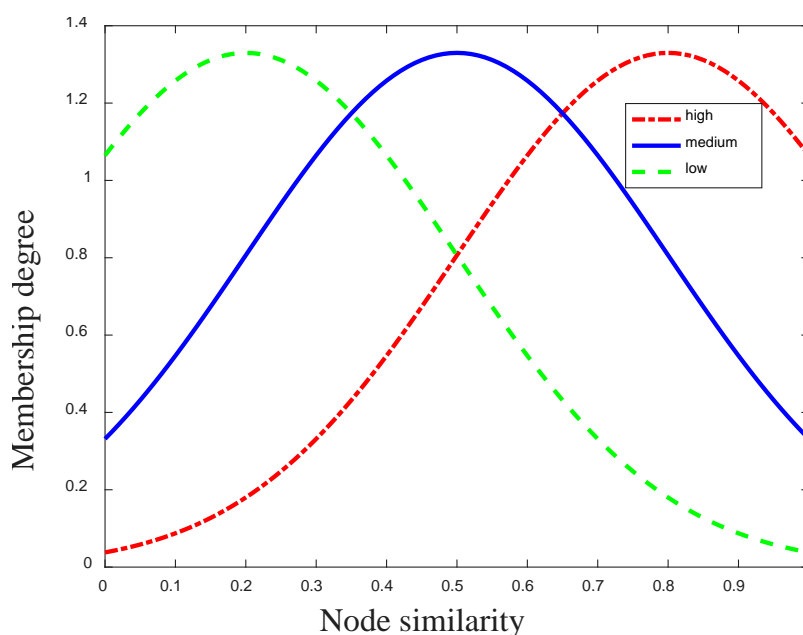


Figure 2. Three different membership functions of the normal distribution.

### 3.3.2. The Fuzzy Inference Component

In the Mamdani fuzzy system, each input variable (mobile similarity or social similarity) corresponds to three different fuzzy sets which are calculated by the three different membership functions, meaning the mobile and social similarities correspond to six membership degrees and possess nine different combinations of the fuzzy sets according to the fuzzy logic. Additionally, the nine fuzzy rules are concisely presented in Table 1. Based on the “If-Then” rules in Table 1, the node, in an opportunistic social network, evaluates the levels of data transmission preference by analyzing the comprehensive consideration of  $MS(u, v)$  and  $SS(\gamma_u, \gamma_v)$ , and the message carrier executes the suitable forwarding operation for messages.

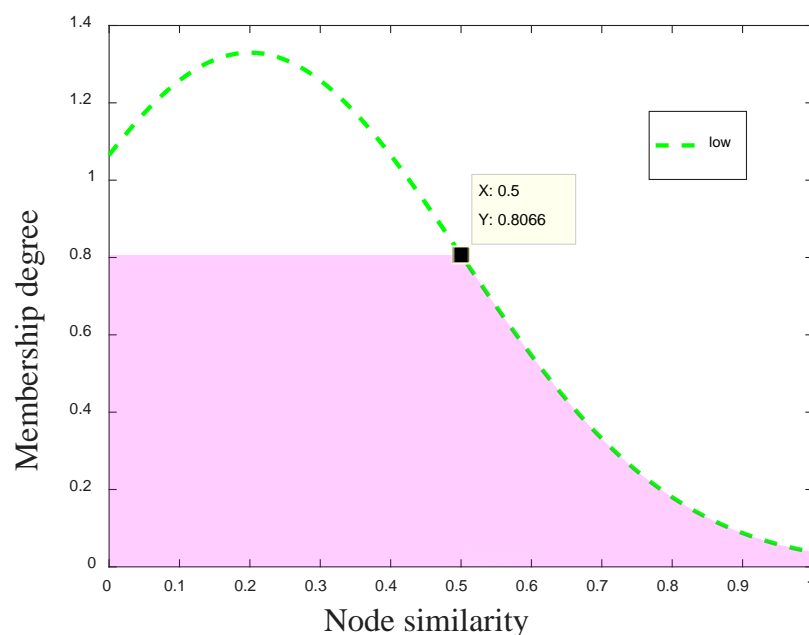
Table 1. If-Then rules for our proposed fuzzy inference system.

Rule No.	If (And)		Then (Transmission Preference)
	$MS(u, v)$	$SS(\gamma_u, \gamma_v)$	
1	High	High	Definitely strong (level 1)
2	High	Medium	Very strong (level 2)
3	High	Low	Stronger (level 3)
4	Medium	High	Strong (level 4)
5	Medium	Medium	Middle (level 5)
6	Medium	Low	Weak (level 6)
7	Low	High	Weaker (level 7)
8	Low	Medium	Very weak (level 8)
9	Low	Low	Definitely weak (level 9)

As shown in Table 1, the impact of the mobility similarity of nodes on data transmission is obviously deeper than that of the social similarity between them, because the mobility similarity mainly reflects the degree of overlap between the moving trajectory for a pair of nodes. Moreover, a pair of nodes may obtain more meeting opportunities to transmit messages to each other if there is a relatively high node similarity between them. Otherwise, social similarity is usually used to divide nodes with a similar attribute characteristic into the same community, and the probability of transmitting messages between the nodes belonging to the same community is also relatively high.

### 3.3.3. The Defuzzifier Component

The proposed algorithm adopts the AND operation and the OR operation to compute the transmission preference of the node in an opportunistic social network according to the Mamdani fuzzy system. In our inference system, we first maximize each value of different fuzzy sets by the OR operation, and then calculate the minimum combination of these fuzzy values through the AND operation. Consequently, the OR operation is employed to obtain the maximum shaded area of each fuzzy value, and the AND operation is used to get the overlap shadow area of the six maximum fuzzy shadows of two input variables. For example, as shown in Figure 3, when  $MS(u, v) = 1/2$ , it corresponds to  $f_{low}(1/2) = 0.8066$ . After the OR operation, the corresponding maximum shadow of the low membership function is shown as the pink area in Figure 3.

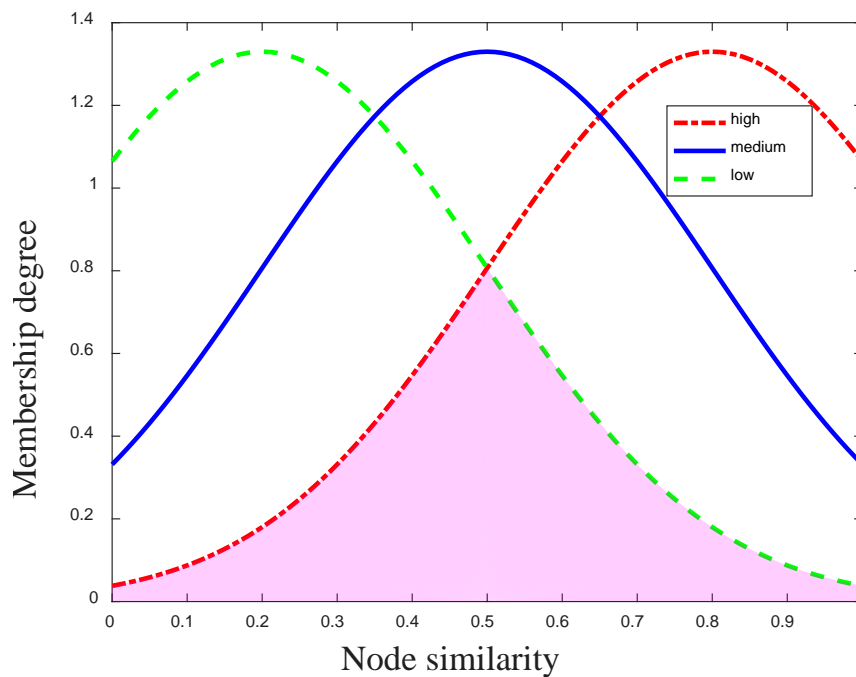


**Figure 3.** The controlling result of low membership function.

In the Mamdani fuzzy system, the maximum shaded area accurately represents the controlling result of each membership function. As mentioned above, there are six membership functions for two input variables (the mobile and social similarities). Hence, there are six controlling results with different sizes for six membership functions in our fuzzy inference system. Moreover, the transmission preference can be gained by the minimum overlapping shadow of these six controlling results, and it represents the final recommendation result after the fuzzy inference of the mobile and social similarities.

Following the above example, when  $MS(u, v) = 1/2$  and  $SS(\gamma_u, \gamma_v) = 1/2$ , the overlapping area of six controlling results is shown as the pink area in Figure 4 after the AND operation.

$$T_{(u,v)} = \frac{\sum_{i=1}^n y_i \times x_i}{\sum_{i=1}^n y_i} \quad (17)$$



**Figure 4.** The controlling result of three membership functions.

Finally, because the movement of the node in an opportunistic social network commonly obeys the normal distribution [28], we creatively used the centroid of the overlap shaded area to represent the transmission preference value. Moreover, the transmission preference value can be calculated by the above Equation (17), where  $x_i$  is the value of the node similarity between nodes  $u$  and  $v$ ,  $y_i$  is the membership degree of the node similarity between nodes  $u$  and  $v$ , and  $n$  is the number of coordinates at the boundary of the overlapping shaded area in Figure 4. During the data transmission process in opportunistic social networks, the message carrier makes the optimal message delivery decision by collecting and comparing the transmission preference values of nodes in the network.

### 3.4. Forwarding Messages by the Feedback Mechanism

As mentioned in the previous sections, each node in the network first gathers enough and accurate information about the state of the network, and then begins to calculate the transmission preference with the destination through the fuzzy evaluation of mobile and social similarities. If there exists a higher similarity between a pair of nodes, there will be a higher probability of successful data transmission between them. Therefore, we proposed that the source node, or message carriers, would seek the future two-hop nodes that possess a higher transmission preference with the destination, rather than their own, as the suitable relay nodes. In the data transmission process of opportunistic social networks, this efficient method is called the feedback mechanism.

Ordinarily, traditional routing algorithms in opportunistic social networks select a suitable relay node only from the neighbors, but the feedback mechanism in our proposed algorithm pays attention

to not only the neighbor nodes, but also the second-hop nodes. As shown in Figure 4, the source node  $A$  needs to send a message to the destination  $N$ . After the warming-up phase, each node obtains detailed information about the state of the network and computes the transmission preference value using the destination  $N$ . After that, the source node  $A$  starts to collect and compare the transmission preference values from its future two-hop nodes. The source node forwards messages to the future two-hop nodes, which have a higher transmission preference with the destination than its own. In Figure 5, the nodes  $B$ ,  $C$ , and  $D$  are neighbors of the source node  $A$ ; the nodes  $E$ ,  $F$ , and  $G$  are neighbors of the node  $B$ . Because  $T(E, N) > T(A, N)$  and  $T(B, N) > T(A, N)$ , the source node  $A$  only sends messages to the neighbor node  $B$ , and the node  $B$  then forwards messages to the second hop  $E$ . Finally, data packages will be delivered to the destination node  $N$  based on the feedback mechanism.

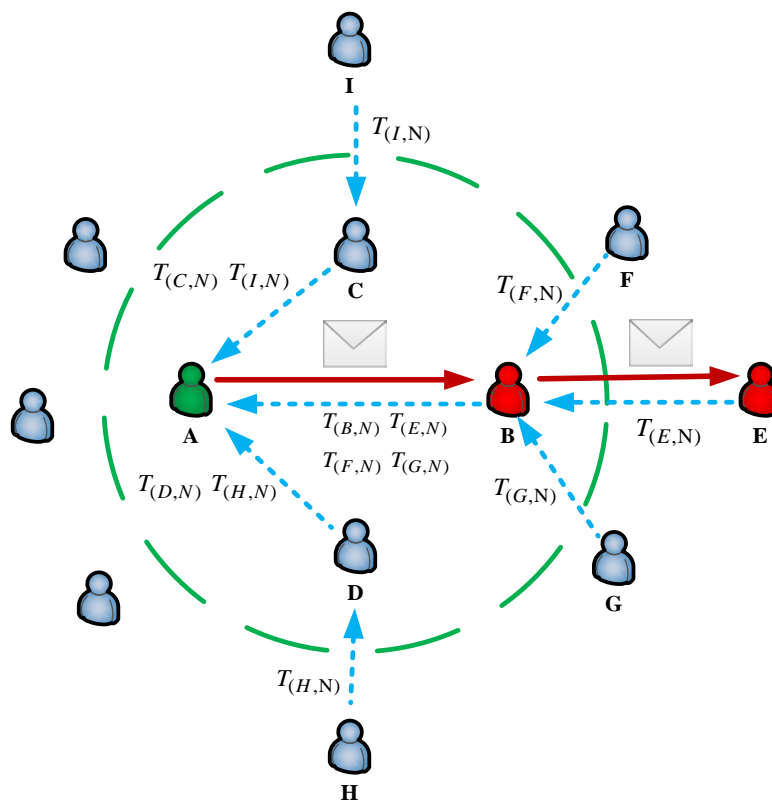


Figure 5. The selection diagram of the future two-hop nodes.

### 3.5. Complexity Analysis of this Algorithm

In summary, we proposed a fuzzy routing-forwarding algorithm exploiting comprehensive node similarity in opportunistic social networks. Meanwhile, to improve the understanding of the proposed algorithm, concrete steps of the data transmission process in this algorithm are listed as follows:

1. During the warming-up phase, each node in the network collects its social attributes and mobile information, constructs a state sequence about these collected statistics, and shares the state sequence with others it encounters.
2. Each node begins to calculate the mobile and social similarities with the destination by the updated state sequence. In opportunistic social networks, if there is a node possessing a relatively high social or mobile similarity with the destination, it is more likely to be a friend of that destination or to have a mobile trajectory highly similar to the destination, meaning the node must be a reliable relay node for the data transmission process.

3. According to the Mamdani fuzzy system, each node evaluates the membership degrees of the similarity with other nodes, determines the level of the transmission preference with the destination, and calculates the transmission preference value with the destination.
4. The source node or message carriers make appropriate message delivery decisions by collecting and comparing the transmission preference values from its neighbors and the second hop relay nodes. The message carrier will transmit data packages to its future two-hop nodes which have a higher transmission preference with the destination than its own. Ultimately, data packets are forwarded to the destination node by the feedback mechanism, which is beneficial to keep the continuity and high efficiency of the data transmission in opportunistic social networks.

For the readability of the FCNS algorithm, we rigorously constructed Algorithm 1 to introduce our proposed algorithm in detail. Specifically, during the warming-up phase, each node shares its state sequence with the other nodes it encounters and updates the encounter status matrix through the obtained state sequences, so the time complexity of this phase is  $O(\log_2 n)$ . Based on the theory of the fuzzy inference system, each node in the network is able to accurately compute the transmission preference values—thus, the time complexity of this process is  $O(n)$ . Finally, message carriers collect and compare the transmission preference values from its future two-hop relay nodes, and then make the suitable messages delivery decisions—hence, the time complexity of messages forwarding is  $O(\log n)$ . On the whole, through rigorous mathematical analysis, the time complexity of the FCNS algorithm can be computed as  $O(\log_2 n + n + \log n) = O(n)$ . In retrospect, the time complexity of Spray and Wait [30] is  $O(n \log n)$  and the time complexity in the Epidemic routing algorithm [31] is  $O(n)$ .

---

**Algorithm 1** The fuzzy routing-forwarding algorithm exploiting comprehensive node similarity

---

**Input:** node  $A$ , node  $B$ , node  $C$  and the destination node  $D$

**Output:**  $T(A, D)$ ,  $T(B, D)$  and  $T(C, D)$

---

```

1  Begin
2      The calculation of transmission preferences:
3      Collect enough information about all nodes;
4      If ( $A.isMessageCarrier()$ ) then
5          Give different weights for each sub-vector of the feature vector;
6          Compute  $MS$  and  $SS$  between each node and the destination;
7          For (each membership function of the fuzzifier component) do
8              Compute the membership degrees of mobile and social similarities;
9          End for
10         Determine the level of transmission preferences of the node based on “If-Then” rules;
11          $T = fuzzy(MS, SS)$ ;
12         Output  $T$ ;
13     End if
14     Forwarding messages through the feedback mechanism:
15     If ( $B.isNeighbor(A) \wedge C.isNeighbor(B)$ ) then
16         If ( $T(B, D) > T(A, D) \wedge T(C, D) > T(A, D)$ ) then
17              $A$  sends data packets to  $B$ ;
18              $B$  sends data packets to  $C$ ;
19         End if
20     End if
21 End

```

---



## 4. Simulations

### 4.1. Simulation Parameters and Simulation Process

The simulation adopts the Opportunistic Network Environment (ONE) to test in the real environment and to evaluate the experimental performance of the FCNS routing algorithm. There are different mobile models that are adopted to depict the moving trajectory and to analyze the state of the network in the ONE. Additionally, in our work, the FCNS routing algorithm will be compared with other four algorithms: Spray and Wait [30], Epidemic [31], EIMST (Effective Information Transmission Based on Socialization Nodes) [2], and ICMT (Information Cache Management and Data Transmission Algorithm) [13]. EIMST and ICMT are the latest routing algorithms for opportunistic social networks, while the other two algorithms are typical and traditional methods.

In this experiment, the parameters are set as follows: The communication area is  $3000\text{ m} \times 3000\text{ m}$ , and it is divided into 9 equal geographical areas, each of which occupies the range of  $1000\text{ m} \times 1000\text{ m}$ . The warming-up time  $\Delta T$  is 20, 25, 30, 35, 40, 45 and 50 min, and it is initialized to 20 min. The total simulation time is 2–6 h and the number of nodes is 500. Moreover, the movement of all nodes in the communication area follows the HCMM (Health Capability Maturity Model) mobility model [32]. Because the HCMM model is based on community division, the communication between nodes of the same community is more frequent, and the weight value  $\theta$  can reasonably be set to 0.8 [15,16]. Additionally, the time precision  $\Delta t$  is usually set to 1 h in the prediction of the coincidence between geographics and time between users. The initial energy for a node is 100 J, and each node carrying 10 data packets consumes 0.25 J energy to send a data packet. The cache space of a node is set to 10, 15, 20, 25, 30, 35, and 40 Mb, and its initial value is 10 Mb. The speed of a node ranges from 1 to 9 m/s, which may be the widespread movement speed of humans, animals, or vehicles. The social attribute for each node includes occupation, interest, residence, workplace, and physical characteristics. In addition, each of the social attributes consists of 5 different feature words. For example, the social attribute of interest may contain the feature words: basketball, game, music, film, and reading. Consequently, there are 5 different sub-vectors and 25 different feature words for the feature vector of a node in the network. Moreover, the 25 feature words are randomly assigned to each node in the simulation area. In order to improve the readability of the above simulation settings, we constructed Table 2 to demonstrate the simulation environment of this experimentation.

**Table 2.** The simulation environment for the fuzzy routing-forwarding algorithm (FCNS) algorithm.

Simulation Environment	Description
Simulator	Opportunistic Network Environment (ONE)
Mobility model	HCMM (Health Capability Maturity Model)
Communication area ( $\text{m}^2$ )	$3000 \times 3000$
Warming-up time (min)	20 (Initial value), 25, 30, 35, 40, 45, 50
Total simulation time (h)	2–6
Number of nodes	500
Cache space of a node (Mb)	10 (Initial value), 15, 20, 25, 30, 35, 40
Speech of a node (m/s)	1–9
Initial energy for a node (J)	100
Number of feature words	25

The FCNS routing algorithm is compared with the four algorithms mentioned above to evaluate its performance. Additionally, the simulation result mainly concentrates on the following parameters:

**Delivery ratio:** This parameter represents the probability of selecting a proper relay node during the data transmission process. The delivery ratio can be computed as  $D_{node} = D_{rec} / D_{sen}$ , where  $D_{sen}$  is the number of messages forwarded by nodes and  $D_{rec}$  is the number of messages received by the surrounding neighbors.

**Overhead on average:** This parameter represents the network overhead of a successful message delivery between a pair of nodes. Overhead on average  $O_{node}$  can be calculated as  $O_{node} = (T_{tr} - T_s) / T_{tr}$ , where  $T_{tr}$  is the total time of message forwarding and  $T_s$  is the total time of successful data transmission between nodes.

**Average end-to-end delay:** This parameter refers to the delay of routing selections, of relay nodes waiting for messages, and of message forwarding. Average end-to-end delay can be shown as  $\theta_{del} = \theta_{sum} / \chi_{node}$ , where  $\theta_{sum}$  is the total delay that is accumulated by each node and  $\chi_{node}$  is the number of nodes that successfully obtain the message.

## 4.2. Simulation Result Analysis

### 4.2.1. Performance Evaluation of the FCNS Algorithm

This section mainly describes the performance of the FCNS algorithm with various warming-up times. During the warming-up phase, all nodes in the communication area move randomly and use the meeting opportunity to exchange state sequence information between them. Therefore, the length of the warming-up time determines the node's understanding of the network state, which has a significant influence on the transmission performance of the FCNS algorithm. According to the simulation results, when the total experimental time is set to 2–6 h and the warming-up time is at 35–45 min, the FCNS algorithm shows the optimal performance in the transmission environment.

Firstly, we explore the impact of warming-up time on delivery ratio. As shown in Figure 6, the broken line first rises sharply, and then falls rapidly. In other words, when the value of the warming-up time is increasing, the delivery ratio rises at first and then decreases, and reaches the maximum value 0.95 when the length of the warming-up phase is 40 min. As the warming-up time increases, all nodes in the network may obtain more information about the state of the network. Accurate calculation results of node similarity will be acquired by the message carrier, and the source node can select suitable and reliable relay nodes through these precise transmission statistics. However, a long warming-up phase means each node in the network needs to carry, transport, and process more irrelevant information in the data transmission process. Meanwhile, frequent communication of information between nodes will consume a lot of the energy of mobile devices in the network, and a severe lack of energy may result in the slow movement of users. Additionally, both information collection and message forwarding require some computing tasks, so these computing tasks executed by nodes also have to consume the cache spaces and energy of users, which may cause large resource consumption and reduced delivery ratios in the network.

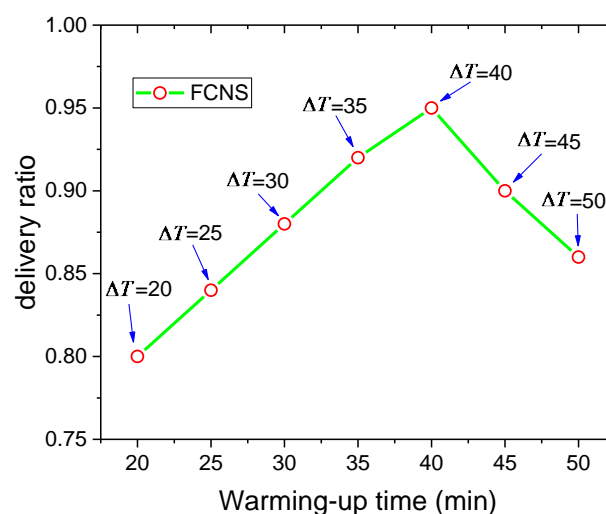


Figure 6. Delivery ratio with various warming-up times.

Secondly, the influence of the warming-up time on average end-to-end delay is demonstrated in Figure 7. As presented, the broken line first declines rapidly and then rises slowly. When the value of the warming-up time is 35 min, the average end-to-end delay of the FCNS algorithm achieves the minimum value of 80. With the increase of the warming-up time, frequent information exchanges between nodes can establish sufficient statistics about the attribute characteristics of the node in the network. Moreover, the message carrier is able to choose the optimal two-hop nodes and less time is spent on successful data transmission. However, for an overly long warming-up time, excessive information carried by nodes in the network may lead to the lack of energy and memory of nodes, implying that it takes more time to route and transmit messages in the data transmission process, and the average end-to-end delay then begins to go up.

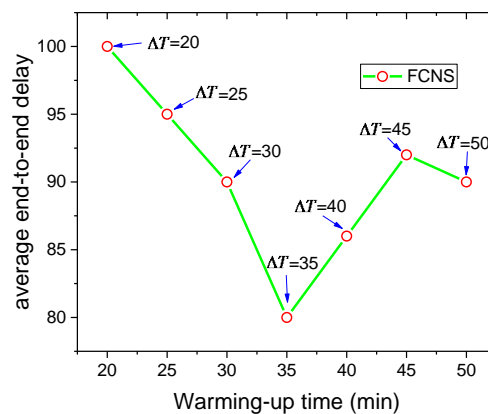


Figure 7. Average end-to-end delay with various warming-up times.

Finally, Figure 8 shows the impact of the warming-up time on the overhead on average. When the value of the warming-up time increases constantly, the corresponding value of the overhead on average continues to decline and never exceeds 300. The later the warming-up time occurs, the deeper the understanding of the network state obtained by the node will be. Meanwhile, more nodes participate in the routing and forwarding of messages. It is efficient and stable to transmit data packages between nodes through the feedback mechanism, and the number of hops decreases significantly in successful data transmission. Consequently, less time is spent and shorter distances are experienced in complete end-to-end communication between nodes. Moreover, the computing resources and cache spaces of nodes can be utilized effectively, and the overhead on average in the network continues to decrease from 300 to 140.

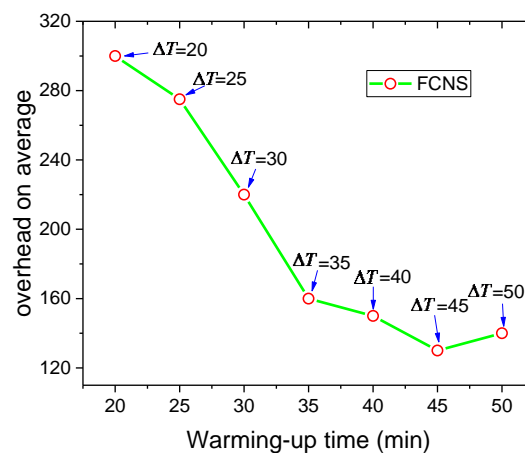


Figure 8. Overhead on average with various warming-up times.

#### 4.2.2. Comparison Result Analysis between Algorithms

This section mainly focuses on the comparison and analysis between the five different algorithms above. For the algorithms based on contextual information in opportunistic social networks, nodes in the network have to carry, transport, and forward large amounts of text information about the routing state, but the insufficient cache space of nodes is a big limitation in the data transmission process. Therefore, in the experiment, we reasonably set the cache space as a variable to explore the transmission performance of each algorithm. Additionally, the comparison results show that the FCNS algorithm always performs better than the other four algorithms in the delivery ratio, average end-to-end delay, and network overhead. More specifically, in the FCNS algorithm, the delivery ratio is 0.85 on average, and the average end-to-end on delay and the overhead on average are the lowest among these algorithms.

Above all, we evaluated the delivery ratio of each algorithm with various cache spaces, and the comparison results between the five different algorithms in terms of delivery ratio are shown in Figure 9. On the whole, where the cache space of nodes continues to increase, the delivery ratio of all algorithms is growing constantly. This is because the node in the network is able to operate complex computing tasks and to carry more information. In addition, the delivery ratio in the FCNS algorithm is always the highest among these algorithms, because successful data transmission in the proposed scheme can be implemented by node similarity and contextual information. In the FCNS strategy, nodes in the network are divided into different communities through social similarity, and a pair of nodes in the same community with a high mobile similarity may communicate with each other frequently. Meanwhile, the message carrier executes the fuzzy routing and forwarding operation for messages, which is an efficient method to synthetically consider the mobile and social similarities of nodes, so the delivery ratio of the FCNS algorithm is the highest all the time. As for the Epidemic algorithm and the Spray and Wait algorithm, they are all typical flooding methods in opportunistic social networks and a large number of message group copies cause a low transmission efficiency in the network, so the delivery ratios of the two algorithms are relatively low. Moreover, the EIMST and ICMT algorithms perform the transmission operation of messages through the cooperation of multiple nodes, but this is not an efficient transmission scheme when the cache space of nodes is limited. Consequently, it can be concluded that the delivery ratios of the EIMST and ICMT algorithms are inferior to the FCNS algorithm.

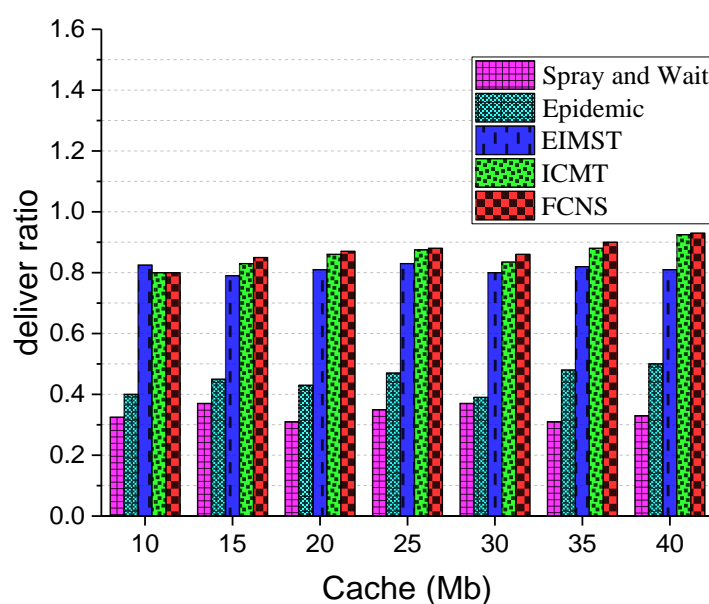


Figure 9. Delivery ratio with various cache spaces.

Next, Figure 10 exhibits the comparison results between the five different algorithms in terms of average end-to-end delay. As shown in the figure, the average end-to-end delay of each algorithm is rising with the increase of the cache space of nodes. Clearly, when the cache space of the node rises from 10 Mb to 40 Mb, the average end-to-end delay of the FCNS algorithm remains relatively stable and is always the lowest among these algorithms. This is due to the fact that the FCNS algorithm makes message delivery decisions by determining and comparing the similarity degree of the node, and the transmission preference plays a decisive role in the data transmission process. Additionally, an increase in the cache space does not have a significant effect on the mobile information and social attributes of nodes—thus, the delay of routing selections, of relay nodes waiting for messages, and of message forwarding are only a slight change. Comparatively speaking, the Epidemic algorithm manufactures a large number of message group copies in the network, which causes a sharp increase in the routing and forwarding delays of the message. The Spray and Wait algorithm effectively limits the number of message copies, so the average end-to-end delay of the algorithm is lower than the Epidemic one. Besides, the EIMST algorithm implements community division and information management, while the ICMT algorithm adopts a cooperative mechanism to realize the reasonable utilization of cache spaces of the node. Therefore, the average end-to-end delays in the EIMST and ICMT algorithms are obviously lower than the traditional algorithms. To sum up, compared with other algorithms, the FCNS algorithm is the optimal method to improve average end-to-end delays in the network.

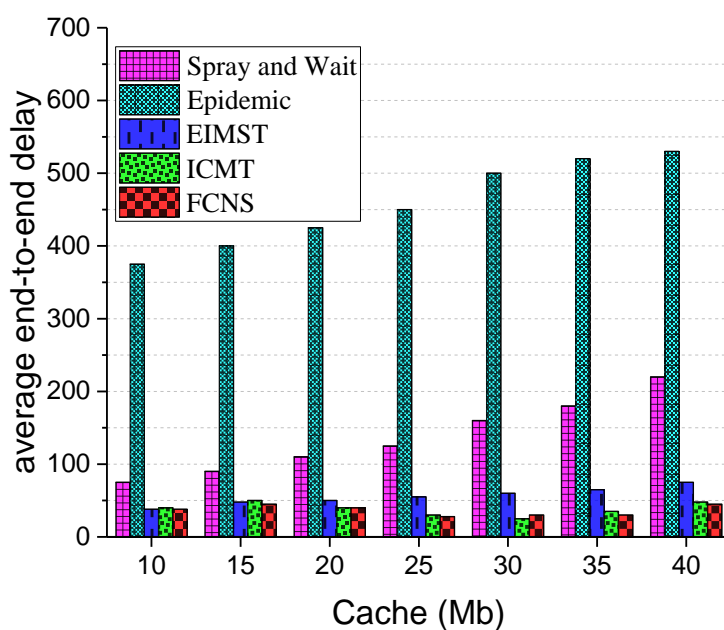


Figure 10. Average end-to-end delay with various cache spaces.

Eventually, the comparison results between the five different algorithms in terms of overhead on average are demonstrated in Figure 11. Overall, when the cache space of the node increases from 10 to 40 Mb, nodes in the network may possess sufficient resource spaces to store statistics and to process the computing tasks, meaning the overhead on average of each algorithm is continuously and dramatically decreasing. Additionally, the overhead on average in the FCNS algorithm always keeps the lowest, because it adopts a comprehensive node similarity to implement the optimal process for data transmission in opportunistic social networks. In the FCNS algorithm, mobile and social similarities mainly reflect the special social relations between nodes, and there is a high probability of successful message transmission between two closely related nodes. Therefore, in the FCNS algorithm, the total number of hops for a complete communication between the source node and the destination is relatively less. The routing and forwarding of messages take only a small amount of time and resources

in the network, and the overhead on average has been significantly improved in the data transmission process. In the Spray and Wait algorithm and the Epidemic algorithm, redundant message group copies need to consume masses of time and resources, meaning the overhead on average is obviously higher than the other algorithms. As for the EIMST and ICMT algorithms, the efficient management for the information and cache space of nodes conduces to allocate resources properly and to control transmission time, so the overhead on average of the two algorithms is at an intermediate level all the time. In conclusion, in terms of the overhead on average, the FCNS is always the best performing one among the five algorithms.

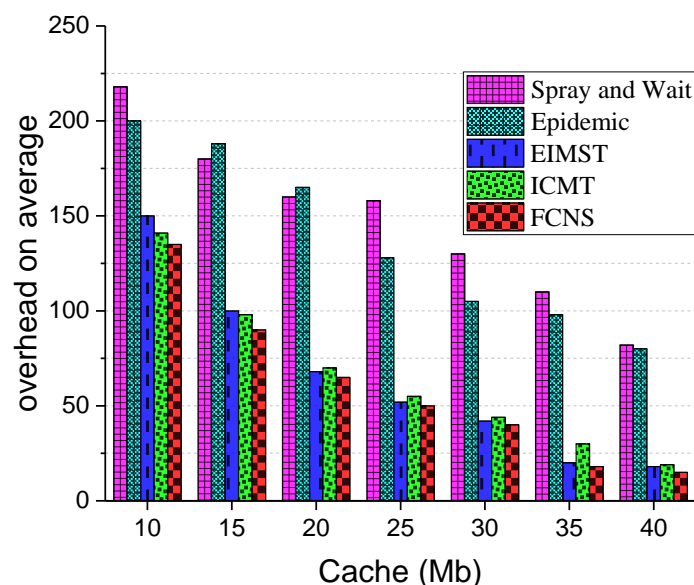


Figure 11. Overhead on average with various cache spaces.

## 5. Discussion

In this study, the length of warming-up time is a significant element that affects the performance of the FCNS algorithm in opportunistic social networks. In real life, during a day, the user is more likely to repeat the same mobile route, such as from places of residence to workplaces, and there is a high probability of them meeting the same person in the day. Therefore, the activity cycle of the user is usually one day, and the optimal value of the warming-up time in the real social scene should be set to 24 h.

The second issue to be discussed is the weight setting of social attributes of nodes. In the FCNS algorithm, the information entropy is used to determine the weight of social similarities of each node in the network, and this method also effectively considers the historical experience of the message carrier. Therefore, the FCNS algorithm can reasonably and accurately evaluate the similarity degrees between nodes in opportunistic social networks.

At last, in opportunistic social networks, although many routing-forwarding algorithms are based on the node similarity, there is no one focusing on mobile similarity. The higher the mobile similarity between a pair of nodes, the higher the probability of them meeting in the future will be. Consequently, the FCNS algorithm creatively pays attention to the mobile information of nodes in the network, and evaluates the comprehensive impact of the mobile and social similarities on data transmission through the fuzzy inference system. Meanwhile, the experimental results show the feasibility and improvement of this algorithm in the transmission environment.



## 6. Conclusions

This work proposed a fuzzy routing-forwarding algorithm (FCNS) exploiting comprehensive node similarity for opportunistic social networks. The algorithm is based on a pair of nodes that will have more opportunities to meet each other if there exists a relatively high node similarity between them. Unlike other routing algorithms based on node similarity, the FCNS synthetically considers the similarity degree between nodes, which is a combination of mobile and social similarities. When selecting a suitable relay node from the surrounding neighbors, this algorithm not only evaluates the social similarities between these neighbors and the destination, but also highlights the importance of the mobile similarity in data transmission. Meanwhile, in order to determine the transmission priority of the node, the Mamdani fuzzy system is employed to calculate the transmission preference value of each node, which is efficiently used in the message forwarding process through the feedback mechanism. In future works, as long as the computing capacity and cache spaces of mobile devices in opportunistic social networks are improved, the FCNS algorithm can be applied to the transmission environment of 5G and big data networks.

**Author Contributions:** K.L., Z.C. and J.W. conceived the idea of the paper. K.L., Z.C., J.W. and L.W. designed and performed the experiments; K.L. and L.W. analyzed the data; Z.C. contributed reagents/materials/analysis tools; K.L. wrote and revised the paper.

**Funding:** This research was funded by [The Major Program of National Natural Science Foundation of China] grant number [No. 71633006]; [The National Natural Science Foundation of China] grant number [No. 616725407]; [China Postdoctoral Science Foundation funded project] grant number [2017M612586]; [The Postdoctoral Science Foundation of Central South University] grant number [185684]; [The Fundamental Research Funds for the Central Universities of Central South University] grant number [No. 2018zzts615].

**Acknowledgments:** This work was supported partially by “Mobile Health” Ministry of Education—China Mobile Joint Laboratory.

**Conflicts of Interest:** The authors declare that they have no competing interests.

## Abbreviations

OPPNET	opportunistic network
DTNs	delay-tolerant networks
BNN	backpropagation neural network model
FCNS	fuzzy routing-forwarding algorithm
MANETs	mobile ad hoc networks
ICNs	intermittently connected networks
HCMM	Health Capability Maturity Model
ONE	Opportunistic Networking Environment
TRSS	Secure Routing Based on Social Similarity
PaSS	an adaptive routing algorithm considering position and social similarities
GD-CAR	a genetic algorithm based dynamic context aware routing protocol
EIMST	Effective information transmission based on socialization nodes
ICMT	Information cache management and data transmission algorithm

## References

1. Trifunovic, S.; Kouyoumdjieva, S.T.; Distl, B.; Pajevic, L.; Karlsson, G.; Plattner, B. A Decade of Research in Opportunistic Networks: Challenges, Relevance, and Future Directions. *IEEE Commun. Mag.* **2017**, *55*, 168–173. [[CrossRef](#)]
2. Jia, W.U.; Chen, Z.; Zhao, M. Effective information transmission based on socialization nodes in opportunistic networks. *Comput. Netw.* **2017**, *129*, 297–305.
3. Bocquillon, R.; Jouglet, A. Robust routing in deterministic delay-tolerant networks. *Comput. Oper. Res.* **2018**, *92*, 77–86. [[CrossRef](#)]
4. Wang, Y.; Liu, Y.; Zhang, J.; Ye, H.; Tan, Z. Cooperative Store-Carry-Forward Scheme for Intermittently Connected Vehicular Networks. *IEEE Trans. Veh. Technol.* **2017**, *66*, 777–784. [[CrossRef](#)]

5. Pirozmand, P.; Wu, G.; Jedari, B.; Xia, F. Human mobility in opportunistic networks: Characteristics, models and prediction methods. *J. Netw. Comput. Appl.* **2014**, *42*, 45–58. [\[CrossRef\]](#)
6. Chen, Z.; Guo, L.; Zhang, D.; Chen, X. Energy and channel transmission management algorithm for resource harvesting body area networks. *Int. J. Distrib. Sens. Netw.* **2018**, *14*. [\[CrossRef\]](#)
7. Ahmad, A.; Doss, R.; Alajeely, M.; Al Rubeaai, S.F.; Ahmad, D. Packet integrity defense mechanism in OppNets. *Comput. Secur.* **2018**, *74*, 71–93. [\[CrossRef\]](#)
8. Wang, L.; Chen, Z.; Wu, J. An Opportunistic Routing for Data Forwarding Based on Vehicle Mobility Association in Vehicular Ad Hoc Networks. *Information* **2017**, *8*, 140. [\[CrossRef\]](#)
9. Kumar, P.; Chauhan, N.; Chand, N. Security Framework for Opportunistic Networks. In *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications*; Springer: Singapore, 2018.
10. Lenando, H.; Alrfaay, M. EpSoc: Social-Based Epidemic-Based Routing Protocol in Opportunistic Mobile Social Network. *Mob. Inf. Syst.* **2018**, *2018*, 6462826. [\[CrossRef\]](#)
11. Wu, J.; Chen, Z. Sensor communication area and node extend routing algorithm in opportunistic networks. *Peer-to-Peer Netw. Appl.* **2016**, *11*, 90–100. [\[CrossRef\]](#)
12. Jang, K.; Lee, J.; Kim, S.K.; Yoon, J.H.; Yang, S.B. An adaptive routing algorithm considering position and social similarities in an opportunistic network. *Wirel. Netw.* **2016**, *22*, 1537–1551. [\[CrossRef\]](#)
13. Wu, J.; Chen, Z.; Zhao, M. Information cache management and data transmission algorithm in opportunistic social networks. *Wirel. Netw.* **2018**, *8*, 1–12. [\[CrossRef\]](#)
14. Wu, J.; Chen, Z. Human Activity Optimal Cooperation Objects Selection Routing Scheme in Opportunistic Networks Communication. *Wirel. Pers. Commun.* **2017**, *95*, 3357–3375. [\[CrossRef\]](#)
15. Mayer, J.M.; Hiltz, S.R.; Barkhuus, L.; Väänänen, K.; Jones, Q. Supporting Opportunities for Context-Aware Social Matching: An Experience Sampling Study. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 2430–2441.
16. Nguyen, H.A.; Giordano, S. Context information prediction for social-based routing in opportunistic networks. *Ad Hoc Netw.* **2012**, *10*, 1557–1569. [\[CrossRef\]](#)
17. Zulfikar, W.B.; Prasetyo, P.K.; Ramdhani, M.A. Implementation of Mamdani Fuzzy Method in Employee Promotion System. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Guangzhou, China, 8–12 October 2018; Volume 288, p. 012147.
18. Nguyen, H.A.; Giordano, S.; Puiatti, A. Probabilistic Routing Protocol for Intermittently Connected Mobile Ad hoc Network (PROPICMAN). In Proceedings of the 2007 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), Espoo, Finland, 18–21 June 2007; pp. 1–6.
19. Sharma, D.K.; Dhurandher, S.K.; Woungang, I.; Bansal, A.; Gupta, A. *GD-CAR: A Genetic Algorithm Based Dynamic Context Aware Routing Protocol for Opportunistic Networks*; Springer: Cham, Switzerland, 2017.
20. Talipov, E.; Chon, Y.; Cha, H. User context-based data delivery in opportunistic smartphone networks. *Pervasive Mob. Comput.* **2015**, *17*, 122–138. [\[CrossRef\]](#)
21. Borah, S.J.; Dhurandher, S.K.; Woungang, I.; Kumar, V. A game theoretic context-based routing protocol for opportunistic networks in an IoT scenario. *Comput. Netw.* **2017**, *129*, 572–584. [\[CrossRef\]](#)
22. Nguyen, P.; Nahrstedt, K. Crowdsensing in Opportunistic Mobile Social Networks: A Context-aware and Human-centric Approach. *arXiv*, **2017**, arXiv:1704.08598.
23. Li, D.; Ma, L.; Yu, Q.; Zhao, Y.; Mao, Y. An community detection algorithm based on the multi-attribute similarity. In Proceedings of the IEEE International Conference on Communications, Circuits and Systems, Chengdu, China, 15–17 November 2013; pp. 118–121.
24. Liu, Y.; Jing, W.; Xu, L. Parallelizing Backpropagation Neural Network Using MapReduce and Cascading Model. *Comput. Intell. Neurosci.* **2016**, *2016*, 2842780. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Qin, X.; Wang, X.; Lin, Y.; Wang, L.; Zhang, L. An Efficient Routing Algorithm Based on Interest Similarity and Trust Relationship between Users in Opportunistic Networks. In *China Conference on Wireless Sensor Networks*; Springer: Singapore, 2018.
26. Allen, S.M.; Chorley, M.J.; Colombo, G.B.; Jaho, E.; Karaliopoulos, M.; Stavrakakis, I.; Whitaker, R.M. Exploiting user interest similarity and social links for micro-blog forwarding in mobile opportunistic networks. *Pervasive Mob. Comput.* **2014**, *11*, 106–131. [\[CrossRef\]](#)
27. Yao, L.; Man, Y.; Huang, Z.; Deng, J.; Wang, X. Secure Routing Based on Social Similarity in Opportunistic Networks. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 594–605. [\[CrossRef\]](#)

28. Coll-Perales, B.; Gozálvez, J.; Friderikos, V. Context-Aware Opportunistic Networking in Multi-Hop Cellular Networks. *Ad Hoc Netw.* **2016**, *37*, 418–434. [[CrossRef](#)]
29. Li, A.; Zhao, Z. An Improved Model of Variable Fuzzy Sets with Normal Membership Function for Crane Safety Evaluation. *Math. Probl. Eng.* **2017**, *2017*, 3190631. [[CrossRef](#)]
30. Guan, J.; Chu, Q.; You, I. The Social Relationship Based Adaptive Multi-Spray-and-Wait Routing Algorithm for Disruption Tolerant Network. *Mob. Inf. Syst.* **2017**, *2017*, 1819495. [[CrossRef](#)]
31. Ma, J.; Ma, Z. Epidemic threshold conditions for seasonally forced seir models. *Math. Biosci. Eng. MBE* **2017**, *3*, 161–172. [[CrossRef](#)]
32. Pak, J.; Song, Y.T. Health Capability Maturity Model: Person-centered approach in Personal Health Record System. In Proceedings of the Americas Conference on Information Systems, San Diego, CA, USA, 11–13 August 2016.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).