Research Article

Effective Data Optimization and Evaluation Based on Social Communication with AI-Assisted in Opportunistic Social Networks

Limiao Li,1 Fangfang Gou,2 Huiyun Long1,3 Keke He,1 and Jia Wu1,2

1School of Computer Engineering and Applied Mathematics, Changsha University, Changsha 410003, China
2School of Computer Science and Engineering, Central South University, Changsha 410083, China
3State Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University, Guiyang 550025, China

Correspondence should be addressed to Huiyun Long; hylong@gzu.edu.cn and Jia Wu; jiawa5110@163.com

Received 22 February 2022; Revised 9 May 2022; Accepted 14 June 2022; Published 1 July 2022

Academic Editor: Muhammad Ismail

Copyright © 2022 Limiao Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Billions of people around the world send and receive data over online networks daily. Sufficient and redundant data are transmitted over social platforms with AI-assisted in 5G networks. In opportunistic social networks, the main challenge faced by traditional methods is that numerous user nodes participate in data transmission, causing a lot of message copy redundancy and node cache consumption. As a result, the transmission delay of the algorithm is high, the node energy consumption is too large, and even information is lost. To solve these problems, this study establishes an artificial intelligence-based optimization multiple evaluation method. The main purpose of this method is to avoid information loss caused by data loss when reducing data noise, reasonably select communication nodes in opportunistic social network scenarios, optimize data transmission performance, and avoid network congestion. Moreover, our method can effectively identify and exclude potential malicious nodes, reducing the situation that packets are intercepted and discarded. The experiment confirms that the optimized transmission evaluation scheme can effectively reduce routing overheads and energy consumption of a user node, improve the delivery ratio of node data transmission, and ensure the reliability and security of data transmission.

1. Introduction

With the development of social media, online social networks have become an essential platform for information dissemination. For the support of 5G networks with high reliability and low latency, billions of users share information on social networking sites such as Facebook, Twitter, and Google+ every day [1]. The 5G network has created abundant conditions for the rapid development of massive data on the Internet [2]. People exchange ideas, share photo albums, upload videos, and live stream videos in real time through these platforms anytime and anywhere [3, 4]. All of this information will be converted into data that can be stored, sent, and received by users [5]. With people’s activities, everyone could become a source of data center and social communication.

When the amount of data is large, the communication link of the traditional social network method is easily interrupted, and a stable end-to-end path cannot be maintained, which significantly affects the success rate of information delivery [6]. Opportunistic social networks are intermittently connected mobile ad hoc networks, which usually implement communication between nodes in a “storage-carry-forward” routing mode, which can provide strong support for the effective transmission of data [7, 8]. This transmission mechanism relies on node movement to achieve, and human mobility can be used to carry messages between separate parts of the network physically. Therefore, human mobility plays a vital role in the performance of forwarding protocols in the network. In addition, 5G realizes the interconnection of all things and make people in the network of intelligent interaction [9]. In modern times,
people’s communication is carried out in such a social network environment, and each activity is affected by the social interaction with each other in the network [10].

In the application scenario of opportunistic social networking, when people transfer data through portable mobile devices, considering that there is no suitable transmission target within a certain transmission range, or the transmission target does not respond, it will occupy a large amount of buffer in the data storage area of the device, which affected the data transmission rate [11]. On the one hand, the information has not been received and responded for a long time, and it is always in a waiting state, which causes a large amount of reserved information to be stored in the device, resulting in transmission delay [12]. On the other hand, when there is a large amount of buffer space when new urgent data information is received, the information release sequence needs to be updated to achieve real-time release [13]. However, when a large amount of storage space is occupied, the information cannot be updated in time.

However, in algorithms commonly used to solve the problem of message transmission in opportunistic social networks, the transmission and screening methods of messages have been difficult to adapt to the complex big data environment and the variability of data-carrying mobile devices [14]. For flood-based methods or social perception methods, they may cause a massive amount of information stored in the network and may not be effectively dealt with in time [15]. In the flood-based routing algorithm, each node forwards the received data packets to all nodes that can meet by broadcasting. This can cause copies of the message to flood the network, blocking data transmission [16]. For the social awareness method, messages wait infinitely for transmission in the process of filtering nodes, occupying limited buffer space, and mobile nodes cannot adjust the storage and transmission of messages according to environmental pressure [17]. In addition, a large number of redundant replicas in the network will lead to increased node energy consumption, increased data overhead, and buffer overflow [18]. This in turn leads to problems such as network congestion and packet loss. We can conclude that facing numerous communication nodes in a big data environment will bring significant pressure to the entire network, thus resulting in a decrease in transmission success rate, transmission delay, and data packets. It may even cause information leakage and security problems due to the existence of pseudo nodes.

In fact, the era of big data is also an era of “big noise.” The information transmitted by people is increasing exponentially, but the useful information is minimal, causing information “noise” to increase much faster than useful information [19]. The pending information in the process of social network message transmission is noise. Although it may become an information resource, the preservation of its potential state also requires resources [20]. In other words, the value of data is contingent, but the loss of data processing is inevitable. In addition, data loss may result in the loss of important messages, creating risks such as information leakage [21]. Therefore, we must establish appropriate models and algorithms to solve a series of problems such as low transmission rate and increased data noise caused by massive data communication.

In this study, a multievaluation method for noise reduction optimization is established with AI-assisted. The nodes must be transmitted under the opportunistic network scenario in 5G, and the transmission must be optimized. This process is convenient for finding the law of research. The design of an effective data optimization and evaluation algorithm (EDOE) effectively divides the information. The noise optimization model ensures reliable transmission of valid data and solves the problem of important information loss when data is lost. Simultaneously, the task of information transmission is conducted using a community node, which effectively reduces the overhead of the user node and improves the performance of the algorithm. In addition, we provide privacy protection for nodes in the form of public key encryption, and exclude malicious nodes through signatures, thereby resisting routing attacks and ensuring data transmission security. The EDOE method proposed in this paper effectively avoids network congestion and packet loss and improves the overall operational efficiency of the network.

The contributions of this study are presented as follows:

(1) A function optimization model in a noisy environment is established using social user nodes to measure user noise with AI-assisted. And different noise reduction strategies are determined according to the data transmission situation, which effectively prevents the loss of important information in the data.

(2) Various effective methods for optimizing the measured noise are established. The use of effective evaluation methods adjusts the adaptability between user nodes, effectively improves the communication ability between users, reduces the influence of noise, and improves the reliability of data transmission.

(3) The experiment shows that the optimized transmission evaluation scheme can effectively reduce the routing overhead and energy consumption of the user nodes and improve the delivery ratio of node data transmission.

2. Related Work

In recent years, in opportunistic social networks and how to use the least resources, the least delay to transmit the sent messages to their desired destinations has been the core of research. This section mainly briefly introduces some research work related to data transmission optimization in the current opportunity social network, from which we can learn about different ways to deal with transmission problems for different application scenarios.

In literature [22], Xiao et al. proposed a distributed optimal community-aware opportunistic routing-CAOR algorithm, which uses a family-aware community model. This algorithm, based on community information, transforms the routing problem between mobile nodes into a routing
problem between static communities to reduce information communication costs and maintenance costs.

Zhou et al. [23] proposed an effective opportunistic mobile network data forwarding strategy based on time closeness and centrality, called TCCB (temporal closeness and centrality-based). The key innovation is to capture data within the valid time and use temporal correlation to infer possible future social contact. This strategy aims to improve the data transmission rate when the transmission cost is similar. In literature [24], Cheol Jeong et al. introduced a hierarchical collaboration (HC) protocol of network decomposition to maximize the throughput-latency trade-off. In a dense network, compared with the case of a network without social relations, as the density of social groups increases, the performance of throughput delay trade-off can be significantly improved.

Abderrahmen Mtibaa [25] and others have proposed a set of social-based trust filters that can use explicit social information and real people movement trajectories to establish reliable communication between relay nodes. The real-time tracking-driven method is used to achieve a fair trade-off between trust and a successful delivery rate. That way, the trusted filter can achieve trusted communication in the opportunity network based on a successful delivery rate of more than 35%. Kr. Sharma et al. [26] proposed a routing protocol-Ant Router based on ant colony optimization algorithm, which stores the relevant information in the buffer of each node and then uses the “ant” algorithm to find a reasonable next hop of the data packet. It reaches the final destination through the shortest channel from the source. The protocol overcomes the problems of frequent disconnection between nodes and low routing efficiency and ensures effective and reliable message delivery.

In literature [27], Xia et al. proposed an interest-based socially aware networking message forwarding scheme BEEINFO (BEE-colony-inspired INterest-based Forwarding). It consists of BEEINFO-D (community density), BEEINFO-S (social tie), and BEEINFO-D&S. Classify communities into specific categories based on interests, introduce message scheduling strategies and buffer management algorithms, and use swarm intelligence to enhance adaptability to dynamic environments and improve forwarding performance. In literature [28], Lin Yao et al. proposed to incorporate social trust into the routing decision-making process and design trust routing based on the social similarity (TRSS) scheme based on the common interests of nodes or social similarity movements. Based on direct and recommended trust, untrusted nodes are removed from the next jump sender, and only data packets from trusted nodes are forwarded.

Recently, much AI-related wireless communication work has been discussed by researchers. To realize the rapid response of node communication and meet the requirements of the bandwidth and response time of network transmission for a large amount of data and applications, Pratik Goswami et al. [29] proposed a method for implementing distributed artificial intelligence using neural networks. The method realizes node prediction by considering the number of observations of node power consumption when executing a specific application and finally effectively saves the response time of routing. Amrit Mukherjee [30] et al. established a method for high-speed communication of large capacity data in wireless multimedia sensor networks. The main idea is to distribute power through distributed artificial intelligence and analyze the real-time spectrum sensing output system. This method not only achieves fast communication between nodes, but also guarantees cost benefits.

In opportunistic social networks, many user nodes participate in data transmission, leading to message replication redundancy, information noise growth, and node cache consumption. Previous research mostly relies on user nodes' social attributes to find a suitable next hop to improve transmission efficiency, without considering the transmission delay and information loss caused by user information noise. To solve these problems, we have established a multiple evaluation method for noise reduction optimization in opportunistic social networks, reasonably communicating nodes in an opportunistic social network scenario, optimizing transmission, and increasing users’ transmission rate.

3. Data Decision and Transmission Model Design

The community, between the media and social [31–34], satisfies the connection between a person and a group. From the BBS when the Internet started to the current social networks (Facebook, WeChat Moments, etc.) [35–37], social media (Instagram, Twitter, etc.), and some vertical communities (Tiger, Blued, etc.), the community has always existed around the users, and the users are related to each other through various communities [38, 39], as shown in Figure 1. Social networks take people as the core and communicate with each other in the 5G network through smart terminals and wireless networks. It is obvious from Figure 1 that in the same community, people usually have similar hobbies and social characteristics, so people are connected more frequently. According to this property of opportunistic social networks, we can apply the identified communities to various forms of social applications such as collaborative recommendation, information dissemination, and knowledge sharing [40–42].

Now, the continuous development of the Internet and mobile devices has led to more and more people joining online social networks. Unlike social interaction, the online community we propose reduces physical distance and time limitations and has more vital interaction and timeliness. When users use social networks, they show different social attributes, such as their interests and hobbies, thereby forming multiple communities. Users with more of the same social attributes are usually divided into the same online community and become nodes in the community, which relies on the analysis of the relationship between people [43, 44]. The nodes in the community can communicate with each other, which makes communication more convenient and faster.

With the increase of social network users, data transmission between users will retain a large amount of information, and data noise will also increase exponentially, which interferes with transmission quality [45–47]. Certain information
may be in a state of no user acceptance and response for a long time. When new data information is received, and urgent information needs to be released in real-time, the information transmission sequence cannot be updated due to a large amount of cache space, resulting in data delay. Reducing the noise between the user data and improving the transmission success rate is the key [48–50]. To solve these problems, we built an optimization function to simulate the noise problem in the social network environment.

Social networks are constantly evolving with time. Many interference factors will affect the evaluation of node relationships. In the opportunistic social network, the process of data transmission is affected by normal random noise. Reducing the impact of noise in the network helps to evaluate node relationships better. Therefore, we must establish a function optimization model in a noisy environment. Model sets:

\[ f_\sigma(X) = f(X) + k \times \delta_x, \delta_x \sim N(0, \sigma^2), \]

where \( f_\sigma(X) \) is the function value after noise interference and is the original multimodal function value, \( X \) is a random variable, \( \delta_x \) is the Gaussi \( f(X) \) an noise with a mean value of 0 and variance of \( \sigma^2 \), and \( k \) is the noise intensity.

For any two users \( u_1 \) and \( u_2 \), we can set two functions; \( X_1 \) and \( X_2 \) are the two solutions of the function \( f(X_1) > f(X_2) \). When noise interference occurs, \( f_\sigma(X_1) = f(X_1) + k \times \delta_1 \), and its probabilities are \( P(f_\sigma(X_1)) \) and \( P(f_\sigma(X_2)) \).

\( X_1 > X_2 \) is better than \( X_1 \) for \( X_2 \). When no noise interference occurs, the probability of \( X_1 > X_2 \) is \( p(X_1 > X_2) = 1. \)

When noise interference occurs and if \( \delta_1 \) and \( \delta_2 \sim N(0, \sigma^2) \) are independent of each other and \( (\delta_1 - \delta_2) \sim N(0, 2\sigma^2) \), the probability of \( X_1 > X_2 \) is:

\[ p_\sigma(X_1 > X_2) = P[f_\sigma(X_1) > f_\sigma(X_2)] = \left[ P(\delta_1 - \delta_2 > \frac{f(X_1) - f(X_2)}{k}) \right]. \]

(2)

If \( X_1 \) and \( X_2 \) are deterministically competitive in the opportunistic social network, then no noise interference occurs when \( p(X_1 > X_2) = 1. \) When noise interference occurs, the probability of \( X_1 \) remaining in the next user group is less than 1. Noise affects the selection operation of the relay node of the network change algorithm and reduces the search performance of the algorithm.

A common problem in AI-assisted social networks is that users are subject to noise interference, which reduces transmission quality. The fitness evaluation is divided into two processes, namely, objective function sampling and fitness calculation, to analyze the influence of normal random noise on fitness evaluation. Obtaining the value of \( f_\sigma(X) \) is a sampling process AI-assisted, and converting one or more sample values into node fitness is a fitness calculation process.

3.1. One Evaluation and One Sample (OO). In opportunistic social network, the traditional network change algorithm’s fitness evaluation uses a one-time evaluation sampling method. This method samples a node’s objective function and evaluates the fitness once. The user’s large population
size AI-assisted is the method of fitness evaluation, as expressed in

\[ f_{\sigma}^{\text{OO}}(X) = f(X) + k \times \delta, \delta \sim N(0, \sigma^2), \]  

(3)

where \( f_{\sigma}^{\text{OO}}(X) \) is the fitness of \( X \) in a noisy environment, \( f(X) \) is the fitness of \( X \) in the environment, and \( \delta \) is the sampling noise when evaluating nodes that obey normal distribution with a mean of 0 and a variance of \( \sigma^2 \).

3.2. One Evaluation and Multisamples (OM). Resampling user data is a method for evaluating multiple sampling times. This method evaluates the fitness of each node only once, and the evaluation process samples the objective function AI-assisted multiple times, as defined in

\[ f_{\sigma}^{\text{OM}}(X) = \frac{1}{N} \sum_{i=1}^{N} (f(X) + k \times \delta_i), \delta_i \sim N(0, \sigma^2), \]  

(4)

where \( f_{\sigma}^{\text{OM}}(X) \) is the fitness of \( X \) in a noisy environment; \( f(X) \) is the fitness of \( X \) in the environment; \( \delta_i \) is the \( i \)-th sampling noise; its mean is 0 and variance is \( \sigma^2 \); and \( N \) is the number of samples. One evaluation multiple sampling method reduces the standard deviation of noise to \( \sigma/\sqrt{N} \).

3.3. Evaluation of every Generation and One Sample Test (EO). For the opportunistic social network, multiple evaluations of the sampling method reevaluate the fitness of nodes in each generation of social networks, and each objective evaluates the objective function once. A node only evaluates once. When a low fitness node is misjudged as a high fitness node, the node will be widespread in the user group, and the search time of the misleading algorithm will be extended. When two nodes in the network compete for survival opportunities at the same time, the sampling method is evaluated multiple times to reevaluate the adaptability of the selectable nodes in a noisy environment. The probability of the inferior node being evaluated as excellent is small, thus eliminating this node from the AI-assisted user group. The sampling method is expressed in

\[ f_{\sigma}^{\text{EO}}(X, t) = f(X) + k \times \delta(t), \delta(t) \sim N(0, \sigma^2), \]  

(5)

where \( f_{\sigma}^{\text{EO}}(X, t) \) is the fitness of \( X \) at time \( t \) in a noisy environment, \( f(X) \) is the fitness of \( X \) in the environment, and \( \delta(t) \) is the sampling noise at time \( t \) and obeys normal distribution with a mean of 0 and a variance of \( \sigma^2 \).

Assuming no noise interference, \( f(X_1) > f(X_2) \), and let \( f(X_1) - f(X_2) = \Delta > 0 \). When using the EO evaluation method, the probability of \( X_1 > X_2 \) in a noisy environment is:

\[ P_{\sigma}^{\text{EO}}[(X_1 > X_2)|f(X_1) > f(X_2)] = P[f_{\sigma}^{\text{EO}}(X_1) > f_{\sigma}^{\text{EO}}(X_2)] = P[(\delta_1^* - \delta_2^*) < (-\Delta/k)]. \]  

(6)

When using the OO evaluation method AI-assisted, the probability of \( X_2 > X_1 \) in a noisy environment is:

\[ P_{\sigma}^{\text{EO}}[(X_2 > X_1)|f(X_1) > f(X_2)] = P\left[f_{\sigma}^{\text{EO}}(X_2) > f_{\sigma}^{\text{EO}}(X_1)\right] = P[(\delta_1^* - \delta_2^*) < (-\Delta/k)]. \]  

(7)

Therefore, when using the OO evaluation method AI-assisted in a noisy environment, the probability difference between the correct and the incorrect judgments is:

\[ P_{\sigma}^{\text{EO}}[(X_1 > X_2)|f(X_1) > f(X_2)] - P_{\sigma}^{\text{EO}}[(X_2 > X_1)|f(X_1) > f(X_2)] = P[\Delta/k > (\delta_1^* - \delta_2^*) > -\Delta/k]. \]  

(8)

Similarly, when using the OM evaluation method in the noise environment, the probability difference between the correct and the incorrect judgments is:

\[ P_{\sigma}^{\text{OM}}[(X_1 > X_2)|f(X_1) > f(X_2)] - P_{\sigma}^{\text{OM}}[(X_2 > X_1)|f(X_1) > f(X_2)] = P[k/\Delta > (\delta_1^* - \delta_2^*) > -k/\Delta]. \]  

(9)

Let the \( t \)-th generation be \( P(t) \). Moreover, let \( X_1^* \) and \( X_2^* \) be the optimal solution, and \( |f(X_1^*) - f(X_2^*)| = \epsilon > 0 \).

When using the EO evaluation method to evaluate the competition of nodes in two generations of social networks, the nodes in the previous generation network must be reevaluated. Therefore, \( f_{\sigma}^{\text{EO}}[X_1^*(t+1)] = f(X_1^*) + k \times \delta_1^*(t+1) \), and \( f_{\sigma}^{\text{EO}}[X_2^*(t+1)] = f(X_2^*) + k \times \delta_2^*(t+1) \). \( \delta_1^*(t+1) \) and \( \delta_2^*(t+1) \) are the sampling noises of the \( t+1 \) generation. Therefore, \( \delta_2^*(t+1) - \delta_1^*(t+1) \sim N(0, 2\sigma^2) \).

When using the EO evaluation method, the probability that the optimal solution is correctly updated (i.e., \( f(X_1^*) < f(X_2^*) \)) is:

\[ P_{\sigma}^{\text{EO}}[(X_2^* > X_1^*)|f(X_1^*) < f(X_2^*)] = P[\delta_2^*(t+1) - \delta_1^*(t+1) > \epsilon/k]. \]  

(10)

When using the EO evaluation method, the probability of the optimal solution error update (i.e., \( f(X_1^*) > f(X_2^*) \)) is:

\[ P_{\sigma}^{\text{EO}}[(X_1^* > X_2^*)|f(X_1^*) > f(X_2^*)] = P[\delta_1^*(t+1) - \epsilon/k]. \]  

(11)

Therefore, when using the EO evaluation method, the probability difference between a correct and an incorrect update of the optimal solution is:

\[ P_{\sigma}^{\text{EO}}[(X_2^* > X_1^*)|f(X_1^*) < f(X_2^*)] - P_{\sigma}^{\text{EO}}[(X_1^* > X_2^*)|f(X_1^*) > f(X_2^*)] = P[\epsilon/k > (\delta_2^*(t+1) - \delta_1^*(t+1)) > -\epsilon/k]. \]  

(12)
The OM evaluation method does not need to reevaluate the nodes in the previous generation of social networks. Thus, $f_{OM}(X^*_1) = f(X^*_1) + k \times \delta^*_1$, where $\delta^*_1$ is the sampling noise of $X^*_1$ and $\delta^*_1$ is a relatively large value that already exists given the role of the selection operator. Furthermore, $f_{OM}[X^*_2(t + 1)] = f(X^*_2) + k \times \delta^*_2(t + 1)$ is the sampling noise in the $t + 1$-th generation. Therefore, $\delta^*_2(t + 1) \sim N(0, \sigma^2)$. When using the OM evaluation method, the probability that the optimal solution is correctly updated (i.e., $f(X^*_1) < f(X^*_2)$) is:

$$P_{OM}[(X^*_1 > X^*_1)] = P[\delta^*_2(t + 1) > \delta^*_1 - \epsilon/k].$$

(13)

When using the OM evaluation method, the probability of the optimal solution error update (i.e., $f(X^*_1) > f(X^*_2)$) is:

$$P_{OM}[(X^*_1 > X^*_1)] = P[\delta^*_2(t + 1) > \delta^*_1 + \epsilon/k].$$

(14)

Therefore, when using the OM evaluation method, the probability difference between a correct and an incorrect update of the optimal solution is:

$$P_{OM}[(X^*_1 > X^*_1)] - P_{OM}[(X^*_1 > X^*_1)] = P[\delta^*_1 + \epsilon/k > \delta^*_2(t + 1) > \delta^*_1 - \epsilon/k].$$

(15)

When using the OM evaluation method, the probability difference between a correct and an incorrect update of the optimal solution is $[\epsilon_L, \epsilon_R]$. $\epsilon_R = \delta^*_1 + \epsilon/k$, and $\epsilon_L = \delta^*_1 - \epsilon/k$. When $\delta^*_1 = \delta_{eq}$, $\epsilon/k$ is very small.

$$P[\delta^*_1 + \epsilon/k > \delta^*_2(t + 1) > \delta^*_1 - \epsilon/k] = P(\epsilon/k > \delta^*_2(t + 1) - \delta^*_1(t + 1) > \delta^*_1 - \epsilon/k).$$

(16)

$$\lim_{t \to \infty} P(\delta^*_1 > \delta_{eq}) = 1.$$

Therefore,

$$P(\epsilon/k > \delta^*_2(t + 1) > \delta^*_1(t + 1) > \delta^*_1 - \epsilon/k).$$

(17)

In opportunistic social networks, the EO method can be used to reduce noise when data transmission must clarify the impact of previous generation historical data information. When the current data transmission does not need to consider the influence of the historical data information of the previous generation, the OM method can effectively improve the normal randomness. By adjusting the adaptability between user nodes, the impact of noise on data transmission is effectively reduced. This study has good noise reduction performance in the noisy environment of opportunistic social network. When dealing with data noise, it solves the problem of important message loss caused by data loss and ensures the security and effectiveness of data transmission.

We can effectively reduce user data transmission noise and improve data transmission efficiency through the evaluation method. We can establish the EDOE algorithm in accordance with the evaluation method.

Effective data optimization and evaluation algorithms are shown in Algorithm 1.

To improve the comprehensibility of our proposed algorithm, as shown in Algorithm 1, we have further explained the algorithm. Since the total time for the node to complete
the task must be less than or equal to the time for the node to exist in the online community, the subtasks will be allocated to the node to power from strong to weak. Then, all the nodes that have been assigned to the task perform the task and return the result to X, and X evaluates a suitable relay node for data transmission based on the evaluation result. The time complexity of this process is \( O(n) \). Besides, the time complexity of spray and wait is \( O(n \log n) \), and the time complexity in the epidemic routing algorithm is \( O(n^2) \). Therefore, our proposed algorithm has low time complexity.

Suppose that the nodes defined in this article meet the following conditions. When using social networks, users will show different social attributes such as their interests and hobbies, and users with more similar social attributes are usually divided into the same online community. Through the “storage-carry-forward” routing model, communication can be achieved between nodes in an opportunity social network. To analyze the nodes in the opportunistic social network, we first need to understand the characteristics of the nodes.

We use a representation called a module network, which divides variables into modules. Each module represents a set of variables with the same statistical behavior. By implementing this constraint on the learned network, we can significantly reduce the model’s complexity and the number of parameters. These reductions allow for more reliable estimates and better generalizations of the data. Thereby, the utilization rate of resources in the network and the overall operation efficiency are guaranteed.

Defined at the time \( p \), we express the modularity of the community as:

\[
B(p) = \frac{h_m + \Delta h}{H + \Delta h} - \frac{(r_n + 2\Delta h)^2}{4(H + \Delta h)^2}. \tag{18}
\]

where \( B \) is the modularity of the community, \( h_m \) indicates the edge total weight in community \( M \), \( r_n \) is the node \( n \) total level, and \( H \) is the total edge weight.

**Condition 1.** If any node belongs to a node in the opportunistic social network, with the increase of edge weight formed with other nodes in the network, the community’s total edge weight \( h_m \) will increase. It will lead to a concomitant increase in the relevance of the community in opportunistic social networks.

**Proof.** At time \( p \), the modularity \( B(p) \) in the community is:

\[
B(p) = \frac{h_m + \Delta h}{H + \Delta h} - \frac{(r_n + 2\Delta h)^2}{4(H + \Delta h)^2}. \tag{19}
\]

Then, the modular change after time \( p + 1 \) can be expressed as:

\[
B(p + 1) - B(p) = \frac{h_m + \Delta h}{H + \Delta h} - \frac{(r_n + 2\Delta h + \Delta \delta h)^2}{4(H + \Delta h + \Delta \delta h)^2} - \frac{h_m - r_n^2}{4H^2} \geq \frac{4H^2 \Delta h - 6H^2 r_n \Delta h + 2H^2 r_n \Delta h - 2H^2 r_n \Delta h + (r_n \Delta h)^2}{4H^2 (H + \Delta h)^2} = \Delta h \frac{(2H^2 - 2Hr_n - r_n \Delta h)(2H - r_n)}{4H^2 (H + \Delta h)^2}. \tag{20}
\]

We can get \( \Delta h > 0 \). So, we only need to prove \( (2H^2 - 2Hr_n - r_n \Delta h)(2H - r_n) < 0 \), and then, we can get \( B(p + 1) - B(p) > 0 \).

In other words,

\[
\begin{cases}
2H^2 - 2Hr_n - r_n \Delta h > 0 \\
2H - r_n > 0 \\
\Delta h > 0
\end{cases}
\Rightarrow
\begin{cases}
0 < \Delta h < 2H \left( \frac{H}{r_n} - 1 \right) \\
2H \left( \frac{H}{r_n} - 1 \right) > 0 \\
r_n < 2H
\end{cases}
\Rightarrow
\begin{cases}
0 < \Delta h < 2H \left( \frac{H}{r_n} - 1 \right) \\
r_n < H
\end{cases}. \tag{21}
\]

**Condition 2.** In the opportunistic social network, if node \( N_a \) satisfies condition \( r_{a\beta}/2H < h_{a\beta} < \Delta h + (r_{a\beta} + r_n \Delta h + \Delta \delta h)/(2(H + \Delta h)) \), it will then be separated from the community of node \( N_\beta \).

**Proof.** Firstly, supposing that community \( M \) is divided into two subcommunities \( M_a \) and \( M_\beta \), where nodes \( N_a \) and \( N_\beta \) belong to different communities. With the increase of the
When the total weight decreases, the formula can be expressed as follows:

\[
H_{a}^* + H_{\beta}^* > H^*,
\]
\[
h_{a\beta} < \Delta h + \frac{r_a r_{\beta} + r_n \Delta h + \Delta h^2}{2(H + \Delta h)}.
\] (23)

Therefore, when the communities \(M_a\) and \(M_\beta\) where the two nodes are located satisfy the condition \(r_a r_{\beta}/2H < h_{a\beta} < \Delta h + (r_a r_{\beta} + r_n \Delta h + \Delta h^2)/2(H + \Delta h)\), the communities are divided.

**Condition 3.** If the unique edge of node \(N_a\) in the opportunistic network is connected to node \(N_\beta\), then, node \(N_a\) exists. When the weight between nodes \(N_a\) and \(N_\beta\) drops, node \(N_\beta\) will still not be separated from the community.

**Proof.** If community \(M\) is divided, then it must meet the following three conditions:

\[
\begin{align*}
H_a + H_\beta &< H \\
& \quad \text{for} \quad \frac{b_a}{H} - \frac{r_a^2}{4H^2} + \frac{b_\beta}{H} - \frac{r_\beta^2}{4H^2} < \frac{r_a + r_\beta + h_{a\beta}}{H} - \frac{(r_a + r_\beta)^2}{4H^2} \\
h_{a\beta} &> \frac{r_a r_\beta}{2H}
\end{align*}
\] (24)

As the weight changes, the formula can be expressed as:

\[
H_{a}^* + H_{\beta}^* > H^*,
\]
\[
h_{a\beta} < \Delta h + \frac{r_a r_{\beta} + r_n \Delta h + \Delta h^2}{2(H + \Delta h)}.
\] (25)

It can be understood as:

\[
\frac{r_a r_{\beta}}{2H} < h_{a\beta} < \frac{r_a (r_\beta + \Delta h)}{2(H + \Delta h)} = \frac{r_a r_{\beta} + r_a \Delta h}{2(H + \Delta h)},
\] (26)

As:

\[
\frac{r_a r_{\beta}}{2H} < h_{a\beta} < \frac{r_a (r_\beta + \Delta h)}{2(H + \Delta h)} = \frac{r_a r_{\beta} + r_a \Delta h}{2(H + \Delta h)} < 0(\Delta h < 0). \] (27)
Among them, the generation and distribution of all private keys is based on the verifiable secret sharing (VSS) architecture. All mobile nodes do not need to authenticate the destination node, so as to protect the privacy of the node. Relay nodes can match the information of the destination node, and encrypt the information of the destination node. The signature of the node is exposed to the untrusted node, or by the message forwarding mechanism. The newly added node obtains part of the signature certificate through the chance to meet the initial node.

In conclusion, we can prove that a node belonging to the opportunistic social network will not be separated from the community if its only edge is connected to another node when the weight between the two nodes drops.

In the previous analysis, we designed efficient algorithms. Below, we consider the routing security issues caused by the message forwarding mechanism.

In the process of data transmission, if the private information of the node is exposed to the untrusted node, or the malicious node attacks the routing, the security of the routing algorithm will not be guaranteed. Public key encryption is one of the important means to protect the confidentiality of information [51]. Therefore, to ensure the security of the routing algorithm, on the one hand, we use the public key to encrypt the message, so as to ensure the confidentiality of the message and protect the privacy information from being exposed to untrusted nodes. On the other hand, we use the threshold signature method to avoid malicious nodes in the network from attacking the routing and improve the security of the routing algorithm.

We first set the corresponding trapdoor for each node and encrypt the information of the destination node. The relay node can match the information of the destination node, but cannot obtain any attribute information of the destination node, so as to protect the privacy of the node during data transmission. In addition, a public-private key pair is generated from the information of the node, and only the destination node can decrypt the information to ensure the confidentiality of the message. Among them, the public key generator is responsible for generating and maintaining the private key of the property, only distributing its trapdoor to the nodes. It avoids the private key exposure problem that may be caused by the capture of a single node [52].

In addition, due to the lack of security mechanisms, the authenticity of node information cannot be detected, and there are malicious nodes in opportunistic social networks. By forging and publishing false information, malicious nodes deceive the encountering nodes to forward data packets to them, so as to tamper and delete the data packets, thus forming a black hole attack in the network [53–55]. To resist the attack of malicious nodes, the nodes jointly generate corresponding private keys for the social attributes they possess and distribute these private keys to each node. Among them, the generation and distribution of all private keys are based on the verifiable secret sharing (VSS) architecture [56]. The newly added node obtains part of the signature service through the chance to meet the initial node and then reconstructs the signature certificate of its own attributes. All mobile nodes do not need to authenticate the identity of the encountering node when forwarding messages. They judge the validity of the social attributes of the encountering nodes by verifying whether the neighbor nodes have valid signatures, thereby resisting potential routing attacks in opportunistic social networks.

Our approach can effectively identify and exclude potentially malicious nodes, reducing the number of packets being intercepted and dropped, thereby protecting the network against opportunistic social network routing attacks.

### 4. Performance Evaluations

#### 4.1. Datasets

We use real datasets, the Cambridge experimental dataset of haggle for simulation experiments to evaluate the algorithm performance [57]. The dataset is the collection and statistics of at least one month’s user groups’ social activities carrying portable mobile devices in environments such as cities, conferences, and offices. The simulation experiment is carried out in a computer laboratory.

#### 4.2. Simulation Setup

This article uses one simulator to perform simulation training on the proposed algorithm and compares it with some message transmission algorithms applied to opportunistic social networks, such as EIMST algorithm [46], epidemic [58], FCNS [47], and spray and wait [59]. In the simulation, the mobile nodes in the map are based on the shortest path movement (SPMBM); that is, the actual map data is collected through the minimum distance coordinates. We determine the simulation parameters through the random model in the opportunistic network. Specifically, the movement parameters used in the experiment are shown in Table 1.

#### 4.3. Experiments and Discussion

In the evaluation experiment, the experimental results mainly focus on the impact of a node’s cache space and the number of nodes on the...
performance of five opportunistic network routing algorithms. Therefore, this chapter constantly changes these two variables to observe the changes in delivery ratio, end-to-end delay on average, and overhead on average.

Figure 2 shows the delivery ratio comparison results of different algorithms with different node cache. As shown in the figure, the five algorithms’ delivery ratio shows an upward trend as the node cache increases. This is because the task can obtain the target data from the node cache and return it, thereby effectively improving the response speed and allowing limited resources to serve more user nodes. It can be seen that the transmission rate of the EDOE algorithm is better than the other four algorithms in most cases, even as high as 0.91. Because the EDOE algorithm measures user noise through social user nodes, adjusts the adaptability between user nodes, and reduces node noise information. In opportunistic networks, popular algorithms and Spray and Wait algorithms are typical routing algorithms based on flooding strategies. So they generate many message copies to forward messages during the information transmission process, resulting in a reduction in network transmission efficiency. It is not difficult to see from the figure that these two algorithms’ transmission rate is always lower than 0.55. Besides, the EIMST algorithm conducts message transmission through the collaboration of multiple nodes. When the buffer space of the node is limited, the efficiency is low. When the nodes of the FCNS algorithm transfer information, due to the nodes’ low movement efficiency and the insufficient computing power of the message carrier, the transmission efficiency of this algorithm is not high. Based on the above analysis, it can be concluded that the transmission rate of the EDOE algorithm is better than the other four algorithms when the cache is increased.

Figure 3 compares the end-to-end delay on average of the five algorithms with different node cache. As the node cache space increases from 5 Mb to 30 Mb, the end-to-end delay of various algorithms shows an inevitable downward trend. As shown in the figure, the end-to-end delay decline curve of the spray and wait algorithm and EIMST algorithm tends to be similar, and the epidemic algorithm delay is higher than the other four algorithms. This is because its information diffusion capability is not as good as the other four algorithms, and a large amount of copied information is generated when transmitting on the community, which causes a high transmission delay. The end-to-end delay on average of the EDOE algorithm and the FCNS algorithm decreases to a small degree, and the delay is always less than 120. The FCNS algorithm can analyze the transmission preference before data transmission by comparing node movement similarity, so the end-to-end delay on average is lower than traditional routing algorithms. Our proposed algorithm can reduce inefficient nodes that are not conducive to the transmission process through community reduction strategies and information screening, reducing the end-to-end delay on average. In summary, among these five algorithms, the end-to-end delay on the average performance of EDOE is the best.

Figure 4 shows the comparison of the overhead on average of the five algorithms with different node cache spaces. As the node’s cache space increases, the spray and wait algorithm and the epidemic algorithm of the five algorithms have more massive network overhead, with the highest values approaching 280 and 260, respectively. Because in the epidemic algorithm and the spray and wait algorithm, a large amount of repeated data information will occupy the node’s cache space, and the node with a large amount of data is likely to consume more energy during the movement. Besides, the network overhead of the EIMST algorithm and the FCNS algorithm are both in the midstream
position. As the node’s cache space increases from 5 Mb to 30 Mb, the network overhead drops by nearly 150. The EIMST algorithm can effectively manage the memory space and node messages and realize the reasonable allocation of resources. For the FCNS algorithm, the carry storage required for information transmission is relatively small, so it is relatively stable in network overhead performance. The fluctuation range of the network overhead of the EDOE algorithm is the smallest. In message transmission, the algorithm can effectively limit the number of message copies in the network through community optimization and neighbor node prediction, reduce the number of message hops from the source node to the destination node, and thus control the unnecessary energy consumption of the node.

Figure 5 compares the delivery ratio of the five algorithms when the number of nodes is different. The delivery ratio of the EIMST algorithm, the FCNS algorithm, and the spray and wait algorithm are in the midstream position, while the epidemic algorithm’s delivery ratio is at a relatively low level. The EIMST algorithm and the FCNS algorithm use cache management and fuzzy inference mechanisms to effectively evaluate the social relationship between the relay node and the destination node, thereby improving message routing and forwarding efficiency. Also, it may take a long time for the relay node to meet the destination node in the second stage of the spray and wait algorithm, which will cause an inevitable decrease in the transmission rate in the network. The average delivery ratio of the EDOE algorithm is about 0.70, which is nearly 30% higher than the epidemic algorithm. The EDOE algorithm can effectively improve the relay node’s probability of meeting the destination node in the communication area through reasonable detection and division of the node community and information noise evaluation.

Figure 6 shows the comparison results of the end-to-end delay on average of five algorithms with different number of nodes. As the number of nodes increases, the average end-to-end delay performance of all algorithms shows a downward trend overall. The flooding mechanism often leads to a higher end-to-end delay on average in the network environment, so the epidemic algorithm’s delay is always higher than the other four algorithms. The EIMST algorithm can maximize the rational use of each node’s cache space through node collaboration and cache management mechanisms, thereby reducing the end-to-end delay on average. As shown in the figure, the end-to-end delay on average of the EDOE algorithm and the FCNS algorithm is always between 20 and 65. As a routing algorithm based on traditional fuzzy inference, the FCNS algorithm increases the success rate of data routing and forwarding by comparing the social and mobile similarities between nodes, thereby making reasonable data transmission decisions. Besides, the EDOE algorithm can also predict the probability that the relay node will reach the destination node’s community at the next moment, evaluate the mobility relationship between the relay node and the destination node, and reduce the number of message routing hops in the network, thereby significantly reducing the end-to-end delay on average. In summary, the EDOE algorithm and the FCNS algorithm perform better in terms of the end-to-end delay on average performance among the five algorithms.

Figure 7 shows the performance comparison of the five algorithms’ overhead on average under different numbers of nodes. As the number of nodes increases from 100 to 600, the epidemic algorithm’s overhead on average has shown an upward trend. This is because too many participants generate duplicate data information and may cause relatively high network overhead during the movement. As shown in the figure, the other four algorithms’ overhead
on average fluctuates little, tending to be flat or slightly rising, because these algorithms are relatively more stable. The network overhead of the EDOE algorithm is always kept to the lowest level because it is committed to storing data information and transmitting messages through the cache cooperation between multiple nodes, reducing the energy consumption of nodes.

5. Conclusions

In this study, a multiple evaluation method (EDOE) for noise reduction optimization is established. Through effective data optimization and evaluation algorithms to deal with noisy data, the problem of important information loss when data noise is lost is avoided. In addition, the information transmission task is performed by community nodes, which effectively reduces the overhead of user nodes and improves the performance of the algorithm. The EDOE method effectively avoids network congestion and packet loss and improves the overall operational efficiency of the network.

In the future, with the development of network communication technology and artificial intelligence, we will further study the trusted interaction of nodes and analyze network rules through machine learning methods according to the characteristics of nodes and the characteristics of social networks. The trust routing table is established according to the selected nodes, and the time stamp mechanism is used to prevent the routing table from being tampered with by malicious nodes during the feedback process. By preventing the competition of malicious nodes and rationally allocating network resources, the security and reliability of data transmission are further improved.

Data Availability

“Data Availability” statement data used to support the findings of this study are currently under embargo, while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The Natural Science Foundation of Hunan Province (Grant No. 2020JJ4647, Grant No. 2020JJ6064) supports this work. The outstanding youth project of the Hunan Education Department (Grant No. UQ19B056) supports this work. The general project of Hunan Education Department (Grant No. 19C0171) supports this work. The general project of Changsha Technology Bureau (Grant No. KC1809015) provides this work with support.

References


Website of CRAWDAD project,” http://crawdad.org/cambridge/haggle/20090529/.
