

Optimal Relay Selection Algorithm for Combining Distance and Social Information in D2D Cooperative Communication Networks

Kaijian Li¹(✉), Jianxin Dai², Chonghu Cheng¹, and Zhiliang Huang³

¹ College of Telecommunications and Information Engineering,
Nanjing University of Posts and Telecommunications, Nanjing 210023, China
15651036692@163.com, chengch@njupt.edu.cn

² School of Science, Nanjing University of Posts and Telecommunications,
Nanjing 210023, China
daijx@njupt.edu.cn

³ College of Mathematics, Physics and Information Engineering,
Zhejiang Normal University, Jinhua 321004, China
zlhuang@zjnu.cn

Abstract. With the rapid growth of mobile data traffic demand, D2D relay technology is becoming an essential technology for the next generation mobile network. In order to select the optimal node in a shorter time, a cooperative D2D relay model considering the physical distance and social information is proposed. And then a threshold based on distance and social information is introduced, which is used to filter out the nodes with poor performance to get a relatively small candidate relay set. According to the optimal stopping theory, this paper presents a D2D relay optimal selection algorithm in order to weigh the consumption of exploration and system performance. The simulation results show that the algorithm proposed is superior to the traditional algorithm in system performance and algorithm complexity.

Keywords: Device-to-Device (D2D) communications · Relay selection
Social information · Optimal stopping theory · Cooperative communication

1 Introduction

The demand for mobile data is growing over the next decade, so it poses a huge challenge to mobile networks. In general, there are three main ways to improve the capacity of wireless networks: increasing the spectrum resources, improving the spectrum utilization and improving the spatial multiplex ratio [1–7]. However, due to the lack of spectrum resources, the capacity of the wireless communication system improved by increasing the spectrum is limited, and the price of spectrum is very expensive. In recent years, with the continuous development of mobile communications, spectrum utilization continues to increase and has gradually close to the Shannon limit, so just by improving the spectrum utilization is difficult to meet the huge business needs. On this basis, the increase in spectral space reuse rate will be the inevitable choice to improve the capacity of wireless network. D2D (Device-to-Device)

technology, refers to allowing cell users to directly perform end-to-end communication under the control of the base station, so it has great potential to solve the problem. D2D communication can improve the spectral efficiency of the cellular communication system, reduce battery energy consumption of the mobile device. In contrast, Bluetooth, WiFi and other traditional end-to-end communication has many drawbacks. First, their communication distances are short in many cases they can't meet the needs of users. Second, Bluetooth needs to manually set the terminal pairing, and WiFi access point requires user-defined settings. In addition, both work in the non-authorized band, which will lead to unstable communications and poor communication quality. In contrast, the D2D communication technology assisted by the cellular network has a wider application prospect, and the related research has very important theoretical significance and application value. D2D relay technology is a key implementation of D2D communication. When the channel quality between D2D users is not ideal, a suitable relay user can be found to establish a connection, which enables users to communicate and greatly improve the capacity of the base station.

In addition, because the energy and storage space of the mobile terminal are limited, the design of the relay algorithm must be simple enough. Although the base station can obtain the channel information of all potential relay users, it will not only increase the burden of the base station, but also take many time slots to complete the process of finding the relay. Therefore, in the D2D relay network, the choice of optimal relay and encouraging relay nodes to forward data are worth studying.

Because the carriers and users of mobile devices are people, it forms a mobile social network. A series of parameters in social networks, such as social relationships, centrality, and communities, can reflect relationships among mobile users. Using the behavior of people in social networks can help solve the problem of relay selection in D2D communications. Because, for a mobile user, there are usually family members, neighbors, friends, or colleagues nearby. Therefore, most of the potential relay nodes of the user have a social relationship with him, so that they can choose the trusted node to forward data for themselves, thus improving the information security, and because of social relations, these relay nodes are more willing to forward data.

There are already many options for D2D relay schemes. In [8], a random selection scheme is proposed, which does not filter the potential nodes, and then randomly takes a node as a relay. In [9], a terminal device with energy acquisition function is proposed as relay, and an optimal relay selection is proposed to minimize the probability of interruption. In [10], a scheme is proposed to select the relay scheme for D2D users to maximize the signal-to-noise ratio. In [11], an optimal stopping scheme is proposed based on the SNR threshold structure. Similarly, the document [12] combines social information with an optimal stop scheme to stimulate relay users to relay information through intrinsic factors. However, these schemes have a common drawback that they must obtain information about each potential relay node, which requires not only a large amount of signaling overhead, but also a lot of energy consumption.

The rest of this paper is arranged as follows. D2D relay system model will be proposed in Sect. 2. In Sect. 3, we introduce the algorithm proposed in this paper. Section 4 presents the simulation results, we give conclusions in Sect. 4.

2 System Model

In the section, a social networking model to quantify the social relationships between the users firstly is built. Besides, a D2D relay system model on social relations network is proposed.

2.1 Social Network Model

Today's social network architecture has subverted the original network structure, the architecture of the social layer network has been added to the original network architecture, which is based on social relationships. As shown in Fig. 1, the scenario considered in this article is data transmission within a single cell D2D mode in a cellular network. The model is a two-tier model that includes a social layer and a physical layer, where the physical layer includes multiple D2D users and cellular users. D2D users reuse uplink resources of cellular users. Sometimes, even in the same cell, the distance between the two D2D users is far away, and the connection cannot be established directly, so communication must be achieved by relay, for example, DUE1 and DUE2 in Fig. 1.

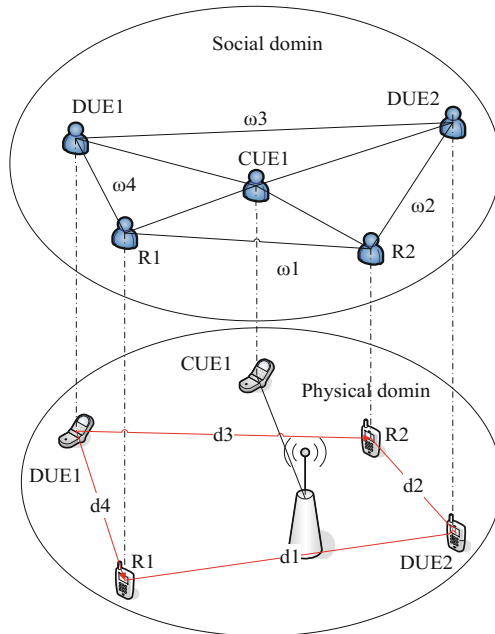


Fig. 1. Two-tier model

There may be multiple idle relays between DUE1 and DUE2, which can be used as potential relays, and we need to select the best nodes as the relay. Suppose there are n D2D user devices (DUE) in the diagram, which can be represented as a set

$D = \{D_1, D_2, \dots, D_n\}$, and assuming that there are m adjacent users near D_i , the adjacent user set can be expressed as $R_i = \{R_{i,1}, R_{i,2}, \dots, R_{i,m}\}$. We exploit an indirect weighted graph $G^P = \{V, E^P\}$ to model the physical relation between DUE D_i and its each neighbor user. In graph G^P , we build the vertex set V , and the set of edges $E^P = \{(D_i, R_{i,j}) : e_{D_i, R_{i,j}} = 1, R_{i,j} \in R_i\}$, where $e_{D_i, R_{i,j}} = 1$ if $R_{i,j}$ can communicate directly with D_i . Moreover, in the social layer, social relationships are quantified with $\omega_{D_i, R_{i,j}}$ ($\omega_{D_i, R_{i,j}} \in [0, 1)$) between users, and the greater the $\omega_{D_i, R_{i,j}}$, the stronger the social relationship between the two users, it also means that they are more willing to forward data to each other. But in fact, we can't directly measure the social relationships between two users, so we have to quantify social relationships. Typically, there are two factors that directly affect social relationships, namely, time factors and interaction factors. This paper uses the average length of time $T_{D_i, R_{i,j}}$ between two users to express the time factor, which can visually reflect the frequency of contact between two users:

$$E(T_{D_i, R_{i,j}}) = \frac{\int_0^\phi \delta_{D_i, R_{i,j}}(t) dt}{S_{D_i, R_{i,j}}} \tag{1}$$

where ϕ denotes the observation time, $S_{D_i, R_{i,j}}$ is the number of communications between two users during ϕ . $\delta_{D_i, R_{i,j}}(t) = 1$ if D_i and $R_{i,j}$ communicate within the observation, otherwise $\delta_{D_i, R_{i,j}}(t) = 0$.

According to [13], we get the time factor as follows:

$$F_{D_i, R_{i,j}} = \exp\left(-\frac{E(T_{D_i, R_{i,j}})^2}{2\sigma^2}\right) \tag{2}$$

where σ denotes the length of the previous communication. The second interactive factor is the number of common friends between two users. The more friends there are between two users, the more intimate their social relationships are. According to [14], a common friend index is proposed to represent interaction factors:

$$K_{D_i, R_{i,j}} = \frac{|k_{D_i} \cap k_{R_{i,j}}|}{|k_{D_i} \cup k_{R_{i,j}}|} \tag{3}$$

where $|\cdot|$ is the cardinality of set, k_{D_i} and $k_{R_{i,j}}$ indicate the friends set of D_i and $R_{i,j}$.

Use the combination of two factors, we can indicate social relationships with $\omega_{D_i, R_{i,j}}$ as:

$$\omega_{D_i, R_{i,j}} = \alpha F_{D_i, R_{i,j}} + (1 - \alpha) K_{D_i, R_{i,j}} \tag{4}$$

where α ($\alpha \in [0, 1]$) is used to adjust the impact of two factors on social relationships. In different scenarios, the two factors may have different effects on social relationships, so α may be different.

In physical layer, the distance between users will affect the quality of communication. It is obvious that the quality of communication is inversely proportional to the distance between users, and in the relay selection process, users prefer to select nodes closer to their own.

2.2 Social Network Analysis in Specific Scenarios

As mentioned above, distance and social factor are the two major factors that affect the choice of relay, so we are talking about it in social networks. As shown in Fig. 2, we consider a single cell model for mixed communication of CUE and DUE. CUE is represented later with C, while D2D users can be divided into sender and receiver, represented with DT and DR respectively. Suppose there are N candidate relay nodes between DT and DR , which can be represented as sets $N = \{n_1, n_2, \dots, n_N\}$. These candidate relay nodes are the intersection of adjacent nodes of DT and DR , represented by the shadow part in Fig. 2. Because D2D users reuse the downlink resources with cellular users, D2D communications will inevitably interfere with cellular users. The channel gain between the node i and the node j can be represented by $|h_{i,j}|^2$, which follows the exponential distribution of the parameter $\lambda_{i,j}$. To ensure that the transmission power P_{DT} does not exceed the maximum interference threshold I_{th} of CUE, while ensuring that the DT can achieve the optimum QoS, we define transmission power as $P_{DT} = \frac{I_{th}}{|h_{C,DT}|^2}$. As mentioned earlier, the stronger the social relationship between the two nodes, they are more willing to help each other. So the SNR received by the relay node $n_i (n_i \in N)$ from the DT is proportional to the social relationships weight coefficient between them, which can be expressed as:

$$S_{n_i} = \omega_{DT,n_i} \frac{I_{th} |h_{DT,n_i}|^2}{\sigma^2 |h_{C,DT}|^2} \tag{5}$$

where ω_{DT,n_i} stands for the intensity of social relationships between DT and node n_i , and σ^2 represents Gaussian noise variance. According to [12], $|h_{DT,n_i}|^2 = \frac{1}{d_{DT,n_i}^\alpha}$,

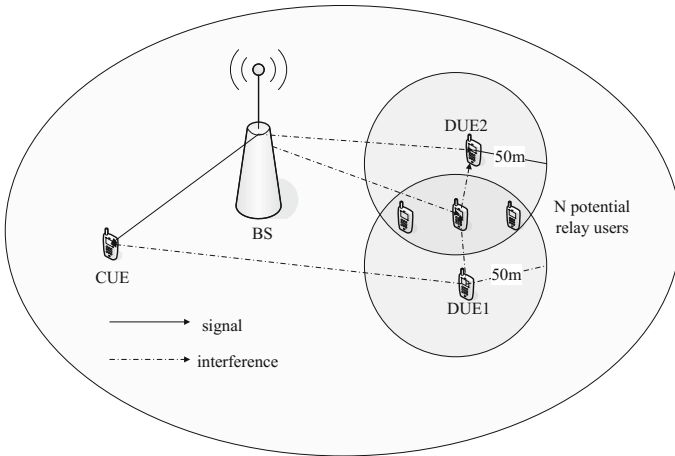


Fig. 2. D2D relay network

where d_{DT,n_i} is the distance between the DT and the relay node n_i , so that it can be obtained:

$$S_{n_i} = \omega_{DT,n_i} \frac{I_{th}}{d_{DT,n_i} \sigma^2 |h_{C,DT}|^2} \tag{6}$$

Thus, sequence $\{S_{n_i}\}$ is a sequence of i.i.d random variables. Assuming $G = \frac{\lambda_{C,DT} I_{th}}{\sigma^2}$. The probability density function of S_{n_i} can be calculated based on [15], that is:

$$f(S_{n_i}) = \omega_{DT,n_i} \frac{\lambda_{DT,n_i} G}{(\lambda_{DT,n_i} S_{n_i} + G)^2} \tag{7}$$

Based on the Shannon formula, the channel capacity of S_{n_i} can be obtained as:

$$C(S_{n_i}) = W \log_2(1 + S_{n_i}) \tag{8}$$

where W is the channel bandwidth.

In Fig. 2, the quality of the channel between DT and DR is poor and the relay node must be sought to forward the data, so it is important to find a suitable relay. In order to find the optimal relay, DT must explore all candidate relays, which is not only inefficient, but also high energy consumption. Therefore, in order to balance performance and consumption, we propose a D2D relay optimal selection algorithm based on distance and social information to solve this problem.

3 The Proposed OSRS Algorithm

First, a threshold based on distance and social information is introduced, which is used to filter out the nodes with poor performance to get a relatively small candidate relay set, then the optimal relay node is found by the optimal stopping theory.

3.1 Threshold Based on Distance and Social Information

From the formula (6) we can see that the SNR S_{n_i} of the relay n_i is proportional to the ratio of the social factor and the distance $\frac{\omega_{DT,n_i}}{d_{DT,n_i}}$. So we can first screen the ratio of social factors and distances, which can not only reduce the number of searches, but also save energy consumption. We use φ to represent the binding threshold:

$$\varphi = \frac{\rho \sum_{i=1}^N \frac{\omega_{DT,n_i}}{d_{DT,n_i}}}{N} \tag{9}$$

where ρ ($\rho \in (0, 1)$) is a percentage value used to adjust the size of φ . ω_{DT,n_i} in the DT communication list can be obtained, and d_{DT,n_i} through the base station to obtain.

Therefore, when the base station gets the relevant information, it does not immediately traverse all the relays, but instead filters out the nodes with poor performance through this information. Through φ , DT filters the nodes based on the following principles:

- Abandon node n_i if $\frac{\omega_{DT,n_i}}{d_{DT,n_i}} < \varphi$,
- Reserve node n_i if $\frac{\omega_{DT,n_i}}{d_{DT,n_i}} \geq \varphi$

Thus, a relatively small candidate relay set can be obtained from the above filter, which can be represented as $M = \{m_1, m_2, \dots, m_M\}$, but then we still need to select the best relay nodes. Because we have to weigh the consumption and system performance, based on the optimal stopping theory, we propose a solution to select optimal node as relay.

3.2 Relay Selection Algorithm

As can be seen from Fig. 3, Assumed that the total length of a slot is T , and this time can be divided into two parts: the exploration time and the transmission time. Where the exploration time $i\tau$ is used to find the optimal relay time, i on behalf of the number of exploration, each time slot τ is a time to explore, and the total transmission time for the $T - i\tau$. So we define the immediate reward r_{m_i} that the DT can achieve from node m_i :

$$r_{m_i} = \mu_i C(S_{m_i}) \tag{10}$$

where $\mu_i (\mu_i \in (0, 1))$ is a discount factor that decreases as the number of explorations increases and can be expressed as:

$$\mu_i = 1 - \frac{i\tau}{T} \tag{11}$$

The larger the instantaneous reward, the better the performance of the node. In the process of probing, we are not sure whether the node is the best node. So a reward

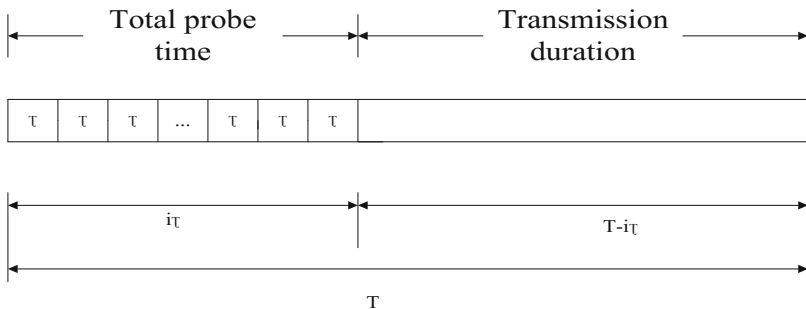


Fig. 3. A slot of the relay

threshold expressed as $\theta_{m_i}^* = E\{R_{m_{i+1}}\}$ in each probing stage is set. Here $E\{R_{m_{i+1}}\}$ is reward expected while DT is probing next node. Based on (10), when the user relay node is explored, the maximum reward that DT can obtain can be expressed as:

$$R_{m_i} = \max\{r_{m_i}, \theta_{m_i}^*\} \tag{12}$$

It is clear that when R_{m_i} equals to r_{m_i} , we can get the best node m_i , and stop probing, otherwise DT will explore the next potential node m_{i+1} .

When the i -th exploratory phase, DT must make the following choices:

- If $r_{m_i} < \theta_{m_i}^*$, ignore the node and explore next node
- If $r_{m_i} \geq \theta_{m_i}^*$, stop exploring and selecting the node as relay to forward data

And the reward threshold $\{\theta_{m_i}^*\}_{i=1}^M$ is defined as follows:

$$\theta_{m_M}^* = -\infty \tag{13}$$

$$\theta_{m_i}^* = \theta_{m_{i+1}}^* \int_0^{\frac{\theta_{m_{i+1}}^*}{\mu_{i+1}}} f(S_{m_i})dS_{m_i} + \mu_{i+1} \int_{\frac{\theta_{m_{i+1}}^*}{\mu_{i+1}}}^\infty W \log_2(1 + S_{m_i})f(S_{m_i})dS_{m_i} \tag{14}$$

Proof: We can use reverse derivation to prove this problem. In the last step of the exploration, DT does not have the remaining nodes to explore, but can only choose the last node as a relay, so $\theta_{m_M}^* = -\infty$. Then we can obtain:

$$\begin{aligned} \theta_{m_i}^* &= E\{R_{m_{i+1}}(C(S_{m_{i+1}}))\} \\ &= E\{\max\{r(S_{m_i}), \theta_{m_{i+1}}^*\}\} \\ &= \theta_{m_{i+1}}^* \int_0^{\frac{\theta_{m_{i+1}}^*}{\mu_{i+1}}} f(S_{m_i})dS_{m_i} + \mu_{i+1} \int_{\frac{\theta_{m_{i+1}}^*}{\mu_{i+1}}}^\infty W \log_2(1 + S_{m_i})f(S_{m_i})dS_{m_i} \end{aligned} \tag{15}$$

For the above proved sequence $\{\theta_{m_i}^*\}_{i=1}^M$, we can also prove that it is monotonically decreasing. Combine (13) and (14) we can know $\theta_{m_M}^* \geq \theta_{m_{M-1}}^*$. According to Sect. 2, we have the sequence of $\{S_{m_i}\}_{i=1}^M$ is i.i.d. sequence. So the following can be deduced as:

$$\begin{aligned} \theta_{m_{M-2}}^* &= E\{\max\{r_{m_{M-1}}, \theta_{m_{M-1}}^*\}\} \\ &= E\{\max\{\mu_{M-1} W \log_2(1 + S_{m_{M-1}}), \theta_{m_{M-1}}^*\}\} \\ &\geq E\{\max\{\mu_M W \log_2(1 + S_{m_{M-1}}), \theta_{m_{M-1}}^*\}\} \\ &= E\{\max\{\mu_M W \log_2(1 + S_{m_M}), \theta_{m_{M-1}}^*\}\} \\ &\geq E\{\max\{\mu_M W \log_2(1 + S_{m_M}), \theta_{m_M}^*\}\} \\ &= E\{\max\{r_{m_M}, \theta_{m_M}^*\}\} \\ &= \theta_{m_{M-1}}^* \end{aligned} \tag{16}$$

Then we can deduce $\theta_{m_1}^* \geq \theta_{m_2}^* \geq \dots \geq \theta_{m_M}^*$. So the sequence $\{\theta_{m_i}^*\}_{i=1}^M$ is monotonically decreasing. According to these analysis, a relay selection algorithm based on threshold which combine distance and social information is proposed, as shown in Algorithm 1.

Algorithm 1. Relay selection algorithm based on threshold which combines distance and social information

1. Initialize the D2D transmission system, including DT and DR
 2. Obtain their neighbor set K_{D_T}, K_{D_R}
 3. Obtain the intersection N of the DT and DR neighbor sets, and obtain their social coefficients ω_{D_T, n_i} with DT, and obtain their distance d_{D_T, n_i} from the DT through the base station
 4. Calculate φ by (9)
 5. In the set N, select the M nodes that satisfy $\frac{\omega_{D_T, n_i}}{d_{D_T, n_i}} \geq \varphi$
 6. $\theta_{m_M}^* = -\infty$
 7. for $j = M - 1$ to 1
 Calculate $\theta_{m_j}^*$ according to (14)
 end for
 8. for $j = 1$ to M
 Calculate r_{m_j} according to (10)
 if $r_{m_j} \geq \theta_{m_j}^*$ then
 Stop the exploration process, the optimal node j is obtained
 end if
 end for
-

4 Simulation

The relay selection algorithm over this section will be analyzed. This section considers the distribution of D2D users and cellular users in a single cell scenario. We first consider the reward of the relay and compare it with the traditional algorithm. Besides, we simulate the number of the whole relay exploration. This algorithm is mainly compared with the social relation based optimal stop algorithm (SARS) proposed by [12]. The main simulation parameters are given in Table 1.

In Fig. 4, we can see clearly that the reward value provided by D2D increases as the number of candidate relays increases. Because as the number of candidate relays increases, the probability of obtaining a better relay increases. In addition, we can find that the algorithm proposed in this paper is superior to the traditional algorithm in terms

Table 1. Simulation parameters

Parameter	Value
τ	0.01 s
T	1 s
N	10
λ	1
G	5

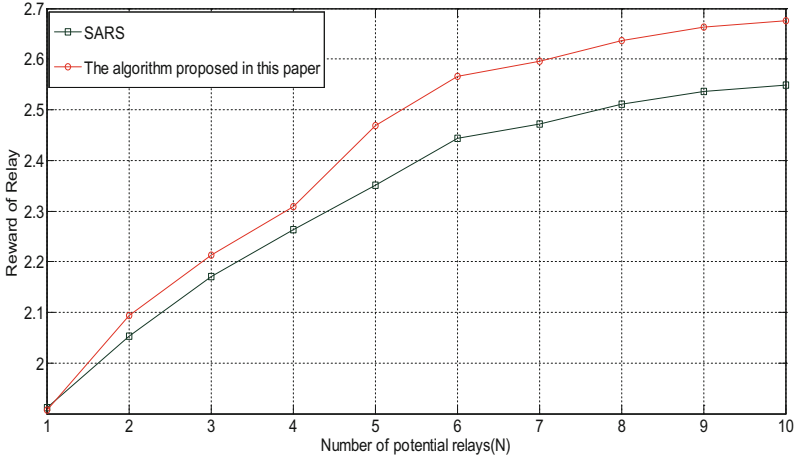


Fig. 4. Returns of different numbers of potential relays

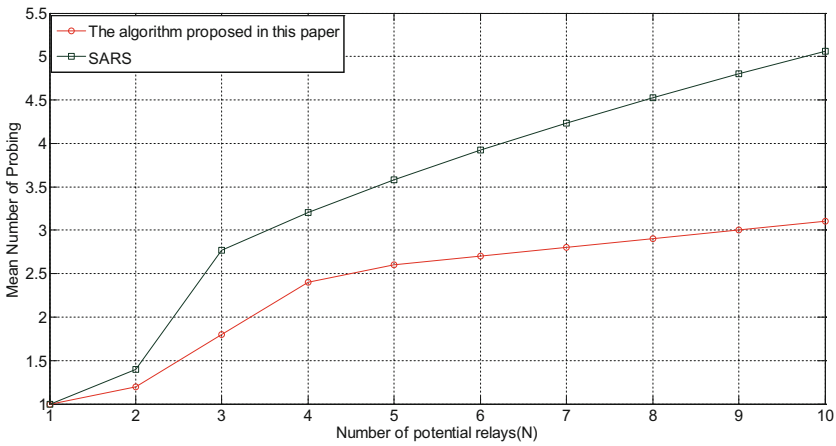


Fig. 5. The probing numbers with different number of potential relays

of relay reward, because the algorithm of this paper has screened some poorly performing nodes according to the social coefficient and distance before the relay exploration began.

In Fig. 5, we can see that as the number of candidate relay increases, the average number of explorations increases, and the average number of explorations proposed in this paper is significantly less than that of traditional algorithms. This is because we have filtered the poorly performing candidate nodes through thresholds before exploring the iterations so that we can find the optimal relay nodes faster.

5 Conclusion

In this paper, a threshold is obtained by combining the social coefficients and distances between D2D users, and an optimal relay selection algorithm is proposed based on the optimal stopping theory. And the simulation results show that the algorithm is superior to the traditional algorithm. However, there are some areas that are not fully considered. For example, only a single cell scene is considered, but not in a multi cell scenario. This is the future research direction.

Acknowledgments. The research was supported in part by Postdoctoral Research Funding Plan in Jiangsu Province (Grant No. 1501073B), Natural Science Foundation of Nanjing University of Posts and Telecommunications (Grant No. NY214108), Natural Science Foundation of China (NSFC) (Grant No. 61401399), and the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (Grant No. 2016D05).

References

1. Shanmugam, K., Golrezaei, N., Dimakis, A.G., Molisch, A.F., Caire, G.: Femtocaching: wireless content delivery through distributed caching helpers. *J. IEEE Trans. Inf. Theory* **59**(12), 8402–8413 (2013)
2. Xie, H., Gao, F., Zhang, S., Jin, S.: A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model. *IEEE Trans. Veh. Technol.* (2016). Earlier access available online
3. Xie, H., Gao, F., Zhang, S., Jin, S.: An overview of low-rank channel estimation for massive MIMO systems. *IEEE Access* **4**, 7313–7321 (2016)
4. Xie, H., Gao, F., Zhang, S., Jin, S.: A full-space spectrum-sharing strategy for massive MIMO cognitive radio. *IEEE J. Sel. Areas Commun.* **34**(10), 2537–2549 (2016)
5. Ma, J.P., Li, H.Y., Zhang, S., Zhao, N., Arumugam, N.: Pattern division for massive MIMO networks with two-stage precoding. *IEEE Commun. Lett.* **21**(7), 1665–1668 (2017)
6. Zhao, N., Yu, F.R., Leung, V.C.M.: Opportunistic communications in interference alignment networks with wireless power transfer. *2015 IEEE Commun. Soc.* **22**, 88–95 (2015). MWC 2015
7. Li, X.H., Zhao, N., Sun, Y., Yu, F.R.: Interference alignment based on antenna selection with imperfect channel state information in cognitive radio networks. *IEEE Trans. Veh. Technol.* **65**, 5497–5511 (2016)
8. Pan, P., Zheng, B.: Outage probability analysis of a cooperative diversity method based on random relay selection. *Acta Electron. Sin.* **38**(1), 79–82 (2010)

9. Yang, H.H., Lee, J., Quek, T.Q.S.: Heterogeneous cellular network with energy harvesting-based D2D communication. *IEEE Trans. Wirel. Commun.* **15**(2), 1406–1419 (2016)
10. Ryu, H.S., Lee, J.S., Kang, C.G.: Relay selection scheme for orthogonal amplify-and-forward relay-enhanced cellular system in a multi-cell environment. In: 71st 2010 IEEE Vehicular Technology Conference (VTC 2010-Spring), pp. 1–5. IEEE (2010)
11. Zheng, D., Ge, W., Zhang, J.: Distributed opportunistic scheduling for ad-hoc communications: an optimal stopping approach. In: Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing, pp. 1–10. ACM (2007)
12. Zhang, M., Chen, X., Zhang, J.: Social-aware relay selection for cooperative networking: an optimal stopping approach. In: 2014 IEEE International Conference on Communications (ICC), pp. 2257–2262. IEEE (2014)
13. Cai, Y., Wu, D., Yang, W.: Social-aware content downloading mode selection for D2D communications. In: 2015 IEEE International Conference on Communications (ICC), pp. 2931–2936. IEEE (2015)
14. Von, L.U.: A tutorial on spectral clustering. *Stat. Comput.* **17**(4), 395–416 (2007)
15. Berggren, F., Jantti, R.: Multiuser scheduling over rayleigh fading channels. In: Global Telecommunications Conference (GLOBECOM 2003), vol. 1, pp. 158–162. IEEE (2003)