

RESEARCH

Open Access



Improving cellular downlink throughput by multi-hop relay-assisted outband D2D communications

Kai Zhou¹, Jinsong Gui^{1*}  and Naixue Xiong²

Abstract

One goal of the fifth-generation (5G) cellular network is to support much higher data capacity (e.g., 1000 times higher than today), where device-to-device (D2D) communication is one of the key enabling technologies. In this paper, we focus on the D2D relaying functionality to improve cellular downlink throughput. Based on the shortages of the latest relevant work, we propose a new scheme that leverages multi-hop relay-assisted outband D2D communications. First of all, by extending two-hop connection to three-hop connection, our scheme can cut down receiving bit error ratio (BER) of cell edge nodes far away from a cellular base station (BS), which improves cellular downlink throughput. Then, it balances network lifetime and throughput by the proposed ratio of income and expenditure (RIE) metric with respect to remaining energy and throughput. Moreover, it reduces the computational overhead of searching relay and also ensures that the optimal relay is selected by the adjustment of searching scope. Compared with the most relevant work, our scheme outperforms it in terms of throughput, delay, and network lifetime.

Keywords: Cellular downlink throughput, Multi-hop relay-assisted, Outband D2D communication, Improvement

1 Introduction

The rapid evolution of global information and communication technology enables innumerable wireless devices (e.g., laptops, smartphones, tablets, phablets, wearable devices, embedded devices, and sensors [1–3]) to connect to each other simultaneously. Thus, more and more people are involved into an integration of physical, social, and mental space [4–6], which leads to an exponential rise in traffic demand and generates a significant burden on current cellular networks [7, 8]. The fifth-generation (5G) cellular network, with larger data rate, higher spectrum and energy efficiency, and smaller delay, supports a large number of connected devices and provides flexible wireless interfaces. In theory, it supports a 1000 times higher data capacity, accommodates different mobile devices 1000 times than today, and has lower cost for infrastructures and devices.

The method of D2D communication is one of the key enabling technologies of a 5G wireless network.

Meanwhile, it is an essential part of the next-generation cellular communication system. The D2D communication mode not only improves spectral efficiency but also potentially enhances throughput, energy efficiency, delay, and fairness. This mode has several possible application scenarios, and we employ its relaying functionality in our scheme to improve throughput and extend communication range. D2D communication mode can be divided into inband D2D mode and outband D2D mode. The former occurs in a licensed spectrum that is allocated to a cellular operator, while the latter uses an unlicensed spectrum adopted by other wireless technologies that support direct communication (e.g., WiFi).

In the outband D2D mode, there is no interference between a D2D link and a cellular link. However, with the increasing usage of unlicensed spectrums, it is difficult to control the co-channel interference level between different D2D links. In addition, the outband D2D communications can be only used by wireless devices which are equipped with two types of radio interfaces (e.g., long-term evolution (LTE) interface and WiFi interface). Although the device performance in outband D2D communications is demanded more strictly, with more

* Correspondence: jsgui06@163.com

¹School of Information Science and Engineering, Central South University, South Road of LuShan, Changsha 410083, China

Full list of author information is available at the end of the article

components highly integrated into wireless devices, the limitation of interfaces of wireless devices will not become the bottleneck of the outband D2D communications. Besides this, 802.21 (see [9]), the standard of wireless interfaces, is going to integrate different heterogeneous networks (e.g., WiFi, LTE, and WiMAX), which brings a smooth and hidden switchover between various network interfaces. Therefore, more and more researchers start to explore outband D2D communications.

The majority of the existing works only propose concept, structure, or simple heuristic algorithms for outband D2D schemes. Nevertheless, some literatures formulate their schemes about optimizing problems of outband D2D communications, where a heuristic idea is usually considered due to non-deterministic polynomial-time (NP)-hardness. For example, in [10], the authors design a channel-opportunistic framework with the purpose of enhancing network throughput under delay constraints, which is implemented by designing a greedy algorithm. But, there is enough room for us to improve this scheme. Firstly, it does not consider the energy constraints; the network lifetime could not be guaranteed. Then, it adopts repeated iteration and global search strategy which causes huge communication complexity and computational complexity. Moreover, it could not meet the demands of cell edge nodes in receiving bit error ratio (BER).

In this paper, we pay special attention to multi-hop relay-assisted outband D2D communications to reduce the BER of cell edge nodes with the objective of improving cellular downlink throughput. Therefore, cellular communication and the relay-assisted D2D communication should be jointly investigated and analyzed. The contributions of our scheme are summarized into the following points:

1. In order to avoid interfering neighboring cells, the transmission power of a cellular base station (BS) is no more than the specified maximum value. And the process of direct data transmission (from BS to anyone of user equipments (UEs) near the edge of cells) usually causes the problem of excessively high BER. However, the existing scheme (e.g., the work in literature [10]), using one relay to form a two-hop communication path, still hardly solves the problem of high BER. Therefore, a scheme of three-hop relay-assisted outband D2D communication is proposed in this paper to reduce receiving BER and improve throughput efficiently.
2. Our scheme constructs a new formula which could balance the performance of the first selected relay between downlink throughput and remaining energy. By setting a threshold of ratio of income and expenditure (RIE) and adjusting this threshold value, we could emphasize either of the two different performance metrics. On this basis, via energy constraints, we could ensure that the remaining energy of the second selected relay is not less than the minimum value between the first relay's energy and the energy threshold, and thus, the transmission path lifetime is also guaranteed when the throughput is improved.
3. To obtain related link throughput and construct transmission paths, the typical existing scheme uses the method of actual measurements of link throughput, which are then reported to the related components of infrastructure. After that, the throughput values of all potential transmission paths are estimated and they are regarded as the input parameters of the relay selection algorithm. However, since the calculation of co-channel interference of outband D2D links involves the number of co-channels, the distribution, and other information of interference sources, the above scheme needs several times of iteration to obtain a convergence value of transmission path throughput. Hence, the computational and communication complexity increase accordingly. The algorithm in this paper firstly adopts the strategy of relay preselection and then verifies whether the selected relay improves throughput or not. Finally, the relatively stable values of link throughput and communication paths are determined at a time instead of several times of iteration. Meanwhile, the communication complexity is reduced because of the one-round decision.
4. Based on the elaboration of the third point, we only consider cellular link throughput between BS and potential relays or receiving UEs, instead of uncertain throughput of outband D2D links, to construct relay preselection metrics. Meanwhile, in order to avoid too low D2D link throughput, which might become the throughput bottleneck of the entire transmission path, an estimation method of relay preselection is proposed in this paper to reduce the possibility of abandoning the preselected relays in the phase of relay verification. By adjusting the related parameter to control the scope of relay preselection, the distance between two communication UEs in an outband D2D link can be limited to an appropriate range, which reduces the number of abandoned preselected relays.
5. Different from the global search strategy for potential relays in the typical existing scheme, local search strategy is adopted in this paper. By reasonably setting a searching scope, the local computational complexity is reduced when the optimal relays are selected.

The remainder of this paper is organized as follows: An overview of related works on relay-assisted D2D communications is presented in the “Related work” section. This section is followed by the detailed scheme for improving cellular downlink throughput in the “The scheme for optimizing cellular downlink throughput” section. The performance of our scheme is evaluated, and the results are discussed in the “Performance evaluation” section. Finally, the conclusions are given in the “Conclusions” section.

2 Related work

In [11], an overview of papers primarily focusing on inband D2D communications in OFDMA cellular networks is provided, where the authors introduce the research status and major problems existing in D2D communications from different perspectives and then provide a number of existing solutions concerning these problems. Another survey on D2D communications is presented in [12], where the advantages and disadvantages of inband D2D and outband D2D are discussed, and also, the authors point out the major issues which occur in overlay D2D, underlay D2D, and outband D2D. Based on the shortcomings of present research analyzed in [12], the authors summarize the research directions in the future.

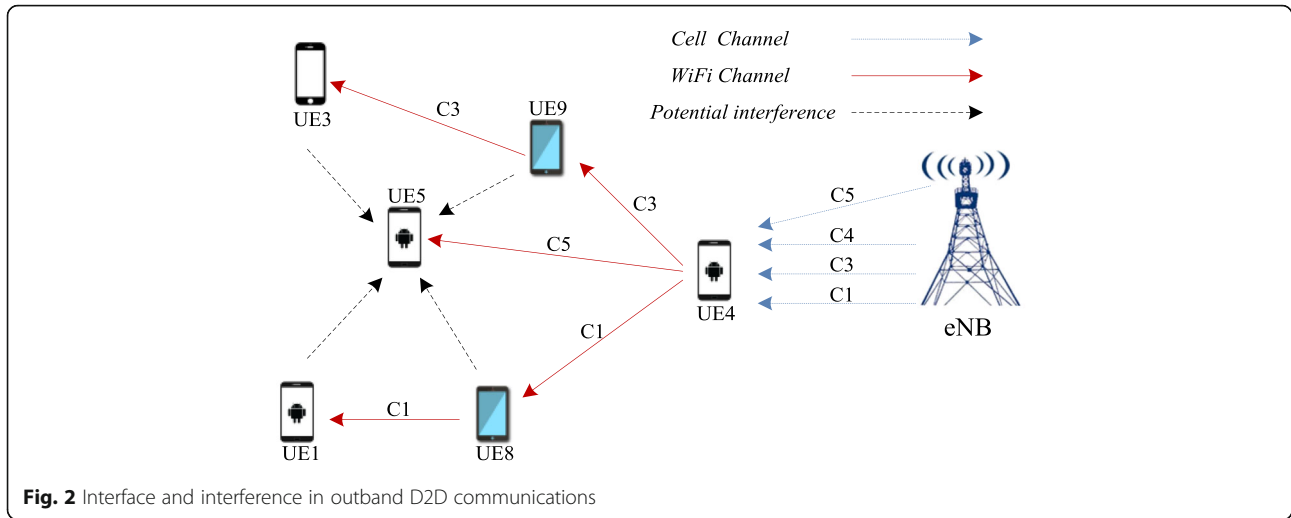
Focusing on the problems of D2D clustering, the authors in [13] propose a cluster-based scheme over D2D users to reuse spectrum resource. Meanwhile, by using an interference alignment (IA) transmission mode within small scale clusters, D2D users could transmit its data in higher spectrum efficiency. The authors in [14] propose a channel-opportunistic architecture to enhance throughput, energy efficiency, and fairness, in which clusters are formed by cellular UEs and then the UE owning the best channel status is selected as the cluster head to communicate with base station on behalf of the whole cellular UEs. Within clusters, the unlicensed spectrums are used between D2D pairs. Similar to [14], literature [15] proposes a clustering scheme and specifies how to use WiFi spectrum in D2D communications to offload data traffic from LTE-based cellular networks.

Another important issue of relaying functionality in D2D communications has also been investigated. Literature [16] utilizes the topology-controlled D2D communications to achieve better performance in capacity and coverage. It particularly points out that the edge UEs have to use mobile UEs to improve system’s capacity and extend the wireless network coverage because of its poor network transferring condition. And if more D2D UEs act as relays for the extension of network coverage, how are the relays selected? This problem is tackled in literature [17], where an algorithm for relay selection based on graph theory is proposed to address it. A spectrum-

sharing protocol is given in [18], where BS allows two UEs in proximity to initiate bidirectional D2D communications in the process of two-way communications, and the channel of bidirectional D2D link is reused in two time slots. A D2D UE is selected as a bidirectional relay only if it helps to achieve a higher bit rate for a cellular UE; hence, a higher system throughput is also guaranteed. Based on relay functionality of D2D communications, literature [19] proposes an approach to extend battery life, where the main objective is to help UEs with a low-battery level to save energy by relaying its data, and the relaying UE is selected according to the metric of the neighbors’ remaining battery and distance.

Aiming at reducing interference of inband communications and outband D2D communications, a radio resource allocation method about time hopping (TH) which has lower signal expenditures is proposed in [20]. The authors in [21] have formulated a semi-distributed resource allocation method to maximize spatial reuse by using the same radio resources in simultaneous transmission. Literature [22] investigates several resource allocation schemes, in which D2D communications and cellular communications share radio resources. Similar problem is also considered in [23, 24], where resource sharing mode selection and transmit power allocation are jointly considered to fulfill some target signal-to-interference-plus-noise ratio (SINR) values for each link. The authors in [25] prove that the reduced complexity of interference management and the improved energy efficiency could be implemented by outband D2D relay communications.

By studying the latest research findings, we find that most researchers focus on addressing the related problems of two-hop outband D2D communications. For example, the authors in [10] introduce a cellular downlink throughput optimization problem under delay constraints, which is solved by designing a greedy algorithm since it is NP-hard problem. In this paper, a UE is either in two-hop outband D2D communication mode or one-hop legacy cellular communication mode. Literature [26] solves that how edge mobile UEs improve fairness and throughput via idle UEs playing role of a full duplex relay, which is only equipped with two network interfaces. Meanwhile, this scheme restricts the number of relays to one in any transmission path and compares three kinds of relay selection criterion with the standard connectivity. However, literature [27] investigates a selling relays service based on tokens exchange in the two-hop outband D2D communications. As described in [11], interaction among conventional cellular communications, two-hop outband D2D communications and multi-hop outband D2D communications should be investigated and analyzed. In summary, an attractive option is to extend two-hop relay communications to



There is no doubt that lots of expenditures spent in user devices are saved while the range of selecting relays is limited.

Based on the above analysis, the proposed system model can be concluded as how to select communication mode reasonably to improve cellular downlink throughput according to the status information of each receiving UE, distribution of interference sources, neighboring UEs' status information and so on.

3.2 Problem formulation

By considering communication mode selection and the number of relays under end-to-end delay constraints and remaining energy constraints, we propose a method for downlink throughput optimization by multi-hop relay-assisted outband D2D communications (DTOMROD). In order to determine relays, the *eNB* in [10] utilizes real-time throughput (i.e., measured by real devices) and the way of repeated iteration. Different from [10], we use the way of relay preselection to determine both the throughput values and the relays just at a time, and all the phases are centrally processed by the *eNB*. Thus, according to the architecture proposed by 3GPP, we also assume that user devices report the channel state information (CSI) of their D2D links as well as their remaining energy to the *eNB*, and it is also available for the ProSe server to acquire user devices' location information.

CSI is the channel property of communication links, which describes signal's attenuation factor in every communication link. In particular, CSI is the value of every element in channel gain matrices, such as the information about signal scattering, environment fading (e.g., multipath and shadowing fading), and power decay of distance. In general, CSI is sent to the sending side after the receiving side evaluates and quantifies CSI, and we

call the quantified value as channel attenuation coefficient (CAC). Literature [28] describes how to acquire full CSI and get CAC.

3.3 Capacity of direct and relay-assisted transmission

The receiving UEs use three types of communication modes to receive data from the *eNB*: (1) the receiving UEs directly communicate with the *eNB*, which refers to the one-hop cellular communication, (2) the receiving UEs use one relay to assist themselves to receive data from the *eNB* through two-hop transmission, and the transmission path is composed of one cellular link and one D2D link, (3) similarly, two relays are utilized in a three-hop transmission path composed by one cellular link and two D2D links to forward data for the receiving UEs, which makes it possible for the *eNB* to extend its cellular coverage.

According to the Shannon capacity formula, when user *i* directly receives its data from the *eNB*, the throughput of user *i* could be expressed as follows.

$$T_{\text{cell}}^i = b_{\text{cell}}^i \cdot \log_2(1 + \gamma_{\text{cell}}^i) \quad (1)$$

In (1), T_{cell}^i and γ_{cell}^i are the throughput and the signal to interference noise ratio (SINR) for user *i* respectively when the *eNB* transmits data to it over a cellular channel, and b_{cell}^i is the allocated bandwidth for user *i*. User *j* receives data for user *i* over a cellular channel b_{cell}^i , where the throughput (i.e., that from the *eNB* to user *j*) T_{cell}^{ji} is computed by the following formula.

$$T_{\text{cell}}^{ji} = b_{\text{cell}}^i \cdot \log_2(1 + \gamma_{\text{cell}}^j) \quad (2)$$

γ_{cell}^j appearing in (2) is the SINR for user *j* when the *eNB* transmits data to it over a cellular channel b_{cell}^i .

User j relays data for user i over a WiFi channel b_{wifi}^{ji} ; hence, the throughput (i.e., that from user j to user i) T_{wifi}^{ji} could be computed by the following formula.

$$T_{\text{wifi}}^{ji} = b_{\text{wifi}}^{ji} \cdot \log_2(1 + \gamma_{\text{wifi}}^{ji}) \quad (3)$$

In (3), $\gamma_{\text{wifi}}^{ji}$ is the SINR for user i when user j transmits data to user i over a WiFi channel, and b_{wifi}^{ji} is the available WiFi bandwidth to user i from user j . Based on (2) and (3), the throughput of user i that receives its traffic from a relaying user j is expressed as follows.

$$T_{\text{d2d}}^{ji} = \min(T_{\text{cell}}^{ji}, T_{\text{wifi}}^{ji}) \quad (4)$$

T_{d2d}^{ji} showing up in formula (4) is the throughput of user i when the eNB transmits data to it via relay j , and its value depends on the bottleneck of throughput over two communication links. Thus, in a similar way, if user i receives its traffic from the eNB via relay k and relay j in turn, its throughput (denoted as T_{d2d}^{kji}) is computed as follows.

$$T_{\text{d2d}}^{kji} = \min(T_{\text{cell}}^{ki}, T_{\text{wifi}}^{kj}, T_{\text{wifi}}^{ji}) \quad (5)$$

The eNB uses the transmission power p_c to transmit data to user i directly over a cellular channel, and the SINR for user i is estimated by the following formula.

$$\gamma_{\text{cell}}^i = \frac{g_i \cdot p_c}{N_i + F_{i,\text{cell}}} \quad (6)$$

In (6), N_i is the noise power perceived by the user i ; $F_{i,\text{cell}}$ is the interference power perceived by the user i over a cellular channel, which could be ignored if a single cellular network is just considered or the co-channel interference of neighboring cells can be controlled effectively when a cell is considered in multi-cell network. g_i is the channel attenuation coefficient of the link from eNB to the user i , which includes the path loss, multipath fading, and shadowing fading effects; as we say before, it is usually evaluated and quantified by the receiver and then sent to the eNB .

When user j relays data to user i by using the transmission power p_{ji} (set as a fixed value p_{ue} in this paper) over a WiFi channel, the SINR for the user i is estimated by the following formula.

$$\gamma_{\text{wifi}}^{ji} = \frac{g_{ji} \cdot p_{ji}}{N_i + F_{i,\text{wifi}}} \quad (7)$$

In (7), g_{ji} is the channel attenuation coefficient of the link from the user j to i , which includes the same factors as that of g_i ; $F_{i,\text{wifi}}$ is the interference power perceived by the user i over a WiFi channel, which mainly caused by

its surrounding potential transmitter using the same or overlapped frequency band, and it is estimated by the following formula.

$$F_{i,\text{wifi}} = \sum_{k \in I_i} g_{ki} \cdot p_k \quad (8)$$

In (8), g_{ki} is the channel attenuation coefficient of the link from the interfering user k to the interfered user i , which includes the same factors as that of g_i ; p_k is the transmission power of the interfering user k ; I_i is the set of the interfering users of user i .

When user j act as a relay (if $j = 0$, there is no relay), the delay over the air from eNB to the user i could be denoted as d^{ji} . Also, the over air delay from eNB to the user i is denoted by d^{kji} when k and j act as its relays (if $k = 0$ and $j = 0$, there is no relay). However, in order to balance the throughput and lifetime of the transmission path, we simultaneously select two potential relays just according to the throughput improvement in the pre-selection process of the first relay (i.e., user j), which lead to the optimal and sub-optimal improvement for the downlink throughput of user i respectively. Therefore, we call them optimal relay o and sub-optimal relay s . Based on the definitions we have listed above and the constraint thresholds, we could model our focused problem as a downlink throughput optimization problem for M users as follows:

$$\begin{aligned} \text{Max} \sum_{i \in \mathcal{R} \setminus \{0\}} \sum_{j \in \mathcal{R} \setminus \{i\}} (\alpha_{ij} T_{\text{cell}}^i \\ + \alpha_{ji} \left(\max(T_{\text{d2d}}^{ji}, \alpha_{kji} T_{\text{d2d}}^{kji}) \right)) \end{aligned} \quad (9)$$

$$\text{s.t. } \alpha_{ij} \in \{0, 1\}, \forall i, j \in \{0, 1, \dots, N\}$$

$$\alpha_{ii} = 0, i \neq 0$$

$$\alpha_{00} = 1$$

$$\sum_{i \in \mathcal{R}} \alpha_{ij} = 1$$

$$\sum_{j \in \mathcal{R}} \alpha_{ij} - (N-1 + 2\delta_{0i}) \alpha_{0i} \leq 0$$

$$\alpha_{kji} \in \{0, 1\}, \forall i, j, k \in \{0, 1, \dots, N\}$$

$$\left(\sum_{k \in \{1, \dots, N\} \setminus \{j, i\}} \alpha_{kji} \right) \in \{0, 1\}, \text{ if } \alpha_{ji} = 1$$

$$d^{ji} \leq d_{th}; d^{kji} \leq d_{th}$$

$$j = \begin{cases} o & RIE_i \geq RIE_{th} \\ s & RIE_i < RIE_{th} \end{cases}; e^k \geq \min(e_j, e_{th})$$

In (9), α_{ij} and α_{kji} are binary decision variables, the value of α_{ij} equals 1 which means that user i transmits to user j , and there is no communication link between user i and user j if α_{ij} equals 0. Also α_{kji} determines

whether user k relays the data of user j to user i or not; δ_{ij} equals 1 if $i=j$ and 0 on the contrary. To facilitate the description, the eNB is marked as user 0. The variables in the first constraint and the sixth constraint are enforced to be a binary value. In the second and third constraint, the situation that users transmit data to themselves are forbidden except for the eNB , and this exception does not mean the eNB transmits data to itself, it is just for notational convenience. The users could receive data either from a relay or the eNB under the fourth constraint. And the number of relays of each receiver is confined within $N-1$. The seventh constraint limits that a user can select one relay at most when itself acts as a relay already. The eighth constraint keeps the delay of transmission path that consists of one relay (i.e., j) or two relays (i.e., k and j) below a given threshold value.

In addition, the ninth constraint is to preselect the first relay between candidates (i.e., UE o and UE s) based on RIE_i and the threshold value (e.g., RIE_{th}) for the RIE metric. Moreover, it also restricts that the residual energy for relay k must be not less than the minimum value between the residual energy of relay j (e.g., e^j) and the energy constraint (e.g., e_{th}). That is, in order to ensure the preselection scope, e^k is only considered to be not less than e_{th} when e^j is more than e_{th} . And when e^j is not more than e_{th} , as long as e^k is not less than e^j , the relay k will not become the energy bottleneck of the transmission path. For another, the RIE metric weighs the superiority degree of being a relay of relaying UE o over relaying UE s , and we will detail the meaning of RIE in the following text.

In the scenarios that the paper focuses on, it is relatively easy to determine the transmission power of reaching the maximum throughput for a cellular channel, if the desired value for the receiving bit error ratio (BER) and environment noise power in the receiving side, allocated bandwidth, and channel attenuation coefficient (based on the CSI's quantification result) are available for us. Therefore, we could estimate transmission power by using formulas (1) and (6). If the evaluated transmission power is not beyond the maximum transmission power of transmitter, the evaluated transmission power will be adopted. Otherwise, if the energy efficiency is not to be considered, the maximum transmission power will be adopted to send data, which could simplify our scheme. But for a WiFi channel, determining the transmission power of reaching the maximum throughput might be more difficult compared with a cellular channel. This is because the uncertain interferences of WiFi co-channels exist, such as the frequently used 2.4 GHz frequency band, in which there are only three non-overlapping channels. Thus, the co-channel interference is inevitable when a large number of WiFi channels are used for D2D communications, which makes

each transmitter adjust the transmission power with the concern of the co-channel interferences in its neighboring area. However, the adjusted transmission power becomes the interference source of other WiFi links. Therefore, the solution to this optimization problem belongs to a NP problem. In order to make our scheme less complicated so that the most attention could be paid to 'What effects multi-hop outband D2D communications will have on the trend of cellular downlink throughput,' we set that each transmitter of a WiFi link (i.e., relaying UE) adopts a uniform transmission power. The interference source of the receiving UE (e.g., UE i) mainly consists of the transmitters of co-channel communication links in UE i 's neighboring area, and the transmitters in this area constitute its neighboring interference set (i.e., $I_{i,nei}$). For example, in Fig. 2, the set $I_{5,nei}$ of UE5 includes UE1, UE3, UE4, UE8, and UE9. When UE4 relays data for UE5, UE4 does not cause any interference to UE5, but UE1, UE3, UE8, and UE9 are potential interference sources of UE5, and they cause interference to UE5 only when they use the same or overlapped spectrum as the UE5's receiving spectrum. Making statistics about the actual co-channel interference sources of WiFi links is relatively complicated and difficult. Thus, in this paper, we adopt the strategy of relay preselection and allocate a WiFi channel randomly as a relay channel. Then, each receiving UE estimates the throughput gain after the assistance of the preselected relays. If the throughput is worse than that in a direct communication mode, this receiving UE will abandon both the preselected relays and the allocated WiFi channels. This behavior reduces interference to other receiving UEs. Therefore, the receiving UEs that estimate the throughput gain afterwards are more likely to retain the preselected relays. At some degree, this relay preselection strategy also takes the advantage of reducing the statistical difficulty caused by the actual interference source except the advantages mentioned before.

3.4 Approximate coverage under the given BER

Each wireless receiver has a certain requirement on its BER. As for the receiving UEs in our scheme, if the BER of the cellular link from the eNB to any of the receiving UEs is higher than the fixed value (set it as BE_{th}), a relay will be needed to forward data for the receiving UE.

From the formula for estimating frame success rate (i.e., $f(\gamma_j) = (1 - e^{-0.5\gamma_j})^M$, where M denotes the length of frame with respect to a given length of packet, and γ_j is the SINR value of receiving-end j [29]), we know that $e^{-0.5\gamma_j}$ is the BER value of receiving-end j . Therefore, when given the value for BE_{th} , the corresponding SINR (i.e., γ_{th}) could be estimated by the formula in (10), which is deduced from the formula $BE_{th} = e^{-0.5\gamma_{th}}$.

$$\gamma_{th} = -2 \ln BE_{th} \tag{10}$$

Then, the relay will be determined in receiving UEs' neighboring scope under the delay constraints and the hybrid constraints of energy and throughput. For any receiving UE i , how broad its neighboring scope should be, in which the UEs send data to it in a given transmission power (e.g., p_{ue}), to ensure its BER within BE_{th} ? For simplicity and without loss of generality, we could estimate the receiving UE i 's approximate coverage D_i according to the following formula (11).

$$D_i = \begin{cases} \sqrt{\frac{p_{ue} \cdot G_t \cdot G_r \cdot \lambda^2}{(4\pi)^2 \cdot L \cdot \gamma_{th} \cdot (N_i + F_{i,wifi})}} & D_i < d_{crossover} \\ \sqrt[4]{\frac{p_{ue} \cdot G_t \cdot G_r \cdot h_t^2 \cdot h_r^2}{\gamma_{th} \cdot (N_i + F_{i,wifi})}} & D_i \geq d_{crossover} \end{cases} \tag{11}$$

In (11), G_t and G_r are the transmitting and receiving antenna's gains respectively; h_t and h_r are denoted as the height of the transmitting and receiving antenna on the ground respectively; λ is the wavelength of the signal carrier; and L is the system loss coefficient which is not related to propagation; $d_{crossover}$ is the crossover distance, which is computed by formula (12) as described in [30].

$$d_{crossover} = \frac{4\pi\sqrt{Lh_t h_r}}{\lambda} \tag{12}$$

The approximate coverage of UE i is estimated by formula (11). Due to the number of actual interference sources and their location distributions, transmission powers and other information about interference estimation may not be available for UE i ; thus, $F_{i,wifi}$ is an uncertain value. Sometimes, it could not be guaranteed that the receiving UE i 's BER is below the given threshold level when the UE near the edge of cellular coverage sends data to UE i . Multiplying the result in (11) by a discount coefficient η improves the probability of a lower BER of UE i that is not more than the given threshold BER, which potentially improves the throughput of UE i to avoid becoming the throughput bottleneck of the transmission path, and thus more preselected relays could be retained. The approximate coverage after discounting is as follows.

$$D_i = \eta \cdot D_i \tag{13}$$

In (13), $0 < \eta \leq 1$, and all UEs within this approximate coverage of UE i are denoted as a set (e.g., V_i).

3.5 Energy constraints of relays

In our scheme, for any receiving UE (e.g., UE i), different energy constraints are adopted in the selection process of its first relaying UE and second relaying UE. As mentioned in the previous text, in the preselection process of the first relaying UE, we have determined two candidate relays, that is, optimal relay o and sub-optimal relay s , where their remaining energy values are denoted as e_{oi} and e_{si} . In order to conveniently weigh these two relays' performance in terms of link throughput and link lifetime, we propose formula (14) as follows.

$$RIE_i = \begin{cases} \infty & e_{oi} \geq e_{si} \\ \frac{T_{cell}^{oi} - T_{cell}^{si}}{T_{cell}^{oi}} \cdot \frac{e_{oi}}{e_{si} - e_{oi}} & e_{oi} < e_{si} \end{cases} \tag{14}$$

In (14), RIE_i denotes the RIE value for UE i , which measures superiority degree of being a relay of relaying UE o over relaying UE s . The higher value of RIE_i is, the greater superiority relaying UE o has over relaying UE s . If e_{oi} is not less than e_{si} , it indicates that not only the relaying UE o has advantage in throughput improvement but also in remaining energy when compared with relaying UE s . Thus, in this case, we set RIE_i as ∞ to show that relaying UE o has more advantage in both improving link throughput and prolonging link lifetime when compared with relaying UE s . Therefore, UE i chooses relaying UE o to relay data for itself. On the contrary, when e_{oi} is less than e_{si} , it indicates that relay o has an advantage in throughput improvement but it is located in an inferior position in extending link lifetime. Therefore, if RIE_i is less than the RIE threshold we set before, user i chooses relay s instead of relay o to transmit data.

After a receiving UE (e.g., i) has determined its first relay, and if there is still a need (i.e., the BER from the eNB to its first relay is still higher than the given BER value) for UE i to choose the second relay, we adopt direct energy constraints and delay constraints to select the second relay. That is to say, the remaining energy of the selected second relay must not be less than the minimum value between the first relay's energy and the energy threshold.

In the extensively used protocol of 802.11b/g/n, WiFi interface usually works in the 2.4G frequency band of ISM, which is divided into 14 basic sub-channels and labeled as from 1 to 14. The overlapping coefficient of two neighboring sub-channels is approximately regarded as 80% of one sub-channel frequency band. Therefore, the overlapping coefficient of frequency band decreases by 20% if the difference between any pair of tags increases by 1. When the difference between any pair of tags is not less than 5, the corresponding sub-channel frequency bands do not overlap (e.g., sub-

channels 1, 6, and 11). If there are no more than three WiFi links in a transmission path, non-overlapping WiFi channel could be allocated to every WiFi link. For a certain WiFi link, other WiFi links using the same or overlapped spectrum are its interference sources. Therefore, the corresponding coefficient of overlapping area should be multiplied when calculating interfering power of receiving UE or relaying UEs. Based on the number of actual interference sources, their location distributions and transmission powers, we should adjust transmission power of this WiFi link to ensure link throughput, and the adjusted transmission power becomes the regulation basis of transmission power of other WiFi links, which is a significant way to improve cellular throughput. As we have mentioned above, in this paper, the effects of multi-hop outband D2D communications on cellular throughput is what we mainly focus on. Therefore, the adjustment of transmission power is not considered in this paper.

3.6 Centralized algorithm for DTO-MROD

To solve the problem (i.e., DTO-MROD) described in the “Problem formulation” section, we propose a centralized heuristic algorithm for scene (c) of Fig. 1, which could be divided into the three subroutines (i.e., the preselection for the first relay, the preselection for the second relay, and the relay verification, which are executed at the *eNB* one after another). The corresponding pseudocodes are given in Algorithm 1~3.

Lines 2 to 7 are to determine the receiving UEs which cannot satisfy the given BER when the *eNB* transmits data directly to it. That is to say, these receiving UEs require the help of relaying UEs. Because of the unavailability for calculating the D2D link throughput, the difference values between the link throughput from the *eNB* to the relaying UE and the link throughput from the *eNB* to the receiving UE are used to select potential relays.

In order to prevent the throughput of WiFi links from becoming the throughput bottleneck of transmission path, we adopt the scope computed by formula (13) to limit the selecting scope of potential relays. First of all, we select the two optimal potential relays (lines 9 to 17), then estimate the value of RIE_i by formula (14) and compare with the given RIE_{th} . Finally, the optimal relaying UE or sub-optimal relaying UE will be chosen to act as the first relay according to the evaluated value (lines 18 to 20).

It is noted that c_{ji} is the allocated WiFi channel's sequence number of the WiFi link from UE j to UE i , but the value that equals 0 means that WiFi channel has not been allocated to this WiFi link. If scene (c) in Fig. 1 is replaced with scene (b) in Fig. 1, one constraint (i.e., UEs that both have not requested for data transmitting

service and have not acted as a relay could be preselected as relays) should be added to line 10.

At the phase of the second relay preselection, the purpose of lines 2 to 7 is to calculate the metric of selecting the second relaying UE. This metric is the difference value between the link throughput from the *eNB* to the second potential relay and the link throughput from the *eNB* to the first potential relay. The purpose of lines 8 to 15 is to preselect the second potential relay which has the optimal metric and satisfies the delay and energy constraints.

At last, due to the uncertainty of co-channel interference in the process of preselecting relays, the throughput gains need to be rechecked. Thus, lines 2 to 27 are to check whether the selected potential relays are able to improve throughput or not after all potential relays and WiFi channels are determined. If the relay can meet the above requirement, no action will be done. Otherwise, the flag of relay will be deleted (i.e., set α_{ji} and c_{ji} to 0 or set α_{kji} , c_{kj} and c_{ji} to 0), and the interference will also be recalculated. The *Abs* (in lines 7, 12, and 21) is the function that solves the absolute value of the argument, and the polynomial value in the bracket means overlapping coefficient for frequency band.

Only one UE with the optimal throughput improvement is selected as the receiving UE's relay in the existing algorithm DORE in [10]. If this relaying UE is infeasible under the delay constraints, the receiving UE will give up the relay selection and receive data directly from the *eNB*. Thus, the whole process shows typical greed features. But in our scheme, if the optimal relay is infeasible, the sub-optimal relay will be considered until a relay is able to improve throughput or all candidates are compared.

3.7 DTO-MROD procedures ProSe-compliant framework

The integration of DTO-MROD in 3GPP Proximity-based Services (ProSe) is elaborated in this subsection. The access network and main ProSe elements (i.e., ProSe Function and ProSe Application Server) mentioned in [31] are depicted in Fig. 3. As shown in Fig. 3, the framework includes the process of registration, collection, decision, activation, communication, and termination.

The first step for cellular D2D communications is registration which is essential for D2D users because the *eNB* needs the whole information in centralized cellular network. As we can see from Fig. 3, UE_{*j*} sends registration request to the ProSe Function, and then, this message is handled in the ProSe Application Server. Meanwhile the confirmation and setting message consist of an application ID that is used to make requests, the information of location updating period, discovery beacon, and discovery channel are returned to UE_{*j*}. After a registration, the necessary information (e.g., location,

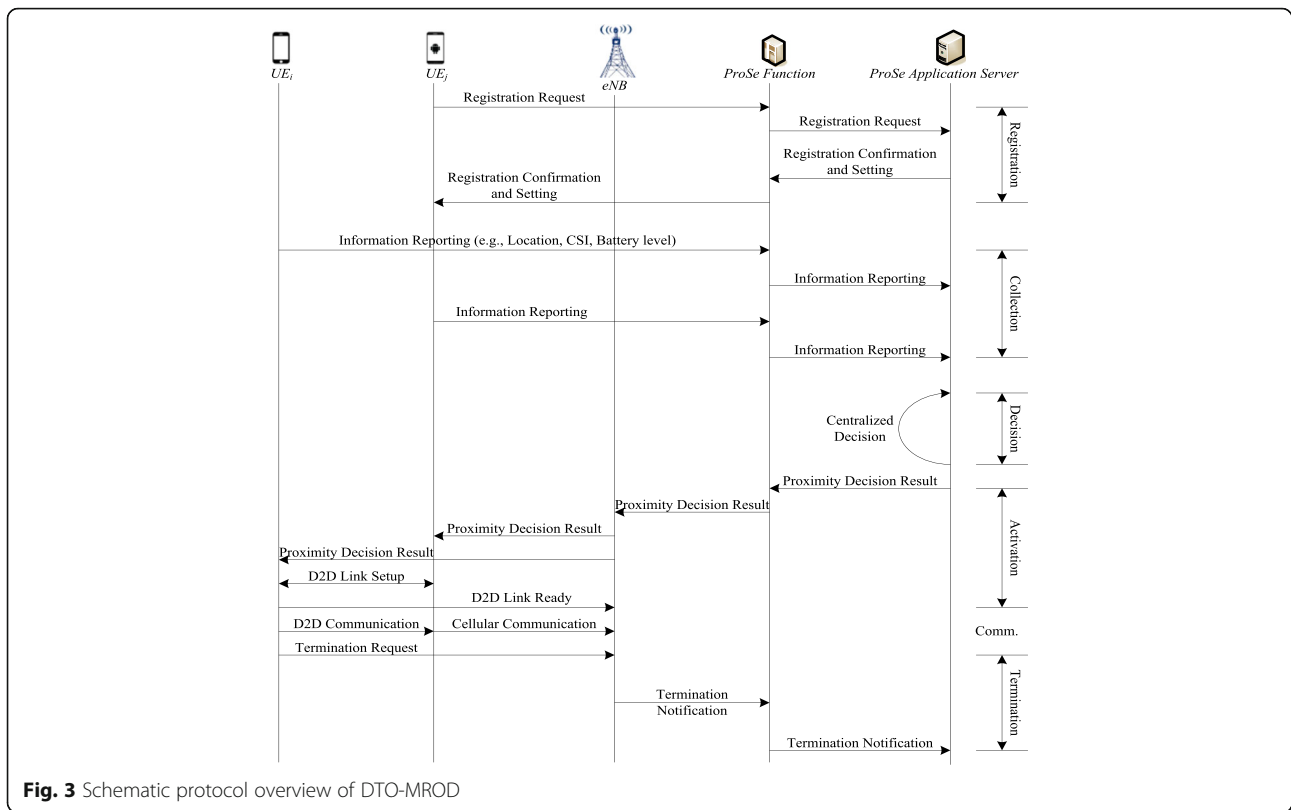


Fig. 3 Schematic protocol overview of DTO-MROD

channel state information (CSI), battery level) must be collected before the stage of decision. A decision is made by running the Algorithm 1~3 at the ProSe Application Server. Based on the results of decision, a proximity decision result message is sent to the ProSe Function from ProSe Application Server. Then, after receiving the proximity decision result message from the ProSe Function, the eNB broadcasts it to all UEs in turn.

The UEs activate the D2D links according to certain procedures. Once establishing the connection with the eNB, the UEs will start to send a Link Ready message to the eNB. Then, the eNB begins to serve the D2D UEs when it takes notice of the activation of D2D links. In order to ensure that the frame (should be received or be relayed) could be identified by every UE, the eNB labels the frame of D2D UEs. Once the relay frames are received, the relaying UEs will handle the packet from the physical layer to the Packet Data Convergence Protocol (PDCP) layer, and then, the PDCP Service Data Unit (SDU) will be encapsulated into a WiFi frame that will be sent via the D2D link. After the receiving UEs decapsulate the relayed frames, the PDCP SDU will be handled by the receiving UEs via the remaining LTE stack.

The eNB terminates the process of relaying when it receives a termination request message from a D2D UE (depicted in Fig. 3). Then, a Termination Notification message will be sent to the ProSe Function to notify that

this communication has been terminated and this message will be forwarded to the ProSe Application Server in the same purpose.

3.8 Theoretical analysis

We explore the communication and computational complexity of the DTO-MROD algorithm in this subsection. The maximum number of UEs in the network is denoted as $|\mathcal{R}|$, and the number of elements in any set V_j is denoted as $|V_j|$. We first give the theorem with respect to computational complexity of the DTO-MROD as follows.

3.8.0.1 Theorem 1 *The computational complexity of the DTO-MROD algorithm executed by a node on the infrastructure is $O(|\mathcal{R}| \cdot \max\{|I_{i,nei}| \cdot |V_j|\} \mid i \in \mathcal{R}, j \in I_{i,nei}\})$, where $I_{i,nei} \subset V_i \subset \mathcal{R} = \{0, 1, \dots, N\}$, $V_j \subset \mathcal{R} = \{0, 1, \dots, N\}$.*

Proof For a node on the infrastructure (e.g., the ProSe Application Server), determining the metrics of the first relay selection (line 5 in Algorithm 1) takes time $O(\max\{|V_j|\} \mid i \in \mathcal{R})$, where $V_i \subset \mathcal{R} = \{0, 1, \dots, N\}$. Hence, the computational complexity of executing the first part of Algorithm 1 (lines 2 to 7) in the node is $O(|\mathcal{R}| \cdot \max\{|V_i|\} \mid i \in \mathcal{R})$ since the *for* loop (lines 2 to 7 in Algorithm 1) iterates $|\mathcal{R}|$ times. As for the second part of Algorithm 1 (lines 8 to 21), the process of selecting two potential relays (lines 9 to 17 in Algorithm 1) and finally

determining the first relay (lines 18 to 20 in Algorithm 1) require a total time of $O(|\mathcal{R}|)$.

By executing the first (lines 2 to 7) and the second (lines 8 to 15) part of Algorithm 2, we could determine the metric of the second relay selection (line 5 in Algorithm 2) and select the second relay (lines 8 to 15 in Algorithm 2) respectively. And the computational complexity of this two parts is $O(|D| \cdot \max\{|V_j| \mid i \in \mathcal{R}, j \in V_i\})$ and $O(|D|)$, respectively, where D is a set of points which constitutes by the sequence number of the receiving UE and its first relay. Due to UEs that are uniformly distributed, there is no significant difference in the number of UEs between V_i and V_j ; hence $|V_i|$ could be simply regarded as $|V_j|$.

Based on the above analysis, in the loop body (lines 2 to 27 in Algorithm 3), both the interference calculation of the receiving UE (lines 7 and 21 in Algorithm 3) and the first relay (line 12 in Algorithm 3) are iterated $|\mathcal{R}|$ times, which could be completed in $O(|\mathcal{R}| \cdot \max\{|I_{i,nei}| \cdot |V_j| \mid i \in \mathcal{R}, j \in I_{i,nei}\})$ time.

Actually the elements in set D is so small that it could be ignored when compared with $|\mathcal{R}|$; hence, $O(|D| \cdot \max\{|V_j| \mid i \in \mathcal{R}, j \in V_i\})$ is better than $O(|\mathcal{R}| \cdot \max\{|V_j| \mid i \in \mathcal{R}, j \in V_i\})$, while $O(|\mathcal{R}| \cdot \max\{|V_j| \mid i \in \mathcal{R}, j \in V_i\})$ is smaller than $O(|\mathcal{R}| \cdot \max\{|I_{i,nei}| \cdot |V_j| \mid i \in \mathcal{R}, j \in I_{i,nei}\})$ by almost one order of magnitude, and thus, it is the maximum computational complexity of Algorithm 1~3. Therefore, the computational complexity of the DTO-MROD is $O(|\mathcal{R}| \cdot \max\{|I_{i,nei}| \cdot |V_j| \mid i \in \mathcal{R}, j \in I_{i,nei}\})$.

Then, we analyze the communication complexity of the DTO-MROD algorithm and give the following theorem.

3.8.0.2 Theorem 2 *The communication complexity of the DTO-MROD algorithm executed by a node on the infrastructure is $O(|\mathcal{R}|)$, where $\mathcal{R} = \{0, 1, \dots, N\}$.*

Proof The DTO-MROD algorithm could be executed in a node on the infrastructure. For any UE, it must send its information about location, CSI, and battery level in an Information Reporting package to the ProSe Function. Then the package is sent to the ProSe Application Server as an input of the DTO-MROD algorithm. Thus, the communication complexity of each UE is $O(1)$, and the node running DTO-MROD algorithm receives $|\mathcal{R}|$ Information Reporting packages from all UEs as shown in Fig. 3. Therefore, its communication complexity is $O(|\mathcal{R}|)$.

Although the existing scheme (i.e., DORE) has advantages on computational complexity compared with DTO-MROD, DTO-MROD outperforms it in communication complexity. It is obvious that the computational power of smart terminals gains a sustained growth, and thus, the increased computational complexity of our scheme does not have a great impact on UEs' performance. Meanwhile, the usage of spectrum resources is

becoming increasingly tense. Therefore, in general, the reduced communication overhead of our scheme plays a stronger role in performance improvement of future mobile networks when compared with the increased computational overhead.

4 Performance evaluation

4.1 Description of the compared scheme

The most relevant scheme to DTO-MROD is the scheme of DORE mentioned in literature [10]. The DORE has given a definition of link throughput which is evaluated by real devices, and it is regarded as an input for the greedy algorithm discussed in [10]. As discussed in the "Problem formulation" section, the greedy algorithm in DORE needs a repeated iteration to get a convergent value. Moreover, the performance of communication process after relays selection is what our scheme tries to improve, that is, the performance metrics of throughput, network connectivity ratio, and so on are what our scheme should focus on. Meanwhile, the actual measurements could not be done in experiment simulations. Thus, in order to compare DTO-MROD with DORE fairly, we adapt DORE by introducing relay preselection and call it LIKE-DORE. Its pseudocode description for scene (c) of Fig. 1 is given in Algorithm 4, where the corresponding WiFi channels will be allocated when the relays of D2D communications are selected, and also, different distribution of WiFi channels has a different impact on the interference distribution which affects the throughput.

Table 1 presents a comparison of the communication and computational complexities between DTO-MROD and LIKE-DORE. According to the theoretical analysis in the "Energy constraints of relays" section, we know that the computational and communication complexities of DTO-MROD are $O(N \cdot \max\{|I_{i,nei}| \cdot |V_j|\})$ and $O(N)$ respectively.

LIKE-DORE involves a communication complexity of $O(N)$, since the eNB receives N Information Reporting packages from all UEs. The computational complexity of executing the lines 2~6 in Algorithm 4 is $O(N^2)$ since there are double-nested loops which repeat $N + 1$ and N times respectively. Also, the computational complexity of executing the lines 15~26 is $O(N^3)$ due to the above similar case. The other part of Algorithm 4 only has a single loop (lines 7~14), so the corresponding computational complexity is $O(N)$. Therefore, the computational

Table 1 The performance comparison between DTO-MROD and LIKE-DORE

Algorithm	Computational overhead	Communication overhead
DTO-MROD	$O(N \cdot \max\{ I_{i,nei} \cdot V_j \})$	$O(N)$
LIKE-DORE	$O(N^3)$	$O(N)$

complexity of LIKE-DORE is $O(N^3)$. Based on the above analysis, the DTO-MROD and LIKE-DORE have the same communication overhead. However, the DTO-MROD outperforms the LIKE-DORE in terms of the computational complexity since $|I_{i,nei}|$ and $|V_j|$ are far less than N , which are benefiting from the local search strategy for potential relays.

4.2 Simulation setting

The definition of our simulation metrics are as follows: (1) Throughput refers specifically to the data rate, that is, after the allocation of all cellular channels, the data successfully and simultaneously transmitted from the eNB to the UEs that are assigned cellular channels during a given time. (2) Delay refers specifically to the forward capability of the relaying UEs in a transmission path when the eNB transmits high-capacity data flow to any destination UE, and this delay depends on the relaying UEs with the worst forward capability in a transmission path. Hence, if there is no relaying UE, no delay is considered. (3) Network connectivity ratio refers specifically to the ratio between the number of the remaining connected transmission paths and the number of the total transmission paths, where the total transmission paths refer specifically to paths from the eNB to the UEs allocated with cellular channels, and the remaining connected transmission paths refer specifically to part of the total transmission paths, which is still connected after several rounds of data transmission. (4) Average downlink continuous service capability refers specifically to the amount of average downlink data, that is, under the given transmission power for relaying UEs, the amount of average data that the receiving UEs have received from the eNB before each downlink transmission path is disconnected. (5) Average downlink data transmission time refers specifically to the average time of transmitting the given number of packets from the eNB to the receiving UEs during the transmission paths are still connected.

For simplicity but without loss of generality, we only consider distance-dependent path losses but no fading losses in our simulation experiments. According to the energy consumption formula for transmitting and receiving in [30], for a link from UE j to i , Δe_{ji}^t is the total energy consumption when UE j transmits 1 bit and Δe_{ji}^r is the total energy consumption when UE i receives 1 bit, and it could be expressed as follows.

$$\begin{cases} \Delta e_{ji}^t = \phi_{11} + \phi_2 \cdot (d_{ji})^{\alpha_{ji}} \\ \Delta e_{ji}^r = \phi_{12} \end{cases} \quad (15)$$

In (15), d_{ji} is the distance from UE j to i ; ϕ_{11} , ϕ_{12} , and ϕ_2 are transmitter electronics energy, receiver electronics

energy and radio amplifier energy respectively; α_{ji} is path loss exponent from UE j to i . For simplicity, α_{ji} takes 2 when d_{ji} is less than $d_{crossover}$; otherwise, it takes 4.

The simulation network is a circular region with a radius of 500 m, and the eNB is located in the center of the region. In addition, the parameter values employed in our simulation are given in Table 2.

4.3 Simulation results

We use the first group of experiments (from Figs. 4, 5, 6, 7, 8, and 9) to compare DTO-MROD with LIKE-DORE in the performance metrics of average cellular downlink throughput, average delay, network connectivity ratio, average downlink continuous service capability, and average downlink data transmission time. The simulation parameters of Figs. 4, 5, 6, 7, and 8 are as follows: the number of the allocated cellular channels N is assigned to 120, the range of the UEs' number is from 400 to 2200, and each UE obtaining cellular channel should follow the constraints of scenario (b) in Fig. 1 when it selects relays. In DTO-MROD, the discount coefficient η of formula (13) is set as 0.7, and the threshold values of RIE , delay, and energy take 1, 110 ns, 0.1 J respectively. Meanwhile, the three strategies of the verification order of relay preselection are adopted in

Table 2 Simulation parameters

Description	Parameter	Value
Transmitting antenna gain	G_t	1
Receiving antenna gain	G_r	1
Transmitting antenna height	h_t	1 m
Receiving antenna height	h_r	1 m
Transmitting power for UE	p_{ue}	0.1w
Transmitting power for eNB	p_c	10w
BER threshold for UE	BE_{th}	10^{-10}
Environment noise power for UE	N_i	For any UE <i>i</i> , 2×10^{-11}
Carrier signal wavelength	λ	0.1224 m
System loss factor	L	1
Crossover distance	$d_{crossover}$	103 m
Path loss exponent	α_{ji}	For any link $j \rightarrow i$, $\alpha_{ji} = 2$ when $d_{ji} < d_{crossover}$; $\alpha_{ji} = 4$ when $d_{ji} \geq d_{crossover}$
Initial battery capacity	$e_{i,int}$	For any UE <i>i</i> , $e_{i,int}$ is randomly distributed between 0.05 J and 0.2 J
Transmitter electronics energy	ϕ_{11}	26.5 nJ/bit
Receiver electronics energy	ϕ_{12}	59.1 nJ/bit
Radio amplifier energy	ϕ_2	For any link $j \rightarrow i$, 10 pJ/bit/m ² when $\alpha_{ji} = 2$; 0.0013 pJ/bit/m ⁴ when $\alpha_{ji} = 4$
Packet forwarding capacity	t_i^b	For any UE <i>i</i> , t_i^b is 100 ns/bit
Data packet length	M	2500 bit

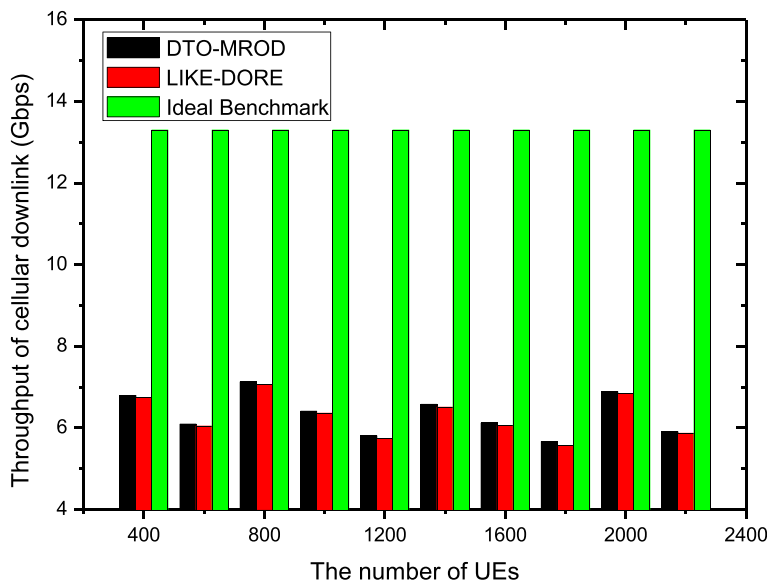


Fig. 4 Throughput of cellular downlink for the two schemes and ideal benchmark under the different numbers of UEs

DTO-MROD, that is, far to near, near to far, and random selection (i.e., check the throughput of receiving UEs according to the order of the distance from the *eNB* to receiving UEs). In this group of experiments, we adopt the strategy of random selection.

As can be seen from Fig. 4, although there is a big gap with the ideal outcome, our DTO-MROD does improve LIKE-DORE. The throughput value of the ideal benchmark solution is got by formulas (1) and (10) based on the assumption that each receiving UE can achieve the BER threshold value (i.e., $BER_{th} = 10^{-10}$). In practice, it is

almost impossible to get such an ideal outcome due to the uncertainty of environmental noise and interference.

Meanwhile, we see that, no matter how the UEs' number M within the cellular network changes, the throughput of DTO-MROD is always better than LIKE-DORE. The main reasons are as follows: if a receiving UE is far away from the *eNB* (i.e., located in the cellular edge), not only the receiving UE has an excessively high BER when *eNB* transmits data directly to it, but also the BER may not be less than the given maximum value with a great possibility in spite of using one relay. Thus, we know

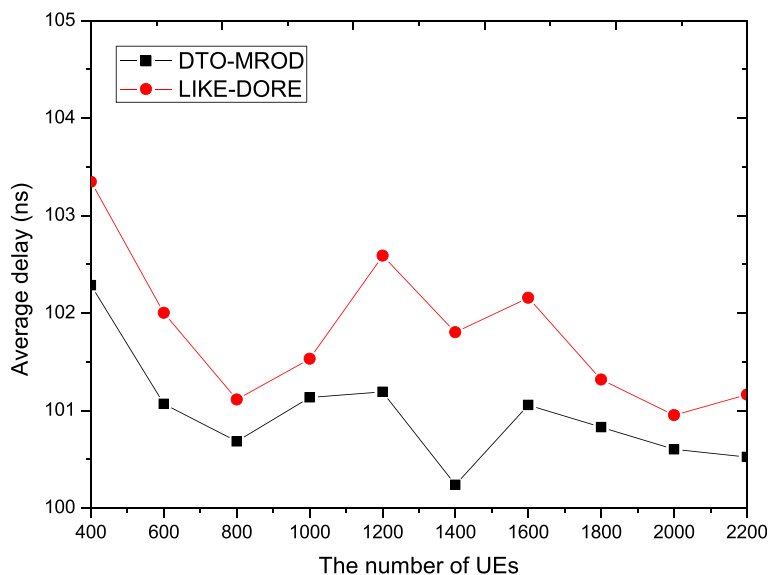


Fig. 5 Average delay for the two schemes under the different numbers of UEs

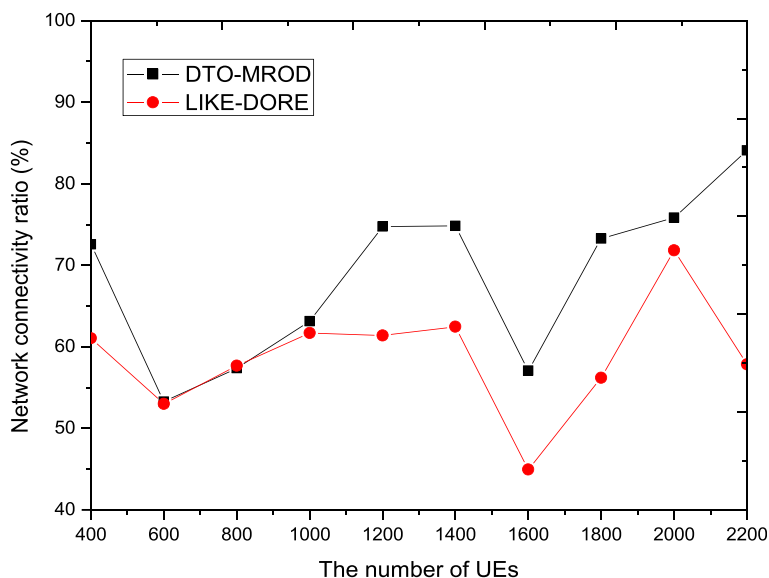


Fig. 6 Network connectivity ratio for the two schemes under the different numbers of UEs

that the SINR of the receiving UE is relatively small based on formula (10). Moreover, from formulas (1)~(4), the throughput of the transmission link will be lower. LIKE-DORE allows each receiving UE to select one relay at most, but two relays are allowed to forward data at most in DTO-MROD. Therefore, if the above-mentioned situation appears in DTO-MROD, by shortening the transmission distance, the BER becomes lower and then the cellular throughput is improved.

Another phenomenon we should notice is that the cellular throughput fluctuates randomly when the number of UEs (i.e., M) changes, this is because of the random variation of the distribution relationship between

the UEs obtaining cellular channels and those that are not. Different distribution relationship leads to different potential candidates of each receiving UE and mutual interference (i.e., the co-channel interference of WiFi links); thus, there is a fluctuation in cellular throughput.

We could see from Fig. 5 that the average delay of DTO-MROD has advantage in that of LIKE-DORE no matter how the number of UEs (i.e., M) within a cellular network changes, which is familiar with Fig. 4. According to the definition of delay metric in this paper, delay of any communication path depends on the relay with the worst forwarding capability over this path, while the forwarding capability of this relay directly influences the

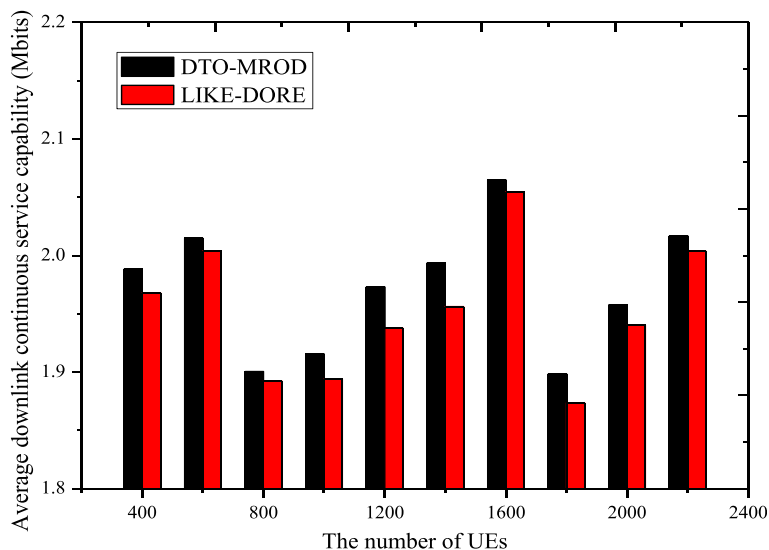
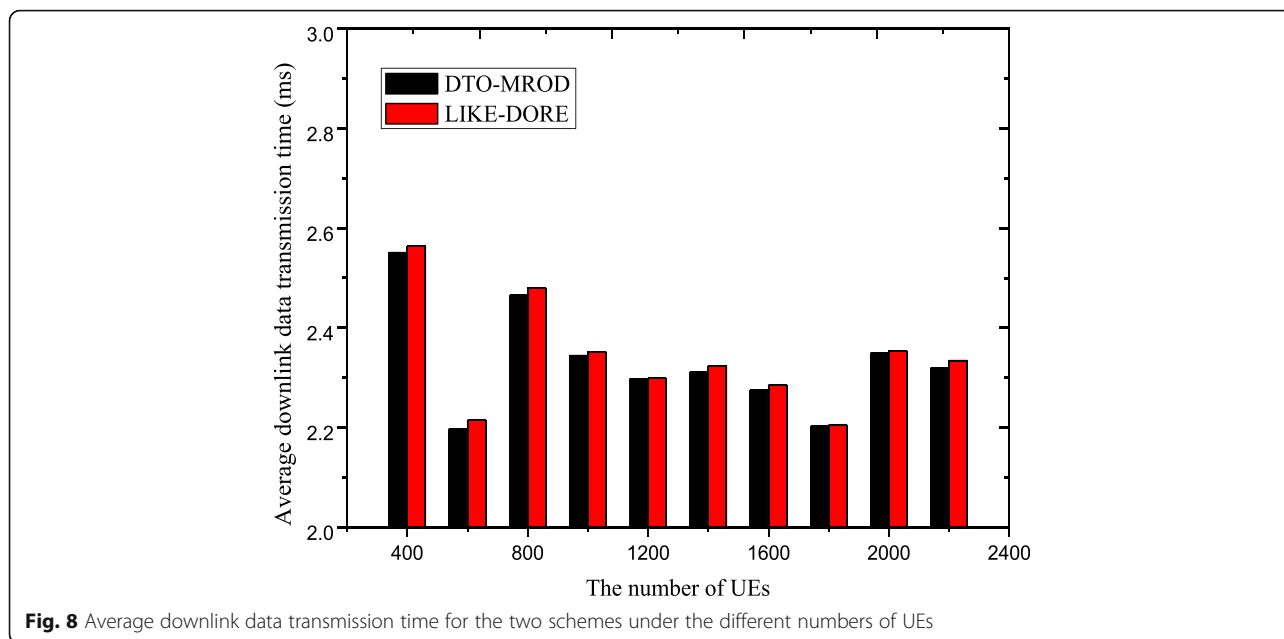


Fig. 7 Average downlink continuous service capability for the two schemes under the different numbers of UEs

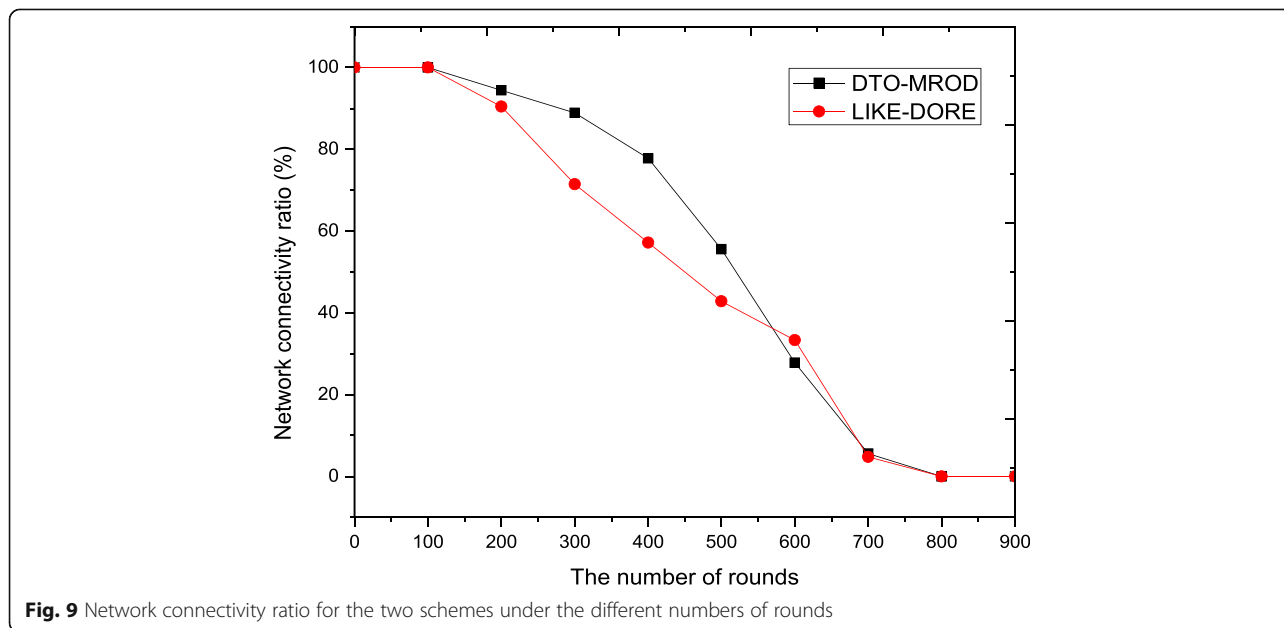


throughput of this path. Thus, the trend of Fig. 5 is similar to that of Fig. 4.

Figure 6 illustrates how the network connectivity ratio of cellular network changes over the number of UEs after 400 rounds' data transfers. We can find that the network connectivity ratio of DTO-MROD is higher than that of LIKE-DORE from Fig. 6. The main reason is that we synthetically consider link throughput and link lifetime in establishing the standard of the first relay selection (see formula (14)). Based on this standard, if another relay is needed in data transfers, we can ensure that the remaining energy of the second relay is not

lower than the minimum value between the first relay's energy and the energy threshold via energy constraint. Therefore, a longer average link lifetime is ensured by these measures based on the relays' remaining energy, while it cannot be guaranteed in LIKE-DORE that does not have energy constraint on its relays.

The fluctuation of network connectivity ratio in Fig. 6 is also caused by the random variation of the distribution relationship, which is familiar with the fluctuation in Fig. 4. But it is worth noting that the network connectivity of LIKE-DORE is slightly better than that of DTO-MROD when the number of UEs is 800. This



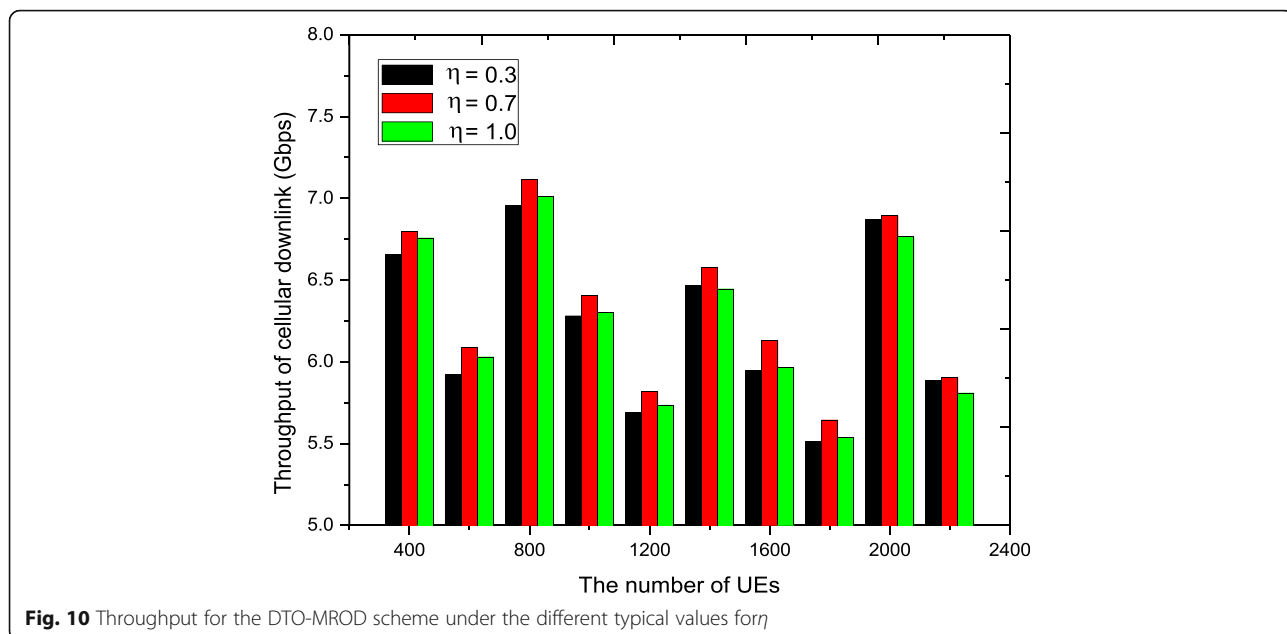


Fig. 10 Throughput for the DTO-MROD scheme under the different typical values for η

unusual phenomenon could attribute to random distribution of UEs' relay coordinates. Furthermore, LIKE-DORE tends to give up relay-assisted transmission if the optimal relay is unfeasible under delay constraint. On the contrary, DTO-MROD has two first relay candidates. Once the optimal relay is abandoned, the sub-optimal relay follows. After that, the second relay will be determined if necessary. When the number of UEs is 800, LIKE-DORE almost gives up all relaying UEs, but DTO-MROD has at least one relaying UE. Thus, a larger disconnected probability of transmission paths occurs in the scheme of DTO-MROD.

Figure 7 illustrates how average downlink continuous service capacity changes under the different number of UEs in a fixed area. From Fig. 7, we find that DTO-MROD outperforms LIKE-DORE in terms of this performance index. We attribute this advantage to two points. Firstly, the energy constraints are considered in DTO-MROD, and thus, the selected relays have higher battery level than those in LIKE-DORE. Secondly, for DTO-MROD, the UEs near the edge of cells trend to select two relays to transfer data, which leads to the shorter transmission links than that in LIKE-DORE. Therefore, the receiving-ends of transmission links have

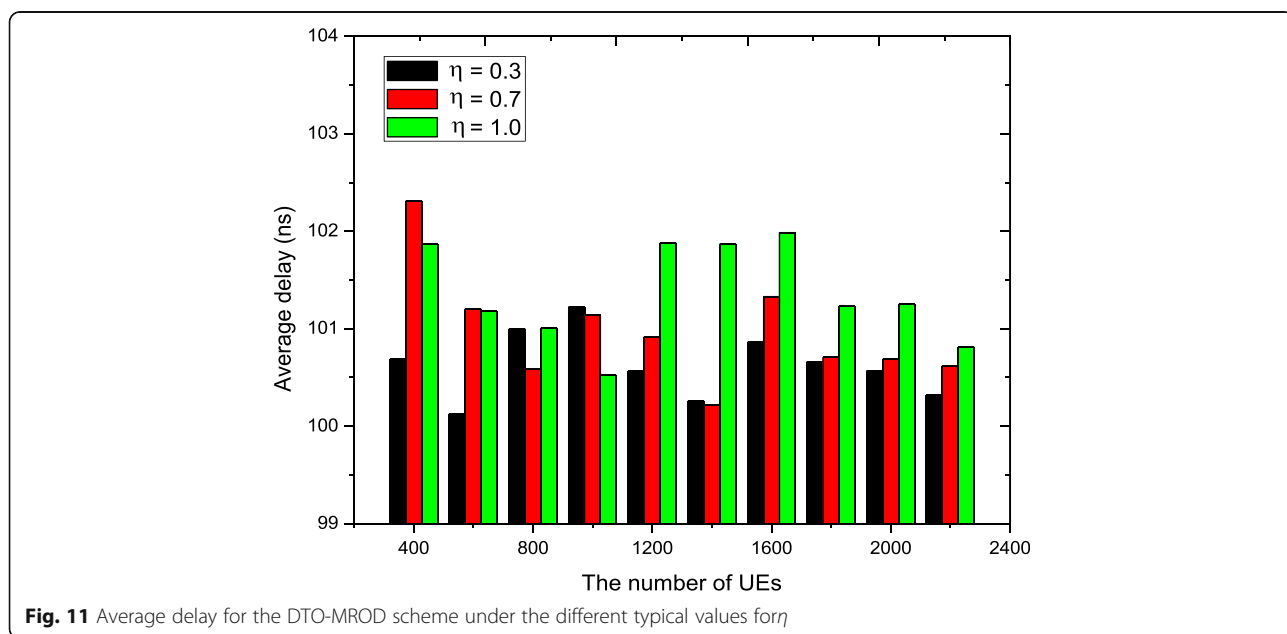
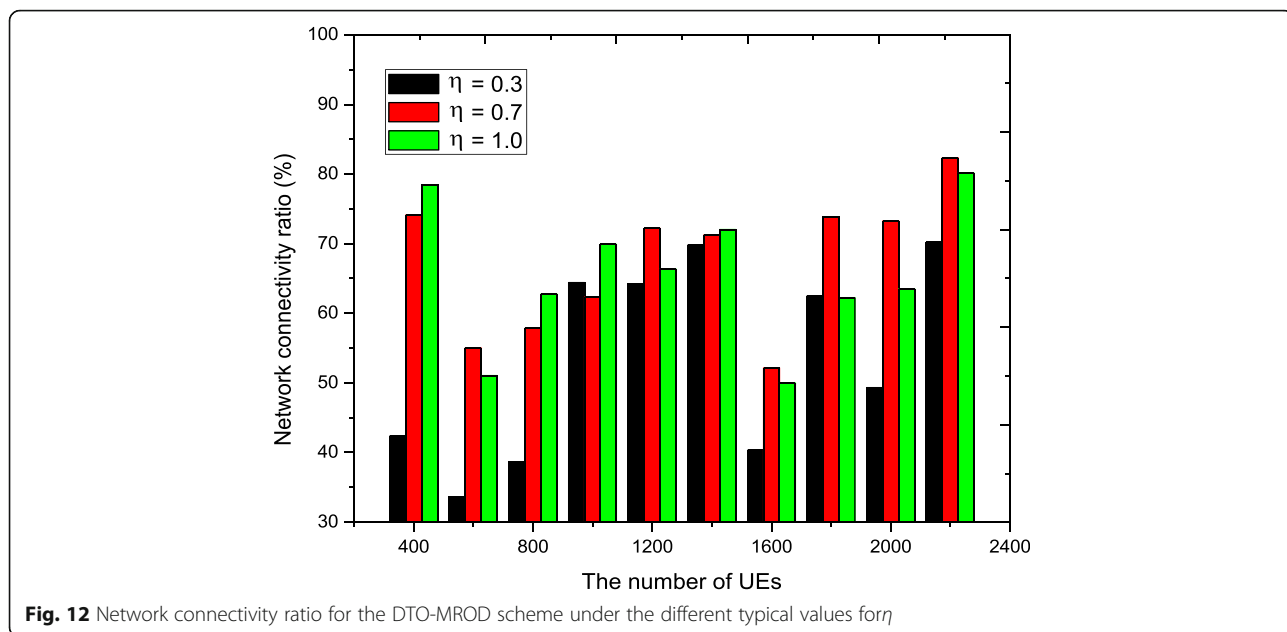


Fig. 11 Average delay for the DTO-MROD scheme under the different typical values for η

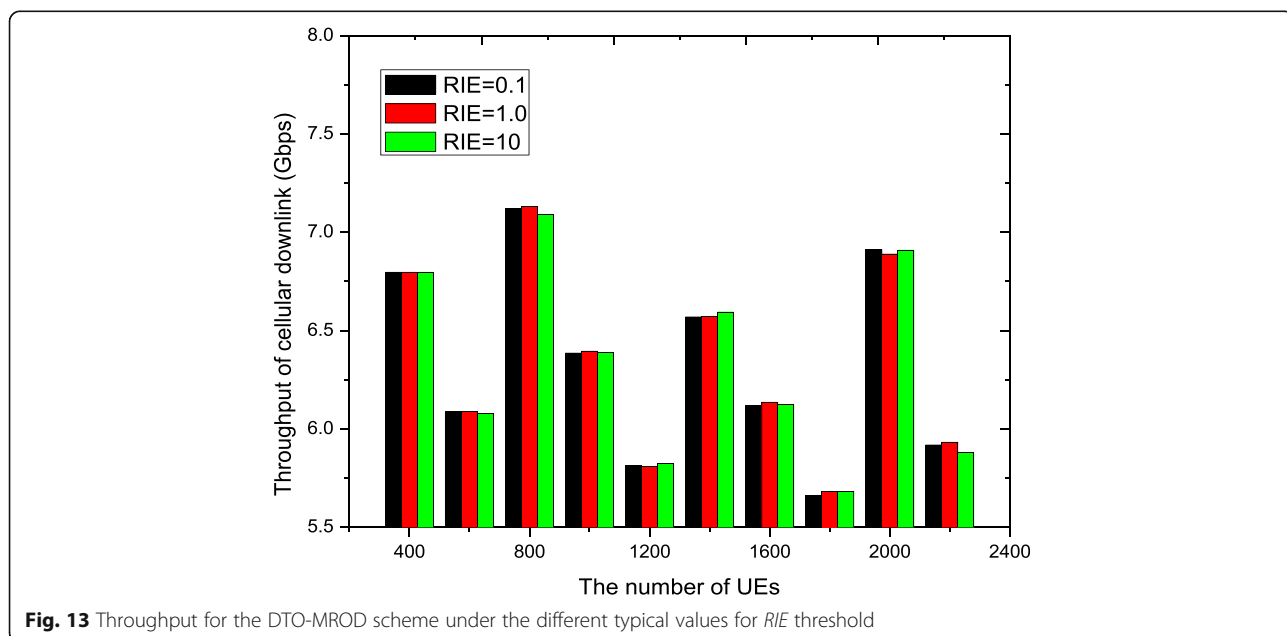


lower BER and thus higher energy efficiency in the process of data transfer.

Figure 8 illustrates how average downlink data transmission time changes under the different number of UEs in a fixed area. From Fig. 8, we see that DTO-MROD has an advantage over LIKE-DORE in terms of average downlink data transmission time. Similar to the second reason mentioned in Fig. 7, the receiving UEs in DTO-MROD have the lower BER due to the shorter transmission links. And according to formulas 1~5 and 10, the average throughput of transmission path in DTO-MROD is larger than that in LIKE-DORE. Therefore, its

average downlink data transmission time is less than that of LIKE-DORE when transmitting a given amount of data.

With the changes of data transfers' round, the variation of network connectivity ratio is depicted in Fig. 9. In every round, the *eNB* only sends a data package with a fixed length to every receiving UE. In this simulation, besides the number of UEs is fixed to 1200, other parameters are the same as Fig. 6. As shown in Fig. 9, the network connectivity ratio of two schemes reaches 100% before the rounds of 100, which means there is no disconnected path in two schemes. From the rounds of



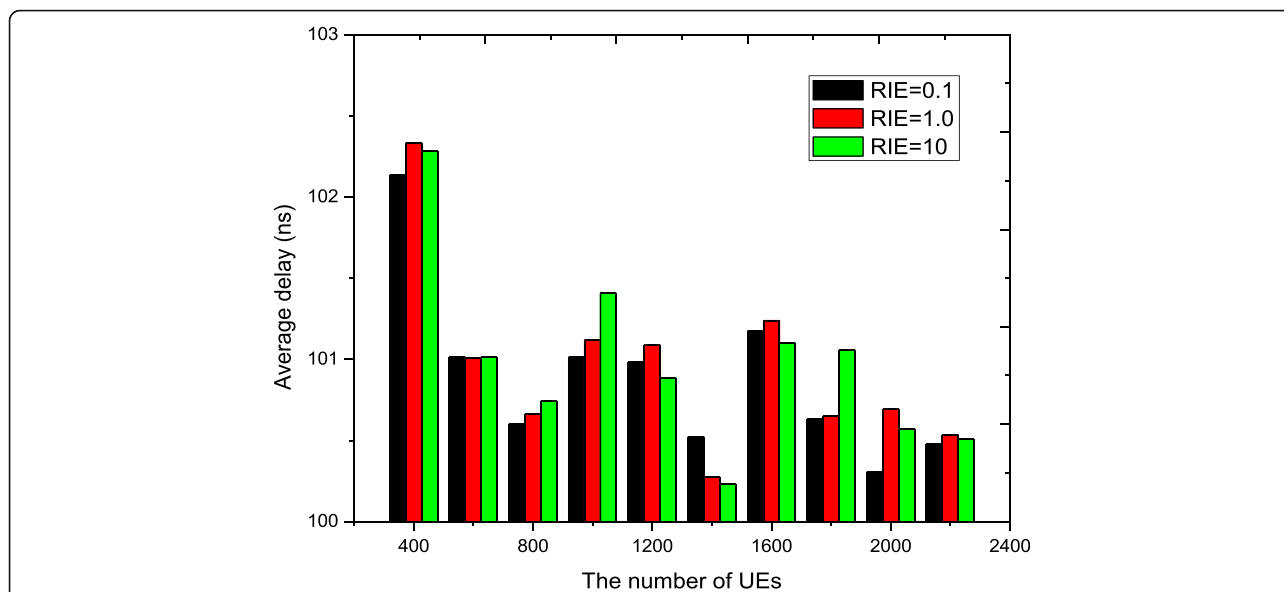


Fig. 14 Average delay for the DTO-MROD scheme under the different typical values for RIE threshold

100 to 560, the network connectivity ratio of DTO-MROD is higher than that of LIKE-DORE, the main reason refers to the explanation of Fig. 6. But from the rounds of 560 to 700, the network connectivity ratio of DTO-MROD is lower than that of LIKE-DORE, the main reason is that some communication paths with two relays appear in DTO-MROD, but it does not appear in LIKE-DORE, and with a large amount of energy consumed, the paths with two relays are more likely to be disconnected. Hence, the network connectivity ratio of DTO-MROD decreases significantly in this scope of rounds. After the rounds of 700, the difference

between two schemes is getting smaller and smaller with the disappearance of all three-hop (i.e., two relays) communication paths.

The second group of experiments is used to compare the performance metrics of cellular downlink throughput, average delay, and network connectivity ratio of DTO-MROD under several typical values of discount coefficient η and RIE threshold. At the same time, we also explore the impact of different verification order of preselected relay on the above performance metrics. In this group (Figs. 10, 11, 12, 13, 14, 15, 16, 17, and 18), besides the changed parameters (e.g., different discount

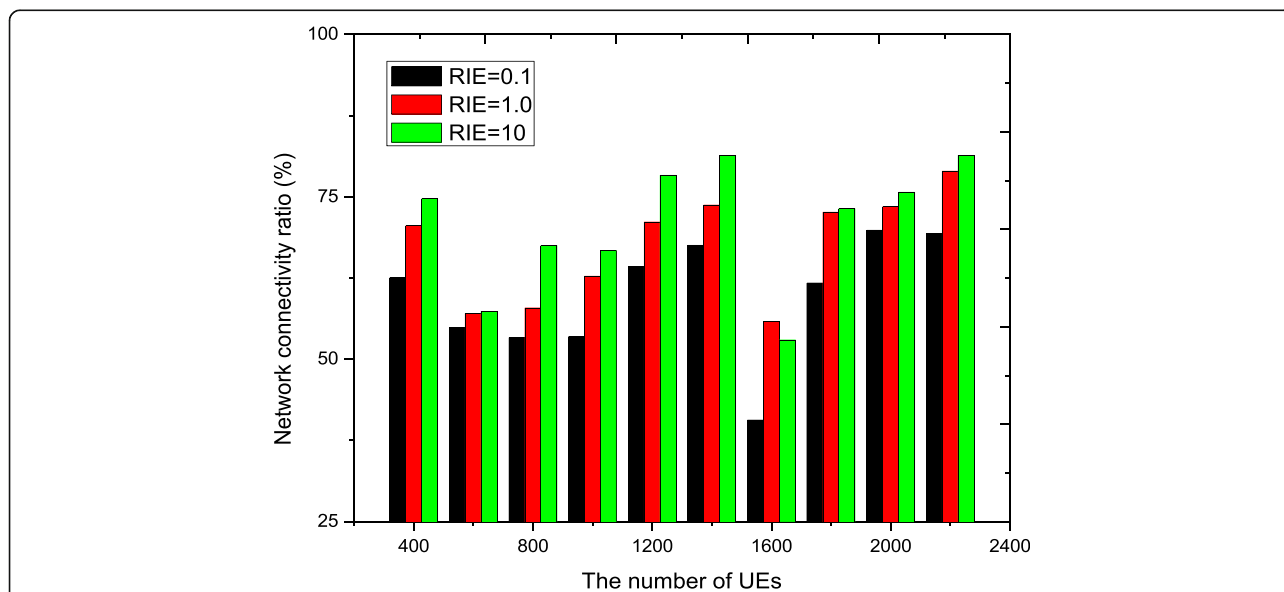


Fig. 15 Network connectivity ratio for the DTO-MROD scheme under the different typical values for RIE threshold

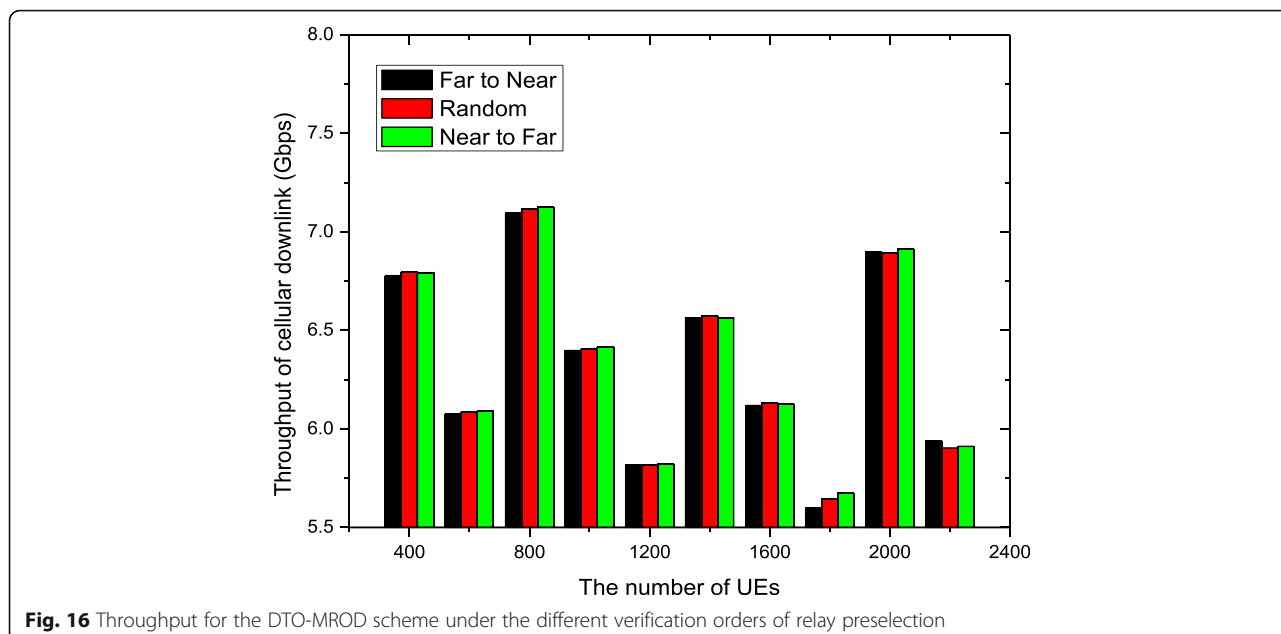


Fig. 16 Throughput for the DTO-MROD scheme under the different verification orders of relay preselection

coefficient, *RIE* threshold, and verification order), other simulation parameters are the same as those of the first group of experiments (Figs. 4, 5, 6, 7, and 8).

Figures 10, 11, and 12 show the impacts of relay selection scope on cellular downlink throughput, average delay, and network connectivity ratio. The value of discount coefficient (i.e., η in Figs. 10, 11, and 12) equaling 0.3 means the relaying distance of WiFi links between the receiving UEs and its relay is relatively short, while η equaling 1.0 means this distance is relatively long. The short relaying distance means that the distance from *eNB* to relaying UEs is relatively long unless the

receiving UEs select the second relay. From Fig. 10, when η equals 0.7, its throughput is relatively high regardless of the changes of UEs' number within cellular area. The main reason is that when η equals 0.3, the cellular links distance is more likely to be longer. However, the relaying distance of WiFi links is more likely to be longer when η equals 1.0. Both two factors become the throughput bottleneck of the whole transmission path; thus, it limits the throughput improvement.

The average delay is relatively short when η equals 0.3 in Fig. 11; this is because the relaying distance of WiFi links is relatively short which is beneficial for data

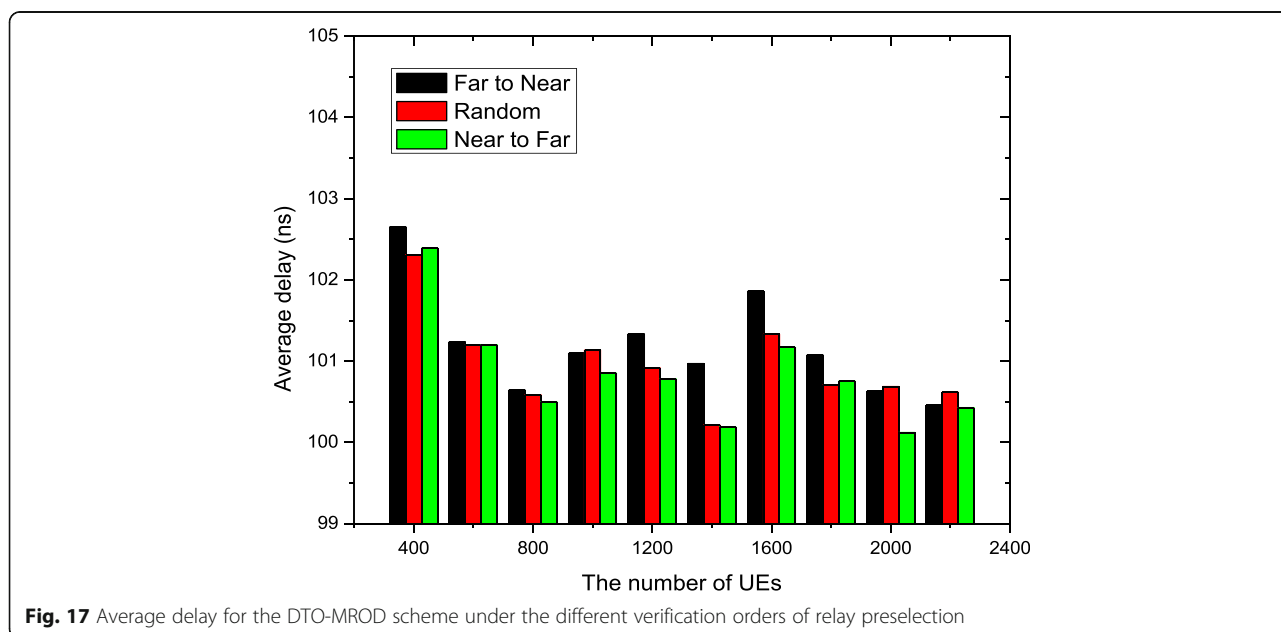
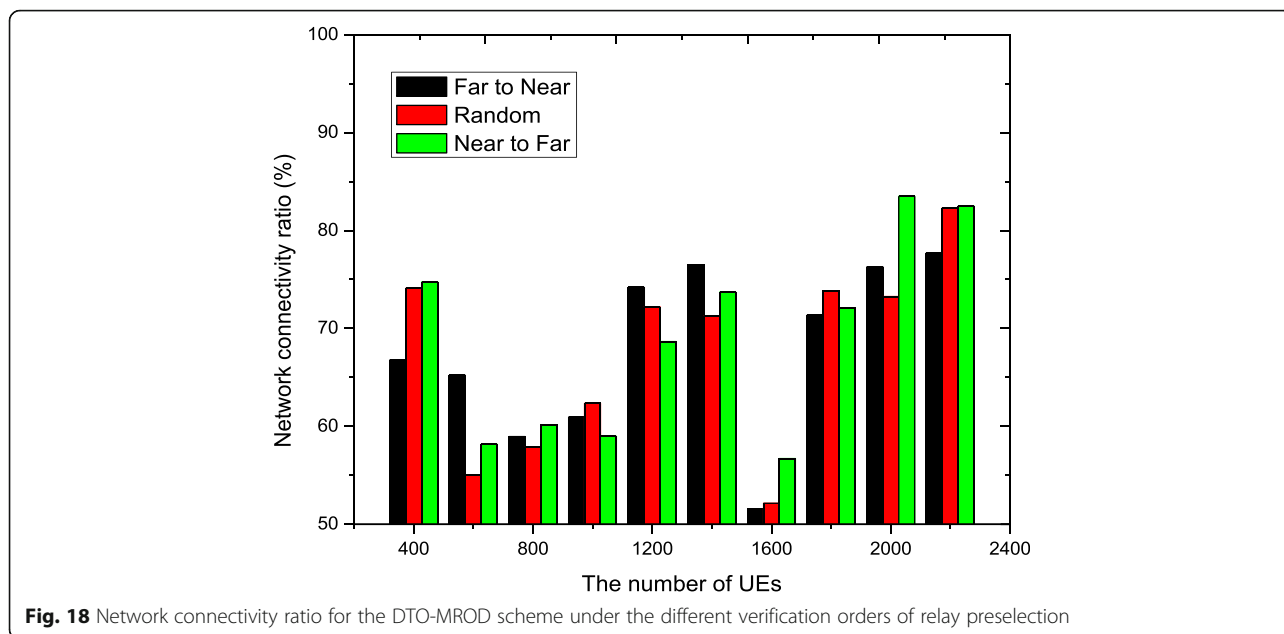
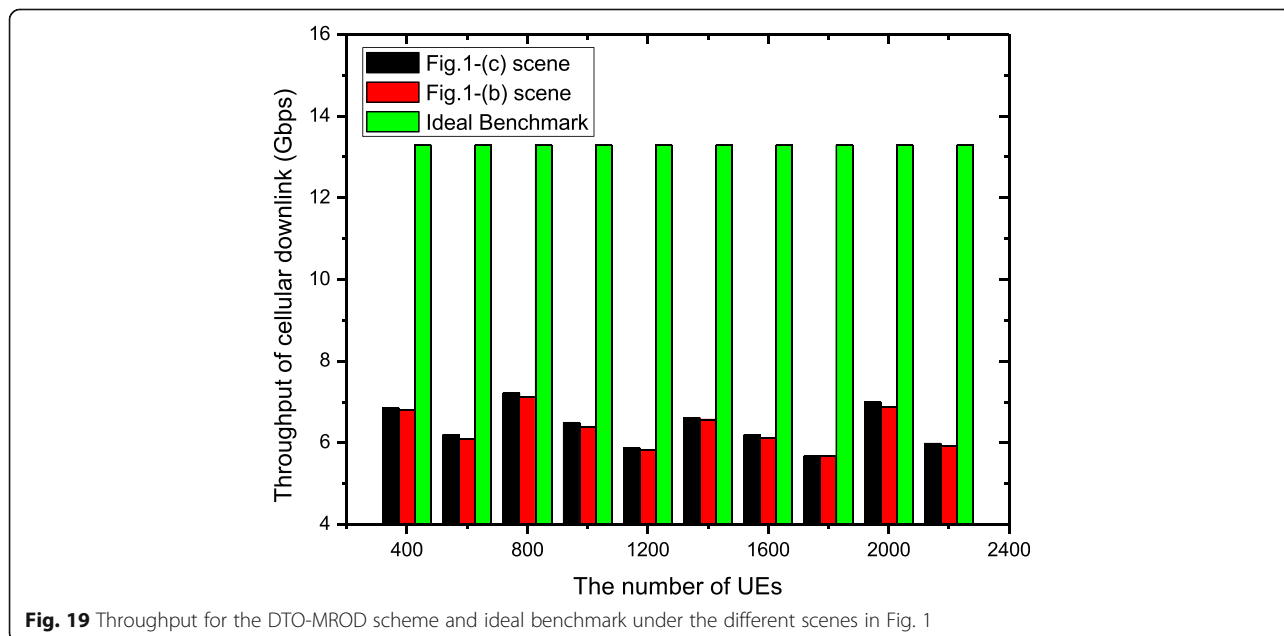


Fig. 17 Average delay for the DTO-MROD scheme under the different verification orders of relay preselection



forwarding. Although the cellular link distance may be relatively long, eNB is the transmitter of this link, whose transmission power is greater than normal UEs. Meanwhile, due to the longer cellular link distance, the receiving UEs are most likely to select the second relay when η equals 0.3. Therefore, the bottleneck of forwarding capability is less likely to exist in the whole path. On the other hand, in Fig. 12, due to the number of communication paths owning two relays is the largest when η equals 0.3, these paths are most likely to be disconnected compared with the other two conditions.

Figures 13, 14, and 15 show the effects of three typical RIE threshold values on a cellular downlink throughput, average delay, and network connectivity ratio respectively. From formula (14), we could know that the relay with sub-optimal throughput is more likely to replace the relay with optimal throughput, when RIE threshold becomes larger. As we can see in Fig. 13, the changes of RIE threshold have little impact on cellular throughput, especially in low density of UEs. Also, as the number of UEs (i.e., M) within the cellular network increases, the throughput does not show obvious regularity. This is because in some cases, the replacement of optimal relay



with sub-optimal relay decreases the cellular throughput. But in other cases, the replaced optimal relays are used in other transmission links which increase the cellular throughput conversely. Thus, the trend of throughput shows a certain degree of randomness. Familiar with the above reason, the trend of average delay depicted in Fig. 14 also shows a certain degree of randomness. From Fig. 15, we could know that the larger the *RIE* threshold is, the higher the network connectivity ratio will be. Based on the above-mentioned analysis, this phenomenon is mainly caused by relay replacement in the preselection of the first relay, which is more likely to happen when *RIE* threshold gets larger. Also, the relay has higher remaining energy after relay replacement; thus, the network connectivity ratio is improved.

Figures 16, 17, and 18 show the effects of different verification order of preselected relay on cellular downlink throughput, average delay, and network connectivity ratio respectively. The order of near to far has advantage in improving throughput, the reason is that the action of verifying the receiving UEs that is closer to *eNB* firstly makes the receiving UEs abandon their relay because of large interference. And this behavior reduces interference when the distant receiving UEs start to verify their throughput; thus, the distant receiving UEs are more likely to retain their relays. On the other hand, by using relays, the throughput of the distant UEs improves greatly compared with the closer UEs. The throughput improvement of far to near is contrary to near to far; however, the order of random verification combines the features of the above two orders; thus, its throughput improvement is between near to far and far to near. But the effects of different verification orders on average

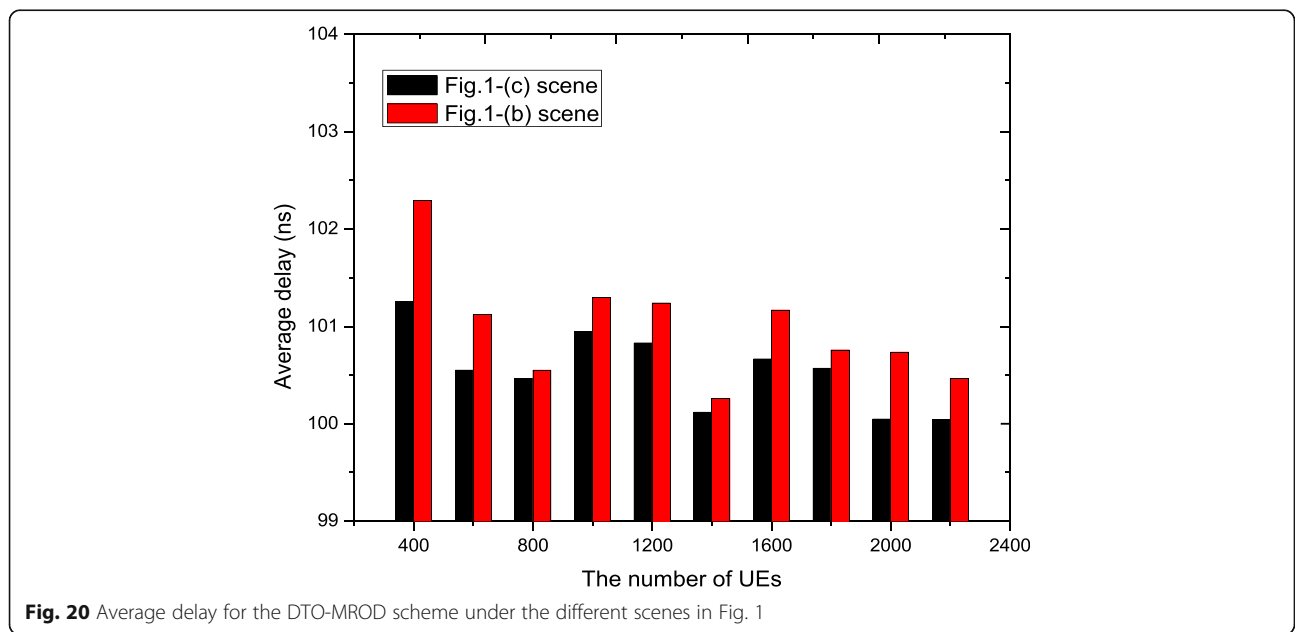
delay and network connectivity ratio show no obvious regulations.

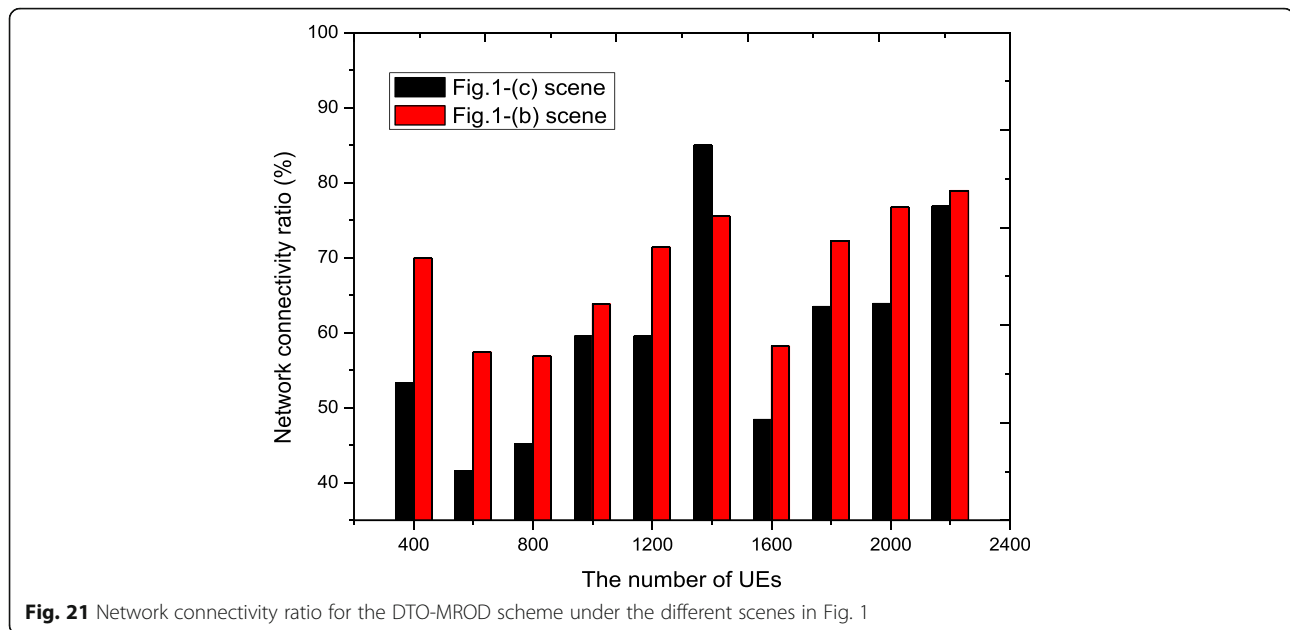
The third group of experiments (Figs. 19, 20, and 21) shows the comparison among cellular downlink throughput, average delay, and network connectivity ratio between scene (b) and scene (c) in Fig. 1. The simulation parameters are set as those in the first group of experiments. In Fig. 19, throughput in scene (c) is better than that in scene (b). This is because the relay selection in scene (c) is not under constraints; thus, the receiving UEs could select an UE as a relay which could optimize link throughput. The interpretation of the ideal result in Fig. 19 is similar to that in Fig. 4.

Based on the explanation of Fig. 5 that is the definition of delay in this paper; the data transfer delay of any path depends on the relay with the worst forwarding capability in this path, while the forwarding capability directly influences the throughput of this path. Thus, the trend of Fig. 20 is familiar with Fig. 19, which is reasonable. Figure 21 shows that the network connectivity ratio in scene (c) is worse than that in scene (b), and the reason is that some UEs (e.g., UE4 in Fig. 1 (c)) located in a key position provide relaying services for other receiving UEs, which leads to high energy consumption; thus, these links are disconnected in shorter time.

5 Conclusions

To deal with tremendous traffic demand in a cyberspace on account of innumerable wireless smart devices and a number of multimedia applications, we propose the DTO-MROD scheme to optimize downlink throughput in a cellular system. We extend two-hop communication links to three-hop communication links to reduce





receiving bit error ratio efficiently and thus improve cellular downlink throughput. The first relay is selected according to the proposed *RIE* metric and the related estimation formulas, which can balance network lifetime and throughput. The remaining energy of the second relay is not less than the minimum value between the first relay's energy and the energy threshold.

In addition, the DTO-MROD scheme firstly adopts relay preselection and then verifies whether the preselected relays improve throughput or not. Therefore, it can determine the ultimate relays and the corresponding transmission paths at a time instead of several times of iteration, which reduces the communication overhead. Also, the DTO-MROD scheme adopts a local searching strategy; thus, it reduces the local computational overhead and also ensures that the optimal relays are selected by reasonably setting a searching scope. The simulation results and theoretical analysis reveal that our scheme outperforms the compared scheme in different perspectives of downlink throughput, delay, network lifetime, and communication overhead.

Acknowledgements

The authors would like to thank the National Natural Science Foundation of China under Grant Nos. 61272494, 61379110, 61379058, 61379057, 61073104 and Central South University postgraduate independent exploration and innovation project (2017zts502) for sponsoring this research work.

Authors' contributions

KZ is the main author of the current paper. KZ and JG contributed to the conception and design of the study. NX contributed to the structuring and reviewing of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹School of Information Science and Engineering, Central South University, South Road of LuShan, Changsha 410083, China. ²School of Computer Science, Colorado Technical University, 4435 North Chestnut Street, Colorado Spring, CO 80907, USA.

Received: 26 June 2017 Accepted: 25 November 2017

Published online: 08 December 2017

References

1. JS Gui, K Zhou, NX Xiong, A cluster-based dual-adaptive topology control approach in wireless sensor networks. *Sensors* **16**(10), 1576 (2016)
2. X Chen, YX Xu, AF Liu, Cross layer design for optimizing transmission reliability, energy efficiency, and lifetime in body sensor networks. *Sensors* **17**(4), 900 (2017). <https://doi.org/10.3390/s17040900>
3. JS Gui, J Deng, A topology control approach reducing construction cost for lossy wireless sensor networks. *Wireless Pers Commun.* **95**(3), 2173–2202 (2017)
4. T Li, AF Liu, CQ Huang, A similarity scenario based recommendation model with small disturbances for unknown items in social networks. *IEEE Access.* **4**, 9251–9272 (2016)
5. JS Gui, K Zhou, Flexible adjustments between energy and capacity for topology control in heterogeneous wireless multi-hop networks. *J. Netw. Syst. Manag.* **24**(4), 789–812 (2016)
6. YX Xu, X Chen, AF Liu, CH Hu, A latency and coverage optimized data collection scheme for smart cities based on vehicular ad-hoc networks. *Sensors* **17**(4), 888 (2017). <https://doi.org/10.3390/s17040888>
7. JS Gui, LH Hui, NX Xiong, A game-based localized multi-objective topology control scheme in heterogeneous wireless networks. *IEEE Access.* **5**(1), 2396–2416 (2017)
8. JS Gui, YJ Lu, XH Deng, AF Liu, Resource allocation adaptive to communication strategy selection for cellular clients using Stackelberg game. *Ad Hoc Netw.* **66**, 64–84 (2017). doi.org/10.1016/j.adhoc.2017.08.003
9. IEEE draft standard for local and metropolitan area networks, *Media independent handover services—amendment for Security Extensions to Media Independent Handover Services and Protocol, 802.21a/D05* (IEEE, Piscataway, 2011), pp. 1–90
10. A Asadi, V Mancuso, R Gupta, *An SDR-based experimental study of outband D2D communications* (Paperpresented at the Proceeding of IEEE

- International Conference on Computer Communications (IEEE INFOCOM 2016), San Francisco, 2016), pp. 10–15
11. P Mach, Z Becvar, T Vanek, Inband device-to-device communication in OFDMA cellular networks: a survey and challenges. *IEEE Communications Surveys and Tutorials*. **17**, 1885–1922 (2015)
 12. A Asadi, AQ Wang, V Mancuso, A survey on device-to-device communication in cellular networks. *IEEE Commun. Surveys Tuts*. **16**, 1801–1819 (2014)
 13. HE Elkotby, KMF Elsayed, MH Ismail, Shrinking the reuse distance: spectrally-efficient radio resource management in D2D-enabled cellular networks with interference alignment. *IEEE Wireless Days*. **139**, 1–6 (2013)
 14. A Asadi, V Mancuso, Network-assisted outband D2D-clustering in 5G cellular networks: theory and practice. *IEEE Trans. Mob. Comput.* **16**(8), 2246–2259 (2017)
 15. A Hematian, W Yu, C Lu, DW Griffith, NT Golmie, *A clustering-based device-to-device communication to support diverse applications* (Paper presented at the Proceedings of the International Conference on Research in Adaptive and Convergent Systems, Odense, 2016), pp. 11–14
 16. K Vanganuru, S Ferrante, G Sternberg, *System capacity and coverage of a cellular network with D2D mobile relays* (Paper presented at the Proceedings of Military Communications Conference, Orlando, 2012), p. 28 Oct-01 Nov.
 17. L Wang, T Peng, Y Yang, W Wang, *Interference constrained relay selection of D2D communication for relay purpose underlaying cellular networks* (Paper presented at the in Proceedings of International Conference on Wireless Communications, Networking and Mobile Computing, TBD Shanghai, China, 2012), pp. 21–23
 18. Y Pei, YC Liang, Resource allocation for device-to-device communications overlaying two-way cellular network. *IEEE Trans. Wirel. Commun.* **12**, 3611–3621 (2013)
 19. T Ta, JS Baras, C Zhu, *Improving smartphone battery life utilizing device-to-device cooperative relays underlaying LTE networks* (Paper presented at the Proceedings of IEEE International Conference on Communications, Sydney, 2014), pp. 10–14
 20. T Chen, G Charbit, S Hakola, *Time hopping for device-to-device communication in LTE cellular system* (Paper presented at the Proceedings of Wireless Communications and Networking Conference, Sydney, 2010), pp. 18–21
 21. DH Lee, KW Choi, WS Jeon, DG Jeong, *Resource allocation scheme for device-to-device communication for maximizing spatial reuse* (Paper presented at the Proceedings of Wireless Communications and Networking Conference, Shanghai, 2013), pp. 07–10
 22. CH Yu, K Doppler, CB Ribeiro, O Tirkkonen, Resource sharing optimization for device-to-device communication underlaying cellular networks. *IEEE Trans. Wirel. Commun.* **10**, 2752–2763 (2011)
 23. T Koskela, S Hakola, T Chen, J Lehtomaki, *Clustering concept using device-to-device communication in cellular system* (Paper presented at the Proceedings of Wireless Communications and Networking Conference (IEEE, 2010), Sydney, 2010), pp. 18–21
 24. S Hakola, T Chen, J Lehtomaki, T Koskela, *Device-to-device (D2D) communication in cellular network performance analysis of optimum and practical communication mode selection* (Paper presented at the Proceedings of Wireless Communications and Networking Conference (IEEE, 2010), Sydney, 2010), pp. 18–21
 25. C Vitale, V Sciancalepore, A Asadi, V Mancuso, *Two-level opportunistic spectrum management for Green 5G radio access networks*, Paper Presented at the Proceedings of Green Communications (Online Green Comm) (2015), pp. 10–12
 26. O Delgado, F Labeau, *D2D relay selection and fairness on 5G wireless networks* (Paper presented at the Proceedings of Globecom Workshops (GC Wkshps), Washington DC, 2016), pp. 4–8
 27. N Mastrorarde, V Patel, J Xu, LJ Liu, MVD Schaar, To relay or not to relay: learning device-to-device relaying strategies in cellular networks. *IEEE Trans. Mob. Comput.* **15**, 1569–1585 (2016)
 28. K Doppler, CH Yu, CB Ribeiro, P Janis, *Mode selection for device-to-device communication underlaying an LTE-advanced network* (Paper presented at the Proceedings of Wireless Communications and Networking Conference (IEEE, 2010), Sydney, 2010), pp. 18–21
 29. JS Gui, AF Liu, A new distributed topology control algorithm based on optimization of delay and energy in wireless networks. *Journal of Parallel and Distributed Computing*. **72**(8), 1032–1044 (2012)
 30. WB Heinzelman, Application-specific protocol architectures for wireless networks, PH.D. Thesis, Massachusetts Institute of Technology (2000)
 31. 3GPP, TR 23.703 Technical Specification Group Services and System Aspects; Study on architecture enhancements to support Proximity-based Services (ProSe) (Release 12). 2013.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com