

PAPER

Joint User Grouping and Resource Allocation for NOMA Enhanced D2D Communications

Jin XIE[†] and Fangmin XU^{†a)}, *Nonmembers*

SUMMARY To mitigate the interference caused by frequency reuse between inter-layer and intra-layer users for Non-Orthogonal Multiple Access (NOMA) based device-to-device (D2D) communication underlying cellular systems, this paper proposes a joint optimization strategy that combines user grouping and resource allocation. Specifically, the optimization problem is formulated to maximize the sum rate while ensuring the minimum rate of cellular users, considering three optimization parameters: user grouping, sub channel allocation and power allocation. However, this problem is a mixed integer nonlinear programming (MINLP) problem and is hard to solve directly. To address this issue, we divide the problem into two sub-problems: user grouping and resource allocation. First, we classify D2D users into D2D pairs or D2D NOMA groups based on the greedy algorithm. Then, in terms of resource allocation, we allocate the sub-channel to D2D users by swap matching algorithm to reduce the co-channel interference, and optimize the transmission power of D2D by the local search algorithm. Simulation results show that, compared to other schemes, the proposed algorithm significantly improves the system sum rate and spectral utilization.

key words: NOMA, user grouping, D2D NOMA group, D2D pair, swap matching

1. Introduction

With the increasing popularity of intelligent terminal devices, the demand for wireless network resources continues to grow. Several initial technologies for 5G networks have been identified by industry experts, among which D2D communication is considered an effective solution to the shortage of spectrum resources [1]. D2D technology allows users to communicate directly within the cellular network without the need for base station relay. This approach helps to reduce base station load and can also improve spectrum efficiency through reuse of radio resources. However, D2D technology also introduces co-channel interference between cellular and D2D links, which presents new challenges for interference suppression strategies [2], [3].

As another method for improving network spectrum efficiency, NOMA technology has also attracted the attention of researchers, opening up new directions for large-scale network access with its unique power domain multiplexing technology. Unlike traditional Orthogonal Multiple Access (OMA) techniques [4]–[8], NOMA allows multiple users to share the same resource block. In NOMA, user signals are multiplexed using superposition coding at the transmitter

and demultiplexed using successive interference cancellation (SIC) at the receiver. This improves rate, user fairness, and scheduling flexibility while reducing intra-user interference caused by multiplexing. Therefore, NOMA is considered one of the key technologies for future mobile communication networks [9]–[21].

1.1 Related Work

In recent years, the combined application of NOMA and D2D has received extensive attention [9]–[14]. Generally, the resource allocation model that combines NOMA and D2D can be divided into two categories. The first category is NOMA groups composed of cellular users and D2D pairs [9]–[11]. The second category is D2D NOMA groups consisting of Device Transmitter (DT) and multiple Device Receivers (DRs) [12]–[14].

In the first category, NOMA is generally used for up-link cellular users and D2D communication. Reference [9] proposes a D2D-interlay communication mode that adopts a graph-based multi-mode selection and channel allocation strategy in the NOMA-D2D system. This mode supports power domain multiplexing of user signals and eliminates strong interference between D2D and cellular users through SIC decoding. Reference [10] proposes a convex approximation algorithm to solve the channel allocation and power control problems and designs a convolutional neural network algorithm to reduce computational complexity. Unlike references [9]–[11] applies the SIC scheme to both the Base Station (BS) and D2D receivers to further reduce co-channel interference.

In the second category, NOMA is only used in D2D NOMA groups. Reference [12] proposes the concept of D2D NOMA groups. Unlike traditional D2D pairs, the D2D transmitters in D2D NOMA group can use NOMA technology to communicate with multiple receivers using the same resources, thereby improving spectrum efficiency. Reference [13] establishes a three-dimensional matching game model for joint user group association and sub-carrier allocation and optimizes the power of D2D using branch-and-bound techniques. In [14], the authors propose a DT-DR grouping scheme and then design resource allocation for cellular mobile users and D2D mobile groups using a many-to-many mapping scheme. Simulation in [14] shows that the pairing of DT and DR significantly affects resource allocation and thus the performance of the D2D system. Inspired by the advantages of user pairing, channel resource allocation and

Manuscript received July 20, 2023.

Manuscript revised September 6, 2023.

Manuscript publicized September 20, 2023.

[†]The authors are with Hangzhou Dianzi University, China.

a) E-mail: xufangmin@hdu.edu.cn

DOI: 10.1587/transfun.2023EAP1086

power control in improving system performance, we investigate a novel joint optimization scheme considering DT-DR pairing based on greedy algorithm, sub-channel allocation and power allocation.

1.2 Contribution

Although the resource allocation problem for D2D communication underlying cellular systems has been extensively studied in the existing literature, most of them are based on the assumption that D2D users have been paired. Therefore, this paper studies how to jointly optimize user pairing and wireless resource allocation to improve the system more effectively. Different from the existing work, DTs and DRs are matched into D2D NOMA group or D2D pair (collectively called the D2D) according to a given algorithm. By designing user grouping and resource allocation between D2D and cellular users (CUEs), we maximize the sum rate of the network while ensuring the minimum rate of CUEs. The main contributions of this paper are as follows:

1. We introduce a novel network model that allows for the simultaneous existence of D2D NOMA groups and D2D pairs. For D2D NOMA groups, each DT can communicate with multiple DRs simultaneously through the NOMA protocol, requiring fewer spectrum resources compared to the OMA model. Then, based on this model, a joint optimization problem is formulated to improve the system sum rate by introducing two types of 0-1 integer variables, one for the DT-DR user matching and the other for sub-channel assignment.

2. To solve the formulated mixed integer nonlinear optimization problem, we divide the problem into three sub-problems: DT-DR grouping, D2D sub-channel allocation, and DT power control. What's more, We prove its convergence and stability theoretically.

3. For DT-DR grouping, we propose a greedy algorithm to find the appropriate matching relationship and improve DR access rates. Based on the maximum communication distance of DT-DR, each DT is paired with DR in the form of a NOMA group or a pair. When there is no DR within the range of DT, DT is treated as a cellular user for uplink transmission and will not reuse cellular user's sub-channel.

4. When DT-DR grouping and DT transmission power are fixed, we convert the problem into a many-to-one matching problem. To maximize the total system rate, a D2D swap matching algorithm is proposed to reduce co-channel interference.

5. With fixed DT-DR grouping and sub-channel allocation, a power control scheme based on local search is designed to further improve the sum rate. In addition, we evaluate the performance of the proposed algorithm through simulations.

2. System Description

In this section, the system scenario and D2D communication model are introduced, and the optimization problem is

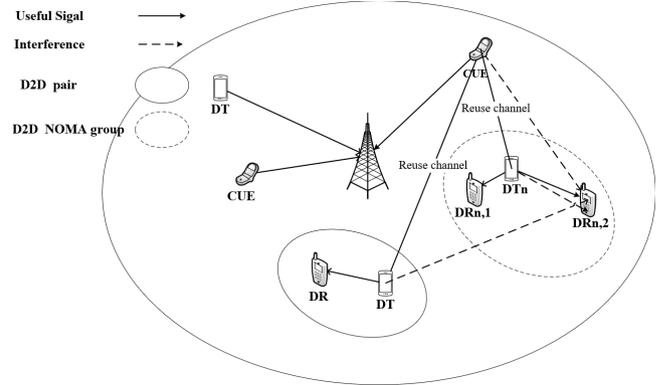


Fig. 1 The system model of NOMA D2D underlay cellular networks.

formulated based on the interference analysis.

2.1 System Model

We consider a time-division duplex (TDD) NOMA-D2D uplink communication system, as shown in Fig. 1. The system model consists of one BS, I cellular users, J DTs, and F DRs. There are two combination modes of DT and DR, namely D2D NOMA group (i.e. one DT to two DRs) and D2D pair (i.e. one DT to one DR). We define $\mathcal{I} = \{1, \dots, i, \dots, I\}$, $\mathcal{J} = \{1, \dots, j, \dots, J\}$, $\mathcal{F} = \{1, \dots, f, \dots, F\}$, $\mathcal{N} = \{1, \dots, n, \dots, N\}$ as the sets of CUEs, DTs, DRs, and RBs (Resource Blocks), respectively. We use ρ to represent the pairing relationship between DT and DR. If the f -th DR is paired with the j -th DT, then $\rho_{j,f} = 1$, otherwise $\rho_{j,f} = 0$. We represent the pairing matrix of the j -th DT and DR as $\beta_{j \times \mathcal{F}} = [\rho_{j,1}, \dots, \rho_{j,f}, \dots, \rho_{j,F}]$. Furthermore, we represent all DT-DR pairing matrices as $\beta_{\mathcal{J} \times \mathcal{F}} = [\beta_{1 \times \mathcal{F}}, \dots, \beta_{j \times \mathcal{F}}, \dots, \beta_{J \times \mathcal{F}}]^T$. If $\sum_{f=1}^F \rho_{j,f} = 2$, DT and DRs are paired as a D2D NOMA group. Otherwise, if $\sum_{f=1}^F \rho_{j,f} = 1$, DT and DR are paired as a D2D pair.

2.2 SINR Analysis

1) **SINR analysis for cellular users:** In the uplink scenario, the signal received at the BS from the i -th cellular user on the n -th RB is given by:

$$y_{i,B}^n = |h_{i,B}^n| \sqrt{P_i} x_{i,B}^n + \sum_{j=1}^J \phi_{j,i}^n |h_{j,B}^n| \sqrt{P_j} x_{j,B}^n + \xi_{i,B}^n \quad (1)$$

where ϕ identifies the RBs used by the user or not. If the i -th CUE and the j -th DT reuse the n -th RB, $\phi_{j,i}^n = 1$, otherwise $\phi_{j,i}^n = 0$. P_i and P_j represent the transmit power of i -th CUE and j -th DT, respectively. $\xi_{i,B}^n$ is the Gaussian white noise. $h_{i,B}^n$ denotes the channel gain from i -th CUE to BS on the n -th RB, which can be expressed as:

$$h_{i,B}^n = G g_{i,B}^n \beta_{i,B}^n d_{i,B}^{-\theta} \quad (2)$$

where G is the path loss constant, $g_{i,B}^n$ and $\beta_{i,B}^n$ are the

log-normal shadowing fading and small-scale fading respectively. $d_{i,B}$ is the distance between i -th CUE and BS, θ is the path loss exponent. Similarly, $h_{j,B}^n$ is channel gain from j -th DT to BS on the n -th RB. Therefore, the SINR of the i -th CUE can be expressed as:

$$r_i^n = \frac{P_i |h_{i,B}^n|^2}{\sum_{j=1}^J \phi_{j,i}^n P_j |h_{j,B}^n|^2 + N_0} \quad (3)$$

2) SINR analysis for D2D NOMA groups: In each D2D NOMA group, power signal reuse is achieved at the DT by using superimposed coding, while SIC is used at the DR to reduce intra-group interference. Specifically, the superimposed signal transmitted by DT can be represented as $\alpha_{j,f_s} x_{j,f_s}^n + \alpha_{j,f_w} x_{j,f_w}^n$, where f_s and f_w represent the stronger DR and the weaker DR associated with the j -th DT, respectively. $\alpha_{j,\epsilon}$ and $x_{j,\epsilon}^n$ represent the power allocation coefficient and signal, respectively, where $\epsilon = f_s$ or $\epsilon = f_w$. If $|h_{j,f_s}^n| > |h_{j,f_w}^n|$, the signal received at the DR f_w can be expressed as:

$$\begin{aligned} y_{j,f_w}^n &= \underbrace{\sqrt{P_j \alpha_{j,f_w}} |h_{j,f_w}^n| x_{j,f_w}^n}_{\text{Expected signal}} + \underbrace{\sqrt{P_j \alpha_{j,f_s}} |h_{j,f_s}^n| x_{j,f_s}^n}_{\text{intra-group interference } (I_{j,f_w}^{in})} \\ &+ \underbrace{\sum_{j' \neq j} \phi_{j',j}^n \sqrt{P_{j'}} |h_{j',j,f_w}^n| x_{j',j,f_w}^n}_{\text{Cochannel interference } (I_{j,f_w}^{out})} + \underbrace{\phi_{j,i}^n \sqrt{P_i} |h_{i,j,f_w}^n| x_{i,j,f_w}^n}_{\text{CUE interference } (I_{j,f_w}^c)} \\ &+ \xi_{j,f_w}^n \end{aligned} \quad (4)$$

The five terms on the right side of Eq. (4), from left to right, are the expected signal, the intra-group interference signal, the co-channel interference signal from other DTs, the co-channel interference from the cellular user and the Gaussian white noise ξ_{j,f_w}^n . Hence, the SINR of the DR f_w can be expressed as:

$$r_{j,f_w}^n = \frac{|h_{j,f_s}^n|^2 P_j \alpha_{j,f_s}}{I_{j,f_w}^{in} + I_{j,f_w}^{out} + I_{j,f_w}^c + N_0} \quad (5)$$

The interference cancellation is successful if DR f_s received SINR for DR f_w 's signal is not smaller than the received SINR at DR f_w for its own signal [12], that is:

$$\frac{|h_{j,f_s}^n|^2 P_j \alpha_{j,f_s}}{I_{j,f_s}^{in} + I_{j,f_s}^{out} + I_{j,f_s}^c + N_0} \geq \frac{|h_{j,f_w}^n|^2 P_j \alpha_{j,f_w}}{I_{j,f_w}^{in} + I_{j,f_w}^{out} + I_{j,f_w}^c + N_0} \quad (6)$$

(6) can be rewritten as:

$$\begin{aligned} Q_j(\phi) &\triangleq |h_{j,f_s}^n|^2 \alpha_{j,f_s} (I_{j,f_w}^{in} + I_{j,f_w}^{out} + I_{j,f_w}^c + N_0) - \\ &|h_{j,f_w}^n|^2 \alpha_{j,f_w} (I_{j,f_s}^{in} + I_{j,f_s}^{out} + I_{j,f_s}^c + N_0) \geq 0 \end{aligned} \quad (7)$$

$Q_j(\phi)$ indicates successful SIC. By SIC, the stronger DR

removes intra-group interference and decodes its own information successfully, so the SINR of DR f_s can be expressed as:

$$r_{j,f_s}^n = \frac{|h_{j,f_w}^n|^2 P_j \alpha_{j,f_w}}{I_{j,f_w}^{out} + I_{j,f_w}^c + N_0} \quad (8)$$

3) SINR analysis for D2D pairs: For traditional D2D pairs, DRs receive interference from cellular users and other DTs on the same RB, so the SINR of the f -th DR can be expressed as:

$$r_{j,f}^n = \frac{|h_{j,f}^n|^2 P_j \alpha_{j,f}}{\sum_{j' \neq j} \phi_{j',j}^n |h_{j',j,f}^n|^2 P_{j'} + \phi_{j,i}^n |h_{i,j,f}^n|^2 P_i + N_0} \quad (9)$$

2.3 Problem Description

According to Shannon Theory and (3), the data rate of the i -th CUE is given by:

$$R_i^n = \log_2(1 + r_i^n) \quad (10)$$

From (5) and (8), the total rate of DRs in the D2D NOMA group mode can be obtained as:

$$R_{j,f_s+f_w}^n = \log_2(1 + r_{j,f_s}^n) + \log_2(1 + r_{j,f_w}^n) \quad (11)$$

Similarly, the data rate of the f -th DR in D2D pair mode can be expressed as:

$$R_{j,f}^n = \log_2(1 + r_{j,f}^n) \quad (12)$$

Therefore, the sum rate of CUEs and the sum rate of DRs in D2D NOMA group mode are given by:

$$R_i = \sum_{n=1}^N R_i^n \quad (13)$$

and

$$R_{f_s+f_w} = \sum_{n=1}^N \sum_{j=1}^J m_j R_{j,f_s+f_w}^n \quad (14)$$

respectively. Here, m_j indicates whether the D2D user is in D2D NOMA group mode or not. Specifically, for any $j \in \mathcal{J}$, if $\sum_{f=1}^F \rho_{j,f} = 2$, then $m_j = 1$, otherwise $m_j = 0$. The sum rate of DRs in D2D pair can be obtained as follows:

$$R_f = \sum_{n=1}^N \sum_{j=1}^J v_j R_{j,f}^n \quad (15)$$

Here, v_j indicates whether the D2D user is in D2D pair mode. For any $j \in \mathcal{J}$, if $\sum_{f=1}^F \rho_{j,f} = 1$, then $v_j = 1$, otherwise $v_j = 0$. Therefore, the sum rate of the system can be expressed as:

$$R_{i+f}(\rho, \phi, \alpha) = R_i + R_{f_s+f_w} + R_f \quad (16)$$

To maximize the sum-rate of the system while ensuring the

minimum rate of CUEs, we formulate the optimization problem as:

$$\max_{\rho, \phi, \alpha} R_{i+f}(\rho, \phi, \alpha) \quad (17)$$

$$\text{s.t. } \sum_{j=1}^J \rho_{j,f}^n \leq 1, \forall f \in \mathcal{F} \quad (17a)$$

$$\rho_{j,f} \in \{0, 1\}, \forall j \in \mathcal{J}, f \in \mathcal{F} \quad (17b)$$

$$\sum_{f=1}^F \rho_{j,f} \in \{0, 1, 2\}, \forall j \in \mathcal{J} \quad (17c)$$

$$Q_j(\phi) \geq 0, \forall j \in \mathcal{J} \quad (17d)$$

$$R_i^n > R_i^{n, \min}, \forall i \in \mathcal{I}, n \in \mathcal{N} \quad (17e)$$

$$\phi_{j,i}^n, \phi_{j',j}^n \in \{0, 1\}, \quad (17f)$$

$$\forall j \in \mathcal{J}, f \in \mathcal{F}, n \in \mathcal{N} \quad (17f)$$

$$d_{j,f} < d_{\max}, \forall j \in \mathcal{J}, f \in \mathcal{F} \quad (17g)$$

$$\alpha_{j,f} \geq 0, \alpha_{j,f_s} \geq 0, \alpha_{j,f_w} \geq 0, \quad (17h)$$

$$\forall j \in \mathcal{J}, f, f_s, f_w \in \mathcal{F} \quad (17h)$$

$$\alpha_{j,f_s} + \alpha_{j,f_w} \leq 1, \quad (17i)$$

$$\forall j \in \mathcal{J}, f_s, f_w \in \mathcal{F} \quad (17i)$$

The constraint (17a) indicates that each DR can be assigned to at most one DT. (17b) represents the integer constraint on the DT-DR pairing identifier. Constraint (17c) means that each DT can be paired with DR to form a D2D NOMA group, D2D pair, or perform OMA transmission. Constraint (17d) ensures successful SIC at stronger DR in D2D NOMA groups. Constraint (17e) guarantees the minimum rate of CUEs. Constraint (17f) ensures each CUE and D2D can only use one RB. Constraint (17g) represents the maximum constraint value for the distance between DTs and DRs. Constraint (17h) is the non-negative constraint for the DT power allocation coefficient. Constraint (17i) represents the transmission power constraint of DTs.

3. Sum-Rate Optimization Scheme

Due to the presence of binary and continuous variables, the optimization problem (17) is a MINLP problem and cannot be solved directly. Therefore, we convert this problem into three sub-problems: DT-DR grouping, D2D channel allocation, and DT power allocation. For DT-DR grouping, we match DRs and DTs into D2D NOMA group or D2D pair by greedy algorithm, with the goal of maximizing the DR access rate. To solve the D2D channel allocation problem, we exchange D2D channels to reduce intra-layer and inter-layer interference based on matching game theory. Finally, in the phase of power allocation, a local power search algorithm is designed to optimize the DT's transmission power.

3.1 DT-DR Grouping

In this sub-section, we optimize the binary variables $\rho_{j,f}$ with fixed channel allocation and power allocation factors,

and rewrite the optimization problem of (17) as:

$$\max_{\rho} R_{i+f}(\rho) \text{ s.t. (17a)–(17c), (17g)} \quad (18)$$

To solve problem (18), we pair DRs and DTs into D2D NOMA groups or D2D pairs by greedy algorithm, as shown in Algorithm 1. Firstly, each DT obtains the list of all DRs that can be paired with it within the range of d_{\max} . We give priority to pairing DTs with lists of length greater than or equal to 2. Next, the two DRs with the highest rate are selected and paired with the corresponding DT. DTs and DRs that are successfully paired do not participate in subsequent pairing. This pairing process does not stop until all DT lists are less than 2 in length. For the remaining DTs, if the length of the DT list is 1, the DT is paired with its corresponding DR to form the D2D pair. If the length of the DT list is 0, the DT is treated as a cellular user for uplink transmission. If the DR is not paired with any DT, it will perform OMA transmission.

Algorithm 1 The proposed Greedy-based DT-DRs Grouping Algorithm (GDGA)

Input: Randomly generate locations for CUEs, DTs and DRs. Initialize $\beta_{\mathcal{J} \times \mathcal{F}} = \mathbf{0}$ and $\mathbf{Z} = \{\emptyset\}$, where \mathbf{Z} is used to record DTs with DT list length larger than 2.

Output: ρ^* and $\beta_{\mathcal{J} \times \mathcal{F}}^*$

```

1: while true do
2:   By detecting the DRs in each DT range, we obtain  $Y = \{Y_1, \dots, Y_j, \dots, Y_J\}$ , where  $Y_j$  is a list which records the DRs within  $d_{\max}$  meters of DT $_j$ . Here,  $d_{\max}$  is the maximum distance between a DT and a DR.
3:   for  $j=1$  to  $J$  do
4:     if length of  $Y_j \geq 2$  then
5:        $\mathbf{Z} \leftarrow \mathbf{Z} \cup \{Y_j\}$ 
6:     end if
7:   end for
8:   For all DTs in  $\mathbf{Z}$  and DRs in their corresponding DT lists, calculate  $R_{j,f_s+f_w}^n$  according to (5), (8) and (11);
9:    $\rho_{j^*,f_s^*}^*, \rho_{j^*,f_w^*}^* = 1 \leftarrow (j^*, f_s^*, f_w^*) = \text{argmax}_{j,f_s,f_w} R_{j,f_s+f_w}^n$ ;
10:   $\beta_{\mathcal{J} \times \mathcal{F}} \leftarrow \beta_{\mathcal{J} \times \mathcal{F}} \cup (\rho_{j^*,f_w^*}^*, \rho_{j^*,f_s^*}^*)$ ;
11:   $\mathcal{J} \leftarrow \mathcal{J} \setminus \{j^*\}$ ,  $\mathcal{F} \leftarrow \mathcal{F} \setminus \{f_s^*, f_w^*\}$ ;
12:  if  $\mathbf{Z} = \{\emptyset\}$  then
13:    break
14:  end if
15: end while
16: for  $j=1$  to  $J$  do
17:   if length of  $Y_j=1$  then
18:      $\rho_{j^*,f^*}^* = 1$ ;
19:      $\beta_{\mathcal{J} \times \mathcal{F}}^* \leftarrow \beta_{\mathcal{J} \times \mathcal{F}} \cup \rho_{j^*,f^*}^*$ .
20:   end if
21: end for

```

3.2 Sub-Channel Allocation for D2D Users

In the previous section, the parameter ρ is optimized, i.e., DTs and DRs are paired into D2D NOMA group or D2D pair. In this section, we optimize the D2D channel with fixed DT power. The optimization problem of (17) is rewritten as:

$$\max_{\phi} R_{i+f}(\phi) \text{ s.t. (17d)–(17f)} \quad (19)$$

In order to solve the 0-1 integer programming problem (19), we propose a swap matching algorithm to reduce the cross-layer co-channel interference between CUEs and D2D users, while reducing the co-layer interference between D2D users. Let $\mathbf{X}_{\mathcal{J} \times \mathcal{J}}$ be the sub-channel matching matrix, D2D $_j$ is the j -th D2D pair or D2D NOMA group that has been paired successfully. Each CUE is assigned a sub-channel in advance, if D2D $_j$ multiplexes i -th CUE's channel, then the i -th row and the j -th column of $\mathbf{X}_{\mathcal{J} \times \mathcal{J}}$ is 1, otherwise, it's 0. The utility function of D2D $_j$ is denoted as $u_j(\mathbf{X})$, which represents the sum rate of DRs with D2D $_j$ in the matching state \mathbf{X} . The utility function of the n -th RB is denoted as $u_n(\mathbf{X})$, which represents the sum rate over the n -th RB in the matching state \mathbf{X} . The swap-matching algorithm is defined as follows:

Definition: For a matching \mathbf{X} , the corresponding matching set is denoted as $M_X = \{(j, n)\}$, where the matching pairs i -th CUE and D2D $_j$ satisfy $\phi_{j,i}^n = 1$. For any two elements $(j, n), (j', n') \in M_X$, we denote the matching matrix after swapping as \mathbf{X}' and the corresponding matching set is $M_{X'}$, which is defined as $M_{X'} = M_X \cup \{(j, n'), (j', n)\} \setminus \{(j, n), (j', n')\}$. It is worth noting that swap matching is only performed when certain conditions are satisfied in order to effectively improve system performance.

Exchange condition: For an existing matching \mathbf{X} , where $\phi_{j,i}^n = 1$ and $\phi_{j',i'}^{n'} = 1$, D2D users $(j, j') \in \mathcal{J}$ are allowed to be exchanged if and only if the following two constraints are satisfied.

- 1) $\forall j, j', n, n'$, we have $u_j(\mathbf{X}') \geq u_j(\mathbf{X})$, $u_{j'}(\mathbf{X}') \geq u_{j'}(\mathbf{X})$ and $u_n(\mathbf{X}') + u_{n'}(\mathbf{X}') \geq u_n(\mathbf{X}) + u_{n'}(\mathbf{X})$.
- 2) $\exists j, j', n, n'$, s.t. $u_j(\mathbf{X}') > u_j(\mathbf{X})$, $u_{j'}(\mathbf{X}') > u_{j'}(\mathbf{X})$ or $u_n(\mathbf{X}') + u_{n'}(\mathbf{X}') > u_n(\mathbf{X}) + u_{n'}(\mathbf{X})$ holds.

Constraint 1 indicates that the utility of participating users or RB should not be reduced in a swap operation, and constraint 2 indicates that the utility of at least one participating user or RB is improved after the exchange. If the above conditions hold, the exchange request of the D2D is allowed, otherwise the current matching remains unchanged. The D2D channel allocation process is described in Algorithm 2.

Algorithm 2 D2D channel allocation algorithm based on swap matching (DASM)

Input: Randomly allocate the channels to D2D user and obtain an initialize the matching matrix \mathbf{X} , and donate Θ as the set of D2D users without available channels.

Output: \mathbf{X}^* , $\phi_{j,i}^*$ and $\phi_{j',j}^*$

- 1: For any D2D $_j \in \mathcal{J}$, search for another D2D $_{j'} \in \mathcal{J} \setminus \{j\} \cup \Theta$;
 - 2: If (j, j') satisfies the exchange condition and the constraint (17e) also satisfies, let $\mathbf{X} \leftarrow \mathbf{X}'$, and update Θ , otherwise the current channel matching remains unchanged.
 - 3: Repeat 1 and 2 until there is no D2D that satisfies the exchange condition.
-

Notice that the final channel matching is a stable exchange matching. This is because, assuming that there still

exists a D2D that satisfies the exchange condition in the last output matching state of Algorithm 2, the algorithm continues to iterate according to Algorithm 2. Therefore, there are no exchange pairs that satisfy the exchange condition in the final channel matching, and the algorithm reaches a stable state in the final channel matching. This also indicates that the initialization of the algorithm does not affect the final allocation result. Moreover, the exchange matching algorithm converges within a finite number of iterations. This is because exchange operations are only performed when the exchange condition is met. Furthermore, if $u_j(\mathbf{X}') > u_j(\mathbf{X})$, $u_{j'}(\mathbf{X}') > u_{j'}(\mathbf{X})$ and $u_n(\mathbf{X}') + u_{n'}(\mathbf{X}') > u_n(\mathbf{X}) + u_{n'}(\mathbf{X})$ holds, it implies that the system throughput is not decreased. On the other hand, the number of users is limited, which indicates the number of swap operations is limited. Therefore, when the exchange operation stops, the sum rate of the system is saturated. In the next section, we will evaluate the convergence of Algorithm 2.

3.3 DT Power Allocation

In the previous two sections, we optimize the integer variable ρ and ϕ by greedy algorithm and swap matching respectively, and obtain the optimized value ρ^* and ϕ^* . In this section, we further optimize the power allocation of DTs based on the given DT-DRs grouping and channel allocation, and decompose the power optimization into n subproblems where the power of DT is assigned in each RB separately. The original problem is rewritten as:

$$\max_{\alpha} R_{i+f}(\alpha) \text{ s.t. (17h), (17i)} \quad (20)$$

To solve the problem (20), we propose a local search algorithm to optimize the power of DTs. Define L_n as the number of D2Ds in n -th RB, the total power allocation vector of DTs is represented by $P = [P_1, \dots, P_j, \dots, P_J]$, and $P_j = [P_{j1}, P_{j2}, \dots, P_{jL_n}]$ is the power allocation vector of DTs in RB n . Denote the sum rate in n -th RB of the current step as $\eta(\mathbf{P}_n)$, $\eta'(\mathbf{P}_n)$ is sum rate of the last power search in RB n , and ε indicates the iteration threshold. We solve the transmission power of DTs by $P_j^* = \underset{P_j}{\operatorname{argmax}} \eta(\mathbf{P}_n)$, where the transmit power of other DTs is fixed. After P_j^* is obtained, we continue to optimize the intra-group power allocation of DT in the D2D NOMA group. The proposed power allocation algorithm based on local search is described in Algorithm 3.

3.4 Joint Optimization for Sum Rate

Combining Algorithms 1, 2 and 3, we obtain the joint user grouping and resource allocation framework, as shown in Algorithm 4, which is used to improve user access rate and sum-rate in dense ultra-dense communication environments.

4. Simulation Results and Analysis

Here, we analyze the performance of the proposed scheme

Algorithm 3 Local Power Search Algorithm (LPSA)

Input: Power distribution vector $\alpha_{j,f} = 1$, $\alpha_{j,f_s} = 0.5$, $\alpha_{j,f_w} = 0.5$.
Power search list for DTs.

Output: α^*

```

1: while true do
2:   for  $n=1$  to  $N$  do
3:     for  $j=1$  to  $L_n$  do
4:        $P_j^* = \underset{P_{j,R_i^n \geq R_i^{n,min}}}{\operatorname{argmax}} \eta(\mathbf{P}_n)$ 
5:     end for
6:   end for
7:   if  $\eta(\mathbf{P}_n) - \eta'(\mathbf{P}_n) < \varepsilon$  then
8:     break
9:   end if
10: end while
11: for  $j=1$  to  $J$  do
12:   if  $m_j = 1$  then
13:      $(\alpha_{j,f_s}^*, \alpha_{j,f_w}^*) = \underset{(\alpha_{j,f_s}, \alpha_{j,f_w})}{\operatorname{argmax}} \eta(\alpha_{j,f_s} P_j^*) + \underset{(\alpha_{j,f_s}, \alpha_{j,f_w})}{\operatorname{argmax}} \eta(\alpha_{j,f_w} P_j^*)$ 
14:   end if
15: end for

```

Algorithm 4 The Joint Optimization of System Sum Rate Based on DT-DR Grouping and Resource Allocation (DGRA)

Input: Randomly generate user locations. Initialize $\beta_{\mathcal{J} \times \mathcal{F}} = 0$. The channel allocation vector $\phi_{j,i} = \{\}$ and $\phi_{j',j} = \{\}$. Power distribution vector $\alpha_{j,f} = \{\}$.

Output: ρ^* , ϕ^* and α^*

- 1: Find ρ^* according to algorithm 1, where ϕ and α are fixed.
- 2: Find ϕ^* according to algorithm 2, where $\rho = \rho^*$ and α is fixed.
- 3: Find α^* according to algorithm 3, where $\rho = \rho^*$ and $\phi = \phi^*$.

Table 1 Simulation parameters and values.

Parameters	Values
number of CUEs / I	4
number of DTs / J	6-10
number of DRs / F	16:2:24
number of RB / N	30
noise power density	-174dBm/Hz
maximum DT-DR distance / d_{max}	40m
DT maximum transmit power / P_j^{max}	23dBm
CUE transmit power / P_i^{max}	23dBm
CUE minimum rate / $R_i^{n,min}$	2bps/Hz
bandwidth / B	15kHz
path loss constant / G	0.01
path loss exponent / θ	3
iterative threshold / ε	0.01

for D2D-NOMA communication underlying cellular systems by simulations. In the simulation, the BS is located at the origin, CUEs, DTs and DRs are randomly distributed between 20 m and 200 m from the BS [21], [22]. The path-loss is calculated by $d^{-\theta}$ where d is the distance between the transmitter and the receiver [22]. The specific simulation parameters are shown in Table 1.

The five comparative algorithms are presented below, where all algorithms adopt the local search power allocation in Algorithm 3:

1. DGRA: The proposed Algorithm 4.
2. Greedy: joint GDGA and greedy channel allocation

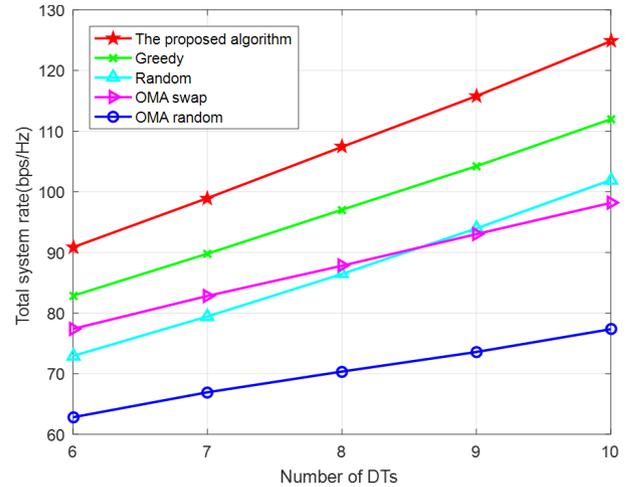


Fig. 2 System sum rate versus different number of DTs, with $F=20$.

algorithm.

3. Random: joint GDGA and random channel allocation algorithm.

4. OMA swap: the DT-DR grouping algorithm employing the OMA model and channel allocation through swapping match algorithm.

5. OMA random: the DT-DR grouping algorithm employing the OMA model and random channel allocation algorithm.

Figure 2 illustrates that the proposed algorithm (DGRA) has the highest sum rate compared with the other four comparison algorithms. When the number of DTs is 10, the proposed algorithm outperforms the Greedy algorithm, random algorithm, OMA swap algorithm, and OMA random algorithm by 11.5%, 22.5%, 27.16% and 61.4%, respectively. It is also observed that when $J=10$, the sum rate of the random algorithm is higher than that of the OMA swap algorithm. This is because the random algorithm accesses more DRs by adopting NOMA. Although the OMA swap algorithm reasonably allocates resources, the achievable rate is still not good enough. Next, in order to verify the advantages of the proposed grouping algorithm in Algorithm 1, the required RB number and the spectrum utilization ratio are compared in Figs. 3–5. In the figures, GDGA is the proposed Algorithm 1, where DTs can form D2D pairs or D2D NOMA groups with DRs. In Case 1, DTs are only allowed to form D2D NOMA groups with DRs, while in case 2, DTs can only form D2D pairs with DRs. In both case 1 and case 2, the greedy algorithm is used during the DT-DR grouping stage.

Figure 3 illustrates that the CUE rate of our proposed algorithm decreases as the number of DTs increases, and the CUE rate is always greater than the threshold value we set, 2 bps/Hz. This shows that although the CUE rate decreases with increasing interference, the performance is still guaranteed.

Figure 4 plots the CDF of the required RBs (the required subchannel) in the network when $J = 8$ and $F = 20$, based

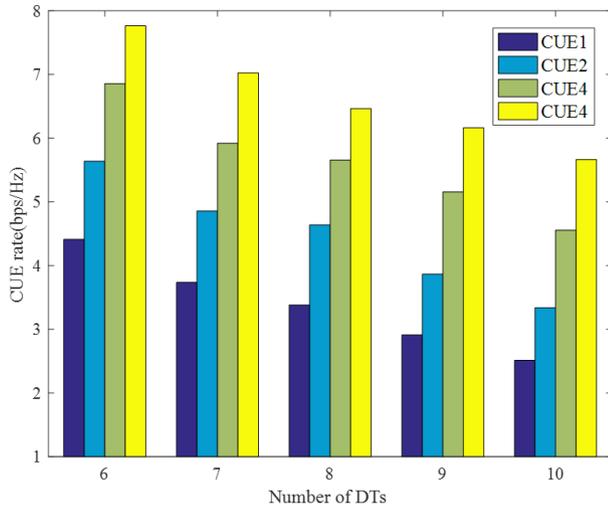


Fig. 3 CUE rate versus different number of DTs, with $F=20$.

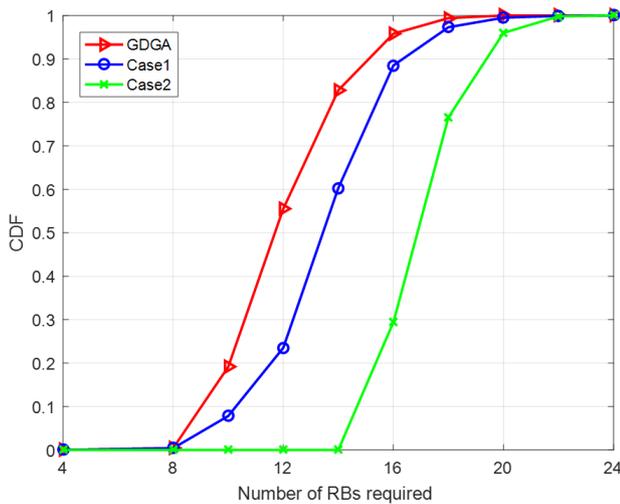


Fig. 4 CDF of required RBs.

on 1000 simulation results. It is observed that the RBs required by the proposed GDGA are significantly less than those required by case 1 and case 2, indicating the proposed model requires fewer spectrum resources. Therefore, GDGA has an absolute advantage compared with case 1 and case 2 in the RB tight scenario, because it can accommodate more users to access the network.

Figure 5 illustrates the variation of the required RBs in the network with the increase of DRs. It can be observed that as the number of DRs increases, the required RBs also increase. This is because, after the pairing of DRs and DTs in algorithm 1, the number of idle DRs increases, requiring separate spectrum resources for transmission, hence increasing the required RBs. It also shows that the required RBs of our proposed GDGA is significantly lower than case 1 and case 2. Furthermore, compared to case 1, in the proposed model, after DTs and DRs are paired into the D2D NOMA group, the remained DT continues to be paired into D2D pair with the DRs within its range. The D2D pairs use the

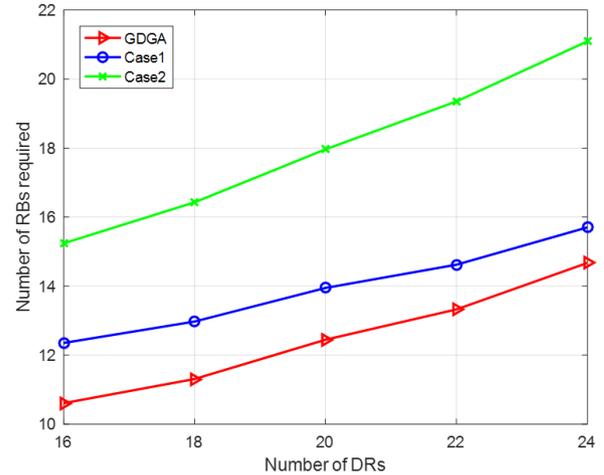


Fig. 5 Number of RBs required versus number of DRs.

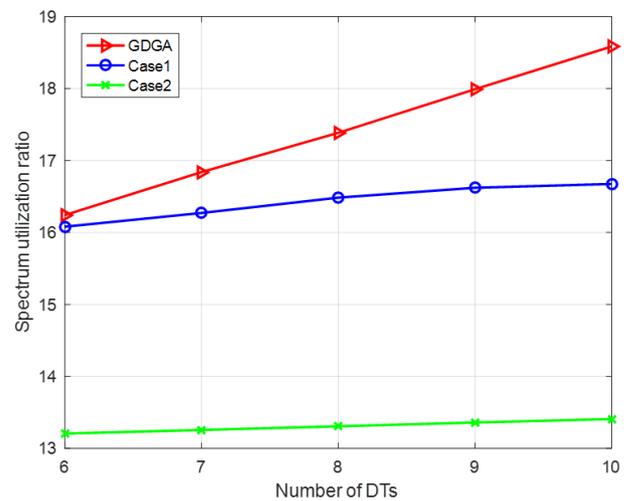


Fig. 6 Spectrum utilization ratio versus different number of DTs, with $F=20$.

resources of CUE using multiplexing mode without occupying additional idle spectrum, which saving resources for the network.

Figure 6 plots the spectrum utilization ratio in three different scenarios. Here, the spectrum utilization ratio is defined as the total system rate (including DTs and DRs under OMA mode) divided by the total number of RBs occupied. It can be observed that the spectrum utilization ratio increases with the increase of the number of DTs. This is because as the number of DTs increases, the number of DRs under the multiplexing mode also increases, resulting in a decrease in the number of RBs occupied. When $J=10$, the proposed GDGA significantly outperforms the other two cases because more DRs are connected to DTs, and the CUE spectrum can be reused to access the network, resulting in a decrease in the total number of occupied RBs. Additionally, it is observed that the growth rate of case 1 and case 2 is relatively slow. This is because, although increasing DTs brings about the access gain of DRs, the pairing patterns of DT and DR in case

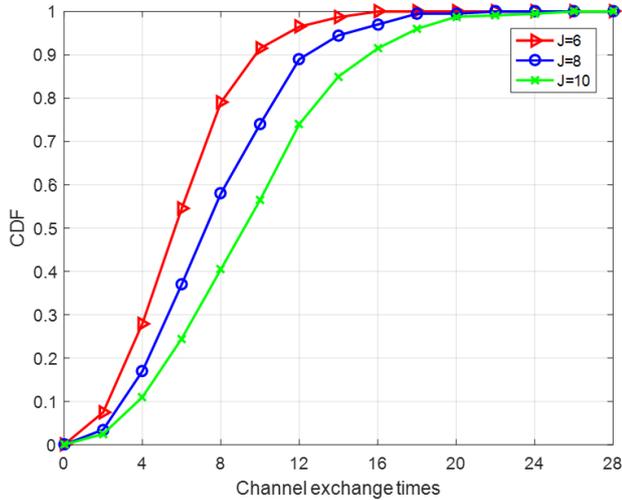


Fig. 7 CDF of the exchange times.

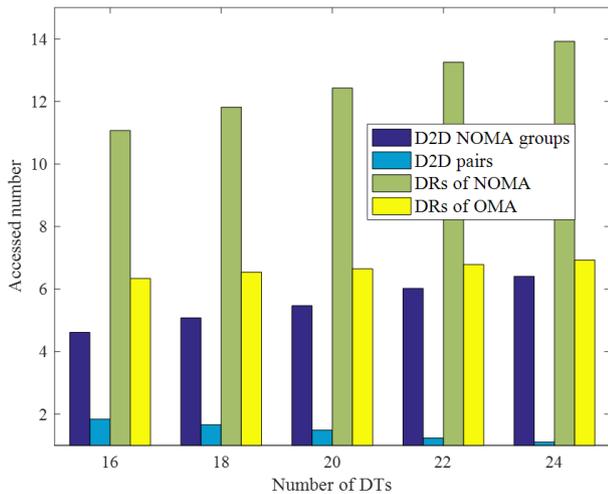


Fig. 8 Accessed number versus different number of DRs, with $J=8$.

1 and case 2 are fixed, and the unpaired DTs and DRs occupy numerous idle RBs, resulting in low spectrum utilization.

Figure 7 shows the CDF of the number of exchange time in Algorithm 2, where $F=20$. It shows that the proposed exchange matching algorithm converges after a finite number of iterations. In addition, as J increases, the number of exchanges also increases, because the more D2D users there are, the more D2D users meet the exchange conditions.

Figure 8 shows the access number of D2D NOMA groups, D2D pairs, DRs of NOMA, and DRs of OMA with $J=8$. As the number of DRs to be connected increases, the access number increases except for the D2D pairs. This is because there are more DRs available to be selected for DTs within range. In order to increase the access rate of DRs, DT and DR are preferentially paired into D2D NOMA group, which leads to a corresponding decrease in the number of D2D pairs. It also shows that, the proposed new model with NOMA has a significantly higher DR access rate than the OMA model. This is because, in the NOMA model, DTs

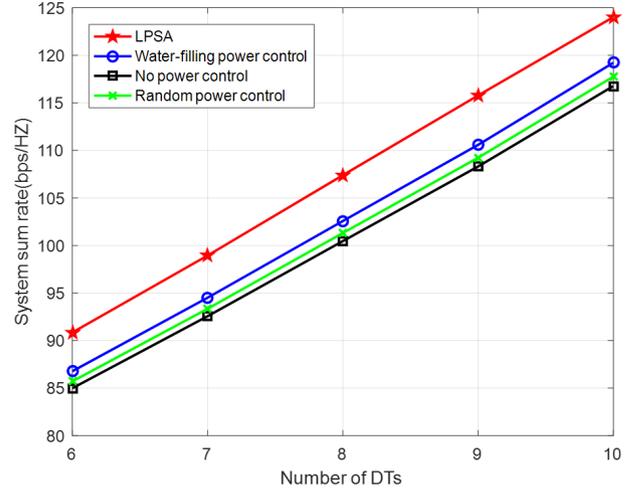


Fig. 9 Comparison of different power control algorithms.

can pair with DRs to form D2D NOMA groups or D2D pairs, while in the OMA model, DTs can only pair with DRs to form traditional D2D pairs.

In Fig. 9, we compare the performance of the proposed LPSA in Algorithm 3, water-filling power control algorithm, no power control algorithm and random power control algorithm. It can be observed that the proposed LPSA power algorithm significantly improves the sum rate compared to the other three methods. When $J=10$, the performance of the proposed LPSA is 4.9%, 6.5%, and 7.1% higher than the water injection power control algorithm, random power control algorithm, and no power control algorithm, respectively.

5. Conclusion

This paper investigates the uplink wireless resource allocation problem in NOMA-D2D communication systems and proposes a system model that allows both D2D pairs and D2D NOMA groups to coexist. For the case where multiple D2D pairs can reuse the sub-channel of the cellular user, we propose a problem of maximizing the sum rate while ensuring the minimum rate of CUEs, where we focus on multi-variable design including user classification parameter, sub-channel allocation and power allocation. To address this problem, we propose a joint user grouping and resource allocation scheme. Specifically, we first use a greedy algorithm to pair DTs and DRs, and then use the exchange matching algorithm to find the optimal RB for D2D. Finally, the power of DT is optimized by the local power search algorithm. Simulation results show that the proposed scheme improves the spectrum utilization ratio and DR access rate significantly, and achieves a higher sum rate than the OMA scheme.

References

- [1] R.M. Radaydeh, F.S. Al-Qahtani, A. Celik, K.A. Qaraqe, and M.S. Alouini, "Generalized imperfect D2D associations in spectrum-shared cellular networks under transmit power and interference con-

- straints,” *IEEE Access*, vol.8, pp.182517–182536, 2020.
- [2] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, “Design aspects of network assisted device-to-device communications,” *IEEE Commun. Mag.*, vol.50, no.3, pp.170–177, 2012.
 - [3] J. Lee and J.H. Lee, “Performance analysis and resource allocation for cooperative D2D communication in cellular networks with multiple D2D pairs,” *IEEE Commun. Lett.*, vol.23, no.5, pp.909–912, 2019.
 - [4] D. Feng, L. Lu, Y. Yuan-Wu, G.Y. Li, G. Feng, and S. Li, “Device-to-device communications underlying cellular networks,” *IEEE Trans. Commun.*, vol.61, no.8, pp.3541–3551, 2013.
 - [5] G. Yu, L. Xu, D. Feng, R. Yin, G.Y. Li, and Y. Jiang, “Joint mode selection and resource allocation for device-to-device communications,” *IEEE Trans. Commun.*, vol.62, no.11, pp.3814–3824, 2014.
 - [6] L. Lei, Y. Kuang, N. Cheng, X. Shen, Z. Zhong, and C. Lin, “Delay-optimal dynamic mode selection and resource allocation in device-to-device communications — Part II: Practical algorithm,” *IEEE Trans. Veh. Technol.*, vol.65, no.5, pp.3491–3505, 2015.
 - [7] J. Liu, Y. Kawamoto, H. Nishiyama, N. Kato, and N. Kadowaki, “D2D-to-device communications achieve efficient load balancing in LTE-advanced networks,” *IEEE Wireless Commun.*, vol.21, no.2, pp.57–65, 2014.
 - [8] H. Nishiyama, M. Ito, and N. Kato, “Relay-by-smartphone: Realizing multihop device-to-device communications,” *IEEE Commun. Mag.*, vol.52, no.4, pp.56–65, 2014.
 - [9] Y. Dai, M. Sheng, J. Liu, N. Cheng, X. Shen, and Q. Yang, “Joint mode selection and resource allocation for D2D-enabled NOMA cellular networks,” *IEEE Trans. Veh. Technol.*, vol.68, no.7, pp.6721–6733, 2019.
 - [10] H. Sun, D. Zhai, Z. Zhang, J. Du, and Z. Ding, “Channel allocation and power control for device-to-device communications underlying cellular networks incorporated with non-orthogonal multiple access,” *IEEE Access*, vol.7, pp.168593–168605, 2019.
 - [11] C. Ma, W. Wu, Y. Cui, and X. Wang, “On the performance of successive interference cancellation in D2D-enabled cellular networks,” 2015 *IEEE Conference on Computer Communications (INFOCOM)*, pp.37–45, IEEE, 2015.
 - [12] J. Zhao, Y. Liu, K.K. Chai, Y. Chen, and M. ElKashlan, “Joint subchannel and power allocation for NOMA enhanced D2D communications,” *IEEE Trans. Commun.*, vol.65, no.11, pp.5081–5094, 2017.
 - [13] I. Budhiraja, N. Kumar, and S. Tyagi, “ISHU: Interference reduction scheme for D2D mobile groups using uplink NOMA,” *IEEE Trans. Mobile Comput.*, vol.21, no.9, pp.3208–3224, 2021.
 - [14] I. Budhiraja, N. Kumar, and S. Tyagi, “Cross-layer interference management scheme for D2D mobile users using NOMA,” *IEEE Syst. J.*, vol.15, no.2, pp.3109–3120, 2020.
 - [15] L.P. Qian, Y. Wu, H. Zhou, and X. Shen, “Joint uplink base station association and power control for small-cell networks with non-orthogonal multiple access,” *IEEE Trans. Wireless Commun.*, vol.16, no.9, pp.5567–5582, 2017.
 - [16] D. Zhai, R. Zhang, L. Cai, B. Li, and Y. Jiang, “Energy-efficient user scheduling and power allocation for NOMA-based wireless networks with massive IoT devices,” *IEEE Internet Things J.*, vol.5, no.3, pp.1857–1868, 2018.
 - [17] B. Di, L. Song, Y. Li, and Z. Han, “V2X meets NOMA: Non-orthogonal multiple access for 5G-enabled vehicular networks,” *IEEE Wireless Commun.*, vol.24, no.6, pp.14–21, 2017.
 - [18] Y. Pan, C. Pan, Z. Yang, and M. Chen, “Resource allocation for D2D communications underlying a NOMA-based cellular network,” *IEEE Wireless Commun. Lett.*, vol.7, no.1, pp.130–133, 2017.
 - [19] C. Kai, Y. Wu, M. Peng, and W. Huang, “Joint uplink and downlink resource allocation for NOMA-enabled D2D communications,” *IEEE Wireless Commun. Lett.*, vol.10, no.6, pp.1247–1251, 2021.
 - [20] M.M. Al-Wani, A. Sali, S.D. Awad, A.A. Salah, Z. Ding, N.K. Noordin, S.J. Hashim, and C.Y. Leow, “Interference cancellation via D2D CSI sharing for MU-MISO-NOMA system with limited feedback,” *IEEE Trans. Veh. Technol.*, vol.70, no.5, pp.4569–4584, 2021.
 - [21] S. Solaiman, L. Nassef, and E. Fadel, “User clustering and optimized power allocation for D2D communications at mmWave underlying MIMO-NOMA cellular networks,” *IEEE Access*, vol.9, pp.57726–57742, 2021.
 - [22] C. Kai, H. Li, L. Xu, Y. Li, and T. Jiang, “Joint subcarrier assignment with power allocation for sum rate maximization of D2D communications in wireless cellular networks,” *IEEE Trans. Veh. Technol.*, vol.68, no.5, pp.4748–4759, 2019.



Jin Xie received the B.Eng degree in communication engineering from Northeast Electric Power University, Jilin, China, in 2021. He is currently pursuing the M.E. degree in information and communication engineering in Hangzhou Dianzi University, Hangzhou.



Fangmin Xu received the Ph.D. degree in communication engineering from Beijing University of Posts and Telecommunications in 2008. Currently, she is a faculty member at Hangzhou Dianzi University, China.