

# Expand the Range of Traffic Offloading with Multi-Hop D2D Communication

Xuewen Yan<sup>1, a</sup>, Chao Dong<sup>1</sup> and Yuben Qu<sup>2</sup>

<sup>1</sup>Army Engineering University of PLA, Nanjing, Jiangsu, 210007, China

<sup>2</sup>Rocket Force University of Engineering, Xi'an, Shanxi, 710025, China

<sup>a</sup>yxw522501@163.com

**Abstract.** Device-to-device (D2D) communication is a promising technique for traffic offloading in next-generation cellular systems, in response to the sharp increase in traffic and uneven distribution. In this paper, in the case of meeting the user's delay, we propose a multi-hop D2D traffic offloading scheme to expand the range of traffic offloading. In consideration of the load balancing problem, we introduce load fairness of Base Stations to restraint the traffic that is offloaded to each BS. And, based on the constraints of delay and fairness, we formulate an optimization model that maximizes the range of offloading. We also proposed joint algorithm to achieve traffic offloading. And simulation results show that our scheme can further expand the range of traffic offloading under the condition of ensuring load fairness.

**Keywords:** Multi-hop D2D; Traffic Offloading; Load Fairness.

## 1. Introduction

According to Cisco's annual Visual Networking Index, over the next five years (2016-2021), the number of connected devices will exceed threefold the world population, by 2021. The mobile data traffic, for its part, is expected to reach 48.3 exabytes per month by the same year, accounting for 17% of the total traffic [1]. Especially when some hot events happen, some base stations (BSs) have to face high traffic requests, which might lead to BS congested and cause uneven distribution of traffic across the cellular network.

To handle the sharp increase and the uneven distribution in traffic, operators need an economical plan that exempt them from colossal investment in building up new networks. One of the promising solutions for improving network performance is traffic offloading, which exploits complementary network technologies to deliver data more efficiently [2]. Device-to-Device (D2D) communication is the key technique in traffic offloading, where devices communicate with each other directly by reused cellular network spectrum resources [3].

Nowadays, in the cellular network, there are always several BSs that are congested due to users surge, while other nearby BSs are idle, which can cause local congestion of the cellular network and severely unbalancing traffic distribution. In the current study, multi-hop D2D communication can effectively solve this challenge of cellular networks [4]. Through multi-hop D2D communication, traffic can be offloaded from the congestion BSs to the adjacent idle BSs (IBSs) to alleviate the pressure on the cellular network and achieve load balancing. However, with the further increase of users, the neighboring cellular network resources cannot meet the demand. It is necessary to further expand the offloading range, which would offload the traffic to the more distant IBSs. At the same time, after the traffic is offloaded, the load balancing of each BS is also guaranteed. Otherwise, the traffic is easily aggregated to some BSs, resulting in uneven traffic distribution.

Therefore, this paper proposes a multi-hop D2D traffic offloading scheme based on constraints load balancing. In the case of meeting the requirement of system delay, the users in the congested BSs can transfer the traffic to the idle BSs of a longer distance as much as possible through the multi-hop D2D relay communication, in order to enhance the range of offloading traffic. Meanwhile, the scheme also ensures that the traffic offloaded to each IBS is fair, which would make the system load balanced.

This article has three main contributions:

- Based on the constraint of system delay, a multi-hop D2D relay communication scheme is proposed. The hop count of the D2D communication link takes the maximum value as much as possible under the condition that the delay constraint is satisfied, so that the traffic is transferred in larger range.

- In addition, in order to ensure that traffic will not be only offloaded to some IBSs, we introduce fairness constraint. This condition ensures that traffic is offloaded to each reachable IBS as fair as possible, which can achieve load balancing.

- The simulation results show that the proposed scheme can effectively offload traffic to farther BSs meeting the fairness of the BSs load, which not only effectively expands the range of traffic offloading, but also achieves load balancing on larger scale.

The rest of this paper is organized as follows. Section 2 is the related work. Section 3 presents the system model. Section 4 introduces our algorithm. Then, in Section 5, we give the simulation result and performance analysis. Furthermore, we give the conclusion and further work in Sections 6.

## 2. Related Work

In this section, we'll cover some of the related work in this article. There has been a lot of researches on how to relieve cellular network pressure through D2D communication. Jiang et al. [5] used D2D relay communication to achieve load balancing between a macro base station (MBS) and several femto-based stations (FBSs) in HetNets. In [6], author proposes a metric to ensure the stability of the D2D relay offloading process, which is a very reusable performance metric in D2D relay communications. Liu et al. [7] first proposed using D2D relay communication between different BSs to achieve load balancing. And a detailed algorithm design has been proposed. The experimental results show that the method can effectively improve the throughput and access rate of cellular networks. The authors in [8] describe a communication process for load balancing between BSs. Users of the overloaded BS are relayed to the nearby idle BSs by the two-hop relay communication. The load balancing between the BSs can be achieved. Results show that D2D relay between BSs can be beneficial for the network performance enhancement in terms of per-user throughput and network load balancing. By an actual experimental observation, Deng et al. [9] described in detail how D2D relay communication achieves load balancing between different BSs and how to improve cellular network spectrum efficiency. The above work implements traffic offloading through D2D communication, but there is no specific analysis of the network performance.

In [10], authors formulate the problem of jointly optimizing the power of small BSs (SBSs), SBS density, and the fraction of spectrum allocated for D2D communication in order to maximize system throughput for emergency situations such as rescue missions. Alwan et al. [11], authors address the joint multicast routing and OFDM resource allocation problem in the D2D network to accomplish the offloading task. In [12], we consider a scenario in which a common file is requested by a subset of users in different times and with different maximum tolerable delays, which ensures that traffic is offloaded to farther area. However, the above work does not take into account the load fairness of each BS after traffic offloading, so we propose multi-hop D2D traffic offloading scheme based on system delay constraints.

## 3. System Model

We consider an OFDMA network with  $S$  BSs randomly distributed in a region, where some BSs are congested, and the rest are idle, as illustrated in Fig.1. Users mainly perform uplink communication. All NCBS congested BSs form set  $CBS$ , and all  $N_{IBS}$  idle BSs constitute set  $IBS$ . The congested BSs need to offload traffic of some users to idle BSs to relieve their own pressure and achieve load balancing between BSs. The traffic offloading technology is mainly D2D relay communication, and user performed the D2D relay communicates by multiplexing the spectrum resources of the cellular network users in the idle BS. And we use set  $DUE$  to represent all  $N_{DUE}$

users in congested BS that need to perform D2D relay communication. As described in Fig.2,  $d_s$  is a user who needs to perform D2D communication in congested BS,  $IBS$  is the target BS for traffic offloading, and  $d_1, \dots, d_k$  are the relay nodes.  $L_1, \dots, L_k$  are each hop D2D link in a multi-hop D2D relay communication link, and  $L_{k+1}$  is the last hop cellular link in the multi-hop link. Traffic is offloaded from the congested BS to idle BS through this D2D relay communication link.

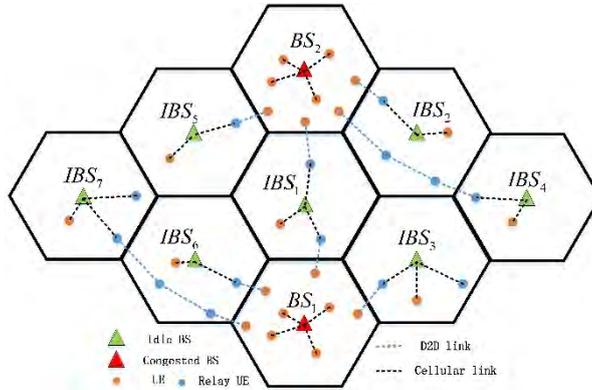


Fig.1 Multi-hop D2D relay communication offload traffic.

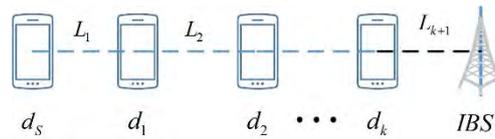


Fig.2 Traffic offload link.

In this scenario, the D2D users need to manage two kinds of interferences caused by spectrum multiplexed: (a). Interference from cellular users at the D2D receivers; (b). Interference from D2D transmitters at the cellular BSs.

Under Rayleigh fading channels, the link outage probability between a D2D transmitter,  $d_t$ , and a D2D receiver,  $d_r$ , is given by [13]:

$$p_{tr}^{out} = 1 - \exp\left(-\frac{10^{\frac{P_r^{int} + 10\eta\log D_{tr} + 10\log\gamma_{th} - P_t^{max}}{10}}}{2\sigma^2}\right) \quad (1)$$

where  $P_r^{int}$  is the cellular interference and noise power measured at  $d_r$ ,  $D_{tr}$  is the distance between  $d_t$  and  $d_r$ ,  $t \in \{d_s, d_1, \dots, d_k\}$  and  $r \in \{d_1, \dots, d_k, IBS\}$ .  $P_t^{max}$  is the maximum allowed transmit power for node  $d_t$  and is determined by  $\gamma_{bs}$ ,  $p_b$  and the distance to the target BS, as follows:

$$P_t^{max} = 10\log\left(-\frac{10^{\frac{\gamma_{bs}}{10}} D_{t,b}^\eta}{2\sigma_{t,b}^2 \ln p_b}\right) \quad (2)$$

where  $\gamma_{th}$  is the least SINR at D2D receivers,  $\eta$  denotes the path loss exponent and  $D_{t,b}$  denotes the distance between the  $t^{th}$  D2D node and other BS.  $\sigma_{t,b}^2$  is average power gains. Interference caused by a D2D transmitter at other BS should be below a threshold  $\gamma_{bs}$  dB with probability at

least  $1 - p_b$ .  $p_b$  is interference outage probability, that is, the probability that the cellular network user providing spectrum multiplexed will be interrupted by the interference of the D2D transmitters when this user communicates with the BS.

By the link outage probability of D2D communication, we can find the average delay of a received packet in the link of traffic offloading in Fig.2. We begin by finding the following conditional probability of packet being delivered across the  $k^{th}$  link in the  $k^{th}$  attempt, given that the packet is successfully received within the available  $\theta$  attempts.

$$P_{k,n}^{sus} = \frac{(1 - p_k) p_k^{n-1}}{\sum_{n=1}^{\theta} (1 - p_k) p_k^{n-1}} \quad (3)$$

where  $p_k$  is the link outage probability  $p_{tr}^{out}$  in Eq.(1).

So the average delay on the  $k^{th}$  link is

$$w_k^{recv} = \sum_{n=1}^{\theta} n P_{k,n}^{suc} \quad (4)$$

Hence, the total delay of the received packet due to all the  $K + 1$  links in the route from  $d_s$  to  $IBS$  is given by

$$w_k^{recv} = \sum_{n=1}^{\theta} n P_{k,n}^{suc} \quad (5)$$

And, by the given the least SINR of link  $\gamma_{th}$ , we can deduce the throughput of each link of traffic offloading.

$$T = B \log_2(1 + \gamma_{th}) \quad (6)$$

Now, by the delay and throughput expressions (4) and (5), we will give optimization model of the system to maximize the range of traffic offloading. We use  $y_{l,i}$  to represent a multi-hop D2D communication link from a user  $UE_l$  in the congested BS that needs to offload traffic to the idle BS  $IBS_i$ . We assume that  $\varepsilon$  is the lowest delay to meet the user's communication needs, so if  $w^{recv} < \varepsilon$ ,  $y_{l,i} = 1$  indicates that the link is successfully established, and otherwise  $y_{l,i} = 0$ . Next, we use  $x_i$  to indicate whether the congested BS is associated to the idle BS  $i$  via a multi-hop D2D communication link. If  $\sum_{l=1}^{N_{DUE}} y_{l,i} > 0, \forall i \in IBS$ , then  $x_i = 1$ , otherwise  $x_i = 0$ . In other words, if

there is a user in the congested BS associated with the idle BS  $i$ , it means that the congested BS can offload the traffic to the idle BS  $i$ . Our goal is to expand the number of idle BSs that can be associated with the congested BS to increase the range of traffic offloading.

This goal has two important constraints:

- By traffic offloading, we should ensure that the congested base BS is no longer congested. We assume that the data congestion BS needs to process is  $S_r$ , and the data that the BS is able to process is  $S_{th}$ . This constraint can be derived by Eq.(5). As follows:

$$S_r - \sum_{i=1}^{N_{IBS}} \sum_{l=1}^{N_{DUE}} T \cdot y_{l,i} \leq S_{th} \quad (7)$$

• Ensure that the data offloaded to each of the idle BSs is balancing. Here we introduce the Jains Fairness Index mentioned in [14]. The data offloaded to each idle BS can be expressed as follows:

$$f_i = \sum_{l=0}^{N_{DUE}} T \cdot y_{l,i}, \quad i \in IBS \quad (8)$$

The fairness metric is expressed as follows:

$$J = \frac{\left( \sum_{i=1}^{N_{IBS}} f_i \right)^2}{\sum_{i \in IBS} x_i \cdot \sum_{i=1}^{N_{IBS}} f_i^2} \quad (9)$$

The value of  $J$  is between 0 and 1, and the larger the  $J$ , the more balancing the traffic offloaded to each IBS. Moreover, we set fairness threshold  $J_{th}$ , and only when  $J$  is greater than  $J_{th}$ , the traffic offloading process is balancing.

Our optimization goal is to maximize the range of traffic offloading, which is to maximize the number of idle BSs that can be associated with congested BS. Combined with (8),(9), we can get the optimization model as follows:

$$\max \sum_{i=1}^{N_{IBS}} x_i \quad (10)$$

$$\text{s.t.} \quad S_r - \sum_{i=1}^{N_{IBS}} \sum_{l=1}^{N_{DUE}} T \cdot y_{l,i} \leq S_{th} \quad (11)$$

$$J \geq J_{th} \quad (12)$$

$$y_{l,i} \in \{0, 1\}, l \in DUE, i \in IBS \quad (13)$$

$$x_i \in \{0, 1\}, i \in IBS \quad (14)$$

The objective (10) maximizes the number of idle BSs that can be associated with congested BS. Constraint (11) ensures that the BS is no longer congested after traffic offloading. Constraint (12) ensures that the traffic offloaded to each of the idle BSs is balancing. Constraint (13) and constraint (14) determine the range of values for the two-dimensional variables  $y_{l,i}$  and  $x_i$ . The optimization formulization that expand range of traffic offloading by maximizing the numbes of idle BSs that can be associated with congested BS is an mixed integer programming (MIP) problem. MIP problem is always NP-complete due to the binary variables [7]. So in the next section, we will give an algorithm of traffic offloading to find solution for this model.

## 4. Joint Algorithm Achieves Efficient Traffic Offloading

In this part, we propose the following traffic offloading algorithm to alleviate the congested state of the system. The main idea is to achieve traffic offload by enumerating all possible multi-hop D2D communication links, and then selecting the appropriate D2D links by constraints. Algorithm 1 mainly establishes multi-hop D2D links from each user with poor communication quality in congested BS to all IBSs. And, we calculate the delay of the possible multi-hop link for each multi-hop link, where links with delay greater than  $\varepsilon$  will be cancelled. All BSs constitute set  $B(b_1, b_2, \dots, b_n)$ , and users in each BS constitute set  $C(c_1, c_2, \dots, c_m)$ .

Algorithm 2 mainly selects the most suitable D2D communication links by the load fairness constraint. First, for each user with poor communication quality, Algorithm 2 selects the multi-hop D2D link with maximum delay to offload traffic and calculates fairness  $J$ . Next, if  $J < J_{th}$ , the user selects the second largest delayed communication link until  $J > J_{th}$ .

---

**Algorithm 1** Construct multihop D2D links.

---

```

1:  $B(b_1, b_2, \dots, b_n), C(c_1, c_2, \dots, c_m), S, S_r, \varepsilon.$ 
2: for  $\forall j \in B$  do
3:   if  $S_{b_j} > S_{th}$  then
4:      $b_j$  is congested.
5:     for  $\forall c_i \in C$  do
6:       if  $SINR_{i,b} < SINR_{th}$  then
7:          $c_i$  is the user who needs to perform D2D relay communication.
8:         These users constitute set  $D_{tr}$  of D2D senders.
9:       end if
10:    end for
11:  end if
12: end for
13: for  $\forall i \in D_{tr}$  do
14:   for  $\forall j \in IBS$  do
15:     Find all possible links from D2D senders  $i$  to the idle BSs  $j$ .
16:     Calculate the delay  $w_{i,j}$  for each link.
17:     if  $w_{i,j} > \varepsilon$  then
18:       Cancel this link.
19:     end if
20:   end for
21: end for

```

---



---

**Algorithm 2** Construct multihop D2D links.

---

```

1: for  $\forall i \in D_{tr}$  do
2:   Select the available links with the longest delay.
3: end for
4: Calculate the  $J$  according to Eq.(9).
5: if  $J < J_{th}$  then
6:   The constructed links meet the model needs.
7: else
8:   for  $\forall i \in D_{tr}$  do
9:     Select the communication link with the next largest delay.
10:    Go to 4.
11:    Jump out of the loop until the selected links meets the system's interference requirements.
12:   end for
13: end if

```

---

## 5. Simulation and Performance Analysis

In this section, we will provide simulation for the performance of the proposed joint load balancing scheme. We consider an OFDMA-based cellular which comprises of one 2 central BSs and 17 adjacent idle BSs, as depicted in Fig.3. The distance between BSs is 433 m. And, the coverage of the BS is regular hexagon with radius of 250m. Users of central BS are 300. The users of idle BS are 100. All the users are randomly distributed within the coverage of the BS. The total number of cellular resources of BS is 200, and the size of each resource is 250KHz. These simulation parameters are chosen in accordance with 3GPP LTE regulation for OFDMA system [15]. Transmit power of user is 23dBm, and noise power spectrum density is 174dBm/Hz.

In Fig.4a, the horizontal axis represents the number of each idle BS, and the vertical axis represents the number of users associated with each idle BS through multi-hop in the congestion BS. And, “MaxDelay” we proposed indicates that in the optional communication link, the user selects the multi-hop link with the longest delay to perform traffic offloading, while fairness threshold  $J_{th} = 0$ . “Fairness” indicates that the users uses the fairest traffic offloading scheme to offload traffic, which is the optimizing scheme. “MinDelay” indicates the user selects the multi-hop link with the lowest delay to offload traffic, in the optional links of user.

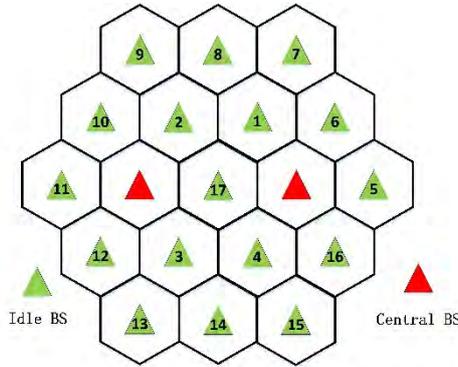


Fig.3 Simulation scenario.

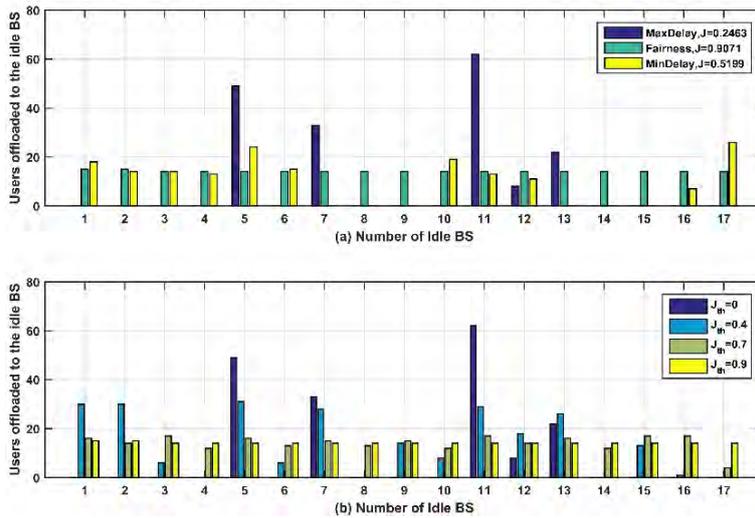


Fig.4 a) Comparison of offloading performance of different scenarios; b) Performance comparison of the proposed scheme under different fairness thresholds.

From the figure, we can know that when “MaxDelay” is selected for traffic offloading, the traffic is offloaded to the IBS {5,7,11,12,13}, and Load Fairness Index  $J = 0.2463$ . in addition, traffic of many users is offloaded to some idle BS, e.g., IBS 11. The offloading process is very unbalancing. When we offload traffic by “MinDelay”, the traffic is offloaded to the IBS {1,2,3,4,5,6,10,11,12,16,17}, and  $J = 0.5199$ , that is, it is fair to offload traffic to these IBSs. The offloading process is balancing. When we use “Fairness” to offloading traffic of central BSs, the traffic is offloaded to all IBSs, and  $J = 0.9071$ , which indicates traffic offloaing is very balancing. However, our goal is to expand the range of traffic offloading while maintaining load balancing. In scheme “MinDelay”, although the traffic is offloaded to many IBSs, these BSs are all in close proximity to the central BSs. But, in the “Max Delay” we proposed, the users can offload traffic to more distant BSs, such as, IBS7, IBS13, which effectively expands the range of traffic offloading. And the fairness threshold is only 0. If we increase the fairness threshold  $J_{th}$ , we can foresee that the traffic offloading would become more balanced.

In Fig.4b, we use scheme “MaxDelay” to offload traffic with different fairness thresholds  $J_{th}$ . As can be seen from the figure, when  $J_{th}=0$ , the traffic can be offloaded to the IBS {5,7,11,12,13}. When  $J_{th}=0.4$ , the traffic can be offloaded to the IBS {1,2,3,5,6,7,9,10,11,12,13,15,16}. When  $J_{th}=0.7$ , the traffic can be offloaded to all IBSs. At last, when  $J_{th}=0.9$ , the performance of traffic offloading is very close to the optimizing scheme “Fairness”. Therefore, our proposed scheme “MaxDelay” can effectively expand the range of traffic offloading under the condition of ensuring fairness.

## 6. Summary

This work presents a delay-based multi-hop D2D communication traffic offloading scheme. The traffic of the congested BSs is offloaded to the adjacent IBSs by the multi-hop D2D link with the largest delay to expand the range of traffic offloading. Under the constraint of BS load fairness, an optimization model for maximizing the range of traffic offloading is proposed, and a joint traffic offload algorithm is proposed to find the optimal solution. By this algorithm, in congested BS, users with poor communication quality select the most suitable multi-hop link to offload traffic to the IBSs based on delay and fairness constraints. The results show that our algorithm can maximize the range of traffic offload, relieve the pressure on the congested BSs, and achieve fair traffic offloading under the condition of ensuring load fairness. In a future extension of this work, we will combine the D2D communication in the BSs and the multi-hop D2D communication between the BSs to offload the traffic.

## References

- [1]. Cisco Visual Networking Index, “Cisco Visual Networking Index: Forecast and Methodology, 2016–2021,” Tech. Rep., 2017. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/serviceprovider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf>.
- [2]. F. Rebecchi, M. D. De Amorim, V. Conan, A. Passarella, R. Bruno, and M. Conti, “Data offloading techniques in cellular networks: a survey,” *IEEE Commun. Surveys & Tutorials*, vol. 17, no. 2, pp. 580–603, Nov. 2015.
- [3]. Arash Asadi, Qing Wang, Vincenzo Mancuso. A Survey on Device-to-Device Communication in Cellular Networks [J]. *IEEE Communications Surveys and Tutorials*, 2014, 16(4): 1801-1819.
- [4]. G. Nardini, G. Stea, and A. Virdis, “A Fast and Reliable Broadcast Service for LTE-Advanced Exploiting Multihop Device-to-Device Transmissions,” *Future Internet*, vol. 9, no. 4, p. 89, 2017.
- [5]. Yuqing Li, Bingyu Shen, Jinbei Zhang, et al. Offloading in HCNs: Congestion-Aware Network Selection and User Incentive Design [J]. *IEEE Transactions on Wireless Communications*, 2017, 16(10): 6479-6492.
- [6]. Yu Zeng, Honglin Hu, Tianheng Xu, et al. User Pairing Stability in D2D-Relay Networks [J]. *IEEE Communications Letters*, 2017, 21(10): 2278-2281.
- [7]. Liu J, Kawamoto Y, Nishiyama H, et al. Device-to-device communications achieve efficient load balancing in LTE-advanced networks[J]. *IEEE Wireless Communications*, 2014, 21(2): 57-65.
- [8]. Juan-Bautista Tomas-Gabarron; Yacine Ghamri-Doudane. Distributed load balancing by two-hop relaying in LTE-Advanced networks[C]//2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), 2014: 337-345.
- [9]. Lei Deng, Ying Zhang, Minghua Chen, et al. Device-to-Device Load Balancing for Cellular Networks[C]//2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems, 2015: 19-27.

- [10]. Ravichandran A, Alnoman A, Sharma N, et al. Traffic offloading problem in two-tier HetNets with D2D support for emergency communications[C]// Humanitarian Technology Conference. IEEE, 2017:128-132.
- [11]. Safwan Alwan, Ilhem Fajjari, Nadjib Aitsaadi, et al. Joint multicast routing and OFDM resource allocation in LTE-D2D 5G cellular network[C]// NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium, 2018: 1-9.
- [12]. Guangsheng Feng, Teng Li, Dongdong Su, et al. A joint optimization method for data offloading in D2D-enabled cellular networks[C]// 2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2018: 1-6.
- [13]. S. Madabhushi, G. Gopal, and C. R. Murthy, "Optimal routing and data transmission for Multi-Hop D2D communications under stochastic interference constraints," in National Conference on Communications, Mar. 2017, pp. 394–399.
- [14]. Tomas-Gabarron J B, Ghamri-Doudane Y. Distributed load balancing by two-hop relaying in LTE-Advanced networks[C]// Eleventh IEEE International Conference on Sensing, Communication, and NETWORKING. IEEE, 2016:337-345.
- [15]. Further advancements for E-UTRA physical layer aspects. 3GPP, Tech. Rep. TR 36.814 V9.0.0, 2010.