



Cross-Layer Joint Scheduling for D2D Communication in Cellular Systems

Bi Xixi¹ , Qin Zhiliang² , and Ma Ruofei^{1,2}  

¹ Harbin Institute of Technology, Weihai 264209, Shandong, China
2213715472@qq.com, maruofei@hit.edu.cn
² Beiyang Electric Group Co. Ltd., Weihai, Shandong, China
qinzhiliang@beiyang.com

Abstract. The introduction of device-to-device (D2D) communications into cellular system can improve spectrum efficiency and reduce traffic load of the core network, but it can also increase the scheduling complexing at the base station (BS). Hence designing an efficient scheduling algorithm has become a hot topic for developing D2D communications in cellular systems. But most existing work on the scheduling algorithm commonly assume that all the channel state information (CSI) are known at the BS, i.e., the BS can obtain all the CSI that it needs to perform scheduling, but the scheduling algorithms designed based on the assumptions usually corresponds to a very high overhead on channel measurement when they are implement into actual systems. With this background, a cross-layer joint scheduling scheme that can jointly perform power coordination, channel allocation for D2D capable users by taking into account both the transmission rate criteria of physical layer and the transmission delay criteria of link layer under the condition that the CSI are partly known at the BS are proposed, also the supported D2D mode for D2D-capable users in the cellular system is only the reuse direct D2D mode. Meanwhile, according to the channel probability and statistical characteristics for the interference channels, the end-to-end transmission delay estimation method for the reuse direct D2D mode is presented. Simulation results validate the effectiveness of the joint scheduling scheme and the increase of the system performance brought in the reuse direct D2D mode, in addition, the influence of the major system parameters of the performance on the joint scheduling scheme are also analyzed.

Keywords: Device-to-device (D2D) communication · Channel state information · Queue theory · Transmission delay · Joint scheduling

This work was supported partially by National Natural Science Foundation of China (Grant No. 61801144, 61971156), Shandong Provincial Natural Science Foundation, China (Grant No. ZR2019QF003, ZR2019MF035), and the Fundamental Research Funds for the Central Universities, China (Grant No. HIT.NSRIF.2019081).

1 Introduction

Device-to-Device (D2D) communication is a promising technique for allowing two proximity user equipment to communicate directly [1]. And it can be classified into two categories according to the frequency band which can work, one is licensed-band D2D and the other is unlicensed-band D2D [2]. For the reason that the interference produced by the licensed-band D2D is easily controlled by the BS, hence the most existing works commonly focus on the licensed-band D2D [3]. But meanwhile, there are existing problem on the resource scheduling of the whole system. Hence designing an efficient resource scheduling scheme is very important.

There are three aspects that the scheduling scheme needs to concern, i.e., mode selection, resource allocation and power coordination. Especially, the earlier scheduling schemes mainly focus on the resource allocation or power coordination for the reuse direct D2D mode, and particularly, scheduling schemes on resource allocation are more based on graph theory [4–7]. For the scheduling scheme on mode selection and channel allocation, [8, 9] complete the channel allocation during the mode selection under the fully loaded cellular network and relatively lightly loaded cellular network based on the three D2D mode, i.e., cellular mode, direct D2D mode and relay D2D mode, especially when the cellular network is fully loaded, i.e., the cellular network only considers the reuse D2D mode, the thought that letting the integer programming transform linear programming is proposed under the condition that the relay is considered or not. In addition, there are existing mode selection and resource allocation which apply the D2D communication to the vehicular communication based on the cellular mode, dedicated mode and reused mode [10], and more special in the paper, the mode selection among the cellular mode and dedicated mode is selected based on the distance between the vehicular to vehicular. And the scheduling schemes based on the reinforcement learning and graph theory on mode selection and resource allocation also exist [11, 12]. For the resource allocation and power coordination, in [13], the paper transforms the constraint into feasible region, and derive the optimal power allocation that the D2D users reuse the cellular users and then using the Hungarian algorithm to solve the final problem. In [14], the paper also first delimits the feasible region, and then deduces the optimal power allocation that single D2D users reuse the cellular user, and then using the KM algorithm to solve the multi-pair D2D users to reuse the spectrum resources of the cellular users. For the jointly considering the mode selection, resource allocation and power coordination, in [15], the optimal power allocation of the reuse direct D2D mode and reuse relay D2D mode is derived first, and then the paper compares the transmission rate of the above two modes under the optimal power allocation, and then based on the results, the optimal working mode is determined and the Hungarian algorithm is finally adopted to derive the final results. Integrating the above papers, the core work is power allocation, and when finish the power allocation, the final results can be derived by adopting some casual algorithm.

But the above algorithms possess two limitations. One is that the scheduling algorithms mainly assume that the BS can acquire the perfect channel state information (CSI) of the channel that the BS needs, but lacking the consideration of the partly-known CSI of the channel and the other is that the most algorithms only consider the criteria on the physical layer and lacking the consideration of the criteria on the link layer. Thus it is important to propose a joint scheduling scheme which jointly considers the criteria on

physical layer and link layer under the condition that the CSI is partly-known. Accordingly, in this paper, scheduling algorithms that can jointly perform power coordination, channel allocation for the D2D users which can only work on the reuse direct D2D mode by taking into account both the transmission rate criteria of the physical layer and the transmission delay criteria of the link layer are proposed, and the end-to-end transmission delay estimation schemes for the reuse direct D2D mode are also presented when only the statistical characteristics for the interference channel are known.

The rest of the paper is organized as followed. In Sect. 2, the considered system model is displayed. Section 3 is the problem formulation, and delay estimation, optimization objective solving are also introduced in the Sect. 3. Section 4 is the performance evaluation, Sect. 5 concludes the paper.

2 System Model

As illustrated in Fig. 1. The considered system model is a single-cell system scenario, and M cellular user equipments (CUEs), K pairs of D2D user equipments (DUEs) coexist in the cell. $\{C_1, C_2, \dots, C_m, \dots, C_M\}$ denote the M CUEs, $\{S_1, S_2, \dots, S_k, \dots, S_K\}$ and $\{D_1, D_2, \dots, D_k, \dots, D_K\}$ denote the source and destination DUE of the K pairs of DUEs. For emphasizing the effect on introducing D2D communications, the system model is assumed as a fully loaded system model where DUEs can only access the system via the channel reuse, moreover, the DUEs can only work on the reuse direct D2D mode and each CUE's channel can only be reused by one D2D pair. DUE must reuse the uplink channel of the selected CUE.

Owing to the CSI of the channel between CUE-DUE which needs much signal overhead to acquire, thus BS can't acquire the perfect CSI of the channel between CUE-DUE and only masters the pass loss based on the distance, thus the channel gain in the system is build by two situations. The channel gain between transmit a and receive b except for the CUE-DUE is expressed as followed:

$$g_{a,b} = k_0 \zeta_{a,b}^f \zeta_{a,b}^s l_{a,b}^{-\alpha} \quad (1)$$

where k_0 and α denote the path loss constant and path loss exponent, $\zeta_{a,b}^f$ and $\zeta_{a,b}^s$ denote the fast and slow fading components. And the channel gain of the CUE-DUE is expressed as:

$$h_{C_m, D_k} = k_0 \beta_{C_m, D_k} l_{C_m, D_k}^{-\alpha} \quad (2)$$

where β_{C_m, D_k} denotes channel fading component and l_{C_m, D_k} denotes the distance between CUE-DUE.

3 Problem Formulation

The purpose of the paper is to maximize the transmission rate of the system while guaranteeing the delay constraint of the DUEs and QoS of the DUEs and CUEs under the condition that the CSI are partly known by the BS. For the reason that the BS can

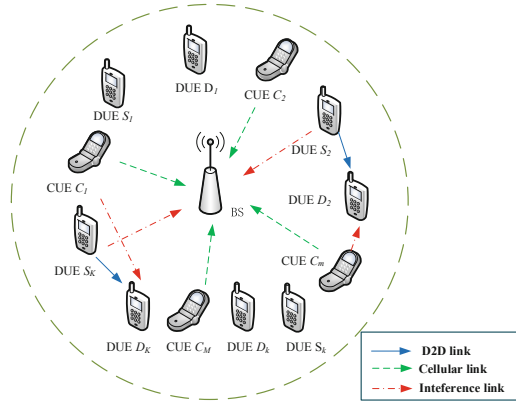


Fig. 1. System model

not master the perfect CSI of the interference channel CUE-DUE, thus the minimum SINR requirement of the DUE can't be guaranteed, hence in this paper, we refer related literature to use the statistical characteristics for the interference channels to guarantee the QoS of the interference channel CUE-DUE, i.e., the probability of the actual SINR of DUE which lower than SINR requirement lower than a given threshold:

$$\Pr \left\{ \frac{P_{k,m}^{(D)} g_{S_k, D_k}}{\sigma^2 + p_{k,m}^{(C1)} h_{C_m, D_k}} < \xi_{\min} \right\} \leq \psi \quad (3)$$

where $p_{k,m}^{(D)}$ and $p_{k,m}^{(C1)}$ denotes the transmit power of DUE and reused CUE, g_{S_k, D_k} denotes the channel gain of DUE, σ^2 and ξ_{\min} denote the noise power and minimum requirement of SINR, and ψ denotes the maximal outage overflow probability of the system.

For the reason that BS can't master the perfect CSI of the interference channel CUE-DUE, hence the transmit outage is inevitable, so we also refer the related literature to use the indicator function Φ to denote the outage event of the DUE. When Φ equals to 1, which denotes the received SINR lower than the given SINR requirement and the receiver of the DUE can't demodulate the signal from the transmit, otherwise, the receiver of the DUE can modulate the signal from the transmit, thus the transmit rate can be denoted as:

$$r_{k,m}^D = \mathbb{E} \left\{ \left[B_0 \log_2 \left(1 + \frac{P_{k,m}^{(D)} g_{S_k, D_k}}{\sigma^2 + p_{k,m}^{(C1)} h_{C_m, D_k}} \right) \right] | \Phi = 0 \right\} \quad (4)$$

where B_0 denotes the bandwidth of the channel.

According to the related literature, above equation can also denoted as followed:

$$r_{k,m}^D = \int_0^l B_0 \log_2 \left(1 + \frac{P_{k,m}^{(D)} g_{S_k, D_k}}{\sigma^2 + \beta_{C_m, D_k} p_{k,m}^{(C1)} H_{C_m, D_k}} \right) \frac{f(\beta_{C_m, D_k})}{F(l)} d\beta_{C_m, D_k} \quad (5)$$

where l denotes the critical value of β_{C_m, D_k} and equals to $\frac{p_{k,m}^{(D)} g_{S_k, D_k} - \xi_{\min} \sigma^2}{p_{k,m}^{(C1)} \xi_{\min} H_{C_m, D_k}}$, H_{C_m, D_k} equals to $k_0 l_{C_m, D_k}^{-\alpha}$, $f(\cdot)$ and $F(\cdot)$ denote the probability density function and cumulative distribution function when the channel obey Rayleigh fading, and the value can be denoted as:

$$f(\beta_{C_m, D_k}) = \begin{cases} \lambda e^{-\lambda \beta_{C_m, D_k}} & \beta_{C_m, D_k} > 0 \\ 0 & \beta_{C_m, D_k} \leq 0 \end{cases} \quad (6)$$

$$F(\beta_{C_m, D_k}) = \begin{cases} 1 - \lambda e^{-\lambda \beta_{C_m, D_k}} & \beta_{C_m, D_k} \geq 0 \\ 0 & \beta_{C_m, D_k} < 0 \end{cases} \quad (7)$$

where λ denotes the mean value of β_{C_m, D_k} , and assumed that it equates to 1.

When CUE is not reused by DUE, the rate of the CUE can be denoted as:

$$r_m^{(C)} = B_0 \log_2 \left(1 + \frac{p_m^{(C)} g_{C_m, B}}{\sigma^2} \right) \quad (8)$$

where $p_m^{(C)}$ and $g_{C_m, B}$ denote the transmit power of CUE when the CUE is not reused by the DUE and the channel gain between CUE and BS.

When the CUE is reused by the DUE, the rate of the CUE can be denoted as:

$$r_{k,m}^{(C1)} = B_0 \log_2 \left(1 + \frac{p_{k,m}^{(C1)} g_{C_m, B}}{p_{k,m}^{(D)} g_{S_k, B} + \sigma^2} \right) \quad (9)$$

where $g_{S_k, B}$ denotes the channel gain between DUE and BS.

Hence the joint scheduling algorithm can be mathematically modelled as:

$$(\mathbf{p}^*, \mathbf{x}^*) = \arg \max_{\mathbf{p}, \mathbf{x}} \left\{ \sum_{k=1}^K \sum_{m=1}^M x_{k,m}^{(1)} (r_{k,m}^{(C1)} + r_{k,m}^{(D)}) + \sum_{m=1}^M \left(1 - \sum_{k=1}^K x_{k,m}^{(1)} \right) r_m^{(C)} \right\} \quad (10)$$

$$\text{subject to: } x_{k,m}^{(1)} \in \{0, 1\}, \forall k, m \quad (11)$$

$$\sum_{k=1}^K x_{k,m}^{(1)} \leq 1, \forall m \quad (12)$$

$$\sum_{m=1}^M x_{k,m}^{(1)} \leq 1, \forall k \quad (13)$$

$$\sum_{m=1}^M x_{k,m}^{(1)} u_{k,m} \leq u_{\max}, \forall k \quad (14)$$

$$\sum_{m=1}^M x_{k,m}^{(1)} p_{k,m}^{(D)} \leq P_{\max}^D, \forall k \quad (15)$$

$$\sum_{k=1}^K x_{k,m}^{(1)} p_{k,m}^{(C1)} + \left(1 - \sum_{k=1}^K x_{k,m}^{(1)}\right) p_m^{(C)} \leq P_{\max}^C, \forall m \quad (16)$$

$$\sum_{m=1}^M x_{k,m}^{(1)} \Pr \left\{ \frac{p_{k,m}^{(D)} g_{S_k, D_k}}{\sigma^2 + p_{k,m}^{(C1)} h_{C_m, D_k}} < \xi_{\min} \right\} \leq \sum_{m=1}^M x_{k,m}^{(1)} \psi, \forall k \quad (17)$$

$$\sum_{k=1}^K x_{k,m}^{(1)} \frac{p_{k,m}^{(C1)} g_{C_m, B}}{p_{k,m}^{(D)} g_{S_k, B} + \sigma^2} + \left(1 - \sum_{k=1}^K x_{k,m}^{(1)}\right) \times \frac{p_m^{(C)} g_{C_m, B}}{\sigma^2} \geq \xi_{\min}, \forall m \quad (18)$$

where $x_{k,m}^{(1)}$ denotes the channel select factor, and when the k -th DUE reuse the channel of the m -th CUE, $x_{k,m}^{(1)} = 1$, otherwise, $x_{k,m}^{(1)} = 0$. $u_{k,m}$ and u_{\max} denotes the delay of the channel of DUE and delay requirement. P_{\max}^D and P_{\max}^C denote the maximal transmit power of the DUE and CUE. (12) and (13) restrict that a DUE can most be reused by one DUE. (14) is the delay requirement constraint. (15) and (16) are the transmit power constraint. (17) and (18) are the QoS constraint of DUE and CUE.

In order to solve the above optimization problem, the first thing is to derive the delay of the reuse direct D2D communications based on the channel probability and statistical characteristics for the interference channels.

3.1 Delay Estimation

D2D link can be treated as a queue server, and it can be denoted by a queuing model, as illustrated in Fig. 2. DUE S_k and D_k denote the source and destination DUE. $A_{k,t}$ and $Q_{k,t}$ denote the number of new arrival packet and queue packets at source DUE at t -th time slot, $U_{k,t}$ denotes the number of the transmitted packets.

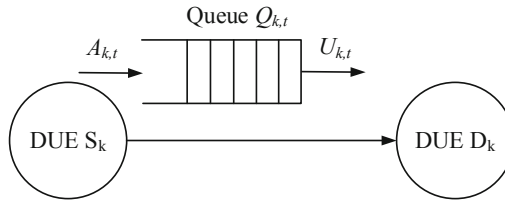


Fig. 2. Queuing model of D2D link

For better depicting the model, the time axis is divided into time slots with length of ΔT . Assuming that packet arrival process is stationary and follows passion distribution, and the arrive rate is λ_k , then the probability of j packets arriving at the source DUE at t -th time slot is depicted as:

$$p(A_{k,t} = j) = \frac{(\lambda_k \Delta T)^j}{j!} e^{-\lambda_k \Delta T}, j \in (0, 1, 2, \dots) \quad (19)$$

Assuming that the transmitted packets always happen before the new arrival packets at every slot, then the relationship between $Q_{k,t}$ and $Q_{k,t-1}$ can be described as followed:

$$Q_{k,t} = \min[B, \max(0, Q_{k,t-1} - U_{k,t}) + A_{k,t}] \quad (20)$$

where B denotes the buffer size, i.e., the number that the source DUE can store at most.

For the reason that the BS can not master the perfect CSI of the interference link CUE-DUE, hence the transmitted rate $U_{k,t}$ is uncertain. Under the condition, the probability of transmitting n packets at t -th time slot can be obtained based on the statistical characteristics of the interference channels, and the process of the solving is presented as followed:

$$\begin{aligned} \Pr\{U_{k,t} = n\} &= \Pr\left\{\left\lceil \frac{\Delta T r_{k,m}^{(D)}}{L} \right\rceil = n\right\} \\ &= \Pr\left\{\frac{nL}{\Delta T} \leq B_0 \log_2\left(1 + \frac{p_{k,m}^{(D)} g_{S_k, D_k}}{\sigma^2 + p_{k,m}^{(C1)} k_0 \beta_{C_m, D_k} l_{C_m, D_k}^{-\alpha}}\right) < \frac{(n+1)L}{\Delta T}\right\} \\ &= \Pr\left\{2^{nL(\Delta TB_0)^{-1}} - 1 \leq \frac{p_{k,m}^{(D)} g_{S_k, D_k}}{\sigma^2 + p_{k,m}^{(C1)} k_0 \beta_{C_m, D_k} l_{C_m, D_k}^{-\alpha}} < 2^{(n+1)L(\Delta TB_0)^{-1}} - 1\right\} \end{aligned} \quad (21)$$

where L denotes the number of data bits in a packet.

Further simplify (21), the above equation can be depicted as:

$$\begin{aligned} \Pr\{U_{k,t} = n\} &= \Pr\left\{\beta_{C_m, D_k} \leq \frac{a-(c-1)\sigma^2}{(c-1)b}\right\} - \Pr\left\{\beta_{C_m, D_k} < \frac{a-(d-1)\sigma^2}{(d-1)b}\right\} \\ &= \Pr\{\beta_{C_m, D_k} \leq e_1\} - \Pr\{\beta_{C_m, D_k} < f_1\} \end{aligned} \quad (22)$$

where $a = p_{k,m}^{(D)} g_{S_k, D_k}$, $b = p_{k,m}^{(C1)} k_0 l_{C_m, D_k}^{-\alpha}$, $c = 2^{nL(\Delta TB_0)^{-1}}$ and $d = 2^{(n+1)L(\Delta TB_0)^{-1}}$, $e_1 = \frac{a-(c-1)\sigma^2}{(c-1)b}$, $f_1 = \frac{a-(d-1)\sigma^2}{(d-1)b}$.

and then (22) can be further depicted as:

$$\Pr\{U_{k,t} = n\} = \begin{cases} 0 & e_1 \leq 0, f_1 < 0 \\ 1 - e^{-e_1} & e_1 > 0, f_1 \leq 0 \\ e^{-f_1} - e^{-e_1} & e_1 > 0, f_1 > 0 \end{cases} \quad (23)$$

According to the expression of $A_{k,t}$ and $U_{k,t}$, the value are not related to the time, hence the following will let the A_k and U_k express the number of new arrival packets and transmitted packets at arbitrarily slot.

Based on the relationship between the number of buffers at two slots, the queue state transition probability matrix can be depicted as:

$$P(Q_{k,t} = \theta_k | Q_{k,t-1} = \varphi_k) = \begin{cases} \Pr\{A_k = \theta_k\} & \varphi_k = 0, 0 \leq \theta_k \leq B - 1 \\ \sum_{j=B}^{\infty} \Pr\{A_k = j\} & \varphi_k = 0, \theta_k = B \\ \sum_{n=\varphi_k}^{\infty} \Pr\{A_k = 0\} \Pr\{U_k = n\} & 1 \leq \varphi_k \leq B, \theta_k = 0 \\ \sum_{n=\varphi_k}^{\infty} \Pr\{A_k = \theta_k\} \Pr\{U_k = n\} \\ + \sum_{n=\varphi_k - \theta_k}^{\varphi_k - 1} \Pr\{A_k = \theta_k + n - \varphi_k\} \Pr\{U_k = n\} & 1 \leq \varphi_k \leq B, 1 \leq \theta_k \leq \varphi_k, \theta_k \neq B \\ \sum_{n=\varphi_k}^{\infty} \Pr\{A_k = \theta_k\} \Pr\{U_k = n\} \\ + \sum_{n=0}^{\varphi_k - 1} \Pr\{A_k = \theta_k + n - \varphi_k\} \Pr\{U_k = n\} & 1 \leq \varphi_k \leq B - 1, \varphi_k < \theta_k \leq B - 1 \\ \sum_{j=B}^{\infty} \sum_{n=\varphi_k}^{\infty} \Pr\{A_k = j\} \Pr\{U_k = n\} \\ + \sum_{n=0}^{\varphi_k - 1} \left(\Pr\{U_k = n\} \sum_{j=B+n-\varphi_k}^{\infty} \Pr\{A_k = j\} \right) & 1 \leq \varphi_k \leq B, \theta_k = B \end{cases} \quad (24)$$

And the packet queuing state transition can be depicted by a Finite state Markov Chain (FSMC), then the stationary probability vector $\Omega_{k,m}$ can be denoted by:

$$\begin{cases} \Omega_{k,m} = \Omega_{k,m} P \\ \sum_{i=0}^B \Omega_{k,m}^i = 1 \end{cases} \quad (25)$$

where $\Omega_{k,m}^i$ denotes the stationary probability of the i packets stored in the source of the DUE. The average packet queue length $\overline{Q_{k,m}}$ can be obtained by:

$$\overline{Q_{k,m}} = \sum_{i=0}^B (i \times \Omega_{k,m}^i) \quad (26)$$

Let $u_{k,m}$ denote the delay of the D2D link, and it can be denoted as followed based on the Littles law:

$$u_{k,m} = \overline{Q_{k,m}} / ((1 - \phi_{k,m}) \times E[A_{k,t}]) \quad (27)$$

where $E[A_{k,t}] = \lambda_k \Delta T$ denotes average number of the packets arriving at the source DUE at a slot and $\phi_{k,m}$ denotes the packet loss rate. And $\phi_{k,m}$ can be obtained by:

$$\phi_{k,m} = D_{k,m} / (E[A_k]) \quad (28)$$

where $D_{k,m}$ denotes the loss number caused by the limited length of the buffer at source DUE, and A_k denotes the stationary distribution of $A_{k,t}$. The calculation of the $D_{k,m}$ can be obtained by:

$$\begin{aligned}
 D_{k,m} &= \sum_{m=1}^B \Omega_{k,m}^m \sum_{n=0}^{m-1} \Pr\{U_k = n\} \sum_{j=B+n-m}^{\infty} (m+j-n-B) \Pr\{A_k = j\} \\
 &+ \sum_{m=1}^B \Omega_{k,m}^m \sum_{n=m}^{\infty} \Pr\{U_k = n\} \sum_{j=B}^{\infty} (j-B) \Pr\{A_k = j\} \\
 &+ \Omega_{k,m}^0 \sum_{j=B}^{\infty} (j-B) \Pr\{A_k = j\}
 \end{aligned} \tag{29}$$

Based on the derivation process, the end-to-end delay can be obtained.

3.2 Optimization Objective Solving

From (10), we can know that the optimization objective is not the linear function and the optimization problem contains the integer variable, hence the optimization problem belongs to the MINLP. For the MINLP, it is difficult to derive the final resolution. Hence we will divide the original problem into two problem to solve, that are power allocation and channel allocation, but the first thing to do first is to control the DUEs to access the network, i.e., access control.

Access Control The accessing control based on the delay requirement, power constraint requirement and QoS requirement will be presented as followed. Only when the DUE satisfy the above three requirements, the DUE can access the network, otherwise, the DUE can't access the network.

When the k -th DUE can reuse the m -th CUE to access the network, the DUE must satisfies the following equations:

$$u_{k,m} \leq u_{\max} \tag{30}$$

$$P_{k,m}^{(D)} \leq P_{\max}^D, P_{k,m}^{(C1)} \leq P_{\max}^C \tag{31}$$

$$\frac{P_{k,m}^{(C1)} g_{C_m,B}}{P_{k,m}^{(D)} g_{S_k,B} + \sigma^2} \geq \xi_{\min} \tag{32}$$

$$\Pr \left\{ \frac{P_{k,m}^{(D)} g_{S_k,D_k}}{\sigma^2 + P_{k,m}^{(C1)} h_{C_m,D_k}} < \xi_{\min} \right\} \leq \psi \tag{33}$$

when the constraints don't consider the delay constraint, i.e., don't consider the (30), whether the k -th DUE can reuse the channel of m -th CUE to work on the reuse direct D2D communications can judge by the distance of CUE-DUE. And only when the real distance of CUE-DUE exceeds the shortest distance, DUE can reuse the channel of the CUE to work on the reuse direct D2D communications. Moreover when the distance

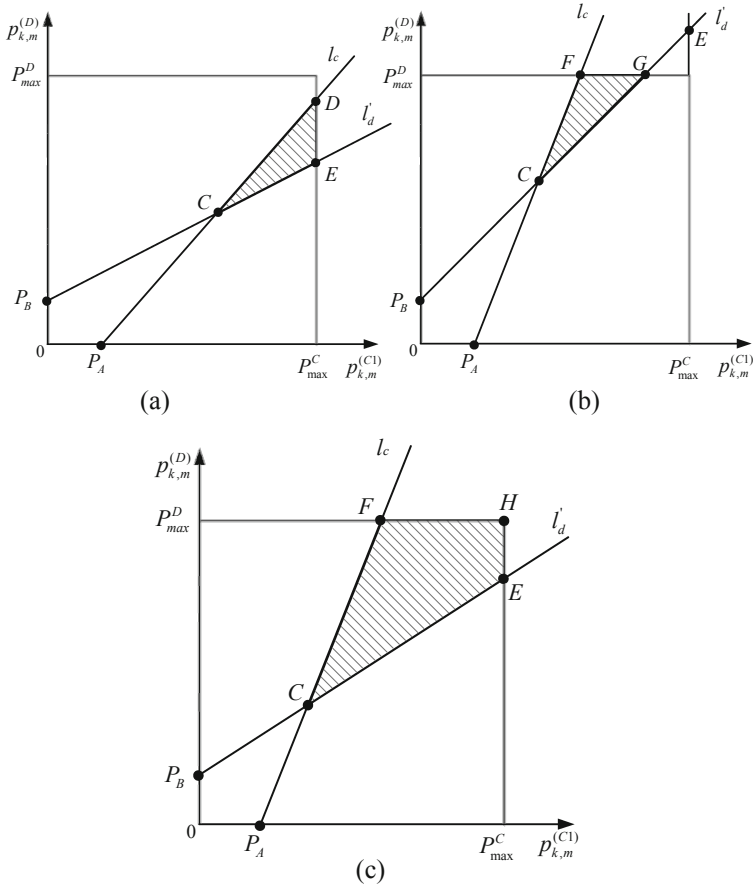


Fig. 3. Three kinds of feasible access regions which don't consider delay constraint

satisfy the distance requirements, there are multiple transmit power pairs $(p_{k,m}^{(D)}, p_{k,m}^{(C1)})$ which can satisfy the (31)-(33) and they can construct three kinds of the feasible access areas, as illustrated in Fig. 3. Three kinds of feasible access regions which don't consider delay constraint where l_c denotes that the (32) takes the equation, and l'_d denotes that the (33) takes the equation.

Assuming that the set of the CUE which satisfies the requirement of distance is CS, if the CS is null, the DUE can't work on the reuse direct D2D communications, otherwise, the delay requirement of the DUE which reuses the CUE in the set need to verify in addition. According to the literature [13, 14], we verified the minimum delay in the feasible access region are D and F. Hence if the D or F satisfy the delay requirement, the DUE can reuse the channel of the CUE to work on the reuse direct D2D communications, otherwise, delete the CUE from the CS, and then justify the CS whether is null, if the CS is null, the DUE can't work on the reuse direct D2D communications, otherwise, D2D can work on the reuse direct D2D communications.

Power Allocation and Channel Allocation The following will allocate the power for the DUEs which can access the network and reused CUEs to maximize the sum rate of them, i.e., $y = r_{k,m}^{(C1)} + r_{k,m}^{(D)}$ and the optimization problem can be described as:

$$\left(p_{k,m}^{(C1)*}, p_{k,m}^{(D)*} \right) = \arg \max_{(p_{k,m}^{(C1)}, p_{k,m}^{(D)})} y \tag{34}$$

subject to: (30)–(33).

When the optimization problem doesn't consider the delay requirement, the heuristic solution of the optimization is:

- (1) In Fig. 3 (a), the optimal allocation point is D and E;
- (2) In Fig. 3 (b), the optimal allocation point is F and G;
- (3) In Fig. 3 (c), the optimal allocation point is F, H and E.

D and F must satisfy the delay requirement, but the E, G and H may not satisfy delay requirement, hence in this paper, we let the D and F to be the power allocation point, but it is also a heuristic solution and may not an optimal solution. The power of D and F are $(P_{max}^C, \frac{P_{max}^C g_{Cm,B} - \xi_{min} \sigma^2}{\xi_{min} g_{Sk,B}})$ and $(\frac{(P_{max}^D g_{Sk,B} + \sigma^2) \xi_{min}}{g_{Cm,B}}, P_{max}^D)$.

When finishing the power allocation, the original optimization problem can be denoted as followed:

$$(\mathbf{x}^*) = \arg \max_x \left\{ \sum_{k=1}^K \sum_{m=1}^M x_{k,m}^{(1)} \left(r_{k,m}^{(C1)} + r_{k,m}^{(D)} - r_m^{(C)} \right) \right\} \tag{35}$$

subject to: (11)–(13).

And the resolution can be solved by related software.

4 Performance Evaluation

The simulation scenario is a single cell with radius 500m. CUE and DUE randomly distribute with the center of BS, and the destination D2D randomly distributes with the center of source DUE. The major parameters are listed in Table 1.

Table 1. Major parameters used in performance evaluation

Parameters	Value
Cell radius	500 m
Path loss exponent (α)	4
Pass loss constant (k_0)	10^{-2}

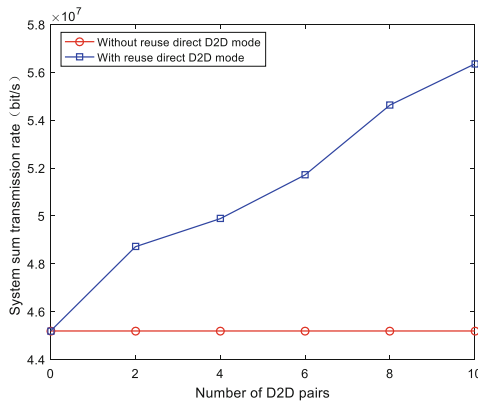
(continued)

Table 1. (continued)

Parameters	Value
Noise power spectral density	-174 dBm/Hz
Channel bandwidth	0.9 MHz
Maximal transmit power of CUE P_{max}^C	24 dBm
Maximal transmit power of DUE P_{max}^D	21 dBm
Minimum SNR/SINR ξ_{min}	10 dB
Number of data bits in a packet L	1024
Duration of a time slot ΔT	1 ms
Fast fading	Exponential distribution with unit mean
Slow fading	Log-normal distribution with standard deviation of 8 dB

The following will validate the effective of the algorithm, the promotion of the system performance and the influence on the performance of methods by some major parameters in the system.

Based on the proposed algorithm, the first step is the evaluation of system transmit performance promotion of introducing the reuse direct D2D mode, as illustrated in Fig. 4. In the simulation, the parameter set as: $M = 10$, $\psi = 0.2$, $\lambda = 7000$, $B = 30$ and $u_{max} = 4$. From Fig. 4, we can see that comparing to the situations without considering the reuse direct D2D mode, considering the reuse direct D2D mode can better promote the system performance, moreover the promotion of the performance increases with the D2D pairs. The results also validate the effective of the algorithm and the importance of considering the reuse direct D2D mode.

**Fig. 4.** System sum transmission rate versus number of D2D pairs

The major characteristic of the paper is considering the transmission delay, thus the influence on the delay by the major parameters is evaluated as followed. Randomly generating a pair of DUE and CUE, observing the influence on the delay by the major parameters. The parameters of the channel generate randomly with fixed distribution and the position of the CUE and DUE generate randomly, hence the change of the delay is wide. In order to observe the influence by the parameters better and provide the reference for the latter, we only retain the relatively small delay results, as illustrated in Fig. 5. The parameters set as: $p_{k,m}^{(D)} = 21$ dBm and $p_{k,m}^{(C1)} = 24$ dBm. From the figure, we can see that when the buffer fixes, the delay will increase with the packets arrival rate, because the increasing of the packets arrival rates will lead to the increasing of average packet queue length and loss rate, hence the delay will increase. In addition, the delay will also increase with the buffer size, that is because increasing of the buffer size will lead to the increasing of the average packet queue length.

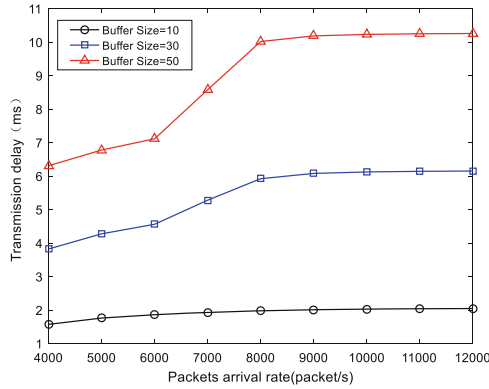


Fig. 5. Transmission delay versus packets arrival rate and buffer size

Above simulation evaluates the delay performance, the following will evaluate the influence by major parameters in the system on the performance of the system, as illustrated in Fig. 6. The parameters in the simulations set as: $M = 10$, $K = 5$, $\psi = 0.2$ and $u_{\max} = 4$. From the results we can see that when the buffer is 30 and 50, the rate of the system is more easily influenced by the packets arrival rate and decreasing with the packets arrival rate. The reason is that the delay will change more obvious with the larger buffer, and the number of the DUE satisfied the delay constraint will decrease, hence the sum rate of the system will decrease.

The following will evaluate the influence by maximal outage overflow probability and delay requirement on the system sum transmission rate. The simulation results illustrate in Fig. 7. And the parameters set as: $M = 15$, $K = 10$, $\lambda = 7000$ and $B = 30$. From the figure, we can see that the system sum transmission rate will increase with the maximal outage overflow probability and delay requirement. The reason is that the number satisfied the access constraint will increase no matter the increasing of the maximal outage overflow probability and delay requirement.

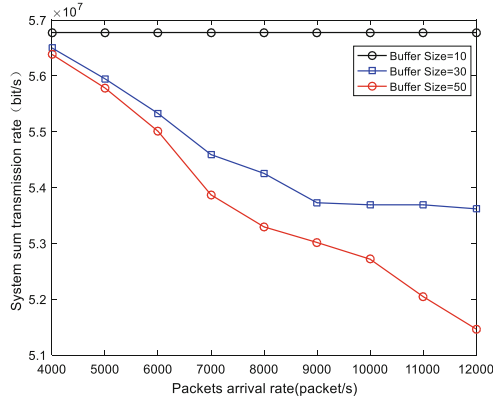


Fig. 6. System sum transmission rate versus packets arrival rate and buffer size

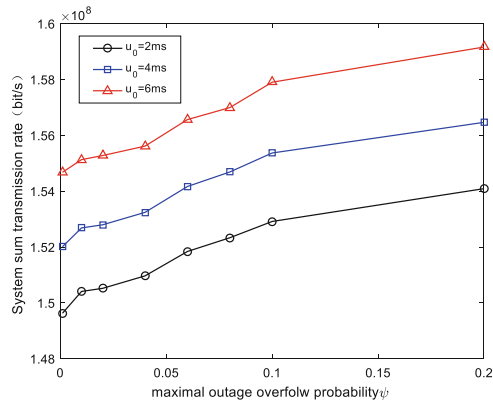


Fig. 7. System sum transmission rate versus maximal outage overflow probability and delay requirement

5 Conclusions

In this paper, we consider the CSI are partly known at the BS and propose a cross-layer joint scheduling scheme that can jointly perform power coordination, channel allocation for D2D capable users by taking into account both the transmission rate criteria of physical layer and the transmission delay criteria of link layer, and meanwhile, the delay estimation method of the reuse direct D2D mode is given. At last, the effectiveness of the joint scheduling scheme and the increase of the system performance brought in the reuse direct D2D mode are validated, in addition, the simulations also analyze the influence of the major system parameters on the performance on the joint scheduling scheme.

References

1. Sakr, A.H., et al.: Cognitive spectrum access in device-to-device-enabled cellular networks. *IEEE Commun. Mag.* **53**(7), 126–133 (2015)
2. Mach, P., Becvar, Z., Vanek T.: In-band device-to-device communication in OFDMA cellular networks: a survey and challenges. *IEEE Communications Surveys & Tutorials* **17**(4), 1885–1922 (2015)
3. Liu, J., Kato, N., Ma, J., Kadowaki, N.: Device-to-device communication in LTE-advanced networks: a survey. *IEEE Communications Surveys & Tutorials* **17**(4), 1923–1940 (2015)
4. Cai, X., Zheng, J., Zhang, Y.: A graph-coloring based resource allocation algorithm for D2D communication in cellular networks. In: 2015 IEEE International Conference on Communications (ICC), pp. 5429–5434 (2015)
5. Mondal, Neogi, A., Chaporkar, P., Karandikar, A.: Bipartite graph based proportional fair resource allocation for D2D communication. In: 2017 IEEE Wireless Communications and Networking Conference (WCNC), pp. 1–6 (2017)
6. Hoang, T.D., Le, L.B., Le-Ngoc, T.: Resource allocation for D2D communication underlaid cellular networks using graph-based approach. *IEEE Transactions on Wireless Communications* **15**(10), 7099–7113 (2016)
7. Zhang, R., et al.: Interference Graph-Based Resource Allocation (InGRA) for D2D communications underlying cellular networks. *IEEE Trans. Veh. Technol.* **64**(8), 3844–3850 (2015)
8. Chen, C., Sung, C., Chen, H.: Capacity maximization based on optimal mode selection in multi-mode and multi-pair D2D communications. *IEEE Trans. Veh. Technol.* **68**(7), 6524–6534 (2019)
9. Chen, C., Sung, C., Chen, H.: Optimal mode selection algorithms in multiple pair device-to-device communications. *IEEE Wirel. Commun.* **25**(4), 82–87 (2018)
10. Jiang, L., Yang, X., Fan, C., Li, B., Zhao, C.: Joint mode selection and resource allocation in D2D-enabled vehicular network. In: 2020 International Conference on Wireless Communications and Signal Processing (WCSP), pp. 755–760 (2020)
11. Zhang, T., Zhu, K., Wang, J.: Energy-efficient mode selection and resource allocation for D2D-enabled heterogeneous networks: a deep reinforcement learning approach. *IEEE Transactions on Wireless Communications* **20**(2), 1175–1187 (2021)
12. Jeon, H.-B., Koo, B.-H., Park, S.-H., Park, J., Chae, C.-B.: Graph-theory-based resource allocation and mode selection in D2D communication systems: the role of full-duplex. *IEEE Wireless Communications Letters* **10**(2), 236–240 (2021)
13. Guo, C., Liang, L., Li, G.Y.: Resource allocation for vehicular communications with low latency and high reliability. *IEEE Transactions on Wireless Communications* **18**(8), 3887–3902 (2019)
14. Feng, et al.: Device-to-device communications underlying cellular networks. *IEEE Transactions on Communications* **61**(8), 3541–3551 (2013)
15. Hoang, T.D., Le, L.B., Le-Ngoc, T.: Joint mode selection and resource allocation for relay-based D2D communications. *IEEE Communications Letters* **21**(2), 398–401 (2017)