Resource sharing and scheduling in device-to-device communication underlying cellular network

Hücresel kullanıcı ile eş zamanlı kanal yaşanılan cihaz-cihaz haberleşmesi için kanal paylaşımı ve zamanlanması

Bilge KARTAL ÇETİN*1, Nuno K. PRATAS*2

1Department of Electrical and Electronics Engineering, Engineering Faculty, Ege University, İzmir, Turkey. 
bilgekortal@ege.edu.tr

2Department of Electronic Systems, The Faculty of Engineering and Science, Aalborg University, Aalborg, Denmark
nuno.pratas@gmail.com

Received/Geliş Tarihi: 28.02.2020
Revision/Düzelme Tarihi: 11.02.2021
Accepted/Kabul Tarihi: 12.02.2021

Abstract

Device-to-Device (D2D) communication is one of the promising technology for the future 5G networks. Utilizing D2D in cellular networks has advantage in terms of capacity and delay. However, in D2D underlay cellular setting, the main concern is quality of service (QoS) for the cellular user due to the mutual interference between D2D user and the cellular user (CU). To utilize the gain brought by D2D setting without violating QoS of the CU, resource sharing is an important design criteria. To this end, we present an optimization model to investigate a resource sharing problem combined with scheduling in a D2D uplink underlay setting. We have used the proposed model to investigate an example resource sharing scenario, in which multiple D2D pairs share the uplink resource of CU, and identified delay and sum throughput for different parameter settings. We observed that there is a significant gain in terms of sum-throughput in allowing a small number of D2D pairs to re-use the cellular resources.

Keywords: D2D communication, Resource sharing, Cellular network, Sum throughput, Delay constraints.

1 Introduction

In conventional cellular networks, a base station (BS) plays a relay role for all communications between the user devices, instead in Device-to-Device (D2D) communication, there is a direct communication link between two users and communication between them takes place without passing through the BS.

There are two types of implementation model for D2D communication to the cellular networks: the underlay and overlay models [1]. In the underlay model, D2D users and cellular users (CUs) can simultaneously share the uplink or downlink channels while in the overlay model there exist a dedicated resource for D2D either in uplink or downlink channel. In the literature, most of the studies considered underlay D2D communication due to its higher spectral gain over to overlay model [2]-[5]. In D2D communications that underlay the communication links of a cellular network, D2D users communicate directly with each other by sharing the radio resources with cellular users, either in a non-orthogonal or orthogonal manner [1]. In this study, we consider the uplink channel and non-orthogonal resource sharing. In such a scenario, there is a mutual interference between CU and D2D pairs, therefore transmission power control is very important to maximize overall network performance. In addition we also consider the queuing dynamics at each device for satisfying required delay constraints.

In the literature, combining the information theoretical bounds and queuing dynamics has been done either by converting a latency constraint into an equivalent rate constraint [6] or relying on elimination [7] or minimization [8] of the interference. In [6], a delay aware resource allocation problem was defined as a sum throughput maximization problem and two statistical approaches were used for expressing a delay constraint in terms of physical layer parameters. In the first approach, a large deviation theory was used to convert the delay constraint into the equivalent rate constraint and in the second approach the characteristics of the power control policies was analyzed by using Lapunov drift stochastic stability analysis. In both approaches, the sum throughput utility optimization problem is converted to a weighted sum rate maximization. In [7], authors formulated a problem of joint

*Corresponding author /Yazıılan Yazar
mode selection, spectrum and power allocation problem with an objective to minimize the maximal buffer size of user devices and evaluated the resulting delay of different resource allocation schemes. In [8], stochastic optimization problem of delay aware dynamic power control was formulated as an infinite horizon average cost Markov Decision Process. To overcome the huge complexity of associated Bellman equations, considering an interference filtering property of the MAC protocol, authors obtained a closed form approximate priority function and introduced a delay-aware power control algorithm. To summarize, [7] and [8] investigated the delay with respect to the average data arrival rates with an assumption of specific stochastic arrival distributions. The authors in [8] also investigated the delay effect of increased number of D2D users without considering the spectral gain trade-off. In [9], a stochastic multi-objective optimization problem is formulated aiming at maximizing spectral efficiency and energy efficiency simultaneously subject to the queue stability constraint. Authors consider a scenario where few D2D devices exist to share many CU’s channel therefore they don’t consider the cellular user’s delay. In [10], authors aims to minimize the average delay and average drop rate, but does not consider the spectral efficiency. In [11], authors focus on distributed flow control and power allocation strategies for delay-aware D2D communication and instead of maximizing overall sum throughput they aim to maximize individual D2D throughput. In [12], dynamic power control is proposed to minimize the average transmit power and average delay but only for overlay D2D communication model.

Resource sharing problem in D2D underlay setting is studied extensively in the literature [2],[3],[14]-[17]. However all these studies focused on increasing sum rate of the network and all of them investigate the case where only one D2D pair is allowed to share the channel with CU. In addition, only few of the existing studies considered QoS limitations of CU in terms of both delay and spectral efficiency [11],[17]. Our contribution is twofold; first, we apply Little’s law for delay calculation which allow us to do more general analysis without relying on any specific data arrival regime. Second, we investigate a scenario where more than one D2D users are allowed to be scheduled for resource sharing. When the channel access is more essential than the data rate this scenario would be an option for the delay tolerance applications.

In this study, we proposed an optimization model that allows to determine the average delay in D2D communication underlaying cellular uplink by joint resource sharing and scheduling. In order to consider the queuing dynamics, we use the Little’s Law which defines a formula to relate packet arrival and departure rates with the average delay for a packet waiting in the queue. Calculated average delay does not depend on any stochastic data arrival regime. Our goal is to maximize the sum data rate of cellular networks, in which one CU shares the uplink resources with the scheduled D2D pairs according to a trade-off between throughput and delay constraints. We formulate the resource sharing problem as a mixed integer nonlinear programming (MINLP) where the transmission powers of the devices and the selection of D2D pairs are determined as decision variables and an objective function is set as maximizing the sum throughput subject to interference and delay constraints. Finally, we present an example case study that considers a D2D communication scenario where multiple D2D pairs can share the sub-channels with cellular users and investigate the sum throughput and delay as a function of increasing number of D2D pairs on the sub-channel.

2 System model

We focus our analysis on a single cell composed by a base station (BS), one CU and 2N devices, as depicted in Figure 1a. The D2D pairs are denoted as $D = \{1, 2, \ldots, N\}$. We assume that the cellular device has an ongoing uplink communication while the 2N devices establish direct D2D links between themselves in pairs, forming N D2D pairs. The set of devices are denoted as $D = \{1, 2, \ldots, (2N + 1)\}$ (i.e. 2N user devices and one cellular user.)

We also assume that communications take place over a frame composed by $M$ time-slots of duration $t_s$ (i.e. the duration of frame is $\{t_f = Mt_s\}$). Furthermore, the distribution of $n$ D2D communication links across the multiple time-slots within a frame is scheduled at the beginning of each frame. As depicted in Figure 1b, in each time-slot $n$ D2D pairs are allowed to share the uplink channel with CU. Note that Figure 1b illustrates an example scenario for a case where only one D2D pair ($n = 1$) shares the resource with CU. The respective mutual interference from the active links occurs as illustrated in Figure 1a.

![Figure 1a](image1a.png) Figure 1a: Uplink resource sharing (desired signals are full lines while interfering signals are traced lines. (b): Scheduling example of two D2D pairs in a frame.

In each slot, $n$ number of D2D pairs are scheduled to share the uplink to maximize the sum rate. SINR levels at each receiver and delay tolerance of user devices are the main criteria for the decision of the scheduling.

2.1 Queue dynamics

We assume that all transmitting devices have to send a volume of data from which they receive from the application layer with a mean rate of $\lambda$ bps. These data are saved in the buffer. The queue length of the buffer for the $i^{th}$ device is $Q_i(m)$ bits at the $m^{th}$ slot. The data in the queue are transmitted in a FIFO fashion. The queue length for the device $i$ at the slot $m$ is calculated as

$$Q_i(m) = Q_i(m - 1) + \lambda t_s - \beta_i(m) R_{i,nom}(m) t_s \tag{1}$$

Where $Q_i(m - 1)$ is the queue length in the former slot and $\lambda t_s$ is total bits arrived in the current slot $m$. The third expression in the right side of the Equation (1) shows the number of bits that are transmitted in one slot from transmitter $i$. Scheduling decision is symbolized with a Boolean variable $\beta_i(m)$. If the $m^{th}$ time slot is assigned to $i^{th}$ transmitter $\beta_i(m) = 1$, otherwise $\beta_i(m) = 0$. The packets in the queue of $i^{th}$ transmitter, $Q_i(m)$ are transmitted at a rate of $R_i(m)$ bits per second. However, in order to calculate the residual queue length we need to
consider a rate without overhead data. We call this rate as a nominal data rate, \( R_{\text{nom}}(m) \) and calculate with
\[
R_{\text{nom}}(m) = R_l(m) - R_{\text{over}}. \tag{2}
\]
The transmission rate, \( R_l(m) \) is calculated with the Shannon formula as follows:
\[
R_l(m) = B \log_2 (1 + \varphi_l) - R_{\text{over}}. \tag{3}
\]
Where \( B \) is the uplink bandwidth and \( \varphi_i \) is the SINR at the receiver of \( i^{th} \) D2D pair. Note that for Equation (3) to be valid, we assume that no bits are dropped and there is always enough buffer to store generated data at the application layer. The SINR value for \( i^{th} \) transmitter is obtained as
\[
\varphi_i(m) = \frac{g_i(m) p_i(m)}{\sigma^2 + \sum_{j \neq i}^{M} p_j(m) h_{ij}(m)} \tag{4}
\]
Where \( g_i(m) \) and \( p_i(m) \) are the channel gain and transmission power at the slot \( m \), respectively, and the white Gaussian noise power is shown as \( \sigma^2 \). At the denominator of Equation (4), channel gain and transmission powers are denoted as \( p_j(m) \) and \( h_{ij}(m) \) for the interfering D2D pairs. There is a minimum SINR threshold \( \varphi_{\text{thr}} \) for \( i^{th} \) transmitter and receiver pair to be scheduled for resource sharing; \( \varphi_i \geq \varphi_{\text{thr}} \). For scheduled devices this property is formulated with the help of Boolean variable, \( \beta_i(m) \) as
\[
\varphi_i(m) \geq \beta_i(m) \varphi_{\text{thr}}, \quad \beta_i(m) \in \{0,1\}. \tag{5}
\]
Minimum required SINR level is determined based on required minimum data rate values.

### 2.2 Scheduling model and delay

We consider estimated delay for new generated packets at the application layer. To do this we utilize Little’s Law. The delay at the slot \( m \), \( D_i(m) \) is the delay experienced by the \( i^{th} \) transmitter at the \( m^{th} \) slot, denoted as \( D_i(m) \), is obtained using Little’s Law as
\[
D_i(m) = \frac{Q_i(m)}{\lambda_{i} \tau_s} \leq D_{\text{max}}. \tag{6}
\]
Where \( D_{\text{max}} \) is the maximum delay tolerance and given in the form of the maximum number of frames \( L \). The impact of the delay in the scheduling decision at each link is modeled by the Boolean variable \( \beta_i(m) \) in the queue length Equation (1).

### 3 Resource sharing problem formulation

In this section, we define the optimal resource sharing for a frame \( \ell \) by taking into account the channel conditions and delay of all communication links and formulate the problem as MINLP. For this to be possible, we assume that all this information is made available to the BS. In a given frame, \( \ell \), the sum rate of CU and the scheduled D2D pairs are given by
\[
R_{\text{sum}}^\ell = \sum_{i=1}^{N} \sum_{m=1}^{M} \beta_i^\ell(m) R_i^\ell(m) + R_l^\ell(m) \quad \forall \ell \in D. \tag{7}
\]
The sum rate \( R_{\text{sum}}^\ell \) is chosen as the objective function of the mathematical programming. In each slot, data rate for the scheduled D2D pairs are determined using the Boolean variable \( \beta_i(m) \), and added to the data rate of the cellular device.

In each slot, \( n \) number of D2D pairs (out of \( N \) D2D pairs who wait for transmission) are chosen to be scheduled to share the channel with CU. The chosen set of D2D pairs are determined to achieve maximum sum throughput. In the optimization model, we maximize \( R_{\text{sum}}^\ell \) by choosing the appropriate transmit power values and a scheduling matrix. Optimum transmitter power levels may be represented by two vectors as follow
\[
P_\ell^\ell = (p_i^\ell)_{m=1}^{M}, \tag{8}
\]
\[
P_\ell^\ell = (p_i^\ell)_{m=1}^{M} \quad \forall \ell \in D. \tag{9}
\]
The scheduling matrix shows that which D2D transmitter scheduled in a given slot for the frame \( \ell \) and shown as follow
\[
B^\ell = (\beta_{i}^\ell(m), ..., \beta_{i}^\ell(m))_{m=1}^{M}. \tag{10}
\]
We formulate the resource sharing and scheduling problem for frame \( \ell \) in the following:

\[
\begin{align*}
\text{max} & \quad 0 \leq P_i \leq P_{\text{max}} \quad R_{\text{sum}}^\ell \\
\text{subject to} & \quad R_{\text{min}} \leq R_i^\ell(m) \leq R_{\text{max}} \tag{11a} \quad \beta \in \{0,1\} \\
& \quad \sum_{i=1}^{N} x_i^\ell(m) = n, \forall m \in M, \forall i \in D^P, \tag{11b} \\
& \quad R_i^\ell(m) = B \log_2 (1 + \varphi_i) - R_{\text{over}} \tag{11c} \\
& \quad \varphi_i(m) = \frac{g_i(m) p_i(m)}{\sigma^2 + \sum_{j \neq i}^{M} p_j(m) h_{ij}(m)} \tag{11d} \\
& \quad \varphi_i(m) \geq \beta_i(m) \varphi_{\text{thr}}, \quad \beta_i(m) \in \{0,1\} \tag{11e} \\
& \quad \frac{Q_i(m)}{\lambda_{i} \tau_s} \leq D_{\text{max}} \tag{11f}
\end{align*}
\]

Where the constraint (11a) denotes the upper and lower bound of data rate while constraint (11b) ensures that maximum numbers of D2D pair can be scheduled simultaneously with CU. (Refer Figure 1b which illustrates the case for \( n = 1 \)). The constraint (11c) denotes the Shannon limit for the data rate and constraint (11d) determines the SINR value for of \( i^{th} \) transmitter. Minimum SINR threshold for D2D pairs to be scheduled is determined by the constraint (11e). Finally the constraint (11f) provides that all users are served within a maximum number of frames, \( D_{\text{max}} \). However, the constraint (11b) does not guarantee that a particular D2D pair should be served in the current frame. To solve the formulated MINLP problem, we use extensive Monte Carlo simulations. Although the solution provides an upper bound on the achievable sum throughput with D2D underlay setting, this bound is not tight due to two reasons. The first, as the delay constraint we utilize the Little’s law which provides the average delay, hence there might be higher delay bounds which may result higher sum throughput values, i.e. even though there exist a set of D2D pairs with better channel gain, another set of D2D pairs with worse channel gain might be scheduled due to their delay constraint. The second, we solve the problem by using extensive Monte Carlo simulations. Hence, the solution might not be global due to non-convex nature of the problem.
4 Delay aware scheduling

In this section, we introduce an algorithm to obtain delay aware scheduling of \( n \) number of D2D pairs in cellular underlay setting, which uses the proposed optimization model. To investigate the sum throughput and delay trade-off, we focus on the following three metrics:

I) The maximum delay experienced by the D2D pairs in cellular network \( D_{\text{max}} \) is calculated as

\[
D_{\text{max}}(n) = \max(D^1_{\text{max}}, ..., D^L_{\text{max}}), \quad \forall i \in D,
\]

where \( D^i_{\text{max}}(n) = \max(D^i(1), ..., D^i(M)) \).

II) The maximum delay experienced by the cellular user defined as

\[
D_{\text{max}}^c(n) = \max(D^i_{\text{max}}(1), ..., D^i_{\text{max}}(M)).
\]

III) The mean sum throughput for \( L \) frame duration when \( n \) D2D pairs share the CU’s resource on the sub-channel.

\[
R_{\text{mean}}(n) = \text{mean}(R^1_{\text{sum}}, ..., R^L_{\text{sum}}).
\]

The solution of the problem (11) provides the scheduling of the \( n \) number of D2D pairs across the multiple time slots within a single frame, and the resulting transmission powers are optimum for the maximum sum throughput.

In order to determine the problem (11) is solved for \( L \) successive frames using a delay calculation algorithm (Table 1). We then obtained and evaluated the delay and sum throughput as a function of increasing number of D2D pairs who share the CU’s uplink.

At the frame number one, queue is empty, but for any successive frames the problem is solved by considering residual queue sizes that are calculated with Equation (1). Outline of the algorithm is given in Table 1.

| Steps | Operation |
|-------|-----------|
| Step 1 | Initialize \( n = 1, \ell = 1, D^L_{\text{max}} = L, \lambda_i = \lambda \)
|       | Set \( n \) in Equation (10) \( \forall i \in D \) |
| Step 2 | Obtain the optimal transmission powers and scheduling matrix \((P^i_c, P^i_k, B^\ell)\) |
| Step 3 | Return \( R^\ell_{\text{sum}}(n), D^i_{\text{max}}(n) \) |
| Step 4 | Update the delay constraints with residual queue size |
| Step 5 | Update \( \ell = \ell + 1 \) until \( \ell \leq L \) go to Step 2 |
| Step 6 | Calculate \( R_{\text{mean}}(n) \) and \( D_{\text{max}}(n) \) |

The proposed resource sharing formulation can be used with any resource sharing scenarios. As an example study, we investigate fixed maximum power scheme, where the CU and D2D pairs communicate with fixed maximum power.

5 Numerical results and discussions

In this section, we present a case study for the application of proposed optimization model in a D2D underlying cellular network. We evaluated the performance of the network in terms of sum throughput and delay experienced by CU and D2D pairs. Formulated MINLP problem is solved for \( L \) number of frames using the parameters in Table 2.

| Parameters | Values |
|------------|--------|
| \( P_{\text{max}}, P_{\text{min}} \) | 23 dBm, -40 dBm |
| \( R_{\text{max}}, R_{\text{min}} \) | -114 dBm |
| Number of slots per frame, \( M \) | 2 |
| Slots length, \( t_s \) | 0.5 ms |
| Number of frames, \( L \) | 50 |
| Max delay tolerance | 100 slots |
| SINR threshold, \( \varphi_{\text{thr}} \) | -80 dB |
| Bandwidth | 5 MHz |
| Maximum Rate, \( R_{\text{max}} \) | 10 Mbps |

We assume that there are 10 D2D pairs and one CU in a cell and the CU shares its uplink resources with \( n \) number of D2D pairs out of 10 D2D pairs. It is assumed that D2D devices and CU are randomly located on a circle with a diameter of \( R = 500 \) m. Note that scheduling of \( n \) D2D pairs corresponds to setting the right hand side of (11b). The numerical results correspond to an average of 1000 fading channel realization in which channel coefficients follow exponential distribution with unit mean. The path loss exponent is chosen as 3.

Figure 2 shows the maximum D2D delay experienced in the network for increased number of scheduled D2D pairs on the sub-channel. We investigated two different arrival rates and two different D2D distances. When no D2D pairs are allowed to be scheduled, the maximum experienced delay is equal to the evaluation period \((D_{\text{max}})\). But, when we let more D2D pairs be scheduled and share the channel with CU then the maximum delay experienced by the D2D pairs is decreasing. As shown in Figure 2, increased data arrival rates \((\lambda = 1 \text{ to } 5 \text{ Mbps})\) cause higher delay, and the distance between D2D pairs has a limited effect on the delay.

Figure 3 shows the maximum delay experienced by CU with different number of D2D pairs on the sub-channel. As shown in Figure 3, as the number of scheduled D2D pairs increases, the maximum delay experienced by the CU increases. It can be also seen from Figure 3 that distance between D2D pairs and CU does not have a significant effect on the CU’s delay.

The sum throughput as a function of the number of scheduled D2D pairs on the sub-channel is illustrated in Figure 4, where two different distance between user equipments are considered to see the effect of interference. We also obtain the sum throughput for different data arrival rates, but they are not illustrated because of negligible difference on sum throughput for the same distance. It can be seen from Figure 4 that, higher distance between D2D pairs result higher sum throughput, but in case a sub-channel is shared by 8 or more D2D pairs, due to the higher mutual interference, the effect of distance between D2D pairs on the sum throughput becomes negligible. As shown in the Figure 4, when D2D pairs are 50m apart, sharing the uplink channel with up to 2 D2D pairs provides the throughput gain compared to the case of no resource sharing (i.e. zero number of D2D pairs on the sub-channel). However, for the distance of 200m the throughput gain exists for the case of up to 3 D2D pairs on the sub-channel. In both distance, for the higher number of D2D pairs on the sub-channel, D2D communication does not enhance the performance in terms of sum throughput. Because cross interference is too strong, the effect of distance on the throughput becomes invisible after
having 8 nodes on the subchannel. However, we can see that significant delay performance (delay reduction) in regards to D2D can only be obtained at the cost of decreased sum throughput and higher cellular delay.

Figure 2. Maximum delay experienced by the D2D pairs.

Figure 3. Maximum delay experienced by cellular user.

Figure 4. Mean sum throughput for increasing number of D2D pairs on the subchannel.

6 Conclusion

In this paper, we have studied D2D communication underlaying a cellular network where we focus the delay aware resource allocation problem for maximum sum rate. The transmission powers and scheduling order of D2D pairs are the main decision variables for maximizing the sum rate subject to delay constraints. By solving the formulated problem, throughput and delay trade-off for a scenario in which multiple D2D pairs are allowed for concurrent communication with a CU is investigated. The results showed that the maximum experienced delay by D2D devices decrease with increasing number of allowed D2D pairs on a sub-channel, while it produces higher delay for CU. However, when the number of D2D pairs on a sub-channel increase, it can be obtained up to 10% gain in terms of sum throughput for only a certain number of D2D pairs, for more D2D pairs the gain starts to decrease due to inevitable mutual interference.

7 Author contribution statements

In the scope of this study, Bilge KARTAL ÇETİN, in the formation of the idea, the design of the system, performing analyzes and the writing and checking the article; Nuno K PRATAS, in the formation of the idea and the assessment of obtained results, were contributed.

8 Ethics committee approval and conflict of interest statement

There is no need to obtain permission from the ethics committee for the article prepared.

There is no conflict of interest with any person/institution in the article prepared.

9 References

[1] Asadi A, Wang Q, Mancuso V. "A survey on device-to-device communication in cellular networks". IEEE Communications Surveys & Tutorials, 16(4), 1801-1819, 2014.

[2] Kazemi Rashed S, Shahbazi R, Ghorashi SA. “Learning-based resource allocation in D2D communications with QoS and fairness considerations”. Transactions on Emerging Telecommunications Technologies, 2018. https://doi.org/10.1002/ett.3249.

[3] Feng Zebing, Feng Zhiyong, Gulliver T Aaron. "Discrete location-aware resource allocation for underlay device-to-device communications in cellular networks". Communications IET, 11(16), 2482-2489, 2017.

[4] Pratas Nuno K, Popovski P. "Network-Assisted device-to-device (D2D) direct proximity discovery with underlay communication". IEEE 2015 Global Communications Conference, SanDiego, CA, USA, 6-10 December, 2015.

[5] Safdar GA, Ur-Rehman M, Muhammad M, Imran MA, Tafazolli R. "Interference mitigation in D2D communication underlaying LTE-A network". Access IEEE, 4(1), 7967-7987, 2016.

[6] Li J, Xia W, Shen L. "Delay-Aware resource control for device-to-device underlay communication systems". Transactions on Emerging Telecommunications Technologies 2017. https://doi.org/10.1002/ett.2949 2017.

[7] Asheralieva A, Miyanaga Y. “Dynamic buffer status-based control for LTE-A network with underlay D2D communication”. IEEE Transactions on Communications, 64(3), 1342-1355, 2016.

[8] Wang W, Zhang F, Lau VKN. “Dynamic power control for delay-aware device-to-device communications”. IEEE Journal on Selected Areas in Communications, 33(1), 14-27, 2015.
[9] Hao Y, Ni Q, Li H and Hou S. "A general framework for spectral efficiency, energy efficiency and delay tradeoff in D2D networks". *IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Exeter, 21-23 June 2017.

[10] Mi X, Zhao M, Xiao L, Zhou S and Wang J. "Delay-Aware resource allocation and power control for device-to-device communications". *IEEE 2015 Wireless Communications and Networking Conference Workshops*, New Orleans, LA, 9-12 March 2015.

[11] Huang S, Liang B and Li J. "Distributed interference and delay aware design for D2D communication in large wireless networks with adaptive interference estimation". *IEEE Transactions on Wireless Communications*, 16(6), 3924-3939, 2017.

[12] Sheng M, Li Y, Wang X, Li J and Shi Y. "Energy efficiency and delay tradeoff in device-to-device communications underlaying cellular networks". *IEEE Journal on Selected Areas in Communications*, 34(1), 92-106, 2016.

[13] Wang F, Li Y, Wang Z, Yang Z. "Social-Community-Aware resource allocation for D2D communications underlaying cellular networks". *IEEE Transactions on Vehicular Technology*, 65(5), 3628-3640, 2016.

[14] Huynh T, Onuma T, Kuroda K, Hasegawa M, Hwang WJ. "Joint downlink and uplink interference management for device to device communication underlaying cellular networks". *IEEE Access*, 4(1), 4420-4430, 2016.

[15] Yang Z, Huang N, Xu H, Pan Y, Li Y, Chen M. "Downlink resource allocation and power control for device-to-device communication underlaying cellular networks". *IEEE Communication Letters*, 20(7), 1449-1452, 2016.

[16] Zhang X and Zhu Q. "D2D offloading for statistical QoS provisionings Over 5G multimedia mobile wireless networks". *IEEE 2019 Conference on Computer Communications*, Paris, France, 29 April-2 May 2019.

[17] Xu Y. "On the performance of device-to-device communications with delay constraint". *IEEE Transactions on Vehicular Technology*, 65(11), 9330-9344, 2016.