FRESH: Push the Limit of D2D Communication Underlaying Cellular Networks

Yang Yang, *Student Member, IEEE*, Tingwei Liu, Xiaoqiang Ma, *Member, IEEE*, Hongbo Jiang, *Senior Member, IEEE*, and Jiangchuan Liu, *Senior Member, IEEE*

Abstract—Device-to-device (D2D) communication has been recently proposed to mitigate the burden of base stations by leveraging the underutilized cellular spectrum resources, where high overall network throughput and D2D access rate are critical for its service performance and availability. In this paper, we study the resource allocation problem to push the limit of D2D communication underlaying cellular networks by allowing multiple D2D links to share resource with multiple cellular links. We propose FRESH, a full resource sharing scheme where each subchannel can be shared by a cellular link and an arbitrary number of D2D links. In particular, FRESH first divides the communication links into so-called full resource sharing sets such that, within each set, all D2D link members are able to reuse the whole allocated resources. Thereafter, it allocates a sum of spectrum resources to each obtained full resource sharing set. As compared with state-of-the-art schemes, FRESH provides fine-grained resource allocation, resulting in throughput improvements of up to one order of magnitude, and D2D access rate improvements of up to 5 times with a moderate node density (e.g., on the order of 1 user per 400 square meters).

Index Terms—Cellular networks, D2D communication, resource sharing

1 Introduction

With the rapid development of smart mobile devices and the proliferation of data-intensive mobile applications, mobile data traffic has increased explosively in recent years, and this trend will undoubtedly continue in the next decade. It is predicted that the mobile data traffic will grow at a compound annual growth rate (CAGR) of 57 percent from year 2014 to 2019, reaching 24.3 exabytes per month by year 2019 [1], nearly a tenfold increase over 2014. Such unprecedented growth of mobile data traffic has put significant pressure on the cellular network infrastructures. Yet cellular networks operate in a centralized fashion, so cellular base stations (BSes) would remain to be the bottlenecks under heavy data traffic.

To address this problem, device-to-device (D2D) communication has been recently proposed to mitigate the burden of base stations by smartly leveraging the *underutilized* cellular spectrum resources [2]. Also, D2D communication is one of the most promising techniques to improve energy and spectrum efficiency, especially for Proximity-based Services (ProSe) [3], [4] and Group Communication System Enablers for LTE (GCSE_LTE) [5]. With this new technique, direct communications between nearby D2D devices are enabled to operate on the same licensed spectrum with traditional cellular communications. It can largely mitigate the traffic demand on base stations and backhaul networks.

- *Y. Yang, T. Liu, X. Ma, and H. Jiang are with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Hubei 430074, China. E-mail: {yangyang8795, tingweiliu2013, mxqhust, hongbojiang2004}@gmail.com.*
- *J. Liu is with the School of Computing Science, Simon Fraser University, BC V5A 1S6, Canada. E-mail: jcliu@cs.sfu.ca.*

Manuscript received 26 Jan. 2016; revised 7 June 2016; accepted 13 Aug. 2016. Date of publication 24 Aug. 2016; date of current version 3 May 2017.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2016.2601917

One technique nodus of exploiting D2D communication stems from its introduced interference to cellular networks due to spectrum sharing. To this end, a proper spectrum resource allocation scheme is essential, and has attracted a lot of studies [6], [7], [8], [9], [10], [11], [12], [13], [14]. Yet most of the works adopt a relatively straightforward strategy that one D2D pair reuses the spectrum resources with only one cellular user (for convenience, we denote this line of works as *user-oriented* resource allocation schemes), as shown in Fig. 1a. In this case, a D2D pair prefers to select a relatively remote cellular user to share the spectrum resources such that the interference of D2D communication is minimized. Another line of works, e.g., [15], [16], [17], [18], take a *resource-oriented* approach, where the resource allocation is based on finer spectrum granularity (see Fig. 1b), say subchannels or resource blocks, rather than the whole link resources. Thus, a cellular user can share its spectrum with multiple D2D pairs, with each operating on different subsets of subchannels; one D2D pair can also share the spectrum resources with multiple cellular users.

In the resource-oriented schemes, the spectrum utilization is still limited by the common assumption that one subchannel can only be either exclusively occupied by one cellular/D2D link or shared by a pair of cellular link and D2D link, as shown in Fig. 1b. Although this assumption simplifies the theoretical analysis, it poses the following limitations. First, the admissible number of D2D links is restricted. In resource-oriented schemes, the admissible number of D2D links cannot exceed the number of subchannels, which significantly limits the access rate of D2D links and the overall throughput when the density of D2D users becomes large. Second, this assumption prevents cellular users from sharing the same subchannels with multiple remote D2D pairs.

We are aware that the spectrum resource allocation is a well-developed problem such as in cognitive radio networks (CRN). In CRN, plenty of works [19], [20], [21] have been

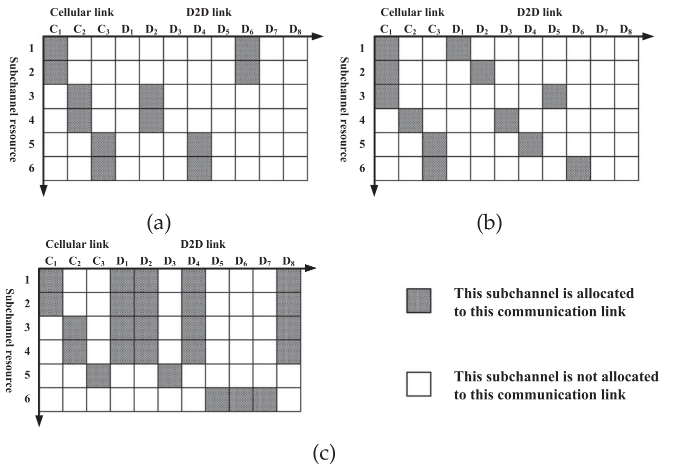


Fig. 1. A simple scenario with three cellular links, eight D2D links, and six subchannels, shows the subchannel assignment by (a) a typical user-oriented scheme where only three D2D links are allocated resources, (b) a typical resource-oriented scheme where six D2D links are allocated resources, resulting in a D2D access rate improvement compared with (a), and (c) our full resource sharing scheme — FRESH, which achieves highest D2D access rate.

proposed for the resource sharing between primary and secondary users based on an observation that up to 85 percent of the spectrum remains unused at a given time and location [20]. However, those works focus on allowing secondary users to access spectrum holes that are unoccupied by the primary users to better utilize the scare resource, which is different from the case in D2D underlaying cellular networks that cellular and D2D users share the spectrum to communicate simultaneously. Also, it is necessary to monitor the receiver of primary server in CRN for sensing the idle resources, which is difficult to be achieved in broadcast systems. Besides, when the cellular cell is overloaded, rare or no idle resource can be perceived. Thus, the schemes in CRN, such as [19], [20], [21], are not suitable for the new scenario.

Intuitively, a natural extension of resource-oriented schemes to address the above limitations is to allow a subchannel to be shared by a cellular user and multiple D2D pairs. The signal strength will decay quickly when the propagation path becomes relatively long; thus a cellular user can operate on the same subchannel together with multiple D2D pairs if they are far away from each other with trivial interference [22]. More specifically, the main idea behind our algorithm is to divide the communication links into different sets and within each set, D2D links can reuse the whole allocated resources. Yet this introduces additional interference in each reuse set which needs careful examination. In other words, the key challenge to implement an efficient full resource sharing scheme is how to identify the cellular users and D2D pairs that can share the same spectrum resources, which trades off between the throughput gain via denser spectrum sharing, and the negative impact caused by accumulative interference. We call our scheme as the *full resource sharing scheme* (FRESH for short) to denote this novel spectrum sharing paradigm, which is shown in Fig. 1c.

1.1 Our Contributions

It is noted that high D2D access rate, referred to as the ratio of the number of admitted D2D pairs to the number of all D2D pairs, is critical for the D2D service availability from the user's perspective. Also, high D2D access rate has the potential to achieve high revenue of network operators by serving more D2D users from the system's perspective [11], [23]. Thus, FRESH offers the potential to simultaneously achieve high network throughput and D2D access rate, and thus pushes the limit of D2D communication underlaying cellular networks. However, it is not a trivial task to find the sweet spot of full resource sharing since bringing new links to the existing spectrum reuse set will possibly introduce additional interference, which could in turn degrade the overall network throughput. The main contributions of our work are as follows:

- We formulate the throughput maximization problem in the scenario that each subchannel is allowed to be shared by a cellular link and an arbitrary number of D2D links, and furthermore, we prove its NP-hardness.
- We develop the theoretical foundations of interference-aware set partitioning on arbitrating the trade-off between the network throughput gain (by allowing further spectrum sharing) and performance degradation (due to accumulative interference). Accordingly, we propose an efficient algorithm for interference-aware set partitioning.
- Based on the obtained partitioning results, we propose two efficient resource allocation algorithms: one devoted to maximizing network throughput (denoted by RA1) and the other devoted to further improving D2D access rate while slightly compromising the network throughput (denoted by RA2).
- Extensive simulations validate that, as compared with state-of-the-art schemes (the user-oriented one [11] and the resource-oriented one [18]), FRESH provides a finer-grained resource allocation, resulting in throughput improvements of up to one order of magnitude and D2D access rate improvements of up to 5 times with a moderate node density (e.g., on the order of one node per 400 square meters).

2 Related Work

In order to ensure efficient usage of resources, two main types of resource allocation schemes have been currently proposed in current version of 3GPP standards: the centralized approaches which allocate unoccupied subchannel resources to D2D users according to the amount of data to be sent and also the radio conditions, when the subchannel resources are sufficient; the distributed approaches in which each D2D transmitter will listen to the resource pool to competitively access to the current unoccupied resource in the resource pool [24]. To further improve the spectrum efficiency and system performance, existing works which focus on resource allocation consider to sharing spectrum resources between cellular users and D2D users.

Most *user-oriented* schemes, e.g., [6], [7], [8], [9], [10], [11], [12], [13], focus on mitigating intra-cell interference by properly pairing cellular and D2D users that share the spectrum resources. Janis et al. [6] proposed a resource allocation scheme to minimize the maximum D2D-to-DL (Download Link) and UL (Upload Link)-to-D2D interference. In [7], the proposed resource allocation scheme is based on uplink resource sharing while avoiding harmful interference by tracking the near-far interference and identifying the interfering cellular users. Yu et al. [10] analyzed the optimal resource allocation and power control between the cellular and D2D links that share the same resources for different resource sharing modes. The authors in [11] formulated the resource allocation problem as a non-linear constraint optimization problem and divided the problem into three subproblems, namely admission control based on SINR requirements, power control, and resource allocation based on maximum weight bipartite matching. Ma et al. [14] jointly consider resource allocation and power control with heterogeneous QoS requirements from the applications, and the proposed user-oriented solution enables better resource utilization for heterogeneous applications with less possibility of underprovisioning or overprovisioning. By introducing an interference price, Zhang et al. [12] model the interference price and power allocation problem between a D2D pair and its allocated cellular uplink as a Stackelberg game. The spectrum resource assignment for different D2D pairs is solved by the Hungarian algorithm. Zhao et al. [13] propose to leverage the social information to assist resource allocation. The resource allocation problem is formulated by forming a model in which the D2D pairs are stimulated to share the resources of the cellular users in the same or outside the community.

Xu et al. [25] proposed a reuse strategy for pairing cellular links and D2D packages (a group of D2D pairs that share the same resources). They introduced an allocation mechanism based on reverse iterative combinatorial auction. This work is similar to ours in the way that multiple D2D pairs can share the same spectrum, yet it does not address the problem of how to divide D2D links into packages and it simply assumes that D2D packages are already known. Thus, the resource allocation problem for the considered scenario is essentially a one-to-one allocation problem, i.e., a D2D package can be regarded as a specific D2D link.

To address the limitation of user-oriented schemes that one D2D user can only share resource with one cellular user, many *resource-oriented* schemes [15], [16], [17], [18] have been proposed to allocate the spectrum resources based on finer granularity. Marco et al. [15] proposed a distributed resource reuse scheme incorporating mode selection and power control. Zhang et al. [16] formulated the interference relationship among different D2D and cellular links as an interference-aware graph, and proposed a resource allocation method based on this graph. Zhou et al. [17] assumed that cellular communication has a higher priority, and introduced an interference coordination mechanism to maximize the number of reliable D2D links by appropriately reusing the cellular resources for D2D links. Su et al. [18] formulated the problem of maximizing the network throughput with minimum data rate requirement. The joint mode selection and resource allocation problem is solved through particle swarm optimization method. However, we reveal that the performance of resource-oriented schemes is limited by the assumption that one subchannel can only be either exclusively occupied by one cellular/D2D link or shared by a pair of cellular link and D2D link.

Malandrino et al. [26] proposed a spectrum reuse scheme wherein the spectrum resources are assigned to multiple pairs of endpoints (I2D and D2D) at the same time, which seems similar with our idea. However, they merely considered the resource allocation for the new coming user which requests service, while the resource allocation works we mentioned above considered how to allocate resource for all the existing users. More specifically, [26] considered the current state (the state of the system which describes how do the existing users occupy the spectrum resources) to be known, and made decision for the new coming user based on current state. The drawback of doing so is that the admission of new coming user will degrade the data rate of existing users (the data rate can even be degraded to be below the lowest threshold). Besides, just considering the optimal choice for the new coming user cannot achieve as good system performance as considering the optimal resource allocation for all the existing users.

Spectrum resource allocation is a well developed problem such as in cognitive radio networks (CRN), while facing the D2D communication underlaying cellular networks scenario, there are still many different challenges. In CRN, plenty of works [19], [20], [21] have been proposed for the resource sharing between primary and secondary users based on an observation that up to 85 percent of the spectrum remains unused at a given time and location [20]. However, those works focus on allowing secondary users to access spectrum holes that are unoccupied by the primary users to better utilize the scare resource. More specifically, the secondary users use the spectrum resources allocated to primary users when there are no primary users' data at a time interval by sensing environment. This is much different from the works for resources sharing in D2D underlaying cellular networks which consider cellular users and D2D users occupying the same spectrum resources can send data at the same time. As the schemes in CRN need to sense the idle spectrum resources, it is necessary to monitor the receiver of primary server, which is difficult to be achieved in broadcast systems. Besides, when the cellular cell is overloaded, rare or no idle resource can be perceived. Thus, when the schemes in CRN, such as [19], [20], [21], are applied to D2D underlaying cellular networks, a very limited number of D2D users can be served, for those schemes do not consider the resource sharing.

3 System Model and Problem Formulation

In this section, we present the system model and formulate the throughput maximization problem of the full resource sharing scheme. We summarize the notations in Table 1.

3.1 System Model

D2D communication can work in either the *dedicated* mode or the *reuse* mode [9]. In the dedicated mode, the cellular network allocates an exclusive portion of spectrum resources for the direct communications between D2D pairs. There is no interference between the cellular and D2D communications; in the reuse mode, D2D communications reuse a portion of or the whole resources allocated to a cellular network. This mode can be further divided into downlink resource sharing (*DLre*) and uplink resource sharing (*ULre*), where the D2D communications reuse the spectrum resources of the downlink/uplink of cellular users. In the reuse mode, D2D communications will cause interference to the cellular communications utilizing the same spectrum. In this paper, we

TABLE 1
Summary of Important Symbol Used

Symbol	Description
N_c	The number of cellular users
N_d	The number of D2D pairs
N_{sub}	The number of subchannels in total
N_{set}	The number of resource reuse sets
C	The set of cellular links, $C = \{C_1, \ldots, C_{N_c}\}$
D	The set of D2D links, $D = \{D_1, \ldots, D_{N_d}\}$
K_i	The ratio of the interference to thermal noise at the receiver of link i
N_0	The power of thermal noise on a single sub-channel
P_i	The transmitting power of link i
$P_B^{\max}, P^{\max}$	The respective maximum transmitting power of BS and UE
$d_{\langle i,j\rangle}$	The distance between the transmitter of link i and the receiver of link j
$g_{\langle i,j\rangle}$	The channel gain between the transmitter of link i and the receiver of link j
ξ_i^c (ξ_j^d)	The SINR of cellular link i (D2D link j)
S_k	The k_{th} resource reuse set in resource sharing set partitioning $\{S_k\}$
S_k^c, S_k^d	The subset of cellular and D2D link members of S_k
$\lvert\cdot\rvert$	The cardinality of a given set
w_k	The number of subchannel resources allocated to S_k
W	The bandwidth of each subchannel
γ_h, γ_l	The upper/lower limit of SINR to satisfy the highest/lowest MCS
n_{cset}	The number of set S_k^c, where $S_k^c = \emptyset$

focus on downlink resource sharing. The analyses can be easily extended to the uplink resource sharing scenario.

We start from a single cell scenario with one base station, N_c cellular users that communicate with the BS, and N_d D2D pairs that communicate within pair using the licensed spectrum resources, as shown in Fig. 2. We use UE to denote a user equipment.

We consider Rayleigh fading channel, which is a widely adopted channel model. The power received at a receiver is expressed by G_0GFP, where P represents the transmitting power; $G > 0$ represents the path gain and is usually assumed to be proportional to $d^{-\alpha}$, where d is the distance between transmitter and receiver, and α is the power decay factor; G_0 stands for the antenna and coding gain; F stands for Rayleigh fading, which is independent and exponentially distributed with unit mean. The distribution of the received power is then exponential with the mean value $E[G_0GFP] = G_0GP$. Thus, the power of useful signal received at receiver can be expressed as $G_0d^{-\alpha}P$. For convenience, we use $g = G_0d^{-\alpha}$ to denote the overall channel gain.

Consider the scenario with only one subchannel, two communication links i and j using the same spectrum as an example, the SINR at the receiver of link j can be expressed as follows:

$$\xi_j = \frac{P_jG_{0\langle j,j\rangle}d_{\langle j,j\rangle}^{-\alpha}}{P_iG_{0\langle i,j\rangle}d_{\langle i,j\rangle}^{-\alpha} + N_0} = \frac{P_jg_{\langle j,j\rangle}}{P_ig_{\langle i,j\rangle} + N_0}, \tag{1}$$

where N_0 is the power of thermal noise on a single subchannel; P_i stands for the transmitting power of link i; the tuple $\langle i,j\rangle$ stands for the transmitter of communication link i and the receiver of communication link j, and $d_{\langle i,j\rangle}$ represents the distance between them. Besides, in most wireless communication systems, there is an upper limit on the spectrum efficiency such that a SINR higher than a maximum value, γ_h, does not further increase the spectrum efficiency that is limited to r_h bps/Hz. A link adaptation technique will select a proper modulation and coding scheme (MCS) from a limited number of options according to the current channel condition [27], and r_h is achieved when the current SINR is high enough to support the highest MCS. On the other hand, the SINR should be no lower than a minimum value, γ_l, to support the lowest MCS that provides the spectrum efficiency of r_l bps/Hz.

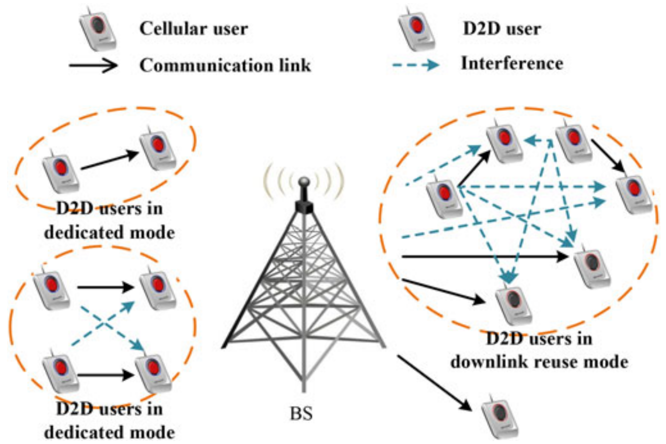


Fig. 2. System model of D2D communication underlaying cellular networks.

3.2 Problem Formulation

In FRESH, a subchannel can be shared by multiple communication links. A *resource reuse set* refers to the set of cellular and D2D links reusing the same spectrum resources, either in dedicated or reuse mode.

Suppose that all the communication links, namely $C \cup D$ (C for cellular links and D for D2D links), are partitioned into a number of discrete resource reuse sets $\{S_k\}$. The kth set contains two subsets S_k^c and S_k^d that stand for cellular links and D2D links of S_k, respectively. In each set S_k, each cellular link occupies an equal share of the subchannel resources allocated to S_k (denoted by w_k), and the D2D links reuse all the available resources.

Fig. 1 illustrates the subchannel assignment of different schemes. For either user- or resource-oriented scheme, each subchannel can be shared by at most two links, namely a cellular link and a D2D link. That is to say, N_{sub} subchannels can only serve up to $2N_{sub}$ links. While for our full resource sharing scheme, three cellular links and eight D2D links are divided into three reuse sets, namely $S_1 = \{C_1, C_2, D_1, D_2, D_4, D_8\}$, $S_2 = \{D_5, D_6, D_7\}$, $S_3 = \{C_3, D_3\}$. Each subchannel can be shared by multiple D2D links and one cellular link, providing the potential to achieve higher spectrum utilization and higher access rate if the interference is well managed.

In the following, we formulate the network throughput maximization problem of the full resource sharing scheme. Let P_i denote the transmit power of link i, and ξ_i^c and ξ_j^d denote the SINR of cellular link i and D2D link j, respectively. In each reuse set S_k, the cellular link members equally split the sub channel resources, and they cause interference to other D2D link members; the D2D link members reuse all the available resources for S_k, causing interference to the other members (as illustrated in Eq. (2)). In this work, we assume that the transmit power of the BS and all the UEs are set to the maximum values P_B^{max} and P^{max}, respectively; we leave the optimal power control in the future work. The overall network throughput maximization problem can be formulated as follows:

$$\max_{\{w_k\}\{S_k\}} \left\{ \sum_k \left(\sum_{i \in S_k^c} \frac{w_k W}{|S_k^c|} \log_2(1+\xi_i^c) + \sum_{j \in S_k^d} w_k W \log_2(1+\xi_j^d) \right) \right\} \tag{2}$$

$$\text{s.t.} \quad \xi_i^c = \frac{P_i g_{\langle i,i \rangle}}{\frac{w_k}{|S_k^c|} N_0 + \sum_{l \in S_k^d} P_l g_{\langle l,i \rangle} / |S_k^c|} \geq \gamma_l, i \in S_k^c, \tag{2a}$$

$$\xi_j^d = \frac{P_j g_{\langle j,j \rangle}}{w_k N_0 + \sum_{l \in S_k^d \cup S_k^d \setminus \{j\}} P_l g_{\langle l,j \rangle}} \geq \gamma_l, j \in S_k^d, \tag{2b}$$

$$\sum_k w_k = N_{sub}, \tag{2c}$$

$$S_k^c = S_k \cap C, \bigcup_k S_k^c = C, \tag{2d}$$

$$S_k^d = S_k \cap D, \bigcup_k S_k^d = D, \tag{2e}$$

$$S_k = S_k^c \cup S_k^d, S_k \cap S_{k'} = \emptyset (k \neq k'), \tag{2f}$$

$$\text{given} \quad P_i = P_B^{\max}, \text{if } i \in C$$

$$P_j = P^{\max}, \text{if } j \in D.$$

Here, constraints (2a) and (2b) ensure that the SNR of both cellular links and D2D links in set S_k can reach the minimum value to enable communication. Constraint (2c) guarantees the amount of resource allocated to different sets should not exceed the total number the cellular cell has. Constraint (2d), (2e) and (2f) guarantee that there is no overlap in different resource sharing sets, and the partitioning result can cover all the links.

3.3 Theoretical Preliminary

To solve problem above, we need to find the suitable reuse set partitioning results $\{S_k\}$, and the resource allocation strategy $\{w_k\}$. This problem can be naturally divided into two subproblems: how to find a *good* set partitioning method, and how to allocate spectrum resources for these reuse sets.

Theorem 1. *The first subproblem, optimal set partitioning problem is NP-complete.*

Proof. See Appendix A for more details. □

Theorem 2. *Given any partitioning results $\{S_k^*\}$, the second subproblem, optimal resource allocation problem is still NP-complete.*

Proof. See Appendix B for more details. □

Theorem 3. *For the first subproblem—optimal set partitioning problem, no polynomial time algorithm can approximate the optimal result within a factor of $O(n^\epsilon)$ ($\epsilon > 0$, n is the number of links), unless $P = NP$.*

Proof. See Appendix C for more details. □

The first subproblem cannot be approximated very well with in polynomial time. An algorithm with exponential running time may approximate the optimal solution well but it cannot be well applied in practice. The channel condition of different links varies due to a variety of reasons, for instance, the users' mobility, and BS need to repeat resource allocation over and over again. Thus, the algorithm with exponential running time will spend too much time and cannot update allocation result right-on-time. Hence, we need to design an efficient algorithm which can obtain good resource reuse set partitioning. Intuitively, bringing more links to share the same spectrum can increase the spectrum utilization; yet the accompanying increased accumulative interference would degrade the data rates of the original links. Thus, a good resource reuse set is supposed to guarantee that the gain through spectrum reusing can surpass the negative performance impact caused by interference, eventually increasing the overall throughput. We next discuss how to obtain such good resource reuse set partitioning in detail.

4 Resource Sharing For Different Modes

In this section we develop the theoretical foundations required to define a good resource reuse set. We then design an interference-aware set partitioning algorithm that divides the communication links into discrete resource reuse sets.

4.1 Resource Sharing Condition for Dedicated Mode

We start from the case of dedicated mode. We use S^d to denote the set of D2D links and n to denote the subchannel resources allocated to these $|S^d|$ D2D links. The bandwidth of each subchannel is W. Here, the D2D links reuse all the available resource allocated to the whole reuse set. And the receiver of each D2D link will receive the interference signals from the transmitters of the other D2D links in the same reuse set. Fig. 2 provides a example for the interference relationship in dedicated mode.

Consider the non-sharing scheme that the D2D links do not share the spectrum resources; instead they equally split the spectrum. The total throughput is

$$R_{non}^{sum} = \max_{\{k_i\} \subset S^d} \left(W \sum_{i=1}^{n} \log_2 \left(1 + \frac{P_{k_i} g_{\langle k_i, k_i \rangle}}{N_0} \right) \right). \tag{3}$$

To achieve the highest spectrum utilization, the link with the maximum SINR can exclusively use the whole spectrum, which gives:

$$R_{non}^{sum} = \max_{i \in S^d} \left(nW \log_2 \left(1 + \frac{P_i g_{\langle i,i \rangle}}{N_0} \right) \right) = nW \log_2 (1 + \xi_h), \tag{4}$$

where

$$\xi_h = \min\left\{\max_{i\in S^d}\left(\frac{P_i g_{\langle i,i\rangle}}{N_0}\right), \gamma_h\right\}. \tag{5}$$

Now consider the full resource sharing scheme. For convenience, we denote the total interference caused by other links to link i as

$$\sum_{j\in S^d\setminus\{i\}} P_j g_{\langle j,i\rangle} = K_i N_0. \tag{6}$$

Thus, the total throughput is

$$\begin{aligned} R_{full}^{sum} &= nW\sum_{i\in S^d}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{nN_0+\sum_{j\in S^d\setminus\{i\}}P_j g_{\langle j,i\rangle}}\right) \\ &= nW\sum_{i\in S^d}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0(K_i+n)}\right). \end{aligned} \tag{7}$$

To ensure that the full resource sharing scheme achieves a higher throughput, we need to have $R_{full}^{sum} \geq R_{non}^{sum}$. By strengthening the inequality, $R_{full}^{sum} \geq R_{non}^{sum}$ holds if

$$\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0(K_i+n)}\right) \geq \frac{1}{|S^d|}\log_2(1+\xi_h), \forall i\in S^d. \tag{8}$$

Proposition 1 (Resource Sharing Condition 1). *Consider that $\left|S^d\right|$ D2D links share n subchannels (here D2D links work in dedicated mode). Denoting the accumulative interference caused by other links to link i by K_iN_0, the full resource sharing scheme will achieve higher throughput than the non-sharing one if*

$$K_i \leq \frac{P_i g_{\langle i,i\rangle}/N_0}{(1+\xi_h)^{\frac{1}{|S^d|}}-1} - n, \tag{9}$$

where

$$\xi_h = \min\left\{\max_{i\in S^d}\left(\frac{P_i g_{\langle i,i\rangle}}{N_0}\right), \gamma_h\right\}.$$

4.2 Resource Sharing Condition for Reuse Mode

We use S^c and S^d to represent the sets of cellular links and of D2D links, respectively, and n to represent the subchannel resources allocated to these links. Here $n > |S^c|$, as we need to allocate at least one subchannel for each cellular link. The cellular link members equally split the sub channel resources, and the D2D links reuse all the available resource allocated to the whole reuse set. There is no interference between cellular links. But the receiver of each cellular link will receive interference signals from the transmitters of all the D2D links in the same set. And the receiver of each D2D links will receive the interference signals from the transmitters of the other D2D links and all the cellular links in the same set. Fig. 2 provides a example for the interference relationship in reuse mode.

Consider the non-sharing scheme that D2D links equally splits the spectrum. The total throughput is:

$$\begin{aligned} R_{no}^{sum} &= W\sum_{i\in S^c}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0}\right) \\ &+ \max_{\{k_i\}\in S^d}\left(W(n-|S^c|)\log_2\left(1+\frac{P_{k_i} g_{\langle k_i,k_i\rangle}}{N_0}\right)\right). \end{aligned} \tag{10}$$

Similarly, if the link with the highest SINR exclusively uses the whole spectrum, the total throughput of the non-resource sharing scheme is

$$R_{no}^{sum} = Wn\log_2(1+\xi_h), \tag{11}$$

where

$$\xi_h = \min\left\{\max_{i\in S^c\cup S^d}\left(\frac{P_i g_{\langle i,i\rangle}}{N_0}\right), \gamma_h\right\}. \tag{12}$$

As for the full resource sharing scheme, for convenience, we denote the accumulative interference caused by other links to link i as

$$\begin{aligned} &\sum_{j\in S^d} P_j g_{\langle j,i\rangle}/n = K_i N_0, \quad i\in S^c, \\ &\sum_{j\in S^c\cup S^d\setminus\{i\}} P_j g_{\langle j,i\rangle} = K_i N_0, \quad i\in S^d. \end{aligned} \tag{13}$$

Thus, the total throughput can be expressed as

$$\begin{aligned} R_{full}^{sum} &= \frac{nW}{|S^c|}\sum_{i\in S^c}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{\frac{n}{|S^c|}N_0+\left(\sum_{j\in S^d}P_j g_{\langle j,i\rangle}\right)/n}\right) \\ &+ nW\sum_{i\in S^d}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{nN_0+\sum_{j\in S^c\cup S^d\setminus\{i\}}P_j g_{\langle j,i\rangle}}\right) \\ &= \frac{nW}{|S^c|}\sum_{i\in S^c}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0\left(K_i+\frac{n}{|S^c|}\right)}\right) \\ &+ nW\sum_{i\in S^d}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0(K_i+n)}\right). \end{aligned} \tag{14}$$

To ensure that the full resource sharing scheme achieves higher throughput than the non-sharing one, we need to guarantee $R_{full}^{sum} \geq R_{non}^{sum}$.

For D2D link $i\in S^d$, we can draw the conclusion for resource sharing condition inspired by Proposition 1. When D2D link $i\in S^d$ satisfies the following condition

$$K_i \leq \frac{P_i g_{\langle i,i\rangle}/N_0}{(1+\xi_h)^{\frac{1}{|S^c|+|S^d|}}-1} - n, i\in S^d,$$

we can obtain

$$(|S^c|+|S^d|)\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0(K_i+n)}\right) \geq \log_2(1+\xi_h), i\in S^d,$$

$$nW\sum_{i\in S^d}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0(K_i+n)}\right) \geq \frac{|S^d|}{|S^c|+|S^d|}nW\log_2(1+\xi_h),$$

thus, $R_{full}^{sum} \geq R_{no}^{sum}$ holds if

$$\frac{nW}{|S^c|}\sum_{i\in S^c}\log_2\left(1+\frac{P_i g_{\langle i,i\rangle}}{N_0\left(K_i+\frac{n}{|S^c|}\right)}\right) \geq \frac{|S^c|}{|S^c|+|S^d|}nW\log_2(1+\xi_h). \tag{15}$$

We strengthen the inequality, and the resource sharing condition for cellular link $i \in S^c$ will be:

$$K_i \leq \frac{P_i g_{\langle i,i\rangle}/N_0}{(1+\xi_h)^{\frac{|S^c|}{|S^c|+|S^d|}}-1} - \frac{n}{|S^c|}, i \in S^c. \tag{16}$$

Proposition 2 (Resource Sharing Condition 2). *Consider that $|S^c|$ cellular links and $\left|S^d\right|$ D2D links share n subchannels (here D2D links work in reuse mode). Denoting the accumulative interference caused by other links to link i by $K_i N_0$, the full resource sharing scheme will achieve higher throughput than the non-sharing one if*

$$K_i \leq \begin{cases} \frac{P_i g_{\langle i,i\rangle}/N_0}{(1+\xi_h)^{\frac{|S^c|}{|S^c|+|S^d|}}-1} - \frac{n}{|S^c|}, i \in S^c, \\ \frac{P_i g_{\langle i,i\rangle}/N_0}{(1+\xi_h)^{\frac{1}{|S^c|+|S^d|}}-1} - n, i \in S^d, \end{cases} \tag{17}$$

where

$$\xi_h = \min\left\{\max_{i\in S^c\cup S^d}\left(\frac{P_i g_{\langle i,i\rangle}}{N_0}\right), \gamma_h\right\}.$$

4.3 Interference-Aware Resource Sharing Set Partitioning

The resource sharing conditions proposed in Propositions 1 and 2 help to define a good resource reuse set on arbitrating the tradeoff between the network throughput gain (by allowing further spectrum sharing) and performance degradation (from accumulative interference). Combining the conditions and the lowest SINR requirements in Eq. (2a) and (2b), we give the definition of *weak interference set* as follows:

Definition 1. *A set of communication links is called a* weak interference set, *if all the links satisfy resource sharing conditions in Propositions 1 and 2 as well as the lowest SINR requirements in Eq. (2a) and Eq. (2b).*

Given n subchannels, the non-sharing scheme provides a *lower bound* of the network throughput

$$\inf R_{full}^{sum} = nW\log_2(1+\xi_h), \tag{18}$$

where the ξ_h stands for the highest SINR of links in the weak interference set without resource sharing, as illustrated in Eq. (12). Thus the lower bound increases while bringing new links into the weak interference set. The concept of weak interference set helps us to design our set partitioning algorithm. The pseudo-code of our interference-aware set partitioning algorithm is shown in Algorithm 1. It iteratively finds the weak interference set with the maximum size from the candidate communication links. We denote the set of candidate links by CL. Initially, CL contains all the communication links in the system. The building blocks of our interference-aware set partitioning algorithm consist of: (1) For the kth resource sharing set S_k, we first set $S_k = CL$ and trim S_k according to the lowest SINR requirements in Eq. (2a) and (2b); (2) Trim S_k according to sharing conditions in Eqs. (9) and (17); (3) The remaining links in S_k constitute a weak interference set, which are then removed from CL. This process will execute iteratively until $CL = \emptyset$.

Algorithm 1. Resource Reuse Set Partitioning

$k = 1$;
while $CL \neq \emptyset$ **do**
  $S_k = CL$;
  //Trim S_k according to the lowest SINR requirements
  for all $i \in S_k$ **do**
    calculate ξ_i;
  end for
  find link $i \in S_k$ with minimum ξ_i;
  while $\xi_i - \gamma_l < 0$ **do**
    remove i from S_k;
    for all $i \in S_k$ **do**
      calculate ξ_i;
    end for
    find link $i \in S_k$ with minimum ξ_i;
  end while
  //Trim S_k according to the resource sharing conditions
  find $S^* \subset S_k$ of links not satisfying Eq. (9) in Proposition 1 and Eq. (17) in Proposition 2;
  while $S^* \neq \emptyset$ **do**
    for all $i \in S^*$ **do**
      calculate ξ_i;
    end for
    find link $i \in S^*$ with minimum ξ_i;
    remove i from S_k;
    find $S^* \subset S_k$ of links not satisfying Eq. (9) and Eq. (17);
  end while
  remove link members of S_k from CL and set $k = k + 1$;
end while

4.4 Complexity Analysis

It is noted that the set partitioning and the resource allocation are both conducted in a centralized fashion. Thus, one important issue of FRESH is its computational complexity. In the worst case, which means there is no communication link can reuse resource with other links perhaps for its poor channel condition or the intensive interference resulted by reusing, our set partitioning algorithm will run $|C| + |D|$ rounds of trimming. The complexity for trimming S_k according to the lowest SINR requirement and reuse condition is $O(|S_k|^2)$ at worst. Hence, the overall time complexity of Algorithm 1 is $O((|C| + |D|)^3)$ in the worst case.

5 Set Partitioning-Based Resource Allocation

In this section we present our resource allocation algorithm based on the obtained partitioning results.

Given $\{S_k\}$ via Algorithm 1, the SINR of individual communication links is also determined. The throughput maximization problem Eq. (2) can be reformulated as

$$\max_{\{w_k\}}\left\{\sum_k\left(\sum_{i\in S_k^c}\frac{w_kW}{\left|S_k^c\right|}\log_2\left(1+\xi_i^c\right)+\sum_{j\in S_k^d}w_kW\log_2\left(1+\xi_j^d\right)\right)\right\}. \tag{19}$$

$$\text{Subject to} \quad \sum_k w_k = N_{sub}, \tag{19a}$$

$$0 \le w_k \le N_{sub}, \tag{19b}$$

We denote the total throughput of S_k by Rs_k, which varies with the number of allocated subchannels. The cellular link members of S_k form a subset S_k^c. We divide the subchannel resource into two parts: $\{w_k^{pre}\}$, which denotes the subchannels reserved for cellular links such that each of them obtains one subchannel; $\{w_k^{rem}\}$, which denotes the allocation for the remaining subchannels. Obviously we have $w_k^{pre} = \left|S_k^c\right|$ and the optimization problem Eq. (19) is further reformulated as

$$\text{Maximize} \quad \sum_k Rs_k\left(w_k^{pre}+w_k^{rem}\right) \tag{20}$$

$$\text{Subject to} \quad 0 \le w_k^{rem} \le N_{sub}-|C|, \tag{20a}$$

$$\sum_k w_k^{rem} = N_{sub}-|C|, \tag{20b}$$

$$\text{given} \quad w_k^{pre} = \left|S_k^c\right|, \sum_k w_k^{pre}=|C|.$$

Note that the cellular link members of set S_k equally split the subchannel resources, and here Rs_k is non-linear with w_k^{rem}. We use $w_k^* = \left\lfloor\frac{w_k^{pre}+w_k^{rem}}{|S_k|}\right\rfloor$ to denote the number of subchannels that each cellular link member of set S_k can obtain while allocating w_k^{rem} remaining subchannels to S_k, where $\lfloor\cdot\rfloor$ is the floor function. The total throughput of S_k is given by

$$\begin{aligned} R_{S_k} = {}& w_k^*W\sum_{i\in S_k^c}\log_2\left(1+\frac{P_ig_{\langle i,i\rangle}}{w_k^*N_0+\frac{w_k^*}{w_k}\sum_{l\in S_k^d}P_lg_{\langle l,i\rangle}}\right)\\ &+ w_kW\sum_{j\in S_k^d}\log_2\left(1+\frac{P_jg_{\langle j,j\rangle}}{w_kN_0+\sum_{l\in S_k/\{j\}}P_lg_{\langle l,j\rangle}}\right). \end{aligned} \tag{21}$$

The optimization problem proposed in Eq. (21) can be simplified to an integer knapsack problem which is NP-complete, as proved in Section 3.3. Thus, we propose two heuristics:

1)Resource Allocation — RA1: We develop an efficient resource allocation algorithm to solve this problem based on dynamic programming, as shown in Algorithm 2. Since Rs_k varies with the number of allocated subchannels, we can calculate an array $R_{sum,k}$ that records the total throughput of the resource reuse set S_k with the number of allocated subchannels w_k^{rem} ranging from 0 to $N_{sub}-|C|$.

If we combine two resource reuse sets S_i and S_j into a larger one S^*, the array R^*_{sum}, which records the total throughput of the resource reuse set S^*, can be given by

$$\begin{aligned} R^*_{sum}(l) &= \max\left\{R_{sum,i}(h)+R_{sum,j}(l-h)\right\},\\ & 0\le h\le l, 0\le l\le N_{sub}-\sum_k w_k^{pre}. \end{aligned} \tag{22}$$

2)Resource Allocation — RA2: Since RA1 aims at finding the resource allocation strategy $\{w_k^{rem}\}$ to achieve the maximum network throughput, it possibly allocates some resource reuse sets zero subchannel (not each reuse set contains cellular links). To this end, we propose an alternative algorithm, denoted by RA2, in order to further improve D2D access rate while just slightly compromising the network throughput. The benefits of a high D2D access rate are twofold. From the user's perspective, it is critical for the D2D service availability and user experience. For example, a high D2D access rate can guarantee more users the basic communication requirement, such as the text information transmission in some social network applications. From the system's perspective, it has the potential to achieve high revenue of network operators by serving more D2D users [11], [23]

Algorithm 2. Resource Allocation (RA1)

for all $S_k \in \{S_k\}$ **do**
  for $i = 0 : N_{sub}-|C|$ **do**
    calculate $R_{sum,k}(i+1)$, the sum date rate of S_k when $w_k = i+\left|S_k^c\right|$;
  end for
end for
$R^*_{sum} = R_{sum,1}$;
for $i = 1 : N_{set}-1$ **do**
  for $j = 0 : N_{sub}-|C|$ **do**
    calculate $R^*_{sum}(j+1) = \max_{0\le p\le j}\left\{R^*_{sum}(p+1)+R_{sum,i+1}(j-p+1)\right\}$;
  end for
end for
calculate $R_{sum,\max} = R^*_{sum}(N_{sub}-|C|+1)$;
calculate $\{w_k\}$ to achieve the sum data rate $R_{sum,\max}$;

Algorithm 3. Resource Allocation (RA2)

calculate n_{cset}, the number of set S_k^c, which satisfy $S_k^c=\emptyset$;
$rsize = 1$;
while $N_{set} > N_{sub}-|C|-n_{cset}$ **do**
  remove S_k, which satisfy $S_k^c=\emptyset$, $\left|S_k^d\right|=rsize$, from $\{S_k\}$;
  calculate N_{set};
  $rsize = rsize+1$;
end while
for all $S_k \in \{S_k\}$ **do**
  for $i = 0 : N_{sub}-|C|-n_{cset}$ **do**
    calculate $R_{sum,k}(i+1)$, the sum date rate of S_k when $w_k = i+\max\left\{\left|S_k^c\right|, 1\right\}$;
  end for
end for
$R^*_{sum} = R_{sum,1}$;
for $i = 1 : N_{set}-1$ **do**
  for $j = 0 : N_{set}-|C|-n_{cset}$ **do**
    calculate $R^*_{sum}(j+1) = \max_{0\le p\le j}\left\{R^*_{sum}(p+1)+R_{sum,i+1}(j-p+1)\right\}$;
  end for
end for
calculate $R_{sum,\max} = R^*_{sum}(N_{sub}-|C|-n_{cset}+1)$;
calculate $\{w_k\}$ to achieve the sum data rate $R_{sum,\max}$;

The main idea of RA2 is to prune small resource reuse sets to make sure each remaining resource reuse set can get at least one subchannel and also guarantee the resources allocated to cellular links. More specifically, we pre-allocate $w_k^{pre} = \max\{|S_k^c|, 1\}$ for each remaining resource sharing set S_k. Here we use $rsize$ to denote the size of sets we need to prune. $rsize$ is set to be 1 at first. We prune the sets which contains only 1 link and check whether the number of subchannels is enough to make sure each remaining set can get one subchannel and also guarantee the resource allocated to cellular links. If the number of subchannels is enough to serve the remaining sets and guarantee the resources for cellular users, RA2 will allocate subchannel resources for those resource sharing sets. If not, the value of $rsize$ will increase and the sets whose size are equal to $rsize$ will be pruned. It is noted that since we keep large reuse sets and guarantee resources allocated to them, RA2 can achieve a higher access rate with the network throughput no less than the lower bound claimed in Eq. (18).

3) Complexity Analysis: We turn to the complexity of the above two algorithms (RA1 and RA2). RA1 can be further divided into two steps: calculating the R_{sum} table, which denotes the sum data rate for R_k with varying number of subchannel resources; applying Eq. (23) to find the maximum resource allocation for $\{S_k\}$. The computational complexity of these two steps are $O((N_{sub}-|C|+1)\ (|C|+|D|))$ and $O\big(\frac{N_{set}}{2}\ \ (N_{sub}-|C|+2)\ (N_{sub}-|C|+1)\big)$ respectively, where N_{set} stands for the number of resource sharing sets. As the number of subchannels of a system is a constant and N_{set} is no more than $|C|+|D|$, the number of communication links, the computational complexity of RA1 is $O(|C|+|D|)$. Compared with RA1, there is an extra step to prune some small resource sharing sets for RA2. The rest two steps are the same with RA1. The computational complexity of the three steps of RA2 are $O(N_{set})$, $O((N_{sub}-|C|-n_{cset}+1)$ $(|C|+|D|))$ and $O\big(\frac{N_{set}}{2}(N_{sub}-|C|-n_{cset}+2)(|N_{sub}|-|C|+1)\big)$ respectively, where n_{cset} stands for the number of sets which do not contain cellular links. Thus, the computational complexity of RA2 is also $O(|C|+|D|)$.

6 Performance Evaluation

We have conducted extensive simulations to evaluate the performance of FRESH. We use FRESH-RA1 and FRESH-RA2 to denote the FRESH scheme based on RA1 and RA2 proposed in Section 5. We compare our proposed FRESH-RA1 (and FRESH-RA2) with the *user-oriented* scheme [11], the *resource-oriented* scheme [18], and an equally splitting scheme serves as the baseline. We use two widely-adopted metrics [11]: the *network throughput* that is the sum of the throughput of all cellular users and D2D pairs; the *D2D access rate* that is the ratio of the number of admitted D2D pairs to the number of all D2D pairs.

6.1 Simulation Setup

We consider a single cell network where the BS is located at the center of the area and UEs are uniformly distributed in the area while the distance between a pair of D2D users ranges from 1 to 50 m. Our simulation parameters are summarized in Table 2, mostly adapted from [8], [28]. We use an advanced link adaptation technique mentioned in [28], where a proper MCS is selected from the available MCSes (e.g., QPSK, 16QAM and 64QAM) with different coding rates ranging from 1/12 to 11/12 according to the estimated SINR value. Each MCS has a SINR threshold value that corresponds to 10 percent BLER (see [28] for more details). Here we take Matlab (version 7.8.0) as the simulator. The simulator was run on a PC with an Intel Core i7-4750HQ CPU at 2.00 GHz, 8 GB of RAM, and the 64bit Windows 10 operating system.

TABLE 2
Simulation Parameters

Parameter	Value
Area size	200 m × 200 m, 500 m × 500m
Carrier frequency	2.0 GHz
System bandwidth	5 MHz
subchannel bandwidth	180 kHz
Number of subchannel	24
Max BS Tx power	20 W (43 dBm)
BS antenna gain	14 dBi
BS noise figure	5 dB
Max UE Tx power	100 mW (20 dBm)
UE antenna gain	0 dBi
UE noise figure	9 dB
Distance between D2D UEs	1 to 50 m
Antenna pattern	Omni
MCS	QPSK: 1/12, 1/9, 1/6, 1/3, 1/2, 3/5 16QAM: 1/3, 1/2, 3/5 64QAM: 1/2, 3/4, 3/5, 5/6, 11/12
Noise density	–174 dBm/Hz
Bandwidth efficiency	0.83
User distribution	Uniform
Number of cellular users	5, 10, 15, 20
Number of D2D pairs	10, 20, 30, . . . , 100

6.2 Network Throughput

Figs. 3a and 3b compare the network throughput of different schemes with different D2D link numbers and also different area size. We run each experiment 10 times, and plot the average value as well as the standard deviation. We have the following two main observations. First, for a fixed area size, the network throughput (both FRESH-RA1 and FRESH-RA2) increases when more D2D links get involved. For example, Fig. 3a shows that the network throughput of FRESH-RA1 has increased from around 39.5 Mbps to 124.7 Mbps when the number of D2D links varies from 10 to 100. The reason is that with more D2D links, both FRESH-RA1 and FRESH-RA2 can find larger resource reuse sets, thereby significantly enhancing the spectrum efficiency. The second observation is that, given the fixed number of users, FRESH-RA1 (or FRESH-RA2) provides higher network throughput when the area size gets larger. The reason is that with larger area size, the distance between users becomes larger and thus the interference gets weaker. Hence FRESH-RA1 and FRESH-RA2 can find larger resource reuse sets. The other schemes, on the other hand, as each subchannel resource can be shared by at most a pair of cellular link and D2D link and thus the spectrum is not efficiently utilized, always achieve network throughput below 40 Mpbs in Figs. 3a and 3b.

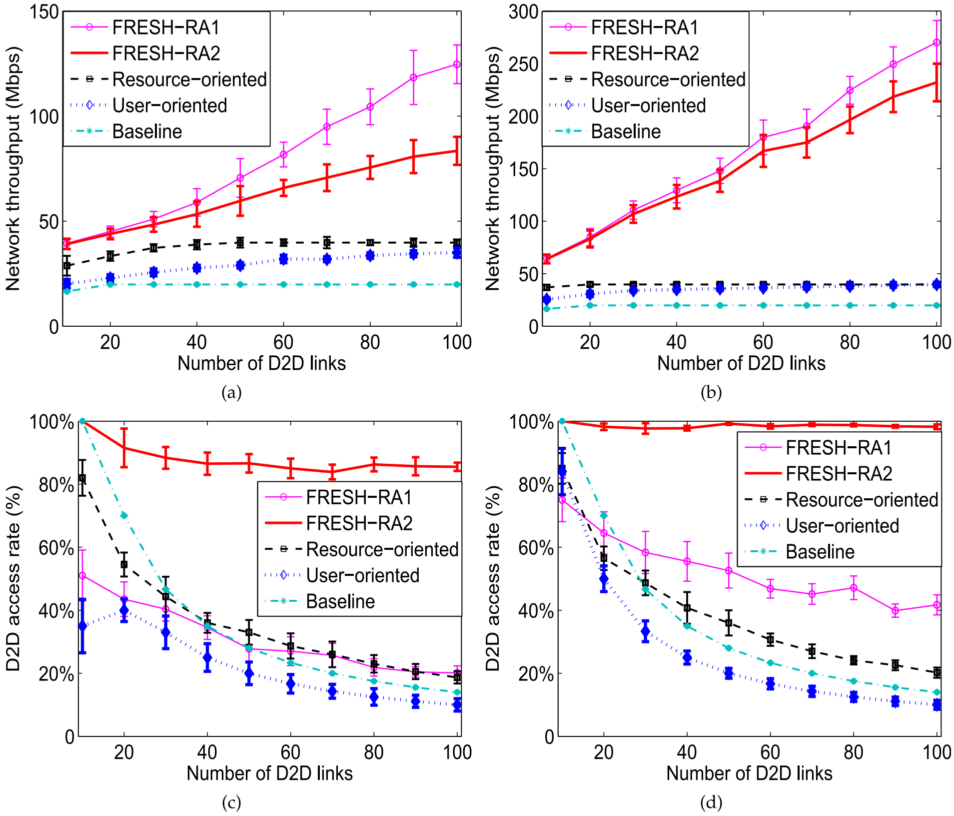


Fig. 3. Network throughput and D2D access rate with different numbers of D2D pairs. (a) Network throughput with 200 m × 200 m area size. (b) Network throughput with 500 m × 500 m area size. (c) D2D access rate with 200 m × 200 m area size. (d) D2D access rate with 500 m × 500 m area size.

We also show the impact of cellular link number on the network throughput in Fig. 4a. Here the network throughput of both FRESH-RA1 and FRESH-RA2 will decrease (and the decreasing will slowdown) with the increasing cellular user number. This is because FRESH will pre-allocate a part of subchannel resource ($\{w_k^{pre}\}$) to the cellular users to guarantee their communication. The increasing cellular user number will reduce the amount of remaining resources ($\{w_k^{rem}\}$) which are allocated by FRESH to the reuse sets which can bring higher throughput and access rate, namely, the flexibility for the resource allocation in FRESH is reduced. For the resource-oriented scheme, the resource allocation is based on say subchannels or resource blocks, rather than the whole link resources. Thus the system throughput will not be much affected, just increase very slightly. For the user-oriented scheme, its resource allocation can be treated as the pairing of D2D pairs and cellular users. When the cellular user number increases, the amount of resources each cellular user can obtain will decrease. For some pairs of cellular user and D2D pair which can bring higher throughput, the resource they can get is reduced. Therefore, the system throughput of user-oriented scheme will decrease with the increasing cellular user number. For the baseline scheme which equally allocate exclusive resource to communication links, each communication link can achieve the highest MCS in their own subchannel. Thus the network throughput will not be influenced by the changing cellular user number.

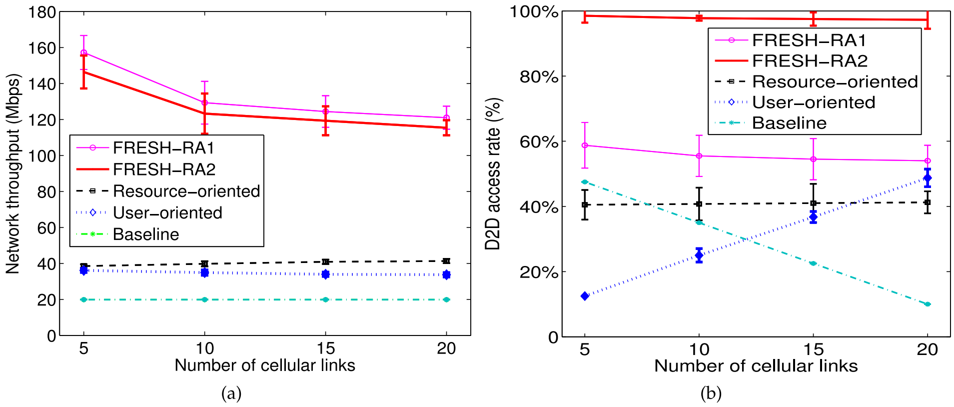


Fig. 4. Network throughput and D2D access rate with different numbers of cellular links. (a) Network throughput with 500 m × 500 m area size. (b) D2D access rate with 500 m × 500 m area size.

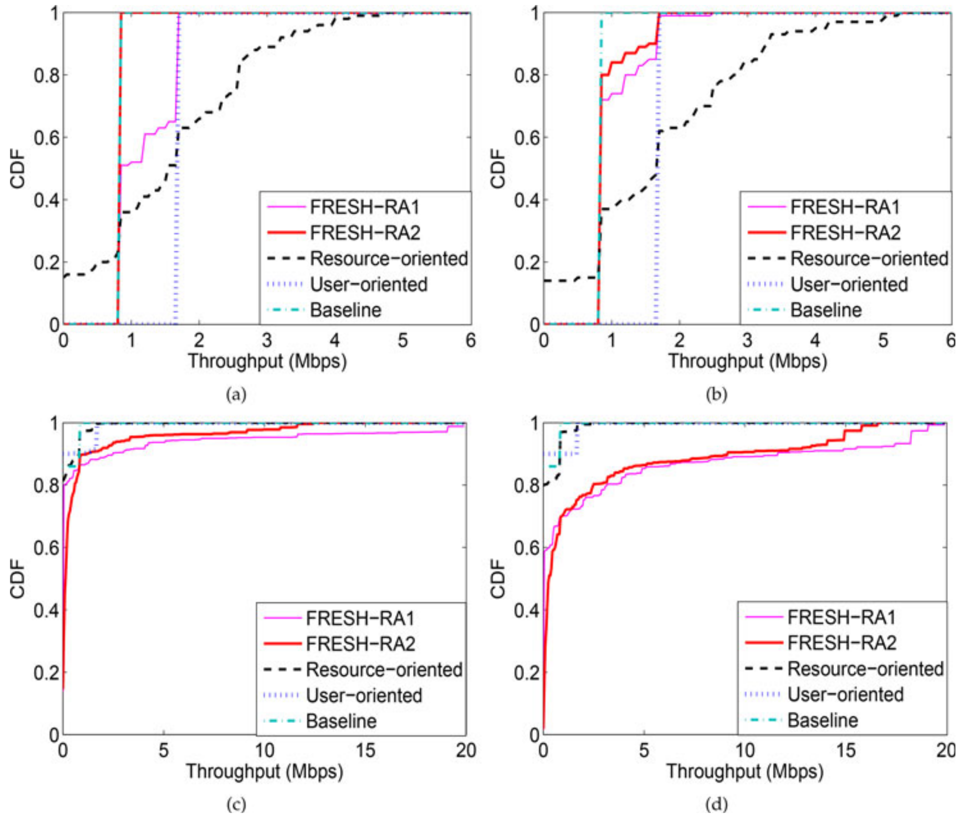


Fig. 5. CDF of data rate for each communication links. a) CDF for cellular users with $N_c = 10$, $N_d = 100$ and 200 m × 200 m area size. b) CDF for cellular users with $N_c = 10$, $N_d = 100$ and 500 m × 500 m area size. c) CDF for D2D pairs with $N_c = 10$, $N_d = 100$ and 200 m × 200 m area size. d) CDF for D2D pairs with $N_c = 10$, $N_d = 100$ and 500 m × 500 m area size.

Fig. 5 depicts the CDF of data rate with 10 cellular links and 100 D2D links in total. Here, FRESH-RA1 and FRESH-RA2 make a sacrifice on the data rate of cellular links for the sake of significant improvements of the data rate of D2D links as compared with other schemes.

6.3 D2D Access Rate

Fig. 3c and Fig. 3d show the D2D access rate of different schemes with a variety of settings. The proposed FRESH-RA1 allocates almost all the subchannels to the largest resource reuse set, and therefore the access rate is limited by the size of maximum resource reuse set. It is observed that FRESH-RA2, by slightly compromising the network throughput, achieves a higher D2D access rate, say nearly 80 percent when the area size is 200m × 200m, and close to 100 percent when the area size is 500m × 500m. Again, the larger area size can help to find larger resource reuse sets, leading to a higher D2D access rate. For the baseline scheme without resource sharing, each subchannel can only serve one link. Thus, the access rate is confined by the number of subchannels. For the user-oriented scheme, the maximum number of accessed D2D links is equal to the number of cellular links. The resource-oriented one provides a more flexible choice for D2D links to reuse a portion of the spectrum resources occupied by cellular links. As both the user-oriented and resource-oriented schemes assign each subchannel to at most a pair of cellular link and D2D link, at most N_{sub} cellular links can be served and at most N_{sub} D2D links can be accessed, as claimed in Fig. 1. On the other side, FRESH desirably allows each subchannel to be shared by multiple D2D links, achieving higher access rate compared with the other three schemes.

The impact of cellular link number on the D2D access rate is illustrated in Fig. 4b. As we have mentioned previously, the increasing cellular user number will decrease the amount of remaining resource that FRESH will allocate to the reuse sets with better D2D access rate. Thus the D2D access rate of FRESH will decrease with the increasing cellular user number. For the resource-oriented scheme, as the resource allocation is based on the subchannels or resource blocks, the D2D access rate will not much affected. For the user-oriented scheme which tries to pair cellular users and D2D pairs, the number of D2D pairs which can be served (at most) is equal to the number of cellular users. Thus the D2D access rate of user-oriented scheme increase linearly as cellular user number increases. For the baseline scheme, as we need to give higher priority to the resource allocation for cellular users, the increasing cellular user number will incur the decreasing of the resources left for D2D links and therefore the D2D access rate.

6.4 Complexity Comparison

As we mentioned previously in Section 3.3, there is no polynomial time algorithm can approximate the optimal result of the overall network throughput maximization problem proposed in Eq. (2). Therefore, considering the difficulty for obtaining the approximate solution and optimal solution, we decompose the overall network throughput optimization problem into two subproblems, which are reuse set partitioning and resource allocation for reuse sets, and propose

TABLE 3
Comparison for the Running Time (Area Size 500 m $\times$ 500 m, $N_c = 10$, $N_d = 10, 20, \ldots, 100$)

		number of D2D links (N_d)									
		10	20	30	40	50	60	70	80	90	100
FRESH-RA1	**avg(sec)**	0.233	0.543	0.988	1.873	2.814	4.209	6.231	8.282	11.084	14.106
	std(sec)	0.062	0.159	0.248	0.317	0.501	0.589	0.643	0.967	0.892	1.415
FRESH-RA2	**avg(sec)**	0.232	0.519	0.963	1.800	2.667	4.052	5.916	7.929	10.520	13.808
	std(sec)	0.056	0.142	0.235	0.273	0.429	0.497	0.548	0.829	0.707	1.231
Resource-oriented	**avg(sec)**	15.167	15.350	17.293	15.693	15.666	15.765	15.806	15.730	15.833	15.864
	std(sec)	0.500	0.358	2.353	0.352	0.431	0.325	0.374	0.261	0.346	0.324
User-oriented	**avg(sec)**	0.701	1.485	2.395	3.431	4.692	6.052	7.841	9.829	11.975	14.500
	std(sec)	0.160	0.193	0.238	0.254	0.346	0.444	0.374	0.578	0.641	0.619
Baseline	**avg(sec)**	0.088	0.091	0.093	0.098	0.100	0.103	0.108	0.110	0.114	0.119
	std(sec)	0.006	0.003	0.002	0.003	0.002	0.003	0.002	0.003	0.004	0.008

heuristic solutions for those two subproblems. Also for the same consideration, we have not compared our solution with the optimal result. The computational complexities of our resource reuse set partitioning problem and our resource allocation algorithms are $O((|C|+|D|)^3)$ and $O(|C|+|D|)$ respectively. Thus the final computational complexity of our FRESH scheme is $O((|C|+|D|)^3)$.

In [11], the optimal resource allocation problem is transformed into a maximum weight bipartite matching problem and solved by the Kuhn-Munkres algorithm. Thus the computational complexity of the user-oriented scheme in [11] is $O((|C|+|D|)^3)$ [29]. In [18], the joint mode selection and resource allocation problem is solved through particle swarm optimization method. The authors propose a coding method to reduce the search space dimension from $2N(|C|+|D|)$ to $2N$. Thus, when the maximal number of iterations is limited, the computational complexity of the resource-oriented scheme in [18] is $O(TMN)$, where T is the maximal number of iterations, and M is the number of particles. In our simulations, $T = 500$ and $M = 30$. However, as the problem which needs to be solved is very complex, the particle swarm optimization algorithm can not well converge and usually takes more time than the user-oriented scheme and our FRESH scheme. Table 3 shows a comparison for the running time of different schemes.

Hence, compared with the user-oriented and the resource-oriented schemes, our FRESH can achieve better network throughput and D2D access rate without increasing the computational complexity.

7 Discussion

Consistent with previous studies [2], [6], [7], [8], [9], [10], [11], [12], [13], [15], [16], [17], [18], [23], [25], [30], we in this paper only consider the static scenario. The resource allocation for dynamic scenario, which considers the factors such as the user mobility, can be simplified into the resource allocation for several static scenarios at different time. The time for resource allocation can be decided by different scheduling algorithms. For example, a straightforward extension to adapt to dynamic scenario is to perform the resource allocation periodically. Here the idea of periodical execution is also widely used in routing schemes [31] .

The second issue is about the privacy of D2D communication. Similarly with previous works, our proposed resource allocation scheme is centralized. Although the resource allocation is done by BS, to make D2D communications not transparent to the cellular network, the source of D2D pair can use the destination's public key to encrypt transmitted message [32].

At last, we envision that our proposed model and algorithms can be applied in a real LTE-advanced system. Several works have been proposed to study how to enable the D2D communication and resource management in LTE-advanced systems. For example, in [33], the authors proposed detailed mechanisms for D2D communication session setup and management, and resource sharing between cellular user and D2D pair.

8 Conclusion

In this paper, we break the limitation for *user-oriented scheme* and *resource-oriented scheme* that each subchannel can only be assigned to at most a pair of cellular link and D2D link, and propose FRESH, a full resource sharing scheme that allows a subchannel to be shared by a cellular link and an arbitrary number of D2D links. We formulate the throughput maximization problem in this scenario and further divide this complex problem into two subproblems, namely interference-aware set partitioning and resource allocation. We develop the theoretical foundations of interference-aware set partitioning on arbitrating the tradeoff between the network throughput gain (by allowing further spectrum sharing) and performance degradation (from accumulative interference). Accordingly, we propose a heuristic algorithm for interference-aware set partitioning such that the users in the same set will operate on the same spectrum. Based on the obtained partitioning results, efficient algorithms for resource allocation are proposed aiming at high network throughput as well as D2D access rate. Extensive simulation results have demonstrated that FRESH significantly improves system performance in terms of both network throughput and D2D access rate, and thus pushes the limit of D2D communication underlaying cellular networks.

Appendix A
Proof of Theorem 1

There are $2^{|C|+|D|}$ possible set partitioning results in all. We use $\{\mathfrak{S}_k\}$ to denote all the possible set partitioning results (i.e., $\{\mathfrak{S}_k\}$ contains all the subsets of $C \bigcup D$). Even if we have prior knowledge of the total data rate $\{\mathfrak{R}_k\}$ of $\{\mathfrak{S}_k\}$ while allocating the same number of subchannel resources w^*, the simplified optimal set partitioning problem is equivalent to the weighted maximum independent set problem which has been proved to be NP-complete in a strong sense [34]. Given any instance I of the weighted maximum independent set partitioning problem, we can construct an instance I' of optimal set partitioning problem by setting a possible set partitioning result $\mathfrak{S}_k \in \{\mathfrak{S}_k\}$ in I' to the vertex $v_k \in V$ in I, $\mathfrak{R}_k$ in I' to the weight of v_k. And if $\mathfrak{S}_i$ and $\mathfrak{S}_j$ in I' have part of the same user communication link members, there exists an edge $e_{i,j} \in E$ between vertex v_i and v_j. Clearly, the reduction can be accomplished in polynomial time. And, it is easy to verify that a feasible solution in I is a feasible solution in I', and vice versa. Thus, the optimal set partitioning problem is NP-complete.

Appendix B
Proof of Theorem 2

If the total throughput of set S_k, which is denoted by $Rs_k(w_k)$, varies with the subchannel resource allocated to S_k, which is denoted by w_k, in an linear fashion (a simplified case), the optimal resource allocation problem later defined in Eq. (21) is equivalent to the bounded integer knapsack problem—a well-known NP-complete problem [34]. Given any instance I of the bounded integer knapsack problem, we can construct an instance I' of the optimal resource allocation problem by setting resource reuse set S_k in I' to the item k in I, w_k, the number of subchannnel allocated to S_k, in I' to the weight of item k in I, the total throughput of S_k in I' to the profit of item k with a weight w_k in I. The rest proof for showing the optimal resource allocation problem to be NP-complete is similar to that of the first problem.

Appendix C
Proof of Theorem 3

As we have discussed above, the simplified first subproblem is still a weighted maximum independent set problem which is NP-complete [34]. [35] proposes that unless $P = NP$, there is no polynomial-time ρ-approximation ($\rho \geq 1$) algorithm for NP-complete problem. [36] proofs a more detailed conclusion that "no polynomial time algorithm can approximate the optimal result of the maximum clique problem within a factor of $O(n^\epsilon)$ ($\epsilon > 0$, n is the number of vertexes in the graph), unless $P = NP$" Here, the clique problem and the independent set problem are complementary. Thus Theorem 3 is proofed.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 61271226, 61272410, 61202460, 61572219, and 61502192; by the National High Technology Research and Development Program ("863" Program) of China under Grants 2014AA01A701 and 2015AA011303; by the National Natural Science Foundation of Hubei Province under Grant 2014CFA040; and by the Fundamental Research Funds for the Central Universities under Grants 2015QN073 and 2016YXMS293. The corresponding author of this paper is Xiaoqiang Ma.

References

[1] Cisco visual networking index: Global mobile data traffic forecast update, 2014–2019. [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html
[2] P. Janis "Device-to-device communication underlaying cellular communications systems," *Int. J. Commun., Netw. Syst. Sci.*, vol. 2, no. 3, pp. 169–178, Sep. 2009.
[3] 3GPP, "3rd generation partnership project," *Technical specification group SA, Feasibility study for proximity services (prose) (release 12). TR 22.803 v12.2.0*, Jun. 2013.
[4] 3GPP, "3rd generation partnership project," *Technical specification group SA, Study on architecture enhancements to support proximity services (prose) (release 12), TR 23.703 v2.0.0*, Mar. 2014.
[5] 3GPP, "3rd generation partnership project," *Technical specification group services and system aspects, Group communication system enablers for LTE (GCSE LTE) (release 12), TS 22.468 v12.0.0*, Jun. 2013.
[6] P. Janis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-aware resource allocation for device-to-device radio underlaying cellular networks," in *Proc. IEEE Vehicular Technol. Conf.*, Apr. 2009, pp. 1–5.
[7] S. Xu, H. Wang, T. Chen, Q. Huang, and T. Peng, "Effective interference cancellation scheme for device-to-device communication underlaying cellular networks," in *Proc. IEEE Vehicular Technol. Conf.-Fall*, 2010, pp. 1–5.
[8] C.-H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2009, pp. 1–5.
[9] K. Doppler, C.-H. Yu, C. Ribeiro, and P. Janis, "Mode selection for device-to-device communication underlaying an lte-advanced network," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2010, pp. 1–6.
[10] C.-H. Yu, K. Doppler, C. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.
[11] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Device-to-device communications underlaying cellular networks," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3541–3551, Aug. 2013.
[12] Y. Zhang, et al., "Resource management in device-to-device underlaying cellular network," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2015, pp. 1631–1636.
[13] Y. Zhao, Y. Li, Y. Cao, T. Jiang, and N. Ge, "Social-aware resource allocation for device-to-device communications underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 12, pp. 6621–6634, Dec. 2015.
[14] X. Ma, J. Liu, and H. Jiang, "Resource allocation for heterogeneous applications with device-to-device communication underlaying cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 15–26, Jan. 2016.
[15] B. Marco, F. Gabor, and A. Andrea, "Performance analysis of a distributed resource allocation scheme for D2D communications," in *Proc. IEEE GLOBECOM*, Dec. 2011, pp. 358–362.
[16] R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference-aware graph based resource sharing for device-to-device communications underlaying cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2013.
[17] W. Zhou, X. Sun, C. Ma, J. Yue, H. Yu, and H. Luo, "An interference coordination mechanism based on resource allocation for network controlled device-to-device communication," in *Proc. IEEE/CIC Int. Conf. Commun. China*, Aug. 2013, pp. 109–114.
[18] L. Su, Y. Ji, P. Wang, and F. Liu, "Resource allocation using particle swarm optimization for D2D communication underlay of cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Apr. 2013, pp. 129–133.
[19] H. Nam, M. Ghorbel, and M.-S. Alouini, "Location-based resource allocation for OFDMA cognitive radio systems," in *Proc. IEEE 5th Int. Conf. Cognitive Radio Oriented Wireless Netw. Commun.*, Jun. 2010, pp. 1–5.

[20] B. Zayen, A. Hayar, and G. Oien, "Resource allocation for cognitive radio networks with a beamforming user selection strategy," in *Proc. IEEE Signals, Syst. Comput.*, Nov. 2009, pp. 544–549.

[21] J.-C. Liang and J.-C. Chen, "Resource allocation in cognitive radio relay networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 3, pp. 476–488, Mar. 2013.

[22] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[23] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng, and S. Li, "Optimal resource allocation for device-to-device communications in fading channels," in *Proc. IEEE Global Commun. Conf.*, Dec. 2013, pp. 3673–3678.

[24] A. Roessler, J. Schlienz, S. Merkel, and M. Kottkamp, "LTE-Advanced (3GPP Rel.12) Technology Introduction (White Paper)", 2015.

[25] C. Xu "Efficiency resource allocation for device-to-device underlay communication systems: a reverse iterative combinatorial auction based approach," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 348–358, Sep. 2013.

[26] F. Malandrino, C. Casetti, C.-F. Chiasserini, and Z. Limani, "Fast resource scheduling in HETNETS with D2D support," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 1536–1544.

[27] K. Baum, T. Kostas, P. Sartori, and B. Classon, "Performance characteristics of cellular systems with different link adaptation strategies," *IEEE Trans. Veh. Technol.*, vol. 52, no. 6, pp. 1497–1507, Nov. 2003.

[28] D. Lopez-Peres, A. Ladanyi, A. Juttner, H. Rivano, and J. Zhang, "Optimization method for the joint allocation of modulation schemes, coding rates, resource blocks and power in self-organizing lte networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 111–115.

[29] D. P. Bertsekas, "A new algorithm for the assignment problem," Mathematical Programming, vol. 21, no. 1, pp. 152–171, Dec. 1981.

[30] Q. Chen, S. Zhao, and S. Shao, "Qos-based resource allocation scheme for device-to-device (d2d) communication underlaying cellular network in uplink," in *Proc. IEEE Int. Conf. Signal Process. Commun. Comput.*, Aug. 2013, pp. 1–4.

[31] J. Pan, M. Khan, I. Sandu Popa, K. Zeitouni, and C. Borcea, "Proactive vehicle re-routing strategies for congestion avoidance," in *Proc. IEEE 8th Int. Conf. Distrib. Comput. Sensor Syst.*, 2012, pp. 265–272.

[32] N. Ferguson and B. Schneier, *Practical Cryptography*. Hoboken, NJ, USA: Wiley, 2003.

[33] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to lte-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[34] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-completeness*. San Francisco, CA, USA: Freeman, 1979.

[35] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to algorithms (third edition)*, MIT press, 2009.

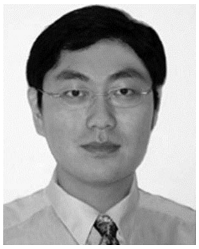
[36] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo, "The maximum clique problem," *Handbook Combinatorial Optimization*, vol. 11, pp. 1–74, 1999.

Yang Yang received the BS and MS degrees from the Wuhan University of Technology, China, in 2009 and 2012, respectively. He is currently working toward the PhD degree in the School of Electronic Information and Communications, Huazhong University of Science and Technology, China. His research interests include algorithms in wireless networks. He is a student member of the IEEE.

Tingwei Liu received the BE degree from the Huazhong University of Science and Technology, China, in 2013. She is currently working toward the ME degree in electronics and information engineering at the Huazhong University of Science and Technology. Her research interests include algorithms and protocols in wireless sensor networks and D2D communication.

Xiaoqiang Ma received the BE degree from the Huazhong University of Science and Technology, China, in 2010. He received the MSc and PhD degrees from Simon Fraser University, Canada, in 2012 and 2015, respectively. He is now with the faculty of Huazhong University of Science as an assistant professor. His areas of interest include wireless networks, social networks, and cloud computing. He is a member of the IEEE.

Hongbo Jiang received the BS and MS degrees from Huazhong University of Science and Technology, China. He received the PhD degree from Case Western Reserve University, in 2008. After that, he joined the faculty of Huazhong University of Science and Technology where he is now a full professor. His research concerns computer networking, especially algorithms and protocols for wireless and mobile networks. He is a senior member of the IEEE.



Jiangchuan Liu received the BE degree (cum laude) from Tsinghua University, Beijing, China, in 1999, and the PhD degree from the Hong Kong University of Science and Technology, in 2003, both in computer science. He is a university professor in the School of Computing Science, Simon Fraser University, British Columbia, Canada, and an NSERC E.W.R. Steacie Memorial fellow. He is an EMC-Endowed Visiting chair professor of Tsinghua University, Beijing, China, (2013-2016). From 2003 to 2004, he was an assistant professor with The Chinese University of Hong Kong. He is a co-recipient of the inaugural Test of Time Paper Award of IEEE INFOCOM (2015), ACM TOMCCAP Nicolas D. Georganas Best Paper Award (2013), ACM Multimedia Best Paper Award (2012), IEEE Globecom Best Paper Award (2011), and IEEE Communications Society Best Paper Award on Multimedia Communications (2009). His students received the Best Student Paper Award of IEEE/ACM IWQoS twice (2008 and 2012). His research interests include multimedia systems and networks, cloud computing, social networking, online gaming, big data computing, wireless sensor networks, and peer-to-peer networks. He has served on the editorial boards of the *IEEE Transactions on Big Data*, the *IEEE Transactions on Multimedia*, the *IEEE Communications Surveys and Tutorials*, the *IEEE Access*, the *IEEE Internet of Things Journal*, the *Computer Communications*, and *the Wiley Wireless Communications and Mobile Computing*. He is the Steering Committee Chair of IEEE/ACM IWQoS from 2015 to 2017. He is a senior member of IEEE.