Sociality-Aware Resource Allocation for Device-to-Device Communications in Cellular Networks

Li Wang*, Lu Liu†, Xianghui Cao†, Xiaohua Tian‡ and Yu Cheng†

*School of Electronic Engineering, Beijing University of Posts and Telecommunications, P.R. China
Email: *liwang@bupt.edu.cn*
†Department of Electrical and Computer Engineering, Illinois Institute of Technology, USA
Email: *lliu41@hawk.iit.edu; {xcao10, cheng}@iit.edu*
‡Department of Electronic Engineering, Shanghai Jiao Tong University, P.R. China
Email: *xtian@sjtu.edu.cn*

Abstract

Exploiting direct transmissions between geographically close mobile users without passing through the base stations (BSs), device-to-device (D2D) communications contribute significant improvement to the spectral efficiency of cellular networks. In D2D assisted cellular networks, the social interaction of mobile users is an important property that will affect the practical performance and should be seriously accounted in the network resource allocation, which is yet to be fully explored. In this paper, we investigate the social interactions for D2D transmissions and develop a contact time model to characterize the D2D links. A D2D link can be considered for resource allocation only when the two users encounter and their contact time is enough long to complete a meaningful transmission. We formulate and compare both sociality-blind and sociality-aware optimization problems for resource allocation in D2D assisted cellular networks. Extensive numerical results are presented, validating that the sociality-aware resource allocation can achieve higher performance than that of the sociality-blind approach.

Index Terms

Device-to-device communication, resource allocation, social-awareness, contact time

I. INTRODUCTION

Aiming to provide high data rate and system capacity for next generation wireless communication systems, the Third Generation Partnership Project (3GPP) has proposed the Long Term Evolution (LTE) and LTE-Advanced (LTE-A) standards to meet the communication requirements [1]. Recently, with the dramatically rising popularity and demands of local area services, Device-to-Device (D2D) communication is proposed as a new component for future cellular networks [2], [3]. Instead of communicating with the help of base stations (BSs), two geographically close D2D users can set up a direct transmission link for local data exchange without the assistance of infrastructures and therefore improve the utilization of spectrum resources. As a promising complement to traditional cellular networks, D2D communications can increase spectral efficiency by exploiting local direct transmission opportunities [4]–[8].

D2D techniques can be implemented as an underlay of cellular networks, such that the interference from D2D transmissions to cellular transmissions should be managed in order to protect cellular users' quality of service [2]. Network resources can be shared by both cellular and D2D transmissions and the above interference can be avoided with LTE's orthogonal frequency-division multiple access (OFDMA) protocol. Such resource sharing will benefit the spectral efficiency and system throughput in the way that D2D can provide relatively higher transmission rate since D2D users are much closer to each other than to the BSs. In addition, limited user demands from BSs imply that it is not necessary for BSs to occupy all the system resources, which further gives opportunities for D2D transmissions.

Unlike the commonly used cellular transmission links where the BSs are operated by service providers and the link availability is guaranteed, the availability of a D2D link relies on the geographical closeness of the corresponding users. Moreover, user density will determine how many D2D links can be set up at the same time. In addition, the distance between a user and its neighbors will affect the link quality since the closer they are, the higher the achievable rate can be. User density and geographical closeness are viewed as physical properties of D2D users.

Aside from the instantaneous closeness and density of mobile users, their social behavior also plays an important role in determining the properties of D2D links. The existence of a neighbor may not necessarily imply that a stable D2D link can be set up because the neighbor, if being a stranger, may not be willing to participate in the D2D transmission or maintain the link over a long time. Taking human mobility into consideration, the neighborhood of a user may vary along time, and so do the D2D links to be established between friends. These observations motivate us to take user sociality as another dimension of user property to characterize the D2D links. The concept of contact time is adopted to describe the sociality-aware link property, which is the length of present time of a D2D link that is available for transmission in view of that this link may not be present all the time. A D2D link between two friends can be considered for resource allocation only when the two users encounter (being geographically close to each other) and their contact time is enough long to complete a meaningful transmission. A longer contact time indicates that the corresponding D2D link can deliver more data. The contact time can be interpreted as the link's strength in light of sociality, as a complement to the physical strength. Both the physical and social properties will affect the D2D transmissions and further influence the network resource allocation results, and thus, should be jointly considered to determine the resource allocation for both cellular and D2D communications. For example, the D2D link between two close users may be preferred since it can provide higher transmission rate; on the other hand, if the link contact time is short, then the resource allocation may not favor this link since the available transmission time is short.

In this paper, we model the sociality of D2D transmission links in terms of contact time as an additional indicator of the link strength. Considering the spectrum resource, we formulate both a sociality-blind and a sociality-aware optimization problems for the resource allocation in D2D assisted cellular networks. In order to maximize resource utilization and network utility, the optimization problems take the sum of flows on all the links including both cellular and D2D links as their objectives under the flow conservation and congestion constraints. The allocation is implemented by OFDMA. The contact time, which is obtained by the sociality model, is imposed as constraints to bound the transmission time of each D2D link. Through extensive numerical results, we show that both physical property (e.g. user density) and social property (contact time) of users will affect the network performance and the resource allocation. With contact time introduced, the resource allocation can be conducted more efficiently and the network performance can be improved.

Our main contributions can be summarized as follows.

1) We develop a contact time model to characterize sociality of D2D transmission links.
2) We formulate both sociality-blind and sociality-aware optimization problems for resource allocation in D2D assisted cellular networks.
3) We present extensive numerical results which demonstrate that the sociality-aware

resource allocation can achieve higher performance than that of the sociality-blind approach.

The remainder of this paper is organized as follows. Section II reviews more related work. Section III and IV present the system model and the contact time model, respectively. The optimization problems for resource allocation are formulated in Section V. Numerical results are presented in Section VI. Finally, Section VII gives the conclusion remarks.

II. RELATED WORK

Many works have attempted to exploit the advantages of D2D technique and improve the spectrum efficiency in cellular networks, where the efforts mainly focus on D2D peer discovery methods, transmission mode selection, power and resource management [9]. The authors in [10] proposed two network-dependent algorithms in discovering D2D pairs, which are based on the signalling message exchange. Mode selection is discussed in [11] to limit the interference from D2D communications to the cellular network, optimal resource sharing mode is selected.

Two major issues faced by D2D communications are the interference management to cellular network and the performance improvement of D2D transmissions [8]. Resource allocation in D2D systems, such as power control and routing, contributes to a large portion of current research in addressing the above issues. The work in [1] provides a solution on the operator side to control the resource allocation and transmission procedure, by targeting on user experience enhancement. By proper power control, [4] proposes the resource allocation where interference is coordinated and resource of cellular user is reused by D2D. The work in [5] effectively improves the throughput without generating interference to cellular networks by optimizing resource allocation and power control under spectrum efficiency restriction and power limitation. In the models of [6] and [7], D2D users opportunistically access the spectrum, whereas the interference from D2D transmissions to primary cellular network is managed based on channel gain information. Authors of [12] propose spectrum sharing protocol that D2D users can assist the transmissions between BS and cellular user, similarly [13] searches for bipartite matching of D2D pair and cellular users in order to maximize network throughput. Xu *et al.* develop an auction-based approach for resource allocation where cellular resources act as bidders and D2D packets are auctioned off as goods, and the method leads to high system sum rate [14]. Authors of [3] propose a game-theoretic approach for radio resource allocation issues, where the problem can be modeled and analyzed with effective mathematical tools and distributed solutions can be obtained. However, throughout existing works, social information has not been exploited for resource allocation.

Sociality-aware approaches are applied in several areas such as data offloading and routing [15]. Traffic offloading process is optimized in [16] based on user encounter histories and content distribution in online social networks, while the offloading is performed by opportunistic communication and social participation in [17]. Social information can also be combined with context information to help user receive the requested content and reduce traffic generated on the network [18]. A forwarding mechanism is proposed in [19] taking advantage of the observation that users with similar interest tend to encounter more often. The data forwarding cost is also reduced in [20] by selecting relay according to a social based capability. Authors in [21] apply social network concepts such as centrality metrics and community formation in designing wireless Ad Hoc network protocols. In our work, social information is exploited to measure the link availability and reliability, based on which resource allocation can be performed to optimize spectrum utilization and improve spectrum efficiency.

III. SYSTEM MODEL

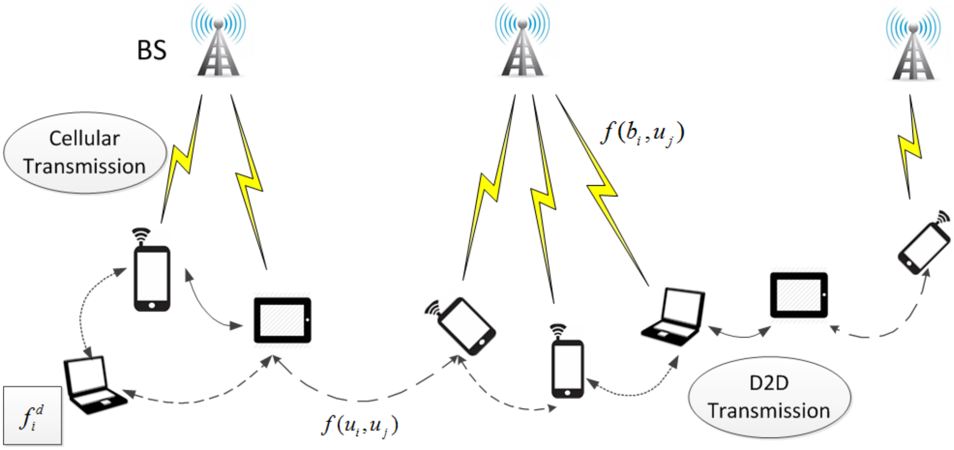


Fig. 1. System architecture. Dashed arrows denote D2D links which are intermittent due to human mobility from a social network perspective.

In a D2D assisted cellular network, suppose user equipments (or users in short) can get data from the BSs by either directly downloading from the BSs or seeking for help from other users' relay in a D2D manner. Therefore, there exist two types of transmissions in the network. One is the direct cellular transmissions from BSs to users which is termed as infrastructure-based transmission or cellular transmission in the rest of this paper. The other type corresponds to D2D transmissions among users. As shown in Fig. 1, the network can be modeled as a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where the vertex set $\mathcal{V}$ consists of all the BSs and users. The edges in $\mathcal{E}$ are the transmission links between BSs and users and between pairs of users. Specifically, if two users have an edge in $\mathcal{E}$, the probability that they will encounter and is able to transmit data between each other is positive; otherwise, there is no D2D link between them. Suppose there is a total number of U users in the network, each of which is denoted as u_j, $j = 1, \ldots, U$. These users are covered by (or associated to) B BSs with each BS denoted as b_i, $i = 1, \ldots, B$.

We consider a slotted system, where a unit time slot represents the time window in which we perform resource allocation for the BSs and mobile users. For simplicity, we ignore the hand-off procedure for a user moving from one BS's service area to another. In each time slot, we assume that each user is associated with (i.e., can download data from) only one BS, which is also its nearest BS. As in the network graph, each user has an incoming edge from only one BS. In this model, we shall focus on downlink transmissions from BSs to users. Therefore, each BS only has outgoing edges (representing the cellular transmissions) toward its associated users.

Unlike direct cellular transmissions, the device power is usually limited so that the D2D transmission only can be established between two approximal users if they are within each other's transmission range. We assume that the D2D transmission ranges are the same, in which case each edge that representing a D2D link is bi-directional. However, our approach can be easily extended to accommodate heterogeneous scenarios where users can use diverse power (which results in directed D2D edges). Define the amount of traffic from vertex i to vertex j as $x_{(i,j)}$. Specifically, $x_{(b_i,u_j)}$ is defined for cellular transmissions and $x_{(u_i,u_j)}$ for D2D transmissions.

Upon receiving a packet, a user u_j, if not the designated destination of the packet, may forward it to its destination via D2D links. However, due to human mobility, the desired D2D links may not always be available. Therefore, we suppose that each user maintains a buffer to store all the packets to be forwarded to others. Once a D2D link is available (i.e., u_j encounters the desired user), it will transmit those buffered data corresponding to it. For each user u_i, denote the amount of traffic destined to it as f_i^d. f_i^d is useful to model the flow balance property as presented later in this paper.

Due to wireless interferences, a D2D transmission cannot take place when a cellular transmission is active in the same band at the same time due to the high power of the BSs. Hence, the spectral resources should be allocated to cellular and D2D transmissions appropriately with techniques such as OFDMA. Denote the percentage of resources allocated to cellular transmissions as $\eta \leq 1$, while that for D2D transmissions as $1-\eta$. Then the cellular and D2D transmissions are carried out in each other's allocated spectrum and there will be no interference between cellular and D2D transmission.

While cellular and D2D communications can be separated in time or spectrum bands, multiple D2D links can deliver data in the same band at the same time. This is because the interference range of each user is limited; and thus a group of D2D links, on which the transmissions are orthogonal and interference-free, can transmit data simultaneously even in the same band. Such groups are usually termed as independent sets [22]. Therefore, the resource allocation should take both interferences between cellular and D2D transmissions and between pairs of D2D links into account. Moreover, in order to account for that the D2D links may not be always available in a slot, the allocation schemes should also take the contact time between user pairs into account.

In this paper, resource allocation is performed in a centralized manner, which can be implemented at a selected BS or through BSs coordination [1]. We assume that the network topology and social information are available at the BSs. Based on this, the BSs can formulate and solve the resource allocation problem. Our goal is to optimize the amount of data obtained by the users via either direct cellular downloading or D2D relaying, while satisfying the constraints stemmed from mutual interference. In the following, we first propose a sociality-aware model for the contact time and then formulate the whole resource allocation problem.

IV. SOCIALITY-AWARE CONTACT TIME MODELING

To account for the effect of sociality of mobile users on the system performance, we model the behavior of D2D links based on the contact time of end users. The D2D link between two users may be disconnected due to the mobility of users and the length of time that they can stay connected is measured by the contact time, in which sense the average contact time is a measure of the link reliability. This model is also used in other related works such as [16].

Due to human mobility, two users may frequently encounter and leave, while they can only contact with and hence forward data to each other through D2D links during their encounter time. Here, we say that two users encounter if they move close into each other's transmission range and can establish a D2D link for data transmission. In this section, we propose a sociality-based network model to characterize the contact time between any two pair of users.

For users i and j, let $\tau^1_{(u_i,u_j)}$ and $\tau^0_{(u_i,u_j)}$ be the encounter time (i.e., the time duration that i and j encounters) and inter-encounter time (i.e., the time duration between two consecutive encounters) between them, respectively. Due to mobility, both $\tau^1_{(u_i,u_j)}$ and $\tau^0_{(u_i,u_j)}$ will vary (randomly) along time. Based on some real experimental studies on human mobilities (e.g., [23]), both the encounter and inter-encounter time can be modeled as power law distributions over $[T_1, \infty)$ and $[T_0, \infty)$ with heavy tail indices α_1 and α_0, respectively [24], [25]. Specifically, we have

$$\mathbb{P}\left[\tau^1_{(u_i,u_j)} > t\right] = \left(\frac{t}{T_1}\right)^{-\alpha_1}, \tag{1}$$

$$\mathbb{P}\left[\tau^0_{(u_i,u_j)} > t\right] = \left(\frac{t}{T_0}\right)^{-\alpha_0}, \tag{2}$$

where $\alpha_1 > 0$ and $\alpha_0 > 0$. T_1 and T_0 represent the minimum encounter and inter-encounter time, respectively. All these parameters can be obtained based on either off-line experiments or data samples in historical time slots. Thus, the average inter-encounter time can be calculated by

$$\begin{aligned}\bar{\tau}^0_{(u_i,u_j)} &= \int_0^\infty \mathbb{P}\left[\tau^0_{(u_i,u_j)} > t\right] dt \\ &= \int_0^{T_0} 1 dt + \int_{T_0}^\infty T_0^{\alpha_0} t^{-\alpha_0} dt \\ &= \begin{cases} \infty, & \text{if } \alpha_0 \le 1 \\ \frac{\alpha_0}{\alpha_0 - 1} T_0, & \text{otherwise.} \end{cases}\end{aligned} \tag{3}$$

Therefore, in order to have a finite mean of $\tau^1_{(u_i,u_j)}$, we must have $\alpha_1 > 1$. Similarly, if $\alpha_0 > 1$, the average encounter time is given by

$$\bar{\tau}^1_{(u_i,u_j)} = \frac{\alpha_1}{\alpha_1 - 1} T_1. \tag{4}$$

Suppose that, if two users i and j encounter, they will attempt to establish a D2D link and contact with each other only if their encounter time is longer than a minimum value, say $\tau_{\min} > 0$, such that a shortest D2D communication can be completed. Let $\tau_{(u_i,u_j)}$ be the contact duration of users i and j. Then,

$$\mathbb{P}\left[\tau_{(u_i,u_j)} > t\right] = \begin{cases} 1, & \text{if } t \le \max\{\tau_{\min}, T_1\} \\ \left(\frac{t}{T_1}\right)^{-\alpha_1}, & \text{otherwise.} \end{cases} \tag{5}$$

If $\tau_{\min} > T_1$, the average contact duration between these two users is

$$\begin{aligned}
\bar{\tau}_{(u_i,u_j)} &= \int_0^{\infty} \mathbb{P}\left[\tau_{(u_i,u_j)} > t\right] dt \\
&= \int_0^{\tau_{\min}} 1 dt + \int_{\tau_{\min}}^{\infty} T_1^{\alpha_1} t^{-\alpha_1} dt \\
&= \tau_{\min}\left[1 + \frac{1}{\alpha_1 - 1}\left(\frac{T_1}{\tau_{\min}}\right)^{\alpha_1}\right].
\end{aligned} \tag{6}$$

Otherwise, $\tau_{\min} \leq T_1$. We have

$$\bar{\tau}_{(u_i,u_j)} = \bar{\tau}^1_{(u_i,u_j)} = \frac{\alpha_1}{\alpha_1 - 1} T_1.$$

Therefore, the average contact time within a unit time slot between users i and j can be written compactly as follows.

$$\begin{aligned}
t_{(u_i,u_j)} &= \frac{\bar{\tau}_{(u_i,u_j)}}{\bar{\tau}^1_{(u_i,u_j)} + \bar{\tau}^0_{(u_i,u_j)}} \\
&= \frac{\tau_{\min}\left[\alpha_1 - 1 + \left(\frac{T_1}{T'}\right)^{\alpha_1}\right](\alpha_0 - 1)}{\alpha_0(\alpha_1 - 1)T_0 + \alpha_1(\alpha_0 - 1)T_1},
\end{aligned} \tag{7}$$

where T' equals to T_1 if $\tau_{\min} \leq T_1$ and $\tau_{\min}$ otherwise. Note that we have assumed $\alpha_1 > 1$ and $\alpha_0 > 1$.

The contact time obtained from this model accounts for the mobility of users, and can be further fine-tuned if additional social information is available. For example, if a user's device will not be turned on for all the time, the contact time should be multiplied by the portion of time that the device is on. If a user is not willing to communicate with a specific neighbor, the contact time of this D2D link will be 0.

Based on this model, the flow allocation can then be performed, where the constraints of resource allocation can be constructed according to the contact time of each D2D link. The details will be discussed in the next section.

V. NETWORK RESOURCE ALLOCATION

In this section, we formulate the optimal resource allocation problem by taking the constraints due to mutual interference into account, where both sociality-blind and sociality-aware link models are analyzed.

A. Problem Formulation

As stated in the previous section, in each slot, the resource allocation aims to maximize the amount of data the users obtained from all the links, including both cellular and D2D links. For user j, the amount of data received from cellular and D2D links are $\sum\limits_{i:(b_i,u_j)\in\mathcal{E}} x_{(b_i,u_j)}$ and $\sum\limits_{i:(u_i,u_j)\in\mathcal{E}} x_{(u_i,u_j)}$, respectively. For the entire network, the objective is to maximize the aggregated throughput of all the users, which is

$$J=\sum_{j=1}^{U}\left(\sum_{i:(b_i,u_j)\in\mathcal{E}} x_{(b_i,u_j)}+\sum_{i:(u_i,u_j)\in\mathcal{E}} x_{(u_i,u_j)}\right) \tag{8}$$

In the following we derive the constraints of the problem. First, the traffic flow should be balanced at each user [26]. The flow conservation constraints are derived from two aspects.

- According to the system model, all the traffic in the network is originated from the BSs and ends at the corresponding destinations of the packets. Therefore, the total amount of outgoing flows from BSs should be equal to that of all incoming flows at all the destinations, i.e.,

$$\sum_{i=1}^{B}\sum_{j:(b_i,u_j)\in\mathcal{E}} x_{(b_i,u_j)}=\sum_{j=1}^{U} f_j^d \tag{9}$$

- Consider the flow balance at each node: the sum of incoming traffic should be equal to the sum of outgoing traffic plus the amount of traffic destined to and received by the node, i.e.,

$$\begin{aligned} & x_{(b_i,u_j)}+\sum_{i:(u_i,u_j)\in\mathcal{E}} x_{(u_i,u_j)} \\ = & \sum_{i:(u_j,u_i)\in\mathcal{E}} x_{(u_j,u_i)}+f_j^d \quad j=1,\ldots,U \end{aligned} \tag{10}$$

Naturally, the amount of traffic on each link should be non-negative, i.e., $\forall b_i, u_j$ and u_i,

$$x_{(b_i,u_j)}\geq 0, \quad x_{(u_i,u_j)}\geq 0. \tag{11}$$

During the optimization, it is possible that a user, if far away from the BSs and the other users, will be isolated without incoming traffic, in order to optimize the total traffic. At the same time, every user in the network demands at least f^{lb} amount of data originated from the BSs. Therefore, we have the following constraint for each user $i\in\{1,2,\ldots,U\}$.

$$f_i\geq f^{lb} \tag{12}$$

For each BS b_i, the transmission rate from b_i to any of its associated user u_j can be calculated by

$$R(b_i, u_j) = W_b \log_2 \left(1 + \frac{P_{b_i} d^{-\gamma}_{(b_i,u_j)}}{N_0}\right) \tag{13}$$

where W_b is the bandwidth allocated to the cellular transmission, P_{b_i} is the transmission power of b_i, $d_{(b_i,u_j)}$ is the distance between b_i and u_j, γ is the path loss exponent and N_0 is the noise power, respectively. If the system bandwidth is W, then $W_b = \eta W$.

Given the transmission rate, the service time from b_i to u_j is $\frac{x_{(b_i,u_j)}}{R_{(b_i,u_j)}}$. Each BS may have to serve multiple users. Since the cellular transmissions are carried out within the same band, a BS cannot transmit to multiple users simultaneously and users associated with the same BS have to share the service time of this BS. Within one time slot, suppose the length of time that is allocated to cellular transmission of b_i is t_b. Then, for any $i = 1, \ldots, B$, the total service time offered to all the associated users of b_i cannot exceed t_b. Therefore, we have the following BS service time constraint:

$$\sum_{j:(b_i,u_j)\in\mathcal{E}} \frac{x_{(b_i,u_j)}}{R_{(b_i,u_j)}} \leq 1 \quad i = 1, \ldots, B \tag{14}$$

where the length of a time slot is assumed to be 1 unit for simplicity.

B. Sociality-Blind Resource Allocation

We first consider the resource allocation problem which is blind of sociality in D2D links. In other words, as long as two users are within each others' transmission rage, the D2D link between them will always exist, with full contact time.

Since all the D2D links share the same spectrum, the links within each others' transmission ranges cannot be active at the same time instant; the links out of each others' transmission ranges may work at the same time but generate interference to each other.

For each D2D link, the other links that share common node with either transmitter or receiver will conflict with the tagged link and these links cannot transmit simultaneously. Define such a set of links which are conflicting with (u_i, u_j) as $E^{Conf}_{(u_i,u_j)}$. During the transmission of (u_i, u_j), the links in $E^{Conf}_{(u_i,u_j)}$ will be silent, but the links out of the conflicting set may be transmitting and generating interference to (u_i, u_j). Based on this, the transmission rate of link (u_i, u_j) will be

$$R(u_i, u_j) \\ = W_u \log_2 \left(1 + \frac{P_{u_i} d_{(u_i,u_j)}^{-\gamma}}{N_0 + \sum_{(u_x,u_y)\in\mathcal{E}\setminus E_{(u_i,u_j)}^{Int}} P_{u_x} d_{(u_x,u_j)}^{-\gamma}} \right) \quad (15)$$

where $W_u = (1-\eta)W$, which is the spectrum allocated to the D2D transmissions and P_{u_i} is the D2D transmission power of u_i.

Due to channel contention, the transmission time of conflicting links should sum up to at most t_u, which is the time allocated for D2D transmissions. In other words, the total transmission time of link (u_i, u_j) and the links in $E_{(u_i,u_j)}^{Conf}$ should be bounded by t_u. This constraint can be expressed as:

$$\frac{x_{(u_i,u_j)}}{R(u_i, u_j)} + \sum_{(u_x,u_y)\in\mathcal{E}_{(u_i,u_j)}^{Conf}} \frac{x_{(u_x,u_y)}}{R(u_x, u_y)} \leq 1 \quad (16)$$

which holds for any $(u_i, u_j) \in \mathcal{E}$.

C. Sociality-Aware Resource Allocation

Now we further consider sociality of D2D links in the resource allocation problem. Based on the sociality model in previous section, we can obtain the contact time of each D2D link. The transmission time of a link (u_i, u_j) cannot exceed $t_u t_{(u_i,u_j)}$. The links in $E_{(u_i,u_j)}^{Conf}$ will contend channel with (u_i, u_j) and therefore the actual transmission time of the link is further decreased. For link (u_i, u_j), the conflicting links will contend with it only when (u_i, u_j) is present. In other words, links in $E_{(u_i,u_j)}^{Conf}$ contend with (u_i, u_j) only in $t_{(u_i,u_j)}$ percent of the time. Therefore, $\forall (u_i, u_j) \in \mathcal{E}$, the sociality-aware D2D congestion constraint can be written as

$$\frac{x_{(u_i,u_j)}}{R(u_i, u_j)} + t_{(u_i,u_j)} \sum_{(u_x,u_y)\in\mathcal{E}_{(u_i,u_j)}^{Int}} \frac{x_{(u_x,u_y)}}{R(u_x, u_y)} \leq t_{(u_i,u_j)}. \quad (17)$$

Then, the formulation of the sociality-aware resource allocation problem (ARAP) is the same as BRAP except that the last constraint in BRAP is replaced by Eq. (17). Finally, the resource allocation problems can be written as the following programming in Fig. 2. In the resource allocation problem, the objective function is the total throughput of all users, where the amount of flow on each cellular link ($x_{(b_i,u_j)}$) and D2D link ($x_{(u_i,u_j)}$) acts as the variable for optimization.

$$\max_{\substack{\{x_{(b_i,u_j)}\}\\ \{x_{(u_i,u_j)}\}\\ \{f_i^d\}}} \quad : \quad J \text{ as defined in (8)}$$

subject to:

$$x_{(b_i,u_j)} + \sum_{i:(u_i,u_j)\in\mathcal{E}} x_{(u_i,u_j)} = \sum_{i:(u_j,u_i)\in\mathcal{E}} x_{(u_j,u_i)} + f_j^d, \quad j = 1,\ldots,U$$

$$f_i \geq f^{lb}, \quad i = 1,2,\ldots,U$$

$$\sum_{j:(b_i,u_j)\in\mathcal{E}} \frac{x_{(b_i,u_j)}}{R_{(b_i,u_j)}} \leq 1, \quad i = 1,\ldots,B$$

$$\textbf{sociality-blind:} \quad \frac{x_{(u_i,u_j)}}{R(u_i,u_j)} + \sum_{(u_x,u_y)\in\mathcal{E}^{Conf}_{(u_i,u_j)}} \frac{x_{(u_x,u_y)}}{R(u_x,u_y)} \leq 1, \quad \forall (u_i,u_j)\in\mathcal{E}$$

$$\textbf{sociality-aware:} \quad \frac{x_{(u_i,u_j)}}{R(u_i,u_j)} + \sum_{(u_x,u_y)\in\mathcal{E}^{Conf}_{(u_i,u_j)}} \frac{x_{(u_x,u_y)}}{R(u_x,u_y)} \leq t_{(u_i,u_j)}, \quad \forall (u_i,u_j)\in\mathcal{E}$$

Fig. 2. Resource allocation problems.

D. *Implementation*

In the formulated optimization problem, the decision variables are the flow on each link. The objective function and constraints are all linear. Therefore the problem falls into the category of linear optimizations. There are many efficient algorithms available in the literature (e.g. delay column generation [27]) to solve linear optimization problems. Then the problem can be solved by off-the-shelf optimization tools such as CPLEX [28].

The global information of network topology and contact time between users can be known by the BSs and therefore will not introduce much communication overhead. After solving the resource allocation problem, the BSs will inform each associated user with the flow allocation result, which can be accomplished by simply perform a downlink transmission to that user. Therefore, the communication overhead can be very low.

In the sociality-blind resource allocation, the amount of flow assigned to each link is only based on the link rate. However, the high-rate link may not be present for a long time if its contact time is short. In this case it will be a waste of resource to assign large amount of flow to the link since the flow may be shut down halfway of the transmission when contact terminates. With the sociality-aware formulation, the contact time is considered in the constraints and more reliable links will be selected for transmission. Such selection is a comprehensive decision balancing both transmission rate and available length of time for transmission.

VI. PERFORMANCE ANALYSIS

We assume mobile users are randomly deployed in a 1000 m $\times$ 1000 m area, which is covered by 4 BSs such that every user is within at least one BS's coverage range. Other network parameters are based on the settings in [14] and summarized in Table I. The user covered by multiple BSs will choose the closest one for cellular transmission. With the network topology, the network graph can be constructed, based on which the resource allocation can be performed. Suppose the contact time of each link is already obtained. Since the D2D links in the network may have different contact time, in this section we use average contact time as a parameter, which is averaged over different links in the network (different from the definition in Section IV). Under a variety of user density, average contact time or resource sharing ratio, the optimization problem is formulated for each case and solved for optimal throughput.

TABLE I

NETWORK PARAMETER SETTINGS

Parameter	Value
Cellular/BS transmission power	46 dBm
BS antenna gain	14 dBi
BS transmission range	400 m
D2D transmission power	23 dBm
User antenna gain	0.5 dBi
D2D transmission range	50 m
Total bandwidth	2MHz
Noise spectrum density	-175 dBm/Hz

A. Contact Time and Throughput

First, the impact of contact time on D2D transmissions is investigated. Since we are focusing on the D2D transmissions, only the throughput on D2D links are measured and displayed. We first fix the number of users on 500 and vary the average contact time from 0.1 to 1 (Fig. 3). For comparison, another set of scenarios is also tested (Fig. 4) where contact time for all the D2D links is assumed to be 1 and the number of users is varied to get comparable throughput as that of the first set.

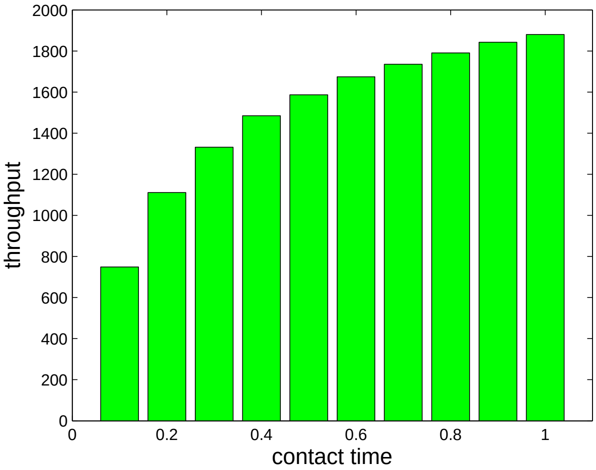


Fig. 3. Throughput of D2D transmissions at different link contact time.

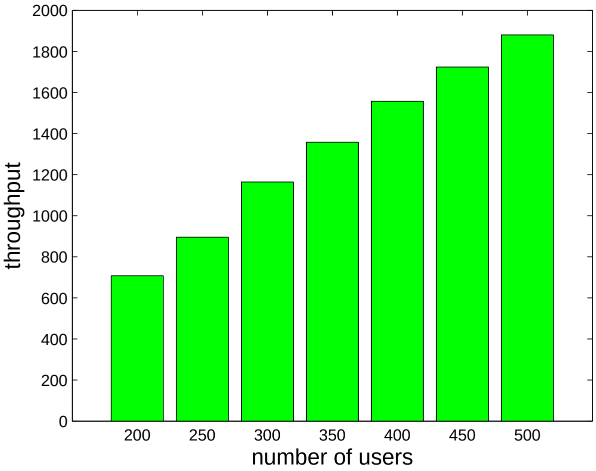


Fig. 4. Throughput of D2D transmissions at different link/user density.

From Fig. 3 and 4 we can observe that both contact time and user density can affect the throughput of D2D transmissions. The user density determines the physical density of D2D links, while the contact time acts as social strength, which affects the D2D throughput in a similar way as that of physical density. For example, the amount of D2D transmission in the case with 400 users and full contact time is close to that of 500 users and half contact time. As a result, the D2D link density in the sociality-aware context is characterized by both physical density and social strength (contact time).

B. Resource Sharing

By varying user density or contact time, we present the relationship between the achieved total throughput (including both cellular and D2D transmissions) and the resource sharing ratio η.

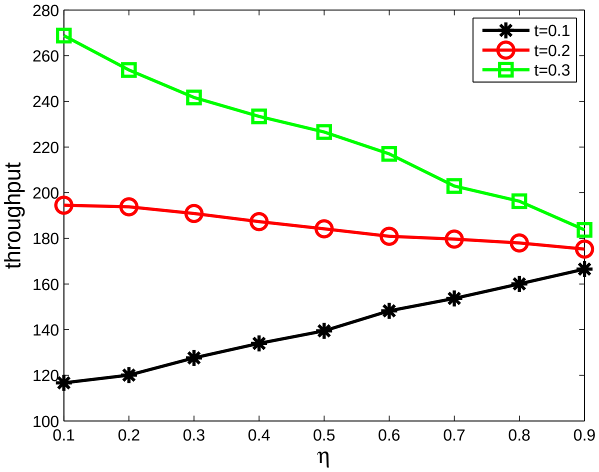


Fig. 5. System throughput with 100 users.

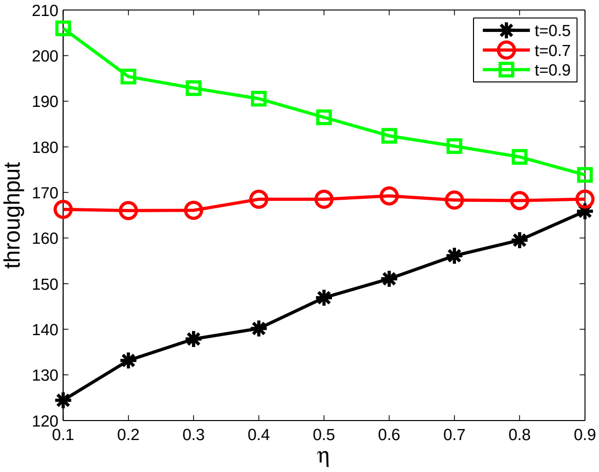


Fig. 6. System throughput with 50 users.

As shown in Fig. 5 and 6, in the scenario either with 100 users and $t = 0.3$ or with 50 users and $t = 0.9$, allocating more resources to D2D transmissions (decreasing the sharing ratio) will improve the throughput. In these cases, there are more users present in the network or the link contact time is long, leading to more D2D transmission opportunities. As a result, D2D transmissions will dominate over cellular transmissions, therefore it is beneficial to assign more resources to the D2D transmissions. On the contrary, in the cases with lower user density or shorter contact time, cellular transmissions become the majority. Consequently, allocating more resources to the cellular transmissions (increasing the sharing ratio) will improve the throughput. In practice, if the sharing ratio is to be decided to achieve higher throughput, then both the user density and D2D link contact time should be taken into consideration to determine which type of transmissions is preferred in order to improve the throughput.

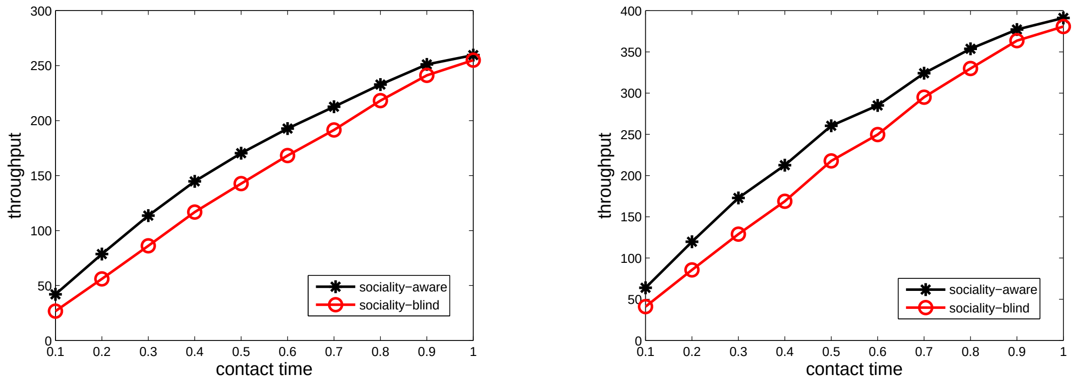
C. Performance Comparison

The sociality-aware and sociality-blind resource allocations are performed and the throughput of D2D transmissions are compared (the throughput for cellular transmission for the two methods are the same). For the sociality-blind method, the resource allocation problem is first solved with full contact time assumption and then the flow on each D2D link is cut down according to its link contact time. The comparisons are carried out under several scenarios with varying contact time and different user numbers or resource sharing ratio, which are shown in the following figures.

From Fig. 7(a) to 7(d), it can be demonstrated that sociality-aware optimization always outperform sociality-blind method. In sociality-blind resource allocation, the amount of flow assigned to each link is only based on the link rate. However, the high-rate link may not be present for a long time if its contact time is short. In this case it will be a waste of resource to assign large amount of flow to the link since the flow may be shut down halfway of the transmission due to the termination of contact time. The drawback of sociality-blind method is even intensified when the contact time is short due to the over-estimation of link availability.

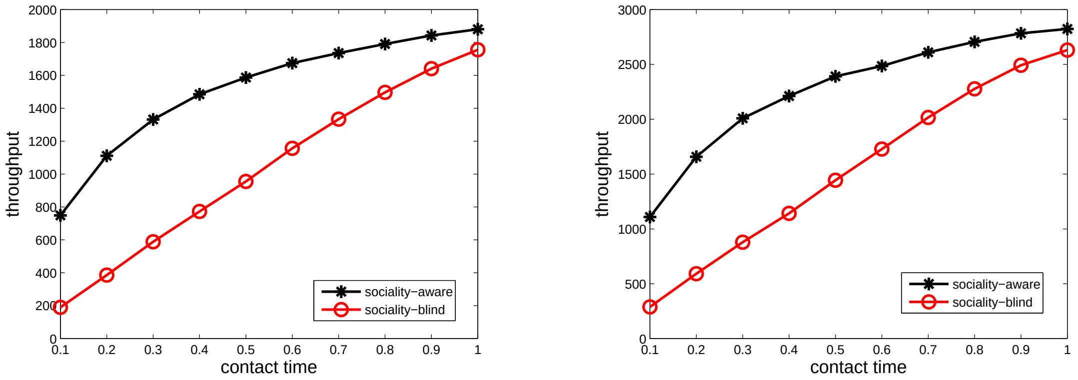
With the sociality-aware formulation, the contact time is considered in the constraints and more reliable links will be selected for transmission. Such selection is a comprehensive decision balancing both transmission rate and available length of time for transmission.

When the number of users is increased or the resource share of D2D transmissions is increased, the advantage of sociality-aware optimization is further enlarged since there are



(a) 100 users, $\eta = 0.6$.

(b) 100 users, $\eta = 0.4$.



(c) 500 users, $\eta = 0.6$.

(d) 500 users, $\eta = 0.4$.

Fig. 7. Performance comparisons between sociality-aware and sociality-blind resource allocation schemes.

more D2D links involved in the resource allocation. In summary, sociality-aware resource allocation can provide better estimation of D2D link quality in the sense of both physical transmission rate and availability, and consequently improve the network performance.

VII. CONCLUSION

In this paper, we have modeled the sociality of D2D transmission links and obtained the contact time as an additional characteristic of link strength, which is then applied in the resource allocation. Two optimization problems for sociality-blind and sociality-aware resource allocation in D2D assisted cellular networks are formulated. In order to maximize the resource utilization and network throughput, the sum of flows on all the links including both cellular and D2D links is optimized by solving the optimization problems, where the contact time obtained by the sociality model, is imposed as constraints to bound the transmission time of each D2D link. Extensive numerical results show that both physical property (e.g. user density) and social property (i.e., contact time) of users affect the network performance and the resource allocation. With the contact time information, the sociality-aware scheme outperforms the sociality-blind one in terms of network throughput.

REFERENCES

[1] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator controlled device-to-device communications in lte-advanced networks," *IEEE Wireless Communnications*, vol. 19, no. 3, pp. 96–104, 2012.

[2] L. Song, Z. Han, and C. Xu, *Resource Management for Device-to-Device Underlay Communication.* Springer Publishing Company, Incorporated, 2013.

[3] L. Song, D. Niyato, Z. Han, and E. Hossain, "Game-theoretic resource allocation methods for device-to-device (d2d) communication," *to appear in IEEE Wireless Communications Magazine*, 2014.

[4] B. Wang, L. Chen, X. Chen, X. Zhang, and D. Yang, "Resource allocation optimization for device-to-device communication underlaying cellular networks," in *IEEE 73rd Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1–6.

[5] C.-H. Yu, K. Doppler, C. B. Ribeiro, and O. Tirkkonen, "Resource sharing optimization for device-to-device communication underlaying cellular networks," *IEEE Trans. on Wireless Commun.*, vol. 10, no. 8, pp. 2752–2763, 2011.

[6] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying lte network," in *IEEE 6th International Conference on Wireless and Mobile Comput., Networking and Communications (WiMob)*, 2010, pp. 368–375.

[7] B. Kaufman, J. Lilleberg, and B. Aazhang, "Spectrum sharing scheme between cellular users and ad-hoc device-to-device users," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 3, 2013.

[8] P. Phunchongharn, E. Hossain, and D. I. Kim, "Resource allocation for device-to-device communications underlaying lte-advanced networks," *IEEE Wireless Communications*, vol. 20, no. 4, 2013.

[9] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 170–177, 2012.

[10] A. Thanos, S. Shalmashi, and G. Miao, "Network-assisted discovery for device-to-device communications," in *Proceedings of the IEEE GLOBECOM 2013 Workshops*, 2013.

[11] K. Doppler, C.-H. Yu, C. B. Ribeiro, and P. Janis, "Mode selection for device-to-device communication underlaying an lte-advanced network," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2010, pp. 1–6.

[12] Y. Pei and Y.-C. Liang, "Resource allocation for device-to-device communications overlaying two-way cellular networks," *IEEE Trans. on Wireless Commun.*, vol. 12, no. 7, pp. 3611–3621, 2013.

[13] D. Feng, L. Lu, Y. Yuan-Wu, G. Li, G. Feng, and S. Li, "Device-to-device communications underlaying cellular networks," *IEEE Trans. on Commun.*, vol. 61, no. 8, pp. 3541–3551, 2013.

[14] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, X. Cheng, and B. Jiao, "Efficiency resource allocation for device-to-device underlay communication systems: a reverse iterative combinatorial auction based approach," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 348–358, 2013.

[15] K. Wei, X. Liang, and K. Xu, "A survey of social-aware routing protocols in delay tolerant networks: applications, taxonomy and design-related issues," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 556–578, 2014.

[16] Y. Zhang, E. Pan, L. Song, W. Saad, Z. Dawy, and Z. Han, "Social network enhanced device-to-device communication underlaying cellular networks," in *IEEE/CIC International Conference on Communications in China-Workshops (CIC/ICCC)*, 2013, pp. 182–186.

[17] B. Han, P. Hui, V. A. Kumar, M. V. Marathe, J. Shao, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Trans. on Mobile Comput.*, vol. 11, no. 5, pp. 821–834, 2012.

[18] C. Boldrini, M. Conti, F. Delmastro, and A. Passarella, "Context-and social-aware middleware for opportunistic networks," *Journal of Network and Computer Applications*, vol. 33, no. 5, pp. 525–541, 2010.

[19] A. Mei, G. Morabito, P. Santi, and J. Stefa, "Social-aware stateless forwarding in pocket switched networks," in *Proceedings of IEEE INFOCOM*, 2011, pp. 251–255.

[20] W. Gao, Q. Li, B. Zhao, and G. Cao, "Social-aware multicast in disruption-tolerant networks," *IEEE/ACM Trans. on Networking (TON)*, vol. 20, no. 5, pp. 1553–1566, 2012.

[21] D. Katsaros, N. Dimokas, and L. Tassiulas, "Social network analysis concepts in the design of wireless ad hoc network protocols," *IEEE Network*, vol. 24, no. 6, pp. 23–29, 2010.

[22] H. Li, Y. Cheng, C. Zhou, and P. Wan, "Multi-dimensional conflict graph based computing for optimal capacity in mr-mc wireless networks," in *IEEE 30th International Conference on Distributed Computing Systems (ICDCS)*, 2010, pp. 774–783.

[23] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and human mobility in conference environments," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, 2005, pp. 244–251.

[24] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. on Mobile Comput.*, vol. 6, no. 6, pp. 606–620, 2007.

[25] T. Karagiannis, J.-Y. Le Boudec, and M. Vojnovic, "Power law and exponential decay of intercontact times between mobile devices," *IEEE Trans. on Mobile Comput.*, vol. 9, no. 10, pp. 1377–1390, 2010.

[26] L. Liu, X. Cao, Y. Cheng, L. Du, W. Song, and Y. Wang, "Energy-efficient capacity optimization in wireless networks," in *IEEE Infocom*, 2014.

[27] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization.* Athena Scientific Belmont, MA, 1997, vol. 6.

[28] IBM ILOG CPLEX Optimizer, http://www.ilog.com/products/cplex/.