A Novel D2D Data Offloading Scheme for LTE Networks

Zehua Wang and Vincent W.S. Wong
Department of Electrical and Computer Engineering
The University of British Columbia, Vancouver, Canada
e-mail: {zwang, vincentw}@ece.ubc.ca

***Abstract*—Downloading remote files (*e.g.*, pictures, videos) from online social networks via smart user equipments (UEs) (*e.g.*, smartphones, tablets) is becoming popular. Friends who are nearby may want to download the same files shared by their mutual acquaintance. People can obtain these files in a device-to-device (D2D) manner via opportunistic connections to reduce their payment for data service. This is referred to as D2D data offloading. However, D2D communications on unlicensed spectrum using Bluetooth or WiFi-Direct may not maintain high data rate when many D2D pairs nearby need to communicate simultaneously. Since D2D connections are transient, it is important to improve spatial reuse of communication resources and increase the data rate of opportunistic D2D communications. In this paper, we propose a scheme to reuse the downlink licensed spectrum of cellular networks for D2D data offloading. Our proposed scheme includes determining the availability of digital files on neighbouring devices, estimating the channel gains, and performing channel allocation and power control for D2D pairs. Simulation results show that our proposed scheme does not affect the existing cellular UEs and it can also offload more data traffic when compared with WiFi-Direct on an unlicensed spectrum.**

I. Introduction

Nowadays, accessing online social networks (OSNs) via smart user equipments (UEs) (*e.g.*, smartphones, tablets) is popular. Mobile users can use their smart UEs to download pictures or videos via Long Term Evolution (LTE) networks. The study in [1] shows that by 2018, the average size of data consumed by a smartphone in each month will reach 2.7 GB and 69% of mobile data traffic will be video streaming. These recreational files may not need to be downloaded at once. Meanwhile, the study in [2] shows that digital files requested by users are closely related with their social relations. Thus, friends who are nearby can obtain delay tolerable files in a device-to-device (D2D) manner by connecting their UEs opportunistically [3]. This is referred to as D2D data offloading.

D2D data offloading using the industrial, scientific and medical (ISM) spectrum has been studied in [4] [5]. On one hand, sharing unlicensed frequency channels without coordination protocols may cause severe interference between multiple transmitter and receiver pairs. On the other hand, when multiple D2D pairs nearby need to communicate simultaneously, distributed coordination usually reserves communication resources for only one D2D pair. Thus, the D2D communication via unlicensed spectrum using Bluetooth or WiFi Direct [6] may not support parallel D2D transmissions nearby. Since D2D connections are transient, it is important to improve spatial reuse of communication resources and increase the data rate for D2D communications.

We refer to the UEs that communicate with the eNB as the *cellular UEs* and refer to the UEs that communicate in D2D manner as *D2D UEs*. The work in [7] allocates licensed spectrum to discover nearby D2D UEs by sending the discovery signals. Yang *et al.* in [8] propose to set up D2D connections based on the UE positions known by LTE networks. Since licensed spectrum in the LTE network is allocated by the base station or the evolved node B (eNB) [9], it is convenient to reuse the licensed spectrum for D2D communications by applying resource allocation and power control [10]–[12]. In particular, the work in [10] uses relay nodes to conduct D2D communications. In contrast, the works in [11] and [12] do not require relay nodes to transfer data between D2D UEs.

In this paper, we use Bluetooth or WiFi-Direct to discover neighbouring UEs and determine the availability of files within the neighbourhood. The size of these signaling packets is actually much smaller than the size of files to be transferred. Thus, the nearby UEs and the available files in a neighbourhood can be discovered efficiently. Moreover, some delay tolerable files may have deadlines. When the deadline is reached, the remaining part of the file has to be downloaded via the LTE network. To offload more data traffic from LTE networks, D2D UEs and the eNB need to sort and schedule these data offloading tasks. When the licensed uplink spectrum in the LTE network is reused between D2D UEs, the cellular UE may be required to increase its transmission power to maintain a target signal to interference plus noise ratio (SINR) at the eNB. Hence, we propose to reuse the licensed downlink spectrum of LTE networks for D2D data offloading.

The licensed downlink spectrum is allocated by eNB in terms of resource blocks (RBs). Both mobile users and wireless service providers can benefit from D2D data offloading. Consider Fig. 1 as an example. There are five UEs (*i.e.,* u_1, u_2, u_3, u_4, and u_5) and one eNB. The payment of data service can be reduced for u_2 and u_3 if they can obtain the digital files from UEs u_1 and u_4, respectively. On the other hand, when the channel gains of communication links between D2D UEs (*i.e.,* $h_{1,2}$ and $h_{4,3}$) are much higher than the channel gains of interference links (*i.e.,* $h_{\mathrm{eNB},2}$, $h_{4,2}$ to u_2, and $h_{\mathrm{eNB},3}$, $h_{1,3}$ to u_3), the RBs used to serve u_5 can be reused by two pairs of D2D UEs with power control [13]. The wireless service provider can allocate these downlink RBs to some

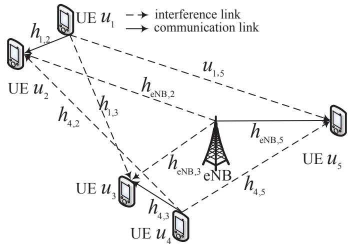

Fig. 1. Reusing the RBs in LTE networks for D2D data offloading.

other cellular UEs and increase its revenue.

In this paper, we propose an optimized scheme to reuse the downlink RBs in LTE networks for D2D data offloading. Our main contributions are summarized as follows:

- We propose to evaluate the utility of offloading demands with their deadlines. We then approve the demands by solving a weighted matching problem in a directed graph.
- We design a scheme to determine the channel gains of communication and interference links by using Zadoff-Chu sequence [14].
- We formulate an optimization problem to jointly allocate RBs and perform power control.
- Simulation results show that our proposed scheme can offload more data traffic when compared with data offloading using WiFi-Direct on an unlicensed spectrum.

The rest of the paper is organized as follows. In Section II, we introduce the system model and present our proposed scheme with modules. We formulate and solve problems for these modules in Section III. The simulation results are presented in Section IV. Conclusion is given in Section V.

II. System Model and Proposed Scheme

We consider each mobile user is equipped with one UE. The terms mobile user and UE are used interchangeably. Let $\mathcal{U}$ denote the set of UEs. Let $\mathcal{F}$ denote the set of all digital files. A popular digital file $k \in \mathcal{F}$ may be requested by multiple users at different time with various delay tolerance. We consider a time slotted system. Let $\mathcal{S}_i^t$ denote the set of digital files that have been completely downloaded and cached by user $i \in \mathcal{U}$ in time slot t. Let $\mathcal{Q}_i^t$ denote the set of files that have been requested by user i and those files have not been completely downloaded in time slot t. Thus, the sets $\mathcal{S}_i^t$ and $\mathcal{Q}_i^t$ change over time. Specifically, when a delay tolerable digital file $k \in \mathcal{F}$ is first being requested by user i in time slot t, file k is included in set $\mathcal{Q}_i^t$ and a deadline $\delta_{i,k} > t$ is specified by user i. If file k has not been completely downloaded via D2D data offloading at time $\delta_{i,k}$, user i will use the LTE network to finish downloading the remaining part of file k. When file k has been downloaded completely at time slot t, the set $\mathcal{Q}_i^t$ is updated by $\mathcal{Q}_i^t \backslash \{k\}$ and the set $\mathcal{S}_i^t$ is updated by $\mathcal{S}_i^t \cup \{k\}$. We assume a UE can cache a number of most recently downloaded digital files in a first-in first-out (FIFO) manner.

Each UE $i \in \mathcal{U}$ can be informed of the existence of other UEs in close proximity by listening to Bluetooth or WiFi-Direct beacons. Let $\mathcal{N}_i^t$ denote the set of neighbouring UEs discovered by UE i in time slot t. For each user $j \in \mathcal{N}_i^t \backslash \mathcal{N}_i^{t-1}$, users i and j exchange the identity of digital files (*e.g.*, uniform resource locator) in sets $\mathcal{S}_i^t$ and $\mathcal{S}_j^t$ to determine the available files on each other, which is referred to as *availability information*. Compared with file transfer, the size of availability information is small and can be exchanged quickly via Bluetooth or WiFi-Direct connections.

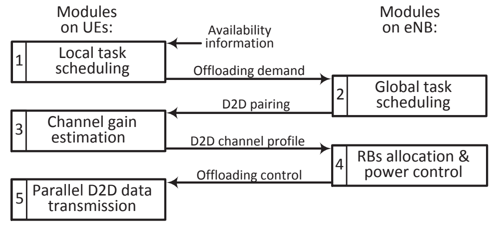

Fig. 2. Basic idea in proposed scheme for D2D data offloading.

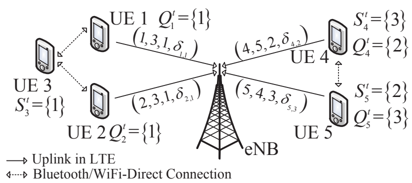

Fig. 3. Examples that not all data offloading demands can be approved.

We now introduce our D2D data offloading scheme, which is compatible with the current LTE networks. The input and output for each module in our proposed scheme are presented in Fig. 2 to show a blueprint. Problems in these modules will be formulated and solved in Section III. With module 1, after determining the availability of digital files in time slot t, each UE $i \in \mathcal{U}$ selects one available file from set $\mathcal{Q}_i^t$ and uploads a *D2D data offloading demand* to eNB. A D2D data offloading demand is defined as a four-tuple in Cartesian space $\mathcal{U}^2 \times \mathcal{F} \times \mathbb{Z}^+$. Consider a data offloading demand $(i, j, k, \delta_{i,k})$ uploaded by UE i as an example. It means UE i requests to download file $k \in \mathcal{Q}_i^t$ from UE $j \in \mathcal{U}$ with a deadline $\delta_{i,k}$. The problem for selecting the digital file will be presented in Section III-A.

Not all D2D offloading demands received by the eNB will be approved. Examples are shown in Fig. 3. Consider UEs 1 and 2 are connected with UE 3 via Bluetooth or WiFi-Direct in time slot t. Assume $\mathcal{Q}_1^t = \mathcal{Q}_2^t = \{1\}$ and both UEs 1 and 2 have been informed $\mathcal{S}_3^t = \{1\}$. Then, the offloading demands $(1, 3, 1, \delta_{1,1})$ and $(2, 3, 1, \delta_{2,1})$ are uploaded by UEs 1 and 2 to eNB, respectively. These two demands cannot be approved simultaneously since UE 3 can only transmit file to one UE. Similarly, if UEs 4 and 5 upload their demands to request files from each other, only one demand can be approved. Let $\mathcal{P}^t$ denote all demands received by eNB in time slot t. Module 2 on the eNB shown in Fig. 2 determines what demands in $\mathcal{P}^t$ are approved. We refer to the output of module 2 as *D2D pairing*, which is a matching [15, Ch. 14] in a weighted directed graph $\mathcal{G}^t = (\mathcal{V}^t, \mathcal{A}^t, w^t)$ with a utility function $w : \mathcal{A}^t \mapsto \mathbb{R}_+$. Specifically, $\mathcal{V}^t = \bigcup_{(i,j,k,\delta_{i,k}) \in \mathcal{P}^t} \{i, j\}$, $\mathcal{A}^t = \bigcup_{(i,j,k,\delta_{i,k}) \in \mathcal{P}^t} \{(j, i)\}$, where (j, i) defined on $\mathcal{U}^2$ denotes a directed edge from j to i. The utility function w^t is defined as $w^t((j,i)) = e^{-\lambda(\delta_{i,k} - t)}$ for each offloading demand $(i, j, k, \delta_{i,k}) \in \mathcal{P}^t$, where $0 < \lambda < 1$ is a system parameter. The proposed utility function is monotonically decreasing since the utility of approving an offloading demand for a file

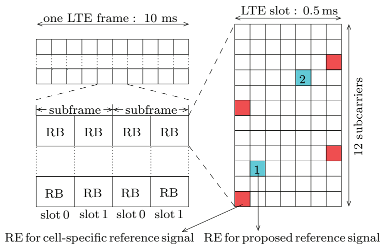

Fig. 4. Proposed reference signals are transmitted on resource elements (REs) in an LTE resource grid.

with small deadline is larger than approving a demand with a longer deadline. This is because approving a demand with a smaller deadline is likely to save more RBs since the file not completely downloaded by its deadline has to be downloaded via the LTE network. Using an exponential function in the utility function is because extending a sufficient large deadline should marginally decrease its utility.

Let directed graph $\hat{\mathcal{M}}^t = (\hat{\mathcal{V}}^t, \hat{\mathcal{A}}^t, \hat{w}^t)$ ($\hat{\mathcal{V}}^t \subseteq \mathcal{V}^t$, $\hat{\mathcal{A}}^t \subseteq \mathcal{A}^t$, $\hat{w}^t : \hat{\mathcal{A}}^t \mapsto \mathbb{R}_+$) denote the D2D pairing information in time slot t, which is determined at eNB by solving a maximum weight matching problem formulated in Section III-B. The directed graph $\hat{\mathcal{M}}^t$ is broadcasted from eNB to UEs. Only those UEs in set $\hat{\mathcal{V}}^t$ are allowed to participate in D2D data offloading on licensed spectrum, either as a D2D data offloading transmitter (if in set $\mathcal{T}^t = \bigcup_{(j,i)\in\hat{\mathcal{A}}^t}\{j\}$) or as a D2D data offloading receiver (if in set $\mathcal{R}^t = \bigcup_{(j,i)\in\hat{\mathcal{A}}^t}\{i\}$).

Let $\mathcal{Y}^t$ denote the set of cellular UEs in time slot t. We now introduce an approach for UEs in set $\mathcal{R}^t \cup \mathcal{Y}^t$ to determine the receiving channel gain from each UE in set $\mathcal{T}^t$ on each channel $b \in \mathcal{B}$. Specifically, the set $\mathcal{B}$ denotes all channels in the LTE network and each channel in $\mathcal{B}$ has the same bandwidth as an RB. We consider flat fading in each channel and consider the fading over different channels to be frequency selective. The idea is to transmit the featured reference signals from UEs in set $\mathcal{T}^t$ to UEs in set $\mathcal{R}^t \cup \mathcal{Y}^t$. Specifically, all reference signals transmitted by the UEs are generated by the same root Zadoff-Chu sequence [14] but with different cyclic shifts. Thus, the correlation of two reference signals transmitted by different UEs in set $\mathcal{T}^t$ is zero. Let c_j^t denote the cyclic shifted sequence generated on UE $j \in \mathcal{T}^t$ in time slot t. The sequence c_j^t for UE j can be generated by a predefined routine based on the received D2D pairing information $\hat{\mathcal{M}}^t$. Thus, by receiving $\hat{\mathcal{M}}^t$ and running the same routine, UEs in set $\mathcal{R}^t \cup \mathcal{Y}^t$ can be certain that sequence c_j^t is transmitted by the UE j. The grid of RBs in the current LTE networks is shown in Fig. 4, where two resource elements (REs) are reserved for the transmissions of our reference signals. We refer to the duration of an RB (*i.e.*, 0.5 ms [9]) as an LTE slot. An LTE slot is much shoter than a time slot t. Let $h_{j,i,b}^{\tau,1}$ denote the channel gain to be determined via RE 1 (Fig. 4) on in channel b from UE $j \in \mathcal{T}^t$ to UE $i \in \mathcal{R}^t \cup \mathcal{Y}^t$ at LTE slot τ. We further denote $g_{i,b}^{\tau,1}$ as the signal received by UE i at τ on RE 1. Then, $g_{i,b}^{\tau,1}$ is the summation of $|\mathcal{T}^t|$ reference signals, *i.e.*, $g_{i,b}^{\tau,1} = \sum_{\ell\in\mathcal{T}^t} h_{\ell,i,b}^{\tau,1} c_\ell^t$. Since the correlation of two Zadoff-Chu sequences with different cyclic shifts of the same root sequence is zero, we have $c_\ell^t c_j^t = 0$ if $\ell \neq j$ and $c_\ell^t c_j^t = C$ if $\ell = j$, where C is a constant determined by the root sequence [14]. Thus, $g_{i,b}^{\tau,1} c_j^t = \sum_{\ell\in\mathcal{T}^t} h_{\ell,i,b}^{\tau,1} c_\ell^t c_j^t = h_{j,i,b}^{\tau,1} C$ and $h_{j,i,b}^{\tau,1} = g_{i,b}^{\tau,1} c_j^t / C$ hold for each UE j. Meanwhile, the channel gain $h_{j,i,b}^{\tau,2}$ from UE j to UE i via RE 2 (Fig. 4) can be determined with similar approach. We first apply interpolation to estimate the channel gains at other REs besides REs 1 and 2. Then, the channel gain from UE j to UE i on the RB of channel b at τ, denoted by $h_{j,i,b}^{\tau}$, is the average value of the channel gains of all REs in channel b. Hence, $h_{j,i,b}^{\tau}$ can be determined at UE i for each $j \in \mathcal{T}^t$.

Let $\boldsymbol{h}_{i,b}^{\tau} \in \mathbb{R}_+^{|\mathcal{T}^t|}$ denote the vector of the current receiving channel gains determined on UE $i \in \mathcal{R}^t \cup \mathcal{Y}^t$ for UEs in set $\mathcal{T}^t$ on the RB of channel b at τ. We refer to $\mathbf{H}_i^{\tau} = [\boldsymbol{h}_{i,1}^{\tau} \cdots \boldsymbol{h}_{i,|\mathcal{B}|}^{\tau}]^T \in \mathbb{R}_+^{|\mathcal{B}|\times|\mathcal{T}^t|}$ as the *D2D channel profile* of UE i at LTE slot τ. The D2D channel profile, which is the output of module 3 on each UE $i \in \mathcal{R}^t \cup \mathcal{Y}^t$, is uploaded via LTE uplink to the eNB. The eNB is aware of the downlink channel gain from eNB to each UE $i \in \mathcal{R}^t \cup \mathcal{Y}^t$ at τ due to the cell-specific reference signals used by the LTE network. Thus, we can formulate an optimization problem in Section III-C to jointly allocate downlink RBs and control the transmission power for each UE in set $\mathcal{T}^t$. The optimization problem is solved by module 4 (Fig. 2) at eNB and the solution is referred to as *offloading control*. The offloading control message contains both the RB allocation and the power control information on licensed spectrum, which is broadcasted to those UEs in set $\hat{\mathcal{V}}^t$. Module 5 (Fig. 2) on each UE in set $\hat{\mathcal{V}}^t$ performs data offloading in parallel by using the allocated RBs with controlled power to improve the spatial reuse.

III. Problem Formulation and Solutions

We have specified the modules 3 and 5 (Fig. 2) of our proposed scheme in Section II. In this section, we formulate and solve problems for modules 1, 2, and 4 mentioned above.

A. Local Task Scheduling Problem

We use the deadline to evaluate the urgency of a data offloading task. For each D2D UE $i \in \mathcal{U}$ in time slot t, the following problem needs to be solved to determine which digital file $k \in \mathcal{Q}_i^t$ should first be downloaded from a neighbouring UE:

$$\underset{k\in\mathcal{Q}_i^t}{\text{argmin}} \quad \delta_{i,k}^t \tag{1a}$$

$$\text{subject to} \quad k \in \mathcal{Q}_i^t \cap \bigcup_{j\in\mathcal{N}_i^t} \mathcal{S}_j^t. \tag{1b}$$

Constraint (1b) implies that a file k selected by UE i to download must be available on one of the neighbouring UEs. We consider each mobile user $i \in \mathcal{U}$ has limited number of delay tolerable files in set $\mathcal{Q}_i^t$. Problem (1) can be solved on UE i by comparing the deadline of all digital files in set $\mathcal{Q}_i^t$, with computational complexity $\mathcal{O}(|\mathcal{Q}_i^t| \sum_{j\in\mathcal{N}_i^t} |\mathcal{S}_j^t|)$.

Algorithm 1: Determine the maximum weight matching $\hat{\mathcal{M}}^t = (\hat{\mathcal{V}}^t, \hat{\mathcal{A}}^t, \hat{w}^t)$ in directed graph $\mathcal{G}^t = (\mathcal{V}^t, \mathcal{A}^t, w^t)$.

1 **init** $\mathcal{E}^t := \emptyset$, $\hat{\mathcal{A}}^t := \emptyset$.
2 **for** $(j,i) \in \mathcal{A}^t$ **do**
3   **if** $\langle i,j \rangle \notin \mathcal{E}^t$ **then**
4     $\mathcal{E}^t := \mathcal{E}^t \cup \langle i,j \rangle$, $\bar{w}^t(\langle i,j \rangle) := w^t((j,i))$.
5   **if** $\langle i,j \rangle \in \mathcal{E}^t$ and $\bar{w}^t(\langle i,j \rangle) < w^t((j,i))$ **then**
6     $\bar{w}^t(\langle i,j \rangle) := w^t((j,i))$.
7 $(\hat{\mathcal{V}}^t, \tilde{\mathcal{E}}^t, \bar{w}^t) := \text{WMA}((\mathcal{V}^t, \mathcal{E}^t, \bar{w}^t))$.
8 **for** $\langle i,j \rangle \in \tilde{\mathcal{E}}^t$ **do**
9   **if** $(i,j) \in \mathcal{A}^t$ and $w^t((i,j)) = \bar{w}^t(\langle i,j \rangle)$ and $(j,i) \notin \hat{\mathcal{A}}^t$ **then**
10     $\hat{\mathcal{A}}^t := \hat{\mathcal{A}}^t \cup (i,j)$, $\hat{w}^t((i,j)) := \bar{w}^t(\langle i,j \rangle)$.
11   **if** $(j,i) \in \mathcal{A}^t$ and $w^t((j,i)) = \bar{w}^t(\langle i,j \rangle)$ and $(i,j) \notin \hat{\mathcal{A}}^t$ **then**
12     $\hat{\mathcal{A}}^t := \hat{\mathcal{A}}^t \cup (j,i)$, $\hat{w}^t((j,i)) := \bar{w}^t(\langle i,j \rangle)$.
13 **output** $\hat{\mathcal{M}}^t := (\hat{\mathcal{V}}^t, \hat{\mathcal{A}}^t, \hat{w}^t)$.

B. Global Task Scheduling Problem

Let $\mathbf{M}^t$ denote a $|\mathcal{V}^t| \times |\mathcal{V}^t|$ binary matrix, where an element $m_{j,i} = 1$ (or $m_{j,i} = 0$) represents that the data offloading demand uploaded by UE i to obtain data from UE j is approved (or not). Thus, the global task scheduling problem at time t can be formulated as

$$\underset{\mathbf{M}^t}{\text{maximize}} \quad \sum_{(j,i)\in\mathcal{A}^t} m_{j,i}^t w^t((j,i)) \tag{2a}$$

$$\text{subject to} \quad \sum_{j=1}^{|\mathcal{V}^t|} m_{j,i}^t \leq 1, \quad \forall\, i \in \mathcal{V}^t, \tag{2b}$$

$$\sum_{i=1}^{|\mathcal{V}^t|} m_{j,i}^t \leq 1, \quad \forall\, j \in \mathcal{V}^t. \tag{2c}$$

Problem (2) can be solved by first casting $\mathcal{G}^t = (\mathcal{V}^t, \mathcal{A}^t, w^t)$ as an undirected graph, where the maximum weight matching can be determined in polynomial time by the weighted matching algorithm (WMA) [16]. Then, we convert the output of WMA to a directed graph which can be proven as the maximum weight matching of the directed graph $\mathcal{G}^t$. We first explain Algorithm 1 and then prove its correctness.

In Algorithm 1, we first cast the directed graph $\mathcal{G}^t$ as an undirected graph $(\mathcal{V}^t, \mathcal{E}^t, \bar{w}^t)$ (Lines 2-6). Specifically, we include a weighted undirected edge $\langle i,j \rangle$ in set $\mathcal{E}^t$ if either directed edge (j,i) or (i,j) is in set $\mathcal{A}^t$ and its weight is either $w^t((j,i))$ or $w^t((i,j))$ correspondingly. If both edges (j,i) and (i,j) are in set $\mathcal{A}^t$, then $\bar{w}^t(\langle i,j \rangle)$ is the larger one of $w^t((j,i))$ and $w^t((i,j))$. The weighted matching algorithm proposed in [16] is used by taking the undirected graph $(\mathcal{V}^t, \mathcal{E}^t, \bar{w}^t)$ as the input. Its maximum weight matching $(\hat{\mathcal{V}}^t, \tilde{\mathcal{E}}^t, \bar{w}^t)$ is determined (Line 7) as the output. For each undirected edge $\langle i,j \rangle \in \tilde{\mathcal{E}}^t$ (Line 8), it is converted to a directed edge (i,j) or (j,i) by comparing $w^t((j,i))$ and $w^t((i,j))$ with $\bar{w}^t(\langle i,j \rangle)$ (Lines 9 and 11). The chosen one is added in the directed edge set $\hat{\mathcal{A}}^t$ with corresponding weight (Line 10 or 12).

Theorem 1: *$\hat{\mathcal{M}}^t = (\hat{\mathcal{V}}^t, \hat{\mathcal{A}}^t, \hat{w}^t)$ is the maximum weight matching of the directed weighted graph $\mathcal{G}^t$.*

Proof: We first prove the following lemma:

Lemma 1: In a directed graph $\mathcal{G}^t = (\mathcal{V}^t, \mathcal{A}^t, w^t)$, if there are two vertices in set $\mathcal{V}^t$ such that both directed edges (i,j) and (j,i) are in set $\mathcal{A}^t$ and if $w^t((i,j)) > w^t((j,i))$, then the directed edge (j,i) is not in its maximum weight matching.

Proof: Assume a matching of $\mathcal{G}^t$ contains directed edge (j,i), then we can compose a new matching by removing directed edge (j,i) and including directed edge (i,j). Due to $w^t((i,j)) > w^t((j,i))$, the new composed matching must have a larger summation of weights, which completes the proof. □

Moreover, since each UE $i \in \mathcal{U}$ selects one digital file to download from one neighbouring UE $j \in \mathcal{N}_i^t$ in time slot t, the indegree of each vertex in the set $\mathcal{V}^t$ is 1. Thus, after removing the directed edge with the smaller weight from all 2-cycles (a cycle [15, pp. 5] with two edges) from $\mathcal{G}^t$, we obtain an oriented graph [15, pp. 27] of the undirected graph $(\mathcal{V}^t, \mathcal{E}^t, \bar{w}^t)$ in Algorithm 1. Since the undirected graph $(\hat{\mathcal{V}}^t, \tilde{\mathcal{E}}^t, \bar{w}^t)$ determined by WMA is the maximum weight matching of the undirected graph $(\mathcal{V}^t, \mathcal{E}^t, \bar{w}^t)$, its oriented graph, $\hat{\mathcal{M}}^t = (\hat{\mathcal{V}}^t, \hat{\mathcal{A}}^t, \hat{w}^t)$ determined in Algorithm 1, is the maximum weight matching of directed graph $\mathcal{G}^t$. ■

The solution of problem (2) is further validated with a random network scenario in Section IV-A.

C. Joint RB Allocation and Power Control Problem

Let $\mathcal{D}^t$ denote the set of cellular UEs in time slot t. Cellular UEs have to be satisfied with higher priority than the D2D UEs. In the LTE networks, a cellular UE in set $\mathcal{D}^t$ can be assigned one or more RBs in set $\mathcal{B}$ at each LTE slot τ, but each RB can be allocated to one cellular UE only. Thus, the eNB is aware of a function $D : \mathcal{B} \mapsto \mathcal{D}^t$. We consider a target SINR on RB $b \in \mathcal{B}$ at τ has to be maintained above a threshold ρ_b^τ for the cellular UE $d \in \mathcal{D}^t$ and $d = D(b)$. If an RB $b \in \mathcal{B}$ is not allocated to any cellular UE, $\rho_b^\tau = 0$.

We assume the maximum transmission power on the eNB is P_{eNB} and denote its power component on RB b at time τ is $p_{eNB,b}^\tau$. For D2D UE pair $(j,i) \in \hat{\mathcal{A}}^t$, the maximum transmission power from the transmitting UE $j \in \mathcal{T}^t$ is denoted by P_j. Let $p_{j,b}^\tau$ denote the power component of the D2D transmitter j on RB b at τ. Then, the joint RB allocation and power control problem to maximize the throughput of D2D data offloading for each LTE slot τ is formulated as

$$\underset{\mathbf{P}_D^\tau,\ \boldsymbol{p}_{eNB}^\tau}{\text{maximize}} \quad \sum_{(j,i)\in\hat{\mathcal{A}}^t} \sum_{b\in\mathcal{B}} W \log_2 \Bigg(1+ \frac{p_{j,b}^\tau h_{j,i,b}^\tau}{WN_0 + p_{eNB,b}^\tau h_{eNB,i,b}^\tau + \sum_{\ell\in\mathcal{T}^t\setminus\{j\}} p_{\ell,b}^\tau h_{\ell,i,b}^\tau}\Bigg) \tag{3a}$$

$$\text{subject to} \quad \frac{p_{eNB,b}^\tau h_{eNB,D(b),b}^\tau}{WN_0 + \sum_{\ell\in\mathcal{T}^t} p_{\ell,b}^\tau h_{\ell,D(b),b}} \geq \rho_b^\tau, \ \forall\, b \in \mathcal{B}, \tag{3b}$$

$$\sum_{b\in\mathcal{B}} p_{j,b}^\tau \leq P_j, \qquad \forall\, j \in \mathcal{T}^t, \tag{3c}$$

$$\sum_{b \in \mathcal{B}} p^{\tau}_{eNB,b} \leq P_{eNB}, \tag{3d}$$

$$p^{\tau}_{eNB,b} \geq 0, \qquad \forall\, b \in \mathcal{B}, \tag{3e}$$

$$p^{\tau}_{j,b} \geq 0, \qquad \forall\, b \in \mathcal{B}, \forall\, j \in \mathcal{T}^t, \tag{3f}$$

where $\boldsymbol{p}^{\tau}_{eNB} = (p^{\tau}_{eNB,1}, \ldots, p^{\tau}_{eNB,|\mathcal{B}|})$, $\mathbf{P}^{\tau}_{D}$ is a $|\mathcal{T}^t| \times |\mathcal{B}|$ matrix with $p^{\tau}_{j,b}$ for each $j \in \mathcal{T}^t$, $b \in \mathcal{B}$, W is the bandwidth of each RB $b \in \mathcal{B}$ at τ, and N_0 is the thermal noise spectral density. Problem (3) is a nonlinear optimization problem, which can be solved by sequential quadratic programming (SQP) method to determine a local optimal solution. An intuitive explanation of the solution is presented in Section IV-A.

IV. Performance Evaluation

We first validate our proposed scheme and then present its performance. The network setting is as follows. The UEs are randomly deployed in a square coverage region of an eNB, where the eNB is located at the centre. The D2D communication range is 30 m. Each UE has a maximum transmission power 23 dBm [17]. The eNB has a maximum transmission power 40 Watts. A user selects a target position to move with random velocity from 1 to 2 m/s. A user stays at a target position from 2 to 20 min randomly and then moves to the next position. Since the velocity is relatively small, we consider 10 sec for each time slot. Each UE has 5 to 35 delay tolerable files under request. These files have random size from 15 to 45 MB and random deadlines from 2 to 90 min. Each UE caches a number of most recently downloaded files. We consider both path loss and channel fading in our simulations. For communication channels from eNB to cellular UEs and all interference channels, Rayleigh fading is considered. The Rician fading is applied for D2D communication channels.

A. System Validation

We consider 30 D2D UEs (the circles) and 25 cellular UEs (the squares) as shown in Fig. 5. Each cellular UE is assigned an RB at an LTE slot. The index of each cellular UE is the same as the ID of RB assigned to it. Thus, up to 25 RB can be reused by each pair of D2D UEs with power control. Fig. 5 is automatically generated. The head end of each solid arrow represents an offloading request submitted by the UE. We also show the utility (given by function w^t in Section II) of the UE that submits the request. The tail end of an arrow represents the UE that is responsible for sending data. We first show the correctness of Algorithm 1, which is used to approve offloading requests. Matching $\hat{\mathcal{M}}^t$ is represented by colouring the approved requests and their utilities in blue. For example, D2D UEs 3 and 10 submit data offloading requests for files from each other with utilities 0.337 and 0.0228, respectively. Meanwhile, D2D UE 9 submits a request with utility 0.388 for the data from D2D UE 10, which is approved. Algorithm 1 is correct since it approves the request that maximizes the utility. We then show the effectiveness of the channel estimation approach (Section II) in Fig. 5. The dotted red arrows, which are from each approved D2D transmitter to three cellular UEs, represent the strongest three interference links estimated in the corresponding RBs. The estimation is valid since the strongest three interference links to cellular UEs are those three with relatively close distance from the D2D transmitter. The solution of problem (3) is shown in Fig. 6. The eNB allocates small transmission power on RBs 9, 20, and 21 for D2D UE 15 due to its high interference channel gains to cellular UEs 9, 20, and 21 (Fig. 5). However, much higher transmission power is allocated on RBs 7, 10, 13, and 17 for D2D UE 15 since its interference channel gains to corresponding cellular UEs are small. Similar results can be found from Figs. 5 and 6 for other D2D transmitters.

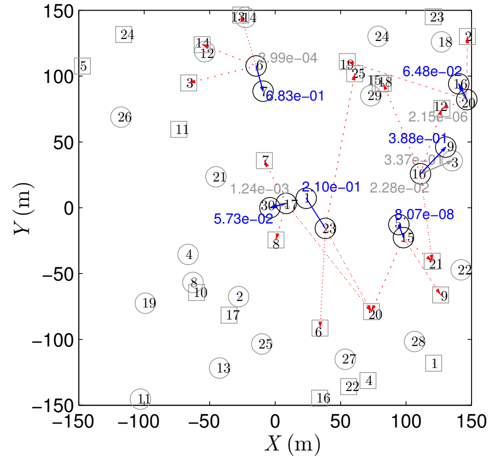

Fig. 5. An automatically generated figure for a random network scenario in an LTE slot, which is used to validate the proposed scheme.

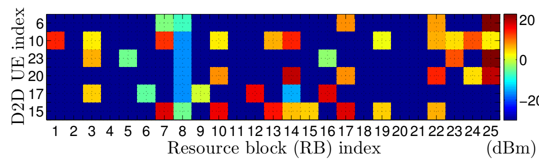

Fig. 6. The power allocation result by solving problem (3).

B. Performance Comparison

We consider 100 RBs for our simulations with $W = 180$ kHz. We consider 802.11g as the base line to show the data offloading ability of WiFi-Direct with 20 MHz unlicensed frequency band, where the carrier sense multiple access with collision avoidance technique is applied. All the simulation results are obtained by 4 hr simulation time.

We deploy 100 cellular UEs and increase the number of D2D UEs from 20 to 100 in a square coverage region of an eNB with length of 200 m. We compare the performance of the proposed scheme with 50 or 100 cached files on each D2D UE. When the required SINR for each cellular UE in each RB is 10 dB, the simulation results are given in Fig. 7. For our proposed scheme and 802.11g, the amount of offloaded data increases with the number of D2D UEs since there are more available files nearby and more D2D communication pairs in the network. Our proposed scheme outperforms the data offloading in 802.11g because multiple D2D pairs nearby cannot communicate simultaneously in 802.11g since the transmission between one pair of D2D UEs suspends others by the collision avoidance technique. Thus, some transient

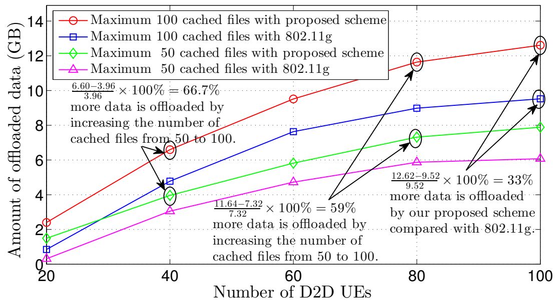

Fig. 7. Results with varying number of D2D UEs and with different maximum number of cached files on each D2D UE.

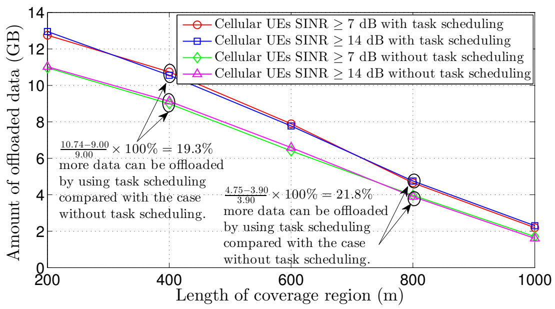

Fig. 8. Results with varying size of the simulation region and with different SINR requirements of cellular UEs.

data offloading opportunities are wasted. However, by jointly allocating the RBs and performing power control, multiple D2D pairs nearby can communicate simultaneously in our proposed scheme. When we have 100 D2D UEs, 33% more data can be offloaded with the proposed scheme compared with WiFi-Direct operated on unlicensed spectrum. We also find that the amount of offloaded data is proportional with the maximum number of cached files on each D2D UE. Almost 60% more data is offloaded when the maximum number of cached files increases from 50 to 100.

We study the effectiveness of task scheduling, *i.e.*, the effectiveness of solving problems (1) and (2) in the proposed scheme. We deploy 100 D2D UEs and 100 cellular UEs in a square coverage region of an eNB with side length varying from 200 m to 1 km. For the case without solving problems (1) and (2), we let each D2D UE randomly select a file to upload its data offloading demand and the eNB randomly approving received offloading demands without concerning their deadlines. We also compare the performance of proposed scheme when different SINR are required by cellular UEs. We assume the required SINR of each cellular UE in each RB is 7 or 14 dB in each set of simulations. When each D2D UE caches 100 files, simulation results are given in Fig. 8. The amount of offloaded data decreases when the length of the coverage region increases from 200 m to 1 km. This follows the explanations for Fig. 7 since the density of D2D UEs is decreased. By solving problems (1) and (2), almost 20% more data traffic can be offloaded compared with the case without task scheduling. We also find different SINR required by cellular UEs marginally changes the amount of offloaded data. This is because joint RB allocation and power control provide high flexibility in data offloading control, which can adapt different SINR requirements of cellular UEs.

V. Conclusion

In this paper, we proposed a scheme to reuse the downlink RBs of LTE networks for D2D data offloading. We presented our system modules, where the local task scheduling problem, global task scheduling problem, channel estimation, and joint RB allocation and power control problem are included. We validated the correctness of global task scheduling and presented the solution of the joint RB allocation and power control problem with a random network scenario. We also conducted simulations to evaluate the performance of our proposed scheme. Results show that, without violating the downlink SINR requirements of cellular UEs, our proposed scheme can offload more data than using WiFi-Direct in an unlicensed spectrum.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2013-2018." [Online]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html
[2] A. Beach, M. Gartrell, S. Akkala, J. Elston, J. Kelley, K. Nishimoto, B. Ray, S. Razgulin, K. Sundaresan, B. Surendar, M. Terada, and R. Han, "Whozthat? Evolving an ecosystem for context-aware mobile social networks," *IEEE Network*, vol. 22, no. 4, pp. 50–55, Jul. 2008.
[3] E. Baştuğ, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 82–89, Aug. 2014.
[4] G. Zyba, G. Voelker, S. Ioannidis, and C. Diot, "Dissemination in opportunistic mobile ad-hoc networks: The power of the crowd," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr. 2011.
[5] A.-K. Pietiläinen and C. Diot, "Dissemination in opportunistic social networks: The role of temporal communities," in *Proc. of ACM MobiHoc*, Hilton Head, SC, Jun. 2012.
[6] WiFi Alliance, "Wi-Fi peer-to-peer (P2P) technical specification," 2010. [Online]. Available: www.wi-fi.org/WiFi_Direct.php
[7] Y. Zhao, B. Pelletier, P. Marinier, and D. Pani, "D2D neighbor discovery interference management for LTE systems," in *Proc. of Int'l Workshop on D2D Communication With and Without Infrastructure*, Atlanta, GA, Dec. 2013.
[8] M. J. Yang, S. Y. Lim, H. J. Park, and N. H. Park, "Solving the data overload: Device-to-device bearer control architecture for cellular data offloading," *IEEE Veh. Technol. Mag.*, vol. 8, no. 1, pp. 31–39, Mar. 2013.
[9] 3GPP TS 36.211 V12.2.0, "Evolved universal terrestrial radio access (E-UTRA) physical channels and modulation," Jun. 2014.
[10] M. Hasan, E. Hossain, and D. I. Kim, "Resource allocation under channel uncertainties for relay-aided device-to-device communication underlaying LTE-A cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2322–2338, Apr. 2014.
[11] 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 Commun. Mag.*, vol. 50, no. 3, pp. 170–177, Mar. 2011.
[12] P. Phunchongharn, E. Hossain, and D. I. Kim, "Resource allocation for device-to-device communications underlaying LTE-Advanced networks," *IEEE Wireless Commun. Mag.*, vol. 20, no. 4, pp. 91–100, Aug. 2013.
[13] 3GPP TS 36.213 V12.2.0, "Evolved universal terrestrial radio access (E-UTRA) physical layer procedures," Jun. 2014.
[14] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. 18, no. 4, pp. 531–532, Jul. 1972.
[15] D. Jungnickel, *Graphs, Networks and Algorithms, Fourth Edition*. Springer, 2012.
[16] H. N. Gabow, "Data structures for weighted matching and nearest common ancestors with linking," in *Proc. of ACM-SIAM Symposium on Discrete Algorithms*, San Francisco, CA, Jan. 1990.
[17] 3GPP TS 36.101 V12.4.0, "Evolved universal terrestrial radio access (E-UTRA) user equipment (UE) radio transmission and reception," Jul. 2014.