R2D2: Embracing Device-to-Device Communication in Next Generation Cellular Networks

Tarun Bansal[‡*], Karthikeyan Sundaresan[†], Sampath Rangarajan[†] and Prasun Sinha[‡]
[†]Mobile Communications & Networking Research, NEC Labs America, Princeton, USA
[‡]Department of Computer Science, Ohio State University, Columbus, USA

***Abstract*—Device-to-device (D2D) communications is being pursued as an important feature in next generation cellular networks. D2D can improve resource utilization in two ways: *Offloading* cellular traffic to D2D, and *Reuse* of resources used by conventional cellular transmissions for D2D communication. In this paper, we show that in multi-cell environments that employ FFR (Fractional Frequency Reuse), the benefits from D2D toward reuse are limited. We then propose *R2D2*- a holistic approach to efficient offloading with D2D traffic. *R2D2* leverages the flexible nature of D2D traffic (in using downlink/uplink resources) to cater effectively to the spatial and temporal asymmetry in traffic load both across and within cells. *R2D2* incorporates a two time-scale solution: a coarse time-scale dynamic FFR scheme that leverages D2D traffic to determine the FFR patterns for downlink and uplink jointly among interfering sectors; and a fine time-scale scheduling solution that intelligently schedules cellular and D2D traffic jointly across DL (Downlink) and UL (Uplink) resources. We establish the hardness of the scheduling problem and present efficient and low complexity algorithms with approximation guarantees. Through extensive evaluations, we confirm that *R2D2* delivers the offloading benefits of D2D, with its proposed algorithms performing very close to the optimal.**

I. Introduction

Device-to-device (D2D) communications is being pursued as an important feature [1] for the next generation cellular networks (LTE-advanced). Being an underlay to cellular networks, the goal is to leverage the physical proximity of communicating devices to improve cellular coverage in sparse deployments, provide connectivity for public safety services and improve resource utilization in conventional deployments [2], [3]. We focus on D2D's ability to improve resource utilization in this work.

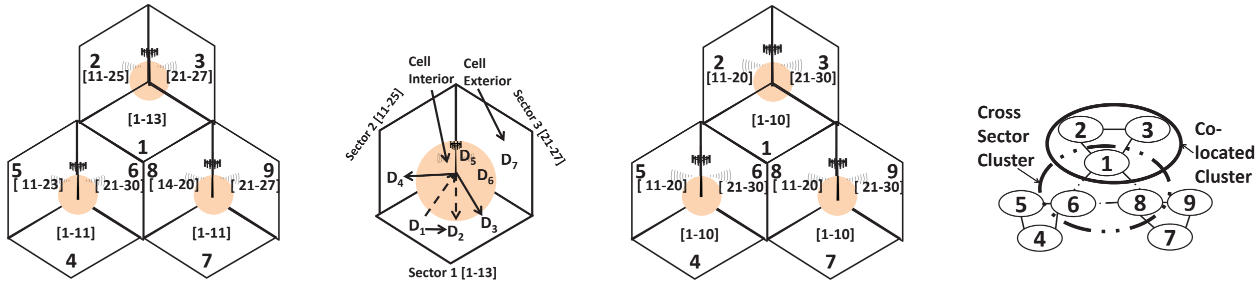
D2D can improve resource utilization in two ways: (i) *offload:* a data session between two devices (D1, D2) in the same sector which conventionally incurs two hop transmissions in the cellular mode (D1→BS, BS→D2) now requires only a single hop transmission (Fig. 1b) in D2D mode (D1→D2), and (ii) *reuse:* leveraging the physical proximity, the D2D communication can further operate on resources on which conventional cellular users are already scheduled. While existing works [4], [2], [5] have highlighted the benefits of both these components, the study was not conducted under practical multi-cell deployments with an inherent pattern of resource reuse (called fractional frequency reuse, FFR) for the cells. Indeed, we show that in multi-cell deployments with even a static FFR pattern, the existing reuse provided by the cellular deployment leaves little room for D2D communication to provide additional reuse. Hence, most of D2D's gain is restricted to its ability to offload cellular traffic.

Given the large spatial and temporal variations in traffic load in practice [6], it is important to consider offloading with D2D in the presence of dynamic FFR schemes. Here, D2D brings both an opportunity as well as a challenge. (i) While the uplink (UL) and downlink (DL) radio resources are fixed across different base stations, their traffic load can vary significantly within a cell. With D2D traffic capable of being scheduled in both UL and DL resources, it presents an opportunity in that it serves as a *flexible* load that can be intelligently placed (in DL/UL resources) to efficiently utilize the cell's net radio resources (see example in Fig. 2). This can help both during dynamic FFR pattern determination as well as during scheduling, albeit at the expense of coupling the DL and UL resource allocation problems that have conventionally been addressed separately. (ii) In a sectored deployment with dynamic FFR, while cross-sectors (*e.g.*, 1, 6, 8 in Fig. 1a) need to split resources to alleviate interference; co-located sectors (*e.g.*, 1, 2, 3 in Fig. 1a) can potentially schedule their cellular users (*e.g.*, D_3 and D_4 in Fig. 1b) on overlapping resources (say RB 11) without interference (due to *directional* BS transmission/reception). However, with the introduction of D2D traffic that is *omni-directional*, this creates interference conflicts between co-located sectors on shared resources (*e.g.* between $D_1 \leftrightarrow D_2$ and D_4 on RB 11) that diminishes the resource utilization and reuse capability of dynamic FFR (Fig. 1a), potentially offsetting the offloading benefits from D2D. Hence, the challenge is to alleviate such interference conflicts either as part of the FFR solution and/or through joint scheduling of D2D and cellular traffic *across* the co-located sectors. Thus, we find that intelligent D2D traffic placement (during FFR) coupled with scheduling of D2D and cellular traffic jointly on DL and UL resources as well as across co-located sectors is critical to achieve effective traffic offloading and resource utilization.

Toward addressing these challenges, we propose *R2D2*- a framework for holistic Radio Resource Management (RRM) with D2D communication in cellular networks. *R2D2* operates at two time scales. At the *beginning of every epoch* (lasting several tens of frames), *R2D2* estimates the average traffic (resource) demand from cellular and D2D traffic in each sector in either direction (DL and UL) based on history (from previous epochs). It partitions the network into disjoint (small) clusters of interfering sectors (called cross sectors) and leverages the flexible nature of D2D traffic to *jointly*

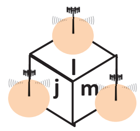
*Tarun Bansal interned at the Department of Mobile Communications and Networking, NEC Laboratories America Inc., Princeton, NJ during this work.

978-1-4799-3360-0/14/$31.00



(a) Dynamic FFR (b) D2D usage (c) Static FFR (d)

Fig. 1. (a)-(c): [] indicates the set of Resource Blocks (RBs, *i.e.* time-frequency allocation units [1]) available for allocation to cell exterior traffic. Total number of RBs to be allocated is 30. (d) Graph for the network shown in Fig. 1a.



(a) Cross-Sector Clique.

Cell\Demand	DL	UL	D2D
Cell *j*	1	2	4
Cell *l*	3	3	9
Cell *m*	5	1	8

(b) Traffic demand at each sector (or cell).

	D2D Oblivious FFR Allocation					D2D Aware FFR Allocation (R2D2)				
Cell	Total Demand		Resources Allocated		Traffic Supported	Total Demand		Resources Allocated		Traffic Supported
	DL	UL	DL	UL		DL	UL	DL	UL	
j	1	2 + 4 = 6	3	3	1 + 3 = 4	1 + 1 = 2	2 + 3 = 5	2	5	2 + 5 = 7
l	3	3 + 9 = 12	8	5	3 + 5 = 8	3 + 6 = 9	3 + 3 = 6	9	6	9 + 6 = 15
m	5	1 + 8 = 9	13	4	5 + 4 = 9	5 + 8 = 13	1 + 0 = 1	13	1	13 + 1 = 14
Total			24	12	**21**			24	12	**36**

(c) The number of resources that can be allocated to UL and DL are 12 and 24, respectively. R2D2 will carefully split D2D traffic across UL and DL, serving **71% higher traffic.**

Fig. 2. A D2D Oblivious dynamic FFR algorithm will put all D2D traffic on UL resources (to avoid interference with the downlink cellular transmissions in the co-located sectors) and allocate resources proportionally. On the other hand, a dynamic FFR allocation scheme aware of D2D traffic (R2D2) will carefully split D2D traffic across UL and DL. When coupled with resource allocation for interior traffic, the D2D-Aware dynamic FFR allocation becomes more challenging (shown in Sec. IV-C) and provides even higher benefits.

determine the dynamic FFR patterns for DL and UL for each of these clusters in a completely distributed and localized manner. Based on the dynamic FFR pattern determined, the resources of operation are determined for each sector in the cluster in DL and UL directions. Then, *in every frame*, for the set of sectors co-located at the same base station and instantaneous traffic demands, *R2D2* solves the coupled problem of D2D traffic placement and scheduling of cellular and D2D traffic jointly on both the DL and UL resources as well as across the sectors. Thus, while the coarse time-scale component in *R2D2* allocates resources and removes interference only between cross sectors (through dynamic FFR) for an entire epoch, the fine time-scale component is responsible for alleviating the interference between co-located sectors generated by D2D traffic (through joint sector scheduling) and maximizing the utilization of allocated resources in every frame in the epoch.

While the first step in *R2D2* is solved efficiently through a light-weight and scalable dynamic FFR scheme, solving the scheduling problem in the second step is an NP-hard problem. Toward solving this per-frame scheduling problem, *R2D2* designs efficient algorithms with performance guarantees that are also amenable to implementation at the time scale of a frame (1 ms in LTE). Our evaluations with realistic LTE settings reveal that intelligent scheduling of D2D and cellular traffic when combined with appropriate D2D traffic placement can provide gains of 2.79x in resource utilization (throughput) compared to existing schemes [4]. Our contributions in this work are multi-fold:

1. Contrary to existing belief, we show that while D2D communications offer offloading benefits, their ability to reuse cellular resources is limited in multi-cell environments, where FFR patterns are already deployed in practice.
2. We propose *R2D2* - A framework for holistic approach to efficient offloading with D2D traffic. *R2D2* incorporates a two time-scale solution: localized computation of dynamic FFR patterns jointly for DL and UL (leveraging flexible nature of D2D traffic) in each cluster of cross sectors at epoch granularity; and intelligent scheduling of cellular and D2D traffic *jointly* on both the allocated DL and UL resources as well as across co-located sectors in every frame.
3. Toward solving the scheduling problem in each frame, we provide a $\frac{1}{2}$ approximation algorithm with a complexity of $O(N^2K^3)$, where N and K are the number of OFDMA resource blocks and users in each sector respectively. We also provide an alternate $\frac{1}{4}$ approximation algorithm that has a lower complexity of $O(N^2K)$. Through extensive evaluations, we show that in practice, the performance of both algorithms are significantly better than their worst-case guarantees and are within 5% of the optimal. While these algorithms apply to TDD systems, for FDD systems, where a D2D traffic session is constrained to not be allocated resources from both DL and UL simultaneously, we provide an $\frac{1}{3}$ approximation algorithm.

II. Background

We consider an OFDMA based next generation cellular network (LTE is used as a reference). The network could be time (TDD) or frequency (FDD) divisioned, where downlink (DL) and uplink (UL) resources for all the cells in the network are determined a priori (same across all base stations). Coverage area of all Base Stations (BSs) is typically divided into three sectors[1] with the help of 120^o directional (sectored) antennas and each sector is considered to be a separate cell in itself for operational purposes (See Fig. 1a). Each frame in LTE is 1 ms long and consists of time-frequency allocation

[1]Discussions and solutions are also applicable to other sectorization models.

units called resource blocks (RBs) on which users' data are scheduled.

LTE advanced users will potentially [1] support D2D communication with their peers that will avoid data having to go through the BS and the mobile core network when it is destined for users in the same sector. We focus on network-assisted D2D communication within the same sector as it is more realistic without requiring neighboring base stations to interact/coordinate with each other. Here, although, data is transferred between the peers directly in D2D communication, the control signaling still goes through the BS and scheduling of D2D traffic is also managed by the BS.

Inter-sector interference in cellular networks is handled with the help of FFR patterns. In the popular 1-3 FFR scheme, the DL (UL) spectrum is divided into four fixed-size frequency bands. One band is used by all the cell-interior clients (in each BS), who do not see interference due to the close proximity to their BS, while the other three bands are split across the three sectors (Fig. 1c) of a BS to mitigate interference with sectors of adjacent base stations.

A. Related Work

D2D in single cell: D2D communication has been considered before [4], [2], [7], [8], [9], [10], [11] in the context of single base stations and with no sectorization. However, cellular deployments are multi-cell and sectored (120^o) in nature and employ FFR patterns. Further, [4], [7], [9], [10], [11] restrict their focus to reusing the D2D transmissions with only the uplink transmissions resulting in under-utilization of resources.

D2D with complete information about path loss: Doppler *et al.* [5] consider the D2D mode-selection problem in multi-cell scenarios. However, to decide which RB should be allocated to a given D2D transmission, their algorithm requires knowledge of path loss between all pairs of D2D transmitters/receivers and non-D2D transmitters/receivers. This constitutes a significant signaling overhead.

Dynamic FFR: Algorithms have been proposed for dynamic FFR algorithm. However, existing algorithms are centralized [12], [13] and/or D2D-oblivious[14]. The centralized approach to computing the FFR allocation and configuring all the base stations (BSs) in every frame is not feasible for cellular networks. Further, the D2D oblivious FFR algorithms are expected to have low throughput when there is an asymmetry in UL and DL traffic load.

III. Benefits and Challenges

As discussed in Section I, D2D can provide gains by offloading and by reusing resources already allocated to cellular users in the same sector. This section highlights the true potential for D2D in practical multi-cell networks and the associated challenges in realizing their benefits. We use simulations from an LTE system with a D2D underlay (details in Section V) to emphasize our inferences.

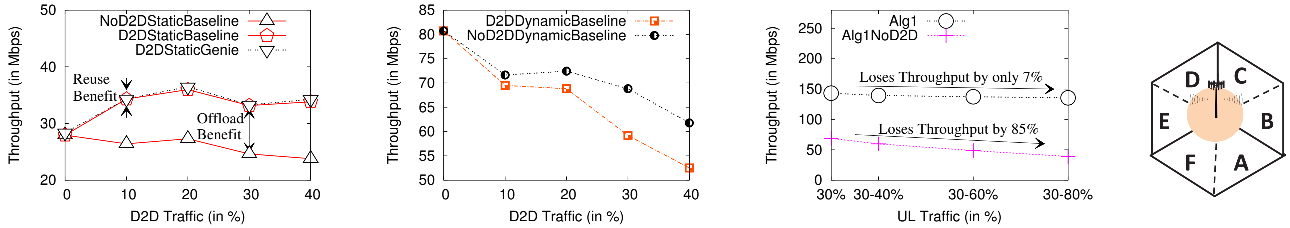
A. Potential for Reuse from D2D

Existing works [4], [2] have attempted to leverage reuse by D2D links in the absence of multiple cells. However, in multicell deployments with FFR schemes, as discussed in Section II, spectral resources are already reused in cells in adjacent BSs. Hence, for a D2D link to reuse resources without impacting existing cellular transmissions in the same sector, the separation between the D2D devices should be much smaller compared to that between the D2D transmitter and cellular user operating on the same resource (on say DL) in either the same or an adjacent sector. Such opportunities are however, not common, especially given the omni directional nature of D2D communications (see Fig. 1b) and small sector sizes. Even if such opportunities arise, to be able to detect and leverage such opportunities, the channel gains between all cellular/D2D and D2D users need to be measured and reported to the BS, which must then use this information to perform per-frame scheduling. Clearly, accomplishing this entire process within a single frame is not feasible due to the prohibitive overhead.

Finally, to understand the potential for reuse from D2D in a hypothetical scenario, we compare three schedulers in the presence of 1-3 static FFR (Fig. 1c): (i) Requires all traffic to go through the BS by classifying all traffic (even D2D) as cellular; (ii) Does not allow D2D traffic to reuse RBs allocated to cellular users in the same sector; and, (iii) A *genie* scheduler that has the channel gain information between all users as well as between users and BS along with power control (details of schedulers, LTE simulation settings, D2D traffic classification etc. are covered in Section V) and uses that to schedule D2D transmissions. In the first two schedulers, each sector independently assigns resources to different transmitters by solving a single cell resource allocation problem [15]. The result in Fig. 3a indicates that *only a marginal gain comes from the reuse of resources by D2D traffic even with a hypothetical genie scheduler, while a large portion of the gain from D2D comes from its offloading capability.*

B. Challenge and Opportunity in D2D Offloading

1) Traffic Variations across Sectors: It is important to align cellular resources to cater effectively to traffic load variations across cells that are common in practice [6]. This is realized with the help of dynamic FFR schemes that allocate spectral resources to interfering cells, taking into account their traffic load. Interestingly, a simple distributed solution can be applied to realize dynamic FFR in a sectored deployment. Each sector belongs to a cluster of cross sectors (located at different cell-sites) and a cluster of co-located sectors (Fig. 1d). Dynamic FFR can be applied independently to each disjoint cluster of three (cross) sectors, whose cell-exterior traffic interfere with each other. Since each sector belongs to exactly two clusters (Fig. 1d), once the operational resources are determined for each sector by dynamic FFR, it is possible that its allocated resources may overlap with those of neighboring co-located sectors in the other cluster. Note that, this is not a problem if all traffic is cellular since the co-located sectors employ directional (sector) antennas to begin with (*e.g.*, sectors 1 and 2 in Fig. 1b can simultaneously transmit to D_3 and D_4, respectively). This allows for applying the above light-weight dynamic FFR scheme for cellular traffic that operates locally on individual clusters of cross sectors alone. However, with the introduction of D2D traffic that is omni-directional, such an approach would not work and would create interference between co-located sectors on overlapping resources (*e.g.* between $D_1 \leftrightarrow D_2$ and D_4 on RB 11 in Fig. 1b) that could



(a) Using D2D with static FFR. (b) Using D2D with dynamic FFR. (c) Asymmetric UL-DL traffic. (d)

Fig. 3. (a)-(c) Evaluation results illustrating challenges in leveraging D2D. (d) Constraints when scheduling transmissions on the same RB.

bring down the benefits from dynamic FFR as well as D2D offloading.

We compare the performance of two systems - one that classifies all traffic as cellular, and another that classifies traffic as cellular and D2D. Both systems first apply dynamic FFR to the clusters with cross sectors and then perform scheduling in each of the sectors independently and then assign resources to different transmitters by solving a single cell resource allocation problem [15]. The result in Fig. 3b clearly indicates the magnitude of the interference created by D2D that did not exist in a pure cellular environment, resulting in the D2D system performing worse than one that classifies all traffic as cellular.

The interference generated by D2D traffic can be addressed as part of the dynamic FFR process, where co-located sectors are incorporated in the FFR process. However, this would couple all the cross and co-located sectors across the network, preventing the dynamic FFR scheme from no longer being local and light-weight. An alternate approach is to *retain the light-weight nature of dynamic FFR (only applied for cross sectors) but to alleviate the interference generated by D2D through intelligent scheduling of cellular and D2D traffic in each cluster of co-located sectors jointly*. We will adopt the latter approach in this work.

2) Traffic Variations within a Sector: In addition to traffic variations across sectors, there is also a lot of asymmetry [6] in traffic load between uplink (UL) and downlink (DL) that changes both spatially and temporally. However, the spectral allocations to the UL and DL cannot be varied dynamically across sectors as this would lead to asymmetric interference between UL and DL traffic across cells, which is a much harder problem to address. For this reason, in practical systems, a resource block can either be used for UL traffic or DL (but not both), and this classification is same across all cells. This makes it difficult to leverage the traffic load variations between UL and DL within each cell. In this regard, we note that placing the D2D traffic on DL or UL resources does not have any relative advantages - placing it on the DL resources may create interference to cellular users from D2D transmitter, while placing it on the UL resources may create interference from the cellular user to D2D receiver. Hence, D2D can be used as a *flexible* load to balance and match the UL and DL traffic load to their available resources, resulting in better resource utilization.

We compare the benefits of a system that employs D2D as a flexible load against one that classifies all traffic as cellular. Both the systems use dynamic FFR but the former computes the FFR resource allocation using D2D as a flexible load (See Sec.V for details). Fig. 3c clearly indicates that the performance of the former system is resilient to variation in traffic disparity while the latter loses 85% throughput when traffic disparity is high. However, to realize these benefits, one would need to solve the D2D traffic placement problem, both as part of dynamic FFR as well as during scheduling. A unique aspect of this problem is that compared to conventional cellular systems that solve the DL and UL resource allocation problems separately, *addressing the D2D traffic placement problem would entail solving the DL and UL resource allocation problems jointly*.

In summary, most of the benefits from D2D arise from its offloading capability and not from reuse. To maximize the offloading gains from D2D, it is important to solve the problem of intelligent D2D traffic placement (during FFR) coupled with schedule of D2D and cellular traffic jointly on DL and UL resources as well as across co-located sectors.

IV. *R2D2*: RRM WITH D2D

A. Overview of *R2D2*

To retain the localized nature of dynamic FFR schemes for scalability without sacrificing performance, we propose *R2D2*- a holistic approach to Radio Resource Management with D2D communication in cellular networks. *R2D2* operates at two time scales.

1. At the beginning of each epoch (lasting several tens of frames), *R2D2* leverages the flexible nature of D2D traffic to determine the dynamic FFR patterns for DL and UL jointly for each of cross-sector clusters (Fig. 1d) in a completely localized manner. Based on the dynamic FFR patterns determined, the radio resources of operation are determined for each sector in the cluster in DL and UL directions for the entire epoch.
2. The burden of alleviating interference generated by D2D traffic between co-located sectors is moved to the frame level time granularity. Here, in every frame, for every cluster of co-located sectors, *R2D2* solves the coupled problem of D2D traffic placement and intelligent scheduling of cellular and D2D traffic on both the DL and UL resources jointly across the sectors.

Note that, *R2D2* carefully assigns the coarse time-scale dynamic FFR process to the cross sectors and the per-frame joint scheduling to the co-located sectors, as coordinating the latter practically comes for free (due to co-location at the BS). *This step ensures that R2D2 is able to effectively leverage spatial reuse of the resources while requiring finer cooperation among only the sectors that are co-located at the same base station.* We now explain each of these components of *R2D2* in detail.

B. D2D Traffic Classification

Flows that originate and end in the same sector can possibly be offloaded to D2D. Let $\hat{r}_{d2d}, \hat{r}_{uplink}$ and $\hat{r}_{downlink}$ be the average physical layer data rates achievable from a given transmitter to a given receiver using D2D, from transmitter to the BS and from the BS to the receiver, respectively. These rates are estimated based on channel quality information from reference signal received power (RSRP) measurements that span over all RBs [1], [16]. At the beginning of every epoch, a BS in *R2D2* offloads a flow to D2D if both of the following conditions are satisfied: (i) Flow originates and ends in the same sector; and, (ii) $\frac{1}{\hat{r}_{d2d}} < \frac{1}{\hat{r}_{uplink}} + \frac{1}{\hat{r}_{downlink}}$. The second condition ensures that the time taken by the flow over D2D is less than the time taken by the flow when routed through the BS. These data rates can be estimated by the BS from the corresponding SINR values. Note that since control messages and signaling overhead are involved in setting up a D2D session, this process cannot be invoked at the same granularity of per-frame scheduling. This justifies the rationale behind employing average data rates for D2D traffic classification at an epoch granularity.

C. Dynamic FFR in R2D2

At the epoch level time granularity (several tens of frames), *R2D2* is executed locally by all the three sectors that form a cross-sector cluster (See Fig.1d).

Step 1 (Traffic Classification): In a dynamic FFR solution, the size of the four bands in the FFR pattern are adapted and chosen to meet the requirements from both cell-interior and exterior traffic. Further, the FFR patterns can be re-configured at coarse time scales (several tens of frames) to track the traffic load variations. In the case of a cellular user, simple SINR thresholds are used to classify the user as a cell-interior or exterior user (*e.g.*, D_5 in Fig. 1b is an interior user due to its high SINR with the BS). For a D2D link, its classification is done taking into account both the devices in the link. We consider a D2D link as interior traffic only if both ends of the link are cell-interior users; otherwise the pair is classified as exterior traffic (*e.g.*, $D_5 \leftrightarrow D_6$ in Fig. 1b is interior traffic while $D_5 \leftrightarrow D_7$ is exterior.) .

Step 2 (Resource Demand Estimation): For each cross-sector cluster, *R2D2* estimates the average traffic (resource) demand from cellular and D2D traffic in each sector for the current epoch based on information from previous epochs. It keeps track of the aggregate resource allocation (R, in RBs) to the interior and exterior traffic of both cellular and D2D users in each epoch (for every sector j) and computes the estimate for average resource demand ($\bar{R}$) for the current epoch in sector j as a weighted (α) moving average: $\bar{R}_{j,xyz}(t) = \alpha R_{j,xyz}(t-1) + (1-\alpha)\bar{R}_{j,xyz}(t-1)$. Here $x = \{C, D\}$ indicates cellular (C) or D2D (D) traffic; $y = \{i, e\}$ indicates interior (i) or exterior (e) traffic; and $z = \{d, u\}$ indicates DL (d) or uplink (u) traffic in sector j. The above equation estimates the average resource demand for a given traffic type in a particular direction.

Step 3 (Determining resources allocated in the cross-sector cluster): Let N_d and N_u be the total available resources in DL and UL spectrum, respectively. This spectrum needs to be shared across the three cross sectors. Further, we also need to determine how much spectral resources would be allocated to interior traffic. Let $F_{0,d}$ and $F_{0,u}$ be the number of RBs allocated to interior DL and UL traffic, respectively that is common to all the cross sectors. $F_{j,d}$ and $F_{j,u}$ denote the number of RBs allocated to exterior DL and UL traffic, respectively for sectors $j = 1, 2, 3$. The allocation of resources (*i.e.* dynamic FFR pattern) can be formalized as a linear optimization problem shown in (1). This formulation exhibits the following properties: (i) Leverages the flexible nature of D2D traffic to carefully split it across UL and DL resources; (ii) Maximizes the total traffic demand satisfied; (iii) Allows interior traffic to be scheduled on exterior traffic resources but not vice-versa as the latter would receive interference from interior traffic in the neighboring cross sectors; and, (iv) Ensures proportional fairness across the three sectors. Using $A_{j,xyz}$ to indicate the resources (number of RBs) allocated to traffic type xy (cellular/D2D, interior/exterior) in direction z (DL/UL) in sector j, we have:

$$\max_{A_{j,xyz}, F_{0,z}, F_{j,z} \forall j,x,y,z} \sum_{x\in\{C,D\}} \sum_{y\in\{i,e\}} \sum_{z\in\{d,u\}} \sum_{j\in\{1,2,3\}} A_{j,xyz} \quad (1)$$

$$\text{subject to } F_{0,z} + \sum_{j\in\{1,2,3\}} F_{j,z} \le N_z \ \forall z \in \{u,d\} \quad (2)$$

$$A_{j,Ciz} \le \bar{R}_{j,Ciz} \ \forall j \in \{1,2,3\} \ \forall z \in \{u,d\} \quad (3)$$

$$A_{j,Ciz} + A_{j,Cez} \le \bar{R}_{j,Ciz} + \bar{R}_{j,Cez} \ \forall j \in \{1,2,3\} \ \forall z \in \{u,d\} \quad (4)$$

$$A_{j,Diu} + A_{j,Did} \le \bar{R}_{j,Di} \ \forall j \in \{1,2,3\} \quad (5)$$

$$A_{j,Did} + A_{j,Diu} + A_{j,Ded} + A_{j,Deu} \le \bar{R}_{j,Di} + \bar{R}_{j,De} \ \forall j \in \{1,2,3\} \quad (6)$$

$$A_{j,Ciz} + A_{j,Diz} \le F_{0,z} \forall j \in \{1,2,3\} \ \forall z \in \{u,d\} \quad (7)$$

$$A_{j,Cez} + A_{j,Dez} \le F_{j,z} \forall j \in \{1,2,3\} \ \forall z \in \{u,d\} \quad (8)$$

$$\frac{\bar{R}_{i,Ceu} + \bar{R}_{i,Ced} + \bar{R}_{i,De}}{F_{i,u} + F_{i,d}} = \frac{\bar{R}_{j,Ceu} + \bar{R}_{j,Ced} + \bar{R}_{j,De}}{F_{j,u} + F_{j,d}} \ \forall i,j \in \{1,2,3\} \quad (9)$$

Here, (1) requires us to maximize the total resource allocation (traffic demand satisfaction) over all the three cross sectors. (2) requires that the total DL (UL) resources allocated to interior and exterior traffic in all the three sectors cannot be more than total DL (UL) resources available in the system. (3) and (4) require that the allocation to interior traffic and net traffic be no more than the interior and net traffic demands respectively in each sector in either direction for cellular traffic. These two constraints together allow for allocation of exterior traffic resources to interior traffic, while at the same time preventing exterior traffic from operating on interior traffic resources. A similar set of constraints apply to D2D traffic in (5) and (6), with the additional *flexibility that the D2D traffic demand can be met by DL and UL resources jointly, thereby allowing for better resource utilization*. (7) and (8) require the net allocation to interior and exterior traffic in each sector in either direction be limited by the interior and exterior traffic resources respectively that would be made available from dynamic FFR. Finally, (9) ensures a proportional allocation across the cross sectors.

Observe that the resource allocation variables $F_{0,z}, F_{1,z}, F_{2,z}$ and $F_{3,z}$ are required to be integers. However, we solve the above optimization problem after relaxing that constraint and later round off the values to the nearest integer such that all resources are allocated to at least one sector. This makes the above optimization a linear optimization problem that can be solved efficiently. Further, it needs to be executed only once every epoch resulting in low overhead. Note that generally the number of resources to be allocated is a large number, and thus, rounding off does not cause significant loss of optimality.

D. Joint Cellular and D2D Scheduling in R2D2

At frame level granularity, *R2D2* resolves the conflicts generated by D2D traffic due to localized dynamic FFR. In the process, it also performs efficient scheduling of cellular (interior and exterior) and D2D traffic (interior and exterior) jointly across DL and UL resources as well as across co-located sectors to maximize resource utilization.

1) Scheduling Model: Scheduling problems are typically formulated as utility maximization problems, where the objective is to maximize the end-to-end system throughput subject to a desired fairness model (captured by the utility function, U_k). We assume proportional fairness (PR, $U_k = \log(\bar{r}_k)$) that is the de-facto fairness model in cellular networks [17]. The problem now reduces to $\max_k \beta_k \log \bar{r}_k$, where β_k captures the priority weight of user k's QoS class and $\bar{r}_k$ its average throughput. The system solution can be shown to converge to the optimum PF allocation at longer (epoch) time scales if the scheduler's decisions at each frame are made to maximize the aggregate marginal utility, $S_{\max} = \arg\max_S \{\sum_{k\in S} \Delta U_k\}$ [18]. ΔU_k denotes the marginal utility received by user k in a valid schedule S and is given by $\frac{\beta_k r_k}{\bar{r}_k}$ for PF, where r_k is the aggregate instantaneous rate received by the user in the frame. Thus, in each frame t, user weight $v_k(t) = \frac{\beta_k}{\bar{r}_k(t)}$ varies with $\bar{r}_k(t)$ and accounts for both fairness and QoS. The scheduling problem at each cell-site of co-located sectors then reduces to determining the frame schedule for the co-located sectors that maximizes the following aggregate weighted rate: $S_{\max}(t) = \arg\max_S \sum_{k\in S} v_k(t) \cdot r_k(t)$

2) Problem Formulation: Next, we formulate our joint DL-UL scheduling problem. We use $x_{k_j,n}$ to denote if user k_j in sector j is scheduled on RB n. $F_{j,z}$ denotes the number of exterior resources available to sector j in direction z (UL/DL). $F_{0,z}^j$ denotes the number of interior resources available to j in direction z. *R2D2* computes both these values for all the sectors in both the directions as explained in Section IV-C. Observe that different sectors in a co-located cluster can differ in the number of interior resources available ($F_{0,z}^j$).

$$
\begin{aligned}
&\max_{x_{k_j,n}\forall k_j,n,j} \sum_{j\in\{1,2,3\}} \sum_{k_j\in\mathcal{K}_j} v_{k_j} \quad \sum_{n\in F_j} x_{k_j,n} r_{k_j,n} && (10)\\
&\text{where,} \quad \mathcal{K}_j \leftarrow \mathcal{C}_{j,d} \cup \mathcal{C}_{j,u} \cup \mathcal{D}_j \\
&\qquad\qquad F_j \leftarrow F_{0,d}^j \cup F_{j,d} \cup F_{0,u}^j \cup F_{j,u} \\
&\text{subject to} \sum_{k_j\in\mathcal{K}_j} x_{k_j,n} \le 1,\ \forall n \in F_j, \forall j && (11)\\
&x_{k_j,n} = 0; \forall n \in F_{0,d}^j \cup F_{j,d} \\
&\qquad \forall k_j \in \mathcal{C}_{j,u}, \forall j && (12)\\
&\qquad = 0; \forall n \in F_{0,u}^j \cup F_{j,u} \\
&\qquad \forall k_j \in \mathcal{C}_{j,d}, \forall j && (13)\\
&\qquad = 0; \forall n \in F_{0,u}^j \cup F_{0,d}^j \\
&\qquad \forall k_j : k_j \text{ is exterior}, \forall j && (14)\\
&\sum_{n\in F_j} x_{k_j,n} r_{k_j,n} \le B_{k_j}; \forall k_j \in \mathcal{K}_j, \forall j && (15)\\
&x_{k_j,n} = \{0,1\}; \quad r_{k_j,n} = h(n, k_j, k_\ell, k_m) && (16)
\end{aligned}
$$

where $\mathcal{C}$ and $\mathcal{D}$ represent the cellular and D2D traffic; and $\mathcal{K}_j$ denotes the set of users in sector j. (11) indicates that in a given sector, at most one transmission can be scheduled on each resource block (RB). (12) and (13) indicate that every RB in the UL and DL can be allocated only to their respective cellular users. (14) prevents an exterior user from using an RB allocated to interior traffic. (15) limits the allocation to each user to be bounded by the finite amount of data in its buffer.

TDD vs. FDD: We note that a D2D user may be allocated RBs both from DL and UL simultaneously in the above formulation. While this can be realized in TDD systems, this would pose a problem for half-duplex clients in FDD systems. Hence, FDD systems would have an additional constraint to ensure that a D2D user be allocated RBs either from DL or UL but not from both.

Reuse across co-located sectors: The rate of a user (k_j) in a sector (j) on a RB (n) (denoted by: $r_{k_j,n}$) depends not only on its individual rate in its sector in the absence of interference ($r'_{k_j,n}$) but also on the other potentially interfering traffic that is co-scheduled on the same RB (due to overlapping dynamic FFR bands) in the other co-located sectors (ℓ and m). While there would be no interference if all traffic were cellular (due to sectorization), the presence of omni-directional D2D traffic could create interference. To estimate the true impact of interference on rate, one would need the channel gain information between all users in all the co-located sectors, which is not feasible to obtain for per-frame scheduling (not to mention the associated overhead). Hence, *R2D2* adopts a conservative approach in identifying users from co-located sectors that can be co-scheduled on the same RB without any interference using information that is already readily available. Note that every user keeps track of the reference signal receiver power (RSRP) from its neighboring cells/sectors. Using RSRP, a user can determine which of its co-located sectors it is closer to, thereby localizing [19] it to a particular half of its sector (Fig. 3d). Now to determine if a user can be co-scheduled on an RB with another user, the receiver of the first link and the transmitter (interferer) of the second link are considered. If both of them are located in opposite halves (e.g. region A and D in Fig. 3d) and are both not interior traffic, then they can be co-scheduled, i.e. $r_{k_j,n} = r'_{k_j,n}$ and 0 otherwise. Note that if either of the two users under consideration is a cellular BS, then there is no interference due to sectorization. When a user is co-scheduled with two other users, the same check can be employed with each of the interfering users independently. Hence, we have $r_{k_j,n} = \{r'_{k_j,n}, 0\}$ depending on if co-scheduled users, *i.e.* k_ℓ and k_m conflict with k_j.

Thus, we see that the *above formulation not only captures the scheduling of cellular and D2D traffic jointly across DL and UL resources but also across co-located sectors (leveraging reuse) in a scalable manner without incurring additional overhead.*

3) Hardness of the Problem: Based on the reduction of 3-bounded 3-dimensional matching problem to a simpler version of (10), we show in our technical report [20] that maximizing (10) is a NP-Hard problem. Further, it admits no $(1-\delta)$-approximation algorithm unless P=NP.

4) Proposed Algorithms: We propose three polynomial-time algorithms with performance guarantees. The first algorithm, *Alg1* has a complexity of $O(N^2K^3)$ and an approximation ratio of 1/2; while the second algorithm *Alg2* has a lower complexity of $O(N^2K)$ to aid in low-latency implementations at the BS (to meet 1ms LTE frame timing) and an approximation ratio of 1/4. Both these algorithms yield close-to-optimal performance in practical evaluations (Section V) and apply to TDD systems. For FDD systems with the additional constraint that a D2D user can be allocated resources in either DL or UL but not both, we provide an alternate algorithm *Alg3* with a different approach that incurs a complexity of $O(N^4K^3)$ and provides an approximation ratio of 1/3. *The proofs for the complexity and approximation ratios for all the three algorithms are included in [20].*

a) Alg1: *Alg1* is a greedy algorithm that requires two inputs: (i) Set of users $\mathcal{K}$ for which the schedule needs to be computed; and, (ii) The set of RBs that can be allocated to each of the user (F_{k_j} for user k_j). To compute the schedule, the BS invokes *Alg1* at the beginning of for every frame by setting $\mathcal{K}$ to the set of users across the co-located sectors: $\mathcal{K} = \mathcal{K}_j \cup \mathcal{K}_\ell \cup \mathcal{K}_m$. For user k_j, F_{k_j} is computed as follows:

$$\begin{aligned}
F_{k_j} &= F^j_{0,z} \cup F_{j,z} \text{ ; if } k_j \in \mathcal{C}_{j,z} \wedge k_j \text{ is interior}, z \in \{u, d\} \\
&= F^j_{0,d} \cup F_{j,d} \cup F^j_{0,u} \cup F_{j,u} \text{ ; if } k_j \in \mathcal{D}_j \wedge k_j \text{ is interior} \\
&= F_{j,z} \text{ ; if } k_j \in \mathcal{C}_{j,z} \wedge k_j \text{ is exterior}, z \in \{u, d\} \\
&= F_{j,d} \cup F_{j,u} \text{ ; if } k_j \in \mathcal{D}_j \wedge k_j \text{ is exterior}
\end{aligned}$$

This allows D2D to be scheduled on both DL and UL resources and allows interior traffic to use exterior resources. *Alg1* works greedily (Line 2-9 of Algorithm 1) and in each iteration, adds that tuple of (n, k_j, k_ℓ, k_m) to the schedule S that maximizes the incremental utility (Line 6-8, $f(n, k_j, k_\ell, k_m)$). Recall (from Section III) that there is negligible benefit to reusing cellular resources by D2D traffic within a sector. Hence, *Alg1* considers only those tuples, where at most three users k_j, k_ℓ, k_m, one from each of the co-located sectors, are scheduled on the same RB (Line 3). This would include tuples with less than three users as well, as these may provide higher utility on a RB depending on the interference conflicts between the co-located sectors on that RB. For feasibility purposes, among all tuples, only those tuples are considered where it is possible to schedule all the users present in the tuple on the associated RB (See Line 4). Further, to ensure that the computed schedule is feasible, before finally adding a tuple to the schedule S, *Alg1* verifies (Line 5) that the updated schedule does not violate any of the constraints specified in (10). Among all the valid tuples (set $\mathbf{T_3}$), the tuple that maximizes the incremental utility is added to S (Line 6-8). *Alg1* returns (Line 9) when it can no longer find a valid tuple that can be added to the schedule.

b) Alg2: *Alg2* is a lower complexity algorithm that decouples the co-located sectors during RB allocation. Similar to *Alg1*, *Alg2* is invoked by the base station by providing two inputs (set of all users across three sectors and the set of RBs for each user). *Alg2* works by greedily (Line 2-7 of Algorithm 2) adding *valid* (Line 3) tuples of (n, k_i) to the schedule that maximize the incremental utility, *i.e.* it allocates an RB to only one user in one of the co-located sectors at a time, although all users from the co-located sectors are considered during the decision process. Further, a new user can also be scheduled on an RB that is already allocated to user(s) from other sectors, subject to interference conflicts and the resulting additional utility the new user would bring to the given RB. This in turn would allow for reuse of RBs across co-located sectors.

Algorithm 1: *Alg1*: Computes the assignment of resources to users that maximizes (10) in a TDD system.

Input : Set of users $\mathcal{K}$, $F_{k_i} \forall k_i \in \mathcal{K}$
1 $S \leftarrow \{\}$
2 **while** *true* **do**
3 $\quad \mathbf{T_1} \leftarrow (n, k_j, k_\ell, k_m) : k_i \in (\mathcal{K}_i \cap \mathcal{K}) \cup \{\phi\} \ \forall k_i \in \{k_j, k_\ell, k_m\}$
4 $\quad \mathbf{T_2} \leftarrow (n, k_j, k_\ell, k_m) : (n, k_j, k_\ell, k_m) \in \mathbf{T_1} \wedge (k_j = \phi \,||\, n \in F_{k_j}) \wedge (k_\ell = \phi \,||\, n \in F_{k_\ell}) \wedge (k_m = \phi \,||\, n \in F_{k_m})$
5 $\quad \mathbf{T_3} \leftarrow (n, k_j, k_\ell, k_m) : (n, k_j, k_\ell, k_m) \in \mathbf{T_2} \wedge$ Adding (n, k_j, k_ℓ, k_m) to S does not violate any constraints
6 $\quad (n^*, k_j^*, k_\ell^*, k_m^*) \leftarrow \text{argmax}_{(n,k_j,k_\ell,k_m) \in \mathbf{T_3}} f(n, k_j, k_\ell, k_m)$
7 $\quad$ **if** $n^* \neq \phi$ **then**
8 $\quad\quad S \leftarrow S \cup (n^*, k_j^*, k_\ell^*, k_m^*)$
9 $\quad$ **else return** S

Algorithm 2: *Alg2*: Computes the assignment of resources to users that maximizes (10) in a TDD system.

Input : Set of users $\mathcal{K}$, $F_{k_i} \forall k_i \in \mathcal{K}$
1 $S \leftarrow \{\}$
2 **while** *true* **do**
3 $\quad \mathbf{T_1} \leftarrow (n, k_i) : n \in F_{k_i} \wedge k_i \in \mathcal{K} \wedge$ Adding (n, k_i) to S does not violate any constraints
4 $\quad (n^*, k_i^*, k_\ell^*, k_m^*) \leftarrow \text{argmax}_{(n,k_i,k_\ell,k_m):(n,k_i) \in \mathbf{T_1} \wedge (n,k_\ell) \in S \wedge (n,k_m) \in S} \leftarrow f(n, k_i, k_\ell, k_m) - f(n, \phi, k_\ell, k_m)$
5 $\quad$ **if** $n^* \neq \phi \wedge (f(n, k_i^*, k_\ell^*, k_m^*) - f(n, \phi, k_\ell^*, k_m^*)) > 0$ **then**
6 $\quad\quad S \leftarrow S \cup (n^*, k_i^*)$
7 $\quad$ **else return** S

c) Alg3: *Alg3* is tailored for FDD systems but can also be applied to TDD systems. In a FDD system, a D2D transmission can either use the UL RBs or the DL RBs (but not both). This additional constraint coupled with finite user buffers makes the problem even more challenging and warrants a new approach to ensure performance guarantees. Unlike *Alg1* and *Alg2* that allocate at the granularity of RBs, *Alg3* instead makes allocations at the granularity of users, *i.e.* allocation of RBs to a user is executed in a single step before moving to another user. *Alg3* first determines the set of RBs that can be allocated to user k_i. Here, $\mathcal{K}'$ denotes the set of users for which the set of RBs that can be allocated to them has not been determined yet (Line 1 of Algorithm 3). While cellular DL (UL) users are restricted to the set of DL (UL) RBs (Line 6-11), D2D users have the flexibility of being allocated either from the set of DL or UL RBs (but not both, Line 3-5). In every iteration of the while loop (Line 12-14), *Alg3* determines the set of RBs that can be allocated for exactly one user. This is done by greedily selecting (Line 13) the tuple (k_i, F_{k_i}) that when added to the possible schedule maximizes the value of the incremental utility. Computing the incremental utility of adding a user (along with its set of potential RBs for allocation) to the current schedule, in turn corresponds to the scheduling problem (10) considered in TDD systems. Hence, *Alg1* is invoked (Line 13) to compute the incremental utility of adding a tuple to the schedule. *Alg3* then determines the user and its set of RBs (whether DL or UL RBs for D2D users) that provides the maximum incremental utility and adds it to the current schedule (Line 14). Note that at each iteration only

Algorithm 3: *Alg3*: Computes the assignment of resources to users that maximizes (10) in a FDD system.

Input : Set of users $\mathcal{K}$
$\mathcal{K}' \leftarrow \phi, S \leftarrow \phi$
forall the $k_i \in \mathcal{K}$ **do**
  if $k_i \in \mathcal{D}_j \cup \mathcal{D}_\ell \cup \mathcal{D}_m$ **then**
    if k_i is exterior **then** $F_{k_i}[0] \leftarrow F_{i,d},\ F_{k_i}[1] \leftarrow F_{i,u}$
    else $F_{k_i}[0] \leftarrow F^i_{0,d} \cup F_{i,d},\ F_{k_i}[1] \leftarrow F^i_{0,u} \cup F_{i,u}$
  else if $k_i \in \mathcal{C}_{j,d} \cup \mathcal{C}_{\ell,d} \cup \mathcal{C}_{m,d}$ **then**
    if k_i is exterior **then** $F_{k_i}[0] \leftarrow F_{i,d}, F_{k_i}[1] \leftarrow \phi$
    else $F_{k_i}[0] \leftarrow F^i_{0,d} \cup F_{i,d}, F_{k_i}[1] \leftarrow \phi$
  else if $k_i \in \mathcal{C}_{j,u} \cup \mathcal{C}_{\ell,u} \cup \mathcal{C}_{m,u}$ **then**
    if k_i is exterior **then** $F_{k_i}[0] \leftarrow F_{i,u}, F_{k_i}[1] \leftarrow \phi$
    else $F_{k_i}[0] \leftarrow F^i_{0,u} \cup F_{i,u}, F_{k_i}[1] \leftarrow \phi$
while $\mathcal{K} \neq \mathcal{K}'$ **do**
  $(k^*, F^*_{k^*}) \leftarrow \text{argmax}_{(k_i,F_{k_i}):k_i \in \mathcal{K}\setminus\mathcal{K}' \wedge F_{k_i} \in \{F_{k_i}[0],F_{k_i}[1]\}} f(\mathit{Alg1}(\mathcal{K}' \cup \{k_i\}, F_{k_j} \forall k_j \in \mathcal{K}' \cup \{k_i\})) - f(\mathit{Alg1}(\mathcal{K}', F_{k_j} \forall k_j \in \mathcal{K}'))$
  $\mathcal{K}' \leftarrow \mathcal{K}' \cup \{k^*\}, F_{k^*} \leftarrow F^*_{k^*}$
Call *Alg1* $(\mathcal{K}', F_{k_j} \forall k_j \in \mathcal{K}')$
return schedule computed by *Alg1*

a user and its operational set (DL or UL RBs if D2D user) are determined and remain fixed as the schedule evolves. The specific RBs themselves that are allocated to existing users in the schedule may change as new users are added to the schedule, due to the re-computation of a TDD schedule in each iteration.

V. Evaluation

A. Setup

A frame-level simulator written in C++ is considered for evaluation of the proposed algorithms. In our evaluation, we set up 19 base stations each with 3 sectors. Further, each sector has number of users varying between 50-150 and they are placed at random location within the sector. Each user generates a request to a randomly selected file server for a file of size 500 KB using an exponential distribution. The default percentage of UL transmissions in each sector was 30%. Further, 30% of the generated requests are D2D, implying that the destination file server is another user in the same sector. For all algorithms, number of RBs available for UL and DL communication were 9 and 21, respectively. Every data session (user) has a finite amount (125 KB) of data (buffer). The Okumara-Hata urban path loss model is employed along with log-normal shadowing and fast (Rayleigh) fading. The doppler fading for each user's Rayleigh channel is equivalent to a velocity of 3-10 Km/hour. The SNR from the model is mapped to a specific data rate [16]. The cells are deployed in the hexagonal fashion with a predetermined sector radius (1000m, by default). Each data point presented here is an average over results from 10 randomly generated topologies.

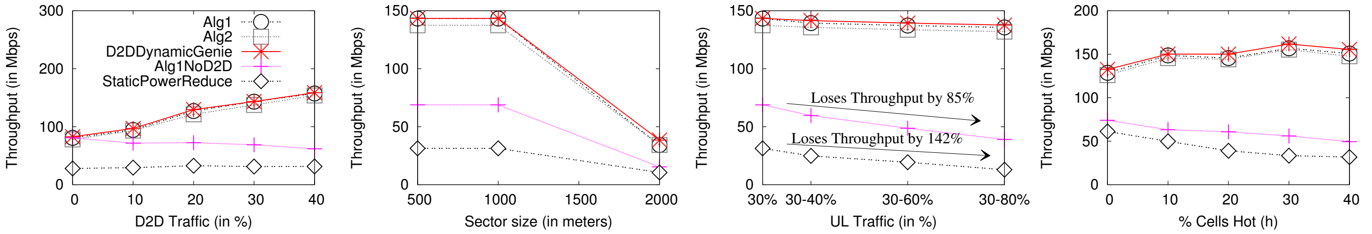
We restrict our evaluations to TDD systems here (since inferences for FDD are similar). Apart from the two flavors of *R2D2* (D2D-aware dynamic FFR with *Alg1* or *Alg2*), we implemented multiple other algorithms for comparison. In all these algorithms (except StaticPowerReduce), at epoch level time granularity, dynamic FFR scheme is executed for determining the FFR bands (as explained in Section IV-C). However, the algorithms behave differently at the frame level time granularity as follows. *a*) **Alg1NoD2D:** Every BS computes the schedule using *Alg1*. Since, this algorithm classifies all traffic as cellular, there is no conflict among transmissions scheduled by co-located sectors. *b*) **D2DDynamicGenie:** Each BS computes the joint schedule for all the three co-located sectors. Further, D2DDynamicGenie is aware of all the path loss values (including, path loss between users), and can potentially leverage that information to schedule multiple transmissions in the same sector on the same RB. It also dynamically controls the transmission power level of D2D transmissions to maximize the reuse of RBs. *c*) **StaticPowerReduce** proposed in [4] works by reducing the transmission power level of all D2D links. It is expected that this reduction will prevent D2D from interfering with UL transmissions, allowing them to be co-scheduled on the same RBs in the same sector. This algorithm does not allow outdoor D2D traffic to coexist with cellular DL traffic (to avoid interference to the cellular users operating on the same RB in the same sector). We extend [4] to multicell deployments through a static FFR scheme that allocates equal number of RBs to all sectors. We evaluate the scheduling algorithms at per-frame granularity based on their objective function (weighted sum rate), which in turn corresponds to throughput with equal user weights.

B. Results

1) Variation in D2D traffic load: With increase in D2D traffic, the throughput for both *Alg1* and *Alg2* increases (See Fig. 4a). This is primarily because of *offloading* benefit from D2D. Further, D2D is also able to leverage the disparity in UL-DL traffic distribution and thus, can be scheduled on the RBs that would have otherwise gone wasted. Fig. 4a shows that *Alg2* gives a throughput of 3.79x and 3.00x compared to StaticPowerReduce [4] and Alg1NoD2D, respectively. Although, *Alg2* has lower complexity than *Alg1*, we observe that on an average, throughput of *Alg2* is within 3.18% of the throughput of *Alg1*, making it an ideal candidate for implementation. This is because when *Alg2* adds a single user to the schedule, it also considers the interference that may arise between this user and the other users that are already scheduled on the same RB.

From the result, we can also see that *Alg2* has throughput within 4.8% of D2DDynamicGenie. This implies that even with all the information about path loss values between users, D2DDynamicGenie is unable to schedule more transmissions than *Alg2*. Dynamic control of transmission power level is also not helpful since D2D transmitters need to transmit at low power to avoid interference to other transmissions in the same sector. This coupled with the interference faced by the D2D receiver, reduces the rate that can be supported on the D2D link. This confirms that D2D provides low reuse benefit in multicell scenarios with FFR. Although, StaticPowerReduce allows D2D transmissions, reducing the transmission power level of D2D transmissions also reduces D2D offload benefit. Further, even after reducing the power levels, D2D transmissions may still interfere with the UL traffic and vice versa. From our evaluations, we observe that at 30% D2D traffic load, 25% of the transmissions scheduled by StaticPowerReduce were unsuccessful due to interference.

2) Variation in sector radius: For larger sectors, all algorithms have lower throughput since the increase in distance results in lower physical layer data rates. When sector radius



(a) Varying D2D traffic. (b) Varying sector radius. (c) Varying UL traffic ratio. (d) Varying number of hot sectors (h).

Fig. 4. Evaluation results illustrating challenges in leveraging D2D.

is less than 1000m, throughput of *Alg2* is within 4.3% of D2DDynamicGenie. However, at sector radius of 2000m, the gap between them is higher at 12.3%, but still within reasonable limits. This is because in larger cells, D2DDynamicGenie is able to leverage the path loss information to schedule multiple transmissions within the same sector.

3) Variation in DL-UL traffic ratio: Fig. 4c shows the variation in throughput with variation in the percentage of UL traffic across different sectors. Here, an x-axis value of 30-60% implies that the UL traffic percentage of all sectors was uniformly distributed between 30% and 60%. For StaticPowerReduce and Alg1NoD2D, the total throughput peaks when the average offered UL traffic (30%) is close to the percentage RBs allocated for UL traffic (30%). The maximum throughput of *Alg2* stays almost constant with variation in UL traffic. On the other hand, for Alg1NoD2D and StaticPowerReduce, their maximum throughput is 85% and 142% more than their minimum throughput. This clearly implies that D2D-aware dynamic FFR algorithm of *R2D2* is able to leverage the flexible nature of D2D traffic and is able to achieve high throughput even when disparity is high.

4) Variation in hotspot D2D traffic: We model the D2D traffic distribution as "hotspot traffic" such that for h% of sectors, 80% of their traffic is D2D, while the remaining sectors have only 20% of their traffic as D2D (to emulate scenarios of stadiums, event centers, etc.). Fig. 4d shows that *Alg2* provides 2.33x and 3.45x throughput compared to StaticPowerReduce [4] and Alg1NoD2D, respectively. The gains here are higher compared to when D2D traffic is uniformly distributed since here the dynamic FFR scheme of *R2D2* is able to better leverage the variation in D2D traffic across sectors.

VI. Conclusions

In this paper, we showed that D2D traffic does not provide significant reuse benefit in multicell networks that employ FFR schemes. To best leverage the D2D in such networks, we proposed that D2D should be used as a flexible traffic so that it can be used on resources which would have otherwise been wasted due the disparity between uplink and downlink traffic across sectors. To that end, we proposed a novel D2D aware dynamic FFR algorithm that is executed at epoch level granularity and is scalable. We also proposed provable approximate scheduling algorithms that maximize the benefit from D2D. Through simulations, we showed that R2D2 improves performance by 2.79x compared to existing D2D algorithms.

References

[1] *LTE Release 12*, 3rd Generation Partnership Project (3GPP), 2013-2014, http://www.3gpp.org/Release-12.

[2] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "D2D Communication as an Underlay to LTE-Advanced Networks," *IEEE Communications Magazine*, vol. 47, no. 12, pp. 42–49, 2009.

[3] L. Lei, Z. Zhong, C. Lin, and X. Shen, "Operator Controlled D2D Communications in LTE-Advanced Networks," *IEEE Wireless Communications*, vol. 19, no. 3, pp. 96–104, 2012.

[4] P. Jänis and et al., "D2D Communication Underlaying Cellular Communications Systems," *IJCNS*, vol. 2, no. 3, pp. 169–178, 2009.

[5] K. Doppler, C.-H. Yu, C. B. Ribeiro, and P. Janis, "Mode Selection for D2D Communication Underlaying an LTE-Advanced Network," in *Proc. of IEEE WCNC*, 2010.

[6] U. Paul and et al., "Understanding Traffic Dynamics in Cellular Data Networks," in *Proc. of IEEE INFOCOM*, 2011.

[7] B. Kaufman and B. Aazhang, "Cellular Networks with an Overlaid Device to Device Network," in *Proc. of IEEE Asilomar Conference on Signals, Systems and Computers*, 2008.

[8] T. Peng and et al., "Interference Avoidance Mechanisms in the Hybrid Cellular and D2D Systems," in *Proc. of IEEE PIMRC*, 2009.

[9] N. Lee, X. Lin, J. G. Andrews, and R. W. Heath Jr, "Power Control for D2D Underlaid Cellular Networks: Modeling, Algorithms and Analysis," *arXiv preprint arXiv:1305.6161*, 2013.

[10] G. Fodor and N. Reider, "A Distributed Power Control Scheme for Cellular Network Assisted D2D Communications," in *Proc. of IEEE GLOBECOM*, 2011.

[11] Q. Duong, Y. Shin, and O.-S. Shin, "Resource Allocation Scheme for D2D Communications Underlaying Cellular Networks," in *IEEE Intl. Conf. on Computing, Management and Telecommunications (ComManTel)*, 2013.

[12] S. H. Ali and V. C. Leung, "Dynamic Frequency Allocation in FFR OFDMA Networks," *IEEE TWC*, vol. 8, no. 8, pp. 4286–4295, 2009.

[13] R. Y. Chang and et al., "A Graph Approach to Dynamic FFR in Multi-Cell OFDMA Networks," in *Proc. of IEEE ICC*, 2009.

[14] A. L. Stolyar and H. Viswanathan, "Self-Organizing Dynamic FFR for Best-Effort Traffic Through Distributed Inter-Cell Coordination," in *Proc. of IEEE INFOCOM*, 2009.

[15] M. Andrews and L. Zhang, "Scheduling algorithms for multi-carrier wireless data systems," in *ACM MOBICOM*, Sept 2007.

[16] M. Arslan, J. Yoon, K. Sundaresan, S. Krishnamurthy, and S. Banerjee, "FERMI: A Femtocell Resource Management System for Interference Mitigation in OFDMA Networks," in *Proc. ACM MOBICOM*, 2011.

[17] E. Liu, Q. Zhang, and K. K. Leung, "Relay-Assisted Transmission with Fairness Constraint for Cellular Networks," *IEEE TMC*, vol. 11, no. 2, pp. 230–239, 2012.

[18] G. Song and Y. Li, "Cross-Layer Optimization for OFDM Wireless Networks - Part I: Theoretical Framework," *IEEE TWC*, vol. 4, no. 2, 2005.

[19] J. Paek and et al., "Energy-Efficient Positioning for Smartphones using Cell-id Sequence Matching," in *Proc. of ACM MOBISYS*, 2011.

[20] "R2D2 Technical Report," Tech. Rep., 2013, http://www.cse.ohio-state.edu/~bansal/R2D2TechRep.pdf.