Singh, Bikramjit; Li, Zexian; Uusitalo, Mikko

Flexible Resource Allocation for Device-to-Device Communication in FDD System for Ultra-Reliable and Low Latency Communications

Published in:
Advances in Wireless and Optical Communications (RTUWO)

DOI:
10.1109/RTUWO.2017.8228531

Published: 01/01/2017

Document Version
Peer reviewed version

Please cite the original version:
Singh, B., Li, Z., & Uusitalo, M. (2017). Flexible Resource Allocation for Device-to-Device Communication in FDD System for Ultra-Reliable and Low Latency Communications. In *Advances in Wireless and Optical Communications (RTUWO)* (pp. 186-191). IEEE. https://doi.org/10.1109/RTUWO.2017.8228531

Flexible Resource Allocation for Device-to-Device Communication in FDD System for Ultra-Reliable and Low Latency Communications

Bikramjit Singh
Department of Communications and Networking,
Aalto University, Espoo, Finland
bikramjit.singh@aalto.fi

Zexian Li, Mikko A. Uusitalo
Nokia Bell Labs,
Espoo, Finland
{zexian.li,mikko.uusitalo}@nokia-bell-labs.com

***Abstract*—*Ultra-reliable and low latency communications* is envisioned to enable new services and applications with high reliability, availability and low latency, e.g, factory automation, Tactile Internet. Device-to-device communication is one such mean that allows devices to experience benefits in terms of shorter communication latency. Currently most device-to-device communication in *frequency-division duplex* system occurs in uplink spectrum. However, due to unbalanced cellular traffic in uplink and downlink bands may result poor spectrum utilization of network resource. In this paper, we investigate resource allocation where network offers reliable communication using normal cellular and device-to-device communications. We consider flexible allocation for device-to-device communication in both uplink and downlink bands in a frequency-division duplex system. The network optimizes resource allocation to improve the network rate with 99.999 % availability target. We demonstrate the usability of the proposed protocol in a factory scenario.**

Keywords*—*5G; ultra-reliable and low latency communications (URLLC); device-to-device (D2D); frequency-division duplex (FDD); reliability; availability

I. Introduction

Ultra-Reliable and Low Latency Communications (URLLC) is an agreed use case scenario in future Fifth generation (5G) wireless communication systems that enables real-time control and automation of dynamic processes in, e.g., factory automation, manufacturing and control, smart grid protection, traffic management and safety, or "Tactile Internet" [1]. Such services has very stricter requirements which consider, e.g., 99.999 % reliability under the End-to-End (E2E) latency of 1 ms [2]. Here, given the maximum allowable E2E latency, including jitter/re-transmissions of 1 ms, the maximum packet error rate must not be higher than 10^{-5}. Long-Term Evolution (LTE) offers guaranteed bit rate that can support packet error rates down to 10^{-6}. However, the delay budget goes up to 300 ms which includes both radio, transport and core network latencies. LTE systems may offer better rate availability, but each Hybrid Automatic Repeat Request (HARQ) round-trip takes at least 8 ms. Therefore, Radio Access Network (RAN) delay budget easily exceeds E2E latency targets of 1-10 ms [3].

Strategies to provide reliable services can be established on based time-varying traffic demands. For example in [4], the rate availability is improved for the worst performing User Equipments (UEs) by scheduling more resource based on max-min strategy. A strategy based on switching between sleeping and transmission mode is considered by Base Stations (BSs) in a cluster [5]. The scheme is based on time slot based transmissions and may not be useful for tight latency services. Different spectrum bands utilization to offer different reliable services is proposed in [6]. Various coordination schemes are discussed in [7]–[9] targeting resource management and interference coordination amongst the neighboring cells in presence of inter-cell interference. However, to provide guaranteed rate, orthogonal sharing serves best for URLLC provisioning as it eliminates co-channel interference [3].

State-of-the-art methods offering reliable services are based on diversity. Diversity in the form of repeated transmissions for contention-based access is analyzed in [10]. However, it is vital to control random collisions to ensure targeted reliability. Macro diversity where a UE is simultaneously connected to multiple BSs is one alternative. Multi-connectivity is another potential solution where a UE has multiple connections on different carriers, possibly to different BSs [11]. Multi-hop communication for creating routing diversity is considered in [12] for URLLC. Spatial diversity can be achieved using Device-to-Device (D2D) communications as well. Further, it curtails tight latency requirements for URLLC by reducing two-hop cellular communication to one-hop. D2D transmissions can be employed in both Time Division Duplex (TDD) or Frequency-Division Duplex (FDD) systems. In TDD, more efficient D2D resource usage can be achieved in comparison to current FDD implementation [13], [14]. However, TDD model has higher requirements of switching between uplink and downlink. Time synchronization for inter-operator D2D support also becomes challenging between the operators [15]. Beside, TDD systems increases the latency requirements and may not be well suited for URLLC or tactile services. On the other hand, currently in LTE FDD, only uplink spectrum is allowed for the D2D resource usage, and may cause inefficient resource usage for URLLC [13].

In this paper, we consider cellular-D2D communication network for URLLC. For latency reduction we adopt a 5G numerology as discussed in [16], [17], with a 0.125 ms Transmission Time Interval (TTI). With such a frame structure, it is a possibility to support cellular and D2D transmissions with 1 ms E2E latency. We concentrate on *URLLC availability* in a multi-user, multi-cell network. The availability metric is defined as a minimum data rate offered in the network's coverage area. However, with the current D2D operation limited to FDD uplink band, coupled with traffic unbalance and differentiated channel conditions in uplink-downlink directions may result lower capacity utilization for cellular-D2D network. We inspect flexible allocation between D2D and cellular communication for URLLC. The allocation between D2D and cellular traffic is optimized in a view by granting resources to the worst performing UEs to improve the rates achieved with 99.999 % availability in the network.

II. System Model

We consider a wireless network with multiple UEs in industrial environments such as factories. We consider both cellular and D2D wireless communication-mode, as shown in Fig. 1. We use FDD system where cellular uplink and downlink communication works independently, but works alongside with D2D communication through flexible allocation. The resource allocation can be classified as Overlay D2D operation where the D2D resource can come from either uplink or downlink or both. To provide communication between the UEs, either of the cellular or D2D link can be selected based on the higher offered rate and the feasibility of D2D links. Therefore both links constitute peer-to-peer communication unlike cognitive scenario where UEs rights are classified as primary and secondary.

Consider a BS with cellular load pairs N_{C} and D2D pairs N_{D2D}. We assume for up to certain optimal Line-of-Sight (LOS) distance, the communicating UE pairs can use D2D communication-mode. Let us assume the number of channels in uplink and downlink frequency bands is equal to W, and $W_{\mathrm{D2D}}^{\mathrm{UL}}$ and $W_{\mathrm{D2D}}^{\mathrm{DL}}$ denotes the number of channels allocated to D2D transmissions in cellular uplink and downlink bands respectively. Resultant, D2D pairs enjoy services on W_{D2D} channels where

$$W_{\mathrm{D2D}} = W_{\mathrm{D2D}}^{\mathrm{UL}} + W_{\mathrm{D2D}}^{\mathrm{DL}}.$$

On the other hand, $W_{\mathrm{C}}^{\mathrm{UL}}$ are the channels available to cellular uplink transmissions, and $W_{\mathrm{C}}^{\mathrm{UL}}$ for downlink where

$$\begin{aligned} W_{\mathrm{C}}^{\mathrm{UL}} &= W - W_{\mathrm{D2D}}^{\mathrm{UL}}, \\ W_{\mathrm{C}}^{\mathrm{DL}} &= W - W_{\mathrm{D2D}}^{\mathrm{DL}}. \end{aligned}$$

Fig. 2 depicts resource sharing between cellular and D2D communications. Each UE uses all the channels in a TTI, i.e., UEs are scheduled based on Time Division Multiplexing (TDM). Each node employs Multiple-Input Multiple-Output (MIMO) communications and the achievable data rate r_i on a link between transmitting node and receiving node pair i is obtained similar to [12]. For simplicity, the allocation between D2D and cellular communication is kept orthogonal. However, more advance model can consider spectrum re-use between D2D and uplink communication where achievable data rate r_i is subjected to inflicted interference. Since a BS is a strong interference source, concurrent transmissions of D2D and downlink may not be allowed.

III. Flexible Resource Allocation

Flexible resource allocation schemes are proposed between cellular and D2D communications in a FDD system. In the network, both cellular and D2D traffic that have the

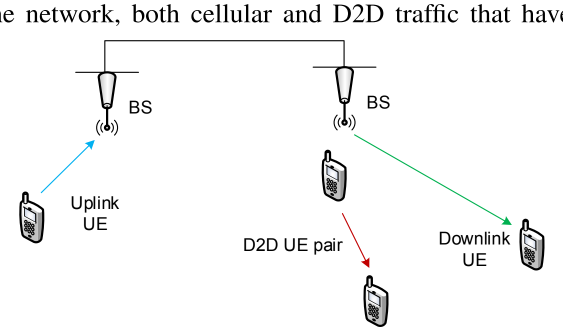

Fig. 1. URLLC in a factory scenario.

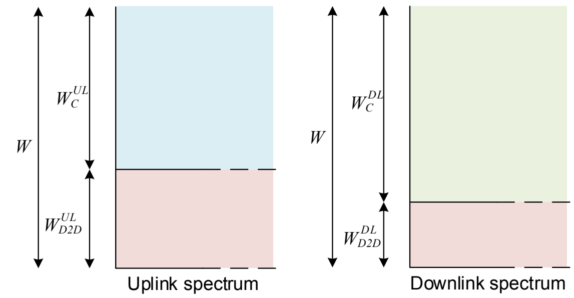

Fig. 2. Resource allocation between cellular and D2D communications in FDD.

same Quality-of-Service (QoS) objective, e.g., all UEs are with URLLC traffic. The use of D2D communication brings spatial diversity, e.g., if cellular link under-performs, D2D connection can be utilized or vice-versa. However, D2D communication can only be utilized if the distance between UE pair is under the certain distance, and the achievable rate estimate is higher than the rate with cellular links. So, either link is utilized that offers better rates and therefore both links are equally emphasized.

Two schemes are discussed, (1) fast allocation and (2) URLLC-optimized allocation. In fast allocation, the spectrum resource is shared between cellular and D2D communications according to the traffic load proportion. The allocation is fast and robust, and no Channel State Information (CSI) is needed. The allocation offers communication improvement to some extent, and can be targeted for medium-capacity region. The second scheme considers maximization of the targeted rate, e.g., the rate available to 99.999 % of the locations and the allocation can be targeted for ultra-low capacity region, e.g., in URLLC.

A. Fast allocation

Let us first consider a single-cell or isolated cell network with N_{C} cellular pairs and N_{D2D} D2D pairs. As all the cellular pairs exist within a cell where $N_{\mathrm{C2C}}^{\mathrm{UL}}$ and $N_{\mathrm{C2C}}^{\mathrm{DL}}$ exists within the same cell, and is related as $N_{\mathrm{C}}^{\mathrm{UL}} = N_{\mathrm{C}}^{\mathrm{DL}} = N_{\mathrm{C}}$. Due to absence of feedback, we consider equal allocation of spectrum resource for cellular pairs in both uplink and downlink direction, i.e.,

$$W_{\mathrm{C}}^{\mathrm{UL}} = W_{\mathrm{C}}^{\mathrm{DL}}. \tag{1}$$

Similarly for D2D communication, its allocation W_{D2D} is

$$W_{\mathrm{D2D}} = 2W - \left(W_{\mathrm{C}}^{\mathrm{UL}} + W_{\mathrm{C}}^{\mathrm{DL}}\right) = 2W_{\mathrm{D2D}}^{\mathrm{UL}} = 2W_{\mathrm{D2D}}^{\mathrm{DL}}. \tag{2}$$

Without channel feedback, one way to allocate resources between different communication modes can be based on traffic load proportion. For example, in [18], dynamic allocation of muted resources between small and its macro-connected cell has been demonstrated. There, it adjusts Almost Blank Sub-frame (ABS) ratio on a fast basis, where ABS resources percentage varies according to the instantaneous load proportion between small and macro cell. However unlike in time domain, here dynamic allocation is pivoted in frequency domain, where spectrum resource is allocated according to their load proportion between cellular and D2D transmissions. For example, if cellular traffic proportion is more, then certainly it has more resource requirement. The allocation then between cellular and D2D communication follows as

$$\frac{W_{\mathrm{C}}^{\mathrm{UL}}}{N_{\mathrm{C}}} = \frac{W_{\mathrm{D2D}}}{N_{\mathrm{D2D}}} \tag{3}$$

where $N_{\mathrm{C}} > 0$ and $N_{\mathrm{D2D}} > 0$. From Eq. (1)-(3), we obtain resultant allocation as

$$\begin{aligned} W_{\mathrm{C}}^{\mathrm{UL}} = W_{\mathrm{C}}^{\mathrm{DL}} &= \frac{2WN_{\mathrm{C}}}{2N_{\mathrm{C}} + N_{\mathrm{D2D}}}, \\ W_{\mathrm{D2D}}^{\mathrm{UL}} = W_{\mathrm{D2D}}^{\mathrm{DL}} &= \frac{WN_{\mathrm{D2D}}}{2N_{\mathrm{C}} + N_{\mathrm{D2D}}}. \end{aligned} \tag{4}$$

In the event of no traffic in either transmission type, the whole bandwidth is utilized by the other. For example, in case of no D2D traffic, cellular uplink and downlink bandwidth usage is $W_{\mathrm{C}}^{\mathrm{UL}} = W_{\mathrm{C}}^{\mathrm{DL}} = W$. On the other hand, with no cellular traffic, D2D transmissions occur in whole spectrum $W_{\mathrm{D2D}} = W_{\mathrm{D2D}}^{\mathrm{UL}} + W_{\mathrm{D2D}}^{\mathrm{DL}} = 2W$.

In a multi-cell network where $N_{\mathrm{C}}^{\mathrm{UL}}$ and $N_{\mathrm{C}}^{\mathrm{DL}}$ may not be the same within a cell. Then each cell allocates spectrum emulating the load proportion for different transmission types accordingly

$$\frac{W_{\mathrm{C}}^{\mathrm{UL}}}{N_{\mathrm{C}}^{\mathrm{UL}}} = \frac{W_{\mathrm{C}}^{\mathrm{DL}}}{N_{\mathrm{C}}^{\mathrm{DL}}} = \frac{W_{\mathrm{D2D}}}{N_{\mathrm{D2D}}}, \tag{5}$$

and subject to the constraints

$$\begin{aligned} W &= W_{\mathrm{C}}^{\mathrm{UL}} + W_{\mathrm{D2D}}^{\mathrm{UL}}, \\ W &= W_{\mathrm{C}}^{\mathrm{DL}} + W_{\mathrm{D2D}}^{\mathrm{DL}}, \\ W_{\mathrm{D2D}} &= W_{\mathrm{D2D}}^{\mathrm{UL}} + W_{\mathrm{D2D}}^{\mathrm{DL}}. \end{aligned} \tag{6}$$

From Eq. (5)-(6), we obtain resultant allocation in a multi-cell scenario as

$$\begin{aligned} W_{\mathrm{C}}^{\mathrm{UL}} &= \frac{2WN_{\mathrm{C}}^{\mathrm{UL}}}{N_{\mathrm{C}}^{\mathrm{UL}} + N_{\mathrm{C}}^{\mathrm{DL}} + N_{\mathrm{D2D}}}, \\ W_{\mathrm{C}}^{\mathrm{DL}} &= \frac{2WN_{\mathrm{C}}^{\mathrm{DL}}}{N_{\mathrm{C}}^{\mathrm{UL}} + N_{\mathrm{C}}^{\mathrm{DL}} + N_{\mathrm{D2D}}}, \\ W_{\mathrm{D2D}}^{\mathrm{UL}} &= W\frac{N_{\mathrm{C}}^{\mathrm{DL}} + N_{\mathrm{D2D}} - N_{\mathrm{C2C}}^{\mathrm{UL}}}{N_{\mathrm{C}}^{\mathrm{UL}} + N_{\mathrm{C}}^{\mathrm{DL}} + N_{\mathrm{D2D}}}, \\ W_{\mathrm{D2D}}^{\mathrm{DL}} &= W\frac{N_{\mathrm{C}}^{\mathrm{UL}} + N_{\mathrm{D2D}} - N_{\mathrm{C2C}}^{\mathrm{DL}}}{N_{\mathrm{C}}^{\mathrm{UL}} + N_{\mathrm{C}}^{\mathrm{DL}} + N_{\mathrm{D2D}}}. \end{aligned} \tag{7}$$

The implementation is fast and robust and does not depend on the CSI of the links. Given the static environment, the allocation may achieve better QoS than if D2D transmission is limited in the uplink band only because here the bottleneck between cellular uplink and cellular downlink transmissions is reduced. In a fast varying environment, the solution may lead to poor performance, especially for the URLLC services as the allocation does not consider CSI availability and QoS requirements.

With CSI availability, more optimal allocation can be achieved, e.g., in [19] where exclusive spectrum allocation for D2D usage is obtained in uplink spectrum. Similar spectrum allocation problem for D2D usage can be formulated, and therefore separate allocations can be achieved both in uplink and downlink. Further improvement can be made, if the allocation is done based on 'sector' of the cell.

Next, we discuss the URLLC-optimized allocation that considers CSI availability and QoS requirements, and improves the rate availability in the fast fading environments.

B. URLLC-optimized allocation

Network is interested in increasing availability in the ultra-low capacity region, e.g., BS can aim to increase the minimal rate, 0.001 % Cumulative Distribution Function (CDF) rate, etc. For this, BS requires the channel estimate of different UEs. Let us first consider a single-cell network where each UE including D2D UEs report their Signal-to-Interference-plus-Noise Ratio (SINR) to the BS. The BS then estimates rates on scheduled TTIs in the certain time window for the channel allocation for both cellular and D2D UEs. Denoting the rate set for cellular UEs as $\{r_{\mathrm{C2C}}\}_i$ for $i = 1, \cdots, N_{\mathrm{C2C}}$, and for D2D UEs as $\{r_{\mathrm{D2D}}\}_j$ for $j = 1, \cdots, N_{\mathrm{D2D}}$. BS allocates channels for D2D communication in uplink and downlink, i.e., $W_{\mathrm{D2D}}^{\mathrm{UL}}$ and $W_{\mathrm{D2D}}^{\mathrm{DL}}$ in a way causing the minimum rate UEs in both cellular and D2D communication to have the same rates. This results in increment of minimal network rate. The optimization problem is then formulated as

$$\begin{aligned} \underset{W_{\mathrm{C}}^{\mathrm{UL}}, W_{\mathrm{C}}^{\mathrm{DL}}, W_{\mathrm{D2D}}^{\mathrm{UL}}, W_{\mathrm{D2D}}^{\mathrm{DL}}}{\text{maximize :}} \quad & \text{minimum } \{\{r_{\mathrm{C}}\}_i, \{r_{\mathrm{D2D}}\}_j\} \\ \text{subject to :} \quad & W = W_{\mathrm{C}}^{\mathrm{UL}} + W_{\mathrm{D2D}}^{\mathrm{UL}} \\ & W = W_{\mathrm{C}}^{\mathrm{DL}} + W_{\mathrm{D2D}}^{\mathrm{DL}} \\ & i = 1, \cdots, N_{\mathrm{C}} \\ & j = 1, \cdots, N_{\mathrm{D2D}}. \end{aligned} \tag{8}$$

Given the isolated cell, for UE i, the achievable rate is

$$r_{\mathrm{C},i} = \min\left(r_{\mathrm{C},i}^{\mathrm{UL}}, r_{\mathrm{C},i}^{\mathrm{DL}}\right)$$

where $r_{\mathrm{C},i}^{\mathrm{UL}}$ and $r_{\mathrm{C},i}^{\mathrm{DL}}$ are uplink and downlink rates for UE pair i over the channels $W_{\mathrm{C}}^{\mathrm{UL}}$ and $W_{\mathrm{C}}^{\mathrm{DL}}$ respectively. Let us assume the minimal achievable rate in cellular transmissions case is $r_{\mathrm{C},k}^{\min}$ for some UE k. Given the TTIs allocation in both uplink and downlink, let us denote by $\rho_{\mathrm{C},k}^{\mathrm{UL}}$ and $\rho_{\mathrm{C},k}^{\mathrm{DL}}$ the uplink and downlink spectral efficiencies for UE pair k. The spectral efficiency is rate per unit of bandwidth, and it is expressed as

$$\rho = r/b. \tag{9}$$

From Eq. (9), for UE pair k, its uplink and downlink spectral efficiencies are related as

$$r_{\mathrm{C},k}^{\min} = \rho_{\mathrm{C},k}^{\mathrm{UL}} W_{\mathrm{C}}^{\mathrm{UL}} \approx \rho_{\mathrm{C},k}^{\mathrm{DL}} W_{\mathrm{C}}^{\mathrm{DL}}. \tag{10}$$

Similarly, for some minimum rate serviced UE l with D2D communication, its rate is

$$r_{\mathrm{D2D},l}^{\min} = \rho_{\mathrm{D2D},l}\left(W_{\mathrm{C2C}}^{\mathrm{UL}} + W_{\mathrm{C2C}}^{\mathrm{DL}}\right) \tag{11}$$

where $\rho_{\mathrm{D2D},l}$ the spectral efficiency estimate of UE l. BS obtains the channel allocation by causing minimum rates to become equal, i.e.,

$$r_{\mathrm{C2C},k}^{\min} = r_{\mathrm{D2D},l}^{\min}. \tag{12}$$

From Eq. (10)-(12), we obtain the resultant allocation as

$$\begin{aligned} W_{\mathrm{C}}^{\mathrm{UL}} &= \frac{2W}{\rho_{\mathrm{C},k}^{\mathrm{DL}}\left(\frac{1}{\rho_{\mathrm{C},k}^{\mathrm{UL}}} + \frac{1}{\rho_{\mathrm{C},k}^{\mathrm{DL}}} + \frac{1}{\rho_{\mathrm{D2D},l}}\right)}, \\ W_{\mathrm{C}}^{\mathrm{DL}} &= \frac{2W}{\rho_{\mathrm{C},k}^{\mathrm{UL}}\left(\frac{1}{\rho_{\mathrm{C},k}^{\mathrm{UL}}} + \frac{1}{\rho_{\mathrm{C},k}^{\mathrm{DL}}} + \frac{1}{\rho_{\mathrm{D2D},l}}\right)}, \\ W_{\mathrm{D2D}}^{\mathrm{UL}} &= W - W_{\mathrm{C}}^{\mathrm{UL}}, \\ W_{\mathrm{D2D}}^{\mathrm{DL}} &= W - W_{\mathrm{C}}^{\mathrm{DL}}. \end{aligned} \tag{13}$$

In a multi-cell network, where load is scattered over multiple cells, the formulation described in Eq. (8) cannot be applied. Here each cell independently allocates the resource in way to maximize the minimum of achievable rates over the different links, i.e.,

$$\begin{aligned} \underset{W_{\mathrm{C}}^{\mathrm{UL}}, W_{\mathrm{C}}^{\mathrm{DL}}, W_{\mathrm{D2D}}^{\mathrm{UL}}, W_{\mathrm{D2D}}^{\mathrm{DL}}}{\text{maximize :}} \quad & \text{minimum } \{\{r_{\mathrm{C}}^{\mathrm{UL}}\}_p, \{r_{\mathrm{C}}^{\mathrm{DL}}\}_q, \{r_{\mathrm{D2D}}\}_j\} \\ \text{subject to :} \quad & W = W_{\mathrm{C}}^{\mathrm{UL}} + W_{\mathrm{D2D}}^{\mathrm{UL}} \\ & W = W_{\mathrm{C}}^{\mathrm{DL}} + W_{\mathrm{D2D}}^{\mathrm{DL}} \\ & \{p, q\} = 1, \cdots, N_{\mathrm{C}} \\ & j = 1, \cdots, N_{\mathrm{D2D}}. \end{aligned}$$

The solution to above optimization problem is that the minimum link rates in all three transmission types (cellular uplink, downlink and D2D) are made equal, i.e.,

$$r_{\mathrm{C},m}^{\mathrm{UL}} \approx r_{\mathrm{C},n}^{\mathrm{DL}} \approx r_{\mathrm{D2D},l},$$

where index m represents UE with minimum cellular uplink rate, n with the UE with minimum downlink rate and l with the UE with minimum D2D transmission rate. The allocation is obtained as

$$\begin{aligned} W_{\mathrm{C}}^{\mathrm{UL}} &= \frac{2W}{\left(1+\frac{\rho_{\mathrm{C},m}^{\mathrm{UL}}}{\rho_{\mathrm{C},n}^{\mathrm{DL}}}+\frac{\rho_{\mathrm{C},m}^{\mathrm{UL}}}{\rho_{\mathrm{D2D},l}}\right)}, \\ W_{\mathrm{C}}^{\mathrm{DL}} &= \frac{2W}{\left(1+\frac{\rho_{\mathrm{C},n}^{\mathrm{DL}}}{\rho_{\mathrm{C},m}^{\mathrm{UL}}}+\frac{\rho_{\mathrm{C},n}^{\mathrm{DL}}}{\rho_{\mathrm{D2D},l}}\right)}, \\ W_{\mathrm{D2D}}^{\mathrm{UL}} &= W - W_{\mathrm{C}}^{\mathrm{UL}}, \\ W_{\mathrm{D2D}}^{\mathrm{DL}} &= W - W_{\mathrm{C}}^{\mathrm{DL}}. \end{aligned} \tag{14}$$

The optimization requires BS capability to perform rate estimation based on the SINR reporting. The improvement in the URLLC availability is at the expense of complexity in the estimation. For this, at regular intervals comparative to channel coherence time, BS broadcasts its request to all active UEs to report their SINRs. After receiving their SINR status, BS estimates the transmission rates, and obtains the allocation for URLLC services. For simplicity, we do not consider errors related to SINR estimation, channel aging due to time variation and hardware imperfectness of FDD communications.

The scope of the work can be extended to D2D pairs of other BSs (perhaps of another operator) visiting in the deep coverage of this BS. In a full re-use network (for small cells), BSs can coordinate on X2 interface and offload D2D UEs. Neighboring BSs can benefit with this coordination, otherwise interference from opponent D2D transmission would be highly damaging. BS schedule the D2D pairs (own and others) in a time domain fashion. For a deep coverage visiting cellular UE of another BS, UE can be connected to macro BS instead of offloading to this small cell BS. Macro BSs usually share spectrum in an orthogonal manner.

IV. Numerical Results

To assess the performance of the allocation protocols for increasing availability, we consider multi-user, multi-cell factory deployment scenario. The scenario, factory channel model and numerical parameters are depicted in the previous work [12]. In addition, we consider D2D communication usage with maximum allowable D2D UEs' distance of 25 m and LOS channel model reported in [20].

First, we present the CDF of UE rate where the spectrum resource between D2D and cellular UEs are allocated on a fast basis as shown in Fig. 3. The cellular rate curve depicts rate distribution with cellular communication-mode, and no D2D transmissions are allowed. With the possibility of D2D transmissions, three rate curves are depicted. Two are with the static allocation for D2D resource where each BS reserves 30 % and 50 % of the spectrum resource both in uplink and downlink for D2D communication use. With different static allocations, the lower tails demonstrate no gains. It appears that D2D UEs enjoy better rates due to the increased allocation of the resources using random static assignment, and the lower tail with cellular UEs suffer more. On the other hand, the flexible allocation shows better performance, especially in the middle and the lower CDF tail in comparison to both static allocations and alone cellular communication case. Here the resources are allocated proportional to the transmission types. The gains at 10 %, 70 % and 90 % of the CDF with respect to cellular communication are 7 %, 17 % and 24 % respectively.

Fast allocation although improves the performance but

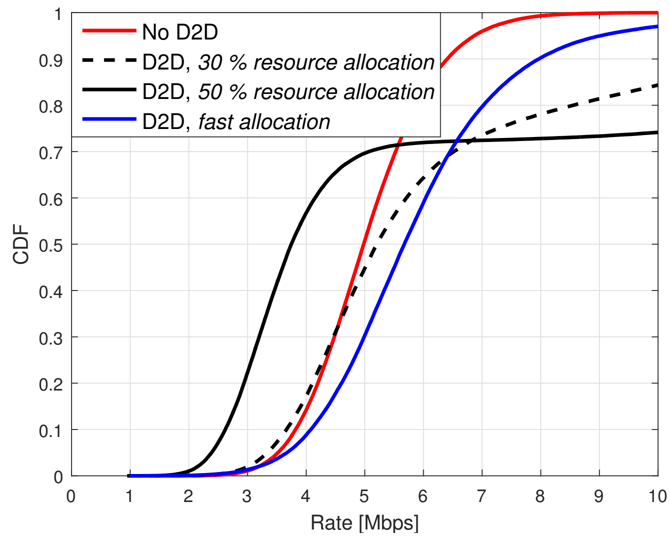

Fig. 3. UE rate CDF with no, static (30 % and 50 %) and fast allocation.

the improvement is negligible in ultra-low capacity region. To achieve better availability in ultra-low capacity region, we consider URLLC-optimized allocation protocol. To do so, first, we consider communication with no D2D support. We collect the rate statistics of uniformly distributed UEs in the network collected over 100 000 network instances, and obtain the rate available to 99.999 % of the UEs (i.e., the 0.001 % point of the UE rate distribution), as with shown in Fig. 4. The URLLC rate is scenario dependent, e.g., on channel model, traffic distribution, BSs availability, etc.

Next, we strive to maximize the URLLC rate, i.e., the rate available to 99.999 % of the UEs, by selectively applying D2D communication use. We restrict the D2D communication usage to the statistics (time units) for the worst UE population. These UEs improve their achievable rates at the expense of better serviced UEs through the flexible allocation of resources to D2D and cellular communication. The targeted population must be chosen suitably above the 0.001 % worst UEs (i.e., $100 - 99.999$ %). Selecting a larger proportion may lead to decrease in gains as the resources are shared amongst the

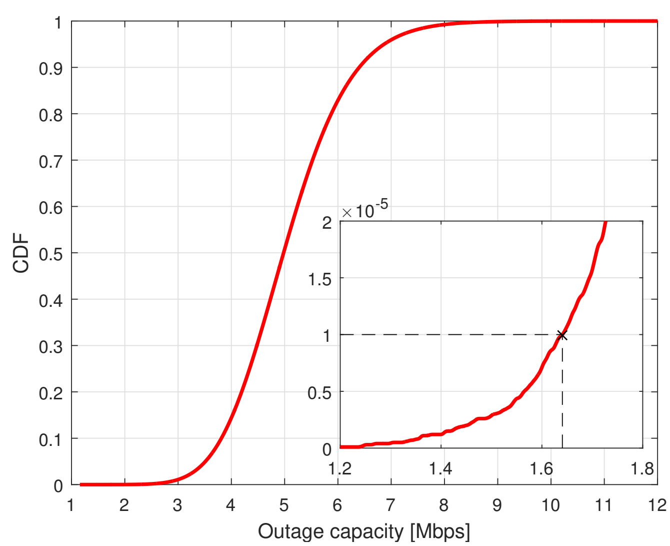

Fig. 4. Capacity distribution of the UEs in the factory scenario. Zoom-in to the URLLC region around 99.999 % reliability.

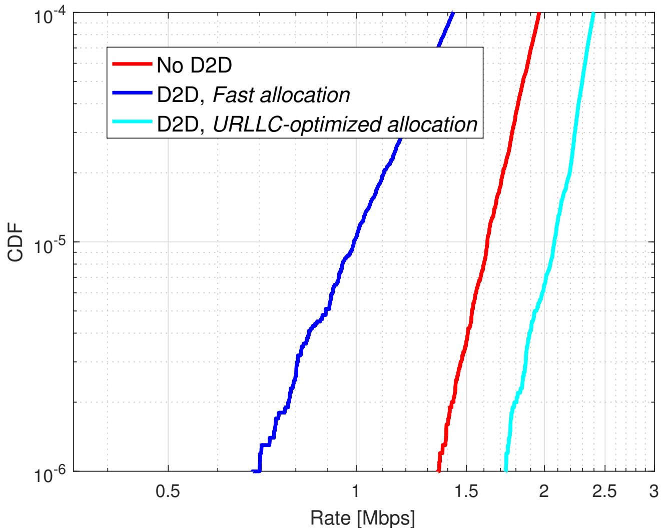

Fig. 5. UE rate CDF with default (no D2D communication allowed) and URLLC-optimized allocation.

bigger population of the worst UEs. In the current simulation studies, we select a random figure of 0.1 % of the worst UEs population. However, a separate problem can be formulated by which the maximum proportion can be deduced in order to achieve maximal gains, which is left for future studies. In every network instance, BS estimate rates based on reported SINR by the UEs and thus, allocate the resource according to the protocol. The gains are depicted in Fig. 5, presenting the lower tail of the CDF of the resulting UE rate distribution and the comparison is made to the rates achieved with no D2D communication availability. Fig. 5 shows 0.001 % outage capacity improvement using the flexible allocation for D2D communication and the gain is 30 %. The UEs in outage boost their rates by choosing new routes in the form D2D communication with less shadowing or utilizing more resources at the expense of high rate D2D UEs.

V. Conclusion

Flexible resource allocation schemes for cellular-device-to-device network in frequency-division duplex system are proposed. Through flexible allocation of cellular resource in both uplink and downlink spectrum for device-to-device resource usage, the rate availability improvement can be realized. Two schemes have been described, fast and ultra-reliable and low latency communications-optimized allocation. Fast allocation requires no channel state information or feedback. It distributes resources between cellular and device-to-device transmissions in accordance to their traffic load. The allocation is well suited for medium-capacity region improvement.

In ultra-reliable and low latency communications-optimized allocation, the objective is the availability improvement in a ultra-low capacity region. We selectively apply the flexible resource allocation to the worst 0.1 % of the users and improve the rate availability at the ultra-reliable communication point (i.e. of 99.999 % of the users). It is demonstrated that the gains may be realized where users in outage boost their rates either by choosing new routes via D2D with less shadowing or by allocating more cellular resources at the expense of other users with better D2D routes. By analyzing end-to-end system performance, the achievable data rate are explicitly identified. The allocation benefits ultra-reliable communication but exhibits complexity-gain trade-off.

VI. Acknowledgment

This work was supported in part by the Finnish Funding Agency for Innovation (TEKES) under the project "Wireless for Verticals (WIVE)". WIVE is a part of 5G Test Network Finland (5GTNF).

References

[1] G. Fettweis *et al.*, "The tactile internet," ITU-T, Tech. Watch Rep., Aug. 2014.

[2] 3GPP, "Study on scenarios and requirements for next generation access technologies," 3GPP TR 38.913 V0.3.0, Tech. Rep., Mar. 2016.

[3] B. Singh *et al.*, "Ultra-reliable communication in a factory environment for 5G wireless networks: Link level and deployment study," in *Proc. IEEE PIMRC Workshop*, pp. 1-5, Sep. 2016.

[4] B. Singh *et al.*, "Resource allocation strategy for ultra-reliable communication in a factory environment," in *Proc. GWS*, Nov. 2016.

[5] K. Adachi *et al.*, "Adaptive coordinated napping (CoNap) for energy saving in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5656-5667, Nov. 2013.

[6] H. Ishii, Y. Kishiyama and H. Takahashi, "A novel architecture for LTE-B: C-plane/U-plane split and phantom cell concept," in *Proc. IEEE GLOBECOM Workshop*, pp. 624-630, Dec. 2012

[7] C. Kosta *et al.*, "On interference avoidance through inter-cell interference coordination (ICIC) based on OFDMA mobile systems," *Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 973-995, 2013.

[8] L. B. Le *et al.*, "QoS-aware and energy-efficient resource management in OFDMA femtocells," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 180-194, Jan. 2013.

[9] B. Singh, K. Koufos and O. Tirkkonen, "Co-primary inter-operator spectrum sharing using repeated games," in *Proc. IEEE ICCS*, pp. 67-71, Nov. 2014.

[10] B. Singh *et al.*, "Contention-based access for ultra-reliable low latency uplink transmissions," *IEEE Wireless Commun. Lett.*, 2017.

[11] M. Simsek *et al.*, "5G-enabled tactile internet," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 460-473, Mar. 2016.

[12] B. Singh *et al.*, "Selective multi-hop relaying for ultra-reliable communication in a factory environment," in *Proc. IEEE PIMRC*, pp. 1-6, Sep. 2016.

[13] 3GPP, "Study on LTE device to device proximity services; Radio aspects," 3GPP TR 36.843 V12.0.1, Tech. Rep., Mar. 2014.

[14] P. E. Omiyi and H. Haas, "Maximising spectral efficiency in 3G with hybrid ad-hoc UTRA TDD/UTRA FDD cellular mobile communications, in *Proc. IEEE ISSSTA*, pp. 613-617, Aug. 2004, pp. 613617.

[15] A. Osseiran, J. F. Monserrat and P. Marsch, "5G mobile and wireless communication technology," Cambridge, UK: Cambridge University Press, Jun. 2016.

[16] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *Proc. IEEE VTC Spring*, pp. 1-6, May 2014.

[17] Nokia and Alcatel-Lucent Shanghai Bell, "Flexible numerology for 5G," 3GPP TSG-RAN WG1 Meeting #85, May 2016.

[18] B. Soret *et al.*, "Fast muting adaptation for LTE-A HetNets with remote radio heads," in *Proc. IEEE GLOBECOM*, pp. 3790-3795, Dec. 2013.

[19] K. Huang *et al.*, "Overlaid cellular and mobile ad hoc networks," in *Proc. IEEE ICCS*, pp. 1560-1564, Nov. 2008.

[20] E. Tanghe *et al.*, "The industrial indoor channel: large-scale and temporal fading at 900, 2400, and 5200 MHz," *IEEE Trans. Commun.*, vol. 7, no. 7, pp. 2740-2751, Jul. 2008.