Single Frequency-based Device-to-Device-enhanced Video Delivery for Evolved Multimedia Broadcast and Multicast Services

Leonardo Militano†, Massimo Condoluci†, Giuseppe Araniti†, Antonella Molinaro†, Antonio Iera†, Gabriel-Miro Muntean‡

†University Mediterranea of Reggio Calabria, Italy, DIIES Department, email: *[leonardo.militano|massimo.condoluci|araniti|antonella.molinaro|antonio.iera]@unirc.it*

‡School of Electronic Engineering, Dublin City University, Ireland, email: *gabriel.muntean@dcu.ie*

Abstract

Despite of the undisputed benefits of the Long Term Evolution-Advanced (LTE-A) networks, offering support for group-oriented services challenges the evolved Multimedia Broadcast Multicast Services (eMBMS) design in LTE-A. This is especially important when delivering video content with high bitrate requirements. The Conventional Multicast Scheme (CMS) is proposed as a radio resource allocation solution for eMBMS to serve all multicast group members with the data rate supported by the receiver with the worst channel conditions. In this paper, we propose a novel Radio Resource Management approach, the Device-to-Device (D2D)-enhanced CMS with Single Frequency (D2D-SF). This proposal extends the CMS with additional D2D communications in order to increase the aggregate data rate of the cell, while also maintaining the typical CMS short-term fairness. D2D-SF makes use of one or more mobile subscriber devices as forwarding devices (FD) to retransmit the data received from the base station (BS) over direct local links to other members of the multicast group. The proposed solution supports both high-rate modulation and coding schemes on the downlink from BS to FDs, and reaches cell-edge devices (hence, experiencing worse channel conditions) through high-performing D2D links (improving this experience). Testing shows how the single frequency-based D2D CMS paradigm proposed, complemented by two novel strategies for selecting FDs, achieves significant enhancements of the overall performance when delivering video content compared to both the state-of-the-art multicast solutions and novel solutions that do not employ a single-frequency paradigm.

Index Terms

LTE-Advanced, Multicasting, Device-to-Device communications, Radio Resource Management, Video delivery

I. Introduction

The growing demand for group-oriented services has led to the definition of new standards and applications for the mobile market. LTE-A [1] is the most promising wireless system to support such services with significant benefits for users and network. For instance, it guarantees higher data rates in both downlink and uplink directions, effective Quality of Service (QoS) management, high spectrum efficiency and increased system capacity. These aspects are very important for the delivery of many service types, including video. Delivery of video content is one of the fastest growing services, especially over wireless and mobile networks. According to the Cisco Visual Networking Index survey made available in September 2014, by 2018 the video traffic carried by mobile networks will be 8 times larger than it is in 2014 [2], reaching 11 exabytes [3]. Given the high interest for video transmissions towards multiple destinations, the 3rd Generation Partnership Project (3GPP) defined a solution to deliver multicast and broadcast services over cellular networks namely the Multimedia Broadcast Multicast Service (MBMS) [4]. From 3GPP's Release 8, MBMS has been extended to the Long Term Evolution (LTE) standard and the extension is called evolved MBMS (eMBMS) [5].

Currently, several research and industrial organizations are actively studying solutions to best handle the increased traffic and the dissimilar channel quality experienced by users in the same multicast group and manage the available resources [6]. A basic solution to the radio resource allocation problem in eMBMS networks is offered by the Conventional Multicast Scheme (CMS) [6], which serves all multicast users in a cell at every transmission time interval (TTI), by constraining the data rate to the user with the worst channel conditions (typically, at the cell edge). This choice translates into poor performance in terms of data rate and low satisfaction levels for users with good channel situations. An alternative approach is offered by the Opportunistic

Multicast Scheme (OMS) [7], which only serves users with the best channel conditions in each time interval. This allows increasing the network data rate, but short-term fairness is no more guaranteed to users and this may affect the delivery of time sensitive services such as video to some users.

In our opinion, multicast delivery schemes in eMBMS networks can benefit from D2D communications. For instance, neighbouring user devices belonging to the same multicast group can activate direct links by using cellular radio resources [8], [9], to cope with adverse cell-edge effects. The use of D2D links can be substantially more efficient than conventional delivery through a Base Station (BS) whenever a communication is inherently local in scope [10], [11]; besides, it can help to either extend the cell coverage, to offload cellular traffic [12], [13], or to support content sharing in a neighbourhood [14], [15].

This paper brings the D2D communications to the multicast communication framework to complement CMS radio resource allocation in order to address some of the limitations of the latter. We retain the CMS's philosophy of serving all the multicast group users in a cell at every TTI, but we release the constraint that all the users are to be served *directly by the BS*. D2D communications are employed to reach the users in the group with poor channel conditions. Mobile devices belonging to same multicast group are clustered around one or more *forwarding devices* (FD) that receive data directly from the BS and forward it to their cluster members. Following careful selection of the FDs, BS can use high-performing modulation and coding schemes (MCS) for data transmission to the FDs, and high-quality D2D from FDs to the nodes with worse channel conditions.

This paper proposes the D2D-enhanced CMS with Single Frequency (D2D-SF) by combining in an innovative manner CMS-based and D2D content delivery in order to increase the aggregate data rate of the cell, while also maintaining the short-term fairness between devices. We assume that the D2D links exploit uplink frequencies, as suggested in [16] and all the FDs in the same cell simultaneously use *the same frequency* to deliver multicast data over the D2D links, as described in [1], [17]. The receivers consider these retransmissions as multipath components of the same signal. D2D-SF is introduced in conjunction with two novel strategies for selecting FDs, based on clustering of the devices: Basic Cluster Formation and Enhanced Cluster Formation. Extensive simulation-based testing has been performed which shows how significant enhancements of the overall performance when delivering video content was recorded compared to both state-of-the-art multicast solutions and novel solutions that do not employ a single-frequency paradigm.

The remainder of the paper is organized as follows. In Section II the related works are discussed and in Section III the reference system model and service configuration are described. The proposed D2D-SF multicast radio resource management is described in Section IV, whereas the problem formulation is presented in details in Section V. The performance evaluation settings and the results are summarized in sections VI and VII respectively, whereas conclusive remarks are made in the last section.

II. Related Work

In the literature, several research contributions are available dealing with the use of only cellular transmissions by the base station to effectively handle the varying channel quality experienced by users in the same multicast group (MG) and to efficiently use the available resources. The Conventional Multicast Scheme (CMS) [6] adopts a conservative *single-rate* approach that selects the MCS for the multicast transmission according to the requirements of the user with the worst channel quality, although this introduces severe inefficiencies. In particular, the potentials of orthogonal frequency division multiple access (OFDMA) are not fully exploited [18], and the performance of the whole set of destinations decreases as the MG size increases [6].

To the same family of single-rate policies belongs the OMS [7] as in a given time slot, the base station feeds multicast users with only one single data rate. In particular, in a time slot OMS only serves the "best" subset of multicast members (i.e., those with the best channel conditions) to maximize their QoS. However, in different time slots, OMS can implement data rate differentiation for multicast users according to the selected transmission parameters. Diverse OMS-based approaches are proposed in LTE environments. Specifically, in [7] the authors propose different OMS algorithms that improve the total data rate by exploiting multi-user diversity; in [19] the best users are selected based on a signal to interference plus noise ratio (SINR) threshold, and the terminals that experience a SINR value below the threshold are not served at all. According to [20], multi-user diversity allows to guarantee a spectral efficiency equal to a pre-defined target value. In general, the price to pay for such data rate improvement in OMS-based solutions (e.g., in [7] and [20]) is a multicast gain reduction (i.e., the reduction of the number of users served in each time slot, the TTI in LTE). Moreover, as the portion of users served by the scheduler dynamically changes over the time, OMS-based solutions need to couple with rateless coding schemes [21], which introduce additional issues of computational burden, buffer size, decoding delay, and short-term fairness [22].

Alternative solutions to single-rate scheduling policies for multicast services have been also proposed. In particular, *multi-rate* approaches deal with the idea to simultaneously serve multicast users with different data rates by taking advantage of the heterogeneity of the channel quality measured by multicast group members. For example, multicast subgroup formation techniques [23], [24] split the multicast members into different subgroups and serve all of them in every scheduling frame at the best conditions allowed by their channel conditions. Similarly, subgroup-based policies can be found, in [25] and [26]. In the former, the authors propose a novel cost function for subgroup formation aiming at guaranteeing a trade-off between throughput and fairness, whereas in the latter a low-complexity subgrouping scheme has been proposed to reduce the complexity load and the scheduling execution time at the BS.

Recently, direct communication between devices has been considered for multicast service delivery to overcome the performance of above mentioned approaches where only cellular transmission by the BS were considered. Most of the conducted studies focus on direct device communications over short links of a different technology than the cellular one. For example, in [27] some mobile devices are selected as anchor points in a cell to forward the multicast data received from the BS to other devices in proximity through multi-hop ad-hoc Wi-Fi links. In [28] cellular users directly communicate to perform cooperative retransmissions using a generic short-range communication capabilities. Nevertheless, the use of heterogeneous wireless interfaces poses several issues in terms of content synchronization which becomes crucial when considering multicast video streaming applications. In addition, as also stated in [29], the use of cellular D2D links introduces several benefits compared to *outband* D2D links, like Wi-Fi, in terms of enhanced user throughput. For these reasons, differently from [27] and [28], in this paper we consider D2D communications over cellular LTE-A links.

In the reference scenario for this paper, a portion of multicast users (i.e., the devices with poor channel qualities) is split into clusters; the cluster members are served via cellular D2D transmissions, whereas the remaining users (i.e., those with better channel quality) are served over cellular transmission from the BS. The D2D-based clustering issues are differently approached in the literature. For instance, in [30] a group of nearby devices create a D2D cluster to share data with other cluster members; on the contrary in our proposal D2D clusters is specifically used to enhance the quality of a multicast service. In [31], similarly to our contribution, UEs are grouped into clusters wherein cluster heads send data to one or more interested devices through D2D communications. However, the focus in [31] is on data retransmissions, when some of the interested nodes did not correctly receive the data. This aspect is addressed also in [32], where the focus is on the resource allocation when in the presence of retransmission over D2D links. Differently, we do not use D2D links just for retransmissions, but consider them as additional means to allow the eNodeB to serve the multicast group as a whole in the most efficient and effective way.

The solutions proposed in the literature that are summarized above suffer in terms of several inefficiencies related to the resource allocation for D2D links. Indeed, as discussed in [29], the main issue deals with the number of resources needed by D2D transmitters to forward the data received from the BS. The works in literature are usually based on the assumption that D2D transmitters use different portions of resources to avoid inter-cluster interference. This aspect meaningfully limits the performance of multicast D2D-based solutions, since it influences the number of D2D transmitters that can be enabled and the cluster configurations that can be enabled. To overcome these issues, in this paper we propose a novel approach for multicast transmissions enhanced by cellular D2D transmissions, namely the single-frequency D2D paradigm. This idea is inspired by the Multicast Broadcast Single Frequency Network (MBSFN) technique [1], [17], [33] and is somewhat similar to the use of *gap filler* in Digital Video Broadcasting (DVB) systems [34]. In particular, in a MBSFN multiple BSs, tightly time-synchronized, simultaneously transmit the same signal over the same frequency to the multicast receivers in their cells. A receiver observes multiple delayed versions of the same signal and, through appropriate synchronization, channel estimation, and equalization techniques, benefits from the multipath diversity, at the only cost of a slight increased computational complexity [35], [18]. Analogously, in our proposal, multiple FDs in a cell simultaneously transmit the same signal received from the BS to their D2D-connected devices over a single uplink frequency. The receivers consider these replications as multipath components of the same signal.

The proposed technique can be applied when the following conditions hold: (a) all the multicast user equipment (UE) devices in the cell are interested in receiving the same content at the same time; (b) the FDs receive the same content from the eNodeB - the LTE BS - and transmit this identically over all the D2D links at the same time, following the eNodeB synchronization. In particular, the eNodeB will: *(i)* select the most suitable MCS for the cellular mode, under the constraint of serving all UEs in the multicast group; *(ii)* identify the number of D2D clusters and the devices to be elected as FDs; *(iii)* identify the best configuration and transmission parameters for the D2D links; *(iv)* execute the resource allocation algorithm on the activated links (eNodeB-to-FD and FD-to-UEs).

To summarize, the main contributions of this paper, that address the limitations of the related work, consist in *(1)* considering the joint use of cellular and D2D modes in LTE-A networks for multicast data delivery; *(2)* defining efficient resource allocation strategies at eNodeB to maximize the aggregate data rate; *(3)* introducing a single-frequency paradigm for D2D-based multicast delivery from the FDs in the cell; *(4)* investigating on possible policies to cluster the D2D-served nodes; *(5)* investigating on the parameters influencing the performance in a wide set of scenarios involving video streaming and video downloading applications.

III. Reference System and Background

In LTE-A systems [1], OFDMA and single carrier frequency division multiple access (SC-FDMA) are used to access the downlink and the uplink, respectively. The available radio spectrum is managed in terms of *resource blocks* (RBs) and, in the *frequency domain*, each RB corresponds to 12 consecutive and equally spaced sub-carriers. One RB is the smallest frequency resource that can be assigned to a UE. The overall number of available RBs depends on the system bandwidth configuration and can vary between 6 (1.4 MHz channel bandwidth) and 100 (20 MHz).

Fig. 1 illustrates the system architecture considered in this paper. This architecture extends the eMBMS standard architecture defined in [4] in order to support D2D-based data communication for efficient video delivery to users in a cell. The classic eMBMS architecture for the access network is composed of eNodeBs, which are the evolved network nodes which communicate directly with UE and a MultiCell/Multicast Coordination Entity (MCE), responsible for transmission parameter configuration in single- and multi-cell mode, respectively. The core network includes: Mobility Management Entity (MME) that is responsible for authentication, security, and mobility management procedures, MBMS Gateway (MBMS-GW), a logical entity whose principal function is data packet forwarding to eNodeBs and Broadcast Multicast-Service Center

(BM-SC) that is the MBMS traffic source, which also accomplishes service announcement and group membership functions. The eNodeB manages the spectrum, by assigning the adequate number of RBs to each scheduled user and by selecting the MCS for each RB. Scheduling procedures are based on the *channel quality indicator* (CQI) feedback, transmitted by each UE to the eNodeB over dedicated control channels. The CQI is associated to the maximum supported MCS [1], as reported in Table I for the LTE-A standard. Transmission parameters (i.e., MCSs) are adapted at every *CQI feedback cycle* (CFC), which can last one or several TTIs (one TTI is equal to 1 ms) [1].

TABLE I
CQI-MCS MAPPING FOR D2D AND CELLULAR COMMUNICATION LINKS

CQI index	Modulation Scheme	Efficiency D2D [bit/s/Hz]	Minimum Rate D2D [kbps]	Efficiency Cellular [bit/s/Hz]	Minimum Rate Cellular [kbps]
1	QPSK	0.1667	28.00	0.1523	25.59
2	QPSK	0.2222	37.33	0.2344	39.38
3	QPSK	0.3333	56.00	0.3770	63.34
4	QPSK	0.6667	112.00	0.6016	101.07
5	QPSK	1.0000	168.00	0.8770	147.34
6	QPSK	1.2000	201.60	1.1758	197.53
7	16-QAM	1.3333	224.00	1.4766	248.07
8	16-QAM	2.0000	336.00	1.9141	321.57
9	16-QAM	2.4000	403.20	2.4063	404.26
10	64-QAM	3.0000	504.00	2.7305	458.72
11	64-QAM	3.0000	504.00	3.3223	558.72
12	64-QAM	3.6000	604.80	3.9023	655.59
13	64-QAM	4.5000	756.00	4.5234	759.93
14	64-QAM	5.0000	840.00	5.1152	859.35
15	64-QAM	5.5000	924.00	5.5547	933.19

A user device in a LTE-A network can either communicate through the serving eNodeB (*cellular mode*) or it can bypass the eNodeB and use direct communications over D2D links (*D2D mode*). The eNodeB is in charge of the D2D session setup (e.g., bearer setup) [9], while power control and resource allocation procedures on the D2D links can be executed either in a distributed or in a centralized way [8]. In this paper we assume that the centralized approach is implemented. Accordingly, the eNodeB is aware of the cell load and the user channel conditions and can efficiently allocate dedicated resources to D2D connections so to improve the session quality and the allocation flexibility. We assume that uplink resources are allocated to D2D communications because *(i)* uplink guarantees a more efficient resources reusing compared to

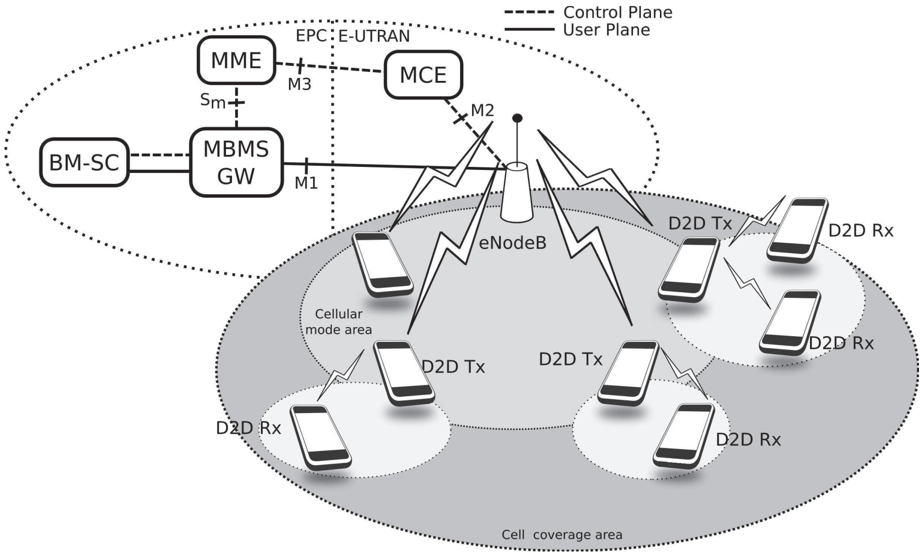


Fig. 1. D2D-enhanced Conventional Multicast Scheme.

downlink, in the worst case of a fully loaded cellular network, as demonstrated in [16], and *(ii)* the use of uplink resources gives the possibility of freeing downlink resources to use for other services within the cell.

D2D connections can be supported on frequency division duplex (FDD) and time division duplex (TDD) bands. The FDD mode poses additional issues in terms of terminal design, cost and complexity [8]; for this reason, we consider TDD, by referring to the frame structure *type 2* foreseen by 3GPP [1] and *configuration 1* which guarantees an equal number of downlink and uplink slots over the frame. The whole radio frame lasts 10 ms and consists of ten sub-frames of 1 ms each, where special fields are used for switching between downlink and uplink transmissions. The communication range between nearby devices can reach tens of meters [36], but the data rate on the D2D link depends on the CQI level and the allocated resources as reported in Table I. In particular, the CQI-MCS mapping for a D2D link can be found in [36] and the values are assumed to be equal to those in a femtocell since the same transmission power is used [37].

IV. The D2D-enhanced Multicast Video Delivery

The reference service scenario for this paper is illustrated in Fig. 1. Let us consider a group of UEs is interested in the same multicast content served by a single LTE-A cell, for instance students on-campus who are accessing a video content of common interest. Under this condition a multicast video delivery can be activated that is able to exploit the enhancements offered by D2D communications among the involved devices. In the remaining of this section we will introduce the details on the system model and the proposed Radio Resource Management (RRM) for the proposed Device-to-Device-enhanced Conventional Multicast Scheme for Video Delivery in LTE-A Systems.

A. System model

In the considered LTE single-cell area, a set of users, denoted by $\mathcal{K}$, is associated to the same MG. The eNodeB performs link adaptation procedures on both cellular and D2D links by handling N available RBs according to the CQI feedbacks collected from each user. Let C be the number of available CQI levels and let $c_k \in \{1, 2, \ldots, C\}$ be the CQI reported by multicast member k, with $k \in \mathcal{K}$. Moreover, let $c_{k,j}$ be the CQI value for each D2D link between nodes $k, j \in \mathcal{K},\ k \neq j$. Each CQI level is associated to a given supported MCS. For a given MCS value m, the attainable data rate depends on the number of assigned RBs and on the spectral efficiency for the given MCS, b_m expressed in bit/s/Hz as reported in Table I. Hence, we denote with b_m^{dl} and b_m^{ul} (where $m = 1, \ldots, C$) the spectral efficiency respectively in downlink and uplink transmissions. Moreover, we represent with $f^{dl}(m, n_m)$ and $f^{ul}(m, n_m)$ the data rate respectively in downlink and uplink transmissions adopting the MCS associated to the CQI m, as a function of m and the assigned RBs n_m[1].

The proposed radio resource management (RRM) scheme is in charge of deciding which multicast configuration to enable, by this meaning: *(i)* the set of UEs directly served by the eNodeB in downlink, *(ii)* the MCS for the downlink transmission, *(iii)* the cluster configuration for D2D relaying, and *(iv)* the resource allocation and the MCS selection for the transmissions of each activated FD.

B. Service configuration

The eNodeB executes the following steps when the service delivery starts. In particular, a single execution of the listed steps is executed. However, when significant variations in the

[1]The admissible throughput values per MCS level are set according to Table 7.1.7.2.1-1 in [37]

channel conditions are observed (e.g., due to UEs' mobility), these steps should be repeated to update the service configuration.

1) *Service registration*: The eNodeB advertises the multicast service and all interested UEs within the cell join this service to form a single MG.
2) *CQI collection*: The eNodeB collects the CQI feedbacks from all UEs belonging to the MG, i.e., $c_k \ \forall k \in \mathcal{K}$.
3) *D2D CQI collection*: The eNodeB collects the $c_{k,j}$ values from all UEs $k, j \in \mathcal{K},\ k \neq j$ belonging to the MG. This information will be used to discover the UEs potentially reachable through D2D links by selected FDs in the MG.
4) *FD selection and cluster formation*: Being $\tilde{C} \subseteq \{1, 2, \ldots, C\}$ the set of CQI levels in downlink for the UEs in the MG, for each $m \in \tilde{C}$ the eNodeB determines: *(i)* the set of UEs that can correctly decode data if served by the BS, i.e., $\mathcal{K}_m^{dl} = \{k \in \mathcal{K} | c_k \geq m\}$; *(ii)* the subset of served UEs $\mathcal{R}_m \subseteq \mathcal{K}_m^{dl}$, that can act as FDs[2]; *(iii)* the remaining UEs that are not served by the eNodeB, but can be served by a FD through D2D connections. To this aim, the eNodeB computes a D2D CQI matrix (DCM) (an example is reported in Table II) based on the $c_{k,j}$ values (where $k \in \mathcal{K}_m^{dl}$ and $j \in \mathcal{K} \setminus \mathcal{K}_m^{dl}$) for all the links between the potential FDs (the matrix rows) and the remaining nodes (the DCM columns). A $c_{k,j} = 0$ value in the DCM indicates that a D2D link cannot be activated between nodes k and j. According to the values in the DCM, the eNodeB will then select the subset of UEs $\mathcal{D}_{m,r} \subseteq \mathcal{K} \setminus \mathcal{K}_m^{dl}$ to be associated to each enabled FD $r \in \mathcal{R}_m$. This association of UEs to a given FD and the choice of the best number of FDs is based on the algorithms presented in Section V-A. Noteworthy, the proposed centralized scheme requires that the eNodeB is aware of the updated DCM, which causes some extra overhead. However, direct device communications are usually based on the assumption of stationary or, at least, semi-static D2D channels due to low mobility and short communication range in the local service scenarios [31]. For this reason, the rate at which the D2D channel conditions are updated can be very low; this implies a significant reduction in the cost of the D2D channel quality acquisition procedure and in the DCM computation. Moreover, as underlined in [31], the

[2]We assume that all the nodes belonging to $\mathcal{K}_m^{dl}$ are willing to act as FD. This assumption is well justified by the data rate improvement obtained by every device in the MG, as shown in the performance evaluation section. Further research related to the increased energy consumption for the D2D forwarding nodes is left for future studies.

overhead can be further reduced by using feedback compression schemes, such as best-M [38], delta compression [38], and DCT significant-M [39].

TABLE II
D2D CQI Matrix.

Downlink-served node \ Other nodes	node 1	node 2	node 3	...	node j
node 4	$c_{4,1}$	$c_{4,2}$	$c_{4,3}$	...	$c_{4,j}$
node 5	$c_{5,1}$	$c_{5,2}$	$c_{5,3}$	...	$c_{5,j}$
...	...	...	...	...	...
node k	$c_{k,1}$	$c_{k,2}$	$c_{k,3}$	...	$c_{k,j}$

5) *D2D link configuration*: For each CQI level $m \in \tilde{C}$ evaluated for downlink transmissions, the eNodeB computes the resource $N^{ul}_{m,r}$ and the MCS level $l_{m,r}$, to be used on the D2D link for each FD $r \in \mathcal{R}_m$. D2D links can be either unicast or multicast. A conservative approach is adopted in the multicast case; thus, the FD serves all UEs in the D2D cluster in a single transmission by using the MCS corresponding to the worst CQI value in the DCM, i.e., $l_{m,r} = \min_{k \in \mathcal{D}_{m,r}} \{c_{r,k}\}$ for FD r. This paper considers two alternative policies according to which the FDs handle the uplink frequencies to transmit data in their own D2D cluster. The first policy associates different resources to the different FDs; the second one (that we will demonstrate is a better choice) implements the novel single-frequency-based D2D paradigm, i.e., all the FDs use the same portion of resources (i.e., the same RBs). In the former case, disjoint sets of RBs are allocated to the D2D links (this means different amounts of resources). In the latter case the amount of resources allocated to the D2D links are constrained by the cluster with the lowest activated MCS (more details are given in section V-B). In general, devices connected on a D2D link are expected to be at a short distance and with good channel conditions, therefore they need a lower amount of resources compared to those needed for a direct cellular communication. This is however not always true as it depends on the node distribution in the cell and the eNodeB choices of the FDs.

6) *Multicast service activation and resource allocation*: Finally, the eNodeB selects the solution to activate, which is the one that maximizes the system data rate under the constraint that all the UEs in a MG are served, either through direct cellular links or through D2D

links. In particular, after the selection of the MCS level m^* to activate in downlink and of the corresponding $\mathcal{K}^{dl}_{m^*}$, $\mathcal{R}_{m^*}$, $\mathcal{D}_{m^*,r}$, $N^{ul}_{m^*,r}$, and $l_{m^*,r}$ values, the eNodeB allocates the available resources.

Fig. 2 shows the whole process and the steps to follow for managing the service. In particular, all values $m \in \tilde{C}$ are considered as potential CQI levels to activate in downlink. For each of the CQI levels a cluster formation algorithm is implemented to define a *configuration* of FDs and corresponding D2D clusters. If a given tested level is *eligible*, then the resulting data rate Ω_m is computed. A cluster configuration is considered *eligible* if the FDs are able to forward the total amount of bits received from the eNodeB over the D2D links to all users not served by the cellular link. This requires two conditions to be met: *(i)* the enabled FDs can successfully serve all the nodes belonging to $\mathcal{K} \setminus \mathcal{K}^{dl}_m$ via D2D links, and *(ii)* the N available resources are enough to relay all data to the D2D receivers. If instead, no cluster configuration for the tested CQI level m can be found, then the iteration on the $m \in \tilde{C}$ value is stopped and the final selection is performed. In particular, the iteration can be stopped since the tested CQI levels follow an order from the minimum to the maximum CQI value, and with higher values for the CQI level in downlink the probability of having an *eligible* configuration is reduced (the number of nodes not able to decode the data in downlink increases).

V. Radio Resource Management for D2D-enhanced Multicast Video

Fundamental steps in the implementation of the proposed RRM discussed in the previous section, are the FD selection and cluster formation (see step 4 in the RRM in Section IV) and the D2D link configuration with the radio resource allocation (see step 5 in the RRM in Section IV). The proposed policies for these two steps are detailed in the remaining of this section.

A. FD selection and cluster formation

Let us consider the generic iteration where the m-th CQI level is tested for downlink transmission. Given $\mathcal{K}^{dl}_m$, the set of UEs that can correctly decode the data according to the considered CQI, and based on the DCM, the eNodeB evaluates which nodes can potentially act as FDs for the remaining $\mathcal{K} \setminus \mathcal{K}^{dl}_m$ nodes. Based on this information the eNodeB can allocate the resources to each D2D link.

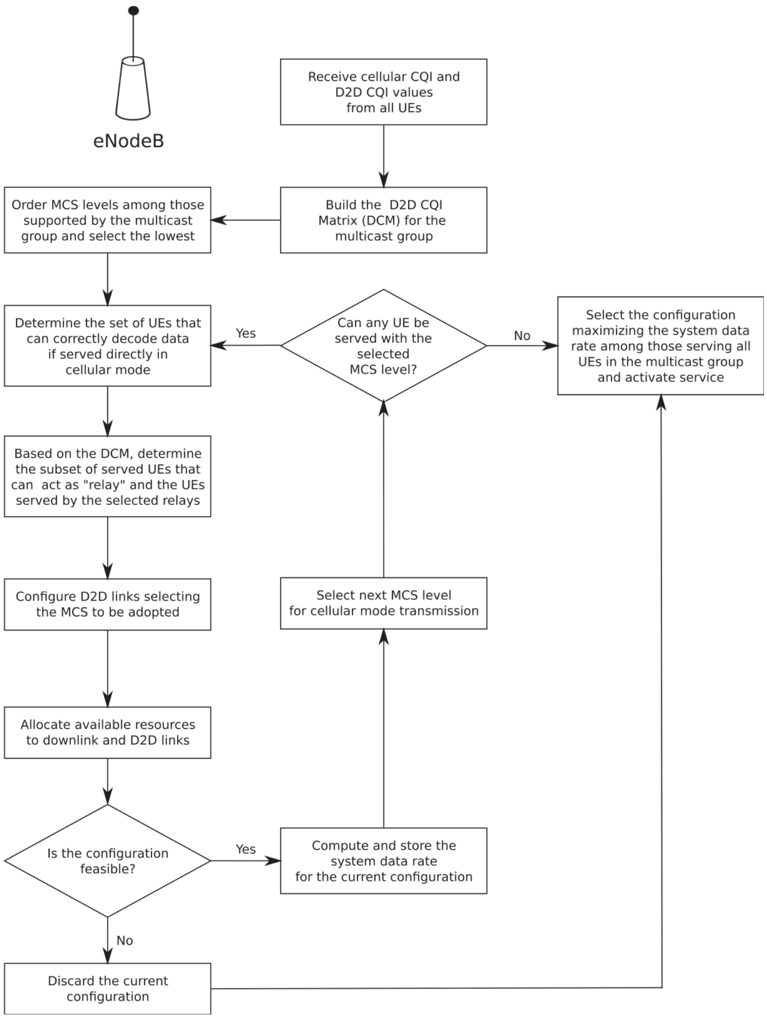


Fig. 2. Flow chart.

The number of cluster combinations to be tested can be very high as it increases exponentially with the size of $\mathcal{K}_m^{dl}$. In particular, this value is expected to be higher for lower CQI values, as more nodes are able to correctly decode the data sent from the eNodeB. An exhaustive search algorithm, whereby all combinations of FDs are tested, would cause unacceptable computational costs. However, it could be not necessary to test them all and however finding several combinations to forward the data in the cluster. Other considerations can be made to reduce the number of configurations to be tested. Two cluster formation strategies are proposed to keep the number of tested solutions low, while still finding a solution: the *Basic Cluster Formation (BCF)* and *Enhanced Cluster Formation (ECF)*.

1) Basic Cluster Formation (BCF): This policy is based on the idea that the eNodeB selects "the best" FD for each UE not served in downlink (as reported in line 2 in Algorithm 1). Specifically, the best FD for each node j belonging to $\mathcal{K} \setminus \mathcal{K}_m^{dl}$ is considered as the node $r \in \mathcal{K}_m^{dl}$ which guarantees the best D2D link conditions. In those cases where more than one FD can guarantee the same CQIs, the eNodeB selects the FD serving more users in order to limit the number of FDs.

Algorithm 1: Implementation of the proposed BCF policy

Data: m, $\mathcal{K}_m^{dl}$, N
Result: $\mathcal{R}_m$, $\mathcal{D}_{m,r}$, $N_{m,r}^{ul}$ and $l_{m,r}$
1 **for all** $j \in \mathcal{K} \setminus \mathcal{K}_m^{dl}$ **do**
2 $\quad r = \arg\max_{k \in \mathcal{K}_m^{dl}} \{c_{k,j}\}$
3 $\quad$ Update $\mathcal{R}_m$ with r
4 $\quad$ Update $\mathcal{D}_{m,r}$ with j
5 **end**
6 **if** $\bigcup_{r \in \mathcal{R}_m} \mathcal{D}_{m,r} \neq \mathcal{K} \setminus \mathcal{K}_m^{dl}$ **then**
7 $\quad$ Solution *not eligible*
8 **else**
9 $\quad$ Compute $N_{m,r}^{ul}$ and $l_{m,r}$
10 **end**

2) Enhanced Cluster Formation (ECF): The idea at the basis of the ECF policy is that if multiple *eligible* cluster configurations exist for a given CQI value m in downlink, then the one with the highest spectral efficiency (i.e., achievable data rate per used radio resource) on the D2D link is chosen. In doing this *(i)* fewer FDs are enabled, and *(ii)* the activated FDs use a lower number of RBs. Details of the proposed ECF policy are pointed out in Algorithm 2. In particular, ECF performs an iterative search for the solution. Since the search for an *eligible* cluster configuration may be processing intensive, countermeasures are adopted to consistently reduce the number of solutions to test:

1) given the value m, only those nodes in $\mathcal{K}_m^{dl}$ which guarantee the highest CQI value at least on one of the D2D links towards users not served in downlink are considered [3] (this is the scope of lines 1-13 in Algorithm 2). This is acceptable because the objective is to select the cluster configuration with the highest spectral efficiency. In fact, a node without the

[3]The simulative analysis showed that FDs usually are selected among those nodes having a close CQI level, not higher than $m + 3$. In fact, the nodes with CQI levels exceeding this threshold are too much close to the eNodeB, thus further away from the nodes the data is to be relayed to, making them less suitable to act as FDs.

mentioned feature, will never be chosen as there is at least another node in $\mathcal{K}_m^{dl}$ performing better for each of the nodes to be served in D2D (note that a special case of this is a node in $\mathcal{K}_m^{dl}$ having all values in the DCM equal to zero);

2) for a downlink CQI level equal to m, the amount of data transmitted by the eNodeB is equal to $f^{dl}(m, N)$. A cluster configuration is considered as *eligible* only if, with the available N RBs, the selected FDs are able to forward via D2D links all the data received from the BS. As a consequence, a cluster configuration can be considered as eligible only if the MCS level to be used on the D2D link is $l_{m,r} \geq b_m^{dl}$ for all FDs r. In fact, if this condition is not met, more than N resources are needed to relay all data. Based on this observation, the number of iterations of the proposed ECF can be reduced (conditions at lines 14-15 in Algorithm 2).

Once the number of solutions to test is reduced, the algorithm starts testing the cluster configurations with only one FD, than it tests those with two FDs and so on until the maximum number of FDs is reached, that is $\mathcal{K} \setminus \mathcal{K}_m^{dl}$. However, the iterations can actually be interrupted under certain conditions discussed later. For each tested cluster configuration, node $j \in \mathcal{K} \setminus \mathcal{K}_m^{dl}$ is associated to a FD r based on the highest $c_{k,j}$ values in the DCM (lines 19-20). In some cases, especially with a dense node distribution in the cell, it might happen that more than one node can act as a FD for another given node. If multiple potential FDs have the same D2D CQI value, then the eNodeB chooses the one that maximizes the sum of the CQI values. A cluster configuration is eligible only if all users are served and all data can be forwarded by using the available N RBs. For each eligible cluster configuration the spectral efficiency is computed (line 25 in Algorithm 2). When multiple solutions are eligible, the chosen combination is the one with the best spectral efficiency (see line 28).

ECF iterations continue by considering the possible cluster configurations obtained when an additional FD is included. If the cluster configuration chosen at the second step outperforms the one selected at the previous step (see line 29), then the algorithm proceeds by adding another FD and by testing the resulting cluster configurations, otherwise it stops (see line 36) and the most performing tested cluster configuration is chosen.[4] The process described goes on until a

[4]According to condition in line 29, the algorithm continues its iterations also in those cases where no eligible cluster configuration was found with previous tested number of FDs, i.e. e_{MAX} remains equal to zero. This ensures the algorithm to test cluster configurations for instance with two or more FDs even if no eligible cluster configuration was found with only one FD.

solution is selected or the maximum cardinality of FDs is reached.

Algorithm 2: Implementation of the proposed ECF policy

Data: m, $\mathcal{K}_m^{dl}$, N
Result: $\mathcal{R}_m$, $\mathcal{D}_{m,r}$, $N_{m,r}^{ul}$ and $l_{m,r}$
for all $k \in \mathcal{K}_m^{dl}$ **do**
  $v = 0$ ▹ Reduce the number of potential FD nodes to consider
  for $j \in \mathcal{K} \setminus \mathcal{K}_m^{dl}$ **do**
    $h_j = \max_{i \in \mathcal{K}_m^{dl}} \{c_{i,j}\}$ ▹ Highest D2D CQI for receiver j
    if $c_{k,j} = h_i$ **then**
      $v = 1$ ▹ Node k can serve node j with the highest CQI h_j
      break ▹ Node k is not deleted from $\mathcal{K}_m^{dl}$
    end
  end
  if $v = 0$ **then**
    $\tilde{\mathcal{K}}_m^{dl} = \mathcal{K}_m^{dl} \setminus \{k\}$ ▹ The set of potential FD nodes is updated removing node k as it will never be selected
  end
end
$l = \min_{j \in \mathcal{K} \setminus \mathcal{K}_m^{dl}} \{h_j\}$
if $l \geq m$ **then**
  $e_{MAX} = 0$ ▹ Spectral efficiency parameter for solution selection
  for all $g = \{1, 2 \ldots, |\mathcal{K} \setminus \mathcal{K}_m^{dl}|\}$ **do**
    Compute FDs sets $\mathcal{S}_n$, with $n = 1, 2, \ldots, \binom{|\mathcal{K} \setminus \mathcal{K}_m^{dl}|}{g}$ ▹ Set of admissible cluster configurations of g FDs
    for all $n = 1, 2, \ldots, \binom{|\mathcal{K} \setminus \mathcal{K}_m^{dl}|}{g}$ **do**
      For each FD $r \in \mathcal{S}_n$ compute its D2D cluster $\mathcal{I}_{n,r}$ ▹ Add D2D receivers to the FD with highest D2D CQI
      For each FD $r \in \mathcal{S}_n$ compute the assigned resources $T_{n,r}$ and MCS $d_{n,r}$
      if *(* $\bigcup_{r \in \mathcal{S}_n} \mathcal{I}_{n,r} \neq \mathcal{K} \setminus \mathcal{K}_m^{dl}$) $\vee$ ($\sum_{r \in \mathcal{S}_n} T_{n,r}$ *not admissable)* **then**
        Discard $\mathcal{S}_n$ ▹ Not all D2D receivers can be served by the current cluster configuration
      else
        $e_n = (f^{dl}(m, N)|\mathcal{K}|)/(\sum_{r \in \mathcal{S}_n} T_{n,r})$ ▹ Spectral efficiency of current cluster configuration
      end
    end
    $\tilde{n} = \text{argmax} \{e_n\}$ ▹ Cluster configuration with the highest spectral efficiency
    if $(e_{\tilde{n}} > e_{MAX}) \vee (e_{MAX} = 0)$ **then**
      $e_{MAX} = e_{\tilde{n}}$ ▹ Store the current best cluster configuration before continuing to the next iteration
      $\mathcal{R}_m = \mathcal{S}_{\tilde{n}}$ ▹ The selected set of FDs
      $\mathcal{D}_{m,r} = \mathcal{I}_{\tilde{n},r}, \ \forall r \in \mathcal{S}_{\tilde{n}}$ ▹ The served nodes in D2D by each FD
      $N_{m,r}^{ul} = T_{\tilde{n},r}, \ \forall r \in \mathcal{S}_{\tilde{n}}$ ▹ The resources used by each FD
      $l_{m,r} = d_{\tilde{n},r}, \ \forall r \in \mathcal{S}_{\tilde{n}}$ ▹ The MCS for each FD
    else
      Stop iterations
      break ▹ No spectral efficiency improvement is obtained
    end
  end
  No eligible cluster configurations can be found with CQI level m ▹ All data cannot be forwarded
end

B. D2D link configuration: The D2D-SF paradigm

Results in the literature show that improved spectral efficiency can be achieved when D2D links within a cell share the same RBs [10], granted that the D2D pairs using the same RBs are sufficiently apart to avoid mutual interference. In the reference scenario for our research, D2D transmissions are synchronized since they are performed in the same TTI. As a consequence, all considered FDs share the same portion of RBs without introducing interference. We refer to this policy as *D2D-enhanced CMS with single frequency* (*D2D-SF*).

A main assumption is that all involved FDs exploit the same MCS to feed the relevant D2D receivers. The choice of the MCS is driven by the minimum $l_{m,r}$ with $r \in \mathcal{R}_m$ among those observed in all the D2D clusters to be activated (i.e., the worst channel conditions in all clusters). The selected value also determines the total amount of resources needed for data relaying. This policy may be used in combination with any of the two cluster formation schemes presented in previous sections. In the remaining sections we will use the acronym $D2D\text{-}SF^B$ when it is coupled to BCF and $D2D\text{-}SF^E$ when it is coupled to ECF.

For performance comparison, we also consider a resource allocation policy, *D2D-enhanced CMS* (hereafter simply referred to as *D2D* in the simulative analyses), which assigns different frequency resources to the different clusters. As a consequence, the eNodeB sets the MCS for each FD (i.e., $l_{m,r}$ $\forall r \in \mathcal{R}_m$) to the one supported by the users in $\mathcal{D}_{m,r}$ with the worst channel conditions. Then, the eNodeB selects the number of resources $N^{ul}_{m,r}$, required by r to forward the $f^{dl}(m, N)$data received from the BS over the cellular link. The considered cluster configuration is *eligible* if the sum of resources assigned to the FDs is equal (or less than) the number of available RBs N. In the remainder of the paper, we will use the name $D2D^B$ when the coupled cluster formation scheme is BCF, and $D2D^E$ when the cluster formation is ruled by ECF.

VI. Simulation Settings and Performance Metrics

An extensive numerical evaluation is conducted using Matlab®. The performance analysis is performed following the guidelines for the LTE system model in [40] and [41]. The main simulation parameters are listed in Table III. The parameters for the LTE system are set according to [1]. We have considered that $R = 100$ RBs are available in the LTE system on a 20 MHz channel bandwidth. Channel conditions for the UEs are evaluated in terms of signal to interference and noise ratio (SINR) experienced on each sub-carrier [42], [43] when path loss and fading phenomena affect the signal reception. The effective SINR is mapped onto the CQI level that ensures a block error rate (BLER) smaller than 1% [42], [44].

Algorithm 3: The proposed radio resource allocation policies

Data: m, $\mathcal{R}_m$, $\mathcal{D}_{m,r}$, N
Result: $N^{ul}_{m,r}$ and $l_{m,r}$
switch *Resource Allocation Scheme* **do**
  case $D2D^B$ *and* $D2D^E$
    for all $r \in \mathcal{R}_m$ **do**
      $l_{m,r} = \min_{j \in \mathcal{D}_{m,r}} \{c_{r,j}\}$
      $N^{ul}_{m,r} = \lceil (b^{dl}_m N)/(b^{ul}_{l_{m,r}}) \rceil$
    end
    if $\sum_{r \in \mathcal{R}_m} N^{ul}_{m,r} > N$ **then**
      $\mathcal{R}_m$ is discarded
    end
  endsw
  case $D2D\text{-}SF^B$ *and* $D2D\text{-}SF^E$
    for all $r \in \mathcal{R}^m$ **do**
      $v_r = \min_{j \in \mathcal{D}_{m,r}} \{c_{r,j}\}$
    end
    $l_{m,r} = \min \{v_r\}, \ \forall r \in \mathcal{R}_m$
    $N^{ul}_{m,r} = \lceil (b^{dl}_m N)/(b^{ul}_{l_{m,r}}) \rceil, \ \forall r \in \mathcal{R}_m$
    if $N^{ul}_{m,r} > N$ **then**
      $\mathcal{R}_m$ is discarded
    end
  endsw
endsw

The following metrics have been considered to evaluate the performance of the proposed solutions with respect to CMS [6] and OMS [7]:

- *mean data rate*, measured as the mean data rate value experienced by the multicast members;
- *aggregate data rate (ADR)*, computed as the sum of the data rates experienced by the multicast users;
- *resource usage*, that is the percentage of RBs used by the eNodeB for the multicast data transmission;
- *served users*, that is the percentage of users which successfully received the multicast content;
- *fairness index*, measured in terms of the Jain's fairness index [45]:

$$FI = \frac{(\sum_{i=1}^{|\mathcal{K}|} d_i)^2}{|\mathcal{K}|(\sum_{i=1}^{|\mathcal{K}|} d_i^2)} \quad (1)$$

  where d_i is the data rate for UE i.

To better assess the behaviour of the proposed schemes, different scenarios have been evaluated by varying the multicast group size $|\mathcal{K}|$ and their distribution within the cell. In addition, the

TABLE III
MAIN SIMULATION PARAMETERS

Parameter	Value
Cell radius	500 m [41]
Frame Structure	Type 2 (TDD) [1]
TTI	1 ms (11 OFDM data symbols plus 3 control symbols)
Cyclic prefix/Useful signal frame length	16.67 μs / 66.67 μs
TDD configuration	1
Carrier Frequency	2.5 GHz
eNodeB Tx power	46 dBm
D2D node Tx power	23 dBm [37]
Noise power	-174 dBm/Hz
Path loss (cell link)	128.1 + 37.6 log(d), d[km]
Path loss (D2D link, NLOS)	40 log(d) + 30 log(f) + 49, d[km], f[Hz]
Path loss (D2D link, LOS)	16.9 log(d) + 20 log (f/5) + 46.8, d[m], f[GHz]
Shadowing standard deviation	10 dB (cell mode); 12 dB (D2D mode)
RB size	12 sub-carriers, 0.5 ms
Sub-carrier spacing	15 kHz
BLER target	1% [44]

number of resources R managed by the eNodeB is also considered as a further variable. The simulations are organized in two studies, focusing on *video streaming* and *Video on Demand (VoD)* analysis, respectively.

The first study evaluates the performance in terms of the metrics indicated when focusing on video streaming towards multicast users. Similar to [46], we simulated a video streaming session lasting 1s, which is considered as the reference time unit for the performance evaluation in this paper. Within this video session time, the transmission parameters (also used in [47] and [46]), are adapted on a frame-basis by the BS. Following the LTE standard [1], we set the scheduling frame duration to 10 ms. During this 1s-long session, 100 data frames are transmitted, with 10 TTIs per data frame, and each TTI lasting 1ms, offering a relative large number of frames for the computation of the average results and good assessment of the proposed solution's improvements in terms of performance. For longer sessions, we expect no significant variations in terms of performance and a similar trend in the results. In fact, the same data frame structure is repeated over time in the system and no important channel fluctuations are expected for the almost static reference scenarios considered in the paper. In this analysis, three different study cases are considered:

- **Case A:** This case studies the impact the channel bandwidth has on the considered policies. In this case we set the multicast group size $|\mathcal{K}|$ to 200, whereas a variable number of resources R (ranging from 10 to 100 RBs, which is the maximum value in the LTE standard) is dedicated to the service. We consider that UEs are distributed within an area of 100m x 100m (e.g., users in a stadium or attending an open space event) and are located near the cell-edge. This is the most challenging scenario for the users as the channel quality to the users decreases;
- **Case B:** This case analyses the impact a varying multicast group size has on the considered policies. The number of available resources R is set to 100 RBs (i.e., the maximum available in the system), whereas the number of UEs $|\mathcal{K}|$ ranges from 20 to 200 (representative of small and relative large groups of users, respectively). The same cell-edge distribution of UEs as in case A is considered;
- **Case C:** This case assesses the impact of user density within the cell as an additional parameter to identify the scenarios where the D2D links introduce benefits compared to traditional approaches. UEs are distributed in an area whose size varies from [100m x 100m] to [1000m x 1000m] (representative of the cases where the multicast group is scattered in a portion of the cell or over the whole cell), the number of UEs $|\mathcal{K}|$ also varies from 20 to 500 (representative of small and large groups of users, respectively). Three sample channel bandwidth deployment scenarios (with 25, 50, and 100 RBs, respectively) are evaluated. Then for the same UEs distribution, we let the number of available resources R range from 10 to 100 RBs, when the density of UEs per square meter is fixed to $0.005\, UE/m^2$. As shown in a previous work in [29], user distributions with such a density offer good D2D communication opportunities. Moreover, this value guarantees a multicast group size of tens of UEs also for the smallest area size considered in the analysis, so that the considered simulation setting is also of interest for these cases.

The second study, which analyses *Video on Demand (VoD)* , considers a typical video delivery application and, consequently, evaluates the performance over the whole time required for the service (i.e., several scheduling frames) for the multicast users to receive the video content from the eNodeB.

VII. Performance Evaluation

In this section the performance evaluation for the proposed solutions is discussed with reference to the two video transmission applications as detailed in the previous section.

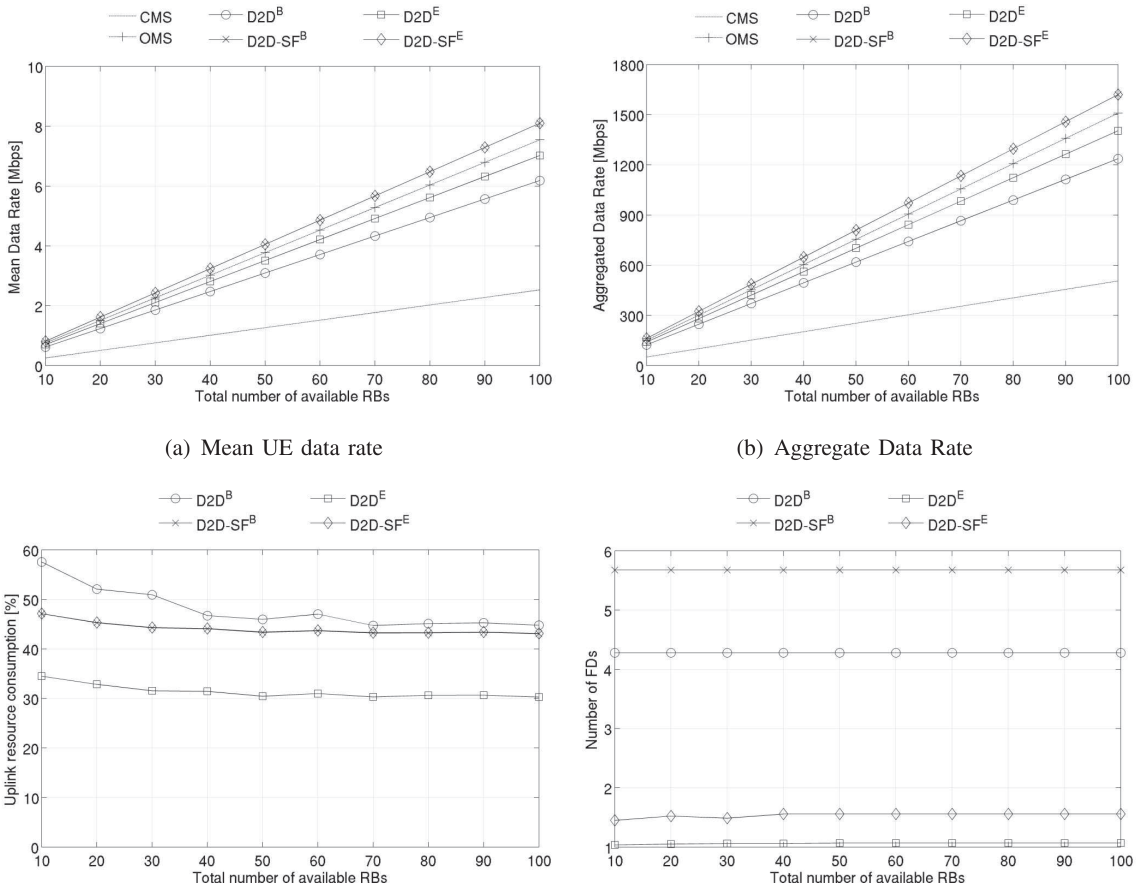
A. Video Streaming Analysis

In this study, we consider a video streaming service transmitted by the BS. Video parameters are set in accordance to [46], where adaptive video coding [48] is assumed to be performed at the BS. This simplification is reasonable, as it is similar with the situation when the adaptation is performed at a remote server and ideal delivery conditions are considered in the core network up to the BS. We tuned the video parameters such that the video stream has an average bit rate between 256 kbps and a maximum value which depends on the channel quality experienced by multicast users.

1) Case A: The focus in this case is on the mean data rate achieved by multicast users and ADR. The results obtained are plotted in Fig. 3. As expected, both mean data rate and ADR increase with the number of available RBs for all solutions. All the proposed D2D-based schemes outperform CMS, with a better performance for $D2D\text{-}SF^B$ and $D2D\text{-}SF^E$. These single-frequency solutions also outperform OMS.

The benefit compared to CMS remains constant with the number of available RBs and equal to 144%, 177%, and 220% for the $D2D^B$, $D2D^E$, and $D2D\text{-}SF^E$ (the same value is obtained for $D2D\text{-}SF^B$) solutions respectively. The mean data rate for CMS is lower than $D2D^B$, $D2D^E$ and $D2D\text{-}SF^E$ (Fig. 3(a)). The data rate of the OMS solution reaches higher values than the $D2D^B$ and $D2D^E$ solutions, but lower values compared to solutions based on the single-frequency paradigm. The price to pay when adopting OMS is the reduction in the number of served users and in short-term fairness (i.e., the fairness measured within one scheduling frame). Moreover, it will be shown later that further drawbacks are observed on the Video on Demand scenarios. A similar trend is found for ADR (Fig. 3(b)).

More details on the behaviour of the cluster formation and resource allocation policies proposed are given by results plotted in Fig. 3(c) and Fig. 3(d) (showing the percentage of uplink resources used and the number of FDs elected). Adopting the ECF clustering policy reduces the number of selected FDs compared to the adoption of BCF. Moreover, a lower number of resources are used in the uplink for the $D2D^E$ compared to the $D2D^B$ case. When comparing



(a) Mean UE data rate

(b) Aggregate Data Rate

(c) Average number of uplink resources used

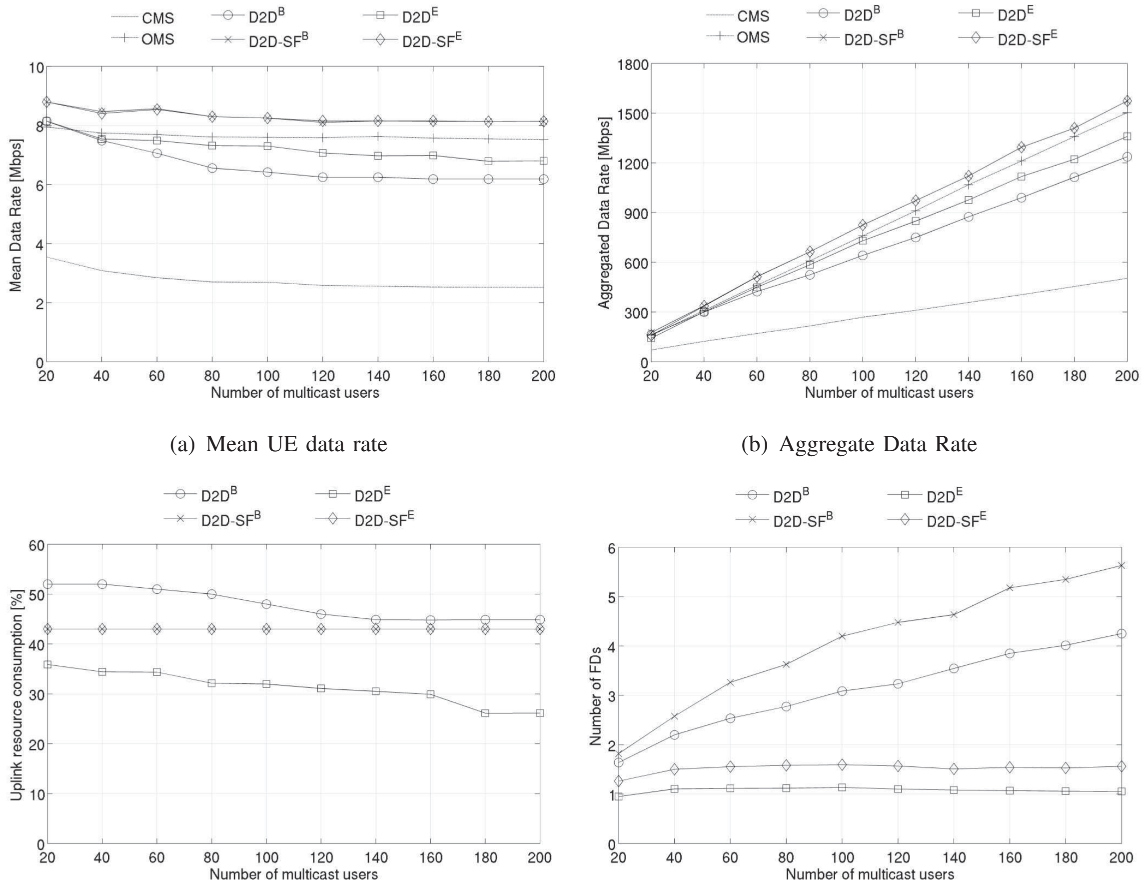
(d) Average number of FDs selected

Fig. 3. Performance analysis for video streaming analysis: Study case A.

the plots for $D2D\text{-}SF^E$ and $D2D^E$, we observe that the latter one activates fewer FDs and uses a lower number of RBs in the uplink. This is not surprising, because $D2D^E$ selects less performing solutions in downlink (as observed from plots in Fig. 3) with the consequence that it has to activate lower CQI levels to obtain an *eligible* solution.

2) Case B: The performance achieved by varying the multicast group size is illustrated in Fig. 4. Also in this case, the novel schemes adopting the single-frequency paradigm, i.e., $D2D\text{-}SF^B$ and $D2D\text{-}SF^E$, outperform the others. Noteworthy, for all the solutions, the mean data rate slightly decreases when the number of users in the cell increases, as shown in Fig. 4(a). In particular, the CMS shows a performance reduction of about 29% when passing from 20 UEs 200 UEs. When considering $D2D^B$, $D2D^E$, and $D2D\text{-}SF^E$, there is a 25%, 16%, and 8%

reduction respectively. This is an expected result since the greater the number of users in the group, the higher the risk of having users with very low channel conditions, which limit the overall performance. Also OMS observes a reduction in its offered mean data rate with the increase in number of UEs. Notwithstanding, this phenomenon is less evident, as the mean data rate decreases from 7.9 Mbps (with 20 UEs) to 7.5 Mbps (with 200 UEs).



(a) Mean UE data rate

(b) Aggregate Data Rate

(c) Average number of uplink resources used

(d) Average number of FDs selected

Fig. 4. Performance analysis for video streaming analysis: Study case B.

As expected, Fig. 4(b) shows how the ADR value increases with the number of multicast members in the cell. Moreover, the gain introduced by $D2D^B$, $D2D^E$, and $D2D\text{-}SF^E$ with respect to the CMS solution is larger when the number of multicast users increases. Regarding the number of uplink resources used in this case (illustrated in Fig. 4(c)) and the number of activated FDs (see Fig. 4(d), similar considerations as in the study case A hold. The only difference is in the reduced number of resources needed. Furthermore, an increase in the number of multicast users increases the number of FDs required by the schemes based on the BCF clustering algorithm (however, the same does not hold for ECF based schemes).

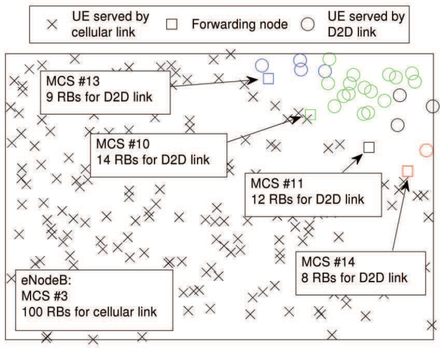
Final comments are on the short-term FI in the data rate assignment. FI = 1 is the maximum fairness value that is achieved when all UEs are served at the same data rate. While the OMS achieves a FI equal to 0.78, the FI of all other solutions is equal to 1.

A meaningful example of service configuration is plotted in Fig. 5. In particular, the role of each UE in the group is shown with reference to a cell-edge scenario. It clearly emerges that different service and cluster configurations are obtained in the four considered cases. The better performance achieved by the proposed single-frequency paradigm is further sustained by the higher number of users served with D2D links and, consequently, the radio spectrum is more efficiently used (i.e., less robust MCSs is adopted on the cellular link).

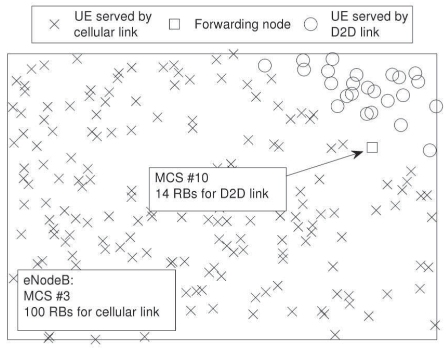
3) Case C: As mentioned, the objective in this case is to assess the performance of the D2D-based solutions for a wide set of UEs distributions within the cell. To do this, CMS is used as a benchmark of minimum performance in the tested scenarios. The area where the UEs are uniformly distributed is progressively extended from the cell-edge scenario until the whole cell of 1000x1000 m is covered.

In details, the average data rate benefits, compared to the CMS, introduced by the $D2D^E$ and $D2D\text{-}SF^E$ schemes (only the best performing D2D-based solutions are considered due to length constraints) are plotted in Fig. 6 and 7 respectively, for three sample values of R and a variable number of UEs. When comparing the cases with 25, 50, and 100 RBs in the subplots, the first observation, is that the number of available resources has almost no influence on the data rate benefit. This benefit increases in general with the number of users in the MG (x-axis in the plots) and decreases with the MG area size (the y-axis in the plots reports the side length of the considered square area). This is an expected behaviour as the D2D coverage range is limited and larger areas with the same number of UEs reduce the possibility to exploit the D2D links. Moreover, based on a general observation of the results, the $D2D\text{-}SF^E$ emerges as the best performing scheme in the video streaming case study.

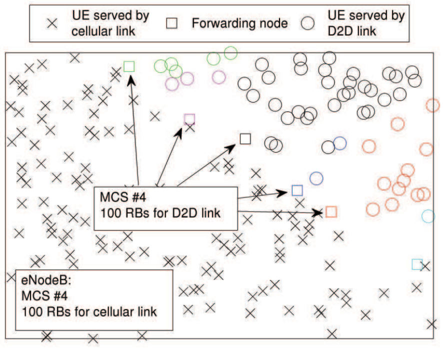
To further investigate the influence of the area covered by the MG, a variable area size is considered in the $[100m \text{ x } 100m - 1000m \text{ x } 1000m]$ range, where the density of UEs in this area is kept at a constant value of $0.005\, UE/m^2$. The focus is on the data rate benefit introduced by



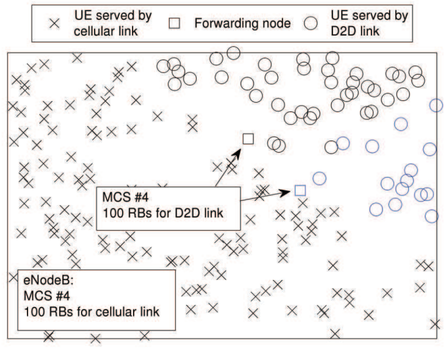
(a) $D2D^B$



(b) $D2D^E$

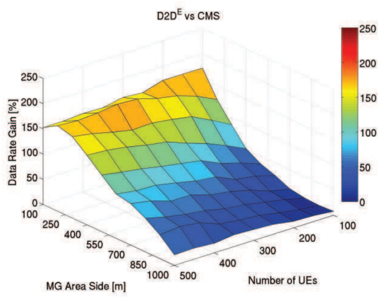


(c) $D2D\text{-}SF^B$

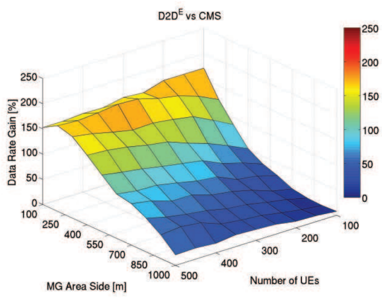


(d) $D2D\text{-}SF^E$

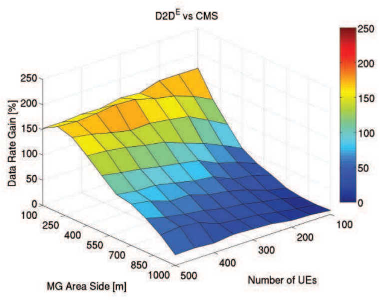
Fig. 5. Sample MG configuration in video streaming analysis for $D2D^B$, $D2D^E$, $D2D\text{-}SF^B$ and $D2D\text{-}SF^E$ solutions (200 UEs, 100 RBs).



(a) 25 RBs



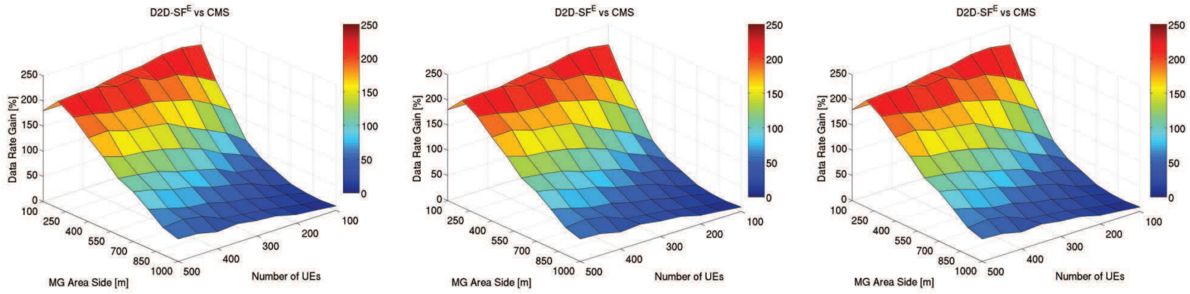
(b) 50 RBs



(c) 100 RBs

Fig. 6. Data rate gain in video streaming analysis for $D2D^E$ *vs.* CMS.

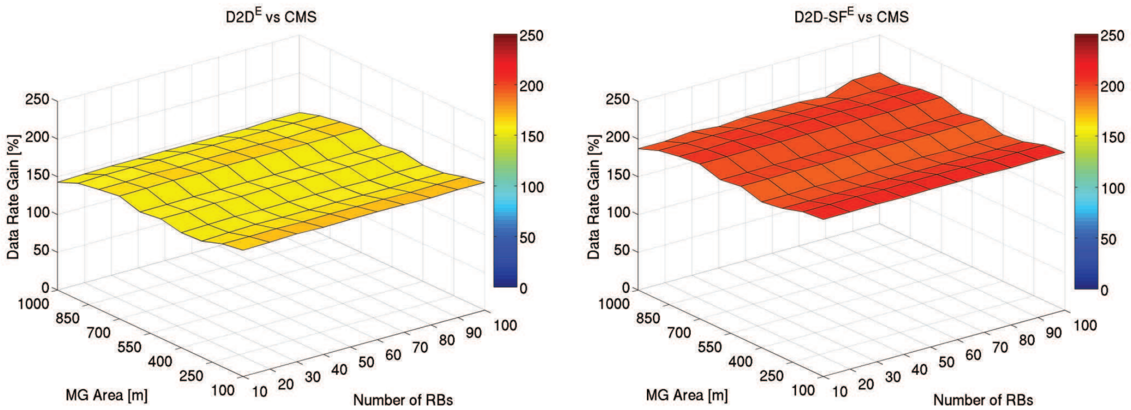
the best performing D2D-based solutions compared to the CMS solution by varying the number of resources R. The results relevant to the $D2D^E$ and $D2D\text{-}SF^E$ solutions are plotted in Fig.



(a) 25 RBs (b) 50 RBs (c) 100 RBs

Fig. 7. Data rate gain in video streaming analysis for *D2D-SF^E vs.* CMS.

8(a) and 8(b), respectively. In both cases the constant node density provides a more or less constant data rate improvement, regardless the number of RBs and the area size. In details, the average benefit for the *D2D*^{E} case is 160% in Fig. 8(a), while this improvement is 200% for the *D2D-SF^E* case plotted in Fig. 8(b). Once again, the best performance is associated to solutions based on the proposed single-frequency paradigm.



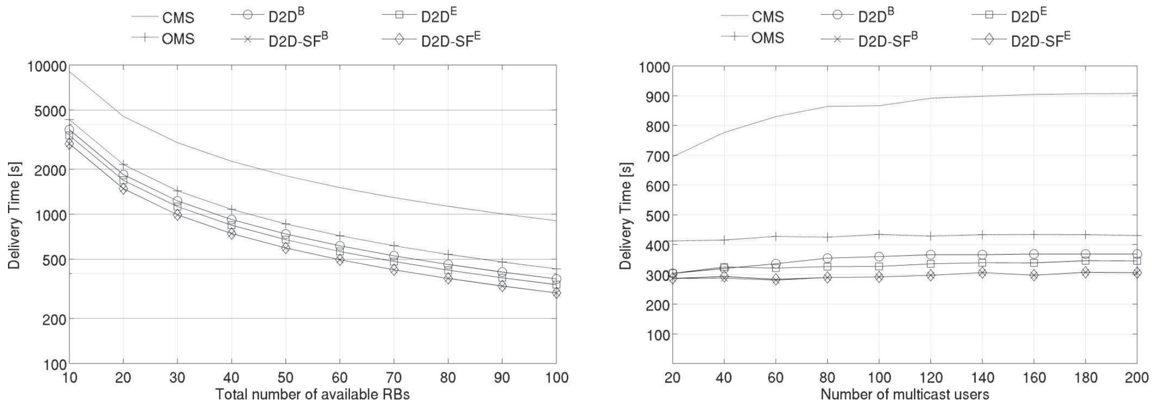
(a) *D2D*^{E} *vs.* CMS. (b) *D2D-SF^E vs.* CMS.

Fig. 8. Data rate gain in video streaming analysis with a constant node density in the MG area.

B. Video on Demand (VoD) Analysis

For this analysis, we consider a Video on Demand (VoD) application [49], where VoD content is simultaneously downloaded by multiple multicast members. To simulate this scenario, we implemented a file-transfer service where all UEs forming the MG subscribe to receive a 1*GByte* file of common interest. The main performance parameters are the average content delivery time, the average data rate, and the long-term (i.e., measured when all multicast members accomplished the VoD download) user fairness.

The first analysis focuses on the mean VoD content delivery time in the same scenarios considered for the study cases A and B. In particular, the results in Fig. 9(a) refer to a multicast group size $|\mathcal{K}|$ set to 200, a variable number of resources dedicated to the service R ranging from 10 to 100 RBs, and a cell-edge distribution of UEs over a concentrated area of $100x100$ m. Instead, the results in Fig. 9(b) refer to a number of available resources R set to 100 RBs, a number of UEs $|\mathcal{K}|$ varying in the range $[20-200]$, and a cell-edge distribution of UEs over an area of $100x100$ m.



(a) Case A, variable number of RBs.

(b) Case B, variable number of UEs.

Fig. 9. Mean delivery time in the VoD analysis.

The results for a varying number of available resources are shown in Fig. 9(a). For all considered policies, an increase in the number of RBs available to the service causes a reduction in the mean delivery time. It is interesting to underline that in this VoD scenario, all D2D-based schemes outperform both CMS and OMS solutions. In Fig. 9(b), one observes that the mean delivery time increases with the multicast group size to a different extent for each considered solution. As expected, CMS is the worst performing policy characterized by the highest delivery time, which ranges from about 695 (for 20 UEs) to about 907s (for 200 UEs). The most interesting behaviour is observed for the novel single frequency schemes $D2D\text{-}SF^B$ and $D2D\text{-}SF^E$ (showing the same trend). Not only the best mean delivery time performance is obtained, with 286 and

305s in the 20 and 200 UEs cases respectively, but almost no increase in the delivery time is observed with the increase in the number of UEs.

Finally, while for all the novel solutions and for CMS the long-term fairness is 1, for the OMS the long-term FI is equal to 0.57. This shows that also in VoD downloading scenarios the OMS manifests its main drawback in terms of low fairness during service delivery. This limitation is effectively overcome by employing our proposed D2D-based approaches.

The main performance findings in cases A and B of both video streaming and VoD analysis are summarized in Table IV. In particular, the minimum, average and maximum data rate benefits introduced by the D2D-enhanced solutions w.r.t. to the baseline CMS solution are reported. Moreover, also the number of FDs and the percentage of uplink resources used for D2D communication is indicated in order to highlight the differences introduced by the two proposed clustering algorithms.

TABLE IV
Comparison of D2D-enhanced solutions and benefits w.r.t. to baseline CMS multicast video delivery

	Scenario	$\mathbf{D2D^B}$	$\mathbf{D2D^E}$	$\mathbf{D2D-SF^B}$	$\mathbf{D2D-SF^E}$
Data rate gain (min/avg/max) [%]	Case A	144/144/144	177/177/177	220/220/220	220/200/220
	Case B	129/141/145	112/158/169	148/197/211	148/197/211
Delivery time gain (1 GByte data) (min/avg/max) [%]	Case A	59/59/59	63/63/63	67/67/67	67/67/67
	Case B	43/56/59	48/60/62	53/61/66	53/61/66
# of FDs[5] (min/avg/max)	Case A	2/4.2/8	1/1.2/3	2/5.6/10	1/1.5/3
	Case B	2/3.1/7	1/1.2/3	2/4.2/13	1/1.5/3
Percentage of uplink resources[5] (min/avg/max) [%]	Case A	17/48/90	11/31/100	11/44/100	11/44/100
	Case B	17/40/100	11/30/100	11/40/100	11/40/100

VIII. Conclusions

In this paper proposes the Device-to-Device (D2D)-enhanced Conventional Multicast Scheme (CMS) with Single Frequency (D2D-SF), a novel strategy for multicast video delivery in LTE-A systems, D2D-SF exploits the advantages introduced by the high-performing D2D links between UEs within the multicast group in order to improve the performance of cell-edge devices and guarantee benefits for the whole multicast group. In particular, the single-frequency D2D links are dynamically activated so that the "best" forwarding devices (FDs) are employed. This allows increasing the spectrum efficiency as all the D2D transmissions exploit the same frequencies

[5]The reported value gives a range of values for all the considered scenarios.

and consequently results in improvement of the overall system throughput, while maintaining the typical CMS short-term fairness. The proposal is compared to both state-of-the-art solutions, such as conventional and opportunistic schemes, and basic novel multi-frequency D2D solutions. As demonstrated through numerical evaluations in a wide set of scenarios, the proposed D2D-enhanced single-frequency paradigm introduces significant enhancements in terms of efficient delivery of multicast services, both for video streaming and for video on demand applications. Future enhancements will focus on the video quality assessment of the schemes proposed in this work.

References

[1] 3GPP, "TS 36.300, Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN), Rel. 11," Tech. Rep., Sept. 2012.

[2] Cisco, "Cisco VNI Forecast Widget 2014," Sept. 2014, http://www.ciscovni.com/forecast-widget/advanced.html.

[3] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2013-2018," June 2014.

[4] 3GPP, "3GPP, TS 36.440, General aspects and principles for interfaces supporting Multimedia Broadcast Multicast Service (MBMS) within E-UTRAN, Rel. 11," Tech. Rep., 2012.

[5] L. Carla, F. Chiti, R. Fantacci, C. Khirallah, and A. Tassi, "Power efficient resource allocation strategies for layered video delivery over eMBMS networks," in *Communications (ICC), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3505–3510.

[6] A. Richard, A. Dadlani, and K. Kim, "Multicast Scheduling and Resource Allocation Algorithms for OFDMA-Based Systems: A Survey," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 1, pp. 240–254, 2013.

[7] T. P. Low, M. O. Pun, Y. W. P. Hong, and C. C. J. Kuo, "Optimized opportunistic multicast scheduling (OMS) over wireless cellular networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, Sept. 2009.

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

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

[10] C.-H. Yu, K. Doppler, C.B. Ribeiro, and O. Tirkkonen, "Resource Sharing Optimization for Device-to-Device Communication Underlaying Cellular Networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2752–2763, Aug. 2011.

[11] M. Condoluci, L. Militano, G. Araniti, A. Molinaro, and A. Iera, "Multicasting in LTE-A networks enhanced by device-to-device communications," in *Globecom Workshops (GC Wkshps), 2013 IEEE*, Dec 2013, pp. 567–572.

[12] S. Andreev, A. Pyattaev, K. Johnsson, O. Galinina, and Y. Koucheryavy, "Cellular traffic offloading onto network-assisted device-to-device connections," *Communications Magazine, IEEE*, vol. 52, no. 4, pp. 20–31, April 2014.

[13] A. Pyattaev, O. Galinina, S. Andreev, M. Katz, and Y. Koucheryavy, "Understanding practical limitations of network coding for assisted proximate communication," *Selected Areas in Communications, IEEE Journal on*, vol. PP, no. 99, pp. 1–1, 2014.

[14] J Seo, Taesoo Kwon, and V Leung, "Social Groupcasting Algorithm for Wireless Cellular Multicast Services," *IEEE Communications Letter*, vol. 17, no. 1, pp. 47–50, Jan. 2013.

[15] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, and A. Iera, "When D2D communication improves group oriented services in beyond 4G networks," *Wireless Networks*, pp. 1–15, 2014.

[16] K. Doppler, M.P. Rinne, P. Jänis, C.B. Ribeiro, and K. Hugl, "Device-to-Device communications; functional prospects for LTE-Advanced networks," *IEEE International Conference on Communications (ICC), Dresden, Germany*, pp. 1–6, June 2009.

[17] 3GPP, "TS 36.440, General aspects and principles for interfaces supporting Multimedia Broadcast Multicast Service (MBMS) within E-UTRAN, Rel. 11," Tech. Rep., Sept. 2012.

[18] A. Alexious, C. Bouras, V. Kokkinos, and G. Tsichritzis, "Communication cost analysis of MBSFN in LTE," *IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Istanbul, Turkey*, pp. 1366–1371, Sept. 2010.

[19] L. Zhang, Z. He, K. Niu, B. Zhang, and P. Skov, "Optimization of coverage and throughput in single-cell E-MBMS," *IEEE 70th Vehicular Technology Conference Fall (VTC-Fall), Barcelona, Spain*, pp. 1–5, Sept. 2009.

[20] A. Alexious, C. Bouras, V. Kokkinos, A. Papazois, and G. Tsichritzis, "Spectral efficiency performance of MBSFN-enabled LTE networks," *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Niagara Falls, Canada*, pp. 361–367, Oct. 2010.

[21] P. K. Gopala and H. E. Gamal, "Opportunistic multicasting," *Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, pp. 845–849, Nov. 2004.

[22] CW Huang, SM Huang, PH Wu, SJ Lin, and JN Hwang, "OLM: Opportunistic Layered Multicasting for Scalable IPTV over Mobile WiMAX," *IEEE Transactions on Mobile Computing*, vol. 11, no. 3, pp. 453–463, Mar. 2012.

[23] L. Militano, M. Condoluci, G. Araniti, and A. Iera, "Multicast Service Delivery Solutions in LTE-Advanced Systems," *IEEE International Conference on Communications (ICC), Budapest, Hungary*, June 2013.

[24] L. Militano, D. Niyato, M. Condoluci, G. Araniti, A. Iera, and G. Molica Bisci, "Radio Resource Management for Group-Oriented Services in LTE-A," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2014.

[25] G. Araniti, M. Condoluci, L. Militano, and A. Iera, "Adaptive Resource Allocation to Multicast Services in LTE Systems," *Broadcasting, IEEE Transactions on*, vol. 59, no. 4, pp. 658–664, Dec 2013.

[26] G. Araniti, M. Condoluci, A. Iera, A. Molinaro, J. Cosmas, and M. Behjati, "A Low-Complexity Resource Allocation Algorithm for Multicast Service Delivery in OFDMA Networks," *Broadcasting, IEEE Transactions on*, vol. 60, no. 2, pp. 358–369, June 2014.

[27] S.C. Spinella, G. Araniti, A. Iera, and A. Molinaro, "Integration of Ad-hoc Networks with infrastructured systems for multicast services provisioning," in *Ultra Modern Telecommunications Workshops, 2009. ICUMT '09. International Conference on*, Oct 2009, pp. 1–6.

[28] Q. Zhang, F.HP Fitzek, and V.B. Iversen, "Design and performance evaluation of cooperative retransmission scheme for reliable multicast services in cellular controlled P2P networks," *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, (PIMRC), Athens, Greece*, pp. 1–5, Sept. 2007.

[29] L. Militano, M. Condoluci, G. Araniti, A. Molinaro, A. Iera, and F.H.P. Fitzek, "Wi-Fi cooperation or D2D-based multicast content distribution in LTE-A: A comparative analysis," in *Communications Workshops (ICC), 2014 IEEE International Conference on*, June 2014, pp. 296–301.

[30] T. Koskela, S. Hakola, Tao Chen, and J. Lehtomaki, "Clustering Concept Using Device-To-Device Communication in

Cellular System," *IEEE Wireless Communications and Networking Conference (WCNC), Sydney, Australia*, pp. 1–6, Apr. 2010.

[31] B. Zhou, H. Hu, S. Huang, and H. Chen, "Intra-Cluster Device-to-Device Relay Algorithm with Optimal Resource Utilization," *IEEE Transactions on Vehicular Technology*, Jan. 2013.

[32] J. Seppala, T. Koskela, T. Chen, and S. Hakola, "Network controlled Device-to-Device (D2D) and cluster multicast concept for LTE and LTE-A networks," *IEEE Wireless Communications and Networking Conference (WCNC), Cancun, Mexico*, pp. 986–991, Mar. 2011.

[33] M. Lanza, A.L. Gutierrez, J.R. Perez, J. Morgade, M. Domingo, L. Valle, P. Angueira, and J. Basterrechea, "Coverage Optimization and Power Reduction in SFN Using Simulated Annealing," *Broadcasting, IEEE Transactions on*, vol. 60, no. 3, pp. 474–485, Sept 2014.

[34] W. Joseph, L. Verloock, D. Plets, E. Tanghe, and L. Martens, "Characterization of Coverage and Indoor Penetration Loss of DVB-H Signal of Indoor Gap Filler in UHF Band," *Broadcasting, IEEE Transactions on*, vol. 55, no. 3, pp. 589–597, Sept 2009.

[35] L. Rong, O. Ben Haddada, and S. Elayoubi, "Analytical Analysis of the Coverage of a MBSFN OFDMA Network," *IEEE Global Communications Conference (IEEE GLOBECOM), New Orleans, LA, USA*, pp. 1–5, Dec. 2008.

[36] M. Iturralde, T.A Yahiya, A. Wei, and A.Beylot, "Interference mitigation by dynamic self-power control in femtocell scenarios in LTE networks," *IEEE GLOBECOM*, pp. 4810–4815, Dec. 2012.

[37] 3GPP, "TS 36.213 Evolved Universal Terrestrial Radio Access (E-UTRA): Physical layer procedures, Rel. 11," Tech. Rep., Dec. 2012.

[38] H. Holma and A. Toskala, "LTE for UMTS: OFDMA and SC-FDMA based radio access," *John Wiley & Sons*, 2009.

[39] Y. J. Li and R. F. Song, "Compression of CQI feedback in adaptive correlated MIMO-OFDM systems," *International Journal of Information and Computer Science*, vol. 8, no. 3, pp. 562–569, Mar. 2011.

[40] 3GPP, "TR 36.210, LTE physical layer; General description, Rel. 11," Tech. Rep., Sept. 2012.

[41] 3GPP, "Physical Layer Aspects for Evolved Universal Terrestrial Radio Access (UTRA), v.7.1.0," Tech. Rep., Sept. 2006.

[42] C. Mehlführer, M. Wrulich, J. C. Ikuno, D. Bosanska, and M. Rupp, "Simulating the Long Term Evolution Physical Layer," Aug. 2009, pp. 1471–1478.

[43] ITU-R, "Guidelines for evaluation of radio interface technologies for IMT-Advanced, Report M.2135," Tech. Rep., 2009.

[44] A. Urie, A. Rudrapatna, C. Raman, and J-M. Hanriot, "Evolved Multimedia Broadcast Multicast Service in LTE: An Assessment of System Performance Under Realistic Radio Network Engineering Conditions," *Bell Labs Technical Journal*, vol. 18, no. 2, pp. 57–76, Sept. 2013.

[45] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared systems," *tech. rep., Digital Equipment Corporation, DEC-TR-301*, 1984.

[46] S. Sharangi, R. Krishnamurti, and M. Hefeeda, "Energy-Efficient Multicasting of Scalable Video Streams Over WiMAX Networks," *Multimedia, IEEE Transactions on*, vol. 13, no. 1, pp. 102–115, Feb 2011.

[47] S. Deb, S. Jaiswal, and K. Nagaraj, "Real-Time Video Multicast in WiMAX Networks," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, Apr. 2008, pp. 1579–1587.

[48] Y. Wang, L.-P. Chau, and K.-H. Yap, "Bit-rate allocation for broadcasting of scalable video over wireless networks," *Broadcasting, IEEE Transactions on*, vol. 56, no. 3, pp. 288–295, Sept. 2010.

[49] G. Muntean, P. Perry, and L. Murphy, "A comparison-based study of quality-oriented video on demand," *Broadcasting, IEEE Transactions on*, vol. 53, no. 1, pp. 92–102, March 2007.