This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

In-Band Device-to-Device Communication in OFDMA Cellular Networks: A Survey and Challenges

Pavel Mach, *Member, IEEE,* Zdenek Becvar, *Member, IEEE,* Tomas Vanek

***Abstract*—Direct communication between two or more devices without the intervention of a base station, known as device-to-device (D2D) communication, is a promising way to improve performance of cellular networks in terms of spectral and energy efficiency. The D2D communication paradigm has been largely exploited in non-cellular technologies such as Bluetooth or Wi-Fi but it has not yet been fully incorporated into existing cellular networks. In this regard, a new proposal focusing on the integration of D2D communication into LTE-A has been recently approved by the 3GPP standardization community as discussed in this paper. In cellular networks, D2D communication introduces several critical issues, such as interference management and decisions on whether devices should communicate directly or not. In this survey, we provide a thorough overview of the state of the art focusing on D2D communication, especially within 3GPP LTE/LTE-A. First, we provide in-depth classification of papers looking at D2D from several perspectives. Then, papers addressing all major problems and areas related to D2D are presented and approaches proposed in the papers are compared according to selected criteria. On the basis of the surveyed papers, we highlight areas not satisfactorily addressed so far and outline major challenges for future work regarding efficient integration of D2D in cellular networks.**

***Index Terms*—D2D communication, D2D mode selection, interference management, D2D energy efficiency, advanced topology for D2D**

I. INTRODUCTION

The ever-increasing requirements and demands of users of mobile wireless networks are the main drivers for further enhancement of the network capacity. To fulfill demands of users in the future, more radical steps have to be taken into consideration as described, e.g., in [1]. One of the most promising approaches is efficient reuse of existing frequency bands. This can be accomplished by several options, such as densification of base stations and deployment of small cells underlying the conventional cellular networks [2], use of a cognitive radio approach and spectrum sharing [3], or direct communication between users without the intervention of a base station, known as device-to-device (D2D) communication. In this paper, we focus on the D2D communication within 3GPP LTE/LTE-A, which has recently attracted the attention of the research community.

As mentioned in [4], D2D communication in cellular networks can be seen as conceptually similar to cognitive radio principles. In both approaches we can distinguish primary users (conventional users of cellular network) and secondary users, who either access the spectrum through cognitive sensing or reuse radio resources by means of D2D communication. Nevertheless, D2D communication is mostly managed by the network whereas cognitive radio is fully autonomous and uses cognitive sensing.

D2D communication can be classified into in-band D2D and out-band D2D. In the case of in-band D2D (in [5], also referred to as LTE direct), D2D communication takes place in a licensed spectrum allocated to the cellular operators. The D2D users (DUEs) can access the licensed spectrum either in a dedicated mode (also described in the literature as an overlay or orthogonal mode) or a shared mode (also known as an underlay or a non-orthogonal mode). In the case of out-band D2D, D2D communication exploits the unlicensed spectrum adopted by other wireless technologies supporting direct communication such as WiFi direct (based on IEEE 802.11) or Bluetooth (based on IEEE 802.15).

A comparison of Bluetooth, WiFi direct and in-band D2D (LTE direct) is provided in Table I. In general, use of an unlicensed spectrum for D2D communication may result in poor quality of service (QoS), because of uncontrolled interference. In addition, the discovery process and setup of connection in out-band D2D or in WiFi direct/Bluetooth is quite complicated since it needs user intervention to establish the connection between the two devices. In contrast to these technologies, in-band D2D uses licensed bands, where the above-mentioned problems are solved in a more efficient way. The reason is that the cellular network commonly controls all D2D communication that is currently ongoing and adopts sophisticated allocation and interference mitigation techniques to provide QoS to its users. On the other hand, since cellular networks are managed by the operators, users are likely to have to pay for the connection.

The advantages of adopting D2D in cellular networks are as follows [6]. First, the proximity of the DUEs promises high bit rates and/or less power consumption (higher energy efficiency) thanks to good channel quality between communicating devices. Second, users are supposed to experience lower packet delays. Third, instead of two-hop transmission via the eNB as in the conventional cellular network, data are sent directly in one hop, saving radio resources. Fourth, reusing the same radio resources with the cellular network can increase spectral efficiency. This, however, depends on how radio resources are shared between the D2D communication

P. Mach, Z. Becvar, and T. Vanek are with the Department of Telecommunication Engineering, Czech Technical University in Prague, Czech Republic, 166 27 Prague, Czech Republic. (e-mail: machp2@fel.cvut.cz, zdenek.becvar@fel.cvut.cz, tomas.vanek@fel.cvut.cz).

Manuscript received xx; revised xx

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

TABLE I
COMPARISON OF WIFI DIRECT AND BLUETOOTH TECHNOLOGIES TO IN-BAND D2D(LTE DIRECT).

Aspect/Technology	Bluetooth	WiFi direct	(LTE direct) in-band D2D
Spectrum	Unlicensed	Unlicensed	Licensed
Interference control	No	No	Yes
QoS	No	No	Yes
Discovery process	Pairing procedure (manually)	Two-steps asynchronous message based discovery	Devices broadcast their services at physical layer [5]
Range	tens of meters [5]	up to 100m [5]	up to 500m [5]
Users' cost	Free of charge	Free of charge/Charging by operator	Charging by operator

and the conventional communication through the eNB, as discussed later in this paper. Despite the advantages of D2D, there are also some limitations implied by the concept. As pointed out in [7], one of the limiting factors is that the probability of direct communication between two devices is relatively low (i.e., D2D communication cannot be used very often because of distance restrictions). However, as indicated in [5], LTE should allow D2D communication distances up to 500m, which is sufficient. Clearly, this aspect heavily depends on the density of UEs in the area. The optimal density and maximum allowable density of D2D devices in a certain area are studied in [8] and [9], respectively. The study in [8] demonstrates that the optimal D2D density is determined by the amount of interference from the cellular network. The authors in [9] show that with an increase in the number of D2D devices, the outage of cellular users (CUEs) decreases. The other crucial problem regarding the D2D concept's exploited in LTE/LTE-A is its immaturity, resulting in many technical challenges. The most critical one is the mode selection and interference management.

The most common use of D2D communication in cellular networks is the offloading of local traffic from an evolved Node B (eNB), resulting in increased network capacity. D2D communication is also suitable for sharing specific contents between close user equipments (UEs) or for gaming purposes. In addition, D2D enables multicasts/broadcasts of user information to several UEs in proximity, or relaying of data using UEs instead of conventional relay nodes. Additional use cases and business models for D2D communication are addressed in [10].

Two surveys on D2D communication have been published. In [11], the authors give an overview of papers addressing in-band D2D in terms of spectral efficiency, energy efficiency, cellular coverage and other performance targets. Further, papers dealing with out-band D2D communication both managed by the network and autonomously controlled are surveyed in [11]. The paper also briefly reviews D2D architecture. Finally, it discusses the advantages of in-band and out-band approaches, the maturity of D2D and its implementation in the real world. The survey presented in [12] is oriented on performance evaluation techniques, application/services for D2D communication, and existing prototypes and experiments. In contrast, our survey focuses mostly on the state of the art approaches regarding interference and radio resource management, which are discussed only briefly in [12].

In our survey we complement both the above-mentioned surveys and provide the missing pieces of information, primarily focusing on in-band D2D communication. To be more precise, with respect to [11][12] our survey:

- gives a more detailed overview of network architecture in order to incorporate D2D communication into cellular networks (Section II)
- provides a comprehensive classification of D2D communication in terms of several aspects, such as D2D management, D2D scenarios and D2D radio resource management, to help the reader with orientation in this domain (Section III)
- gives deep insight into D2D mode selection and contemplates, and notes which mode is more appropriate under specific circumstances (Section IV)
- distinguishes several interference problems regarding D2D, such as interference from D2D in cellular communication, interference of cellular in D2D communication and mutual interference in D2D communication, and comprehensively surveys papers dealing with this topic (Section V)
- analyzes research dealing with power consumption and energy efficiency in more detail (Section VI)
- emphasizes the usability of the D2D communication paradigm for advanced topology scenarios (besides more conventional direct communication) such as multicast/broadcast concepts and D2D relay functionality (Section VII)
- includes a section dealing with the coexistence of D2D communication with small cells (Section VIII)
- gives a more detailed overview of 3GPP standardization activities and discusses related 3GPP documents (Section VIII)
- provides a more detailed overview of future research directions in in-band D2D cellular networks, such as mode selection, interference management, energy efficiency, mobility management and security issues (Section X)

Note that in this survey we provide an overview of papers primarily focusing on OFDMA cellular networks, since OFDMA is applied in 4G and considered also for 5G future mobile networks. Hence, initial technical studies dealing with the D2D communication concept and considering CDMA [13][14], WCDMA [15]-[17], or OFDM [18][19] are intentionally left out to keep the paper more focused. Also, we do not present details of D2D dependent on individual use cases. This information can be found in [20].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

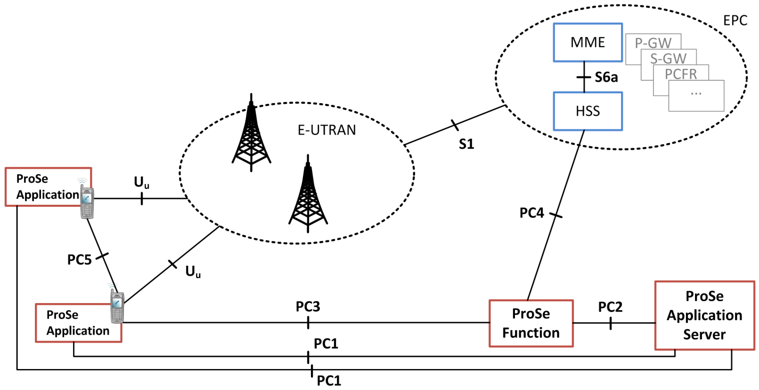

Fig. 1. Enhancement of LTE-A network architecture with D2D communication (new entities are in red color, blue color indicates modified entities) [23].

II. Mobile network architecture supporting D2D communication

To enable exploitation of D2D communication in cellular networks, new features and functionalities related to D2D must be introduced into existing cellular networks. As D2D is intended for future mobile networks, its integration is based on the existing architecture of an LTE-A network presented and defined in [21].

Like a conventional LTE-A network, the LTE-A architecture supporting D2D is composed of the evolved packet core (EPC) and the evolved universal terrestrial access network (E-UTRAN) (see Fig. 1). The requirements on features that should be supported by the architecture are described in [22]. Moreover, the document [22] tackles the problem of the EPC enhancement by means of new interfaces and entities to support D2D functionality. The outcomes of [22] are transformed to specification of 3GPP standard Release 12 [23]. As proposed in [23], the support of D2D is enabled by two new functional entities on the network side, proximity-based service (ProSe) Function and ProSe Application Server, and one new entity on the user side, ProSe Application (see Fig. 1).

The ProSe Function entity is implemented as a logical function which provides three different sub-functions. The first sub-function, Direct Provisioning Function, covers provision of parameters for D2D discovery and D2D communication. These parameters, listed in Section 4.5.1 in [23], are related to authorization policy (e.g., whether the UE is authorized to perform D2D discovery and/or D2D communication), the radio parameters needed to configure the UE to enable D2D discovery and/or communication (e.g., frequency band), and D2D communication parameters (e.g., whether IPv4 or IPv6 should be used). The second sub-function, Direct Discovery Name Management Function, enables us to identify a D2D application and its support in the network of the operator for D2D discovery purposes. The third sub-function, EPC-level Discovery ProSe Function, provides network-related functionalities, such as authorization, charging and subscriber information management. In 3GPP Release 12, only one ProSe Function is expected. This can become a limiting factor if the function becomes overloaded. Therefore, 3GPP has left open doors for potential future enhancements enabling multiple ProSe Functions. Nevertheless, cooperation among multiple ProSe Functions must be defined in the future.

The ProSe Application Server ([24]) provides functionality of ProSe applications and maps users to the individual functions. It also stores information about all available functions. Note that the application itself is deployed in the UE but it is not subject to 3GPP standards. The ProSe Application Server is connected with the ProSe Function by a PC2 interface, which defines interaction between both entities for ProSe discovery purposes as described in [24].

Also, the UEs must be modified in order to enable D2D communication. This assumes extension of the UE with support of D2D discovery and communication by the ProSe Application. In addition, for relaying scenario, also UE's relay functionality must be supported. The authorization policy for D2D discovery and communication is handled over PC3 interface by protocol defined in [25]. The D2D communication between the UEs is performed over PC5 interface, which is also defined in [25] along with the protocol carried over this interface in [25]. The ProSe application in the UE communicates with the ProSe application server through PC1 interface in order to define application layer signaling features and parameters.

All three new entities (ProSe Function, Application, Application Server) introduces security threats and risks related to D2D communication. Therefore, in [26], the authors first analyze the new architecture from the security point of view and propose key management among common LTE-A entities and new introduced entities in order to avoid potential risks from the security side.

Besides new entities, also existing Mobility Management Entity (MME) and Home Subscriber Server (HSS) must be enhanced in order to enable exchange of user's information regarding ProSe services for authorization purposes [23] (see Fig. 1). To that end, new interface PC4 is defined between the ProSe Function and the HSS (see [27]). Also, interface S6a between the HSS and the MME must be enhanced to enable exchange of information related to the ProSe subscription information [23].

Before above-described architecture has been defined by 3GPP, it was heavily addressed in literature. In [28], the authors have proposed a new concept of D2D communica-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

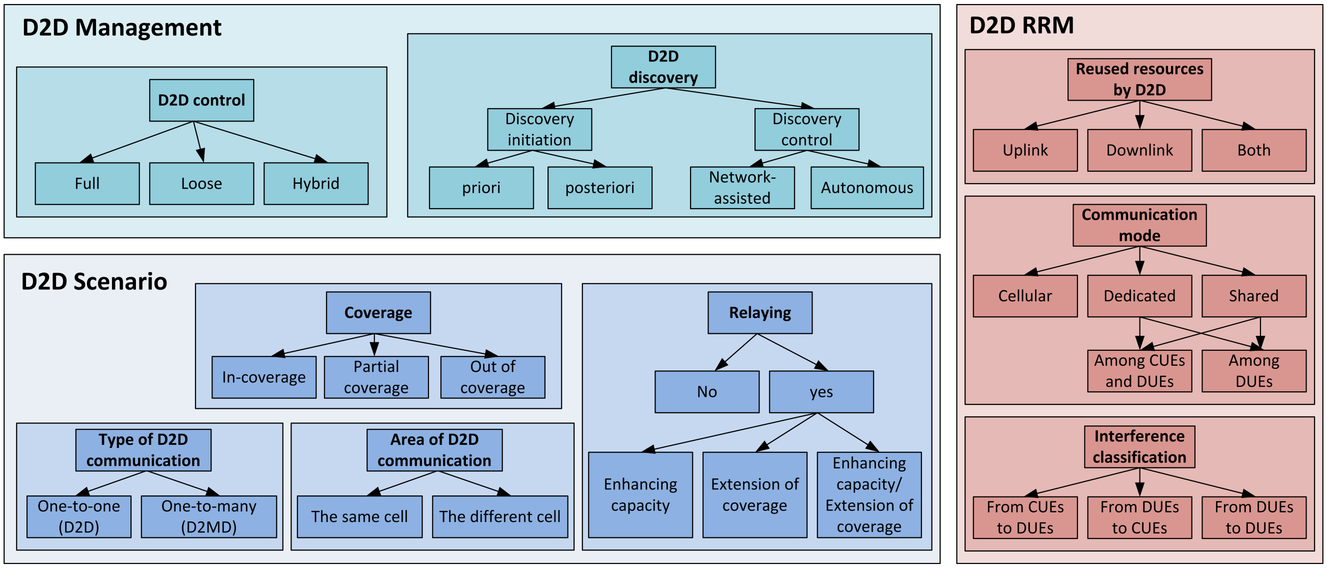

Fig. 2. Classification of D2D communication in cellular networks.

tion underlying LTE-A network and shown how the D2D communication can be established within a system architecture evolution (SAE). The study introduces an exchange of messages to support the D2D functionality within the SAE and contemplates the possible limits of the D2D concerning interference issues both in the downlink (DL) and the uplink (UL) transmission directions. The paper also presents feasibility analysis evaluating performance of network with enabled D2D communication. Functional prospects of the D2D communication and its implementation into LTE-A system are tackled in [29]. The paper describes new features necessary to be added into the SAE architecture in order to support the D2D communication: radio identification and bearer setup, means to exchange the information over a D2D connection and interference management, link adaptation, timing, and mobility issues. Design aspects of network assisted D2D communication is addressed in [30]. The paper firstly provides a brief overview on technical challenges posed by enabling of D2D concept in cellular networks and provides the solutions for individual challenges.

An option of architectural modification in LTE-A networks for D2D is also proposed in [31]. The authors introduce new network entity, called a D2D server (in Fig. 1, this server is not depicted to keep clarity of the figure), and necessary interfaces to connect it to the existing LTE-A architecture. The D2D server is located within the EPC and interfaces with a (MME), a policy and charging rules function (PCRF), peer D2D servers, and with application servers. The logical functions of the D2D server are, for example, device identifier allocation, policy management, assistance in location, call establishment, or mobility tracking. Further, the paper proposes protocol stack describing protocol termination for the D2D communication. Finally, procedures necessary for establishing and maintaining the D2D connection are introduced. With respect to 3GPP architecture, this D2D server is analogous to ProSe application server with ProSe function.

III. D2D CLASSIFICATION

This section describes high level overview on D2D classification. As indicated in Fig. 2, the classification of D2D can be divided into several main distinctive categories; D2D management, D2D scenarios, and D2D radio resource management (RRM). Individual categories are described in the following subsections.

A. D2D Management

From the management point of view, D2D can be classified according to the level of the network's involvement in control of D2D and a how D2D communication is established (denoted as D2D discovery). In this section, the main characteristic of both management aspects are described and their pros and cons are discussed.

1) D2D control: The D2D control indicates how deeply the network is involved in the control and management of D2D communication. As shown in Fig. 2, the control of D2D can be fully managed by the network (full control), partly managed by the network (loosely control) or hybrid (between full and loosely control).

In case of full control, the D2D communication is managed by the network of the operator [10]. To be more specific, the network is fully responsible for D2D authentication process during D2D discovery and initiation phase, it handles D2D connection, and allocates power and radio resources. An advantage of the full control approach is that the network can easily coordinate D2D and cellular communications. Thus, the network can mitigate harmful interference between the CUEs and the DUEs. Moreover, the eNB can effectively perform radio resource management and give priorities to individual transmissions to fulfill various QoS requirements. On the other hand, disadvantage of the full control can be seen in high signaling overhead necessary to manage the D2D underlying communication. For example, the eNB has to know the channel state information (CSI) of all involved links for interference avoidance technique and its exchange is very demanding in terms of signaling [32].

If loose control is applied, D2D terminals can autonomously communicate between themselves with very limited or no intervention from the network's side [10]. The network itself is practically responsible only for authentication of the terminals during network entry. The D2D communication can be initiated autonomously and resource allocation or power control is handled solely by the DUEs. The most prominent advantage of this approach consists in generation of only negligible signaling overhead comparing to full control. The critical issue of loosely controlled D2D, however, is the interference caused by the D2D to the legacy CUEs. Consequently, this approach can be avoided by the operators and mobile providers, since the top priority is to guarantee QoS for the CUEs. One possible way to make loosely controlled D2D more feasible is to use unlicensed spectrum for this kind of communication and sharing the frequencies with WiFi or Bluetooth as suggested in [10].

The last option for D2D control is a hybrid one [33]. In the hybrid control, the most critical aspects are done by the network (similarly as in full control). Among these are authentication process, allocation of radio resources in a large time scale, put restriction on maximal power control allowed at the side of D2D, etc. At the same time, the DUEs are able to manage radio resources and schedule their own transmission and set power control autonomously in a distributive manner according to a short time measurement (similarly as in loosely control). As a consequence, the hybrid control could be seen as a good trade-off for the operators, as generated signaling overhead is reasonable, while QoS offered to the CUEs can be guaranteed.

From the above-mentioned, it could be derived that the D2D control influences the amount of signaling needed to be exchanged between the network and the DUEs during ongoing D2D communication. Besides, the D2D control has an impact on the implementation complexity of D2D and it defines how much the UE has to be modified in order to support D2D functionality (in case of loosely controlled D2D, more modifications are expected at the UE's side).

2) D2D discovery: Essential part of the D2D management is a discovery of the DUEs (in literature also known as a peer discovery procedure). The purpose of D2D discovery process is to find the presence of devices that could potentially communicate directly. The overall discovery process can be split into two stages: discovery initiation and discovery control.

The D2D discovery can be initiated either before the DUEs start to communicate (labeled as "priori") or during ongoing communication (known as "posteriori") [34]. The common use of priory D2D discovery is a sharing of a specific content between two devices. On the other hand, posteriori D2D discovery could be used, for example, by mobile devices that move to vicinity of each other during data exchange and D2D communication becomes more suitable/efficient.

Like overall D2D control (described in Section III-A1), the discovery process can be controlled with different levels of network involvement. Thus, the discovery can be controlled either fully by the network (network assisted D2D discovery) or autonomously by the DUEs, which find potential counterparts on its own (autonomous D2D discovery). The network assisted D2D discovery is more convenient as the network is aware of approximate device locations (e.g., whether the DUEs about to communicate with each other are within the same cell or not). On the other hand, network assisted D2D discovery can result in higher signaling overhead due to its centralized nature. The advantage of autonomous D2D discovery is that it has low signaling overhead because it is fully distributed. Nevertheless, discovery process itself could drain battery significantly as it is fully handled by individual devices.

A procedure for network assisted D2D discovery is proposed in [35]. Firstly, the packet data network gateway (P-GW) detects potential D2D candidates. Secondly, a message exchange involving the MME, the eNB, and the UEs participating in D2D is initiated. After the D2D bearer is established, D2D communication can take place instead of conventional cellular communication. The proposed D2D discovery, however, can result in quite significant signaling overhead. Another network assisted D2D discovery utilizing resources allocated for the device discovery is proposed in [36]. At a specific discovery interval, several UEs perform discovery by means of a discovery message. To avoid contention among them, the UEs take three steps, which are random selection of; i) search/listen state, ii) discoverable interval, and iii) frequency multiplexed discovery channel. The results indicate that the proposed technique is able to increase the amount of discovered D2D within one discovery period. Signature-based D2D discovery, during which the DUEs transmit discovery signal using temporary identity, is proposed in [37]. The paper shows how discovery signal is mapped to the physical resources and how to avoid collisions at discovery channel, which is allocated by the network.

Autonomous D2D discovery technique intended for communication systems based on Qualcomm's defined D2D - FlashLinQ [38] is introduced in [39]. A fully distributed D2D discovery for synchronous OFDM-based system using time division multiplexing (TDM) technique is proposed in [40]. The radio resources contain discovery region, during which the devices receive or transmit discovery signals. This allows each device to advertise its presence and service and to discover other close devices autonomously and continuously in distributed manner. Another autonomous D2D discovery protocol based on dynamic source routing protocol is introduced in [41]. The proposed scheme floods the network with modified discovery packets. These packets include channel number, power used for transmission and interference power measured by transmitting mode. Hence, the receiver is able to calculate the path loss and signal to interference plus noise ratio (SINR), which are used for estimation of minimum transmit power to be heard by the transmitter. As a result, two nodes can create D2D pair if bidirectional link can be established while the constraint on power is set to minimize interference to the CUEs. Using peer discovery resources for beacon signal is proposed in [42]. To ensure low discovery overhead, only small part of physical layer frames are used for the D2D discovery. The discovery of neighboring device is determined according to SINR measured from the received beacon.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

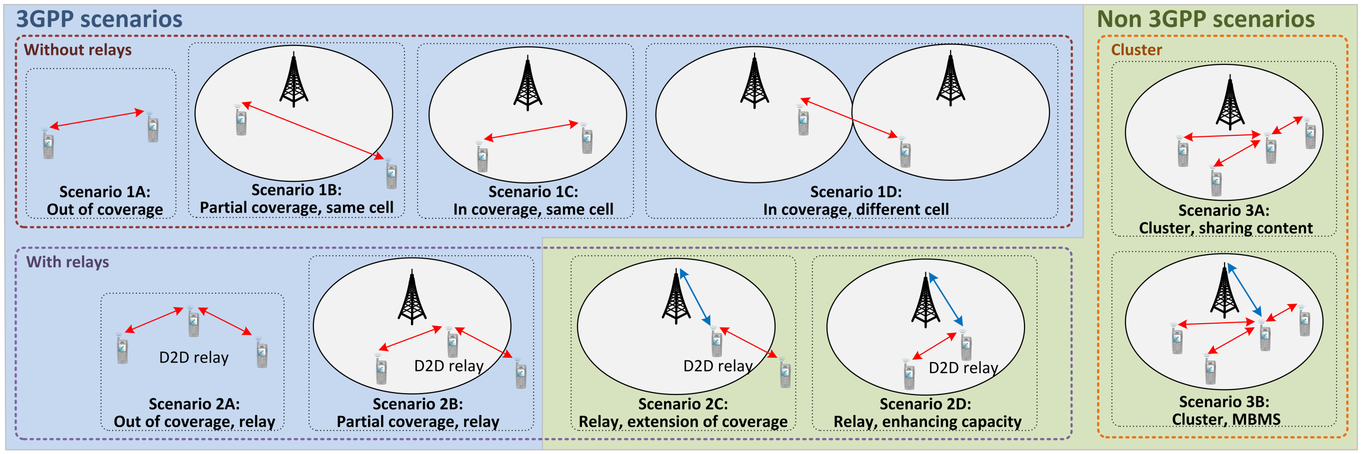

Fig. 3. Scenarios for D2D communication.

B. D2D communication scenarios

This section illustrates individual possible scenarios that could be considered for D2D communication. In general, D2D scenarios can be classified by several aspects (see Fig. 2) as follows.

- *Coverage* - This aspect distinguishes whether the pair of DUEs is under the coverage of a cellular network. In this context we can categorize D2D communication as:
  - *In coverage* - Both DUEs are within the coverage of the cellular network.
  - *Partial coverage* - One DUE is in the coverage of the cellular network whereas the second one is out of coverage (e.g., it could be in a coverage hole caused by interferes in the proximity).
  - *Out of coverage* - Both DUEs are outside the cellular communication network. Note that this scenario is considered mainly in 3GPP for public safety cases, when the network can be temporarily disabled.
- *Type of D2D communication* - This aspect expresses how many DUEs are involved in D2D communication:
  - *One-to-one communication* - Direct communication between two DUEs that creates one D2D communication pair.
  - *One-to-many communication* - One DUE multicasts/broadcasts data to several DUEs in a cluster. This option is also labeled device to multi-device (D2MD).
- *Area of D2D communication* - The aspect distinguishing whether both communicating DUEs are served by the same cell or not:
  - *The same cell* - The DUEs creating a D2D pair or a cluster are attached to the same eNB.
  - *The different cell* - The DUEs belonging to the same D2D pair or cluster are attached to different eNBs.
- *Relaying functionality* - The DUE may have relaying functionality to retransmit data of other DUEs within its proximity. This feature can be used to:
  - *Enhance capacity* The DUE attached to another DUE with relay functionality is usually in coverage of the eNB.
  - *Extend coverage* The DUE out of coverage may use other DUEs in order to reach the eNB.

Depending on the above-mentioned four aspects, several scenarios for D2D communication can be specified (see Fig.3). First, a set of D2D scenarios corresponds to those taken into account in a 3GPP standardization group, according to [22]. Note that scenario labeling is in line with 3GPP and also we consider only scenarios when all eNBs belong to the same public land mobile network (PLMN). In addition to 3GPP scenarios, we include possible options for D2D communication tackled in research studies but not currently standardized in 3GPP.

In general, 3GPP categorizes D2D scenarios into two groups; without relays and with relays. The scenarios without relaying are more conventional, since there are no special requirements imposed on the UE in order to support relaying functionality. The most common scenario used in contemporary studies and standards is a simple communication between two devices underlying cellular communication, where both devices are served by the same eNBs. In 3GPP, this scenario is denoted as "in coverage, same cell" and it is labeled as Scenario 1C. In Scenario 1C, either one device transmits some data to the other device in its proximity or both devices exchange data mutually. The second possibility is to facilitate D2D communication of two devices that belong to two different cells. This scenario is known in 3GPP as "in coverage, different cell" (Scenario 1D). Similarly to Scenario 1C, both directly communicating devices are under coverage of a cellular network. Nevertheless, this scenario is much more complicated in terms of its establishment and in solving interference issues because of the need for cooperation by both involved cells. The third option that could occur in the network is that only one DUE is under coverage of the cellular network and the second one is out of coverage, i.e., only partial coverage is provided (Scenario 1B). If both DUEs are out of coverage of the cellular network, they can independently form a D2D pair without infrastructure and initiate communication (Scenario 1A). This option is used mostly for public safety scenarios in case of emergency and disasters if a network infrastructure is not available.

As already mentioned, 3GPP also considers the option whereby one UE can serve as a relay between two DUEs. The 3GPP assumes two options; the first option assumes all involved DUEs are out of network coverage (Scenario 2A) whereas the second option reflects the case when partial network coverage is available (Scenario 2B). Note that in this case, the UE must be enhanced by relaying capabilities in order to have the means to receive and retransmit the signal in the same way as the conventional relay nodes.

So far, all the above-mentioned scenarios are defined by 3GPP. In addition to these we can find another four viable options to be used for D2D communication. The first two extend the idea of relay functionality. The main objective of the relay UE can be to extend the area served by the eNBs (Scenario 2C) [43]. The extension of the coverage by means of D2D communication can be exploited if a UE is not able to connect to the eNB directly because the UE is of out of the eNB's coverage. In this case, the UE can be served through a relay UE. The communication between the UE and the eNB occurs in two hops instead of one. In the first hop, the UE sends its data to the relay UE through a D2D link. In the second hop, the relay retransmits data to the eNB in the cellular mode. The UE relay can also help enhancing capacity of the eNB (Scenario 2D) [44]. The capacity enhancement is mainly convenient if the UE is experiencing low quality signal from the eNB despite the fact that it is in the coverage of the eNB. Consequently, the attachment to the eNB through the relay UE offers high quality connection resulting in a higher capacity.

The last two scenarios assume a specific case, where more than two UEs in close distance can form a cluster. In the first scenario, several users within the same cluster may share the same content (e.g., video, music, etc.) (Scenario 3A) [45][46]. In such cases one user is usually selected to be a cluster head and acts as a file/content sharer and the other devices in close communication distance become cluster members and are meant to receive file/content shared by the cluster head. For these purposes, multicast or broadcast communication is considered. The other specific scenario profiting from D2D communication is real-time streaming content facilitated by the multicast and broadcast multimedia service (MBMS) currently supported by LTE/LTE-A standards. In this scenario, one device selected as a cluster head receives transmission from the eNB and retransmits it to other devices within the same cluster (Scenario 3B) [48][47]. In this case, the cluster head can be considered as a specific type of relay since it retransmits data from the eNB to several devices in the vicinity.

C. D2D radio resource management

From the RRM perspective, D2D communication can be classified by several criteria depending on: duplexing mode and reuse of resources assigned for D2D communication, the used communication mode, and the type of interference management.

1) Duplexing mode and reuse of resources: The D2D communication underlying cellular networks can access the resources either in TDD or FDD duplexing mode. Resources originally allocated for the UL, the DL, or both can be reused for direct communication. The most common approach in the current literature is to use the UL resources of cellular networks (see, e.g., [49]-[51]). The advantage of the UL is that this direction is mostly underutilized compared with the DL, since most users would rather download data from the network. In addition, the interference situation in the UL is much easier to resolve with respect to cellular transmission because the victim of D2D interference is solely the eNB. Although the problem of underutilization of UL resources can be partly solved by a suitable TDD frame configuration [52], interference in the DL is still a significant obstacle.

2) Communication mode: The communication mode identifies whether the DUEs communicate directly with each other or via the eNB. Furthermore, it distinguishes if D2D communication uses the same radio resources as the conventional cellular communication or not. In this regard, we can recognize the following communication modes used by D2D applied in the current literature (see Fig. 4).

- *Cellular mode (CM)* - The CM corresponds to the conventional cellular communication as the DUEs exchange data through the eNB and no direct exchange of data between the DUEs takes place. This mode is usually utilized if UEs are too far from each other or simply if D2D communication would not pay off. The advantage of the CM mode is that interference can be easily managed by the eNB and no new features have to be implemented. On the other hand, the CM is characterized by a low spectral efficiency.
- *Dedicated mode (DM)* - The DM is a mode allowing two DUEs to transmit data directly between themselves without intermediate eNB, which would relay data. Still, the eNB has to dedicate radio resources for the DUEs' transmission and thus the CUEs cannot exploit the full capacity of the eNB. As can be observed in Fig. 4, the radio resources are used with a higher efficiency than in the case of the CM, since only one transmission direction, either the DL or the UL, is used for the D2D transmission. Note that in some of the literature this allocation mode is also referred to as an orthogonal mode or an overlay mode, as transmission of the CUEs and the DUEs has assigned a non-overlapping orthogonal radio resource. The advantage of the DM is that the eNB does not need to handle interference among the CUEs and the DUEs.
- *Shared mode (SM)* - In the SM, the same radio resources are used both for the DUEs and for the CUEs. In some of the literature this mode is referred to as a non-orthogonal or an underlay mode. Similarly to the DM, the SM can use either the DL or the UL radio resources. From the spectral efficiency perspective this option is even more profitable for the system, since the reuse factor is significantly higher than in the case of the CM or the DM. Nevertheless, the SM also has some disadvantages because strong interference could be generated among the DUEs and the CUEs. To prevent the generation of harmful interference, new techniques and procedures have to be introduced to the system. As a result, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

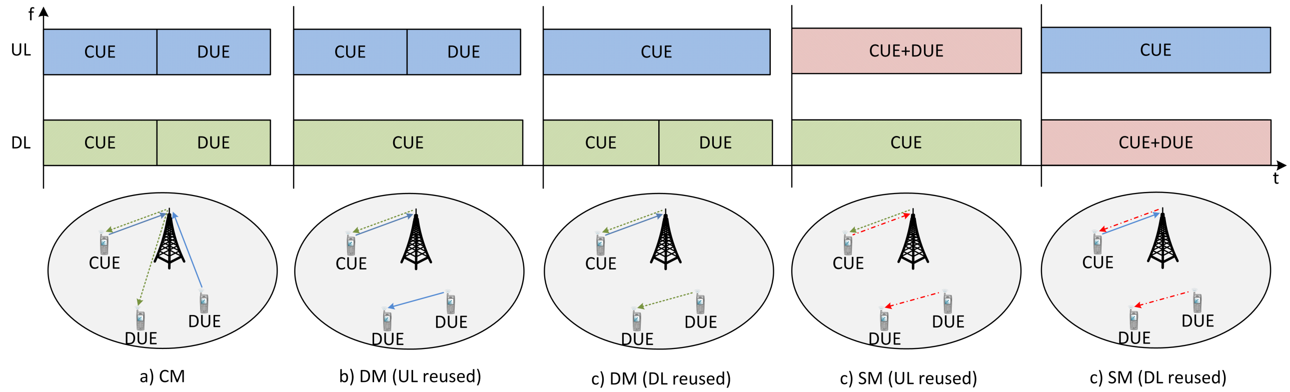

Fig. 4. Communication modes exploited by D2D communication.

TABLE II
COMPARISON OF THE D2D COMMUNICATION MODES WITH RESPECT TO SPECTRAL EFFICIENCY, THE AMOUNT OF INTERFERENCE, AND SYSTEM COMPLEXITY.

Criteria/Type of D2D mode	CM	DM (DUEs use dedicated resources)	DM (DUEs use shared resources)	SM (DUEs use dedicated resources)	SM (DUEs use shared resources)
Spectral efficiency	Low	Medium	High	High	Very high
Interference among CUEs and DUEs	No	No	No	Yes	Yes
Interference among DUEs	No	No	Yes	No	Yes
Implementation complexity	Low	Low	Medium	Medium	High

complexity of whole system is increased (usually, the more sophisticated the interference cancellation technique the more requirements are imposed on the system).

Thus far, we have determined only whether the DUEs use dedicated or shared resources with respect to the CUEs. An important aspect that should also be considered is if several D2D pairs can reuse the same resources or not. This option makes the D2D communication even more spectrally efficient. Nonetheless, this type of allocation is, at the same time, the most challenging to implement in real systems, since not only interference among the DUEs and the CUEs has to be resolved but also interference among individual D2D pairs. The general comparison of individual allocation modes in terms of their spectral efficiency, interference issues and complexity is shown in Table II. Note that in Table II, the individual modes are compared regarding spectral efficiency and implementation. For example, the CM is less spectrally efficient than the DM (DUEs use dedicated resources). In addition, the DM with DUEs using shared resources is even more efficient than the DM with DUEs using dedicated resources. Consequently, the CM is considered to have low spectral efficiency whereas the DM with DUEs using dedicated resources and the DM with DUEs utilizing shared resources enjoy medium and high spectral efficiency, respectively.

As explained above, the communication mode affects overall spectral efficiency of the system and its complexity. The papers addressing the problem of proper selection of communication mode are discussed in detail in Section IV.

3) Interference classification: Interference is the most critical problem regarding D2D communication. The reason is that the cellular communication should not be affected by the introduction of D2D communication. In addition, if the D2D communication is strongly interfered by the cellular communication, its applicability and/or efficiency is significantly reduced. The nodes that are affected by interference (eNB, CUE, DUE) depend on D2D communication mode and on the resources used for D2D communication (UL/DL). As already explained, interference between the CUEs and the DUEs occurs only if the CUEs and the DUEs are in the SM. Furthermore, interference between D2D pairs is introduced only if the same resources are allocated to them. The classification of interference by the D2D is depicted in Fig. 5. We can divide the interference into three cases in terms of who is the interferer and who is the victim of interference.

Case 1 - Interference from the D2D communication to the cellular communication. In the UL direction (hereafter termed Case 1a), the interference is caused to the eNB, which receives data from its CUEs (see Fig. 5). In the given example, the eNB is disturbed by the DUE1 and the DUE4 that are transmitting data to the DUE2 and the DUE3, respectively. Hence, the interference at the eNB caused by the D2D communication could be expressed as:

$$\gamma_{eNB} = P_{D1} \times g_{D1-eNB} + P_{D4} \times g_{D4-eNB} \quad (1)$$

where P_{D1} and P_{D4} are the transmission powers of the DUE1 and the DUE4, respectively and g represents the link gains between particular DUEs. On the other hand, in the DL direction (Case 1b), the CUEs are the victims of the D2D interference as they are receiving data from the eNB at the same time as the DUEs and exchange data among themselves. The amount of interference is expressed, analogously to (1),

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

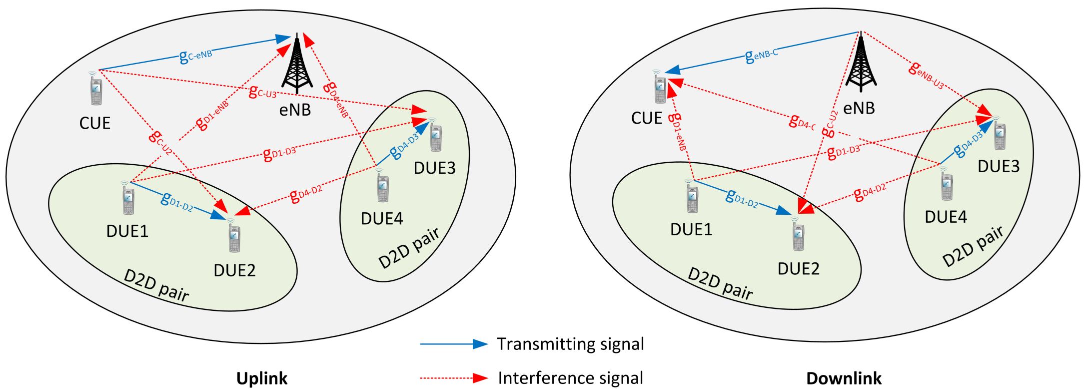

Fig. 5. Classification of interference in D2D communication.

as:

$$\gamma_{CUE} = P_{D1} \times g_{D1-C} + P_{D4} \times g_{D4-C} \tag{2}$$

Case 2 - Interference from the cellular communication to the D2D communication. In the UL direction (Case 2a) the interference to the D2D communication is generated by the CUE, which transmits to the eNB. Hence, the DUE2 and the DUE3 suffer from the interference and their performance may be degraded. The interference from the CUE experienced by both DUE2 and DUE3 in this case is expressed as:

$$\gamma_{DUE2} = P_C \times g_{C-D2} \tag{3}$$

$$\gamma_{DUE3} = P_C \times g_{C-D3} \tag{4}$$

where P_C is the transmission power of the CUE. In the DL direction (Case 2b), the interference also affects the DUE2 and the DUE3 but the source of interference is the eNB. The amount of interference can be expressed as:

$$\gamma_{DUE2} = P_{eNB} \times g_{eNB-D2} \tag{5}$$

$$\gamma_{DUE3} = P_{eNB} \times g_{eNB-D3} \tag{6}$$

Case 3 - Interference between D2D pairs. If more than one D2D pair is reusing the same radio resources, the additional concern is the interference generated between the DUEs. Regardless of the transmission direction, the interference is always caused by the transmitting DUEs to receiving DUEs in different D2D pairs using the same resources. In Fig. 5, the interference is caused by the DUE1 and the DUE4 to the DUE2 and the DUE3, respectively. The level of interference is expressed as:

$$\gamma_{DUE2} = P_{D4} \times g_{D4-D2} \tag{7}$$

$$\gamma_{DUE3} = P_{D1} \times g_{D1-D3} \tag{8}$$

From (1)-(8) it can be seen that the main factors affecting the amount of interference caused to the eNB, the CUEs, or the DUEs depends on the geometry of D2D pairs and the CUEs and on the transmission power of individual stations. The geometry expresses the mutual distance of the DUEs forming the D2D pairs and their distance from the CUEs and the eNB. In general, interference between cellular network and D2D communication is lower if the mutual distance of the D2D pairs and the eNB (in the case of the UL) or the CUE (in the case of the DL) is higher. At the same time, if the DUEs within the D2D pair are close to each other, the transmission power could be decreased proportionally to reduce interference. Moreover, the interference between two D2D pairs is also decreased if their distance increases.

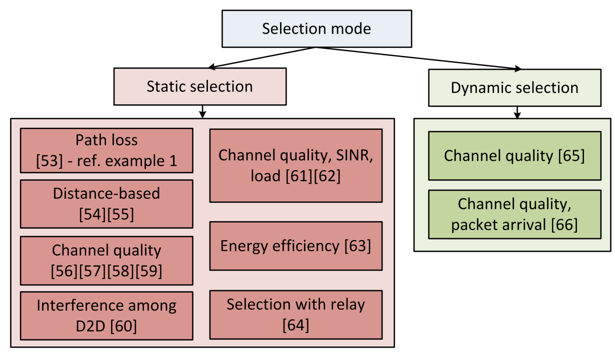

Fig. 6. Overview of mode selection techniques for D2D communication.

Interference poses a significant risk to both cellular and D2D communication. In this regard, the technical papers dealing with the interference problem are surveyed exhaustively in Section V.

IV. Selection of Communication Mode

Proper mode selection plays a crucial role in D2D communication. The reason is that it determines the potentials to increase the frequency reuse factor (spectral efficiency of the system) and, at the same time, it affects the amount of interference among the CUEs and the DUEs (or among the DUEs), as already explained in Section III-C2.

This section contemplates both static mode selection (Section IV-A) and dynamic mode selection (Section IV-B). A summary and comparison of all related papers are then tackled in Section IV-C). For easier orientation of readers, we present a high level overview of individual techniques and approaches for the mode selection published recently in Fig. 6. As can be seen in this figure, static mode selection has attracted much more attention than dynamic mode selection so far. Therefore, we devote more space to static mode selection.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

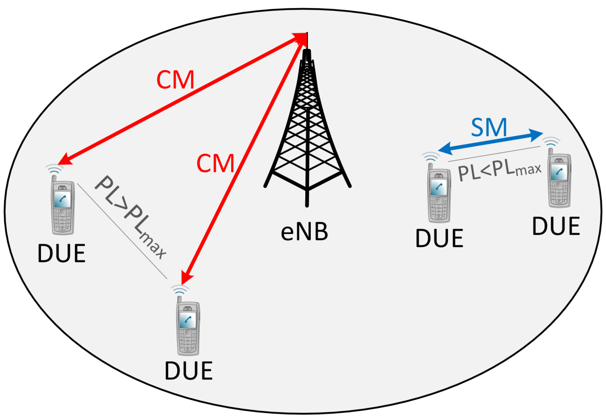

Fig. 7. Mode selection based on path loss [53] (reference example 1).

A. Static mode selection

This section firstly describes a simple reference example for mode selection based on path loss. We then discuss research dealing with mode selection using various metrics, such as distance, channel quality of D2D and cellular links, interference, load of the eNB, and energy efficiency.

Reference example 1: The path loss-based method can be considered as the simplest approach for selection of an appropriate mode for DUEs [53]. The basic principle is shown in Fig. 7. The communication through the eNB (using the CM) is established if the path loss (PL) between the DUEs is above a threshold PL_{max}. If the path loss between both DUEs is lower than the PL_{max}, D2D communication takes place (the SM is selected). Of course, the selection solely according to the path loss measurement is far from optimal as it does not reflect exact channel quality or interference issues.

1) Mode selection according to distance: Extension of the simple reference example 1 based on path loss metric towards more realistic assumptions considering real distance among the DUEs and the eNB is presented in [54] [55]. Mode selection according to the mutual distance of DUEs is taken into account in [54]. The study assumes use of either the DM or the SM. The authors propose an optimal D2D mode selection threshold in order to minimize the transmit power of the DUEs. From analytical and numerical results it can be observed that the optimal threshold is inversely proportional to the eNBs density and linearly increases with the path loss exponent. Hence, with an increasing number of eNBs, the CM becomes more favorable. The study also demonstrates that both the DM and the SM can improve the overall performance of the network with respect to the CM. However, mode selection between these two options is not considered in the paper. Moreover, the selection of mode is performed only according to the distance between the potential DUEs. The inaccuracy of distance derivation is a key aspect that is not addressed in the paper.

A proposal which selects the mode not only according to the distance between the DUEs as in previous papers, but also according to the distance to the eNB, is introduced in [55]. Potential D2D transmitters use direct communication only if the D2D quality is at least of the same quality as the cellular UL. The factor with an impact on the selection is so-called bias factor T_d, which regulates the distance from the eNB at which the DUEs can communicate directly via D2D communication. The results show that the proposal is superior to more simple distance-based mode selection in terms of the outage probability of CUEs, the intensity of admitted D2D links and the average transmit power of DUEs. Nevertheless, it still suffers from a lack of accuracy in the distance determination.

2) Mode selection according to channel quality: Another feasible improvement of reference example 1 is to exploit mode selection according to channel quality. In [56], the authors investigate the maximum benefit in terms of system capacity if D2D is enabled and which mode is the most appropriate. The most efficient mode is evaluated according to a sum rate calculated by means of the Shannon capacity formula. The analysis considers the CM, the DM (strictly half of the resources is used by the CUE and one D2D pair) and the SM (both DL and UL). The one that gives the highest sum rate is selected. The evaluations are performed for varying distance between the D2D pair and the eNB (parameter D) and between DUEs composing the D2D pair (parameter L). If the D is large, that is, if the D2D pair is far from the eNB, it is usually beneficial to use the SM because the interference to the eNB is low and efficiency is at its highest. Conversely, if the D is low, the SM using the UL resources or the DM is the most suitable option.

In [57], the authors also compare individual modes in terms of sum rate for the D2D communication that could be selected by the eNB. They assume the same modes as considered in [56] but, in addition, two optimization power control schemes are utilized for the SM. The first power control scheme is based on greedy sum rate maximization and the second one is rate constrained power control with priority given to the CUEs. Again, the performance is analyzed for different values of D and L distances that are similar to the ones used in [56]. It is shown that the scheme exploiting greedy sum rate optimization is beneficial for small values of D. Regarding the scheme prioritizing the CUEs, the performance is worse than in the case of fixed power but still significantly better than in the case of a single CUE without power control. Further, the study finds that if the power optimization exploiting the greedy sum rate is used, the SM is exploited in 95% of cases. Only in situations when the D2D pair is close to the eNB and the distance between the DUEs is large is the CM selected by the eNB. On the assumption that the CUEs have priority over the DUEs, the probability of the DM and the CM is increased and an SM between 60% (for small value of D) and approximately 90% (for large value of D) is used. Note that in the case of the CM or the DM, the maximum transmission power is used for both schemes. The work in [57] is further extended in [58] by applying sum rate optimization not only to the SM but also to the DM and the CM. It demonstrates that the optimal selection of the allocation mode can significantly outperform the reference example 1 based on path loss selection [53].

The disadvantage of [56] and [57] is that they use only simple scenarios when strictly only one D2D pair shares resources with just one CUE. The authors in [59] investigate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

how the system benefits if the D2D pair is able to reuse resources of more than one CUE (both for the DL and for the UL direction). It is shown that if the D2D pair is able to use resources of more CUEs, its performance is improved. However, the study considers a very simple scenario, where the CUEs are either located uniformly or close to each other (six CUEs are assumed).

3) Mode selection according to interference among D2D pairs: Further improvement of mode selection with respect to reference example 1 is represented by consideration of interference among D2D pairs [60]. The DUEs may either utilize the CM or the DM, which is shared by all the DUEs. Hence, interference among D2D pairs in proximity can occur. To solve this problem, the DUEs use a carrier-sensing threshold to determine their transmission mode as indicated in Fig. 8. While one D2D pair can use direct communication, the other one has to use the CM, since measured energy is above the threshold. Selection of the threshold impacts on the density of D2D pairs and the interference among them. As a result, the optimal threshold is found. The advantage of the approach is that the selection is done in a distributive manner and the signaling overhead is minimized. On the other side, the drawback of the study is that it considers a fixed distance between the DUEs of each D2D pair, which does not reflect the situation in a real network.

4) Mode selection according to channel quality, interference and load of the eNB: So far, the mode for D2D has been selected only according to the path loss, distance of the DUEs, interference among DUEs or quality of D2D link and the channel between the DUEs and the eNB. Nevertheless, the interference between the DUEs and cellular network has to be taken into account together with the load of the eNB (i.e., the number of resource blocks (RBs) used by the CUEs). The load influences the available capacity for D2D if the DM is selected (at higher load, fewer dedicated resources are allocated). An optimal selection algorithm considering all the above-mentioned aspects is proposed in [61]. Whereas the interference for the CUEs is obtained by means of the conventional method used for cellular communication, interference at

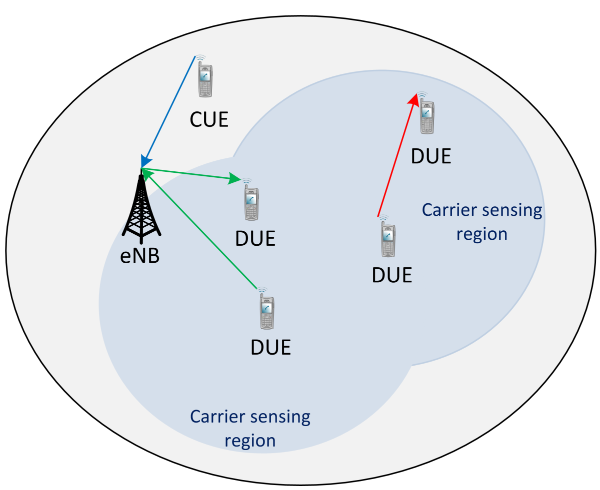

Fig. 8. The principle of mode selection according to interference among D2D pairs [60].

the side of a D2D pair has to be additionally determined. The procedure for optimal mode selection is processed as follows. First, the D2D terminals send probing signals to each other and estimate the received signal powers. Second, the terminals determine the amount of interference plus noise power both in the DL and in the UL. Third, this information is transmitted to the eNB, which sets the amount of resources in the DM depending on the load of the cell together with the maximum transmit power for individual allocation modes (CM, DM, SM). Fourth, the eNB estimates SINR and determines the data bit rate for all allocation strategies and the one offering the highest one is selected. Nonetheless, the optimal selection procedure generates a high amount of signaling overhead, which makes implementation in real networks questionable.

Another scheme considering interference and the load of the eNB is introduced in [62]. The main purpose is to select the mode with respect to the D2D throughput only. As a consequence, the throughput of the CUEs decreases with higher numbers of D2D pairs in the system. On the other hand, the throughput of DUEs increases with the number of D2D pairs.

5) Mode selection according to energy efficiency: In previous papers addressing the mode selection, the main objective is to improve capacity of the whole system. Another criterion for mode selection is energy efficiency, as assumed in [63]. The mode selection is based on a coalition game, where the DUEs cooperate in order to reduce transmission cost. Cooperation means that the DUEs within the same coalition use orthogonal resources and, thus, do not interfere with each other. In general, the DUEs can select from the CM, DM, or SM. The D2D users form coalitions for individual modes (three groups of users are formed) to get the benefit of coalitions in terms of lower transmission cost. The advantage of the proposed scheme is that it can react to changing situations in the network, i.e., to the situation when new D2D connections are created or terminated. Under such circumstances, existing D2D pairs can change the coalition from time to time if they can profit in terms of energy efficiency. However, mode selection according to energy efficiency should be done jointly with selection of capacity.

6) Selection with relay stations: Thus far, all studies have assumed a simple network consisting of the eNBs, the DUEs and the CUEs. How the mode selection is influenced by introduction of relays in the network is studied in [64]. The paper considers that the DUEs should use only the DM or the SM mode. The selection is done according to the SINR experienced between D2D pairs. The analysis is performed first for the scenario without relay nodes. If the SINR is sufficiently high, the SM mode is preferred, since the D2D pair is more resilient to interference caused by the CUEs in the UL. If the SINR of D2D is lower than a selected threshold, the DM mode is used for the D2D communication. In scenarios without relays, the SM mode is often selected only when the DUEs in the D2D pair are far from each other. If the relay nodes are introduced into the network, the CUEs select connection through the relay (denoted also as relay node) if they can reach the same capacity but with lower transmission power. Consequently, the implementation of the relay results in lower interference with the D2D, which can use the SM mode more often to increase the spectral efficiency of the system. The disadvantage of the paper is that it considers only a simple single scenario with one D2D pair and one CUE. Hence, the coexistence relay nodes and D2D communication should be studied in much more detail.

B. Dynamic mode selection

None of the previous studies take into consideration the dynamicity of the network, where channel quality between individual nodes may vary in time. In this regard, the D2D mode can also be changed during its operation. Hence, the solutions focusing on mode selection for the static scenario need to be extended to cope with the network's dynamic environment. This problem is partly addressed in [65], where the authors develop a framework that opportunistically performs mode selection under varying channel conditions. The study itself considers only two possible modes, the CM and the DM, utilizing either the UL or the DL cellular resources. The effectiveness of the dynamic mode switching is demonstrated by means of simulations. The results are compared with the CM and the case when the DM is always used regardless of the distance between the DUEs. It is shown that for all investigated distances between the DUEs, the average sum rate is always highest for the proposal. Although the paper addresses dynamic mode switching, it does not elaborate how often the more appropriate mode should be selected. Also, the paper does not consider the mobility of users but defines the best mode depending on the position of the DUEs within the cell.

Semi-static mode selection and dynamic selection between the CM, the DM, and the SM on a slot-by-slot basis is addressed in [66]. Semi-static mode selection is performed at the timescale of connection establishment/release and can save computational and communication resources. On the other hand, dynamic selection takes into account a dynamic packet arrival process and fast fading wireless channel. The paper shows that dynamic mode selection outperforms the semi-static approach. What is missing in the evaluation is how much signaling overhead is generated by the dynamic selection. The authors only discuss the possibility of reducing overhead by preferring semi-static selection if the gain with the dynamic one is above a certain threshold. In addition, the real UE mobility is not considered in this paper. However, it is expected to have a strong impact on dynamic mode selection.

C. Summary of mode selection approaches

The comparison of individual mode selection schemes and assumptions that have been considered are summarized in Table III. We have chosen several comparable criteria such as selection metric, main objective of the mode selection, assumed allocation modes, etc. Of the decision metrics considered in research papers, the most common one is the channel quality and SINR with the objective of maximizing system capacity [56][57][58][59]. Nevertheless, the selection of allocation mode can also be done with the purpose of minimizing the transmission power of D2D (one of the advantages introduced by D2D) and increasing the energy efficiency of the system [63].

In most of the studies, it is considered that if the SM is used exactly one D2D pair shares radio resources of just one CUE (e.g., [56] or [57]). Even though this assumption makes the selection less complex it does not always result in the optimal solution if there are more CUEs than D2D pairs. All studies except [65] and [66] assume that the selection is made in the static way and there is no mention of what would happen if the channel conditions change and vary in time. In such situations new D2D connections may be established during the communication or some DUEs may be switched back to the CM mode. This supports the need for future research in the area of dynamic mode selection to enable D2D for scenarios with mobile users or with fluctuating channel quality. Additionally, neither [65] nor [66] assume the mobility of users, which could lead to more complex solutions in selection of the proper mode.

Most studies consider the most common D2D scenario where both DUEs are attached to the same eNB (Scenario 1C). Only [54] and [55] consider that the DUEs composing one D2D pair can be located in different cells. Another critical point in most of the papers is consideration of very limited scenarios, which are far from real networks. The other interesting scenario if relay nodes are introduced in the network is analyzed in [64]. However, the impact of relay nodes on D2D communication should be tackled more thoroughly.

From the above studies we can conclude the following. From the performance perspective, the SM should be applied because of its high spectral efficiency and high frequency reuse factor. However, the mode selection depends on the geometry of the D2D pair reflecting its location with respect to the eNB and the mutual distance of the DUEs. The other important factor is whether the D2D pairs reuse the UL or the DL direction as this has a crucial impact on the interference level. To that end, we can draw several conclusions (summarized in Fig. 9) regarding the decision process. When the DUEs are far away from each other, the CM has to be used. If the DUEs are close to each other, the D2D communication is more easily established. Nonetheless, if the UL direction is reused and the D2D pair is close to the eNB, interference to cellular communication can still be significant and a less spectrally efficient DM has to be applied instead. At the same time, if the DUEs are close to the transmitting CUEs interference caused by cellular communication to D2D can forbid the use of the SM. As a result, the SM can be utilized only if the D2D pair is a sufficient distance from the eNB (interference to cellular communication is minimized) or the CUEs (interference to D2D communication is minimized). Similarly, in the DL direction, if the D2D pair is in the proximity of the CUEs, the SM is not a feasible option and the DM or even the CM is selected.

V. Interference mitigation

One of the major challenges in D2D communication is to avoid interference between the DUEs and the CUEs if the DUEs are reusing radio resources of the CUEs (i.e., if the SM

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

TABLE III
COMPARISON OF INDIVIDUAL METHODS FOR SELECTION OF ALLOCATION MODE WITH RESPECT TO SELECTED CRITERIA.

Paper	Selection metric	Objective	Allocation mode	D2D control	Dynamicity of selection	D2D scenario	Scenario	Sharing model	Mobility
[53]	Path loss	Maximize system throughput	CM, SM (UL)	Full	Static	1C	Multicell	Not specified	Low mobility (3 km/s)
[54]	Distance-based	Minimize outage probability and transmit power	CM, DM, SM (UL)	Full	Static	1C, 1D	Multicell	Not specified	No
[55]	Distance-based	Minimize outage probability and transmit power	CM, SM (UL)	Full	Static	1C, 1D	Multicell	Not specified	No
[56]	Channel quality	Maximize system throughput	CM, DM, SM (UL and DL)	Full	Static	1C	Multicell	1 D2D pair and 1 CUE	No
[57]	Channel quality	Maximize system throughput	CM, DM, SM (UL and DL)	Full	Static	1C	Single cell	1 D2D pair and 1 CUE	No
[58]	Channel quality	Maximize system throughput	CM, DM, SM (UL and DL)	Full	Static	1C	Single cell, multicell	1 D2D pair and 1 CUE	No
[59]	Channel quality	Maximize system throughput	CM, DM, SM	Full	Static	1C	Single cell	1 D2D pair and N CUEs	No
[60]	Int. among D2D	Maximize system throughput	CM, DM	Full	Static	1C	Multi cell	N CUEs, M DUEs	No
[61]	SINR, load	Maximize system throughput	CM, DM, SM (UL and DL)	Full	Static	1C	Single cell, Multicell	1 D2D pair and 1 CUE	No
[62]	SINR, load	Maximize D2D throughput	CM, DM (DL,UL), SM (DL, UL)	Full	Static	1C	Multicell	N CUEs and N DUEs	Mobile users (Random walk)
[63]	Energy efficiency	Minimize energy consumption	CM, DM, SM (UL, DL)	Loosely	Static	1C	Single cell	N CUEs, M DUEs	No
[64]	SINR	Maximize usage of the SM	DM (UL, DL), SM (UL, DL)	Full	Static	1C	Single cell	1 D2D pair and 1 CUE	D2D fixed, CUE semi mobile
[65]	Channel quality	Maximize system throughput	CM, DM (DL, UL)	Loosely	Semi-static	1C	Single cell	1 D2D pair and 1 CUE	No
[66]	Packet arrival, channel quality	Maximize system throughput	CM, DM (UL), SM (UL)	Full	Semi-static, dynamic	1C	Single cell	1 D2D pair and 1 CUE	No

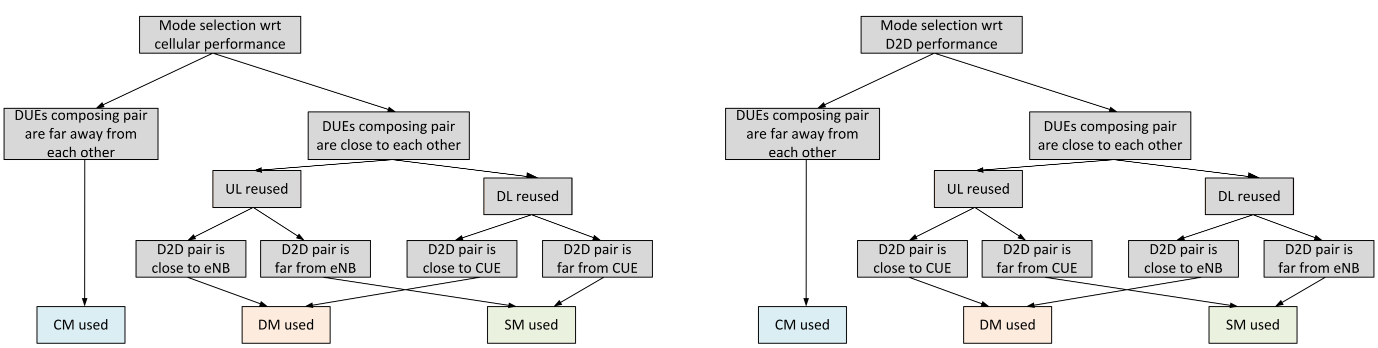

Fig. 9. Decision process for mode selection from cellular communication perspective (left) and from D2D communication perspective (right).

is applied). Of course, to avoid the interference completely, the eNB could dedicate extra radio resources that are exploited only by D2D (i.e., in the DM). Nevertheless, this option lowers spectrum efficiency. Interference may also occur among individual D2D pairs if their transmissions overlap in time and frequency. This section gives a comprehensive survey of the state of the art on the problem of interference among CUEs and DUEs as well as among the DUEs themselves. We divide individual technical papers according to interference scenarios (see section III-C3) as below:

- papers addressing the problem of interference from the D2D to the cellular communication (*case 1*)
- papers addressing the problem of interference from the cellular to the D2D communication (*case 2*)
- papers addressing the problem of both interference from the D2D to the cellular communication (*case 1*) and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

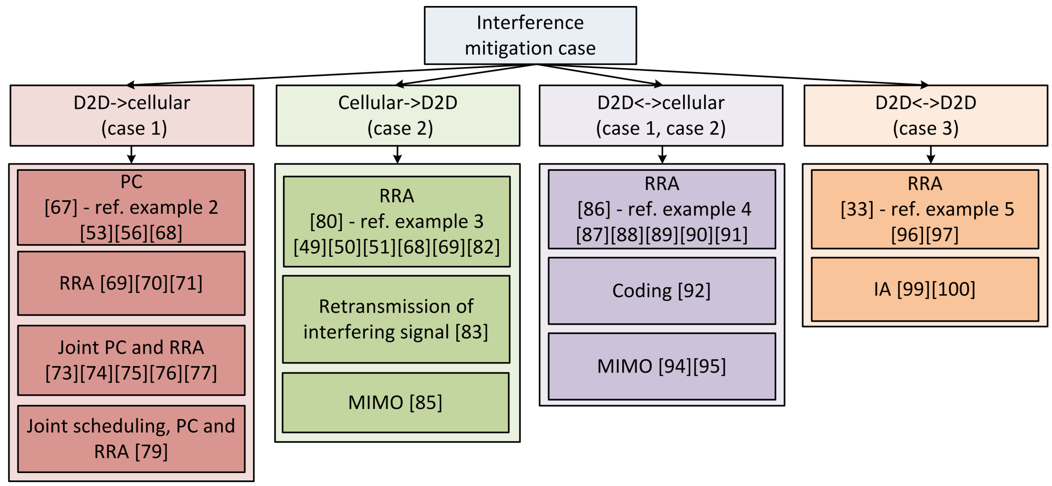

Fig. 10. Overview of interference mitigation techniques (PC = power control, RRA = radio resource allocation, IA = interference alignment).

interference from the cellular to the D2D communication (*case 2*)

- papers addressing the problem of mutual interference among D2D users (*case 3*)

Individual interference cases are tackled in the following subsections. Each subsection firstly introduces a reference example presenting basic solution to specific interference case (similar to Section IV) and, after that, it provides a thorough survey of techniques addressing a particular interference case. An overview of all techniques used for these interference cases is provided in Fig. 10.

A. Mitigation of interference from D2D to cellular communication

The mitigation of interference from D2D communication to cellular communication is the most important one, since the DUEs should not disturb the CUEs or the eNB.

Reference example 2: The most straightforward approach used for mitigation of interference from D2D to cellular networks is a power control. Interference mitigation can be achieved by reducing the DUE's transmission power [67]. The objective here is to set the power of the transmitting DUE

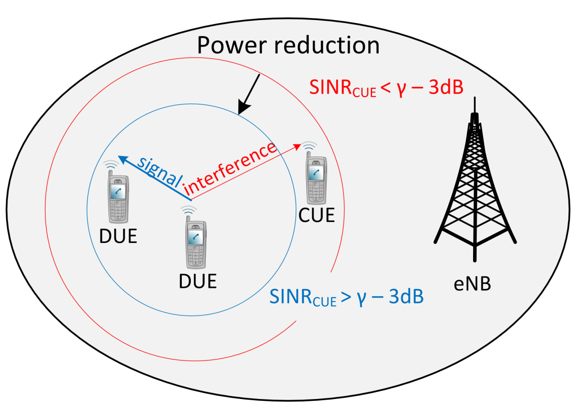

Fig. 11. Interference mitigation from D2D to cellular communication based on power control [67] (reference example 2).

such that the experienced SINR of the CUEs ($SINR_{CUE}$) is not degraded by more than 3dB (see Fig. 11 where γ stands for SINR without D2D). The D2D power reduction is evaluated for several distances between the D2D pair and the eNB and various distances between the DUEs. If the DL is reused by D2D communication, the reduction of transmission power fluctuates between -15dB and -20dB according to the position of the D2D pair and the CUE. If the UL is reused, the power reduction significantly depends on the distance from the eNB and it varies between -7dB and -30dB. The disadvantage of this simple power control is that the probability of D2D communication between DUEs may be very low due to low transmission power.

1) Power control techniques: A similar method as in the reference example 2 is proposed in [68]. Nevertheless, a difference is that the eNB sets the transmission power of the DUEs to achieve target SINR to minimize disruption to cellular communication (i.e., allowed SINR degradation is not defined contrary to [67]).

The applicability of conventional power control schemes for the purpose of D2D communication is analyzed in [53]. With respect to reference example 2, the authors consider: 1) fixed transmission power, 2) fixed SINR target, 3) open loop fractional power control applied in LTE systems, and 4) close loop fractional power control. It is demonstrated that closed loop power control achieves the best result in terms of SINR experienced by the UEs. Nevertheless, the closed loop power control scheme needs additional signaling overhead to adjust the power level at the side of the DUEs.

More sophisticated power control, when compared to reference example 2, is introduced in [56]. The power control is based on the cellular UL power control framework, mitigating the interference to the eNBs in the UL phase. The aim is to achieve a specified SINR target for the CUEs. In this regard, the proposed power control assumes that the eNB can reduce transmission power level of the DUE by means of back-off parameter B. The B is expressed as $B = P_1 g_1 / P_2 g_2$, where P_1 and P_2 are transmission powers of the CUE and the DUEs and g_1 with g_2 are the corresponding link gains. With a high

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

value of B, the interference to the CUEs can be significantly reduced, as the power of the D2D transmitter is considerably reduced. On the other hand, the distance between two DUEs is substantially limited at the same time. To limit this drawback, a power boost factor for the UL cellular transmission is incorporated to compensate for the interference from D2D pairs. If there is no D2D transmission, no power boosting is needed, since there is no interference. At the same time no power back-off has to be applied for D2D transmission if no CUEs use the same radio resources. The results show that with B set to 5dB, 95% of the CUEs experience degradation of SINR of less than 3dB. Although the power boost to the UL transmission increases possible distance between the DUEs, this distance is still limited.

2) Radio resource allocation (RRA) techniques: A different approach for mitigation of interference than power control is to use various RRA techniques. The advantage of RRA over power control can be seen in the fact that the transmission power of the DUEs is not restricted like in Fig. 11.

A simple method for mitigation of interference from the DUEs to the CUEs is suggested in [69]. The eNB calculates a tolerable interference level from the D2D transmission for each RB in the UL and broadcasts this information to the DUEs. According to this information, the DUEs use only those UL resources at which harmful interference does not occur. The results show that the performance of the CUEs improves from 2.65 Mb/s to 3.33 Mb/s. On the other hand, this is accomplished at the cost of a DUE throughput decrease from 3.02 Mb/s to 2.83 Mb/s.

A more advanced RRA method which effectively labels time slots for the DUEs and the CUEs is introduced in [70]. The CUEs are classified into two groups: near-far-risk and non-near-far-risk. In addition, the eNB defines shared time slots and cellular dedicated time slots. Whereas the CUEs from the former group can use both kinds of time slots, the CUEs belonging to the latter group access only cellular dedicated time slots without increase in interference caused by the DUEs. The eNB identifies which CUEs can be in non-near-far-risk (interference from the DUEs is not an issue) or near-far-risk (the CUE suffers interference from the DUEs) according to SINR values received by the D2D during their transmissions. Since the proposed technique implies quite significant overhead for cellular network, the authors consider loosely controlled D2D, where DUEs autonomously determine radio resource allocation. The performed simulations show the average throughput achieved by the CUEs and the DUEs depends on threshold value (from -5dB to 13dB). It is demonstrated that although the performance of the CUEs decreases, overall gain can be increased significantly, especially for high threshold values.

In [71] the problem of radio resource allocation to the D2D users is formulated as a mixed integer nonlinear programming (MINLP) problem. However, the MINLP problem is of high complexity and its implementation in real systems is not possible. The main issue here is that the MINLP problem cannot be solved within a short scheduling period of 1ms, which is considered in LTE(-A). Therefore, a heuristic greedy algorithm that can mitigate interference to the CUEs is employed instead. This algorithm uses the information on channel gain between the DUEs and the CUEs. The DUEs are allowed to reuse resources of those CUEs that have a high channel quality indicator (CQI) reflecting good channel quality. The channel quality is determined according to a predefined threshold. Thus, if both the CUEs and the DUEs achieve SINR higher than the threshold, the DUEs can use the same resources. The authors perform extensive simulations considering conventional scheduling algorithms for the CUEs such as Round Robin, Maximum Carrier to Interference Ratio and Proportional Fair [72]. It is shown that with D2D-enabled communication, normalized throughput is substantially increased. Unfortunately, the paper does not compare the performance of the proposed heuristic algorithm with the optimal solution to demonstrate the optimality of the proposed heuristic solution.

3) Joint power control and RRA techniques: To make the protection of cellular communication against D2D communication even more effective, power control can be combined with various RRA techniques. In [73] the mitigation of interference from DUEs to CUEs is accomplished by a combination of resource allocation and dynamic power control at the side of the DUEs. The first step is executed by the eNB in such a way that the resources are first assigned to the CUEs and the remaining resources are then allocated to the DUEs. If the demands of the DUEs are not met by this allocation, the eNB identifies resources that can be shared by both the DUEs and the CUEs. The interference is subsequently mitigated by dynamic power control of the DUE transmitters. The power control is done by the eNB, which determines channel gain between individual terminals in the first step. In the second step, the channel gain between the D2D pair and the CUEs is measured by the DUEs in the UL and the channel gain between the eNB and the D2D pair is determined by the eNB. The performance of the proposal is compared with the fractional power control scheme described in [53]. The simulation demonstrates that the CUEs experience on average 5.7dB higher SINR compared with [53]. Similarly, the DUEs reach, on average, 2.77dB higher SINR. The disadvantage of the approach is that the power control is managed centrally by the eNB, and the amount of overhead generated is not discussed in the paper.

Similarly to the study introduced in [73], joint power control and resource allocation are assumed in [74]. The main objective is to maximize the number of D2D pairs permitted in the system. The eNB determines the minimum transmission power of the D2D transmitter in order to guarantee a SINR threshold for D2D communication. If the required power of the D2D transmitter would cause interference to cellular communication, the D2D communication is not permitted analogous to reference example 2. The optimization problem is solved by an optimal Hungarian algorithm with complexity equal to $O(MN^2)$ (M is the number of the CUEs and N stands for the number of D2D pairs). To decrease the overall complexity of the algorithm, a suboptimal heuristic algorithm with complexity equal to $O(MN)$ is also developed. The proposed scheme (both optimal and heuristic) is compared with the random reuse algorithm. Although the number of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

permitted D2D pairs in the system can be increased with the proposed scheme, the increase is quite marginal (between 1% and 7%), for both optimal and heuristic algorithms. The disadvantage of the approach is its centralized nature, which will result in a significant overhead. Work based on [74] always assumes that the eNB first allocates data to the CUEs, and the authors in [75] extend the proposal by introducing a heuristic algorithm, which allocates radio resources to the DUEs and the CUEs jointly. Nonetheless, the complexity of the proposed heuristic algorithm is increased compared with [74] to $O(MN^2)$. On the other hand, the results show that this approach results in a higher number of permitted D2D pairs in the system compared with the heuristic algorithm presented in [74].

Two studies considering joint allocation of radio resources and power control in order to guarantee the QoS of the CUEs are presented in [76][77]. In [76] authors propose an optimal resource and power allocation algorithm. Although the DUEs reuse the resources already allocated to the CUEs, the CUEs should not be affected significantly by the D2D transmissions. Instead of using basic power control presented in reference example 2, the authors define an allowable throughput drop of the CUEs determined by a loss rate factor. Since the primary problem is non-convex, the authors solve it by means of less complex Lagrangian dual theory. The complexity of the algorithm is then $O(N(N+M)K)$, where N is the number of D2D links, M is the number of the CUEs, and K is the number of subcarriers. The complexity is, however, still high, especially for large networks. The results are compared with a water-filling algorithm [78] and improvement in performance demonstrated in the paper is between 28% and 50%.

The second QoS-based resource allocation scheme [77] takes into account the difference in QoS requirements of the CUEs and the DUEs. The DUEs use a simple power control mechanism, which limits the upper bound transmission power of the DUEs in order not to interfere with the CUEs. Hence, if the DUEs and the CUEs use the same RBs, the SINR of the CUEs should not be below the SINR target value (the same approach as proposed in reference example 2). Nonetheless, the difference between the proposed scheme and other proposals (e.g., [53][56]) is that the DUEs utilize the RBs that are the most suitable for their QoS rather than those preferred by the CUEs. The complexity of the proposed algorithm is roughly $O(K)$, where K is the number of RBs in the system. The results are compared with a reference scheme, where each D2D pair can reuse RBs of only one CUE. It is illustrated that the QoS of DUEs can be satisfied in most cases. However, the QoS of the CUEs is neglected in the evaluation.

4) Joint scheduling, power and resource allocation: To further enhance reference example 2 for the mitigation of interference caused by the D2D to cellular communication, power control and RRA techniques can be exploited jointly with the scheduling algorithm [79]. The optimization problem is solved by the Stackelberg game, where each CUE can share one orthogonal channel with one D2D pair. In this game, the CUEs act as the leaders of the game and owner of radio resources, and the D2D pairs are the followers charged with a specific fee in order to use the same resources. The paper addresses two optimization problems. The first optimization problem requires the leader to set a price to maximize his utility functions (expressed by throughput performance and the gain earned from the follower). The second optimization problem is to set the transmission power of the follower to maximize her utility function (represented by her throughput and what she pays for the channel). After that, a joint scheduling and resource allocation algorithm is proposed to achieve fairness among D2D pairs accessing radio resources of the CUEs. Consequently, the follower (DUE) has to pay an additional sum to access the same channel in two consecutive scheduling intervals. The proposed algorithm is of a sufficiently low complexity and is similar to that in [74] if M and N are low; it equals $O(MN)$. The results indicate that throughput of the D2D pairs increases with the number of the CUEs in the system, because more resources are available for the D2D pairs.

Summary: This subsection surveyed the papers solely addressing the problem of interference from the D2D to the cellular communication (i.e., *case 1* according to the classification introduced in Section III). In general, the most direct approach is to use simple power control at the side of the D2D transmitter [53][56][67]. The proposed power control techniques suggest lowering the transmission level of the DUE transmitter such that the CUEs' SINR is decreased only by a certain value (e.g., in reference example 2 degradation of 3dB is allowed [67]). The main weakness of power control approaches is that D2D communication cannot always be enabled because of restrictions on the power transmission level. In this instance the use of D2D in cellular networks can be enabled by smart radio resource allocation [69][70][71], joint power control and radio resources allocation [73][74][75][76][77], or even by joint scheduling, power and resource allocation [79]. These solutions are able to achieve more satisfying results in terms of system performance than simple power control techniques. The implementation cost of the proposed solutions, however, increases with the complexity of the proposed algorithm. In this regard, it is necessary to find the optimal trade-off for both a high system performance and a low complex algorithm.

B. Mitigation of interference from cellular to D2D communication

Although interference from D2D communication to cellular systems is primarily studied, as demonstrated in the previous section, the performance of the DUEs also plays an important role when enabling D2D communication.

Reference example 3: The interference from cellular to D2D communication can be easily solved by a distance-based resource allocation scheme [80]. The principle is simple: whenever any DUE in the proximity wants to communicate directly, it sends a request to the eNB. Subsequently, the eNB selects the resources of the CUE to minimize outage probability of the DUEs. The basic principle is depicted in Fig. 12 where the D2D pair shares resources with CUE2, which is in a sufficient distance from the pair. In contrast, sharing resources with CUE1 would result in strong interference. The advantage of the distance-based method is the reduced

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

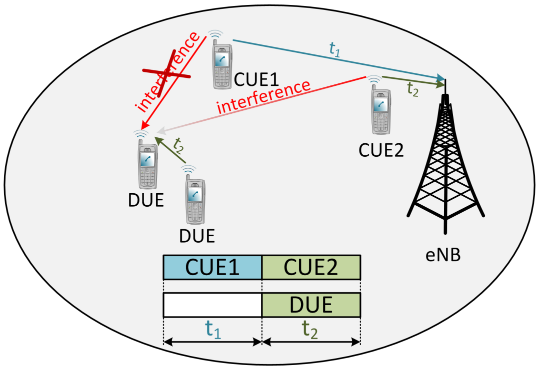

Fig. 12. Interference mitigation from cellular to D2D communication done by simple distance-based RRA [80] (reference example 3).

signaling overhead since the allocation of resources is done only according to the mutual distance of the CUEs and the DUEs, not according to CSI. On the other hand, it requires exact knowledge of the CUE and the DUE positions to allow the eNB to allocate resources appropriately. Unfortunately, the authors do not compare the amount of overhead saved by this approach and do not compare the proposal with any method based on collection of CSI.

1) RRA techniques: Similar method as assumed by reference example 3 is proposed also in [68] where the authors exploit multi-user diversity. To be more specific, the DUEs sense the radio spectrum during the UL transmissions and help the eNB to be aware of radio spectrum environment. Then, the eNB exploits this information in interference aware resource allocation to the DUEs.

Enhancement of the simple reference example 3 presented in [49] allocates resources not only according to distance but also channel quality. In the proposed procedure, the eNB broadcasts information on allocated resources of the CUEs that could cause interference problems for D2D in proximity. This information is obtained by the CUEs via listening to the control channel. Hence, the DUEs are able autonomously to schedule resources to avoid interference from close CUEs. The performed simulations compare the proposed scheme with the case where no interference cancellation technique for the DUEs is considered. It is shown that the throughput of the DUEs can be significantly improved. The extension of [49] to a multi-cell scenario (i.e., Scenario 1D), is addressed in [50]. Similarly to [49], the CUEs have to listen to signaling on the control channel and determine if any DUEs are in the vicinity. If the affected DUE is not in its own cell, the eNB has to exchange information with an adjacent eNB that acts as a serving station for the interfering CUE. The eNB can stop scheduling transmission of interfering CUE until D2D transmission is over. Similarly to the single cell case, the throughput of D2D in the multi-cell scenario is substantially enhanced. The disadvantage of the approaches introduced in [49] and [50] can be seen in the quite significant overhead compared with [80]. Moreover, both works assume that sensing of the control channel is reliable, which is not always the case.

An approach similar to that in [49] is proposed in [69] as well. First, the DUEs decode radio resources management information broadcast by the eNB. In LTE-based systems, this information is scrambled by the Radio Network Temporary Identifier [81]. Second, the DUEs measure the level of interference caused by the CUEs. This means the D2D pairs are able to determine which resources can be reused. The results show that the performance of the DUEs and the CUEs is improved from 3.02 Mb/s to 5.12 Mb/s and from 2.65 Mb/s to 2.9 Mb/s, respectively. The amount of overhead generated by broadcasting can be regulated by means of the trade-off between accuracy and overhead, as suggested by the authors.

An even more sophisticated RRA scheme with respect to reference example 3 is proposed in [51]. The authors introduce an interference limited area (ILA), where the coexistence of the DUEs and the CUEs is not allowed. The ILA is defined as the area in which the interference to signal ratio (ISR) from the CUEs to the DUEs is higher than a predefined threshold. In other words, if the CUE is within the ILA of the DUE receiver, the DUE cannot reuse those resources for its own benefit. The basic principle is illustrated in Fig. 13 where the D2D pair is not able to reuse resources of CUE_1 while the resources allocated to CUE_2 - CUE_N can be utilized. The size of ILA is determined by the eNB on the assumption that it knows the location of the CUEs and the DUEs. It is demonstrated that with increasing radius of ILA, the gain of D2D increases as well and the performance of the CUEs is affected only marginally. The disadvantage of this approach is again seen in the assumption that the eNB has to know the exact location of individual CUEs and DUEs, which can be especially problematic in indoor environments with poor GPS signal quality.

The mitigation of interference from the CUEs to the DUEs using graph coloring is considered in [82]. Similarly to reference example 3, the DUEs reuse resources of those CUEs that are sufficiently far away. In order to determine which CUEs are in the vicinity of a D2D pair, the DUEs have to be able to detect cellular transmission. Hence, at the beginning of the UL transmission period, the DUE listens for a short period (in the paper this is termed as a "quiet period") whether some CUEs close to the potential D2D pair are transmitting or not (see Fig. 14). Then, this information is sent to the eNB through the control channel. This notifies the eNB about

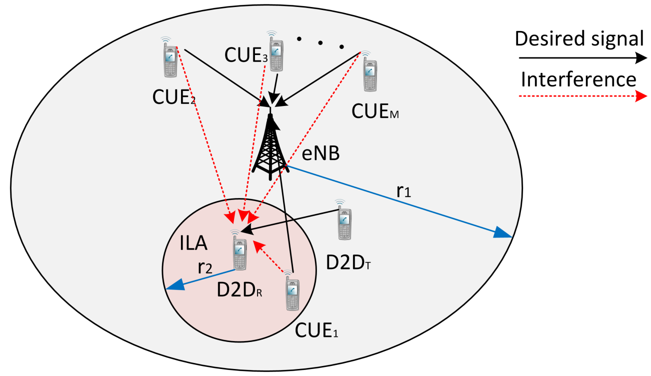

Fig. 13. The principle of ILA area restricting interference from the CUEs to the DUEs in the UL [51].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

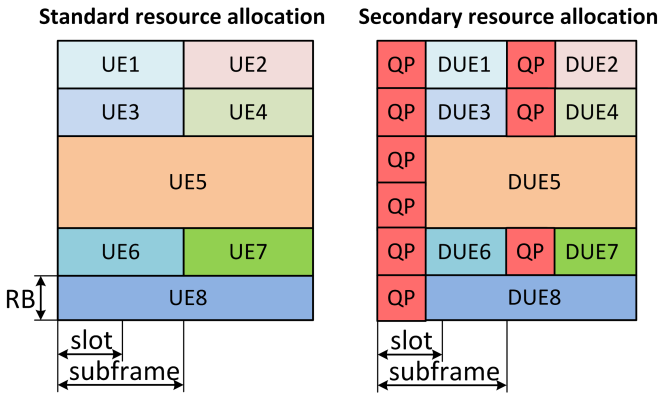

Fig. 14. Introduction of quiet period (QP) for D2D secondary resource allocation when the DUEs share the same radio resources with the CUEs [82].

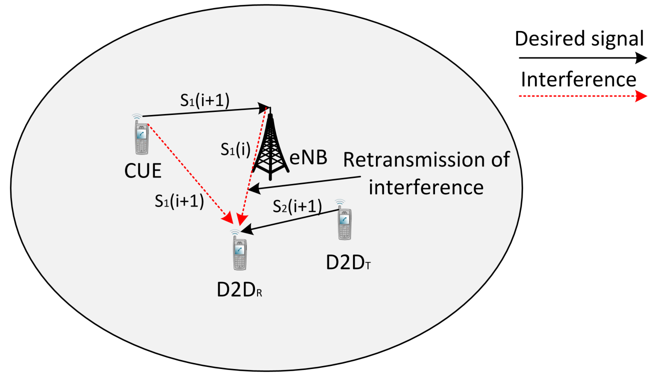

Fig. 15. The principle of retransmitting interference signal according to [83].

which DUEs can reuse which resources of the CUEs. In the paper, the interference among individual UEs is modeled by a node contention graph and resources for the DUEs are allocated by means of several possible algorithms, such as the greedy algorithm, random sequential algorithm and repeat random sequential algorithm. The results illustrate that the last algorithm outperforms the others. The disadvantage of the proposed approach is in reducing radio resources to the DUEs because of the QPs. This could be solved by smart use of QPs if really necessary.

2) Technique based on retransmission of the interfering signal: A novel technique mitigating interference from CUEs to DUEs by using retransmission of the interference signal is proposed in [83]. When compared to reference example 3, the proposed method does not need to schedule radio resources for the DUEs to mitigate interference. For the retransmission of the interference signal, the maximum ratio transmission (MRT) [84] scheme is used. It means, in the i+1 period, the receiving DUE obtains an interference signal of the i-th UL period from the eNB together with UL interference in the i+1 period from the CUE and with data signal from the i+1 UL period from the transmitting DUE (see Fig. 15). Then, the interference signal is demodulated with the interference cancellation technique (IC). In order to implement the proposed technique, the eNB has to be equipped with more than one antenna. Otherwise, the proposed technique cannot be used, since the eNB has to be able to transmit and receive signals simultaneously. The proposal is compared with other two interference cancellation techniques. The first technique is suitable for cases with low interference where the affected DUE demodulates signals directly by treating interference as a noise. The second technique is suitable for high interference when the signal is demodulated after the IC is performed. It is shown that the proposed scheme is most suitable for cases when interference is somewhere between light and heavy, whereby it outperforms the other two techniques in terms of an outage probability. The slight drawback of the proposal is that it requires redundant transmission of interference signals from the eNB to the receiving DUE.

3) Multiple Input Multiple Output (MIMO) technique: The previous mitigation techniques for interference from cellular networks to D2Ds assume only SISO (Single Input Single Output). The utilization of an advanced MIMO technique is considered in [85]. This study assumes application of MIMO transmission scheme only at the DL cellular transmission (i.e., at the side of eNB). To be more exact, the authors propose novel precoders for the DL cellular transmissions that constrain transmission from the eNB in order not to interfere with D2D communication. The proposed scheme can be enhanced by the closed loop techniques when lower rank transmission is used by the DUEs (usually only SISO transmission is applied). This allows an increased number of degrees of freedom for the eNB DL transmission and, thus, more efficient interference mitigation for the DUEs. The performance evaluation is done for two antenna configurations; 4x2 and 4x4. The simulations demonstrate the efficiency of MIMO as it can radically increase the SINR of D2D links (by approximately 15dB) whereas reduction of the CUEs SINR is only marginal (3dB). To implement the proposed scheme, the eNB has to be aware of interference channel state information, which implies additional signaling overhead. However, the authors assume that DUEs will be fixed or slow moving, and the updates can be done relatively infrequently.

Summary: This subsection demonstrated that the interference from the cellular to the D2D communication (interference *case 2* according to the classification introduced in Section III) is not neglected by current researchers. We can see that the methods for interference mitigation differ slightly compared with *case 1*. The power control technique at the side of cellular communication is not always a feasible option in *case 2*. In the DL direction, limitation or adaptation of the transmission power of the eNB is not as convenient as coverage of the network could be heavily affected. In the UL direction, power restriction at the side of the CUEs is not suitable either since CUEs should be able to communicate with the eNB. To that end, the most common approach is to use various techniques using a distance-based algorithm [80] that can be enhanced by channel-based radio resource allocation [49][50][51] or to employ a graph coloring technique [82]. Basically, the objective of these techniques is to allocate CUEs' resources to the DUEs at sufficient distance. Another approach taken into consideration in the current literature is the advanced coding

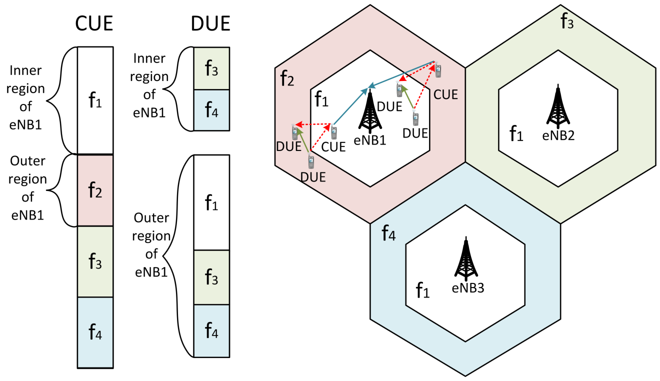

Fig. 16. Mitigation of mutual interference between D2D and cellular communication by means of FFR [86] (reference example 4).

technique [83] or MIMO [85].

C. Mitigation of mutual interference between D2D and cellular communication

Whereas the previous two sections focused on papers proposing mitigation of interference from D2D communication to CUEs or vice versa, this section draws attention to studies addressing both interference problems at the same time.

Reference example 4: Mutual interference between the CUEs and the DUEs could be solved by utilization of the fractional frequency reuse (FFR) approach [86]. The FFR itself divides the whole frequency band into four sub-bands (f_1, f_2, f_3, and f_4). The first sub-band utilizing the frequency $f1$ is reused in every cell by all CUEs located in the so-called inner region (see Fig. 16). The other three sub-bands are used by the CUEs in the outer region. The whole concept of the FFR is further exploited by the D2D pairs. If the DUEs are in the inner region of the eNB, they can use sub-bands allocated for the outer region of other eNBs (i.e., if the DUE is in the cell of eNB1 using f_2 for the outer region, the DUE in the inner region of the eNB1 can use f_3 and f_4). If the DUEs are in the outer region, they can also reuse the f_1 frequency from the inner regions. The results of the FFR approach are compared with the scheme, which assigns resources to the DUEs randomly. It is illustrated that the FFR achieves higher SINR values for both CUEs and DUEs. However, the problem with the FFR is that it does not utilize resources efficiently, as only $1/4$ of the whole bandwidth can be assigned to one user. Another problem is the need for rough estimation of users location, which can influence performance, but it is not evaluated in the paper.

1) Radio resource allocation (RRA) techniques: The method presented by reference example 4 is extended in [87], where, besides the inner and outer regions, the authors specify accessible and reusable regions. Thus, only the DUEs in the accessible region can reuse radio resources of the CUEs in the reusable region. To be more specific, if the D2D pair is located in the inner region, the accessible region is the area within the outer region of adjacent cells. Similarly, if D2D pair is in the outer region, it can reuse resources of the CUEs in the inner region. The results show that both the CUEs and the DUEs can achieve a higher SINR compared with [86]. However, the problem with localization of users is even more significant because of the creation of smaller regions.

A more spectral efficient approach than reference example 4 is proposed in [88]. The interference from CUEs to DUEs is solved by means of ILA (similarly to [51]). The minimization of interference from DUEs to CUEs is accomplished by definition of the coverage area around D2D pairs denoted here as Z_1. If the CUE is located inside this area, whose radius is equal to r_1, the D2D cannot use radio resources allocated here for any CUE. Similarly, the second area, Z_2, is defined around the receiving DUE with radius r_2. If some CUEs are within it, the D2D pair cannot reuse these resources. Nonetheless, the problem with the proposal is that interference from CUEs to DUEs in the DL is caused solely by the eNB, which transmits to the CUEs. Hence, this interference is not affected by the positions of the CUEs but it only depends on the channel between the eNB and the receiving DUE.

More complex techniques for mutual interference mitigation between D2D and cellular communication adopt advanced mathematical tools such as game theory or graph theory. Game theory exploited for mitigating interference between DUEs and CUEs is described in [89]. The paper uses a sequential second price auction to optimize the overall sum rate of the system. During the auction process, the total resources are divided into a certain number of sub-bands whereby each sub-band is already assigned to one CUE. These sub-bands are then auctioned off one by one among the DUEs during each round as a second price auction (the bidder with the highest price obtains the sub-band on payment of the second highest price). The auction process continues as long as there is at least one sub-band to be used by a D2D pair during the auction. The value of a certain sub-band is quantified by the gain of channel capacity acquired by individual DUEs. In other words, the DUEs use the sub-bands that result in the highest utility function. The utility function is defined as the total value of the sub-bands auctioned off minus the total payments. The performance of the proposed technique is compared with a random allocation scheme (DUEs randomly select the sub-bands). The proposal shows its superiority, especially with increasing number of D2D pairs (e.g., for 2 D2D pairs, the sum rate is improved from 26 bit/(sHz) to 78 bit/(sHz)). However, this performance is at the cost of complexity, which is $O(n^m)$, where n represents the number of bidders and m is the number of sub-bands to be allocated.

The work in [89] is further extended in [90], where the aim is to decrease the complexity of the previous sequential second price auction intended for allocation of radio resources to D2D communication. The approach considered in this paper is based on a reverse iterative combinatorial auction (ICA). The reason why the ICA is called "reverse" is that spectrum resources that are composed of the set of RBs compete to obtain business as the bidders whilst D2D pairs are auctioned off as goods in each round. The auction game proceeds as follows. The auctioneer (eNB) announces the initial price and the bidders (resources) submit their bids at the current price. As long as the demand exceeds the supply or vice versa, the auctioning process continues. The authors proved that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

their approach is cheat-proof, converges properly with finite iterative steps and its complexity is not NP-hard and equals $O(n(2^m - 1) + t)$, where t stands for the total number of iterations. Although the complexity of the algorithm is less than that in [89], it could be quite significant for higher values of m and n.

A graph theory-based scheme adopting interference-aware graph-based resource allocation is proposed in [91]. The objective of this paper is to allocate radio resources to the DUEs and the CUEs in such a manner that the system sum rate is maximized. The interference relationships among the DUEs and cellular communication are formulated as an interference-aware graph. First, the interference-aware graph is constructed according to the network topology. Each vertex in the graph has three attributes; 1) the link attribute, distinguishing whether the vertex represents the DUE or the CUE, 2) the resource attribute, containing information on SNR value for individual RBs, and 3) the cluster attribute, representing assignment of RBs to individual vertexes (DUE or CUE). Second, the suboptimal algorithm allocating RBs to individual UEs in the system is performed. The optimal solution needs an exhaustive search of all allocation possibilities and it is very complex and does not scale well with an increasing number of UEs in the system. Hence, the suboptimal algorithm with the complexity $O(\frac{(M+N+1)(M+N)K}{2})$ is proposed (M is the number of the CUEs, N represents the quantity of the D2D pairs, and K is the number of RBs.) The results show that the suboptimal algorithm achieves almost the same results as the optimal one in terms of the system sum rate and it significantly outperforms the greedy orthogonal sharing scheme (allocation of non-overlapping resources to CUEs and DUEs).

2) Coding technique: Besides various RRA techniques, mutual interference mitigation between the D2D and cellular communication can be based on a rate splitting [92]. The paper exploits the Han-Kobayashi scheme [93], where the transmitting data are split into private and public parts. Whereas the private part is decodable only by intended receivers (DUE receiver), the public part is decodable by the UEs (both the DUEs and CUEs), which are subject to the interference. After that, the DUE and the CUE receivers perform a best-effort Successive Interference Cancellation (SIC) algorithm mitigating interference caused by public signals. Note that the SIC represents a special case of Han-Kobayashi scheme rate splitting, where only the public message is sent. In addition, the paper derives the optimal rate splitting factor to optimize the utility function (in the paper it is the sum rate of the system). The advantage of this approach is that interference between the CUEs and the DUEs can be canceled without decreasing transmission power. The performance of the rate splitting technique is compared with the schemes proposed, e.g., in [53] or [58]). It is demonstrated that the proposed scheme outperforms these conventional methods.

3) MIMO techniques: The utilization of advanced antenna techniques such as MIMO, beamforming or IC can have an important edge in interference mitigation. Whether beamforming or the interference cancellation technique at the eNB is more suitable for interference mitigation is studied in [94]. The former is more advantageous for the CUEs, as the DL signal from eNB is more focused and SINR is increased as well. On the other hand, the latter is more profitable for D2D pairs as the IC mitigates interference at the receiving DUE. The paper suggests selecting one of the techniques dynamically depending on the SNR value at the eNB. It is demonstrated that the system performance is maximized if beamforming is used at a low SNR whereas the IC technique at a higher SNR results in a higher sum capacity.

The MIMO technique and beamforming are also considered in [95]. Whereas previous works tend to assume accurate CSI (e.g., [58][68]), the proposed method here is based on a more realistic estimate of CSI. The estimation is done by linear minimum mean-square error. Then, beamforming is applied at the transmitting DUE to direct signals only towards the null of the estimated channel between the transmitting DUE and the eNB. This means interference can be sufficiently mitigated at the side of the eNB. In order to estimate the channel between the transmitting DUE and the eNB and the channel between the CUE and the DUEs receiver, N slots are assigned for a training sequence. The optimal length of the training sequence is numerically evaluated. In general, a longer training sequence results in a more accurate CSI but less time is assigned for the D2D transmission. The results show that the D2D throughput can be enhanced by proper selection of N. Nevertheless, the paper does not compare the results with other proposals to show the benefit of the proposed optimization.

Summary: This subsection offers an overview of research works that address mutual interference between the D2D and cellular communication (i.e., *case 1* and *case 2* according to the classification introduced in Section III). The current research is mainly focused on diverse RRA techniques, such as FFR [86][87] or spatial allocation [88]. The exploitation of game theory [89] or graph theory [91] can be seen as a powerful tool for interference mitigation. These, however, are often characterized by high complexity and suboptimal and less complex algorithms need to be developed. The current research aiming at mitigation of mutual D2D and cellular communication does not completely disregard the use of advanced coding [92] or advanced antenna techniques [94].

D. Methods for mitigation of mutual interference among D2D pairs

So far, we have surveyed papers solving the interference problem between the D2D and cellular communication. This section further focuses on papers where the DUEs and the CUEs use orthogonal radio resources (i.e., no interference occurs among the DUEs and the CUEs), but DUEs use non-orthogonal resources.

Reference example 5: Interference among D2D pairs can be mitigated by allocation of spatial, frequency, and time orthogonal resources [33]. A spatial orthogonality is accomplished in the following manner. The eNB creates groups of DUEs based on their transmission power and estimation of their geographical location (in Fig 17 D2D pair 1 and 4 create one group and D2D pair 2 and 3 compose other group). Then, the mutual interference of close DUEs in one group is solved by assigning orthogonal resources to them. The orthogonal

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

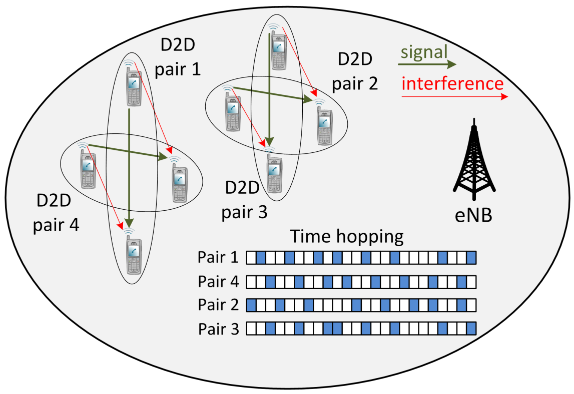

Fig. 17. Mitigation of mutual interference among D2D by spatial, frequency and time orthogonality [33] (reference example 5).

resources are allocated to the DUEs semi-persistently for a large timescale (up to 120 frames) based on long-term measurement in order to reduce signaling. However, interference among the groups may occur, so a time hopping approach is proposed. The eNB applies random sequence offsets at regular times to resources scheduled semi-persistently. This means any interference between the two groups lasts for a shorter time (e.g., in Fig 17 D2D pairs 2 and 4 have allocated some time intervals orthogonally even if they belong to different groups). The results indicate the ability of the time hopping technique to increase SINR.

1) RRA techniques: The time hopping approach proposed in reference example 5 is not able to solve the interference problem completely. Hence, the proposal in [96] tries to further increase spatial reuse for the DUEs. The objective is maximally utilizing the same resources by the DUEs that do not interfere with each other (i.e., DUEs that are sufficiently spatially distant). Since the CUEs and the DUEs are assumed to use orthogonal resources (the DM is considered), the aim is to maximize the number of RBs allocated to the CUEs while satisfying the requirements of DUEs. Similarly to [33] the proposal tries to minimize signaling and computational overhead by utilizing a distributed resource allocation scheme. Consequently, the eNB allocates RBs to the DUEs in a centralized manner with a slow timescale while the DUEs decide on their transmission powers and modulation and coding scheme in a distributive manner with a fast timescale. The D2D links can use the same resources if the probability of interference among them is lower than a specific threshold. The results are compared with [71] and show that the proposal can increase network throughput up to 45% for 25 D2D pairs in the system.

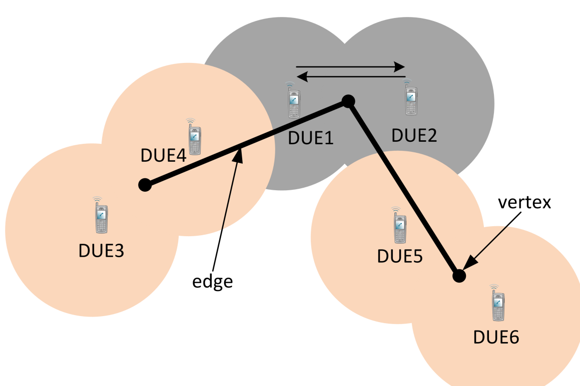

Fig. 18. Allocation of resources to D2D pairs according to [97].

Sharing the same radio resources by several D2D pairs is also tackled in [97]. With respect to reference example 5, the authors exploit the graph coloring approach. The DUEs that are not interfering with each other are assigned the same color in the created graph. The D2D pairs are assumed to be interfering if at least one of the DUEs is located within the transmission zone of the other DUE. The transmission zone is circular in area and it is determined by transmission power (see Fig. 18). The D2D pairs of the same color make so-called D2D reuse groups. The DUEs within the same reuse group use the same radio resources to increase spectral efficiency. The paper further addresses the problem of assigning available radio resources (represented by RBs) to individual D2D reuse groups. This is done either by opportunistic or fair assignment. In the former case, the RBs are scheduled for reuse groups in such a manner that overall throughput is maximized. The latter case tries to satisfy the QoS for all D2D pairs fairly. The drawback of the paper is that it assumes the transmission zone to be circular in area. In the real scenario, the circular area is notably deformed by the obstacles.

2) Interference alignment technique: As an efficient way to improve spectral efficiency, the interference alignment (IA) method can be used [98]. The IA technique is exploited in [99][100] focusing on D2D communication and interference mitigation. In [99], several D2D pairs are grouped together and share a fraction of radio resources to gain an extra degree of freedom offered by the IA. Each IA group is limited to three D2D pairs in order to reduce the complexity of the IA precoding process. While reference example 5 assumes that D2D groups are selected according to transmission power and geographical location, authors in [99] proposes three different grouping algorithms; position-based, channel-based, and distance-based. The results show that even though conventional point-to-point transmission outperforms the IA techniques in terms of bit error rate (BER), a higher throughput is achieved through the IA because of the additional degree of freedom. Regarding individual grouping schemes, the channel-based grouping results in a higher D2D throughput.

Similarly, in [100], the IA technique is exploited together with a clustering of the DUEs to achieve even tighter reuse of radio resources than in [99]. The main idea of the paper is first to group DUEs in several clusters within one cell (see Fig. 19). Clusters are formed from transmitting DUEs that are close to each other but individual clusters are sufficiently spatially separated. Thus they can fully reuse all dedicated resources allocated to them by the eNB. The cluster-forming is based on Fuzzy C-Means Clustering [101]. Within each cluster the D2D pairs are further formed into IA groups (comprising three D2D pairs), where the IA technique is utilized to increase spectral efficiency. The paper considers two grouping schemes: the channel-based and the distance-based as in [99]. It is demonstrated that the channel-based grouping scheme slightly outperforms distance-based schemes in terms

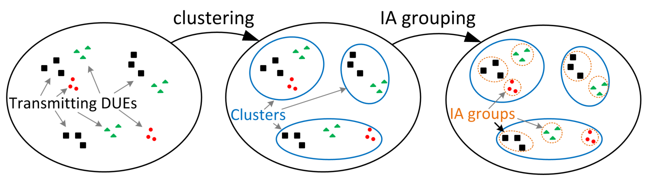

Fig. 19. Cluster and IA group forming according to [100].

of D2D throughput. Unfortunately, the authors do not show how the joint clustering and IA perform with respect to the simple IA presented in [99].

Summary: This subsection surveyed the papers where individual D2D pairs share radio resources among themselves and mutual interference has to be addressed as well (i.e., *case 3* according to the classification in Section III). The most common method uses an RRA technique increasing spatial reuse for D2D pairs [33][96]. Another effective means to improve spectral efficiency of the system is to use the IA technique described in [99] or to combine IA and clustering principles to shrink reuse distance of D2D pairs even more [100].

E. Summary of approaches to interference mitigation

This section gives an overview of a wide spectrum of methods and techniques for suppressing interference caused by D2D communication coexisting with cellular communication. The comparison of individual methods described in this section is summarized in Table IV (note that the priority in Table IV means that users (DUEs or CUEs) have higher priority in the system).

From the papers surveyed in this chapter, we can observe that the methods for interference mitigation differ slightly, depending on the case to be solved. For example, power control techniques are exploited mostly for interference mitigation from the D2D to cellular communication [53][56][67]. The important fact regarding power control is that its applicability is strongly dependent on two essential factors: 1) distance of the D2D pair from the CUEs (in case the DL is reused) or the eNB (in case the UL is reused) and 2) the mutual distance of the DUEs creating the D2D pair. If the D2D pair is far from the CUEs (or the eNB) and, at the same time, the DUEs are close to each other, lessening of power control is not especially significant and QoS to the DUEs can easily be guaranteed. In the opposite case, if the D2D is relatively close to the CUE (the eNB) and the distance between them is large, further decreasing the DUE's transmission power could result in significant degradation of QoS or, in the worst case, in preventing the use of D2D communication at all. As a result, simple power control at the side of the DUEs is applicable only if the D2D pair are close to each other and/or distance from the CUEs or the eNB is sufficient. If the use of power control is out of the question because of the above-mentioned restrictions, radio resource allocation techniques can be used. The allocation of radio resources to the DUEs and the CUEs can be done either by the eNB itself (if full control is used) [71][51][80] or by the DUEs (if loose control is applied) [70][49][50]. Since the eNB needs to know the exact CSI of all involved links in the case of full control, methods considering loosely controlled D2D should be preferred thanks to the lower signaling overhead. On the other hand, the loose control approach is problematic from the operator's perspective because of the loss of control over communication and management of the network. Another way to avoid mutual interference among the CUEs and DUEs by means of radio resource allocation is to utilize the FFR as proposed in [86][87]. This method, however, significantly decreases the system's spectral efficiency.

A standalone power control algorithms or radio resource allocation algorithms are not always able to cope with interference issues. Consequently, more sophisticated joint power control and radio resource allocation algorithms are proposed in the literature. Nevertheless, the most tangible problem here is how to guarantee low complexity of more intricate proposed algorithms while achieving performance close to the optimum. In addition, some of the proposed algorithms are centralized and they can result in excessive signaling overhead [73][74].

To make the protection against interference even more efficient, joint optimization of power allocation, resource allocation and scheduling can be exploited as introduced in [79]. The other possible method for interference mitigation, besides the above-mentioned radio resource management techniques, is utilization of advanced antenna techniques such as MIMO, beamforming, or IC [85][94][95]. These, however, require multiple antennas at individual nodes.

As in Section IV, most of the papers consider the D2D scenario when both DUEs are under coverage of the same cell (Scenario 1C). Only one paper [50] focuses on Scenario 1D, where DUEs creating one D2D pair are attached to different eNBs. Nevertheless, some of the papers take into account interference introduced by neighboring cells and inter-cell interference management (e.g., [50]). In addition, a large majority of papers assumes that one D2D pair can reuse the resources of just one CUE. Although this assumption significantly reduces the complexity of proposed interference mitigation solutions, it also limits the possible benefits of D2D communication. Similarly to the papers discussed in Section IV, a few studies assume the mobility of the users [53][73][69] and the majority consider a static scenario. The papers considering users' movement do not address the problems related to mobility management such as the change of interference pattern because of varying conditions or how power control and radio resource allocation should be updated if the DUEs are mobile.

VI. Power consumption and energy efficiency

One of the merits introduced by D2D communication is a possibility to reduce power consumption of the UEs (prolong battery life). This section is divided between papers analyzing energy efficiency and papers targeting minimization of power consumption. An overview of individual papers focusing on energy efficiency is given in Fig. 20.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

TABLE IV
THE COMPARISON OF INDIVIDUAL INTERFERENCE MITIGATION TECHNIQUES FOR SEVERAL SELECTED CRITERIA.

Paper	Technique	Interference case	D2D reuse direction	Priority	D2D Control	D2D scenario	Scenario	Sharing model	Mobility
[53]	PC (fixed, open loop)	1a	UL	CUEs	Full	1C	Multicell	N/A	Low (3 km/h)
[56]	PC (power back-off)	1a, 1b	UL, DL	CUEs	Full	1C	Multicell	1 D2D pair, 1 CUE	No
[67]	PC (simple power reduction)	1a, 1b	DL or UL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[70]	RRA (labeling of time slots)	1b	DL	CUEs	Loosely	1C	Single cell	N CUEs, M DUEs	No
[71]	RRA	1a, 1b	UL or DL	CUEs	Full	1C	Multicell	1 D2D pair, 1 CUE	Low (5 km/h)
[73]	RRA, PC	1a, 1b	UL and DL	CUEs	Full	1C	Multi cell	1 D2D pair, 1 CUE	Prob. rand. walk
[76]	RRA, PC	1a	UL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[77]	RRA(QoS-based), PC	1a	UL	CUEs	Full	1C	Multicell	1 D2D pair, N CUEs	Low (3 km/h)
[74]	RRA, PC	1a	UL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[75]	RRA, PC	1a	UL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[79]	Joint scheduling, PC, RRA (Stackelberg game)	1a	UL	DUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[49]	RRA	2a	UL	DUEs	Loosely	1C	Single cell	N CUEs, M DUEs	No
[50]	RRA	2a	UL	DUEs	Loosely	1D	Multicell	N CUEs, M DUEs	No
[51]	RRA	2a	UL	DUEs	Full	1C	Single cell	1 D2D pair, M CUE	No
[80]	RRA (distance based method)	2a	UL	DUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[83]	Retransmission of interference signal	2a	UL	DUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[85]	MIMO	2b	DL	DUEs	Full	1C	Multicell	1 D2D pair, 1 CUE	No
[86]	RRA (FFR)	1b, 2b	DL	No priority	Full	1C	Multicell	N/A	No
[87]	RRA (FFR)	1a, 2a	UL	No priority	Full	1C	Multicell	1 D2D pair, 1 CUE	No
[69]	RRA	1a, 2a	UL	No priority	Loosely	1C	Multicell	1 D2D pair, 1 CUE	Low (3 km/h)
[89]	RRA (auctioning)	1b, 2b	DL	CUEs	Full	1C	Single cell	1 D2D pair, , 1 CUE	No
[90]	RRA (auctioning)	1b, 2b	DL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[88]	RRA	1b, 2b	DL	CUEs	Full	1C	Multicell	1 D2D pair, N CUEs	No
[91]	RRA (graph theory)	1b, 2b	DL	No priority	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[82]	RRA (Graph Coloring)	1a, 2a	UL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[68]	RRA, PC	1a, 1b, 2a	UL and DL	CUEs	Full	1C	Multicell	1 D2D pair, 1 CUE	No
[94]	MIMO (Beamforming, IC)	1b, 2b	DL	No priority	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[95]	MIMO (beamforming)	1a, 2a	UL	CUEs	Full	1C	Single cell	1 D2D pair, 1 CUE	No
[92]	Rate splitting	1, 2	DL	No priority	N/A	1C	Single cell	1 D2D pair, 1 CUE	No
[33]	RRA (time hopping)	3	UL	No priority	Hybrid	1C	Multicell	N/A	No
[96]	RRA (maximizing spatial reuse)	3	UL, DL	No priority	Hybrid	1C	Single cell	N/A	No
[97]	RRA (graph coloring)	3	UL, DL	No priority	Full	1C	Single cell	N/A	No
[99]	IA	3	N/A	DUEs	Full	1C	Single cell	N/A	No
[100]	IA	3	N/A	DUEs	Full	1C	Single cell	N/A	No

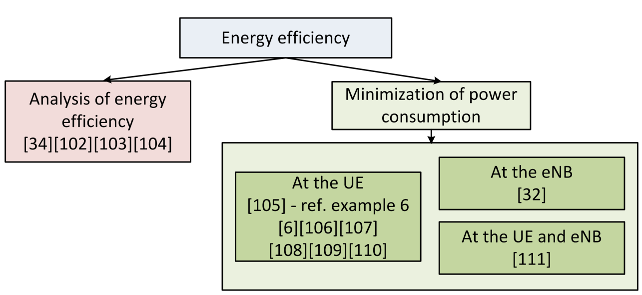

Fig. 20. Overview of papers dealing with energy efficiency (PC = power control, RRA = radio resource allocation, MS = mode selection).

A. Analysis of energy efficiency

The energy perspectives of D2D communication are contemplated in [34]. The paper analyzes several aspects with an impact on energy consumption if D2D is enabled. The first aspect is the effect of D2D control. Network-assisted D2D communication (i.e., full control of the DUEs) is much more energy-efficient than autonomous communication (i.e., loose control), since D2D discovery and communication phases are controlled by the network (usually by the eNB). Hence, there is no need to send beacons consuming additional energy in order to find other devices in proximity. Another aspect influencing the energy efficiency of D2D communication is the selection of duplex mode (TDD, FDD). Typically, TDD is more efficient and thus more suitable for D2D. The authors also perform some simulations to evaluate the energy saving potential of the D2D. As expected, the energy efficiency is higher if the communicating DUEs are close to each other. Also, it is shown that the interference from the eNB has a great impact on the energy efficiency of D2D communication.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

A higher interference by the eNB results in a lower energy efficiency. Further, an impact of the mode selection (DM or SM) on energy efficiency is studied in the same paper. It is demonstrated that the DM results in a higher energy efficiency for the CUEs whereas the SM is more advantageous for the DUEs.

The D2D energy efficiency is also addressed in [102] and [103]. The aim of [102] is to analyze energy efficiency when the data flows are not generated as a full buffer but arrive with Poisson distribution determined through stochastic modeling. The paper investigates energy efficiency for two allocation modes: the SM (in the paper denoted as full reuse strategy) and the DM (presented as an orthogonal sharing strategy). In addition, two topologies differing in the CUE and the D2D pair geometry are considered. It is shown that if the interfering D2D pair and the CUE are at a sufficient distance, interference is also low and, consequently, throughput achieved by SM is higher. Moreover, the DM is shown to be more energy-efficient. If interfering users are close to each other, the DM seems to be a more suitable alternative in terms of both spectral and energy efficiency (i.e., the same conclusions as in [34]). Similarly to [102], the paper [103] analyzes the energy efficiency of the system, where D2D communication is enabled for different D2D pairs and CUE geometry. All three allocation modes (DM, SM, and CM) are considered. For individual strategies, an energy-efficient power allocation scheme is discussed. Again, it is concluded that the energy consumption increases with the distance between communicating DUEs. It is shown that the CM is efficient only if the distance between the DUEs is close to the cell radius. Otherwise, the SM is the most efficient. The DM is preferable in terms of energy consumption only if the DUEs are close to the eNB. At medium and large distances, the SM significantly outperforms the DM.

Investigation of trade-off between energy efficiency and spectral efficiency for both the DUEs and the CUEs is performed in [104]. The maximization of energy efficiency is done by means of a distributed resource allocation algorithm, which exploits nonlinear fractional programming. To guarantee convergence of the proposed algorithm the number of iterations is limited to L_{max}. The complexity of the proposed algorithm is $O(I^d_{i,dual} I^d_{i,loop} K)$ where $I^d_{i,dual}$ is the number of iterations required to solve the dual problem (i.e., maximization of energy efficiency of the DUEs and the CUEs), $I^d_{i,loop}$ stands for the number of iterations to achieve convergence, and K represents the number of cellular links. The results show that the maximal spectral efficiency results in relatively low energy efficiency and vice versa.

B. Minimization of power consumption

Generally, maximization of D2D energy efficiency is achieved by minimization of overall power consumption. In this regard, we can classify the studies dealing with minimization of transmission power at the side of the UEs [6],[105]-[110], eNBs [32] or both [111]. Before surveying various approaches in this area, we present a reference example for minimizing UE power consumption.

Reference example 6: The most logical option for achieving lower power consumption is to introduce some kind of power control. In [105], the objectives of the power control are to; 1) minimize sum power at the UEs (both the DUEs and the CUEs), 2) reach a target SINR (γ_{tgt}) for the UEs, and 3) guarantee specified sum rate c_s. The authors propose an iterative transmit power and power loading optimization. The optimal SINR target is determined using Lagrangian optimization, which is not feasible in practice. In this regard, suboptimal heuristic algorithm based on greedy iterative algorithm is proposed as well. The performed simulations evaluate how significantly the transmission power can be decreased if D2D is used instead of conventional cellular communication. It is shown that the average sum power can be significantly reduced for D2D communication. The suboptimal heuristic scheme performs slightly worse if the distance of the CUE from the eNB is small. The drawback of the paper is that the simulation scenario is quite simple with only one CUE and one D2D pair.

1) Minimization of power consumption at the UE: The authors in [106] analyze two conventional power control strategies with the objective of minimizing the power consumption of the UEs. The first one is analogous to reference example 6 as the aim is to minimize total transmitted power while a fixed minimum target SINR is ensured for all users. The process is done iteratively on a channel inversing principle, where the UEs experiencing good channels are allocated less power and vice versa. The second power control algorithm works with varying SINR targets. Whereas the UEs with high transmission power have a low SINR target, the UEs with low transmission power have a high SINR target. Hence, this principle encourages the UEs to decrease their transmission power. As demonstrated in the paper, the second approach is more suitable from the energy consumption perspective.

The enhancement of power setting with respect to reference example 6 is proposed in [107]. Besides the power control, the authors assume mode selection. Hence, the DUEs can switch from SM to CM if beneficial for the system. Similarly as in [106], the authors consider both fixed SINR setting and adaptive SINR setting and perform exhaustive simulations. The bottom line of simulation results is that the utilization of mode selection further decreases power consumption. In addition, the adaptive SINR setting slightly outperforms the fixed SINR setting. The important fact is that the proposed algorithm is of low complexity.

The power control presented in reference example 6 can be also enhanced by joint power and resource allocation as assumed in [108]. The optimization problem is solved by reverse ICA (also used in [90] for mitigation of interference among DUEs). The DUEs and their transmission powers are considered as items and the resources of the CUEs are assumed to be the bidders competing for the items. The proposed algorithm proceeds as follows: 1) the DUEs are grouped into all possible packages and the auctioneer (eNB) sets an initial price, 2) bidders (i.e., resources of the CUEs) compete for the DUEs as long as all DUEs are auctioned off. The proposed reverse ICA is compared with the optimal energy efficiency (giving the upper bound of energy efficiency). It is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

demonstrated that the energy efficiency increases significantly with the number of DUEs in the system. The advantage of the proposal is that it achieves similar performance to the exhaustive optimal algorithm and its complexity is reduced. On the other hand, the drawback is that only a simple simulation scenario is considered with up to four D2D pairs, which does not fully convince the reader of its merits.

An even more complex solution to minimize power consumption at the UE compared to reference example 6 is proposed in [6], where joint optimization of mode selection (CM, DM, SM), resource allocation, and power assignment are assumed. The paper introduces an optimal centralized approach, which proves to be a strong NP-hard problem impractical for implementation in real networks. Consequently, a distributed suboptimal mode selection and resource allocation scheme are proposed. In this scheme, each eNB selects the proper mode and RBs for each device (DUE as well as CUE). The overall objective is to reach Nash equilibrium, in which no player gains any reward when changing strategy. Since the Nash equilibrium is not guaranteed if the SINR target is fixed, the resource allocation can be unstable. This problem is solved by the introduction of an algorithm for load control (LC). The LC minimizes the SINR target of those users that causes the greatest interference and instability in the system. The simulations investigate the impact of power consumption on the distance of the CUE from the eNB (R_{UE}) and the mutual distance between DUEs (R_{D2D}). It is demonstrated that the suboptimal algorithm performs similarly to the optimal one as long as the distances R_{UE} and R_{D2D} are less than 200m. However, as explained in [5], the R_{D2D} should be larger to maximize the profit from D2D. Hence, the proposed suboptimal algorithm should reflect this fact.

Another way to achieve low energy consumption besides the above-mentioned power control-based approaches is to exploit clusters [109]. The UE is selected to be part of the cluster if overall power consumption is decreased. In the UL direction, the energy consumption is composed of energy consumed by the cluster members transmitting to the cluster head, by the cluster head during reception of these transmissions, and finally again by the cluster head, which transmits data to the eNB. In the DL direction, the power is consumed at the side of the cluster head by reception of data from the eNB, by retransmission to cluster members and, of course, at the side of cluster members during reception of this transmission. Since the condition in the UL and the DL may be different, one DUE can use a different cluster for the DL and UL transmission in order to preserve the battery. The proposed method is divided into two steps. In the first step, each DUE is connected to the eNB (i.e., each cluster is formed by one DUE). In the second step, denoted as coalition formation, individual single DUEs subsequently form clusters on condition that by participating in the cluster they allow the power consumption of the UE to be reduced. Note that the cluster head of each cluster is the UE with the lowest energy consumption on the link to the eNB. The paper proves by means of simulation that the energy consumption can be significantly reduced if the proposal is used. A similar concept is introduced in [110]. The paper considers MBMS services in LTE, where real-time video streaming is broadcast by the eNB. Neither [109] nor [110] consider the impact on the overall capacity of the network. Nevertheless, capacity and energy efficiency should be addressed jointly, especially in the case of more than one hop transmission, since the same data are sent more than once.

2) Minimization of power consumption at the eNB: In the previous subsection, the target for power consumption is solely the UE. Conversely, in [32] minimization of total DL transmission power and overall power consumption at the side of the eNB are considered. The D2D communication is allowed only if the DUEs are close to each other and if they are far away from the eNB. In this case, the eNB would have to transmit with high power to satisfy their QoS. After the subcarrier and bit allocation is done, the transmission mode is subsequently selected. For subcarrier and bit allocation, two existing algorithms presented in [112] and [113] are considered. Further, the authors propose the heuristic algorithm, which selects D2D communication if the total transmission power of the DUEs is lower than in the case of the conventional CM. It is demonstrated that the transmission power reduction is more significant if more UEs can communicate directly. In addition, overall reduction of the power consumption for the eNB is accomplished. However, the power transmission consumption is only a small part of overall power consumption regarding the eNB [114] and this should also be considered in the study.

3) Minimization of power consumption at the UE and the eNB: The saving of power at both the UE and the eNB is assumed in [111]. The authors propose joint operation of D2D communication and green cellular networking. The idea is to form cooperative clusters of D2D users that share content. To that end, one DUE is selected as a cluster head, which communicates with the eNB at long range (i.e., cluster head acts as a relay for other devices as described in III-B). Then, the received data from the eNB are multicast to other DUEs within the cluster at a short distance. The DUEs are added to the cluster if the energy consumption in coalition is lower than the sum of the individual energy consumptions of the coalition members (a similar approach is described in [63]). In addition, the cluster close to the cell boundary can be served by individual eNBs in such a manner that some eNBs can be temporarily switched off (see Fig. 21). Nevertheless, the eNBs can be turned off only if only a small number of UEs is currently connected to the network.

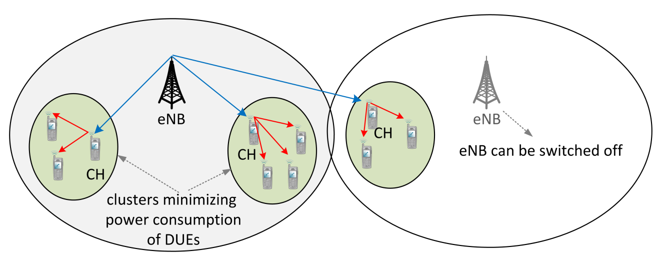

Fig. 21. The principle of D2D clustering method combined with green networking [111] (note that CH stands for a Cluster Head).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

TABLE V
THE COMPARISON OF INDIVIDUAL STUDIES FOCUSED ON POWER CONSUMPTION AND ENERGY EFFICIENCY FOR SELECTED CRITERIA.

Paper	Objective	Technique	D2D reuse direction	Criteria of efficiency	D2D scenario	Target nodes for saving	Scenario
[34]	Analysis of energy efficiency	N/A	UL	bps/Joule	1C	UE	Multicell
[102]	Analysis of energy efficiency	N/A	UL	bits/Joule	1C	UE	Single cell
[103]	Analysis of energy efficiency	Power control	UL	bits/Hz/W	1C	UE	Single cell
[104]	Increase energy efficiency	Distributed resource allocation algorithm	UL	Average energy efficiency [bits/Hz/J]	1C	UE	Single cell
[105]	Minimization of overall power	Power control	UL	N/A	1C	UE	Multicell
[107]	Minimization of power consumption	Power control	UL	N/A	1C	UE	Multicell
[106]	Minimization of power consumption	Power control	UL	N/A	1C	UE	Single cell
[108]	Increase energy efficiency, minimize power consumption	Joint resource and power allocation	UL	bits/(s*Hz*W)	1C	UE	Single cell
[6]	Minimization of overall power	Joint allocation mode selection, radio allocation and power allocation	UL	N/A	1C	UE	Multicell, Single cell
[109]	Reduce energy consumption	Cooperative cluster formation	DL, UL	Joule	3B	UE	Single cell
[110]	Reduce energy consumption	Cooperative cluster formation	DL	Normalized energy [-]	3B	UE	Single cell
[32]	Minimization of power in DL, overall power consumption of eNB	Joint allocation mode selection and resource allocation	DL	N/A	1C	eNB	Single cell
[111]	Reduce energy consumption	Cooperative cluster formation	DL	Joules (UE), No. of switched on eNB	3B	UE, eNB	Multicell

C. Summary of approaches focusing on energy efficiency

The current trend regarding mobile communication is to achieve high energy efficiency of the network (green networking). The D2D communication can be considered as a means to accomplish this goal. The comparison of individual papers dealing with energy efficiency of D2D communication is provided in Table V).

The important factors with an impact on energy efficiency are how the D2D communication is controlled and how DUEs are discovered [34]. To be more precise, network-assisted D2D discovery is much more energy-efficient than autonomous discovery. In addition, the duplexing method has an impact on D2D energy efficiency and TDD is superior to FDD.

Most contemporary studies are focused on the minimization of power consumption at the side of the UE (i.e., the minimization of transmission power in the UL). The reason is that the battery life of mobile devices is very limited. The power savings described in the current literature are often achieved by power control [105][106]. Another means to ensure energy efficiency is to exploit power control together with mode selection [107], joint resource and power allocation techniques [108] and joint mode selection with power and/or radio resource allocation [6][32]. In addition, an interesting way to make power savings is through the clustering concept [109][110][111]. The DUEs are grouped into cooperative clusters in order to decrease their transmission powers.

The minimization of power in DL is not completely disregarded by current research as the policy of major mobile operators is to decrease overall power consumption of the networks [32][111]. This approach is feasible especially if the DUEs are far from the eNB and power requirements are too high. However, the power expended on data transmission is negligible compared with other eNB components [114].

In the above-mentioned papers, the following could be observed. In order to make D2D communication energy-efficient, several conditions have to be guaranteed [34][102][103]. First of all, the DUEs communicating directly need to be close to each other, so transmission power can be decreased sufficiently. Hence, the SM is preferred if DUEs constituting one D2D pair are too far from each other. Moreover, the location of a D2D pair with respect to the eNB plays an important role. In general, the higher interference observed from the cellular transmission, the less energy-efficient the system is. This means that the DM can show higher energy efficiency if the D2D pair is close to the eNB and orthogonal resources are allocated to D2D and cellular communication. Otherwise, the SM outperforms the DM in terms of energy efficiency. This suggests that energy efficiency should be investigated together with a mode selection algorithm (see Section IV) and interference mitigation technique (Section V).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

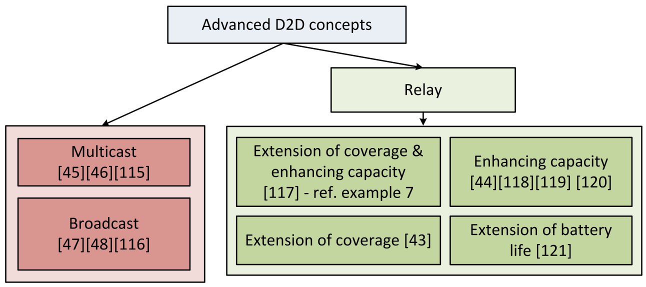

Fig. 22. Overview of papers focusing on advanced topology concepts.

VII. ADVANCED TOPOLOGY CONCEPTS EXPLOITING D2D COMMUNICATION

Introduction of D2D communication can serve not only conventional communication principles such as simple direct communication between two devices but it could also be exploited for multicast/broadcast communications or for relay purposes. This section surveys the papers focusing on these concepts. Each subsection addresses a specific type of advanced topology concept as shown in Fig. 22.

A. *Clustering multicast/broadcast concept*

In case of clustering multicast/broadcast concept, several DUEs in vicinity may create a cluster. For both multicast as well as broadcast, the DUE selected as a cluster head needs to have relay functionality as explained in section III-B. Problems and solutions related to multicast and broadcast are described in the following two subsections.

1) Multicast: The clustering multicast concept introduced by D2D communication enables, for example, efficient sharing of files among multiple DUEs in proximity (Scenario 3A). The multicast can save radio resources required for content delivery, since there is no need to send shared files to the eNB in the UL and then retransmit them to individual devices in the DL. The challenges of multicast transmission are to guarantee reliable transmission, to select a proper cluster head, or to manage interference efficiently.

The challenges of reliable multicast transmission and how to implement D2D clustering in an LTE-A system are considered from a standardization perspective in [45]. The most important factor is to guarantee correct reception of the content by all devices within a cluster. Using different rate transmissions depending on the channel quality of individual receivers is complicated and impractical. Hence, the feasible option is to adapt data rates according to the device experiencing the worst channel quality. For error correction of improperly received data packets, HARQ (Hybrid Automatic Repeat Request) feedback is handled by joining NACK (Negative Acknowledgment) feedbacks of every receiver into a common feedback region. The paper also considers how RLC (Radio Link Control) feedback is handled for multicast transmission. In addition, features such as cooperative retransmission for D2D multicast transmission are introduced here.

The main objective of the authors in [46] is to decide whether to form a cluster or to use the conventional cellular mode for sharing the files. The required service data rates are determined by the SINR of the weakest D2D link within the cluster (similarly as in [45]). In this regard, the transmission power of the cluster head ($P_{t(DUE-CH)}$) is set to such level that the SINR of the DUE (γ_{DUE}) with the weakest quality can receive data correctly. In Fig 23, the transmission power of the cluster head is adjusted with respect to DUE3. The drawback of both the above-mentioned papers is that the transmission depends on the worst channel quality of individual receivers and optimal cluster selection is not described properly.

The other important aspect regarding multicast communication is how to manage interference. This problem is addressed by authors in [115]. The interference from the DUEs to the CUEs is solved by power control applied to the DUEs. The aim is to set the minimal power of the transmitting DUE so data can be successfully received by all receiving DUEs. Interference control from the CUEs to the DUEs is solved by allocation of RBs. The paper proposes two allocation schemes: full set allocation (FSA) and subset allocation (SA). The authors define instantaneous SINRs between the CUEs (in this case considered as interferer) and individual DUEs that are part of the multicast group. Since there are multiple receivers of the D2D communication, the optimal allocation of RBs with respect to the weakest D2D link is found. Although the full set allocation method is the optimal one, it is too complex. Therefore, the authors also propose the SA method, which considers that each RB has a tolerable interference level. The eNB is able to calculate and also update the allocation for D2D transmission for each RB and create an interference value list. The list is then broadcast to the DUEs that facilitate radio resource management. The proposed scheme surpasses the fractional power control in [57] in terms of total throughput.

2) Broadcast: The other application of the D2D clustering concept is retransmission of multimedia services broadcast by the eNB (Scenario 3B). This way, network resources can be saved notably since DUEs within the cluster are able to retransmit data received from the eNB. The most important challenges needed to be addressed are how to efficiently retransmit incorrectly received data from the eNB, to minimize necessary signaling caused by HARQ, and to mitigate interference.

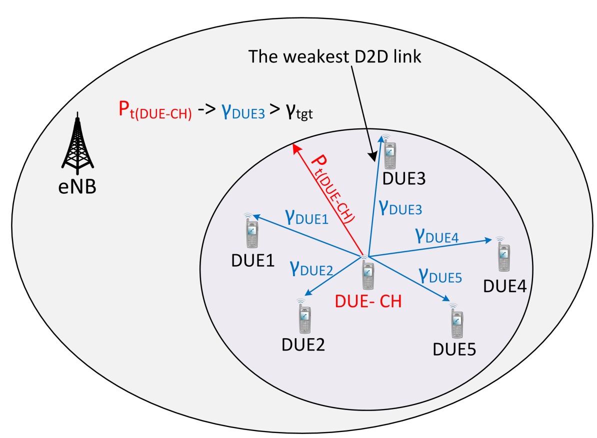

Fig. 23. The principle of cluster head power setting according to [46] (CH=Cluster Head).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

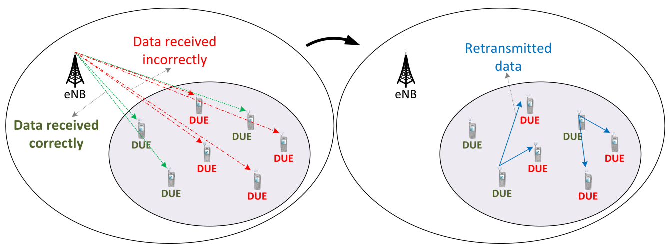

Fig. 24. Example of optimal selection of DUEs for retransmission of data within a cluster according to [47].

The first challenge is addressed in [47]. To be more specific, if one or several DUEs are not able to decode received information correctly, the DUEs which received the data correctly and which belong to the same cluster retransmit the data. The authors address the challenge of selecting the proper number of retransmitting DUEs to achieve maximal resource utilization and to find proper retransmission routes. An example of the situation is shown in Fig. 24. In the left part of Fig. 24, one DUE is selected to retransmit data to all DUEs with NACK acknowledgment; in the right part of Fig. 24, two DUEs are selected to retransmit data.

While proposal in [47] focuses on transmission efficiency, reduction of signaling in the UL caused by HARQ process is the main objective of [48]. The authors propose novel compressed HARQ feedback mechanism for broadcast communication. The foremost assumption in the paper is that if several devices in close proximity receive broadcast packets from the eNB, the packets do not have to be received correctly by all of them. In the conventional HARQ mechanism, all devices have to send feedback to the eNB (ACK/NACK message). To save radio resources, the devices that are close to each other can create a cluster (how the cluster is created and the cluster head selection are not specified in the paper but could be based on [46]). The simulation results illustrate that the proposed method is able to decrease HARQ feedback error probability.

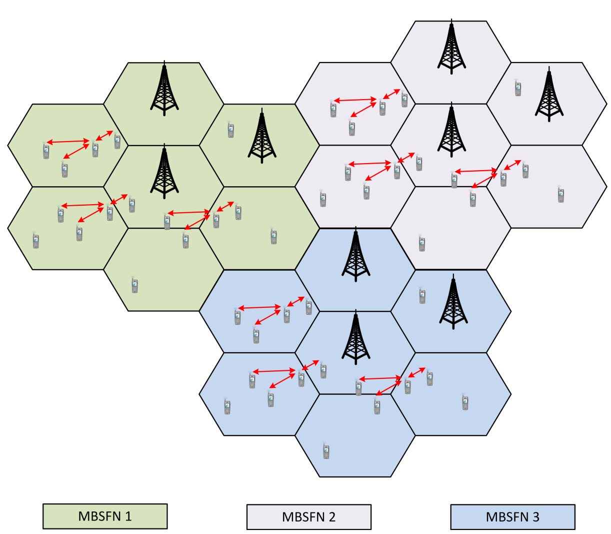

Fig. 25. The principle of EMBMS proposed in [116].

Both [47] and [48] focus on the situation when the DUEs within the cluster receive data incorrectly and no pre-emptive steps are considered (e.g., how to minimize incorrect reception in the first place). The efficient way how to accomplish that is to minimize interference as proposed in [116] where an evolved multimedia broadcast multicast service (EMBMS) incorporating a D2D multicast group is considered. The eNBs belong to different single frequency networks (SFNs) as indicated in Fig. 25. Within each SFN, the same content is broadcast at the same frequency. To avoid interference from the D2D multicast groups to the CUEs, the D2D groups use radio resources of the CUEs in other SFNs. The interference among D2D multicast groups within the same SFN is solved by a round robin resource sharing algorithm. This is done by determination of the safe reuse distance depending on the number of D2D multicast groups. The paper itself, however, does not specify whether the D2D multicast group retransmits data from the eNB or how the group of DUEs communicate among themselves.

B. Relay concept

Another concept exploiting D2D communication is based on enhancement of the UE's functionality by means of relaying. In general, the relay concept can be exploited to extend coverage (Scenario 2C) and/or to enhance capacity (Scenario 2D) or to extend battery life. This section firstly describes a representative reference example and, after that, surveys individual technical papers is delivered.

Reference example 7: A relay concept wherein the DUE serves as a relay and can both extend coverage and improve system capacity is described in [117]. The paper describes protocol architecture and evaluates data and signaling routing if the DUE uses the relay for connection to the eNB. If the DUE is in coverage of the eNB and uses the relay to enhance capacity, only data are relayed and signaling is exchanged directly between the relay and the eNB (see Fig. 26). On the other hand, in the coverage extension scenario (the DUE is not under coverage of the eNB), both data and signaling have to be relayed by the relay because the DUE cannot reach the eNB (Fig. 26). The assignment of the relay to the conventional UE in order to improve capacity is based on end-to-end throughput (if the direct path can offer higher end-to-end throughput, the

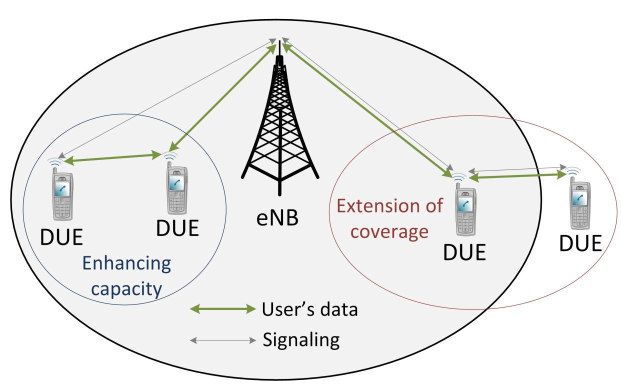

Fig. 26. D2D relay concept [117] (reference example 7).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

relay is not used). On the other hand, in the coverage extension mode, the DUE with the highest channel quality to the eNB is selected as a relay.

1) Extension of coverage: Reference example 7 does not exactly specify how the relay is selected if more DUEs can be used for this purpose. This problem is tackled in [43] where the authors propose an algorithm for relay selection based on graph theory. All DUEs requesting relay and all DUEs which could play a role in the relay are abstracted as vertexes in the graph. If a DUE can relay data of other DUEs, the weighted edge between the two is created. To find the edge with maximum weight, the paper uses the algorithm denoted as KM and greedy algorithm. The KM algorithm always finds the best relay (optimal selection), and the greedy algorithm reduces the complexity of the KM. The results demonstrate that the greedy algorithm performs only slightly worse in terms of data rate.

2) Enhancing capacity: With respect to reference example 7, the authors in [118] suggest to use currently idle UE (i.e., inactive UE) as a relay in order to enhance capacity. The path selection (i.e., whether to use direct path or path through relay) is based on the graph theory approach, where the throughput maximization problem is solved. The results show that the downloaded content can be increased up to approximately 26.5% compared with the conventional cellular communication.

In reference example 7, the source or destination is always the eNB. On the other hand, a relay-assisted D2D network where two DUEs can communicate via idle UE (relay) is considered in [119] (note that this corresponds to modified Scenario 2B with both DUEs in the coverage of the eNB. The DUEs are assumed to communicate either directly or via idle UE relay by means of two-hop communication in order to enhance capacity. The idle relay is used only on the assumption that the relay selection range r is lower than R, which is the maximum value of r. The optimization of the relay selection range is also addressed, since the optimal R vary for different scenario parameters.

Another feasible way to enhance capacity by exploiting D2D communication for relay purposes is introduced in [44]. The eNB allows two UEs in proximity to initiate D2D communication only on the assumption that one of the DUEs will act as a bi-directional relay for the CUE with a weak signal to the eNB (Fig. 27). Since the DUEs can have radio channels of different quality between the eNB and the CUE, the study also proposes a mechanism for relay selection. The DUE is selected to be the relay only if it can help to achieve a higher bit rate for the CUEs while guaranteeing the required bit rate for the D2D communication. To be more specific, the relay is chosen according to the CSI of the relay links as well as the D2D links. If more D2D pairs can be used to help the CUE, the selection of appropriate D2D pairs is also proposed in the study. The results show that both the CUEs and the DUEs can profit from implementation of this algorithm in terms of capacity.

A concept similar to that described in [44] is also assumed in [120]. The DUEs do not have any dedicated resources and have to cooperate with the CUE. The cooperation consists

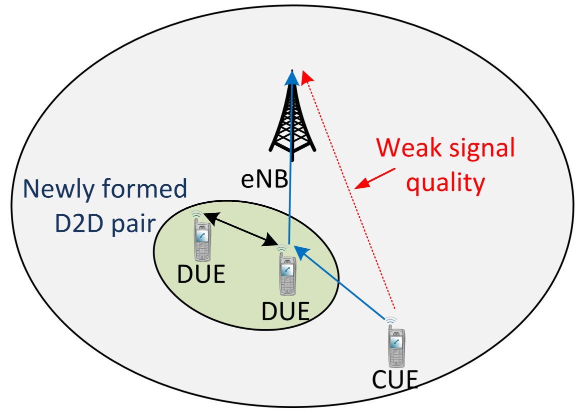

Fig. 27. The principle of bi-directional relay [44].

in the fact that the D2D pair serves as in-band relay for the CUE. Consequently, the DUEs transmit their own data while simultaneously relaying data for the CUE in the DL. To that end, the DUE transmitter exploits a superposition coding scheme as it transmits linear combinations of its information and the CUE's data. Such cooperation is allowed only if the CUE's data rate is not degraded. Although the paper presents an interesting idea, it does not explain how the CUE and DUEs would cooperate in a real system.

3) Extension of battery life: An approach extending battery life using D2D communication with relay functionality is proposed in [121]. The main objective is to help UEs with low battery level by relaying their traffic via relay. Hence, if the battery of the UE is running low, it requests selected neighbors to act as a relay. The UE could be selected either according to remaining neighbor battery status or distance. The main disadvantage of the proposed approach is that it relies on the cooperative nature of the UEs, which cannot be guaranteed in real networks. In addition, the paper completely neglects to analyze how the relay traffic would be multiplexed and scheduled with its own data.

C. Summary of approaches exploiting advanced topologies

This section shows that the usage of D2D communication does not have to be restricted to the direct communication of two users. There are several proposals that benefit from the D2D paradigm beyond conventional D2D communication between two devices in proximity. The individual papers and selected criteria are provided in Table VI and Table VII.

Of the advanced topology concepts, the cluster approach is usually adopted. In this case, several UEs can create a cluster if the system can profit from it. The one possible usage is to share specific content (e.g., music, video, etc.) among several devices without intervention of the eNB. In this case, the data sharer is automatically selected as the cluster head [45][46]. In this scenario the most tangible challenge is to guarantee that all DUEs in the same cluster are able to receive data correctly (e.g., by interference mitigation [115]).

The other attractive use of D2D communication is retransmitting broadcast data by the eNB. In the current literature, the allocation of radio resources to D2D clusters to ensure

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

TABLE VI
THE COMPARISON OF INDIVIDUAL STUDIES FOCUSED ON MULTICAST/BROADCAST CONCEPT.

Paper	Concept	Objective	D2D reuse direction	D2D scenario	CH selection	Scenario
[45]	Multicast	Addressing general challenges of multicast concept	UL	3A	Files sharer is automatically CH	N/A
[46]	Multicast	Selection between cluster and conventional mode	UL	3A	Files sharer is automatically CH	Single cell
[115]	Multicast	Interference mitigation	UL	3A	N/A	Single cell
[48]	Broadcast	Saving signaling overhead due to HARQ	UL	3B	Not specified	Single cell
[47]	Broadcast	Selecting proper no. of retransmitting UEs	UL	3B	No CH considered	Single cell
[116]	Broadcast	Radio resource and interference management	N/A	3B	Not specified	Multi cell

TABLE VII
THE COMPARISON OF INDIVIDUAL STUDIES FOCUSED ON RELAY CONCEPT.

Paper	Objective	D2D reuse direction	D2D scenario	Relay selection	Scenario
[43]	Extension of coverage	UL	2C	Weight of the edge in graph	Single cell
[119]	Enhancing capacity	UL	2B	Relay selection range (idle UE)	Single cell
[118]	Enhancing capacity	DL	2D	Throughput (idle UE)	Multi cell
[44]	Enhancing capacity	UL, DL	2D	CSI of relay links and D2D links	Single cell
[120]	Enhancing capacity	DL	2D	Data rates of the DUEs and CUEs	Single cell
[117]	Enhancing capacity, Extension of coverage	UL, DL	2C, 2D	End-to-end throughput	Multi cell
[121]	Extension of battery life	UL	2D	Remaining neighbor battery status, distance	Single cell

interference mitigation with the CUEs is proposed [116]. Moreover, two studies focus on the HARQ technique and proper retransmission of unsuccessfully delivered data to all DUEs within the same cluster [48][47]. However, these techniques are rather reactive as data not received by the DUEs are retransmitted. The more convenient approach would be to introduce techniques preventing unsuccessful delivery of data in the first place.

Moreover, D2D communication can be utilized for relay purposes without the need to install new expensive eNBs or relay stations. Consequently, the D2D could also be used for extension of coverage [43], capacity enhancement [44][119][118][120], or both [117]. The most critical issue regarding the D2D relay functionality addressed in existing literature is how to select a proper relay node. The most common goal is to maximize system capacity. However, the main problem concerning relay functionality is how to convince users to act as a relay for other users as this can drain the battery of the relaying UEs significantly and, at the same time, it can limit the capacity of relaying users. These issues are partly addressed in [44] and [120], where D2D communication is allowed only if the DUE consents to serve as a relay for a CUE with weak signal quality. Consequently, the users can benefit from D2D only if they serve as a relay. However, this condition may be double-edged. Besides the expected higher number of relays, it can also discourage users from utilizing D2D communication. In [119][118] selected relay nodes are only among the idle UEs, where the capacity is not a critical issue according to the authors. Nevertheless, even for idle UEs, battery life can be a limiting aspect. To solve this problem, the above-mentioned concepts can be done jointly [121], whereby the relay functionality is exploited to extend battery life by using less energy-consuming links.

VIII. D2D COMMUNICATION IN COEXISTENCE WITH SMALL CELLS

Introduction of small cells (SCeNBs), especially femtocells (HeNBs), into contemporary wireless networks brings new challenges if combined with D2D communication. The most prominent challenge is how to tackle interference issues in three-tier networks, where the 1st-tier encompasses the eNBs, the 2nd-tier is created by the SCeNBs, and the 3rd-tier is composed of D2D pairs as shown in (Fig. 28). In this regard, we can distinguish three interference scenarios: 1) interference between the 1st-tier and the 2nd-tier, 2) interference between the 1st-tier and the 3rd-tier, and 3) interference between the 2nd-tier and the 3rd-tier.

The first interference scenario, described in Fig. 28, is considered in [122] where the two-tier eNB/SCeNBs network profits from adoption of D2D communication. The proposal

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

combats the interference caused by the SCeNB to the Macro UE (MUE) in the vicinity through interference forwarding, as shown in (Fig. 29). The D2D functionality enables the setting up of the D2D link with known transmission power and channel measurements for interference combining. Hence, the MUE (in [122] termed MUE relay) can help the "victim" MUE to first decode the interference and then derive the desired signal. The paper does not consider a similar method for interference mitigation from the eNB to the SUEs, which could improve performance of the SUEs as well.

The first two interference scenarios as depicted in Fig. 28) are considered in [123]. The authors propose a flexible spectrum management for dense networks including both small cells and D2D communication. The eNB deployed in the area and its MUEs use licensed spectrum as primary users. The outdoor HeNBs are allowed to utilize licensed shared access, which could be static or dynamic. Finally, the indoor HeNBs and D2D pairs access the spectrum by means of secondary spectrum access with the lowest priority. The paper [123], however, does not address the problem of interference among indoor small cells and D2D communication (i.e., the last interference scenario).

The second and the third interference scenarios defined in Fig. 28 are handled in [124][125]. The authors of [124] address the interference issue via joint resource allocation and power adjustment of the DUEs. The DUEs listen to information broadcast by the eNB and femtocells (HeNBs) containing information on maximum interference tolerance (MIT). Subsequently, to ensure reliable D2D communication, the D2D pairs also determine the maximum allowable transmission power based on MIT information. Then, the D2D pairs autonomously perform radio resource management by allocation of only those RBs which guarantee reliable D2D transmission, and interference to either the eNBs or the HeNBs is avoided. The paper [125] deals with interference caused by the D2D communication to the eNB and HeNBs by means of a Stackelberg game. In the scheduling process, the UEs of the eNB and HeNBs are assumed to be the leaders, and the DUEs the followers. The leaders own radio resources and they charge fees to the DUEs. However, a disadvantage of the proposed approach is that it considers both the HeNBs and the eNB own dedicated channel (i.e., dedicate channel deployment). Nevertheless, a more realistic scenario when the eNB and the HeNBs are in co-channel deployment should be assumed.

Summary: D2D communication can also coexist with heterogeneous networks encompassing small cell deployment. A summary and comparison of individual papers dealing with this problem are provided in Table IX. The most evident challenge here is how to avoid interference in a three-tier network hierarchy. The convenient solution is to use flexible spectrum management, as suggested in [123]. Moreover, various radio resource management techniques have been proposed in the literature. Also, D2D can help to solve interference between macrocell users and small cell users by means of an interference forwarding method as proposed in [122]. Nonetheless, none of the existing papers tries to solve mutual interference among all tiers.

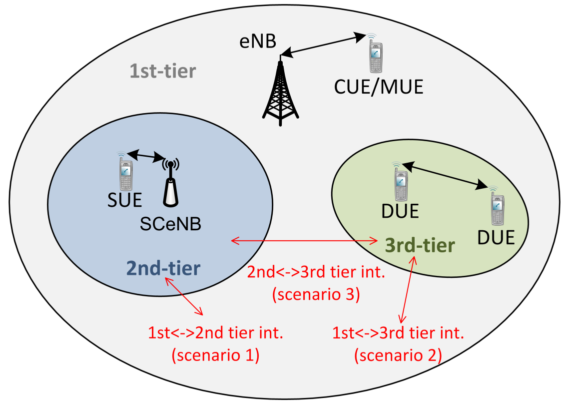

Fig. 28. The three-tier network encompassing macrocell, small cells and D2D (SUE=Small Cells UE).

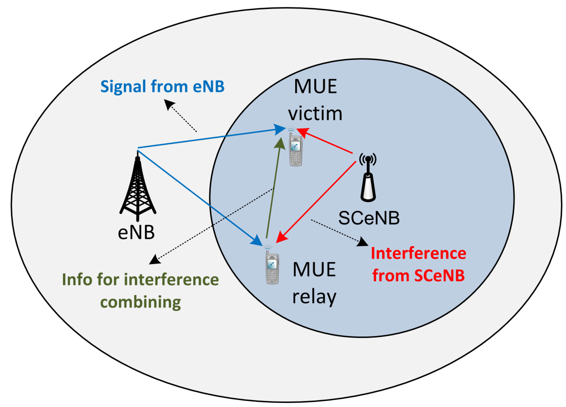

Fig. 29. The principle of interference forwarding for D2D based heterogeneous networks [122].

IX. 3GPP STANDARDIZATION ACTIVITIES ON D2D

D2D communication coexisting with cellular networks is attracting a lot of attention nowadays. A 3GPP standardization group has recently (at the end of 2011) looked at integration of D2D communication in Release 12 [126]. The justification for supporting D2D communication in emerging 3GPP standards is to cope with current socio-technological trend of supporting proximity-based services and applications (ProSe). The feasibility study for ProSe communication within 3GPP and description of use cases is introduced in [127]. This study defines the major requirements and scenarios for direct communication between the DUEs. The conclusions of this study, along with preliminary studies of requirements on evolved packet system [128] and charging and billing [129], are considered as inputs for a study on D2D architecture [22]. In 3GPP, the D2D communication is assumed to be used especially for public safety scenarios but it is not limited to them and could also be exploited in a more conventional sense for other scenarios indicated in Fig. 3.

Standardization activity is divided into two D2D main features: the D2D discovery phase and D2D communication phase. The requirements for both phases are defined in [128] and architectural aspects are studied in [22] and consequently

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

TABLE VIII
THE COMPARISON OF INDIVIDUAL STUDIES FOCUSED ON INTERFERENCE MITIGATION IF D2D COMMUNICATION IS DEPLOYED IN HETEROGENEOUS NETWORKS.

Paper	Method	Interference case	D2D reuse direction	D2D scenario	Scenario
[122]	Interference combining	eNB-HeNB	DL	1C	Single cell
[123]	Flexible spectrum management	eNB-HeNB, eNB-D2D	N/A	1C	N/A
[124]	Joint resource allocation and D2D power adjustment	eNB-D2D, HeNB-D2D	UL, DL	1C	Single cell
[125]	Joint scheduling and resource allocation (Stackelberg game)	eNB-D2D, HeNB-D2D	UL	1C	Single cell

transformed into technical specifications [23]. The purpose of the discovery phase is to identify whether two UEs with D2D functionality (termed ProSe-enabled UE) are in proximity or not. This functionality has to be permitted by the operator. The D2D communication phase enables establishment of a new communication path between two (or more) ProSe-enabled UEs and manages all D2D communication. Both discovery and communication phases are studied also from a radio perspective in [130]. This document states that usage of fully network controlled D2D (i.e., eNB schedules resources) is considered if DUEs are in coverage of the eNB; for cell edge DUEs or DUEs out of coverage, loosely controlled D2D can be adopted (see more detail on D2D control in Section III). The 3GPP also assumes usage of a new synchronization signal (D2D synchronization signal). This signal is transmitted by either the eNB if available for DUEs or by the DUE itself if no D2D synchronization signal from the eNB is available.

Regarding the management of D2D communication, 3GPP initiated work on the definition of objects and parameters for provisioning and authorization of ProSe to the UEs [131]. In parallel, work on the definition of new interfaces and design of protocols carried over those interfaces is in the preliminary stage, as indicated in [25], [24] and [132].

Aspects related to security of D2D communication are addressed by 3GPP in technical reports [133]. The documents contain analysis and possible solutions addressing all potential security risks, including direct communication between DUEs, as well as relaying functionality. Nevertheless, such documents are considered only for evaluation of possible solutions and provide no final definition of security solutions to be adopted by 3GPP. Security remains an open challenge for future research, not only within the frame of 3GPP but for general research on D2D underlying cellular networks.

An overview of standardization activities in 3GPP focusing specifically on ProSe enabled communication is also described in [134]. The study also shows initial evaluation results in terms of achieved CINR and throughput gains for better illustration of the importance of D2D communication in LTE-A.

X. CHALLENGES AND FUTURE RESEARCH DIRECTIONS

D2D communication has attracted attention thanks to its potential to increase the spectral and energy efficiency of the network. However, to maximize the gains of D2D communication, there are many open challenges dealing with mode selection, radio resource management, energy efficiency, advanced topology concepts, and scenarios assuming the coexistence of D2D with small cells. Furthermore, future research challenges concerning mobility management and security mechanisms need to be thoroughly addressed as well. All these challenges are outlined in the following subsections.

A. Mode selection

The potential future research direction regarding mode selection is the dynamicity of selection between the individual allocation modes. With respect to the literature, switching between individual allocation modes depending on the current state of the network should be considered since the wireless environment may be changing often. As already described in Section IV, a semi-static and a dynamic mode selection is partly addressed in [65] and [66], respectively. However, these papers do not consider the mobility of users, which will have a large impact on mode selection. Initial study addressing dynamic mode selection, considering mobility of indoor users attached to femtocells, is addressed in [135]. In addition, the dynamic selection performed on a slot-by-slot basis as proposed in [66] will result in an excessive amount of signaling overhead. Hence, it is necessary to further investigate at what timescale the decision on appropriate mode (initiated either by the network or the DUEs) should be taken. The shorter the interval selected, the more optimal the current allocation mode. At the same time, this also results in a heavier burden on the network in terms of higher generated overhead. Consequently, some trade-off between allocation mode optimality and signaling overhead needs to be investigated and analyzed. Moreover, the reporting mechanism may be periodic, event-driven, or hybrid, as indicated in [30]. In the case of periodic reporting, high overhead could be generated and the network could be overloaded by signaling activity. On the other hand, the reactive approach can lead to a higher delay in selection of the allocation mode and thus be unsuitable for cases when the situation changes frequently.

The other factor only partly tackled in [66] is the activity/inactivity of the CUEs and how this influences mode selection strategy. The most contemporary studies consider a full buffer model where all UEs in the system are still active. Nevertheless, this assumption does not reflect the behavior of users in real life that follows a different pattern. In this regard, interference and load of the network change rapidly and the impact on D2D allocation mode selection has to be evaluated.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

B. Interference management

So far, the research addressing interference in Section V has focused on the case when only the CUEs and the DUEs share the same resources (in Table II termed SM (DUEs use dedicated resources)) or when several D2D pairs use overlapping resources but have dedicated bandwidth with respect to the cellular communication (in Table II termed DM (DUEs use shared resources)). From the capacity perspective it would be beneficial if the DUEs could exploit all CUEs resources and, at the same time, if several D2D pairs could reuse the same resources (in Table II this case is labeled as "SM (DUEs use shared resources)"). In such scenarios, interference management could become complex in terms of computational and communication load. However, it is worth investigating this possibility in more detail as the reward in terms of high spectral efficiency is an important part of satisfying the ever-increasing demands of users.

An attractive way to achieve more efficient utilization of radio resources by D2D communication and to mitigate interference is to use adaptive radio resource allocation methods, selecting the UL and the DL direction for the D2D communication in a dynamic way. The current research considers only the option when the DL, the UL or both are used in a static manner. Dynamic selection means the situation when the DUEs use the DL and the UL in relation to the current load in both directions (i.e., activity of the CUEs and the DUEs) and in dependence on actual interference patterns. The selection itself should be done jointly with an appropriate dynamic mode selection algorithm and interference management technique. Another alternative is to use both the UL and the DL not currently utilized by the cellular network in a similar fashion, as a cognitive radio would do. This solution should result in even higher spectral efficiency and effective utilization of radio resources. Since this option brings additional challenges (e.g., how effectively to determine the activity of primary users in order not to disturb them and the increased complexity of the terminals that have to possess both D2D and cognitive capabilities), new, sophisticated and at the same time simple and low-cost algorithms/schemes need to be developed.

Similarly to mode selection, the existing studies on interference mostly fail to consider the impact of users' mobility on interference management, which is inherent in mobile networks. Nevertheless, this aspect strongly affects the level of interference among individual network entities and should not be disregarded in future research work. Consequently, new advanced algorithms and highly sophisticated techniques have to be developed to cope with dynamic mobile networks adopting D2D communication paradigms.

C. Energy efficiency

The D2D communication advantage rests in its potential to minimize power consumption thanks to the proximity of communication devices (see Section VI). Recent studies mostly employ various radio resource management techniques (e.g., [6][105][106]) or cooperative cluster formation [109][110][111] to accomplish this goal. Another option, not considered in existing research works, is to implement existing power saving mechanisms exploited in the conventional mobile wireless networks and investigate their effect on D2D communication. Hence, power-saving algorithms should be developed for D2D communications.

Furthermore, all recent papers address the problem of energy consumption by considering only radio communication. However, radio communication in today's UEs (smartphones, laptops, etc.) is only part of the overall energy consumption of the whole device. Moreover, the users are more interested in the battery life of their devices than in energy consumption. Therefore, investigation of the overall impact of the D2D on the battery life is necessary in order to show whether a notable gain can be achieved from the user's perspective.

D. Advanced topology concepts exploiting D2D

As described in Section VII, D2D communication can be used for multicast/broadcast purposes using cluster topology. In the case of multicast transmission (i.e., sharing of some data among users in proximity), the problem is that the transmission has to be adapted according to the member experiencing the worst channel quality [45][46]. This approach may, however, significantly decrease spectral efficiency. In this case, the use of multi-hop communication within the cluster can help to increase transmission efficiency, as in [136]. In addition, the papers dealing with multicast/broadcast concept focus primarily on HARQ techniques and on retransmission of data to UEs which do not receive data correctly. The future research in this area should be rather aimed at strategies preventing unsuccessful reception in the first place. This could be done by deploying cooperative strategies among DUEs in the same cluster to cope with interference.

Regarding the relay concept, when the D2D users serve as the relays, a key challenge is how to convince the users to serve as relay nodes for other users, because of battery limitations. One option proposed in the literature is to "buy" and "sell" DL relay services [137]. Another possible way to solve this problem is to use only those devices that are currently plugged in to power outlets. However, battery restriction can significantly reduce the potential of this approach. On the other hand, users utilizing relaying functionality may be worried about security threats as private data will be handled by other users. Consequently, privacy and security have to be assured by future research in this area. The other attractive option regarding the relay concept is to extend two-hop relay communication to multi-hop relay communication. In this case, interaction among cellular transmission, conventional direct (one-hop) D2D communication, two-hop relay, and multi-hop relay communication should be investigated and analyzed.

E. D2D communication in coexistence with small cells

The other attractive option regarding D2D communication is to investigate the scenario when D2D is used in heterogeneous networks, when small cells (pico/micro/femto) are introduced into the system (see Section VIII). Small cells are expected to be heavily deployed in the near future (especially femtocells)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

[138] and one of the challenges here is the interference between two-tier network when one tier is composed of the eNBs and the second tier is represented by small cells. If D2D communication is introduced, the interference management is even more challenging, since DUEs can interfere both with macro users and with users of small cells. So far, several studies focus on interference issues if both small cells and DUEs are deployed in the network under the same eNB [123][124][125][122]. Nevertheless, no paper examines the worst and, at the same time, the most challenging interference case, where all three-tier components of the network (i.e., the eNBs, the HeNBs, and D2D communication) are mutually interfered. Nevertheless, to fully exploit both small cells and D2D communication, these should not be handled separately.

Besides interference management, new emerging scenarios need to be assumed when one DUE composing a D2D pair can be connected to the eNB and the second DUE is located inside the small cell or when both are connected to the same or different small cells. These will have an impact on mode selection strategies and radio resource management in general. On the other hand, the foreseen advantage is that D2D communication can alleviate the backhaul of small cells, which is very welcome, especially in the case of femtocells. Besides, this scenario can also simplify D2D discovery in some cases as users served by a small cell can be assumed to be in proximity because of the small radius of those cells.

F. New potential D2D scenarios

So far, most of the studies dealing with proper mode selection, interference issues or energy efficiency of the system assume that both communicating DUEs are located within the same cell (single cell scenario 1C). Nevertheless, this assumption limits the advantages introduced by D2D. In particular, the scenario whereby two UEs are attached to different cells is worth investigating in more detail (multi-cell scenario 1D). The advantages of this scenario are as follows. First, since the UEs are at the cell boundaries, the battery power of devices is drained significantly because of the high transmission power in the UL. Hence this solution seems to be attractive in terms of energy efficiency. Second, the advantage of D2D results in offloading backhaul of cellular networks as data do not need to be transmitted between serving eNBs of both DUEs [139][140]. On the other hand, this scenario introduces several challenges that need to be addressed. The most significant challenge is how to handle interference management. One obvious solution is to use the DM. However, the selection of appropriate dedicated resources has to be negotiated between both involved cells, which adds to the complexity of the solution compared with Scenario 1C. Moreover, the use of the SM could result in very complex interference mitigation techniques that could be too hard to implement in real networks. Furthermore, the limited quality of backhaul for exchange of information between both cells is a factor which needs to be taken into account in future proposed solutions.

G. New bands for communication

Currently, 3GPP is investigating a possibility to exploit an opportunistic use of unlicensed bands for cellular networks in order to enhance the bandwidths available for specific services, which does not require high QoS. This concept is known as LTE-U [141]. Consideration of D2D as an approach exploiting unlicensed bands could help to relieve the negative impact of problems related to interference among DUEs and CUEs. Analogously, potential exploitation of millimeter waves seems to be a very promising solution [142]. Millimeter waves are intended for short-range communication. In [143], the authors demonstrate, by means of field trials, a way to communicate by using millimeter waves in a range of a few hundred meters. Obviously, DUEs will not be able to transmit with the same power as common base stations because of battery limitations. Nevertheless, even a communication range of tens of meters is a promising way to reduce interference among DUEs and CUEs by smart offloading of part of the communication from common bands to bands of millimeter waves. For both LTE-U and millimeter waves, it is necessary to design algorithms to select the most suitable band to be used for communication with respect to the users' location and demands for services consumed by the users.

H. Mobility management

The mobility of D2D users has an impact on mode selection and/or interference management as explained before, but mobility management also has to be enhanced accordingly. This issue is thus far completely neglected in the current literature. The problem of DUE mobility consists in the fact that if a pair of users exploiting D2D communication in a single-cell scenario (i.e., both users are under coverage of the same cell) crosses the cell edge, handover of the user to a new cell must be performed. The problem is even more complex if small cells are deployed together with D2D communication. This leads to switching of D2D from a single-cell to a multi-cell scenario. In other words, the pair of users under coverage and management of just one cell will be split and each is managed by a different cell. As a result, new algorithms for decisions on the handover of D2D users considering the change of single/multi-cell scenario need to be proposed. In addition, algorithms to cope with the change from single-cell to multi-cell scenario, if change cannot be avoided, need to be proposed.

I. Security mechanism

One of the most important issues regarding D2D communication, in general, is to ensure its security so it can be accepted by the masses. Yet only marginal effort has been devoted to this issue. The only paper addressing the security problems of D2D in LTE-A environments is [26]. Nevertheless, the paper just provides a basic analysis of security threats in three simple topology scenarios and proposes authentication and key management solutions based on a shared key. There are, however, many other types of attack, such as man-in-the-middle attack, denial of service, distributed denial of service and replay attacks, that have not been analyzed at all. For real-life D2D communication, it is necessary to make a deep and complex security analysis not only of the simple scenarios but also of the complex ones. In these complex and more realistic scenarios not all devices belong to the same cell or to the same mobile operator and there are various combinations of devices, topologies, and protocols that should work in unison. Conventional symmetric cryptosystems are not able to meet all the requirements, and it is a chance to use modern and highly efficient cryptographic techniques, which should ensure not only classical confidentiality-integrity-availability properties, but also advanced security properties like anonymity, pseudonymity, secure online reputation, non-repudiation, identity-based encryption and attribute-based encryption.

XI. Conclusion

The D2D communication underlying cellular mobile networks offers several advantages such as offloading of an overloaded mobile cellular network, hop gain or energy efficiency. At the same time, the D2D paradigm introduces several critical challenges that must be addressed in order to profit from the direct communication between mobile devices.

Albeit D2D communication underlying cellular networks is a heavily investigated area, research carried out so far is still in the preliminary stage of studying the performance of D2D in simplified scenarios or under limited conditions. These studies show the potential of D2D in terms of high performance gain in cellular networks and becoming an integral part of future mobile networks. Nevertheless, recent research also presents many new challenges and issues that must be addressed in order to overcome the expected difficulties and obstacles in the management of D2D communication from a technical perspective. The major weakness of recent research is a lack of D2D mobility and more realistic scenarios for future mobile networks such as heterogeneous networks with densely deployed small cells. In addition, more attention should be paid to security and privacy issues as these aspects are becoming a priority and are a potential barrier to the success of many new approaches.

Acknowledgment

This work has been performed in the framework of the FP7 project TROPIC IST-318784 STP, which is funded by the European Community. The Authors would like to acknowledge the contributions of their colleagues from TROPIC Consortium (http://www.ict-tropic.eu).

TABLE IX
Table summarizing the acronym used in the paper.

Acronym	Explanation
CM	Cellular Mode
CSI	Channel State Information
CQI	Channel Quality Indicator
CUE	Cellular User
D2D	Device-to-Device
DL	Downlink
DM	Dedicated Mode
DUE	D2D User
EMBMS	Evolved Multicast and Broadcast Multimedia Services
eNB	Evolved Node B
EPC	Evolved Packet Core
FFR	Fractional Frequency Reuse
FSA	Full Set Allocation
HARQ	Hybrid Automatic Repeat Request
HeNB	Home eNB (Femtocell)
HSS	Home Subscriber Server
IA	Interference Alignment
IC	Interference Cancellation
ICA	Iterative Combinatorial Auction
ILA	Interference Limited Area
ISR	Interference to Signal Ratio
LC	Load Control
MBMS	Multicast and Broadcast Multimedia Service
MIMO	Multiple Input Multiple Output
MIT	Maximum Interference Tolerance
MME	Mobility Management Entity
MRT	Maximum Ratio Transmission
MS	Mode Selection
MUE	Macrocell UE
NACK	Negative Acknowledgment
PC	Power Control
PCRF	Policy and Charging Rules Function
P-GW	Packet data network Gateway
PLMN	Public Land Mobile Network
ProSe	Proximity Services
QoS	Quality of Service
QP	Quiet Period
RB	Resource Block
RLC	Radio Link Control
RRA	Radio Resource Allocation
RRM	Radio Resource Management
SA	Subset Allocation
SAE	System Architecture Evolution
SCeNB	Small Cell
SFN	Single Frequency Network
SIC	Successive Interference Cancellation
SINR	Signal to Interference plus Noise Ratio
SISO	Single Input Single Output
SM	Shared Mode
SUE	Small cell UE
UE	User Equipment
UL	Uplink

References

[1] Small Cell Forum. (2012). Small Cells whats the big idea? Femtocells are expanding beyond the home. Feb 2012.
[2] http://www.smallcellforum.org/.
[3] J. Marinho and E. Monteiro, "Cognitive Radio: Survey on Communications Protocols, Spectrum Issues, and Future Research Directions," *Wireless Networks*, vol. 18, no. 2 , 2012, 147-164.
[4] P. Cheng, L. Deng, H. Yu, Y. Xu, and H. Wang, "Resource allocation for cognitive networks with D2D communication: An evolutionary approach," *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2012, 2671-76.
[5] Qualcomm, "LTE Direct Overview", available online at "http://www.qualcomm.com/media/documents/lte-direct-whitepaper", Aug. 2013.
[6] M. Belleschi, G. Fodor, and A. Abrardo, "Performance analysis of a distributed resource allocation scheme for D2D communications," *2011 IEEE GLOBECOM Workshops (GC Wkshps)*, Dec. 2011, 358-362.
[7] S. Hakola, C. Tao, J. Lehtomak, and T. Koskela, "Device-To-Device (D2D) Communication in Cellular Network - Performance Analysis of Optimum and Practical Communication Mode Selection," *2010 IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2010, 1-6.
[8] Z. Liu, T. Peng, Q. Lu, and W. Wang, "Transmission capacity of d2d communication under heterogeneous networks with dual bands," *7th International ICST Conference on Cognitive Radio Oriented Wireless*

Networks and Communications (CROWNCOM), June 2012, 169174.

[9] Z.-S. Syu and C.-H. Lee, "Spatial constraints of device-to-device communications," *First International Black Sea Conference on Communications and Networking (BlackSeaCom)*, July 2013, 9498.

[10] [L. Lei, Z.i Zhong, Ch. Lin, and Xuemin Shen, "Operator controlled device-to-device communications in LTE-advanced networks," *IEEE Wireless Communications* , vol. 19, no. 3, June 2012, 96-104.

[11] A. Asadi, A. Q. Wang, and V. Mancuso, "A Survey on Device-to-Device Communication in Cellular Networks", *IEEE Communications Surveys and Tutorials*, vol. 16, no. 4, 2014, 1801-1819.

[12] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-Device Communication in LTE-Advanced Networks: A Survey" accepted in *IEEE Communications Surveys and Tutorials*, 2014.

[13] Y. Zhang, L. Du, D. Shang, and Y. Yang, "Performance Analysis of Intra-Cell Spatial Reuse in a P2P-Enabled TDD CDMA Network," *2006 IEEE 64th Vehicular Technology Conference*, 2006, 1-5.

[14] Y. Zhang, L. Du, D. Shang, and Y. Yang, "Performance analysis of a P2Penabled TDDCDMA network with intracell spatial resource reuse," *Wireless Communications and Mobile Computing*, vol. 9, no. 8, 2009,1059-1069.

[15] L. Popova, T. Herpel, and W. Koch, "Enhanced downlink capacity in UMTS supported by direct mobile-to-mobile data transfer," *2007 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet NETWORKING*, 2007. 522-534.

[16] L. Popova, T. Herpel, and W. Koch, "Improving downlink UMTS capacity by exploiting direct mobile-to-mobile data transfer," *Ad Hoc and Wireless Networks and Workshops, Modeling and Optimization in Mobile (WiOpt)*, 2007, 1-8.

[17] L. Popova, T. Herpel, W. Gerstacker, and W. Koch, "Cooperative mobile-to-mobile file dissemination in cellular networks within a unified radio interface," *Computer Networks*, vol. 52, no. 6, 2008, 1153-1165.

[18] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks," *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2010, 514-521.

[19] H. W. Yoon, J. S. Kim, S. J., Bae, B. G. Choi, and M. Y. Chung, "Performance analysis of FlashLinQ with various yielding threshold values," *2012 International Conference on ICT Convergence (ICTC)*, Oct. 2012, 477-478.

[20] S. Mumtaz and J. Rodriguez (eds.), "Smart Device to Smart Device Communication", Publisher: Springer, ISBN 978-3-319-04963-2, Dec. 2014.

[21] 3GPP TS 23.002: Technical Specification Services and System Aspects;Network Architecture; (Release 12), V12.6.0, Dec. 2014.

[22] 3GPP TR 23.703: Technical Specification Group Services and System Aspects; Study on architecture enhancements to support Proximity-based Services (Release 12), V12.0.0, Feb. 2014.

[23] 3GPP TS 23.303: Technical Specification Group Services and System Aspects; Proximity-based services (ProSe); Stage 2 (Release 12), V12.1.0, June 2014.

[24] 3GPP TS 29.343: Technical Specification Group Core Network and Terminals; Proximity-services (ProSe) Function to Proximity-services (ProSe) Application Server aspects (PC2); Stage 3 (Release 12), V0.1.0, Apr. 2014.

[25] 3GPP TS 24.334: Technical Specification Group Core Network and Terminals; Proximity-services (ProSe) User Equipment (UE) to Proximity-services (ProSe) Function Protocol aspects; Stage 3 (Release 12), V1.1.0, July 2014.

[26] M. Alam, D. Yang, J. Rodriguez, and R. Abd-Alhameed, "Secure device-to-device communication in LTE-A," *IEEE Communications Magazine*, vol. 52, no. 4, Apr. 2014, 66-73.

[27] 3GPP TS 29.334: Technical Specification Group Core Network and Terminals; IMS Application Level Gateway (IMS-ALG) IMS Access Gateway (IMS-AGW); Iq Interface; Stage 3 (Release 12), V12.3.0, June 2014.

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

[29] K. Doppler, M. P. Rinne, P. Janis, C. Ribeiro, and K. Hugl, "Device-to-Device Communications; Functional Prospects for LTE-Advanced Networks," *IEEE International Conference on Communications (ICCC)*, June 2009, 1-6.

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

[31] B. Raghothaman, E. Deng, R. Pragada, G. Sternberg, T. Deng, and K. Vanganuru, "Architecture and protocols for LTE-based device to device communication," *2013 International Conference on Computing, Networking and Communications (ICNC)*, Jan. 2013, 895-899.

[32] X. Xiao, X. Tao, and J. Lu, "A QoS-Aware Power Optimization Scheme in OFDMA Systems with Integrated Device-to-Device (D2D) Communications," *2011 IEEE Vehicular Technology Conference (VTC Fall)*, Sept. 2011, 1-5.

[33] T. Chen, G. Charbit, and S. Hakola, "Time Hopping for Device-To-Device Communication in LTE Cellular System," *2010 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2010, 1-6.

[34] S. Mumtaz, H. Lundqvist, K. M. S. Huq, J. Rodriguez, and A. Radwan, "Smart Direct-LTE communication: An energy saving perspective," *Ad Hoc Networks*, vol. 13, 2014, 296-311.

[35] M. J. Yang, S. Y. Lim, H. J. Park, and N. H. Park, "Solving the Data Overload: Device-to-Device Bearer Control Architecture for Cellular Data Offloading", *IEEE Vehicular Technology Magazine*, vol. 8, no. 1, March 2013, 31-39.

[36] P. Nguyen, P. Wijesinghe, R. Palipana, K. Lin, and D. Vasic, "Network-Assisted Device Discovery for LTE-based D2D Communication Systems," *International Conference on Communications (ICC)*, June 2014, 3160-3165.

[37] K. J. Zou, M. Wang, K. W. Yang, J. Zhang, W. Sheng, Q. Chen, and X. You, "Proximity discovery for device-to-device communications over a cellular network," *IEEE Communications Magazine*, vol. 52, no. 6, June 2014, 98-107.

[38] X. Wu, et al., "FlashLinQ: A Synchronous Distributed Scheduler for Peer-to-Peer Ad Hoc Networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 4, Aug. 2013, 1215-1228.

[39] F. Baccelli, et al., "On the design of device-to-device autonomous discovery," *2012 Fourth International Conference on Communication Systems and Networks (COMSNETS)*, 2012, 1-9.

[40] S. Jung and S. Chang, "A discovery scheme for device-to-device communications in synchronous distributed networks," *16th International Conference on Advanced Communication Technology (ICACT)*, Feb. 2014, 815-819.

[41] B. Kaufman, B. Aazhang, and J. Lilleberg, "Interference aware link discovery for device to device communication," *Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers (ASCCS)*, Nov 2009, 297301.

[42] L. Qianxi, M. Qingyu, G. Fodor, and N. Brahmi, "Clustering Schemes for D2D Communications Under Partial/No Network Coverage," *IEEE Vehicular Technology Conference-Spring (VTC-spring)*, May 2014.

[43] L. Wang, T. Peng, Y. Yang, and W. Wang, "Interference Constrained Relay Selection of D2D Communication for Relay Purpose Underlaying Cellular Networks," *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, Sept. 2012, 1-5.

[44] Y. Pei and L. Ying-Chang, "Resource Allocation for Device-to-Device Communications Overlaying Two-Way Cellular Network," *IEEE Transactions on Wireless Communications*, vol. 12, no. 7, 2013, 3611-21.

[45] 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." *2011 IEEE Wireless Communications and Networking Conference (WCNC)*, 2011, 986-991.

[46] T. Koskela, S. Hakola, T. Chen, and J. Lehtomaki, "Clustering concept using device-to-device communication in cellular system," *2010 IEEE Wireless Communications and Networking Conference (WCNC)*, 2010, 1-6.

[47] B. Zhou, H. Hu, S. Huang, and H. Chen, "Intra-Cluster Device-to-Device Relay Algorithm with Optimal Resource Utilization," *IEEE Transaction on Vehicular Technology*, 2013, 2315-26.

[48] J. Du, W. Zhu, J. Xu, Z. Li, and H. Wang, "A Compressed HARQ Feedback for Device-to-Device Multicast Communications," *2012 IEEE Vehicular Technology Conference (VTC Fall)*, 2012, 1-5.

[49] S. Xu, H. Wang, T. Chen, Q. Huang and T. Peng, "Effective Interference Cancellation Scheme for Device-to-Device Communication Underlaying Cellular Networks," *2010 IEEE 72nd Vehicular Technology Conference Fall (VTC 2010-Fall)*, Sept. 2010, 1-5.

[50] S. Xu, H. Wang and T. Chen, "Effective Interference Cancellation Mechanisms for D2D Communication in Multi-Cell Cellular Networks," *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, May 2012, 1-5.

[51] H. Min, J. Lee, S. Park, and D. Hong, "Capacity Enhancement Using an Interference Limited Area for Device-to-Device Uplink Underlaying Cellular Networks," *IEEE Transactions on Wireless Communications*, Dec. 2011, 3995-4000.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

[52] 3GPP TS 36.300: Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 11), V11.7.0, Sep. 2013.

[53] H. Xing and S. Hakola, "The investigation of power control schemes for a device-to-device communication integrated into OFDMA cellular system," *2010 IEEE 21st International Symposium onPersonal Indoor and Mobile Radio Communications (PIMRC)*, 2010, 1775-1780.

[54] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," submitted to *IEEE Transactions on Wireless Communications*, Mar. 2014.

[55] H. ElSawy and E. Hossain, "Analytical Modeling of Mode Selection and Power Control for Underlay D2D Communication in Cellular Networks," submitted to *IEEE Transactions on Communications*, May 2014.

[56] P. Janis, C.-H. Yu, K. Doppler, C. Ribeiro, C. Wijting, K. Hugl, O. Tirkkonen, and V. Koivunen, "Device-to-Device communication underlaying cellular communication systems," *International Journal on Communications, Networking and System Science*, vol. 2, no. 3, June 2009, 169178.

[57] Y. Chia-Hao, O. Tirkkonen, K. Doppler and C. Ribeiro, "Power Optimization of Device-to-Device Communication Underlaying Cellular Communication," *IEEE International Conference on Communications (ICC)*, June 2009, 1-5.

[58] Y. Chia-Hao, K. Doppler, C. Ribeiro, and O. Tirkkonen, "Resource Sharing Optimization for Device-to-Device Communication Underlaying Cellular Networks," *IEEE Transactions on Wireless Communications*, vol.10, no.8, Aug. 2011, 2752-63.

[59] B. Wang, L. Chen, X. Chen, X. Zhang, and D. Yang, "Resource Allocation Optimization for Device-to-Device Communication Underlaying Cellular Networks," *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, May 2011, 1-6.

[60] B. Cho, K. Koufos, and R. Jantti, "Spectrum allocation and mode selection for overlay D2D using carrier sensing threshold" *International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, June 2014, 26-31.

[61] K. Doppler, Y. Chia-Hao Yu, C.B . Ribeiro, and P. Janis, "Mode Selection for Device-To-Device Communication Underlaying an LTE-Advanced Network," *2010 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2010, 1-6.

[62] J. Gu, S. J. Bae, B. G. Choi, and M. Y. Chung, "Mode selection scheme considering transmission power for improving performance of device-to-device communication in cellular networks," *6th International Conference on Ubiquitous Information Management and Communication (ICUIMC)*, Feb. 2012, 30-36.

[63] K. Akkarajitsakul, P. Phunchongharn, E. Hossain, and V. K. Bhargava, "Mode selection for energy-efficient D2D communications in LTE-advanced networks: A coalitional game approach," *2012 IEEE International Conference onCommunication Systems (ICCS)*, Nov. 2012, 488-492.

[64] Z. Liu, T. Peng, S. Xiang, and W. Wang, "Mode selection for Device-to-Device (D2D) communication under LTE-Advanced networks," *2012 IEEE International Conference onCommunications (ICC)*, June 2012, 5563-5567.

[65] M. H. Han, B. G. Kim, and J. W. Lee, "Subchannel and Transmission Mode Scheduling for D2D Communication in OFDMA Networks," *2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sept. 2012, 1-5.

[66] L. Lei, X. Shen, M. Dohler, Ch. Lin, and Z. Zhong, "Queuing Models with Applications to Mode Selection in Device-to-Device Communications Underlaying Cellular Networks," accepted in *IEEE Transactions on Wireless Communication*, 2014.

[67] C. H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "On the Performance of Device-to-Device Underlay Communication with Simple Power Control," *69th IEEE Vehicular Technology Conference (VTC-Spring)*, April 2009, 1-5.

[68] P. Janis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-Aware Resource Allocation for Device-to-Device Radio Underlaying Cellular Networks," *IEEE 69th Vehicular Technology Conference (VTC Spring)*, April 2009, 1-5.

[69] T. Peng, Q. Lu, H. Wang, S.i Xu, and W. Wang, "Interference avoidance mechanisms in the hybrid cellular and device-to-device systems," *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept. 2009, 617-621.

[70] S. Xu, H. Wang, T. Peng, and Q. Huang, "Effective Labeled Time Slots Based D2D Transmission in Cellular Downlink Spectrums," *71st IEEE Vehicular Technology Conference (VTC-Spring)*, May 2010, 1-5.

[71] M. Zulhasnine, H. Changcheng, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying LTE network," *2010 IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2010, 368-375.

[72] P. Tapia, J. Liu, Y. Karimli, and M. J. Feuerstein, "HSPA Performance and Evolution: A Practical Perspective", ISBN 978-0-470-69942-3, 2009.

[73] J. Gu, S. J. Bae, B. G. Choi, and M. Y. Chung, "Dynamic power control mechanism for interference coordination of device-to-device communication in cellular networks," *2011 Third International Conference on Ubiquitous and Future Networks (ICUFN)*, June 2011, 71-75.

[74] T. Han, R. Yin, Y. Xu, and G. Yu, "Uplink channel reusing selection optimization for Device-to-Device communication underlaying cellular networks," *2012 IEEE 23rd International Symposium onPersonal Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2012, 559-564.

[75] Y. Xu, R. Yin, T. Han, and G. Yu, "Dynamic resource allocation for Device to Device communication underlaying cellular networks", *International Journal of Communication Systems*, 2012.

[76] B. Peng, Ch. Hu, T. Peng, and W. Wang, "Optimal Resource Allocation for Multi-D2D Links Underlying OFDMA-Based Communications," *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, Sept. 2012, 1-4.

[77] X. Zhu, S. Wen, G. Cao, X. Zhang, and D. Yang, "QoS-based resource allocation scheme for Device-to-Device (D2D) radio underlaying cellular networks," *2012 19th International Conference on Telecommunications (ICT)*, April 2012, 1-6.

[78] W. Yu, G. Ginis, and J. M. Cioffi, "Distributed Multiuser Power Control for Digital Subscriber Lines," *IEEE Journal on Selected Areas in Communiocation*, vol. 20, no. 5, june 2002, 1105-1115.

[79] F. Wang, L. Song, Z. Han, Q. Zhao, and X. Wang, "Joint scheduling and resource allocation for device-to-device underlay communication," *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, 134-139.

[80] Q. Duong, Y. Shin, and O. S. Shin, "Resource allocation scheme for device-to-device communications underlaying cellular networks," *2013 International Conference on Computing, Management and Telecommunications (ComManTel)*, Jan. 2013, 66-69.

[81] 3GPP TS 36.211: Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 11), V11.4.0, Sept. 2013.

[82] D. Tsolkas, E. Liotou, N. Passas, and L. Merakos, "A graph-coloring secondary resource allocation for D2D communications in LTE networks," *2012 IEEE 17th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Sept. 2012, 56-60.

[83] H. Min, W. Seo, J. Lee, S. Park, and D. Hong, "Reliability Improvement Using Receive Mode Selection in the Device-to-Device Uplink Period Underlaying Cellular Networks," *IEEE Transactions on Wireless Communications*, vol.10, no.2, Feb. 2011, 413-418.

[84] T. K. Lo, "Maximum ratio transmission," *IEEE Transactions on Communications*", vol. 47, Oct. 1999, 1458-61.

[85] P. Janis, V. Koivunen, C. B. Ribeiro, K. Doppler, and K. Hugl, "Interference-avoiding MIMO schemes for device-to-device radio underlaying cellular networks," *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept. 2009, 2385-89.

[86] H. S. Chae, J. Gu, B. G. Choi, and M. Chung, "Radio resource allocation scheme for device-to-device communication in cellular networks using fractional frequency reuse," *2011 17th Asia-Pacific Conference on Communications (APCC)*, Oct. 2011, 58-62.

[87] P. Bao, G. Yu and R. Yin, "Novel frequency reusing scheme for interference mitigation in D2D uplink underlaying networks," *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, July 2013, 491-496.

[88] X. Chen, L. Chen, M. Zeng, X. Zhang, and D. Yang, "Downlink resource allocation for Device-to-Device communication underlaying cellular networks," *2012 IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sept. 2012, 232-237.

[89] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, and B. Jiao, "Interference-aware resource allocation for device-to-device communications as an underlay using sequential second price auction," *2012 IEEE International Conference on Communications (ICC)*, June 2012, 445-449.

[90] C. Xu, L. Song, Z. Han, D. Li, and B. Jiao, "Resource allocation using a reverse iterative combinatorial auction for device-to-device underlay cellular networks," *2012 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2012. 4542-47.

[91] R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference-aware graph based resource sharing for device-to-device communications underlaying

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

cellular networks," *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, 140-145.

[92] Ch. H. Yu and O. Tirkkonen, "Device-to-Device underlay cellular network based on rate splitting," *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, 262-266.

[93] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE T. Inform. Theory*, vol. 27, no. 1, Jan. 1981, 49 60.

[94] W. Xu, L. Liang, H. Zhang, S. Jin, J. C. F. Li, and M. Lei, "Performance enhanced transmission in device-to-device communications: Beamforming or interference cancellation?," *2012 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2012, 4296-4301.

[95] W. Fu, R. Yao, F. Gao, J. C. F. Li, and M. Lei, "Robust null-space based interference avoiding scheme for D2D communication underlaying cellular networks," *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, 4158-4162.

[96] D. H. Lee, K. W. Choi, W. S. Jeon, and D. G. Jeong, "Resource allocation scheme for device-to-device communication for maximizing spatial reuse," *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2013, 112-117.

[97] M. Hajiaghayi, C. Wijting, C. Ribeiro, and M. T. Hajiaghay, "Efficient and practical resource block allocation for LTE-based D2D network via graph coloring," *Wireless Networks*, Aug. 2013, 1-14.

[98] V. R. Cadambe and S. A. Jafar, "Interference Alignment and Spatial Degrees of Freedom for the K User Interference Channel," *2008 IEEE International Conference on Communications (ICC)*, vol. 54, 2008, 971-975.

[99] H. E. Elkotby, K. M. F. Elsayed, and M.H. Ismail, "Exploiting interference alignment for sum rate enhancement in D2D-enabled cellular networks," *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2012, 1624,1629.

[100] H. E. Elkotby, K. M. F. Elsayed, and M.H. Ismail, "Shrinking the reuse distance: Spectrally-efficient radio resource management in D2D-enabled cellular networks with Interference Alignment," *2012 IFIP Wireless Days (WD)*, Nov. 2012, 1-6.

[101] G. Fung, "A Comprehensive Overview of Basic Clustering Algorithms," *IEEE June*, 2001, 1-37.

[102] L. Lei, J. Zhao, Z. Zhong, and K. Zheng, "Flow-Level Analysis of Energy Efficiency Performance for Device-to-Device Communications in OFDM Cellular Networks," *The Computer Journal*, 2013.

[103] X. Qiu, L. Xuewen, K. Dong, and S. Zhu, "Energy efficiency analysis in device-to-device communication underlaying cellular networks," *2013 IEEE Consumer Communications and Networking Conference (CCNC)*, Jan. 2013, 625-630.

[104] Z. Zhou, M. Dong, K. Ota, J. Wu, and T. Sato, "Energy Effciency and Spectral Efficiency Tradeoff in Device-to-Device (D2D) Communications", accepted for publication in *IEEE Wireless Communications Letters*, 2014.

[105] G. Fodor and N. Reider, "A Distributed Power Control Scheme for Cellular Network Assisted D2D Communications," *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, Dec. 2011, 1-6.

[106] M. G. Rego, T. F. Maciel, H. H. M. Barros, F. R. P. Cavalcanti, and G. Fodor, "Performance analysis of power control for device-to-device communication in cellular MIMO systems," *2012 International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2012, 336-340.

[107] N. Reider and G. Fodor, "A distributed power control and mode selection algorithm for D2D communications," *EURASIP Journal on Wireless Communications and Networking*, 2012, 1-25.

[108] F. Wang, Ch. Xu, L. Song, Z. Han, and B. Zhang, "Energy-efficient radio resource and power allocation for device-to-device communication underlaying cellular networks," *2012 International Conference on Wireless Communications and Signal Processing (WCSP)*, Oct. 2012, 1-6.

[109] E. Yaacoub and O. Kubbar, "Energy-efficient Device-to-Device communications in LTE public safety networks," *2012 IEEE Globecom Workshops (GC Wkshps)*, Dec. 2012, 391-395.

[110] E. Yaacoub and A. Abu-Dayya, "Enhancing the QoS of real-time video streaming over LTE MBMS using D2D communications," *8h ACM symposium on QoS and security for wireless and mobile networks*, 2012.

[111] E. Yaacoub, H. Ghazzai, M. S. Alouini, and A. Abu-Dayya, "Achieving energy efficiency in LTE with joint D2D communications and green networking techniques," *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, July 2013, 270-275.

[112] G. Zhang, "Subcarrier and bit allocation for real-time services in multiuser OFDM systems," *IEEE International Conference on Communications (ICC)*, 2004, 2985-2989.

[113] Y. Gao, H. Xu, H. Tian, and P. Zhang, "A QoS-Guaranteed Adaptive Resource Allocation Algorithm with Low Complexity in OFDMA System," *IEEE Wicom06*, 2006, 1-4.

[114] A. De Domenico, E. C. Strinati, and A. Capone, "Enabling Green cellular networks: A survey and outlook," *Computer Communications*, vol. 37, no. 1, Jan. 2014, 5-24.

[115] D. Wang, X. Wang, and Y. Zhao, "An Interference Coordination Scheme for Device-to-Device Multicast in Cellular Networks," *2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sept. 2012, 1-5.

[116] W. Gong, X. Wang, M. Li, and Z. Huang, "Round-robin Resource Sharing Algorithm for Device-to-Device Multicast Communications Underlying Single Frequency Networks," *International Conference on Telecommunications (ICT)*, May 2014, 191-195.

[117] K. Vanganuru, S. Ferrante, and G. Sternberg, "System capacity and coverage of a cellular network with D2D mobile relays," *IEEE Military Communications Conference*, 2012.

[118] Y. Li, D. Jin, F. Gao, and L. Zeng, "Joint Optimization for Resource Allocation and Mode Selection in Device-to-Device Communication Underlaying Cellular Networks," *IEEE International Conference on Communications (ICC)*, June 2014, 2245-2250.

[119] Z. Lin, Y. Li, S. Wen, Y. Gao, X. Zang, and D. Yang, "Stochastic geometry analysis of achievable transmission capacity for relay-assisted Device-to-Device networks", *IEEE International Conference on Communications (ICC)*, June 2014, 2251-2256.

[120] S. Shalmashi and S. B. Slimane, "Cooperative Device-to-Device Communications in the Downlink of Cellular Networks", *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2014.

[121] T. Ta, J. S. Baras, and C. Zhu, "Improving smartphone battery life utilizing device-to-device cooperative relays underlaying LTE networks" *IEEE International Conference on Communications (ICC)*, June 2014, 5263-5268.

[122] S. K. Choi, W. J. Kim, H. S. Lee, and D. I. Kim, "Interference forwarding for D2D based heterogeneous cellular networks," *IEEE/CIC International Conference on Communications in China (ICCC)*, Aug. 2013, 130134.

[123] A. Laya, K. Wang, A. A. Widaa, J. Alonso-Zarate, J. Markendahl, and L. Alonso, "Device-to-device communications and small cells: enabling spectrum reuse for dense networks," *IEEE Wireless Communications*, vol. 21, no. 4, Aug. 2014, 98-105.

[124] A.-H. Tsai, L.-C. Wang, J.-H. Huang, and T.-M. Lin, "Intelligent resource management for device-to-device (D2D) communications in heterogeneous networks," *15th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Sept. 2012, 7579.

[125] Y. He, F. Wang, and J. Wu, "Resource Management for Device-to-Device Communications in Heterogeneous Networks Using Stackelberg Game," *International Journal of Antennas and Propagation*, vol. 2014, May 2014.

[126] 3GPP: Overview of 3GPP Release 12, V0.1.0, Sept. 2013

[127] 3GPP TR 22.803: Technical Specification Group Services and System Aspects; Feasibility study for Proximity Services (ProSe); (Release 12), V12.2.0, June 2013.

[128] 3GPP TS 22.278: Technical Specification Group Services and System Aspects; Service requirements for the Evolved Packet System (EPS), (Release 12), V12.4.0, Sept. 2013.

[129] 3GPP TS 22.115: Technical Specification Group Services and System Aspects; Service aspects; Charging and billing (Release 13), V13.0.0, Dec. 2013.

[130] 3GPP TR 36.843: Technical Specification Group Radio Access Network; Study on LTE Device to Device Proximity Services; Radio Aspects (Release 12), V12.0.1, March 2014.

[131] 3GPP TS 24.333: Technical Specification Group Core Network and Terminals; Proximity-services (ProSe) Management Objects (MO) (Release 12), V1.1.0, July 2014.

[132] 3GPP TS 29.344: Technical Specification Group Core Network and Terminals; Proximity-services (ProSe) Function to Home Subscriber Server (HSS) aspects; Stage 3 (Release 12), V0.2.0, May 2014.

[133] 3GPP TR 33.833: Technical Specification Group Services and System Aspects; Study on Security issues to support Proximity Services (ProSe) (Release 12), V1.0.0, May 2014.

[134] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An Overview on 3GPP Device to Device Proximity Services", *IEEE Communication Magazine*, vol. 52, no. 4, Apr. 2014.

[135] P. Mach and Z. Becvar, "Efficient routing of data for femtocells", *IFIP Wireless Days*, 2011, 1-3.

[136] D. Lee, S.-I. Kim, J. Lee, and J. Heo, "Performance of multihop decode-and-forward relaying assisted device-to-device communication

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/COMST.2015.2447036, IEEE Communications Surveys & Tutorials

underlaying cellular networks," *International Symposium on Information Theory and its Applications (ISITA)*, pp. 455-459, Oct. 2012.

[137] N. Mastronarde, V. Patel, J. Xu, and M. van der Schaar, "Learning relaying strategies in cellular D2D networks with token-based incentives," *IEEE International Workshop on Emerging Technologies for LTE-Advanced and Beyond-4G (IEEE GC Wkshps 2013)*, Dec. 2013.

[138] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, Present, and Future," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 3, Apr. 2012, 497-508.

[139] Skyfiber, "Breaking the backhaul bottleneck: How to meet your backhaul capacity needs while maximizing revenue," whitepaper form series Road to profitable backhaul, 2012.

[140] Ericsson, "It all comes back to backhaul," whitepaper, Feb. 2012.

[141] Qualcomm Technologies, Inc., "Qualcomm Research LTE in Unlicensed Spectrum," *Whitepaper*, June 2014, 1 - 19.

[142] T. Bai, A. Alkhateeb, and R. W. Heath, Jr., "Coverage and Capacity of Millimeter-Wave Cellular Networks," *IEEE Communications Magazine*, Vol. 52, No. 9, Sept. 2014, 70-77.

[143] A.I. Sulyman, A.T. Nassar, M.K. Samimi, G.R. MacCartney Jr., T.S. Rappaport, and A. Alsanie, "Radio Propagation Path Loss Models for 5G Cellular Networks in the 28 GHz and 38 GHz Millimeter-Wave Bands," *IEEE Communications Magazine*, Vol. 52, No. 9, Sept. 2014, 78-87.

Pavel Mach received his MSc and PhD degree from the Czech Technical University in Prague in 2006 and 2010, respectively. At the present, he is a researcher at the Department of Telecommunication Engineering within the same university. He has been actively involved in several national and international projects founded by European Commission in FP6 and FP7 framework such as ROCKET, FREEDOM or TROPIC. In 2015, he has joined 5G mobile research lab funded at Czech Technical University focusing on key aspects and challenges related to the future mobile networks and emerging wireless technologies. He has co-authored more than 50 papers in international journals and conferences. His research interest is mainly related to radio resource management in emerging wireless technologies.

Zdenek Becvar received the MSc and PhD degree in telecommunication engineering from the Czech Technical University in Prague, Czech Republic in 2005 and 2010, respectively. Currently, he is associate professor at the Department of Telecommunication Engineering at the same university.

From 2006 to 2007, he joined Sitronics R&D centre in Prague focusing on speech quality in VoIP. Furthermore, he was involved in research activities of Vodafone R&D center at Czech Technical University in Prague in 2009. He was on internships at Budapest Politechnic, Hungary and CEA-Leti, France in 2007 and 2013, respectively. Since 2007, he continuously participates in projects founded by European Commission. In 2013, he becomes representative of the Czech Technical University in Prague in ETSI and 3GPP standardization organizations. He is a member of more than 15 program committees at international conferences or workshops and he published 3 book chapters and more than 50 conference or journal papers. His current research interests include mobility and radio resource management in future mobile networks.

Tomas Vanek Tomas Vanek received his MSc and PhD degree from the Czech Technical University in Prague in 2000 and 2008, respectively. At the present, he is an assistant professor at the Department of Telecommunication Engineering within the same university. He has been actively involved in several national and international projects founded by European Commission in FP7 framework such as FREEDOM or TROPIC. He has co-authored more than twenty papers in international journals and conferences. His research interests are related to the security in emerging mobile networks and security of multicast and broadcast authentication protocols.