IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 9, 2020, accepted January 27, 2020, date of publication January 31, 2020, date of current version February 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2970829

Statistical Qos Guarantees for Licensed-Unlicensed Spectrum Interoperable D2D Communication

SYED WAQAS HAIDER SHAH[1,2], **(Student Member, IEEE),**
ADNAN NOOR MIAN[2], **(Member, IEEE), AND JON CROWCROFT**[1], **(Fellow, IEEE)**
[1]Computer Lab, University of Cambridge, Cambridge CB3 0FD, U.K.
[2]Department of Computer Science, Information Technology University, Lahore 45000, Pakistan

Corresponding author: Syed Waqas Haider Shah (sw920@cam.ac.uk)

This work was supported by the University of Cambridge. The work of Syed Waqas Haider Shah was also supported by the Commonwealth Scholarship Commission, funded by the UK government.

ABSTRACT In this paper, we propose network-assisted device-to-device (D2D) communication in licensed and unlicensed spectrum interoperable networks, to improve D2D users' throughput while alleviating the spectrum scarcity issue of cellular networks. The idea of licensed and unlicensed spectrum interoperability is based on the findings of the IEEE 1932.1 working group. Conventionally, D2D users were only able to communicate by using either cellular or non-cellular networks and no interoperability mechanism was available. The proposed scheme brings in many benefits including but not limited to higher D2D users' throughput, alleviation in spectrum scarcity issue of cellular networks, and better network management. However, ensuring quality-of-service (QoS) in this dynamic environment is a challenging task. To this end, we analyze the QoS using a well-known analytical tool "Effective Capacity (EC)" for eNodeB-assisted as well as WiFi-assisted D2D communication. Moreover, we also see the impact of neighboring cells' load and full-duplex transceiver at eNodeB and WiFi access point on the EC of D2D users. Simulation results show that EC increases with a decrease in neighboring cell's load and decreases when more stringent QoS constraints are imposed. Results also show that the maximum sustainable source rate at the transmitter's queue increases with an increase in maximum allowed packet delay but converges to a maximum value soon after that.

INDEX TERMS Licensed-unlicensed spectrum interoperatability, D2D communication, effective capacity, quality-of-service.

I. INTRODUCTION

The telecom sector has witnessed an exponential growth in the number of connected devices in the last decade. This trend is expected to thrive even faster with the introduction of upcoming fifth-generation (5G) cellular networks. This huge number of connected devices and spectrum usage associated with it will put a huge strain on the traditional cellular networks. Traditional cellular networks, which were primarily designed for voice-only communication, may not provide high data rate connectivity to new applications. Moreover, the spectrum scarcity issue is a bottleneck for the cellular networks and the telecom community is actively looking for alternate spectrum and technologies which can help in envisioning the true breadth of 5G. This can be seen from the fact that the spectrum available to the cellular network providers remains the same (top five cellular network providers in the United States owns only approximately 140 MHz [1]) while the number of users and their demand of high data rate is increasing exponentially. This has sparked the idea of evolving the current cellular networks to use the unlicensed spectrum. Unlicensed spectrum, on the other hand, is mostly underutilized and can accommodate a large number of users with high data rate requirements.

5G cellular network promises to solve the high data rate as well as huge connectivity issues of the legacy cellular networks [2]. This promise can only come true either with the introduction of an entirely new spectrum (such as above 6 GHz and millimeter-wave communication spectrum) or by incorporating the underutilized unlicensed spectrum. Extensive research is going on in both of the directions, however, utilizing unlicensed spectrum is much easier than that of an

The associate editor coordinating the review of this manuscript and approving it for publication was Bhaskar Prasad Rimal.

entirely new spectrum due to its availability and low cost. There have already been some efforts in realizing the impact of licensed and unlicensed spectrum coexistence such as long term evolution (LTE)-unlicensed [3] also known as LTE-WiFi aggregation (LWA) and licensed assisted access (LAA) [4]. These technologies provide a base for a new paradigm of licensed and unlicensed spectrum interoperability. This paradigm will allow the transfer of higher data volumes with the additional airtime obtained from the unlicensed spectrum. However, this paradigm is still under investigation for its radio access network (RAN) and backhaul network management. Once established, it will help in reducing the spectrum scarcity issue of the 5G cellular networks.

On the other hand, device-to-device (D2D) communication is also one of the key enabling technologies of 5G cellular networks. It allows two devices to communicate with each other without the help of a base station (BS) by reusing the cellular user's bandwidth [5]. By reusing the cellular users' bandwidth, it allows the network to accommodate more and more users without putting a strain on the network. D2D communication can be done in the overlay (underlay) fashion in which cellular and D2D users are assigned orthogonal (nonorthogonal) resources. When utilizing D2D communication in underlay fashion, D2D users', as well as network throughput, decreases due to an increase in interference. This issue can be dealt with by utilizing interference management techniques while performing resource allocation for underlay D2D communication. The authors in [6] have proposed a matching game approach for mode selection and resource allocation in underlay D2D communication which also provides efficient interference management. Similarly, transmit power optimization and D2D user clustering can also be used for interference management. To this end, the authors in [7] have studied the impact of user clustering and transmit power assignment on the overall network performance. They have maximized the sum-rate of the network by jointly optimizing the D2D users' transmit power and user clustering.

To alleviate the spectrum scarcity problem in cellular networks, network densification by deploying small cells is considered as one of the most feasible solutions [8]. However, achieving capacity growth through network densification will ultimately experience severe inter-cell interference. D2D communication in this new paradigm of licensed and unlicensed spectrum interoperability can help in alleviating the spectrum scarcity problem by increasing the licensed spectrum reuse as well as utilizing the unlicensed spectrum. This way, exponential growth in capacity can be achieved without experiencing inter-cell interference. Moreover, it can also help in realizing the energy-efficient (due to reduced transmission power levels) and low-latency (because of the direct link) communication systems, which are the key components of the 5G cellular networks. Various studies have already been done on the utility and benefits of licensed and unlicensed spectrum interoperability. The authors in [9] have proposed an energy-efficient mechanism for LAA. They have investigated the energy-efficient optimization of the LAA system by joint optimization of licensed and unlicensed spectrum allocation. Their simulation results have shown that the energy efficiency of the BS can only be increased when the number of resource blocks from the licensed and unlicensed spectrum is balanced properly, otherwise, it may enhance the throughput of the BS but not the energy efficiency. The authors in [10] have proposed a software-defined heterogeneous network (HetNets)-based multi-scale dynamic allocation of licensed and unlicensed spectrum. They have also presented four potential application scenarios (machine-to-machine communication, multi-access edge computing, ultra-dense networks, and space-air-ground integration) to highlight the benefits of the proposed framework. Their simulation results have shown the efficacy of their proposed framework which can dynamically allocate licensed and unlicensed spectrum by keeping in view the channels state information, interference level, and demand of the users. Nevertheless, licensed and unlicensed interoperability still faces many critical challenges including inter-channel and intra-channel interference, spectrum allocation and network management, and quality-of-service (QoS) provisioning.

QoS provisioning is one of the key challenges of the licensed and unlicensed spectrum interoperable wireless systems. In wireless communication channels, ensuring QoS is a challenging task due to the harsh and rapidly changing environment which have a direct impact on the instantaneous transmission rates of the channel. This task becomes more challenging when licensed and unlicensed channels coexist. This compelled researchers to look for statistical QoS guarantees instead of deterministic QoS guarantees. To find the statistical QoS guarantees, EC is one of the most influential analytical tools. EC is a link-layer channel model in which a channel can be modeled in terms of delay violation probability and the probability of having non-empty buffer [11]. Due to the realistic statistical QoS guarantees it offers, EC has sparked great interest among the researchers as numerous works have been reported in the literature. To date, EC based statistical QoS guarantee analysis has been carried out for various different wireless communication scenarios including but not limited to D2D [12], cognitive radio networks [13], wireless local area networks (WLANs) [14], wireless sensor networks (WSNs) [15], 5G cellular networks [16], to name a few selected works. On the other hand, D2D communication has been widely studied but the integration of D2D communication in licensed and unlicensed spectrum interoperable wireless systems is still a question mark. *To the best of the authors' knowledge, this is the first study which provides statistical QoS guarantees for the licensed and unlicensed spectrum interoperable D2D communication.*

A. CONTRIBUTIONS

This work has the following contributions.

1) We propose a model for licensed and unlicensed spectrum interoperable D2D communication. We also put forward a network-assisted D2D communication model in LTE and WiFi interoperable networks. We investigate device

discovery and handover mechanism for the network-assisted interoperable D2D communication.

2) To ensure QoS in network-assisted interoperable D2D communication, we perform EC analysis which provides statistical QoS guarantees under varying channel conditions. We also investigate the impact of full duplex communication and network cell load on the EC of network-assisted communication.

3) Last but not the least, we calculate the maximum sustainable source rate at the transmitter's queue under varying channel conditions along with different cell load.

B. ORGANIZATION

The remainder of the paper is organized as follows. Section II describes the licensed and unlicensed spectrum interoperability. Section II-A proposes the licensed and unlicensed spectrum interoperable D2D communication. Section II-B presents the D2D device discovery mechanism for interoperable D2D communication. Section II-C presents the handover mechanism for interoperable D2D communication. Section III discusses the QoS guarantees for the proposed interoperable D2D communication. Section III-A presents some background on effective service capacity and Section III-B presents the EC for the eNodeB-assisted as well as WiFi-assisted D2D communication. The EC and maximum sustainable arrival rate at the source' queue has been investigated using simulation results in Section IV. Finally, concluding remarks and some future directions are presented in Section V. For the readers' facilitation, TABLE 1 shows all the mathematical notations used in this paper for convenient referencing.

TABLE 1. **Mathematical notations.**

Notation	Description
P_D	Probability of D2D device discovery
γ^{max}	SINR threshold for device discovery
$S(t)$	Service rate of the transmitter' queue
$r(t)$	Source rate of the transmitter' queue
Q_t	Steady-state queue length
θ	Quality-of-service exponent
P_{out}	Outage probability
$d(t)$	Delay experienced by a packet at time t
d_{max}	Maximum delay bound
$P[Q(t) > 0]$	Probability of non-empty queue at transmitter
$C_{eNB,ul}(t)$	Instantaneous channel capacity of the link D_T →eNodeB
$C_{eNB,dl}(t)$	Instantaneous channel capacity of the link eNodeB→ D_R
$\gamma_{eNB,ul}(t)$	SSINR of the link D_T →eNodeB
$\gamma_{eNB,dl}(t)$	SINR of the link $eNodeB \rightarrow D_R$
ξ_n	Actual cell load of n neighboring cells
ξ_D^c	Minimum resources allocated to device D in a cell C
P_{eNB}^c	Average transmit power of eNodeB of cell C
$r_{max,eNB}^*$	Maximum constant arrival at D_T queue in eNodeB-assisted D2D communicatin
$C_{AP,ul}(t)$	Instantaneous channel capacity of the link $D_T \rightarrow AP$
$C_{AP,dl}(t)$	Instantaneous channel capacity of the link $AP \rightarrow D_R$
$\Gamma_{AP,ul}(t)$	SSINR of the link $D_T \rightarrow AP$
$\Gamma_{AP,dl}(t)$	SINR of the link $AP \rightarrow D_R$
P_{AP}^m	Average transmit power of AP of cell m
$r_{max,AP}^*$	Maximum constant arrival at D_T queue in WiFi-assisted D2D communicatin

II. LICENSED AND UNLICENSED SPECTRUM INTEROPERABILITY

One of the major challenges cellular networks are facing today is spectrum scarcity. It is expected to become more severe if the telecom community only relies on the licensed spectrum to provide connectivity to its massive number of users and their ever-increasing demands of high data rates. Moreover, cellular networks are lagging in providing high QoS due to their limited resources and radio interference [17]. Keeping in view the resource-constrained nature of the cellular networks, the institute of electrical and electronics engineers (IEEE) has formed a working group to investigate the possibility of the interoperation of licensed and unlicensed spectrum. They have proposed a new standard IEEE 1932.1 which will define a mechanism for devices/users operating in licensed and unlicensed spectrum to efficiently communicate with each other. This standard will explain the interoperation among medium access control and physical layer protocols designed specifically for technologies operating in licensed and unlicensed spectrum. Moreover, it will also provide a controller which can coordinate among devices/users communicating using these technologies.

Although interoperation among entities operating in different frequency spectrums brings in many benefits for cellular as well as non-cellular users, it comes with several challenges including coexistence, design, and infrastructure challenges. Coexistence is considered to be one of the major challenges of cellular networks in the unlicensed band because it can cause a significant amount of performance degradation to other non-cellular technologies operating in the same frequency band. In order to get maximum benefits out of licensed and unlicensed interoperability, this coexistence problem must be resolved to ensure QoS of the cellular as well as non-cellular users. The authors in [18] have proposed a game-theoretic approach for fair-coexistence between cellular and non-cellular systems which maximizes the sum-rate of cellular users while considering the QoS requirements as well as the coexistence issue of the non-cellular users. Another study [19] presents a similar approach to the quality of experience-enabled unlicensed sharing in 5G cellular networks. On the other hand, design and infrastructure challenges should also be considered while practically implementing licensed and unlicensed spectrum coexisted networks. Control data separation architecture (CDSA) is one of the promising network architectures [20]. In CDSA, a logical separation between the data plane and the control plane exists. By utilizing CDSA in licensed and unlicensed interoperable networks, we can restrict all the control signaling to only use LTE/5G control channels, whereas a dynamic network controller (DNC) will decide for users/devices to use either licensed or unlicensed spectrum for data transmission. The DNC decision can be based on licensed spectrum congestion or the cell load. If the licensed spectrum is congested, DNC may direct users/devices through control channels to opt unlicensed spectrum for data transmission, or vice-versa. This way licensed carriers can leverage extra bandwidth from

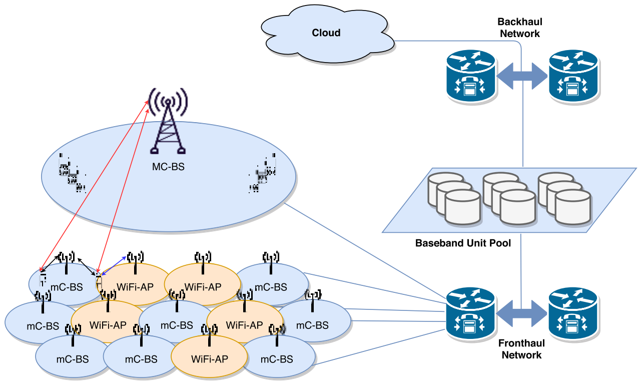

FIGURE 1. **Licensed and unlicensed spectrum interoperable network with cloud radio access network (C-RAN): control data split architecture is used, solid red arrow shows the LTE/5G control signaling, black solid arrow shows the data communication through licensed spectrum, and blue solid arrow shows the data communication through unlicensed spectrum.**

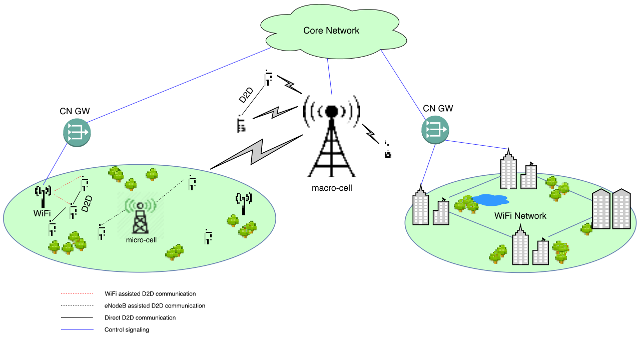

FIGURE 2. **System model of licensed and unlicensed spectrum technologies coexistence: .**

unlicensed spectrum without incurring overhead expenses for obtaining a license. A cloud radio access network (C-RAN) based architecture for licensed and unlicensed spectrum interoperability is shown in Fig. 1. This architecture has four basic units: remote radio head (RRH) comprising both the macro-cell (MC) and micro-cell (mC) deployment, fronthaul network, baseband unit pool which performs computation, storage, processing, and network management, and a backhaul network. This C-RAN based architecture not only allows licensed and unlicensed operating networks to coexists, but it can also reduce network-level energy consumption, the complexity of the RRH network, and cost associated with network deployment and operations [21].

The utility of licensed and unlicensed operable networks can only be recognized when it can accommodate all of the key enabling technologies of the upcoming 5G cellular networks. D2D communication is one of those technologies which have proved its efficacy for 5G cellular networks [22]. Now we will investigate how D2D communication can be efficiently integrated with this new architecture and what benefits can be availed from this integration.

A. LICENSED AND UNLICENSED SPECTRUM INTEROPERABLE D2D COMMUNICATION

Today's cellular networks can be considered as a multiple layered-network with D2D, femto, pico, and wifi networks forming different layers of the communication network underlaid the main MC layer. The introduction of each layer has its benefits just like D2D communication. D2D communication was originally been proposed as a short-range direct communication between two devices without the involvement of the network infrastructure. However, with the ever-increasing demand for high data rate and QoS enabled applications, D2D communication is no more an independent mode of communication without the involvement of access point (in case of WiFi network) or eNodeB (in case of LTE network). D2D communication is mainly characterized by two phases: the device discovery phase and the communication phase. Generally, the device discovery phase is done by the network operator and it shares knowledge of this phase to devices present in close proximity to each other. After this, devices communicate with each other on a direct D2D communication link. However, this way network has a lose control over the D2D communication and it can not ensure reliability and QoS of the communication link [23]. On the other hand, if the network retains full control over both of the device discovery and communication phases, it can ensure reliability as well as QoS of the communication link. In this case, there are two possible ways, one is when the entire cellular infrastructure is used for the communication (which is known as cellular mode of D2D communication) and another possible way is that both D2D devices communicate with each other through eNodeB (but not utilizing the whole infrastructure of the cellular network). This is only possible when a dedicated bearer is established for each device who requests for the establishment of the communication link [24]. This type of D2D communication is known as network-assisted D2D communication. Further details are discussed in Section II-C.

To integrate D2D communication in licensed and unlicensed interoperable networks, network-assisted D2D communication can be considered as one of the most efficient approaches. From the network design point of view, eNodeBs and WiFi access points (APs) are the closest points of interaction for the D2D devices; thus, they can be considered for the integration of all of the D2D functions. When an eNodeB and WiFi AP are equipped with D2D functions, they can perform device discovery as well as allow D2D devices to communicate through eNodeB (in eNodeB-assisted D2D communication) and WiFi AP (in WiFi-assisted D2D communication) as shown in Fig. 2. This way network can ensure service authorization and device authentication aspects in the device discovery phase and QoS and reliability in the communication phase. Network-assisted D2D communication can also allow switching among the networks (switching from licensed to unlicensed network and vice-versa). This switching is done by the DNC depending upon the available spectrum and the data rate requirement of the concerned D2D devices. A detailed discussion on the functions of DNC and how D2D devices can seamlessly switch among different networks is presented in the following subsection.

Recently, a couple of works on D2D communication in cellular and non-cellular networks have emerged. The authors in [25] have proposed a model in which D2D users can access the unlicensed spectrum as an underlay of the uplink LTE network. The authors have also proposed a duty-cycle based protocol for D2D users to utilize the unlicensed spectrum. However, this work does not provide any framework for the interoperation between cellular and non-cellular networks, similarly, there is no mechanism available for D2D users' handover between cellular and non-cellular networks. [26] studies the joint channel and power allocation for D2D communication on licensed and unlicensed bands. The authors have proposed an algorithm using particle swarm optimization to manage the interference and to improve the overall throughput of D2D users. [27] investigate the spectrum access problem for cellular and D2D users by maximizing the total throughput of the network which is an entirely different problem as compared to ours. [28] proposed a new scheme for the medium access control protocol for D2D communication in an unlicensed spectrum. They have proposed the request-to-send/clear-to-send mechanism with the free-to-receive technique to improve the performance of the D2D communication in the unlicensed spectrum. However, none of these works discuss the interoperation and QoS of the licensed and unlicensed-enabled D2D communication.

B. DEVICE DISCOVERY FOR THE INTEROPERABLE D2D COMMUNICATION

Device discovery is one of the most important tasks for establishing D2D communication. The more efficient device discovery is, the more reliable the D2D link will be. It can be done in two different ways, one is autonomous, in which a D2D device itself indicates its presence to other devices present in the network by sending beacons and establishing a neighboring device list. This list is updated regularly based on any change in the network. The second and more efficient way is network-assisted device discovery. In this type of D2D device discovery, eNodeB exploits the information of the network layout, device's distance from eNodeB, and the angle of arrival of the signal to identify the proximity of the D2D device. It is a centralized device discovery mechanism and has the potential to substantially reduce the power consumption of resource-constrained D2D devices, network interference, and signaling required for device discovery [29]. Above that, current cellular networks already have functionalities [30] which can support this type of device discovery, which makes it more suitable for deployment purposes.

Two devices can only be considered nearby when the received signal strength from the transmitting D2D device is greater than or equal to the receiver device's sensitivity. Based on this assumption, the probability of D2D device discovery can be defined as,

$$\P_D = \P[P_t.Z_{D_t,D_r} \geq P_r] \tag{1}$$

where P_t is the transmit power of the transmitting device, Z_{D_t,D_r} are the channel coefficients of the link between transmitting and receiving device, and P_r is the receiver device's sensitivity. P_r is mostly considered as fixed as it depends on the physical parameters of the receiver's antenna. It is impractical to use (1) to find the actual probability of D2D device discovery because it does not consider the impact of signal to interference and noise ratio (SINR) and the varying channel conditions. Moreover, we are also considering full-duplex communication in our scenario so, the impact of self-interference should be considered while defining the probability of D2D device discovery. Nevertheless, (1) can be considered as the upper bound on the performance of $\P_D$. For a more realistic D2D device discovery probability, we consider these factors in our analysis. The probability of device discovery would become,

$$\P'_D = \P[\gamma^{FD} \geq \gamma^{max}]. \tag{2}$$

where γ^{FD} is the signal to self-interference and noise ratio (SSINR) at the receiver, it also incorporates interference caused by the interfering users and γ^{max} is the SINR threshold which is directly dependant on the receiver's sensitivity (P_r). γ^{FD} can be defined as,

$$\gamma^{FD} = \frac{P^i_t.G_t.Z_{D_t,D_r}}{\sum_{m=1,m\neq i}^{M} P^m_t.G_m.Z_{D_m,D_r} + \alpha P^{\beta}_{t_o} + N_o}. \tag{3}$$

where $\sum_{m=1,m\neq i}^{M} P^m_t.G_m.Z_{D_m,D_r}$ is the cumulative interference caused by M interfering devices present in the network, $\alpha P^{\beta}_{t_o}$ is the residual self-interference caused by the full-duplex radio at the receiver, α and $0 \leq \beta \leq 1$ are the self-interference cancellation (SIC) factors and $\beta = 0$ reflects perfect SIC. P_{t_o} is the transmitting power of the receiver device and N_o is the noise variance. When the receiver device is highly sensitive then the probability of D2D device discovery will increase and vice versa. Similarly, increasing P_t will increase P'_D but it will also increase the overall interference in the network; thus, optimization of transmit power is required for better performance. However, such analysis is beyond the scope of this work and thus left for future work. One can also find the probability of D2D device discovery in the unlicensed network by simply calculating SINR in that case. Note that SINR in the unlicensed network will be worse than the one in the licensed network due to higher interference caused by the appliances which are not part of the designated network such as microwave ovens, baby monitors, and etc.

C. HANDOVER MECHANISM FOR THE INTEROPERABLE D2D COMMUNICATION

In network-assisted D2D communication, eNodeB or WiFi AP has full control over D2D device discovery as well as the data communication phase. Once the D2D device discovery phase is done, knowledge of this phase is broadcasted to all of the devices present in the network. Discovered D2D devices who want to communicate with each other then send a request to their respective eNodeB to form a network-assisted D2D communication link. After receiving these requests, eNodeB asks the core network (CN) to form a dedicated D2D radio bearer for each of the devices. Once the D2D radio bearer has been established, eNodeB sends a message to each device that their request for establishing a network-assisted D2D communication link has been accepted. Both the devices then communicate with each other via eNodeB or WiFi AP. A signaling call-flow for the establishment of network-assisted D2D communication is shown in Fig. 3. The first priority of the DNC to provide a licensed spectrum to all of the discovered D2D devices who want to communicate. However, when the licensed spectrum is congested or the D2D devices request for a higher data rate which licensed spectrum-based channel can not provide, DNC allocates an unlicensed spectrum to the respective D2D devices. This is called handover between licensed and unlicensed spectrum [31]. A detailed explanation of how this handover mechanism can be used in a network-assisted D2D communication scenario is presented next.

When two D2D devices which are communicating with each other over the licensed spectrum (eNodeB-assisted D2D communication) wants to switch to the unlicensed spectrum (WiFi-assisted D2D communication), respective eNodeB (source BS) generates a request to the target BS (WiFi AP) and to the mobility management entity (MME)[1]. After verifying that it has a free channel, target BS (WiFi) responds to the core network gateway (CN-GW) with the available channel information. CN-GW then shares it with the source BS, which then passes this information to both of the connected D2D devices and initiates an RRC connection reconfiguration. Once this reconfiguration completes, both the devices contact the target BS (WiFi) on the allocated channel. Target BS then sends a path switch request to the CN-GW which responds with the path switch request acknowledgment and release order of both of the devices to target (WiFi AP) and source (eNodeB) BSs, respectively. Both the newly connected D2D devices can now initiate a WiFi-assisted D2D communication procedure, similar to the eNodeB-assisted D2D communication procedure explained earlier. A complete signaling call-flow for the establishment of eNodeB-assisted D2D communication then switching among licensed and unlicensed spectrum and again establishing a WiFi-assisted D2D communication is shown in Fig. 3.

Fig. 4 presents an overview of the interoperable D2D communication. First, the front-haul network performs D2D device discovery using the suggested device discovery protocols by the MME. Every eNodeB and WiFi-AP performs device discovery mechanism independently and then share this knowledge with core (CN-GW + MME) using the backhaul network. If there are no handover requests by the candidate devices then these devices are assigned physical resource blocks (PRBs) by the PRB allocator in their respective network, otherwise, these devices go through handover mechanisms initiated by the source and target BSs before being assigned PRBs. Once the PRBs are assigned, PRB allocator then passes on the information of transmission power levels and designated channels to the candidate D2D devices, which then start communicating using their respective eNodeB and AP in network-assisted D2D communication mode. Now we will investigate the statistical QoS guarantees for eNodeB-assisted as well as WiFi-assisted D2D communication in the following section.

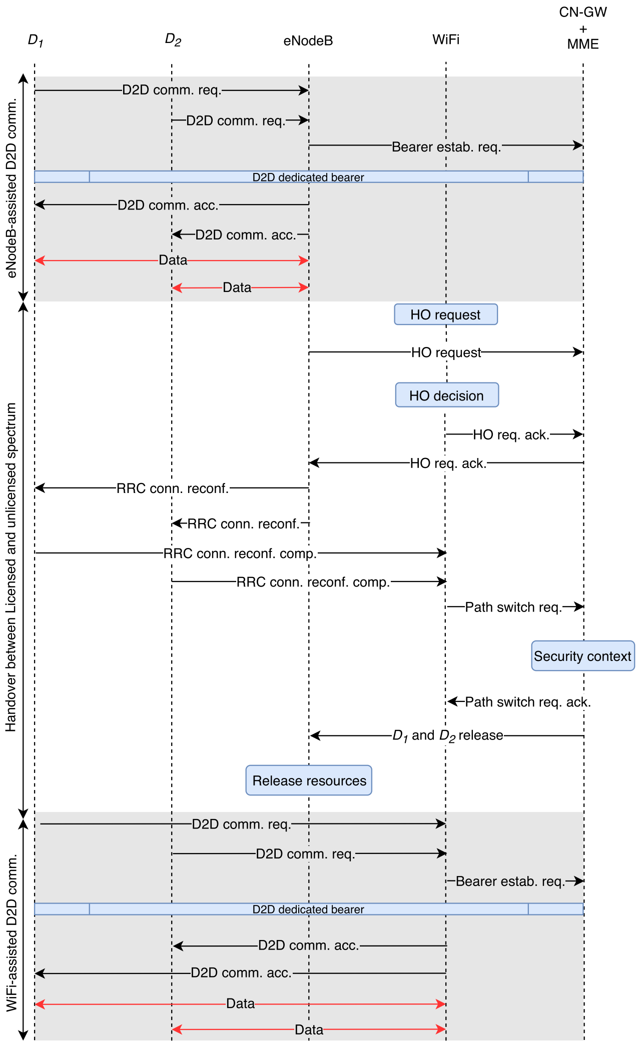

FIGURE 3. Complete signaling call-flow for the establishment of eNodeB-assisted and WiFi-assisted D2D communication as well as handover mechanism for licensed and unlicensed spectrum interoperability.

III. QOS GUARANTEES FOR THE INTEROPERABLE D2D COMMUNICATION

In wireless communication scenarios, ensuring QoS is a challenging task due to the harsh environment which has a direct impact on the instantaneous transmission rates of the channel. In the case where licensed and unlicensed operating

---

[1] Note that control signaling of the WiFi APs are managed by the macro-cell base station (MC BS); thus, RRC messages of the WiFi users are transmitted from the MC BS.

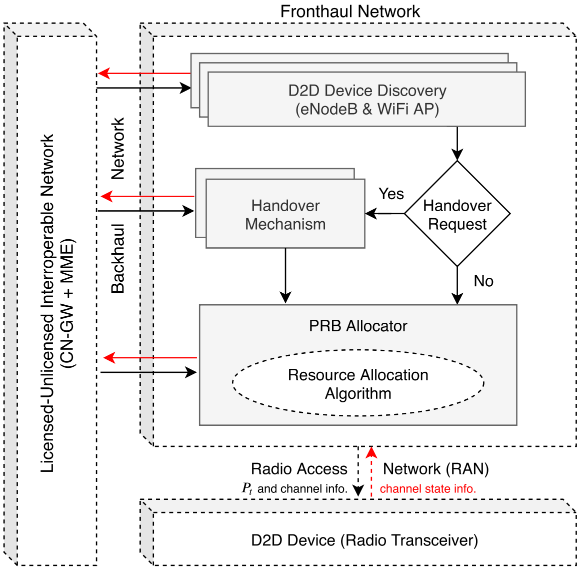

FIGURE 4. **Interoperable D2D communication overview: solid red and black arrows shows the up-link and down-link backaul transmission, respectively and stripped red and black arrows represent the up-link and down-link transmission in RAN.**

networks coexist, this task becomes even more challenging due to inter as well as intra-carrier interference. Moreover, QoS requirements of licensed and unlicensed networks are different and when they coexist ensuring end-to-end QoS guarantees are difficult. Key factors that affect the QoS of any wireless network are a delay and outage probability. Since the channel conditions of a wireless network change rapidly over time, deterministic delay guarantees are difficult to achieve. In such a case, statistical QoS guarantees are preferred because they are easier to map when shadowing and fading affect the channel conditions. To find statistical QoS guarantees, EC is considered as one of the most influential analytical tools.

A. EFFECTIVE CAPACITY

In this subsection, we will present the basic concept of EC. Since EC is a concept that utilizes the effective bandwidth as presented by Wu. [11], we will first explore the idea of effective bandwidth. It defines as the minimum constant service rate which is required to satisfy a queueing delay requirement for a given source rate [32]. Let's say a steady-state queue length at the transmitter Q_t has a source rate $r(t)$ and service rate (which can also be termed as channel capacity) $S(t)$ which changes over time t, then in order for the queue to be stable, the following condition must satisfy,

$$\mathbb{E}[r(t)] \leq \mathbb{E}[S(t)].$$

where $\mathbb{E}[.]$ defines the expectation operator. For any system operating under statistical QoS guarantees, some constraints on Q_t should be imposed so that data to be transmitter should not wait for too long. These constraints are identified by the QoS exponent θ,

$$\lim_{Q_{th}\to\infty} \frac{\log P[Q_t > Q_{th}]}{Q_{th}} = -\theta \tag{4}$$

where Q_{th} is the maximum threshold on the queue length. Packets are generally lost when queue (buffer) becomes full. If (4) is satisfies, then the buffer violation probability at the transmitter can be written as,

$$P[Q_t > Q_{th}] \approx e^{-\theta Q_{th}} \tag{5}$$

From (5), we can say that θ is the exponential decay rate of the buffer violation probability at the transmitter. A lower θ entails a higher queue length violation probability which implies a relax QoS constraint. On the other hand, a large θ refers to a low queue length violation probability which ultimately results in stringent QoS constraint. Moreover, we can also say that $\theta \to 0$ implies delay-tolerant communication, while $\theta \to \infty$ implies delay-sensitive communication. Based on these principles, authors in [11] have presented the idea of EC for wireless communication networks.

Since, the effective bandwidth is defined as the minimum constant service rate required to satisfy a queueing delay requirement, EC, on the other hand, is the maximum sustainable constant arrival rate at a transmitter's queue in the face of a randomly time-varying (channel) service, under given QoS constraints. It can be defined as the log moment generating function (LMGF) of the cumulative service process $S(t)$ in the limit:

$$\mathrm{EC}(\theta) = -\frac{\Lambda(-\theta)}{\theta} = -\lim_{t\to\infty} \frac{1}{\theta t} \log \mathbb{E}(e^{-\theta S(t)}) \text{ [bits/slot]} \tag{6}$$

where $S(t) = \sum_{k=1}^{t} s(k)$, with $s(k)$ as the channel service (i.e., number of bits delivered) during slot k. In order to find the delay experienced by a packet at any time t, the probability for non empty queue/buffer can be used,

$$P_{out} = P[d(t) > d_{max}] \approx P[Q_t > 0]e^{-\theta.\mathrm{EC}.d_{max}}. \tag{7}$$

where P_{out} is the outage probability, $d(t)$ is the delay experienced by a packet at any time t, d_{max} is the maximum delay bound, and $P[Q_t > 0]$ is the probability of a non empty queue.

B. EC OF THE INTEROPERABLE D2D COMMUNICATION

As we know from (6) that EC is,

$$\mathrm{EC}(\theta) = -\frac{\Lambda(-\theta)}{\theta} = -\lim_{t\to\infty} \frac{1}{\theta t} \log \mathbb{E}(e^{-\theta S(t)}) \text{ [bits/slot]} \tag{8}$$

if we assume service rate ($S(t)$) as independent and identically distributed (IID) and no correlation among different samples of $S(t)$ over t where ($t = 1, 2, \ldots, T$) then $S(t)$ can be considered as addition of T random variables. As in (5), $t \to \infty$ then according to law of large numbers [33], a simplification of LMGF can be written as,

$$EC(\theta) = -\frac{\Lambda(-\theta)}{\theta} = \mathbb{E}[S(t)] - \frac{\theta}{2}\mathrm{Var}[S(t)] = m_s - \frac{\theta}{2}\sigma_s^2 \tag{9}$$

where m_s and σ_s^2 are the mean and variance of the service rate $S(t)$. Hence, in order to find the EC of the D2D communication link (either eNodeB-assisted or WiFi-assisted), finding mean and variance of $S(t)$ is enough. As already mentioned in the previous section that in order for the queueing system to be stable, the QoS exponent θ must fulfil this condition, $r < \frac{\Lambda(-\theta)}{\theta}$. Utilizing this along with (9), we can write general expression for θ as $\theta = \frac{2(m_s-r)}{\sigma_s}$. If we utilize the expression $P[d(t) > d_{max}] \approx P[Q_t > 0]e^{-\theta.EC.d_{max}}$ given in (7) and upper bounding probability of non-empty queue ($P[Q_t > 0]$) by 1, then we can find the maximum arrival rate r^* that can be supported by $S(t)$,

$$r^*\theta \approx \frac{-\ln(P_{out})}{d_{max}}. \tag{10}$$

now substituting $\theta = \frac{2(m_s-r)}{\sigma_s}$ in (10),

$$\begin{aligned} r^*\frac{2(m_s-r)}{\sigma_s} &\approx \frac{-\ln(P_{out})}{d_{max}} \\ r^*2(m_s-r) &\approx \frac{-\ln(P_{out})\sigma_s}{d_{max}} \end{aligned} \tag{11}$$

by simplifying it, we can derive the maximum sustainable source rate r^* that can be supported by $S(t)$ [34],

$$r^* = \frac{1}{2}m_s + \frac{1}{2}\sqrt{(m_s)^2 + 2\frac{\ln(P_{out})}{d_{max}}\sigma_s^2}. \tag{12}$$

Fig. 5 shows the system model for eNodeB-assisted and WiFi-assisted D2D communication. First, we will investigate the EC of eNodeB-assisted D2D communication.

1) ENODEB-ASSISTED D2D COMMUNICATION

We assume that eNodeB supports full-duplex communication (i.e., it receives data from D_T and transmits to D_R simultaneously in a single time frame). Although full-duplex communication can enhance the system capacity (double in some cases), it also introduces self-interference [35] as shown in Fig. 5. There are several self-interference cancellation (SIC) techniques that can be used at the physical layer to mitigate this effect. However, complete SIC is not possible using these techniques [36]. We, therefore, incorporate the residual self-interference in our analysis for a realistic scenario. Moreover, we will also include the impact of cell load on the down-link channel capacity. This way, we can see the impact of network densification as well as cell congestion on a user's throughput. Further, we also assume that the transmitter (D_T) has the perfect channel state information (CSI); thus, the service rate ($S(t)$) will be equal to the instantaneous channel capacity of the eNodeB-assisted D2D communication link [2]. Instantaneous channel capacity ($C_{eNB}(t)$) can be written as,

$$C_{eNB}(t) = 0.5\min(C_{eNB,ul}(t), C_{eNB,dl}(t)). \tag{13}$$

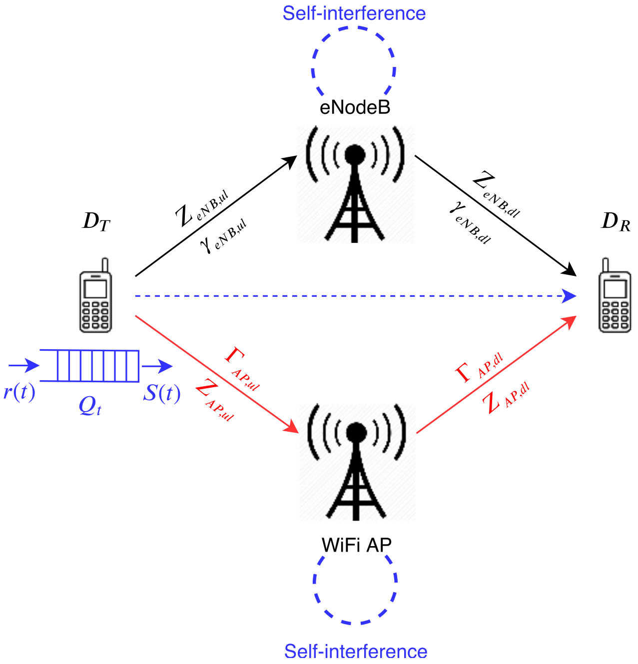

FIGURE 5. Network-assisted D2D communication model in licensed (LTE) and unlicensed (WiFi) interoperable network: D_T communicate with D_R in either eNodeB-assisted D2D communication manner (shown in solid black arrows) or WiFi-assisted D2D communication (shown in solid red arrows), D_T has a queue with steady-state queue length Q_t, $r(t)$ and $S(t)$ are the source rate and service rate at the transmitter, respectively.

where $C_{eNB,ul}(t)$ and $C_{eNB,dl}(t)$ are the instantaneous channel capacities of the $D_T \rightarrow eNodeB$ and $eNodeB \rightarrow D_R$ links, respectively. $C_{eNB,ul}(t)$ can be written as,

$$C_{eNB,ul}(t) = B_c\log_2(1 + \gamma_{eNB,ul}(t)). \tag{14}$$

where $\gamma_{eNB,ul}(t)$ is the SSINR of the uplink transmission and can be calculated as,

$$\gamma_{eNB,ul}(t) = \frac{\bar{P}_t.G_T.Z^c_{eNB,ul}(t)}{N_o + \alpha_1\bar{P}^{\beta_1}_{eNB}}. \tag{15}$$

where $\bar{P}_t$ and G_T are the average transmit power and gain of the D_T, respectively. $\alpha_1\bar{P}^{\beta_1}_{eNB}$ is the residual self-interference due to the poor SIC techniques being used [36]. α_1 and β_1 reflects the quality of SIC techniques used. $\bar{P}^c_{eNB}$ is the average transmit power of the eNodeB of cell c. By substituting (15) in (14), we can find instantaneous channel capacity of the link $D_T \rightarrow eNodeB$. Now the down-link channel capacity can be calculated as follows,

$$C_{eNB,dl}(t) = B_c\log_2(1 + \gamma_{eNB,dl}(t)). \tag{16}$$

where $\gamma_{eNB,dl}(t)$ is the SINR of the down-link channel $eNodeB \rightarrow D_R$. In this case, SINR will also be affected by the cell load of the neighboring cells. This can be expressed as follows,

$$\gamma^c_{eNB,dl}(t) = \frac{\bar{P}^c_{eNB}.G_R.Z^c_{eNB,dl}(t)}{N_o + \sum_{\forall n\in C}\xi_n\bar{P}^n_{eNB}G_RZ^n_{eNB,dl}(t)}. \tag{17}$$

---

[2] If the transmitter in unaware of the perfect instantaneous CSI or only aware of the outdated or average CSI then constant service/transmission rate can be used [33].

where $\bar{P}^{c}_{eNB}$ and $\bar{P}^{n}_{eNB}$ are the average transmit powers of the eNodeB of the connected cell and interfering cells, respectively. G_R is the gain of D_R. $Z^{c}_{eNB,dl}(t)$ and $Z^{n}_{eNB,dl}(t)$ are the channel coefficients of the desired down-link channel and interfering channels, respectively. $\bar{\xi}_n$ is the actual cell load of the interfering cells. In order to calculate cell load of a cell c, we first need to find the minimum resources allocated to a D2D device D. These resources can be calculated by incorporating the throughput requested by the device D (ζ_D) and the actual throughput experienced by the same device [8]. $\xi^{c}_{D} = \frac{1}{w_B}(\frac{\hat{\zeta}_D}{\log_2(1+\gamma^{c}_{D})})$ are the minimum resources allocated to device $D \in U_D$, where U_D represents the set of all active devices connected to cell c. Now by utilizing ξ^{c}_{D}, we can find the total load of cell c, which can be calculated as,

$$\xi_c = \frac{1}{B_c}(\frac{1}{w_B}\sum_{U_D}\frac{\hat{\zeta}_D}{\log_2(1+\gamma^{c}_{D})}). \tag{18}$$

Now by substituting cell load from (18) in (17), we can find SINR of the down-link channel ($eNodeB \rightarrow D_R$) which also incorporates cell load of neighboring interfering cells.

$$\begin{aligned}\gamma^{c}_{eNB,dl}(t) \\ = \frac{\bar{P}^{c}_{eNB}.G_R.Z^{c}_{eNB,dl}(t)}{N_o + \sum_{\forall n \in C}\left[\frac{1}{B_n.w_B}\sum_{U_D}\frac{\hat{\zeta}_D}{\log_2(1+\gamma^{n}_{D})}\right]\bar{P}^{n}_{eNB}G_R Z^{n}_{eNB,dl}(t)}.\end{aligned} \tag{19}$$

By substituting (19) in (16), we can find instantaneous channel capacity of the down-link channel ($eNodeB \rightarrow D_R$). To find the overall instantaneous capacity of the link $D_T \rightarrow eNodeB \rightarrow D_R$, substitute (14) and (16) in (13),

$$\begin{aligned}C_{eNB}(t) &= 0.5B_c\log_2(1+\min\{\gamma_{eNB,ul}(t), \gamma_{eNB,dl}(t)\}) \\ &= 0.5B_c\log_2(1+\gamma_{eNB}(t)).\end{aligned} \tag{20}$$

where $\gamma_{eNB}(t)$ is the overall SINR of the link $D_T \rightarrow eNodeB \rightarrow D_R$ and can be calculated by taking the minimum of the SINR of the individual links $D_T \rightarrow eNodeB$ and $eNodeB \rightarrow D_R$.

Now to find the EC (EC_{eNB}) and maximum sustainable arrival rate in case of eNodeB-assisted D2D communication (r^{*}_{eNB}), we need to find the mean and variance of $C_{eNB}(t)$. Mean can be calculated as follows,

$$\begin{aligned}m_{c_{eNB}} &= \mathbb{E}[0.5B_c\log_2(1+\gamma_{eNB}(t))] \\ &= \frac{B_c}{2\bar{\gamma}_{eNB}}\int_0^{\infty}\log_2(1+x)e^{\frac{-x}{\bar{\gamma}_{eNB}}}dx \\ &= \frac{B_c}{2}\log_2(e)e^{\frac{1}{\bar{\gamma}_{eNB}}}E_1[\frac{1}{\bar{\gamma}_{eNB}}].\end{aligned} \tag{21}$$

where $\bar{\gamma}_{eNB}$ is the average SINR of the link $D_T \rightarrow eNodeB \rightarrow D_R$ and $E_1[.]$ is the exponential integral. Variance can be calculated as shown in (22), as shown at the bottom of this page, where ${}_aF_b\left[\begin{smallmatrix} n \\ d\end{smallmatrix}; z\right]$ is the hypergeometric function. By substituting $m_{c_{eNB}}$ from (21) and $\sigma^2_{c_{eNB}}$ from (22) in (9), we can find EC in case of eNodeB-assisted D2D communication[3] as shown in (23), as shown at the bottom of this page. Similarly, by substituting $m_{c_{eNB}}$ and $\sigma^2_{c_{eNB}}$ in (12), we can find $r^{*}_{max,eNB}$ in case of eNodeB-assisted D2D communication under statistical QoS guarantees θ. After simplification, $r^{*}_{max,eNB}$ is shown in (24), as shown at the bottom of this page.

2) WIFI-ASSISTED D2D COMMUNICATION

In this section, we will evaluate statistical QoS guarantees for WiFi-assisted D2D communication. First, we will compute maximum sustainable source rate r^{*} under explicit QoS constraints (maximum allowed target delay d_{max} and delay violation probability $P[d(t) > d_{max}]$). In this scenario, as the communication will be done on unlicensed frequency spectrum, WiFi AP as well as D_R will experience severe interference from other services co-existing in the industrial, scientific, and medical (ISM) radio band (e.g. baby monitors,

$$\begin{aligned}\sigma^2_{c_{eNB}} &= \mathbb{E}[0.5B_c\log_2(1+\gamma_{eNB}(t))^2] - m^2_{c_{eNB}} = \frac{B_c^2}{4\bar{\gamma}_{eNB}}\int_0^{\infty}\log_2(1+x)^2 e^{\frac{-x}{\bar{\gamma}_{eNB}}}dx - m^2_{c_{eNB}} \\ &= \frac{B_c^2}{4}\log_2(e)^2 e^{\frac{1}{\bar{\gamma}_{eNB}}}[\frac{\pi^2}{6} + \ln(\frac{1}{\bar{\gamma}_{eNB}})^2] - \frac{B_c^2}{4}\log_2(e)^2 e^{\frac{1}{\bar{\gamma}_{eNB}}}(\frac{2}{\bar{\gamma}_{eNB}})_3F_3\left[\begin{matrix}1 & 1 & 1 \\ 2 & 2 & 2\end{matrix}; \frac{-1}{\bar{\gamma}_{eNB}}\right] - (\frac{B_c}{2}\log_2(e)e^{\frac{1}{\bar{\gamma}_{eNB}}}E_1[\frac{1}{\bar{\gamma}_{eNB}}])^2 \\ &= \frac{B_c^2}{4}\log_2(e)^2 e^{\frac{1}{\bar{\gamma}_{eNB}}}\left(\frac{\pi^2}{6} + \ln(\frac{1}{\bar{\gamma}_{eNB}})^2 - \frac{2}{\bar{\gamma}_{eNB}}{}_3F_3\left[\begin{matrix}1 & 1 & 1 \\ 2 & 2 & 2\end{matrix}; \frac{-1}{\bar{\gamma}_{eNB}}\right] - e^{\frac{1}{\bar{\gamma}_{eNB}}}E_1^2[\frac{1}{\bar{\gamma}_{eNB}}]\right).\end{aligned} \tag{22}$$

$$EC_{eNB} = \frac{B_c}{2}\log_2(e)e^{\frac{1}{\bar{\gamma}_{eNB}}}\left[E_1[\frac{1}{\bar{\gamma}_{eNB}}] - \frac{\theta B_c\log_2(e)}{4}\left(\frac{\pi_2}{6} + \ln(\frac{1}{\bar{\gamma}_{eNB}})^2 - \frac{2}{\bar{\gamma}_{eNB}}{}_3F_3\left[\begin{matrix}1 & 1 & 1 \\ 2 & 2 & 2\end{matrix}; \frac{-1}{\bar{\gamma}_{eNB}}\right] - e^{\frac{1}{\bar{\gamma}_{eNB}}}E_1^2[\frac{1}{\bar{\gamma}_{eNB}}]\right)\right]. \tag{23}$$

$$\begin{aligned}& r^{*}_{max,eNB} \\ &= \frac{B_c\log_2(e)}{4}e^{\frac{1}{\sqrt{\bar{\gamma}_{eNB}}}} \\ &\times\left[e^{\frac{1}{\sqrt{\bar{\gamma}_{eNB}}}}E_1[\frac{1}{\bar{\gamma}_{eNB}}] + \sqrt{e^{\frac{1}{\bar{\gamma}_{eNB}}}E_1^2[\frac{1}{\bar{\gamma}_{eNB}}] + \frac{2\ln(P_{out})}{d_{max}}\left(\frac{\pi^2}{6} + \ln(\frac{1}{\bar{\gamma}_{eNB}})^2 - \frac{2}{\bar{\gamma}_{eNB}}{}_3F_3\left[\begin{matrix}1 & 1 & 1 \\ 2 & 2 & 2\end{matrix}; \frac{-1}{\bar{\gamma}_{eNB}}\right] - e^{\frac{1}{\bar{\gamma}_{eNB}}}E_1^2[\frac{1}{\bar{\gamma}_{eNB}}]\right)}\right].\end{aligned} \tag{24}$$

---

[3] we assume that the transmitter (D_T) has the perfect CSI; thus, $S(t)$=$C_{eNB}(t)$.

microwave oven, etc.) as well as from other communicating users in the network [37]. Such uncoordinated interference makes it difficult to ensure QoS on the D2D link, which motivated us to compute the effective service capacity for the WiFi-assisted D2D communication.

Similarly as of eNodeB-assisted D2D communication, we also assume that WiFi-AP supports full-duplex communication; thus, we will incorporate residual self-interference in calculating channel capacity of WiFi-assisted D2D communication as shown in Fig. 5. Due to the presence of DNC, a coordinated WiFi network will also be established so, incorporating cell load of WiFi-cells will also help our analysis depict a more realistic scenario. Moreover, we also assume that D_T has the perfect CSI in this case as well; thus, $S(t)$ will be equal to the instantaneous channel capacity of the WiFi-assisted D2D communication link ($C_{AP}(t)$). This can be calculated as follows,

$$C_{AP}(t) = 0.5 \min\{C_{AP,ul}(t), C_{AP,dl}(t)\}. \tag{25}$$

where $C_{AP,ul}(t)$ and $C_{AP,dl}(t)$ are the instantaneous channel capacities of the link $D_T \rightarrow AP$ and $AP \rightarrow D_R$, respectively. $C_{AP,ul}(t)$ can be calculated as,

$$C_{AP,ul}(t) = B_m \log_2(1 + \Gamma_{AP,ul}(t)). \tag{26}$$

where B_m is the bandwidth allocated by the AP to D_T for up-link communication and $\Gamma_{AP,ul}(t)$ is the SSINR of the up-link transmission and can be calculated as,

$$\Gamma_{AP,ul}(t) = \frac{\bar{P}_t.G_T.Z_{AP,ul}(t)}{N_o + \alpha_2 \bar{P}_{AP}^{\beta_2}}. \tag{27}$$

where $Z_{AP,ul}(t)$ are the channel coefficients of the up-link channel, $\alpha_2 \bar{P}_{AP}^{\beta_2}$ is the residual self-interference and α_2 and β_2 represents the quality of SIC techniques used at the AP. $\bar{P}_{AP}$ is the average transmit power of AP. Now the down-link channel capacity can be calculated as,

$$C_{AP,dl}(t) = B_m \log_2(1 + \Gamma^m_{AP,dl}(t)). \tag{28}$$

where $\Gamma_{AP,dl}(t)$ is the SINR of the down-link transmission and can be calculated as,

$$\Gamma_{AP,dl}(t) = \frac{\bar{P}^m_{AP}.G_R.Z^m_{AP,dl}(t)}{N_o + \sum_{\forall i \in M, i \neq m} \hat{\xi}_i \bar{P}^i_{AP} G_R Z^i_{AP,dl}(t) + \sum^M_{t \neq m} I_t + \sum^L_{l=1} I_l}. \tag{29}$$

where $\sum^M_{t \neq m} I_t$ shows the interference caused by mC BSs, where $t = 1, 2, 3, \ldots, M$ represents all mC BSs present in the network. $\sum^L_{l=1} I_l$ shows the interference caused by all the unlicensed network users present in the network (where $l = 1, 2, 3, \ldots, L$ are the total number of unlicensed users). $Z^m_{AP,dl}(t)$ and $Z^i_{AP,dl}(t)$ are the channel coefficients of the desired downlink channel and interfering channels, respectively. $\hat{\xi}_i$ is the actual cell load of the interfering WiFi cells. To calculate the cell load, in this case, one simply has to follow the way we calculated the cell load in the eNodeB-assisted D2D communication scenario. The final expression for the cell load, in this case, comes out to be,

$$\xi_i = \frac{1}{B_i}(\frac{1}{w_B} \sum_{U_D} \frac{\hat{\zeta}_D}{\log_2(1 + \Gamma^m_D)}). \tag{30}$$

Now by substituting (30) in (29), we can find the final expression of the SINR of the downlink channel $AP \rightarrow D_R$ (as shown in (31), at the bottom of this page) which also incorporates the cell load of neighboring interfering cells as well as the interference caused by neighboring interfering cells and unlicensed users present in the network. Now by substituting (31) in (28), we can find the down-link channel capacity. Now to find the overall capacity of the link $D_T \rightarrow AP \rightarrow D_R$, substitute (26) and (28) in (25),

$$\begin{aligned} C_{AP}(t) &= 0.5 B_m \log_2(1 + \min\{\Gamma^m_{AP,ul}(t), \Gamma^m_{AP,dl}(t)\}) \\ &= 0.5 B_m \log_2(1 + \Gamma_{AP}(t)). \end{aligned} \tag{32}$$

where $\Gamma^m_{AP,dl}(t)$ is the overall SINR of the link $D_T \rightarrow AP \rightarrow D_R$ and can be calculated by taking the minimum of the SINR of the individual link $D_T \rightarrow AP$ and $AP \rightarrow D_R$.

Now to find the EC (EC_{AP}) and $r^*_{max,AP}$, one needs to recompute the mean and variance by substituting $\bar{\Gamma}_{AP}$ in (21) and (22). EC_{AP} and $r^*_{max,AP}$ can then be computed by substituting these new expressions of mean and variance in (9) and (12), respectively.

IV. SIMULATION SECTION

A. SIMULATION SETUP

The simulations were done using MATLAB 2019. We consider a Macro-cell of radius 700 m with multiple heterogenous micro-cells in its coverage area each with a radius of 70 m. D2D users are placed randomly in the coverage area. The average transmit power of D2D users, eNodeB, and WiFi AP are set 1 dBm, 23 dBm, and 10 dBm, respectively. We use following path-loss models for transmission: PL(D) = 128.1+37.6 $\log_{10}$(D) and PL(D) = 140.7 + 36.7 $\log_{10}$(D) for cellular cell and WiFi cell, respectively. Where D is the distance between D2D user and eNodeB and WiFi. We set the range of SINR (γ_{eNB} and Γ_{AP}) between 0-40 dB. β_1 and β_2 are quality of SIC techniques employed at eNodeB and WiFi AP and it ranges from [0,1].

$$\Gamma^m_{AP,dl}(t) = \frac{\bar{P}^m_{AP}.G_R.Z^m_{AP,dl}(t)}{N_o + \sum_{\forall i \in M, i \neq m} [\frac{1}{B_m.w_B} \sum_{U_D} \frac{\hat{\zeta}_D}{\log_2(1+\Gamma^m_D)}] \bar{P}^i_{AP} G_R Z^i_{AP,dl}(t) + \sum^M_{t \neq m} I_t + \sum^L_{l=1} I_l} \tag{31}$$

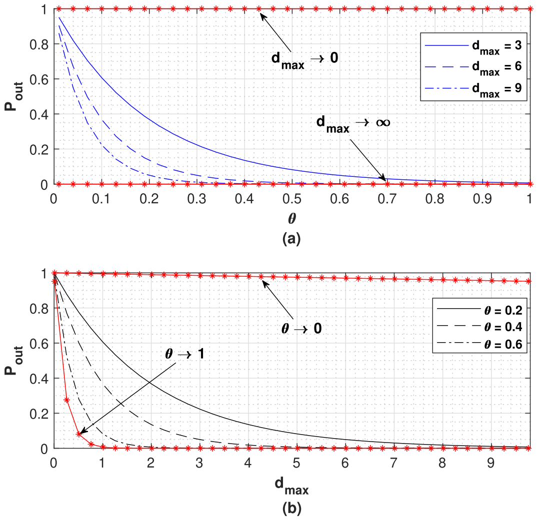

FIGURE 6. **(a) Outage probability (P_{out}) vs QoS exponent (θ) for different maximum allowed delay guarantees (d_{max}), (b) Outage probability (P_{out}) vs maximum allowed delay guarantees (d_{max}) for different QoS exponent (θ).**

B. SIMULATION RESULTS

Fig. 6 presents the outage probability (the probability that delay experienced by a packet is greater than the maximum allowed packet delay) for different parameters of the network. Fig. 6(a) shows that outage probability decreases exponentially fast with an increase in QoS exponent. Lower QoS exponent (which implies a relax QoS constraints) entails a higher outage probability, whereas large QoS exponent leads to a lower outage probability which results in strict QoS constraints. Moreover, the effect of changing the maximum allowed packet delay on the rate of change in outage probability is also observed. When maximum allowed packet delay approaches to zero, outage probability approaches to the maximum. On the other hand, when it approaches infinity, outage probability becomes zero irrespective of the QoS exponent value. Fig. 6(b) presents the effect of maximum allowed packet delay on the outage probability. It is evident from the figure that as the allowed packet delay increases, chances of transmission increase which leads to lesser outage probability and vice versa. Similarly, the effect of the QoS exponent can also be observed from this figure, as QoS exponent approaches zero (delay-tolerant communication environment), outage probability approaches to the maximum. On the other hand, when QoS exponent approaches to maximum (i.e., $\theta \to 1$) (delay-sensitive communication environment), outage probability decreases exponentially fast, in fact, it becomes nearly zero for $d_{max} = 1$.

Fig. 7 demonstrates the EC of the WiFi-assisted D2D communication for different SINR of that link. It shows that the EC increases exponentially fast as the SINR increases for lower QoS exponent and as the QoS exponent increases, this rate of increase decreases rapidly. In fact, for delay-sensitive communication scenario (when $\theta \to 1$), a decline in EC is experienced as the SINR increases for $\bar{\Gamma}_{AP} = 1\ dB \to 6\ dB$ and it becomes zero for $\bar{\Gamma}_{AP} = 6\ dB \to 10\ dB$ and after that it starts increasing though slowly but consistently with the increase in SINR. The highlighted region shows this effect that for higher QoS exponent, EC decreases initially and then starts increasing with the increase in SINR.

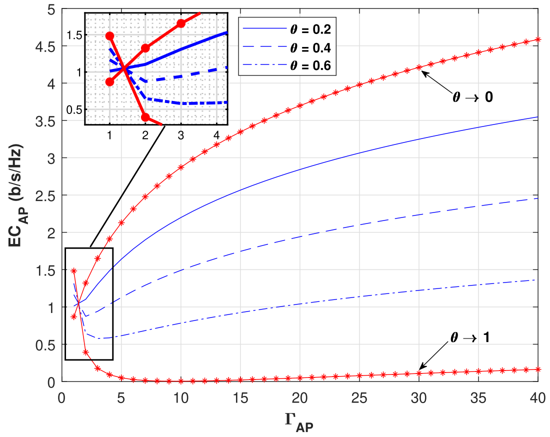

FIGURE 7. **WiFi-assisted D2D communication: effective capacity (EC_{AP}) vs SINR ($\bar{\Gamma}_{AP}$) for different values of QoS exponent (θ).**

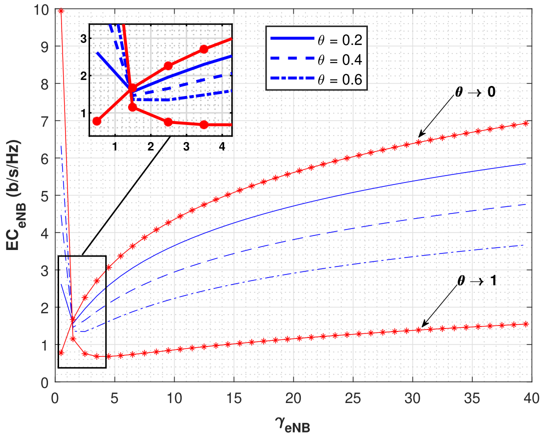

FIGURE 8. **eNodeB-assisted D2D communication: effective capacity (EC_{eNB}) vs SINR ($\bar{\gamma}_{eNB}$) for different values of QoS exponent (θ).**

Fig. 8 presents the EC of the eNodeB-assisted D2D communication. The findings of Fig. 8 are same as for Fig. 7. However, in case of eNodeB-assisted D2D communication, D_T can achieve higher EC when compared to WiFi-assisted D2D communication because a cellular network can provide dedicated channels (in overlay D2D scenario [38]) to connected D2D devices. Transmission using a dedicated channel will allow a D2D device an interference-free communication; thus, the EC will be better.

Fig. 9 presents the maximum arrival rate at the queue of D_T for different values of the maximum allowed packet delay for the eNodeB-assisted D2D communication link. The maximum arrival rate increases exponentially fast with an increase in maximum allowed packet delay but ultimately converges to a maximum value. This maximum value is

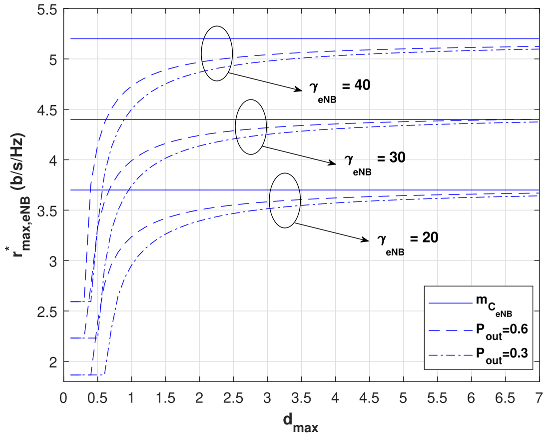

FIGURE 9. eNodeB-assisted D2D communication: maximum sustainable source rate ($r^*_{max,eNB}$) vs maximum delay bound (d_{max}) for different values of SNR ($\bar{\gamma}_{eNB}$).

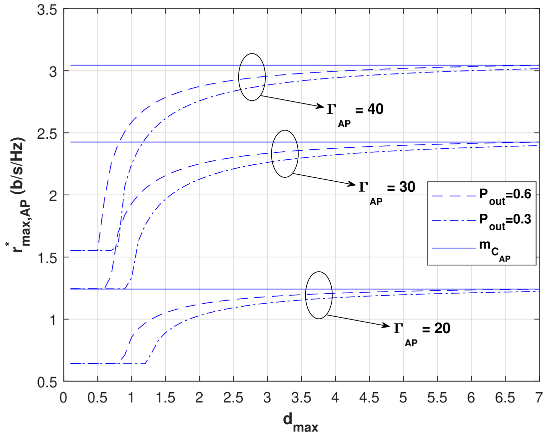

FIGURE 10. WiFi-assisted D2D communication: maximum sustainable source rate ($r^*_{max,AP}$) vs maximum delay bound (d_{max}) for different values of SINR ($\bar{\Gamma}_{AP}$).

different for different SNR of the eNodeB-assisted D2D communication link. For $\bar{\gamma}_{eNB} = 40\ dB$, maximum arrival rate converges to 5.1 $b/s/Hz$ and for $\bar{\gamma}_{eNB} = 30\ dB$ it converges to 4.42 $b/s/Hz$, in short, higher the SNR leads to a higher maximum arrival rate. Moreover, this figure also highlights the effect of outage probability on the maximum arrival rate, higher the outage probability will allow the queue of D_T to support higher arrival rate though the benefit of higher outage probability diminishes with the increase in maximum allowed packet delay.

Fig. 10 also presents the maximum arrival rate at the queue of D_T versus the maximum allowed packet delay but for the WiFi-assisted D2D communication link. The findings of Fig. 10 are same as for Fig. 9. However, in the case of WiFi-assisted D2D communication, the maximum arrival rate at D_T will be less than that of in eNodeB-assisted D2D communication due to the effect of uncorrelated interference caused by the neighboring WiFi cells as well as other devices operating in the unlicensed frequency band. This shows that for a queue at D_T to support the same arrival rate as that of in eNodeB-assisted D2D communication link, D_T must employee better equalization as well as interference-cancellation techniques.

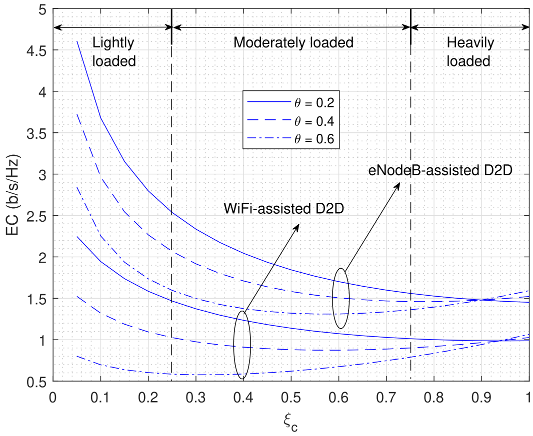

FIGURE 11. Effective capacity (*EC*) vs cell load (ξ_c) for different QoS exponent (θ).

Fig. 11 presents the EC vs different cell load for eNodeB-assisted and WiFi-assisted D2D communication. EC decreases exponentially fast as the cell load increases in the lightly loaded region. The rate of decay of EC reduces in the moderately loaded region, but the decreasing trend remains intact. Whereas, in the heavily loaded region, this rate of decrease becomes almost zero and EC converges to a single value for eNodeB-assisted (WiFi-assisted) D2D communication when cell load is 90% (94%). Moreover, the advantage of delay-tolerant communication over delay-sensitive communication diminishes with the increase in cell load and ultimately becomes negligible when the cell is 90% loaded. This implies that if no proper cell load management techniques are adopted, then ensuring QoS of the D2D link is not achievable.

V. CONCLUSION AND FUTURE DIRECTIONS

In this paper, we proposed a model for the network-assisted D2D communication in licensed and unlicensed interoperable networks. The proposed model provides a coordinated D2D device discovery as well as a handover mechanism in LTE and WiFi networks. Moreover, we also investigate the statistical QoS guarantees for the network-assisted D2D communication in LTE and WiFi networks. The simulation results have shown that the EC, as well as the maximum sustainable arrival rate at the transmitter's queue of eNodeB-assisted D2D communication, is higher than that of WiFi-assisted D2D communication. It is due to the uncoordinated interference experienced by the D2D devices in WiFi networks. This signifies the importance of better equalization and interference cancellation techniques to achieve the same EC and arrival rate. The simulation results also indicate that the traffic load of a cell deteriorates the EC of a D2D link; thus, a better cell load management by offloading devices to the WiFi network when the cell is congested is more appropriate.

In this work, CN-GW and MME played an important role as an entity for coordination among cellular and WiFi networks, however, future work will study the impact of fully decentralized coordination between these networks. Moreover, one can also consider other non-cellular networks like WiMax for the network-assisted D2D communication scenario. This way, decentralized coordination among three different wireless networks might be very interesting. One another future direction could be that independent and identically distributed (IID) assumption about service rates is not realistic in many situations (but obviously is perfectly good to use in the absence of other information). It will be quite intriguing to investigate about how much worse (or even better!) offloading to/from WiFi AP is, if the service demands are non IID.

REFERENCES

[1] T. S. Rappaport, Y. Xing, G. R. Maccartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, Dec. 2017.

[2] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[3] R. Zhang, M. Wang, L. X. Cai, Z. Zheng, X. Shen, and L.-L. Xie, "LTE-unlicensed: The future of spectrum aggregation for cellular networks," *IEEE Wireless Commun.*, vol. 22, no. 3, pp. 150–159, Jun. 2015.

[4] A. Mukherjee, J.-F. Cheng, S. Falahati, L. Falconetti, A. Furuskar, B. Godana, D. H. Kang, H. Koorapaty, D. Larsson, and Y. Yang, "System architecture and coexistence evaluation of licensed-assisted access LTE with IEEE 802.11," in *Proc. IEEE Int. Conf. Commun. Workshop (ICCW)*, Jun. 2015, pp. 2350–2355.

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

[6] S. M. A. Kazmi, N. H. Tran, W. Saad, Z. Han, T. M. Ho, T. Z. Oo, and C. S. Hong, "Mode selection and resource allocation in device-to-device communications: A matching game approach," *IEEE Trans. Mobile Comput.*, vol. 16, no. 11, pp. 3126–3141, Nov. 2017.

[7] S. M. A. Kazmi, N. H. Tran, T. M. Ho, A. Manzoor, D. Niyato, and C. S. Hong, "Coordinated device-to-device communication with non-orthogonal multiple access in future wireless cellular networks," *IEEE Access*, vol. 6, pp. 39860–39875, 2018.

[8] S. W. H. Shah, A. N. Mian, S. Mumtaz, and J. Crowcroft, "System capacity analysis for ultra-dense multi-tier future cellular networks," *IEEE Access*, vol. 7, pp. 50503–50512, 2019.

[9] Q. Chen, G. Yu, R. Yin, A. Maaref, G. Y. Li, and A. Huang, "Energy efficiency optimization in licensed-assisted access," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 723–734, Apr. 2016.

[10] Z. Zhou, X. Chen, and B. Gu, "Multi-scale dynamic allocation of licensed and unlicensed spectrum in software-defined HetNets," *IEEE Netw.*, vol. 33, no. 4, pp. 9–15, Jul. 2019.

[11] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 24, no. 5, pp. 630–643, May 2003.

[12] S. W. H. Shah, M. M. U. Rahman, A. N. Mian, A. Imran, S. Mumtaz, and O. A. Dobre, "On the impact of mode selection on effective capacity of device-to-device communication," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 945–948, Jun. 2019.

[13] S. Akin and M. C. Gursoy, "Effective capacity analysis of cognitive radio channels for quality of service provisioning," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3354–3364, Nov. 2010.

[14] A. Abdrabou and W. Zhuang, "Statistical QoS routing for IEEE 802.11 multihop ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1542–1552, Mar. 2009.

[15] Q. Wang, D. O. Wu, and P. Fan, "Delay-constrained optimal link scheduling in wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4564–4577, Nov. 2010.

[16] W. Cheng, X. Zhang, and H. Zhang, "Statistical-QoS driven energy-efficiency optimization over green 5G mobile wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3092–3107, Aug. 2016.

[17] C. Wu, X. Chen, T. Yoshinaga, Y. Ji, and Y. Zhang, "Integrating licensed and unlicensed spectrum in the internet of vehicles with mobile edge computing," *IEEE Netw.*, vol. 33, no. 4, pp. 48–53, Jul. 2019.

[18] A. K. Bairagi, N. H. Tran, W. Saad, Z. Han, and C. S. Hong, "A game-theoretic approach for fair coexistence between LTE-U and Wi-Fi systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 442–455, Jan. 2019.

[19] A. K. Bairagi, S. F. Abedin, N. H. Tran, D. Niyato, and C. S. Hong, "QoE-enabled unlicensed spectrum sharing in 5G: A game-theoretic approach," *IEEE Access*, vol. 6, pp. 50538–50554, 2018.

[20] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Control-data separation architecture for cellular radio access networks: A survey and outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 446–465, Jun. 2016.

[21] R. Wang, H. Hu, and X. Yang, "Potentials and challenges of C-RAN supporting multi-RATs toward 5G mobile networks," *IEEE Access*, vol. 2, pp. 1187–1195, 2014.

[22] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, May 2014.

[23] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklós, and Z. Turányi, "Design aspects of network assisted device-to-device communications," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 170–177, Mar. 2012.

[24] T. T. Gunes and H. Afifi, "Hybrid model for LTE network-assisted D2D communications," in *Proc. Int. Conf. Ad-Hoc Netw. Wireless*. Cham, Switzerland: Springer, Jun. 2014, pp. 100–113.

[25] H. Zhang, Y. Liao, and L. Song, "D2D-U: Device-to-device communications in unlicensed bands for 5G system," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3507–3519, Jun. 2017.

[26] G. G. Girmay, Q.-V. Pham, and W.-J. Hwang, "Joint channel and power allocation for device-to-device communication on licensed and unlicensed band," *IEEE Access*, vol. 7, pp. 22196–22205, 2019.

[27] F. Wu, H. Zhang, B. Di, J. Wu, and L. Song, "Device-to-device communications underlaying cellular networks: To use unlicensed spectrum or not?" *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6598–6611, Sep. 2019.

[28] M. M. Islam and Z. Zhang, "Device-to-device communications in unlicensed spectrum: Problem identification and performance maximization," *IEEE Access*, vol. 7, pp. 74134–74148, 2019.

[29] D. Xenakis, M. Kountouris, L. Merakos, N. Passas, and C. Verikoukis, "Performance analysis of network-assisted D2D discovery in random spatial networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5695–5707, Aug. 2016.

[30] *Physical Layer Measurements*, document TS 36.214 V11.1.0, 3GPP, Dec. 2012.

[31] H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of Wi-Fi and heterogeneous small cell networks sharing unlicensed spectrum," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 158–164, Mar. 2015.

[32] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.

[33] J. Gross, "Scheduling with outdated CSI: Effective service capacities of optimistic vs. Pessimistic policies," in *Proc. IEEE 20th Int. Workshop Quality Service*, Jun. 2012, p. 10.

[34] B. Soret, M. Aguayo-torres, and J. Entrambasaguas, "Capacity with explicit delay guarantees for generic sources over correlated Rayleigh channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 1901–1911, Jun. 2010.

[35] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw.*, 2010, pp. 1–12.

[36] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. 17th Annu. Int. Conf. Mobile Comput. Netw.*, 2011, pp. 301–312.

[37] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[38] D. Feng, L. Lu, Y. Yuan-Wu, G. Li, S. Li, and G. Feng, "Device-to-device communications in cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 49–55, Apr. 2014.

SYED WAQAS HAIDER SHAH (Student Member, IEEE) received the M.S. degree in electrical engineering from the National University of Science and Technology, Islamabad, Pakistan. He is currently a Ph.D. Research Scholar with Information Technology University, Pakistan. He is also a Visiting Researcher with the Computer Laboratory, Department of Computer Science and Technology, University of Cambridge, U.K. He has published various articles in reputed international conferences and journals. His area of research includes wireless communication, future cellular networks, 5G, device-to-device communication, analytical analysis of mobile networks, and quality-of-service provisioning. He is also a Reviewer of many international journals.

ADNAN NOOR MIAN (Member, IEEE) received the Ph.D. degree in computer engineering from the Sapienza University of Rome, Italy, in 2009. He held a postdoctoral position with the Sapienza University of Rome. From 2018 to 2019, he was a Visiting Scholar with the Computer Laboratory, Department of Computer Science and Technology, University of Cambridge, U.K. He is currently an Associate Professor with Information Technology University, Lahore, Pakistan, and also a Senior Associate in Internet of Things (IoT) with the International Centre for Theoretical Physics (ICTP), Trieste, Italy. He has published more than 40 articles in refereed conferences and journals and serving in several technical program committees of international conferences and a Reviewer of many international journals. His research interests include wireless sensor and ad hoc networks, distributed algorithms, device-to-device communication, and mobile and distributed systems.

JON CROWCROFT (Fellow, IEEE) received the degree in physics from the Trinity College, University of Cambridge, in 1979, and the M.Sc. degree in computing and the Ph.D. degree from UCL, in 1981 and 1993, respectively. He has been the Marconi Professor of communications systems with the Computer Laboratory, since 2001. He has worked in the area of Internet support for multimedia communications for more than 30 years. Three main topics of interest have been scalable multicast routing, practical approaches to traffic management, and the design of deployable end-to-end protocols. Current active research areas are opportunistic communications, social networks, and techniques and algorithms to scale infrastructure-free mobile systems. He is a Fellow of the Royal Society, of the ACM, of the British Computer Society, of the IET, and of the Royal Academy of Engineering.

• • •