



Full length article

A survey on 5G: The next generation of mobile communication

Nisha Panwar^a, Shantanu Sharma^{a,*}, Awadhesh Kumar Singh^b^a Department of Computer Science, Ben-Gurion University of the Negev, Israel^b Department of Computer Engineering, National Institute of Technology, Kurukshetra, India

ARTICLE INFO

Article history:

Received 30 June 2015

Received in revised form 11 October 2015

Accepted 30 October 2015

Available online 11 November 2015

Keywords:

Cloud radio access networks

Cognitive radio networks

D2D communication

Dense deployment

Multi-tier heterogeneous network

Tactile Internet

ABSTRACT

The rapidly increasing number of mobile devices, voluminous data, and higher data rate are pushing to rethink the current generation of the cellular mobile communication. The next or fifth generation (5G) cellular networks are expected to meet high-end requirements. The 5G networks are broadly characterized by three unique features: ubiquitous connectivity, extremely low latency, and very high-speed data transfer. The 5G networks would provide novel architectures and technologies beyond state-of-the-art architectures and technologies. In this paper, our intent is to find an answer to the question: “*what will be done by 5G and how?*” We investigate and discuss serious limitations of the fourth generation (4G) cellular networks and corresponding new features of 5G networks. We identify challenges in 5G networks, new technologies for 5G networks, and present a comparative study of the proposed architectures that can be categorized on the basis of energy-efficiency, network hierarchy, and network types. Interestingly, the implementation issues, e.g., interference, QoS, handoff, security–privacy, channel access, and load balancing, hugely effect the realization of 5G networks. Furthermore, our illustrations highlight the feasibility of these models through an evaluation of existing real-experiments and testbeds.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The evolution of the cellular network generations is influenced primarily by continuous growth in wireless user devices, data usage, and the need for a better quality of experience (QoE). More than 50 billion connected devices are expected to utilize the cellular network services by the end of the year 2020 [1], which would result in a tremendous increase in data traffic, as compared to the year 2014 [2]. However, state-of-the-art solutions are not sufficient for the challenges mentioned above. In short, the increase of 3D (‘D’evice, ‘D’ata, and ‘D’ata transfer rate) encourages the development of 5G networks.

Specifically, the fifth generation (5G) of the cellular networks will highlight and address three broad views, as: (i) user-centric (by providing 24×7 device connectivity, uninterrupted communication services, and a smooth consumer experience), (ii) service-provider-centric (by providing a connected intelligent transportation systems, road-side service units, sensors, and mission critical monitoring/tracking services), and (iii) network-operator-centric (by providing an energy-efficient, scalable, low-cost, uniformly-monitored, programmable, and secure communication infrastructure). Therefore, 5G networks are perceived to realize the three main features as below:

- *Ubiquitous connectivity*: In future, many types of devices will connect ubiquitously and provide an uninterrupted user experience. In fact, the user-centric view will be realized by ubiquitous connectivity.

* Corresponding author.

E-mail addresses: panwar@cs.bgu.ac.il (N. Panwar), sharmas@cs.bgu.ac.il (S. Sharma), aksingh@nitkkr.ac.in (A.K. Singh).<http://dx.doi.org/10.1016/j.phycom.2015.10.006>

1874-4907/© 2015 Elsevier B.V. All rights reserved.

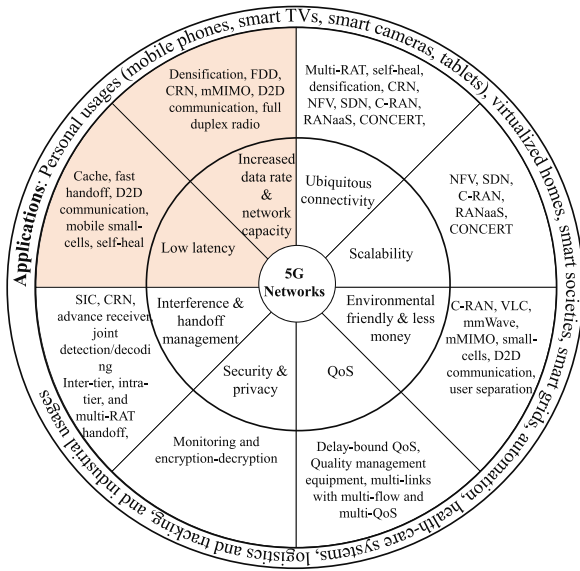


Fig. 1. Requirements and proposed solutions for the development of 5G networks. The inner, middle, and outermost layers present requirements, solutions, and applications of 5G networks, respectively. Two colored wedges highlight primary features of 5G networks.

- **Zero latency:** The 5G networks will support life-critical systems, real-time applications, and services with zero delay tolerance. Hence, it is envisioned that 5G networks will realize zero latency, i.e., extremely low latency of the order of 1 millisecond [3,4]. In fact, the service-provider-centric view will be realized by the zero latency.
- **High-speed Gigabit connection:** The zero latency property could be achieved using a high-speed connection for fast data transmission and reception, which will be of the order of Gigabits per second to users and machines [3].

A few more *key features of 5G networks* are enlisted and compared to the fourth generation (4G) of the cellular networks, as below [5–7]: (i) 10–100x number of connected devices, (ii) 1000x higher mobile data volume per area, (iii) 10–100x higher data rate, (iv) 1 ms latency, (v) 99.99% availability, (vi) 100% coverage, (vii) $\frac{x}{10}$ energy consumption as compared to the year 2010, (viii) real-time information processing and transmission, (ix) $\frac{x}{5}$ network management operation expenses, and (x) seamless integration of the current wireless technologies.

The revolutionary scope and the consequent advantages of the envisioned 5G networks, therefore, demand new architectures, methodologies, and technologies (see Fig. 1), e.g., energy-efficient heterogeneous frameworks, cloud-based communication (software-defined networks (SDN) and network function virtualization (NFV)), full duplex radio, self-interference cancellation (SIC), device-to-device (D2D) communications, machine-to-machine (M2M) communications, access protocols, cheap devices, cognitive networks (for accessing licensed, unlicensed, and shared frequency bands), dense-deployment, security–privacy protocols for communication and data transfer, backhaul connections, massive multiple-input and

multiple-output (mMIMO), multi-radio access technology (RAT) architectures, and technologies for working on millimeter wave (mmWave) 30–300 GHz. Interestingly, 5G networks will not be a mere enhancement of 4G networks in terms of additional capacity; they will encompass a system architecture visualization, conceptualization, and re-designing at every communication layer [8].

Several industries, e.g., Alcatel-Lucent [9], DOCOMO [10], GSMA Intelligence [6], Huawei [11], Nokia Siemens Networks [3], Qualcomm [12], Samsung [13], Vodafone,¹ the European Commission supported 5G Infrastructure Public Private Partnership (5GPPP) [5], and Mobile and Wireless Communications Enablers for the Twenty–Twenty Information Society (METIS) [7], are brainstorming with the development of 5G networks. Currently, the industry standards are yet to be evolved about the expected designs and architectures for 5G networks.

Scope of the paper. In this paper, we will review the vision of the 5G networks, advantages, applications, proposed architectures, implementation issues, real demonstrations, and testbeds. The outline of the paper is provided in Fig. 2. In Section 2, we will elaborate the vision of 5G networks. Section 3 presents challenges in the development of 5G networks. Section 4 addresses the currently proposed architectures for 5G networks, e.g., multi-tier, cognitive radio based, cloud-based, device proximity based, and energy-efficient architectures. Section 5 presents issues regarding interference, handoff, quality of services, load balancing, channel access, and security–privacy of the network. Sections 6–8 present several methodologies and technologies involved in 5G networks, applications of 5G networks, and real demonstrations and testbeds of 5G networks, respectively.

We would like to emphasize that there do exist some review works on 5G networks by Andrews et al. [14], Chávez-Santiago et al. [15], and Gavrilovska et al. [16], to the best of our knowledge. However, our perspective about 5G networks is different, as we deal with a variety of architectures and discuss several implementation affairs, technologies in 5G networks along with applications and real-testbed demonstrations. In addition, we intentionally avoid an mmWave oriented discussion in this paper, unlike the current work [14–16].

We encourage our readers to see an overview about the generations of the cellular networks (see Table 1) and the crucial limitations of current cellular networks in the next section.

1.1. Limitations of the conventional cellular systems

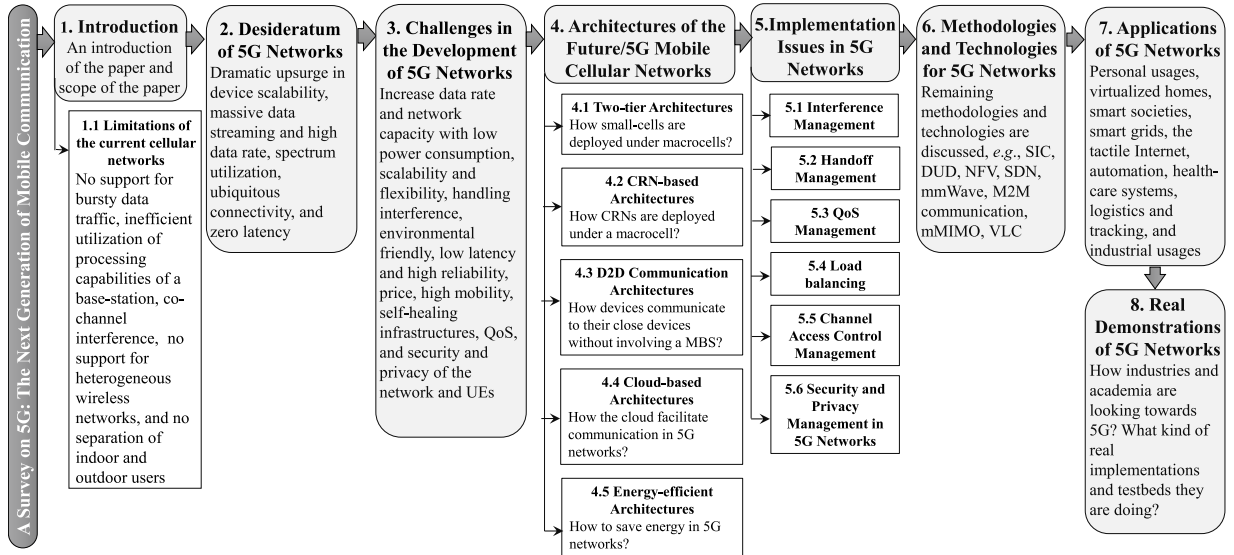
The 4G networks are not substantial enough to support massively connected devices with low latency and significant spectral efficiency, which will be crucial in the future communication and computing. In this section, we discuss a few crucial aspects in which conventional cellular networks lag far behind, thereby motivating the evolution of 5G networks.

¹ <http://www.surrey.ac.uk/5gic/research>.

Table 1

The generations of the cellular networks.

Generations	Year	Features	Limitations
1G	1980s	Analog signals for voice only communications	Very less security
2G	1990s	Digital signals, voice communications, and text messaging	Very less support for the Internet
3G	1998–1999	Voice communications, wireless mobile and fixed Internet access, video calls, and mobile television (TV)	Less support for high-speed Internet
4G	2008–2009	Higher data rate (hundreds of megabits per second)	No support for 50 billion ubiquitous connected devices
5G	2020	Mentioned in Section 1	

**Fig. 2.** Schematic map of the paper.

No support for bursty data traffic. There are several mobile applications that send heartbeat messages to their servers and occasionally request for very high data transfer rate for a very short duration. Such types of data transmission consume more battery life of (mobile) user equipments (UEs) with increasing bursty data in the network, and hence, may crash the core network [17]. However, only one type of signaling/control mechanism is designed for all types of the traffic in the current networks, creating high overhead for bursty traffic [18,19].

Inefficient utilization of processing capabilities of a base-station. In the current cellular networks, the processing power of a base-station (BS) can only be used by its associated UEs, and they are designed to support peak time traffic. However, the processing power of a BS can be shared across a large geographical area when it is lightly loaded. For example: (i) during the day, BSs in business areas are over-subscribed, while BSs in residential areas are almost idle, and vice versa [20], and (ii) BSs in residential areas are overloaded in weekends or holidays while BSs in business areas are almost idle [21]. However, the almost idle BSs consume an identical amount of power as over-subscribed BSs; hence, the overall cost of the network increases.

Co-channel interference. A typical cellular network uses two separate channels, one as a transmission path from a UE to a BS, called uplink (UL), and the reverse path, called

downlink (DL). The allocation of two different channels for a UE is not an efficient utilization of the frequency band. However, if both the channels operate at an identical frequency, i.e., a full duplex wireless radio [22], then a high level of co-channel interference (the interference between the signals using an identical frequency) in UL and DL channels is a major issue in 4G networks [23]. It also prevents the *network densification*, i.e., the deployment of many BSs in a geographical area.

No support for heterogeneous wireless networks. Heterogeneous wireless networks (HetNets) are composed of wireless networks with diverse access technologies, e.g., the third generation (3G), 4G, wireless local area networks (WLAN), WiFi, and Bluetooth. HetNets are already standardized in 4G; however, the basic architecture was not intended to support them. Furthermore, the current cellular networks allow a UE to have a DL channel and a UL channel must be associated with a single BS that prevents the maximum utilization of HetNets. In HetNets, a UE may select its UL and DL channels from two different BSs belonging to two different wireless networks for performance improvement [24,25].

No separation of indoor and outdoor users. The current cellular networks have a single BS installed preferably near the center of the cell and interacts with all the UEs irrespective of the indoor or outdoor location of the UEs; while UEs stay indoors and outdoors for about 80% and 20%

of the time, respectively. Furthermore, the communication between an indoor UE and an outside BS is not efficient in terms of data transfer rate, spectral efficiency, and energy-efficiency, due to the attenuation of signals passing through walls [26].

Latency. When a UE receives an access to the best candidate BS, it takes several hundreds of milliseconds in the current cellular networks [27], and hence, they are unable to support the zero latency property.

2. Desideratum of 5G networks

A growing number of UEs and the corresponding surge in the bandwidth requirement for the huge amount of data transmission certainly necessitate the novel enhancement to the current technology. In this section, we highlight requirements of the future 5G networks.

Dramatic upsurge in device scalability. A rapid growth of smart phones, gaming consoles, high-resolution TVs, cameras, home appliances, laptops, connected transportation systems, video surveillance systems, robots, sensors, and wearable devices (watches and glasses) is expected to continue exponentially in the near future. Therefore, 5G networks are perceived to support massively connected devices [26,1,28].

Massive data streaming and high data rate. A vast growth in a number of wireless devices will of course result in a higher amount of data trading (e.g., videos, audio, Web browsing, social-media data, gaming, real-time signals, photos, bursty data, and multimedia) that will be 100-times more as compared to the year 2014 and would overburden the current network. Thus, it is mandatory to have matching data transfer capabilities in terms of new architectures, methods, technologies, and data distribution of indoor and outdoor users [29,28,30].

Spectrum utilization. The two different channels (one for a UL and another for a DL) seem redundant from the point of view of the spectrum utilization [31]. In addition, the currently allocated spectrums have their significant portions under-utilized [32]. Hence, it is necessary to develop an access control method that can enhance the spectrum utilization. Furthermore, the spectrum utilization and efficiency have already been stretched to the maximum. It definitely requires spectrum broadening (above 3 GHz) along with novel spectrum utilization techniques [33].

Ubiquitous connectivity. Ubiquitous connectivity requires UEs to support a variety of radios, RATs, and bands due to the global non-identical operating bands. In addition, the major market split between time division duplex (e.g., India and China) versus frequency division duplex (e.g., US and Europe) so that UEs are required to support different duplex options. Hence, 5G networks are envisioned for *seamless connectivity* of UEs over HetNets [34].

Zero latency. The future mobile cellular networks are expected to assist numerous real-time applications, the tactile Internet [4,35], and services with varying levels of quality of service (QoS) (in terms of bandwidth, latency, jitter, packet loss, and packet delay) and QoE (in terms of

users' and network-providers' service satisfaction versus feedback). Hence, 5G networks are envisioned to realize real-time and delay-bound services with the optimal QoS and QoE experiences [28,23].

3. Challenges in the development of 5G networks

The vision of 5G networks is not trivial to achieve. There are several challenges (some of the following challenges are shown in Fig. 1 with their proposed solutions) to be handled in that context, as mentioned below:

Data rate and network capacity expansion with energy optimization. The deployment of more BSs in a geographical area, use of the higher frequency bands, and link improvement might support the network capacity expansion, billions of UEs, high data rate, high volume of data, and efficient backhaul data transfer to the core network. However, the implementation of these solutions is a cumbersome task in terms of economy and energy intake. Hence, the network capacity is required to be significantly increased, keeping the energy consumption and cost under strict control.

Proposed solutions: Network densification or small-cell deployment [28,36,26] (Section 4.1), cognitive radio networks (CRNs) [37] (Section 4.2), mMIMO [38–40] (Section 6), network offload using D2D communication [41–43] (Section 4.3), efficient backhaul networks [8,44] (Section 4.1.1), energy-efficient architectures [45,46] (Section 4.5), full duplex radios [22] (Section 6), NFV, and SDN based architectures [47–50] (Section 6).

Scalability and flexibility. These are the most prominent features of the future mobile communication. The future cellular infrastructures and methodologies must be designed to work in HetNets. Moreover, a vast number of potential users might request simultaneously for a set of services. Therefore, 5G networks must be powerful enough to support a scalable user demand across the coverage area [48,51].

Proposed solutions: NFV- and SDN-based architectures [47–50] (Section 6).

Single channel for both UL and DL. A *full duplex wireless radio* [22] uses only a single channel for transmitting and receiving signals at identical time and frequency. Thus, a full duplex system achieves an identical performance as having different UL and DL channels, and hence, increases link capacity, saves the spectrum, and cost. However, the implementation of full duplex systems is not trivial, because now a radio has to use sophisticated protocols for the physical and the data link layers [52], and mechanisms to remove the effects of interference [31]. The advantages of a full duplex radio in 5G networks are given in [53,31, 18].

Handling interference. Handling interference among communicating devices is a well-known challenge in the wireless communication. Due to a growing number of UEs, technologies (e.g., HetNets, CRNs, full duplex, and D2D communication) and applications, the interference will also increase in 5G networks, and the state-of-the-art technique may not perform well in the future cellular networks [29]. In 5G networks, a UE may

receive interference from multiple macrocell base-stations (MBSs), various UEs, and small-cell base-stations (SBSs). Hence, it is required to develop an efficient (in terms of avoiding network overload) and reliable (in terms of perfect interference detection and decoding) interference management technique for channel allocation, power control, cell association, and load balancing.

Proposed solutions: Self-interference cancellation [18, 31], an advance receiver with interference joint detection/decoding, and network-side interference management [23]. We will discuss these solutions in Section 5.1.

Environmentally friendly. The current radio access network (RAN) consumes 70%–80% of the total power [18,54]. The wireless technologies consume lots of energy that lead to huge CO₂ emission and inflate the cost. It is a serious threat to the environment [26]. Thus, it is required to develop energy-efficient communication systems, hardware, and technologies, thereby the ratio between the network throughput and energy consumption is equitable.

Proposed solutions: Cloud-RAN (C-RAN) [54,45], visual light communication (VLC) [54], mmWave [54], separation of indoor and outdoor users [54], joint investigation of spectral efficiency and energy-efficacy [18,45], multi-tier architectures [45], D2D communication [41–43], mMIMO architectures [45], and full duplex radios [18]. Except the above mentioned solutions, we will discuss some special techniques/architectures in the context of energy-efficiency in 5G networks in Section 4.5.

Low latency and high reliability. Low latency and high reliability are critical in several real-time applications, e.g., message transmission by robots monitoring patients, life safety systems, cloud-based gaming, nuclear reactors, sensors, drones, and connected transportation systems. However, it is very challenging to have extremely low latency and reliable delivery of data over a large scale network without increasing the network infrastructure cost, as it requires the development of techniques providing fast connections, quick handovers, and high data transfer rate.

Proposed solutions: Caching methods [24,55], VLC, mmWave, mMIMO (Section 6), fast handover techniques [56–58] (Section 5.2), and D2D communication (Section 4.3).

Network performance optimization. The performance parameters, e.g., peak data rate, geographical area coverage, spectral efficiency, QoS, QoE, ease of connectivity, energy-efficiency, latency, reliability, fairness of users, and implementation complexity, are crucial for a cellular network [26]. Hence, a general framework for 5G networks should substantially optimize these parameters. However, there are some tradeoffs among all parameters, which further emphasize the need of a joint optimization algorithm.

Economical impacts. A revolutionary change in the future mobile communication techniques would have drastic economical impacts in terms of deployment and motivation for user participation. It is critical to provide an entirely new infrastructure due to economical stretch. Therefore, the cost of deployment, maintenance, management, and operation of an infrastructure must be affordable from the perspective of governments, regulating authorities, and network operators. Also, the cost of using

D2D communication should be feasible, so that devices involved in D2D communication should not charge more than using the services of a BS [28,35]. Further, the projected revenue growth is much lower than the traffic growth [47]; hence, it is required to develop 5G networks in a manner that both network operators and users get honey in their hands.

High mobility and handoff. The 5G wireless UEs are meant for retaining an active service connection while frequently moving from one cell to another or from one RAT (e.g., 3G, 4G, 5G, WiFi, Bluetooth, and WLAN) to another. The mobility adaptation for the wireless services should not back-off even at a very high speed as a UE inside a moving vehicle. Moreover, during a particular interval, many UEs move from one place to another; for example, moving to offices from residential areas in the morning. As a result, 5G networks are envisioned to use the spectrum in the best manner and to cope up with pace of the device movement.

Proposed solutions: Inter-tier, intra-tier, and multi-RATs handoff mechanisms, and a mechanism for secure handoff [56–59], which we will discuss in Section 5.2.

Self-healing infrastructures. A self-healing infrastructure finds a *failed* macrocell or small-cell (i.e., a cell that is unable to work because of hardware failures, software failures, or misconfigurations) with the help of neighboring cells and provides a way for communication to the affected users by adjusting the transmission power and operating channels in the neighboring cells [60,61]. The design of a self-healing network insists on the frequent communication among cells; hence, it brings in the following challenges, as: (i) develop an efficient algorithm that can detect and reconfigure a failed cell with insignificant communication and computational overheads in the minimal detection time, and (ii) reconfiguration of a failed cell should not lead to degradation of nearby cells' services.

Proposed solutions: A small-cell network with self-healing property is suggested in [61], which we will discuss in Section 4.1.2.

QoS. QoS guarantee in 5G networks has inherent difficulties, e.g., node mobility, multi-hop communication, resource allocation, and lack of central coordination. In addition, in 5G networks, a huge amount of bursty and multimedia data, multi-RATs, and low latency bound for different applications and services are major hurdles in achieving the desired QoS. Hence, it is challenging to design fast and efficient algorithms to maintain real-time QoS without overloading a BS [17,62].

Proposed solutions: Delay-bound QoS [45,62], intelligent equipment [17], and multi-link with multi-flow and multi-QoS [63] have been suggested, which we will discuss in Section 5.3.

Security and privacy of the network and UEs. The promising features of 5G networks bring in hard challenges in the design of security and privacy oriented 5G networks. For example, a huge number of new types of social (all-time connected) devices may originate several types of attacks like impersonation, denial-of-services (DoS), replay, eavesdropping, man-in-the-middle, and repudiation

Table 2

Summary of methodologies and technologies for 5G networks.

Methodologies/Technologies	Section	Increasing data rate	Increasing network capacity	Massive device support	Energy-efficient	Low latency	Economic	Security and privacy	Interference	Mobility support
Small-cells	4.1	✓	✓		✓		✓ ^P ^a			
Mobile small-cells	4.1	✓	✓		✓		✓ ^P			✓
CRN	4.2		✓				✓			
D2D	4.3	✓ ^P		✓ ^P	✓		✓ ^P		✓ ^P	
C-RANs	4.4	✓	✓	✓	✓		✓		✓	✓
Full duplex radio	3, 5.1		✓						✓	
Advance receiver	5.1		✓						✓	
SIC	5.1, 6		✓			✓ ^P			✓	✓ ^P
DUD	6		✓ ^P		✓					
mmWave	6	✓	✓	✓	✓	✓				
mMIMO	6	✓	✓	✓	✓	✓			✓	
VLC	6	✓			✓	✓	✓	✓		
CCN-based caching	6						✓			

^a ✓^P: Partial support.

attacks. Also, the transfer of a huge volume of data in secure and high speed manners is critical while preventing malicious files to penetrate. In addition, the network densification needs to be secure and requires fast-secure hand-off of UEs. We further highlight challenges in security and privacy of the network and UEs in Section 5.6.

Proposed solutions: Physical layer security [64], monitoring [65–67], secret adaptive frequency hopping [67], encrypted- [68], and policy-based communications [69], which we will discuss in Section 5.6.

All the above mentioned methodologies and technologies are comparatively studied in Table 2.

4. Architectures for the future/5G mobile cellular networks

In this section, we elaborate on the existing architectures for 5G networks, namely multi-tier, CRN-based, D2D communication based, and the cloud-based architectures. These proposed 5G architectures will be explained in the light of relevant advantages, disadvantages, and the challenges that are yet to be resolved.

4.1. Two-tier architectures

Several two-tier architectures have been proposed for 5G networks, where a MBS stays in the top-tier and SBSs work under the supervision of the MBS in the lower tier. A macrocell covers all the small-cells of different types, e.g., femtocell, picocell, and microcell (see Table 3), and both the tiers share an identical frequency band. The small-cell enhances the coverage and services of a macrocell, and the advantages of small-cells are mentioned at the end of this section. In addition, D2D communication and CRN-based communication enhance a 2-tier architecture to a multi-tier architecture; see Fig. 3. Note that in this section, we confine ourselves on the deployment of small-cells under

Table 3

Classification of the cells.

Cells	Range	Users
Femtocell	10–20 m	A few users
Picocell	200 m	20–40
Microcell	2 km	> 100
Macrocell	30–35 km	Many

the cover of a macrocell; the discussion of CRN-based and D2D communications will be carried out in Sections 4.2 and 4.3, respectively.

Wang et al. [26] suggested a way for separating indoor and outdoor users and using a *mobile small-cell* on a train or a bus. For separating indoor and outdoor users, a MBS holds large antenna arrays with some antenna elements distributed around the macrocell and connected to the MBS using optical fibers. A SBS and large antenna arrays are deployed in each building for communicating with the MBS. All UEs inside a building can have a connection to another UE either through the SBS or by using WiFi, mmWave, or VLC. Thus, the separation of users results in less load on a MBS.

Wang et al. [26] also suggested to use a mobile small-cell that is located inside a vehicle to allow communication among internal UEs, while large antenna arrays are located outside the vehicle to communicate with a MBS. Thus, all the UEs inside a vehicle (or a building) appear to be a single unit with respect to the corresponding MBS, and clearly, the SBS appears as a MBS to all these UEs.

In [36], a two-tier architecture is deployed as a process of *network densification* that is a combination of *spatial densification* (increasing the number of antennas per UE and MBS, and increasing the density of BSs) and *spectral aggregation* (using higher frequency bands >3 GHz). A tradeoff between the transmission power of a macrocell and the coverage area of small-cells is presented in [36]; i.e., on the one hand, if the transmission power of a

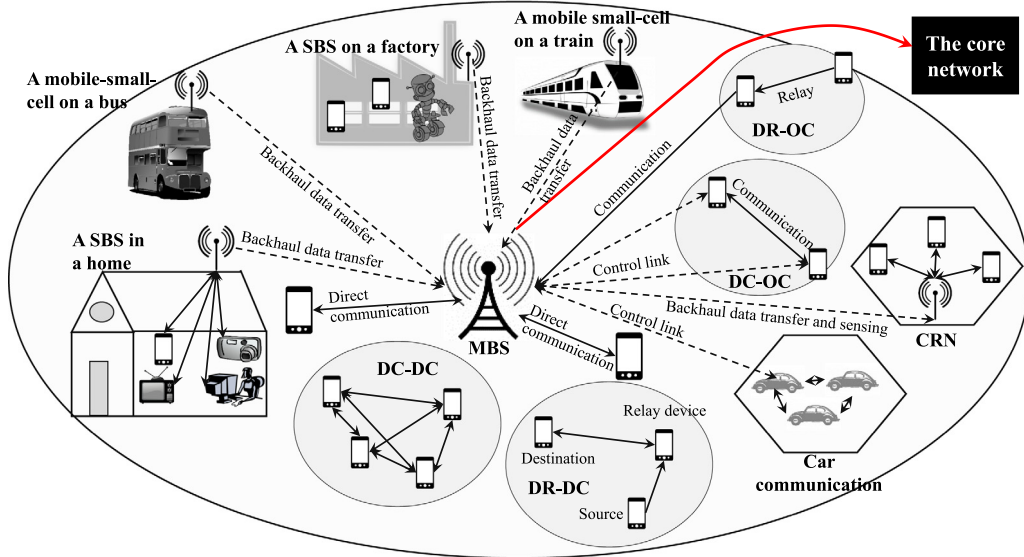


Fig. 3. A multi-tier architecture for 5G networks with small-cells, mobile small-cells, and D2D- and CRN-based communications.

macrocell is high, then many adjacent UEs to a small-cell may find themselves in the service area of the macrocell, and hence, it will decrease the coverage area of that small-cell. On the other hand, if the transmission power of a macrocell is low, then the coverage area of the small-cell will increase. Therefore, *cell range expansion* (i.e., a biased handoff in the favor of small-cells) is carried out to serve more UEs by small-cells to which they are closer. Moreover, SBSs, deployed in offices or homes, can be used to serve outdoor users, e.g., pedestrians and low-mobility vehicles, in their neighborhoods, and the approach is called *indoor-to-outdoor user service* [36].

Hossain et al. [29] presented a multi-tier architecture consisting of several types of small-cells, relays, and D2D communication for serving users with different QoS requirements in spectrum-efficient and energy-efficient manners. Interestingly, all of these architectures consider that UEs spontaneously discover a SBS. Zhang et al. [27] proposed a centralized system in which a MBS assists UEs to have connections to particular SBSs, thereby interference between UEs and SBSs is reduced. However, this approach overburdens the MBS.

Advantage of the deployment of small-cells

- **High data rate and efficient spectrum use:** The small physical separation between a SBS and UEs (served by the same SBS) leads to a higher data rate and a better indoor coverage. Also, the spectrum efficiency increases due to fewer UEs in direct communication with a MBS [61].
- **Energy saving:** The use of small-cells reduces the energy consumption of the network (by not involving MBSs) and of UEs (by allowing UEs to communicate at a shorter range with low signaling overhead) [8].
- **Money saving:** It is more economical to install a SBS without any cumbersome planning as compared to a MBS, and also the operational-management cost is much lower than the cost associated with a MBS [60, 36].

- The plug-and-play utility of small-cells boosts the on-demand network capacity [17].
- **Less congestion to a MBS:** SBSs offload UEs from a MBS so that the MBS is lightly loaded and less congested, and hence, improve the system capacity [60].
- **Easy handoff:** Mobile small-cells also follow the advantages of small-cells. Moreover, they provide an attractive solution to highly mobile UEs by reducing handoff time overheads, since a mobile small-cell is capable to do the handoff on behalf of all related UEs [36].

Disadvantage of small-cells. Despite numerous prominent benefits as mentioned above, there are a few realistic issues such as implementation cost and operational reliability. The small-cells indeed impose an initial cost to the infrastructure, but less than the cost associated with a MBS. Moreover, a frequent authentication is mandatory due to frequent handoff operations. Further, an active or passive (on/off) state update of any small-cell would definitely result in frequent topological updates.

Open issues in the deployment of 2-tier architectures using small-cells

- **Interference management:** The deployment of small-cells results in several types of interferences, as: *inter-tier interference* (i.e., interference from a MBS to a SBS, interference from a MBS to a SBS's UEs, and interference from a SBS to a MBS's UEs), and *intra-tier interference* (i.e., interference from a SBS to other SBSs' UEs). Hence, it is also required to develop models and algorithms to handle these interferences [60,70].
- **Backhaul data transfer:** Though we have models to transfer data from a SBS to the core network, which we will discuss next in Section 4.1.1, an extremely dense-deployment of small-cells requires a huge amount of data transfer, and certainly, requires cost efficient architectures.

4.1.1. Backhaul data transfer from small-cells

Data transfer from a SBS to the core network is a challenging task, and in general, there may be three approaches to transfer (backhaul) data to the core network, as follows:

- *Wired optical fiber*: by establishing a wired optical fiber link from each SBS to a MBS; however, it is time-consuming and expensive.
- *Wireless point-to-multipoint (PTMP)*: by deploying a PTMP-BS at a MBS that communicates with SBSs and transfers data to the core network.
- *Wireless point-to-point (PTP)*: by using directional antennas in line-of-sight (LOS) environments; hence, it provides high capacity PTP links (same as with wired optical fibers), at a significantly lower cost.

Ge et al. [8] presented two architectures based on the wireless PTMP approach. In the first (centralized) architecture, all SBSs send data using mmWave to a MBS that eventually aggregates the received data and forwards the same to the core network using fiber. In the second (distributed) architecture, all small-cells cooperatively forward data using mmWave to a specified SBS that transfers data to the core network using fiber without the explicit involvement of a MBS.

Ni et al. [44] proposed an *adaptive backhaul architecture* based on the wireless PTP approach and frequency division duplex for UL and DL channels. A tree structure is used, where the root node is connected to the core network using fiber, the leaf nodes represent UEs, and other nodes represent SBSs. The data is transferred from the leaf nodes to the root node that transfers the same to the core network. The bandwidth is selected dynamically for backhaul links, as per the current bandwidth requirements, interference conditions, and the network topology. A similar approach is also presented in [36].

4.1.2. Two-tier architectures with self-healing property

An automatic detection and recovery of a failed cell is an important issue in densely deployed multi-tier architectures. Wang and Zhang [61] provided three approaches for designing a self-healing architecture such as below:

1. *Centralized approach*: A dedicated server is responsible for detecting a failed cell by measuring and analyzing abnormal behavior of users, e.g., received signal strengths (RSSs) at users and handoff by several users at any time from a particular cell. The server collects global information and reconfigures the failed cell. However, the approach suffers with a high communication overhead and a high computational cost.
2. *Distributed approach*: Each SBS detects failed small-cells in neighborhoods by measuring and analyzing users' handoff behavior and the neighboring small-cells' signals. Consequently, on detecting a failed cell, a SBS might increase the transmission power in order to incorporate users of the failed cell. However, the approach might not work efficiently in case that users are scattered sparsely.

3. *Local cooperative or hybrid approach*: This approach combines the benefits of both the previous approaches, and therefore, minimizes the drawback. Essentially, two steps are utilized, namely distributed trigger and cooperative detection. In the distributed trigger, each SBS collects information about users' behavior. Subsequently, a trigger message is sent to a dedicated server in case the received information thrives below a certain threshold. Hence, it does not require communication among small-cells. In the cooperative detection, the dedicated server takes the final decision based on the information received from several small-cells, resulting in a high accuracy and lower latency.

4.2. Cognitive radio network based architectures

A cognitive radio network (CRN) [37] is a collection of cognitive radio nodes (or processors), called secondary users (SUs) that exploit the existing spectrum opportunistically. The SUs have the *LEIRA* (learning, efficiency, intelligence, reliability, and adaptively) property for scanning and operating on multiple heterogeneous channels (or frequency bands) in the absence of the licensed user(s), termed as primary user(s) (PUs), of the respective bands [71]. Each PU has a fixed bandwidth, high transmit power, and high reliability; however, the SUs work on a broad range of bandwidth with low transmit power and low reliability.

A CRN in 5G networks is used for designing multi-tier architectures, removing interference among cells, and minimizing energy consumption in the network [72,60,73,30,74,46].

4.2.1. CRN-based architectures for 5G networks

A CRN creates a 2-tier architecture, similar to architectures discussed in Section 4.1; however, it is assumed that either a MBS or a SBS has cognitive properties for working on different channels.

Hong et al. [30] presented two types of CRN-based architectures for 5G networks, as: (i) non-cooperative and (ii) cooperative CRNs. The *non-cooperative CRN* establishes a multi-RATs system, having two separate radio interfaces that operate at the licensed and temporary unoccupied channels by PUs, called cognitive channels. The SUs work only on cognitive channels and form a CRN, which overlays on the existing licensed cellular network. The two networks can be integrated in the upper layers while must be separated in the physical layer. This architecture can be used in different manners, as: (i) the cognitive and licensed channels are used by users near a MBS and users far away from the MBS, respectively, (ii) the cognitive and licensed channels are used for relaxed QoS and strict QoS, respectively. The *cooperative CRN* uses only a licensed channel, where SUs access the channel in an opportunistic fashion when the PU of the channel is absent. This architecture can be used in different manners, as: (i) a SBS communicates with a MBS using the licensed channel and provides service to its UEs via an opportunistic access to the licensed frequency band, (ii) a licensed channel is used to serve UEs by a SBS and the opportunistic access to

the licensed channel is used to transfer backhaul data to the MBS.

In short, the cooperative CRN [30] provides a real intuition of incorporating CRNs in 5G networks, where a SBS works as a SU, which scans activities of a macrocell and works on temporarily unoccupied frequency bands (known as *spectrum holes* [75]) by a PU to provide services to their UEs with minimally disrupting macrocell activities.

A dynamic pricing model based on a game theoretic framework for cognitive small-cells is suggested in [76]. Since in reality SBSs' operators and MBSs' operators may not be identical and small-cells' UEs may achieve a higher data rate as compared to macrocells' UEs, the pricing model for both UEs must be different.

4.2.2. Interference management using CRNs

Huang et al. [72] provided an approach for avoiding inter-tier interference by integrating a *cognitive technique* at a SBS. The cognitive technique consists of three components, as: (i) a cognitive module, which senses the environment and collects information about spectrum holes, collision probability, QoS requirements, macrocell activities, and channel gains, (ii) a cognitive engine, which analyzes and stores the collected information for estimating available resources, and (iii) a self-configuration module, which uses the stored information for dynamically optimizing several parameters for efficient handoff, interference, and power management. Further, the channel allocation to a small-cell is done in a manner to avoid inter-tier and intra-tier interferences, based on Gale–Shapley [77] spectrum sharing scheme, which avoids collisions by not assigning an identical channel to neighboring small-cells.

Wang et al. [73] suggested an approach for mitigating inter-tier interference based on spectrum sensing, spectrum sharing, and cognitive relay, where links between a MBS and its UEs are considered as PUs and links between a SBS and its UEs are considered as SUs. Cognitive techniques are used for detecting interference from a MBS to a SBS and vice versa, and a path loss estimation algorithm is provided for detecting interference from a small-cell's UEs to a macrocell's UEs. After detecting inter-tier interference, a small-cell shares spectrum with a macrocell using either *overlay spectrum sharing scheme* (i.e., SUs utilize unoccupied channels, and it is applicable when a MBS and a SBS's UEs are very close or no interference is required by a macrocell's UEs) or *underlay spectrum sharing scheme* (i.e., SUs and PUs transmit on an identical channel while restricting transmit power of SUs, and hence, resulting in a higher spectrum utilization).

Note that a CRN can be used to support D2D communication and mitigate interferences caused by D2D communication, which we will see in Section 4.3.

Advantages of CRNs in 5G networks

- **Minimizing interference:** By implementing a CRN at small-cells, cognitive small-cells can avoid interference very efficiently by not selecting identical channels as the channels of neighboring small-cells.
- **Increase network capacity:** The spectrum holes can be exploited for supporting a higher data transfer rate and enhancing bandwidth utilization [46].

Open issues. Usually, cellular networks are not energy-efficient as they consume energy in circuits, cooling systems, and also radiate in air. Hence, an energy-efficient deployment of a CRN in a cellular network is of utmost importance [30]. Further, there is a tradeoff between the spatial frequency reuse and the outage probability, which requires the selection of an efficient spectrum sensing algorithm [60].

4.3. Device-to-device communication architectures

Device-to-Device (D2D) communication allows close proximity UEs to communicate with each other on a licensed cellular bandwidth without involving a MBS or with a very controlled involvement of a MBS. The standards and frameworks for D2D communication are in an early stage of research. In this section, we will review D2D communication networks in short. For a detailed review of D2D communication, interested readers may refer to [78].

Challenges in D2D communication

- **Interference management:** UEs involved in D2D communication, say D-UEs, face (or create) interference from (or to) other UEs, or from (or to) a BS, based on the selection of a DL or UL channel, respectively. The following types of interferences are investigated in [43]:
 - When using a DL channel: (i) interference from BSs in the same cell, (ii) interference from other co-channel D-UEs in the same cell, and (iii) interference from BSs and co-channel D-UEs from other cells.
 - When using a UL channel: (i) interference from all co-channel C-UEs² in the same cell and other cells, and (ii) interference from all co-channel D-UEs in the same cell and other cells.

Proposed solutions: A simple solution may exist by implementing CRNs in D2D communication, as: D-UEs are considered as SUs and C-UEs are considered as PUs that should not be interfered. Consequently, any mechanism of CRNs can be implemented in D2D communication for interference removal.

- **Resource allocation:** When UEs involved in D2D communication, it is required to allocate a sufficient amount of resources, particularly bandwidth and channels. However, the allocation of optimum resources to D-UEs must be carried out in a fashion that C-UEs do not have interference from D-UEs, and D-UEs can also communicate and exchange data efficiently [43,79].
Proposed solutions: SARA [41], frame-by-frame and slot-by-slot channel allocation methods [80], and a social-aware channel allocation [79].
- **Delay-sensitive processing:** Audio, video streaming, and online gaming, which are natural in close proximity UEs, require real-time and delay-sensitive processing. Hence, it is required to consider delay-sensitive and real-time processing in D2D communication [81].
Proposed solutions: Solutions based on channel state information (CSI) and QoS are provided in [81].

² A UE that is not involved in D2D communication and communicates to a MBS, we call it a *cellular user equipment* (C-UE) in this section.

- **Pricing:** Sometimes a D-UE uses resources (e.g., battery and data storage) of other UEs for relaying information, where the other UE may charge for providing its resources. Hence, the design of a pricing model is needed, thereby a D-UE is not charged more money than that involved to communicate through a MBS.

Proposed solutions: Some solutions based on game theory, auction theory, and bargaining are suggested in [42].

D2D communication types. D2D communication can be done in the following four ways [42], as follows:

1. **Device relaying with operator controlled link establishment (DR-OC):** A UE at the edge of a cell or in a poor coverage area can communicate with a MBS by relaying its information via other UEs, which are within the stronger coverage area and not at the edge; see Fig. 3.
2. **Direct D2D communication with operator controlled link establishment (DC-OC):** Source and destination UEs communicate directly with each other without involving a MBS, but they are assisted by the MBS for link establishment; see Fig. 3.
3. **Device relaying with device controlled link establishment (DR-DC):** Source and destination UEs communicate through a relay without involving a MBS, and they are also responsible for link establishment; see Fig. 3.
4. **Direct D2D communication with device controlled link establishment (DC-DC):** Source and destination UEs communicate directly with each other without involving a MBS, and they are also responsible for link establishment; see Fig. 3.

Note that DR-OC and DC-OC involve a MBS for resource allocation and call setup, and hence, prevent interference among devices to some extent.

Two types of coding schemes (or communication types) are described in [43]: (i) two-way relay channel (TRC), where a source and a destination communicate through a relay, and (ii) multiple-access relay channel (MRC), where multiple sources communicate to a destination through a relay with direct links. Note that the workings of DR-OC and MRC, and DR-DC and TRC are identical. Two types of node discovery and D2D communication methods are also studied in [79], namely network-controlled approach and ad hoc network approach that work in a similar manner as DC-OC and DC-DC, respectively.

Resource allocation methods. Now, we will review an architecture and some methods for resource allocation in D2D communication.

Social-aware D2D architecture: As D2D communication is very efficient for close proximity UEs, keeping this fact in view, Li et al. [79] suggested a social-aware D2D communication architecture based on the social networking. The architecture, see Fig. 4, has four major components as:

- **Ties:** They are similar to friend relations in a social media, and hence, may be used as a trust measurement between two UEs. Allocating more spectrum and energy resources to UEs with strong ties can increase the peer discovery ratio, avoid congestion, and improve spectral efficiency.

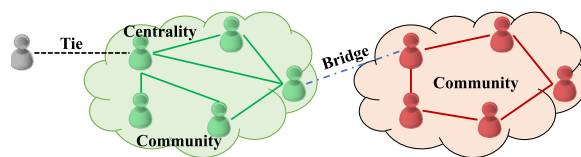


Fig. 4. A D2D communication architecture based on a social networking.

- **Community:** It is similar to a group on Facebook and helps in allocating more resources to all the UEs in a community to decrease content duplication and increase the network throughput.
- **Centrality:** It is similar to a node that has more communication links/friends in a social network. The concept of centrality in D2D communication reduces congestion and increases the network throughput by allocating more resources to a central node.
- **Bridges:** They are similar to a connection between two communities. Hence, two devices forming a bridge can be allocated more resources as compared to other devices.

Channel allocation methods: Two cooperative channel allocation methods, *frame-by-frame* and *slot-by-slot*, are given in [80]. Consider three zones A, B, and C with some UEs such that A and B, B and C intersect, but A and C do not intersect, and B holds a UE that communicates with UEs of A and C. In the frame-by-frame channel allocation method, UEs of A and C do intra-zone communication at different frames, and the UEs of B also communicate at different frames. However, in the slot-by-slot channel allocation method, UEs of A and C do intra-zone communication at an identical time, and of course, UEs of B communicate at a different time. Both the methods improve the efficiency of frequency division multiplexing and increase the network throughput.

Hoang et al. [82] provided an iterative algorithm for subcarrier³ and power allocation such that the minimal individual link data rates and proportionate fairness among D2D links are obtained. A 2-phase service-aware resource allocation scheme, called SARA, is proposed in [41]. In the first phase of SARA, resources are allocated on-demand to meet different service requirements of D-UEs, and in the second phase of SARA, the remaining resources are allocated to D-UEs such that the system throughput increases.

Wang et al. [81] provided a delay-aware and dynamic power control mechanism that adapts the transmit power of D-UEs based on instantaneous values of CSI, and hence, finds the urgency of the data flow. The dynamic power control selects a power control policy so that the long-term average delay and the long-term average power cost of all the flows minimize.

Advantages of D2D communication. D2D communication results in link reliability among D-UEs, a higher data rate to D-UEs, instant communication, an easy way for peer-to-peer file sharing, local voice services, local video

³ A subcarrier is a signal carried on a main radio signal, and hence, two signals are transmitted at an identical time.

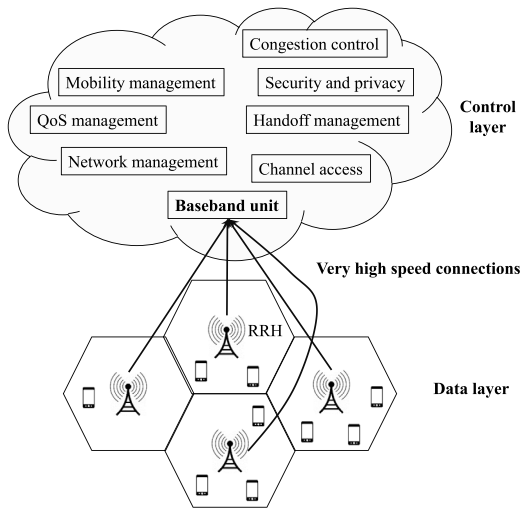


Fig. 5. A basic cloud-based architecture for 5G networks.

streaming, local online gaming, an improved spectral efficiency, decreased power consumption of D-UEs, and the traffic offload from a MBS.

Open issues

- **Security and privacy:** In D2D communication, D-UEs may take help from other UEs as relay nodes; hence, it is required to communicate and transfer data in secure and privacy-preserving manners. Consequently, the designing of energy-efficient and trust-making protocols is an open issue.
- **Network coding scheme:** When D2D communication uses relay nodes, an efficient network coding scheme may be utilized for improving the throughput [43].
- **Multi-mode selection:** In the current design of D2D communication, UEs can do either D2D communication or communication to a BS [79]; however, it is not efficient. Hence, there is a need to design a system that allows a UE to engage the two types of communication modes (i.e., D2D communication and communication to a BS) simultaneously.

4.4. Cloud-based architectures

Cloud computing [83] infrastructure provides on-demand, easy, and scalable access to a shared pool of configurable resources, without worrying about the management of resources. The inclusion of the cloud in the mobile cellular communication can provide its benefits to the communication system. In this section, we will see cloud-based architectures or cloud-based radio access networks (C-RANs) for 5G networks. A detailed review of C-RANs is given in [84].

The main idea of a C-RAN. The first C-RAN is provided by China Mobile Research Institute [47]. The basic idea behind any C-RAN is to execute most of the functions of a MBS in the cloud, and hence, divide the functionality of a MBS into a *control layer* and a *data layer*; see Fig. 5. The functions of the control and the data layers are executed in a cloud and in a MBS, respectively. Thus, a C-RAN provides a *dynamic*

service allocation scheme for scaling the network without installing costly network devices.

Specifically, a MBS has two main components, as: (i) a baseband unit (BBU, for implementing baseband processing using baseband processors), and (ii) a remote radio head (RRH, for performing radio functions). In most of the C-RANs, BBUs are placed in the cloud and RRHs stay in MBSs. Thus, a C-RAN provides an easily scalable and flexible architecture. We will see advantages of C-RANs at the end of this section.

Challenges in the deployment of a C-RAN

- **An efficient fronthaul data transfer technique:** A flexible cloudification of the functions of a MBS comes at the cost of efficient fronthaul data transfer from RRHs to BBUs. The fast and efficient data transfer to the cloud has a proportionate impact on the performance of a C-RAN [49].
- **Real-time performance:** Since C-RANs will be used instead of a MBS that provides all the services to users, it is required to transfer and process all the data in the cloud as fast as a MBS can do; otherwise, it is hard to find solutions to real-time problems using a C-RAN [49].
- **Reliability:** The cloud provider does not ensure any guarantee of failure-free executions of their hardware and software. Thus, it is hard to simulate an error-free MBS using a C-RAN.
- **Security:** The resources of the cloud are shared among several users and never be under the control of a single authority. Hence, a malicious user may easily access the control layer of a C-RAN, resulting in a more severe problem.
- **Manageability:** It is clear that a non-secure C-RAN may be accessed by any cloud user, which poses an additional challenge in manageability of C-RANs. Further, the dynamic allocation of the cloud resources at a specific time interval is a critical issue; otherwise, a C-RAN may face additional latency [85].

Now, we will see some C-RAN architectures in brief.

2-layered C-RAN architectures. The authors [47] provided two C-RAN architectures based on the division of functionalities of a MBS, as: (i) full centralized C-RAN, where a BBU and all the other higher level functionalities of a MBS are located in the cloud while a RRH is only located in the MBS, and (ii) partially centralized C-RAN, where a RRH and some of the functionalities of a BBU are located in the MBS while all the remaining functions of the BBU and higher level functionalities of the MBS are located in the cloud. Thus, the authors [47] proposed the use of only two layers, namely a control layer and a data layer for implementing C-RANs, as follows:

1. **Data layer:** It contains heterogeneous physical resources (e.g., radio interface equipment) and performs signal processing tasks (e.g., channel decoding, demultiplexing, and fast Fourier transformation).
2. **Control layer:** It performs baseband processing and resource management (application delivery, QoS, real-time communication, seamless mobility, security, network management, regulation, and power control); see Fig. 5.

Rost et al. [49] introduced RAN-as-a-service (RANaaS) concept, having the control and the data layers. However, in RANaaS, a cloud provides flexible and on-demand RAN functionalities (such as network management, congestion control, radio resource management, medium access control, and physical layer management), according to the network requirements and characteristics, unlike [47]. Hence, there is no need to split functionalities in advance to the control and the data layers, as a result RANaaS provides more elasticity.

Till now, it is clear that how a C-RAN will work. However, in order to achieve real-time performance, a RRH executing latency-critical applications may connect to a nearby small cloud while other RRHs that are not adhered to real-time applications may connect to a far larger cloud [28]. SoftAir [86] is also a two-layered C-RAN that performs mobility-management, resource-efficient network virtualization, and distributed and collaborative traffic balancing in the cloud.

3-layered C-RAN architectures. The full-centralized C-RAN architecture [47] has some disadvantages, as: continuous exchange of raw baseband samples between the data and the control layers, and the control layer is usually far away from the data layer resulting in a processing delay.

In order to remove these disadvantages, Liu et al. [48] proposed *convergence of cloud and cellular systems* (CONCERT). In this architecture, one more layer, called a *software-defined service layer*, is introduced at the top of the control layer. The functioning of the layers in CONCERT is as follows:

1. *Data layer*: is identical to the full centralized C-RAN's data layer [47], having RRHs with less powerful *computational resources* for application level computations.
2. *Control layer*: works just as a logically centralized entity. The control layer coordinates with the data layer resources and presents them as virtual resources to the software-defined service layer. The control layer provides a few services as: radio interfacing management, wired networking management, and location-aware computing management to the data layer.
3. *Software-defined services layer*: works as a virtual BS and provides services (e.g., application delivery, QoS, real-time communication, seamless mobility, security, network management, regulation, and power control) to the data layer.

Wu et al. [20] enhanced C-RAN architecture [47] and RANaaS [49], by moving the whole RAN to a cloud. The proposed architecture also has three layers, where the data layer and the control layer are same to the respective layers of C-RAN [47,48]. The third layer, called a *service layer*, executes in the cloud and provides some more functionalities than the software-defined services layer of [49], e.g., traffic management, the cell configuration, interference control, allocation of functional components to the physical elements, and video streaming services.

The authors [50] proposed an all-software-defined network using three types of hierarchical network controllers, namely MBS controller, RAN controller, and network controller, where except the MBS controller all the others can be executed in the cloud, as follows:

1. *MBS controller*: usually stays nearby UEs, and performs wireless resource management and packet creation.
2. *RAN controller*: stays at the top of MBS controllers, and performs connectivity, RAT selection, handoff, QoS, policies, mobility management.
3. *Network controller*: stays at the top of RAN controllers, ensures end-to-end QoS, and establishes application-aware routes.

Advantages of C-RANs in 5G networks. C-RANs provide a variety of services as a software, power efficient, flexible, and scalable architecture for the future cellular communication. Here, we enlist some advantages of C-RANs, as follows:

- *An easy network management*: C-RANs facilitate on-demand installation of virtual resources and execute cloud-based resources that dynamically manage interference, traffic, load balance, mobility, and do coordinated signal processing [87,49].
- *Reduce cost*: It is very costly and time-consuming to deploy and install a MBS to increase the network capacity. However, the deployment of C-RANs involves less cost, while it provides usual services like a MBS [20]. As a result, operators are required to only deploy, install, and operate RRHs in MBSs.
- *Save energy of UEs and a MBS*: C-RANs offload data-intensive computations from a MBS and may store data of UEs and MBSs. Consequently, C-RANs allow UEs and MBSs to offload their energy-consuming tasks to a nearby cloud, which saves energy of UEs and MBSs [88].
- *Improved spectrum utilization*: A C-RAN enables sharing of CSI, traffic data, and control information of mobile services among participating MBSs, and hence, results in increased cooperation among MBSs and reduced interference [20].

Open issues. Transferring data from RRHs to BBUs, i.e., from the data layer to the control layer, is a crucial step based on the selection of the functions of a MBS that has to be sent to a cloud, resulting in the minimal data movement in the network [49]. However, the selection of functions to be executed in a cloud and a MBS is a non-trivial affair [49]. The security and privacy issues involved in the cloud computing effect C-RANs, and hence, the development of a C-RAN has to deal with inherent challenges associated with the cloud and the wireless cellular communication simultaneously.

4.5. Energy-efficient architectures for 5G networks

Energy-efficient infrastructures are a vital goal of 5G networks. Researchers have proposed a few ways of reducing energy in the infrastructure. Rowell et al. [18] considered a joint optimization of energy-efficiency and spectral-efficiency. A user-centric 5G network is suggested in [18] so that UEs are allowed to select UL and DL channels from different BSs depending on the load, channel conditions, services and application requirements. In a similar manner, decoupling of signaling and data is useful for energy saving; for example, a MBS may become a signaling BS while SBSs may serve all data requests. Thus,

when there is no data traffic in a SBS, it can be turned off. A similar approach for decoupling of signaling and data is presented in [27]. However, in [27], a UE gets connected to a SBS according to instructions by a MBS, and hence, it results in less energy consumption at UEs' side due to less interference, faster small-cells' discovery, and MBS-assisted handover.

Hu and Qian [45] provided an energy-efficient C-RAN architecture in a manner that RRHs serve almost a same number of UEs. They also present an interference management approach so that the power consumption of SBSs and MBSs can be decreased. Like Rowell et al. [18], Hu and Qian [45] also suggested that the association of a UE cannot be done based on entirely a DL channel or a UL channel, and a UE must consider both the channels at the time of association with a BS. Lin et al. [89] suggested to include an energy harvesting device (to collect energy) and a spectrum harvesting controller (to collect spectrum) at SBSs.

5. Implementation issues in 5G networks

In this section, we will see issues regarding the interference, handoff, QoS, load balancing, and channel access management in the context of 5G networks.

5.1. Interference management in 5G networks

We have already seen challenges in interference management (Section 3). In this section, we will review some techniques/methods for interference management in 5G networks.

Nam et al. [23] handled UE-side interference by using a new type of receiver equipment, called an advanced receiver, which detects, decodes, and removes interference from receiving signals. In addition, the network-side interference is managed by a joint scheduling, which selects each UE according to the resources needed (e.g., time, frequency, transmission rate, and schemes of multiple cells) for its association with a BS. Hence, the joint scheduling, which can be implemented in a centralized or distributed manner, requires a coordination mechanism among the neighboring cells.

Hossain et al. [29] proposed distributed cell access and power control (CAPC) schemes for handling interference in multi-tier architectures. CAPC involves: (i) prioritized power control (PPC), which assumes that UEs working under a SBS have a low-priority than UEs working under a MBS, and hence, low-priority UEs set their power so that the resulting interference must not exceed a predefined threshold; (ii) cell association (CA), which regards dynamic values of resources, traffic, distance to a MBS, and available channels at a MBS for selecting a MBS with the optimum values of the parameters; and (iii) resource-aware CA and PPC (RCA-PPC), which is a combination of the first two approaches and allows a UE to connect simultaneously with multiple BSs for a UL channel and a DL channel based on criteria of PPC and CA.

Hong et al. [31] suggested to use self-interference cancellation (SIC) in small-cells' networks. As we have seen that SBSs require methods to transfer backhaul data to a

MBS (Section 4.1.1), the use of SIC can eliminate the need of such methods and result in *self-backhauled small-cells* (SSCs). SSCs use SIC for providing services and backhaul data transfer, and more importantly, they gain almost the same performance as having a small-cell connected with a wired optical fiber. It works as: in the DL channel, a SBS may receive from a MBS and simultaneously transmit to UEs. In the UL channel, a SBS may receive from UEs and simultaneously transmit data to the MBS. Therefore, a small-cell can completely remove the need of a separate backhaul data transfer method, resulting in reduced cost.

The authors [18] suggested that the measurement of inter-user interference channel, and then, the allocation of UL and DL channels by a MBS can mitigate *inter-user UL-to-DL interference* in a *single-cell* full duplex radio network. However, in the case of a *multi-cell* full duplex radio network, interference mitigation becomes more complex, because of the existence of interference in UL and DL channels between multi-cells' UEs that work on identical frequency and time.

Open issues. Interference cancellation in a full duplex radio network is still an open problem for multi-cell. The design of algorithms for inter-BS DL-to-UL-interference and inter-user UL-to-DL-interference cancellation in a full duplex radio network is still open and to be explored.

5.2. Handoff management in 5G networks

Handoff provides a way to UEs connected to a BS to move to another BS without disconnecting their sessions.

Challenges in the handoff process in 5G networks.

The handoff management in 5G networks has inherent challenges associated with the current cellular networks, e.g., minimum latency, improved routing, security, and less uncertainty of having no services. The network densification, very high mobility, the zero latency, and accessing multi-RATs make handoff management in 5G networks harder. Also, the current cellular networks do not provide efficient load balancing for a BS at the time of handoff. For example, movement of UEs from houses to offices in the morning creates a load imbalance at BSs of respective areas.

Types of handoff in 5G networks. Three types of handoffs are presented in the context of 5G networks, as follows:

1. *Intra-macrocell handoff*: refers to handoff between small-cells that are working under a single MBS [58].
2. *Inter-macrocell handoff*: refers to handoff between macrocells. It may also lead to handoff between two small-cells that are working under different MBSs. Note that if the handover between small-cells of two different MBSs is not done properly, then the inter-macrocell handoff also fails [58].
3. *Multi-RATs handoff*: refers to handoff of a UE from a RAT to other RAT.

Song et al. [58] provided a handoff mechanism for highly mobile users, where a UE sends some parameters (e.g. QoS, signal-to-interference ratio (SIR), and time to handoff) in a measurement report to the current MBS. SIR is considered as a primary factor for finding a situation

for an initiation of the handoff. The Gray system model predicts the $(N + 1)$ th measurement report from the N th measurement report. The predicted value is used for the final decision for the handover process. Zhang et al. [27] proposed a handoff mechanism assisted by a MBS. The MBS collects several parameters from UEs, and if the MBS finds the values of the received parameters below a threshold, then it finds a new SBS or MBS for handoff and informs to the UEs.

For handoff over different RATs, Orsino et al. [57] proposed a handoff procedure so that a UE can select the most suitable RAT without any performance loss. A UE collects received signal strength (i.e., RSRP) or quality (i.e., RSRQ) from the current MBS, and then, it initiates handoff if RSRQ is below a threshold. The UE collects several parameters (e.g., transmitted power, the cell's traffic load, and UE requested spectral efficiency) from adjacent BSs, and then, selects the most suitable BS.

Duan et al. [56] provided an authenticated handoff procedure for C-RANs and multi-tier 5G networks. The control layer holds an authentication handover module (AHM) for monitoring and predicting the future location of UEs (based on the current location) and preparing relevant cells before UEs' arrival in that. The AHM holds a master public-private key pair with each RRH, which are authenticated by the AHM in off-peak hours, and UEs are verified before accessing the network services by RRHs. UEs send ID, the physical layer's attributes, location, speed, and direction to the control layer in a secure manner for the handoff process. The proposed approach reduces the risk of impersonation and man-in-the-middle attacks. Giust et al. [59] provided three distributed mobility management protocols, the first is based on the existing Proxy Mobile IPv6 (PMIPv6), the second is based on SDN, and the third is based on routing protocols.

Open issues. Handoff mechanisms are yet to be explored. It will be interesting to find solutions to extremely dense HetNets. Furthermore, the handoff process may also create interference to other UEs; hence, it is required to develop algorithms while considering different types of interferences in 5G networks and a tradeoff between the number of handoffs and the level of interference in the network [90].

5.3. QoS management in 5G networks

We have already seen challenges in QoS management in Section 3. In this section, we will review some techniques/methods for QoS management in 5G networks.

Zhang et al. [62] provided a mechanism for different delay-bounded QoS for various types of services, applications, and users having diverse requirements, called heterogeneous statistical delay-bounded QoS provisioning (HSP). HSP maximizes the aggregate effective capacity of different types of 5G architectures, reviewed in Sections 4.1–4.3. HSP [62] algorithm claims better performance over other approaches; however, it imposes new challenges in terms of the assignment of different resources for different links under the cover of delay-bounded QoS requirements.

The authors [17] proposed the deployment of a quality management element (QME) in the cloud for monitoring inter-UEs and inter-layer (the control and the data layers in C-RANs) QoS. RRUs send wireless information (e.g., CSI, reference signal received quality, and resource block utilization) to the QME. Consequently, the QME executes service control algorithms (to manage QoS and some other activities like traffic offloading and customized scheduling), and then, sends scheduling strategies to the RRUs to achieve a desired level of QoS.

Kim et al. [63] provided a routing algorithm for multi-hop D2D communication. The algorithm takes into account different QoS for each link so that it can achieve better performance than max-min routing algorithms. The algorithm increases flow until it provides the desired QoS or reaches the maximum capacity of the link. Since the algorithm considers individual links, there is a high probability that some of the channels will serve multiple-links with the desired QoS. Zhou et al. [91] provided a QoS-aware and energy-efficient resource allocation algorithm for DL channels, where UEs are allocated an identical power in one case and non-identical power in the second case. The algorithm maximizes energy-efficiency while minimizes transmit power. Hu and Qian [45] suggested that a UE must consider the source data rate, delay bound, and delay-violation probability before connecting to a MBS.

Open issues. The 5G networks are supposed to satisfy the highest level of QoS. The tactile Internet [4] requires the best QoS, especially, latency of the order of 1 millisecond for senses such as touching, seeing, and hearing objects far away, as precise as human perception. However, the current proposed architectures do not support efficient tactile Internet services. In future, it would be a promising area as to encode senses, exchange data satisfying the zero latency, and enable the user to receive the sensation.

5.4. Load balancing in 5G networks

Load balancing means allocation of resources to a cell such that all the users meet their demands. It is an important issue in the cellular wireless networks. In a 2-tier architecture, discussed in Section 4.1, user offloading to a small-cell is useless if there is no resource partitioning [92]. Singh and Andrews [92] provided an analytical and tractable framework for modeling and analyzing joint resource partitioning and offloading in a 2-tier architecture. Hossain et al. [29] provided a technique for cell association based on dynamic resources and traffic in a cell, as we have already discussed in Section 5.1. In a fast moving vehicle, e.g., a train, it is very hard to allocate resources without any service interruption at all. A distributed load balancing algorithm for fast moving vehicles is presented in [93]. Load balancing architectures for D2D communication are given in [79,41,82], which we discussed in Section 4.3. Interested readers may refer to [94] for further details of load balancing in 5G networks.

5.5. Channel access control management in 5G network

Channel access protocols allow several UEs to share a transmission channel without any collision while utilizing the maximum channel capacity.

Challenges in channel access control management in 5G networks. Channel access control management in 5G networks faces inherent challenges associated with the current cellular networks, e.g., synchronization, fairness, adaptive rate control, resource reservation, real-time traffic support, scalability, throughput, and delay. In addition, providing the currently available best channel in 5G networks is prone to additional challenges, as: high mobility of UEs, working at the higher frequencies (>3 GHz), different RATs, dense networks, high QoS, high link reliability, and the zero latency for applications and services.

The authors [95] proposed a frame-based medium access control (FD-MAC) protocol for mmWave-based small-cells. FD-MAC consists of two phases, as: (i) scheduling phase, when a SBS collects the traffic demands from the supported UEs and computes a schedule for data transmission, and (ii) transmission phase, when the UEs start concurrent transmissions following the schedule. A schedule, which is computed using a graph-edge coloring algorithm, consists of a sequence of topologies and a sequence of time intervals for indicating how long each topology should sustain.

Liu et al. [89] provided MAC protocols for UEs of small-cells and macrocells. Two types of MAC protocols for SBS's UEs are suggested, as: contention-based random channel access (CRCA), where UEs randomly access the channel and send messages if the channel is available, and reservation-based channel access (RCA), which uses time division multiple access. For macrocells' UEs, they provided a MAC protocol for CRN-based 5G networks, where SUs sense a licensed channel until it is free or the residual energy of SUs exceeds a predetermined threshold, for saving their battery. Further, they evaluated a tradeoff between the network throughput and sensing overhead.

The authors [60] suggested two CRN-based channel access techniques for cognitive SBSs. The first channel access scheme is termed as contention-resolution-based channel access (CCA), which is similar to CRCA [89], is based on carrier sense multiple access. The other channel access scheme is termed as uncoordinated aggressive channel access (UCA), which works identically as RCA [89], for aggressively using channels in a small-cell for increasing opportunistic spectrum access performance. Nikopour et al. [96] provided a multi-user sparse code multiple access (MU-SCMA) for increasing DL's spectral efficiency. MU-SCMA does not require the complete CSI, and hence, provides high data rate and the robustness to mobility. Nikopour et al. [96] also provided an uplink contention based SCMA for massive connectivity, data transmission with low signaling overhead, low delay, and diverse traffic connectivity requirements.

Open issues. The current channel access protocols do not regard QoS and latency challenges in 5G networks. Hence, there is a scope of designing algorithms for finding multiple reliable links with the desired QoS and the zero latency.

5.6. Security and privacy management in 5G networks

In this section, we present security and privacy related challenges and a discussion of security and privacy protocols in the context of 5G networks.

Challenges in security and privacy in 5G networks. *Authentication* is a vital issue in any network. Due to the zero latency guarantee of 5G networks, authentication of UEs and network devices is very challenging, since the current authentication mechanisms use an authentication server that takes hundreds of milliseconds delay in a preliminary authentication phase [56]. A *fast and frequent handover* of UEs over small-cells requires for a robust, efficient, and secure handoff process for transferring context information [56]. Security to *multi-RATs* selection is also challenging, since each RAT has its own challenges and certain methods to provide security; clearly, there is a need to provide overlapped security solutions across the various types of RATs. C-RANs also inherit all the challenges associated with the cloud computing and wireless networks. In addition, several other challenges (e.g., authorization and access control of UEs, availability of the network, confidentiality of communication and data transfer, integrity of communication and data transmission, accounting and auditing of a task, low computation complexity, and communication cost) require sophisticated solutions to make a secure 5G network. The security and latency are correlated as a higher level of security and privacy results in increased latencies. Therefore, the communication satisfying the zero latency is cumbersome when combined with secure and privacy-preserving 5G networks.

Monitoring is suggested for securing the network and detecting intruders [65,66]. However, monitoring a large number of UEs (by a trusted authority) is not a trivial task; hence, we do not see monitoring as a preferred way to secure networks. Yang et al. [64] focused on the physical layer security, which is independent of computational complexity and easily handles any number of devices. The physical layer security protocol considers locations of UEs and provides the best way for UEs for securely selecting a MBS or a SBS without overloading the network. Tehrani et al. [42] provided a method for secure and private D2D communications, called close access, where D-UEs have a list of other trusted D-UEs devices, and all such UEs can communicate directly using an encryption scheme while the remaining UEs not in the list utilize a MBS-assisted communication. An encryption based video sharing scheme is also presented in [68]. Kantola et al. [69] proposed a policy-based scheme that can prevent DoS and spoofing attacks. A solution to secure handoff is given in [56] and discussed in Section 5.2.

Open issues. The current security and privacy solutions to 5G networks are not impressive and, hopefully, unable to handle massive connections. We can clearly visualize a potential scope for developing *latency-aware protocols* along with security awareness that must consider secure data transmission, end-to-end security, secure and private storage, threats resistant UEs, and valid network and software access.

6. Methodologies and technologies for 5G networks

It is already mentioned in Section 1 that the development of 5G networks requires the design and implementation of new methodologies, techniques, and architectures. We have reviewed some of the methodologies and technologies in the previous sections, such as: (i) full duplex radios in Sections 3 and 5.1, (ii) CRNs in Section 4.2, (iii) D2D communication in Section 4.3, (iv) multi-tier heterogeneous deployment or dense-deployment techniques in Sections 4.1–4.3, (v) C-RANs in Section 4.4, (vi) ‘green’ communication systems in Section 4.5, and (vii) techniques related to interference, QoS, handoff, channel access, load balancing in Section 5.

In this section, we will briefly describe some techniques that are mentioned but not explained in earlier sections.

Self-Interference Cancellation (SIC). When a full duplex radio receives signals from another radio, it also receives interference signals by its own transmission, resulting in self-interference. Hence, a full duplex radio has to implement techniques to cancel self-interference [22,53,52]. SIC techniques are classified into passive and active cancellations; see [52]. As advantages, the implementation of SIC enables seamless global roaming, high-throughput services, and low-latency applications in a cost effective manner [31].

Downlink and Uplink Decoupling (DUD). In the current cellular networks, a UE is associated with a BS based on the received signal power in its DL channel, and then, uses the same BS for UL channel transmission [97]. DUD allows a UE to select a DL channel and a UL channel from two different BSs, based on the link quality, the cell load, and the cell backhaul capacity [25,98]. Therefore, a UE may have the DL channel connected through a BS and the UL channel connected through a different BS, resulting in a user-centric 5G architecture [18] and improving the capacity of UL channels, which is a prime concern.

Network Function Virtualization (NFV). NFV [85] implements network functions such as network address translation, firewalls, intrusion detection, domain name service, the traffic load management, and caching through software running on commodity servers. However, the conventional networks implement these functions on dedicated and application specific servers. Hence, NFV decreases the burden on network operators by not updating dedicated servers/hardware, thereby saves cost.

Software-Defined Networking (SDN). SDN architectures [99,100] partition network control functions and data forwarding functions, thereby the network control functions are programmable, and the network infrastructure handles applications and network services. SDN architectures can be divided into three parts, as: (i) *the software controller*: holds network control functions such as the network manager, APIs, network operating system, and maintaining the global view of the network; (ii) *the southbound part*: provides an interface and a protocol between the controller and SDN-enable infrastructure, where OpenFlow [101] is the most famous protocol that provides communication between the controller and the southbound part; (iii) *the*

northbound part: provides an interface between SDN applications and the controller [102]. Interested readers may refer to [103–105,100,106,107] to find details of SDN, challenges in SDN, and applications of SDN.

Note that SDN, NFV, and C-RANs offload functionalities to software running on commodity servers. However, SDN separates network control functions from data forwarding functions, while NFV implements network functions in software [108]. Besides that C-RANs integrate both SDN and NFV to meet the scalability and flexibility requirements in the future mobile networks [108].

Millimeter Waves (mmWave). The current wireless bandwidth is not able to support a huge number of UEs in 5G networks. Hence, researchers are looking at 30–300 GHz frequency bands, where mmWave communication is proposed for achieving high-speed data transfer. The current research focuses on 28 GHz band, 38 GHz band, 60 GHz band, and the E-band (71–76 GHz and 81–86 GHz) [109]. However, mmWave has several challenges at the physical, MAC, and network layers. There are a number of papers about mmWave, and hence, we are not discussing mmWave in detail. Interested readers may refer to [110,24,109].

Machine-to-Machine (M2M) communication. M2M communication refers to the communication between (network) devices (e.g., sensors/monitoring-devices with a cloud) without human intervention. Some examples of M2M communication are intelligent transport systems, health measurement, monitoring of buildings, gas and oil pipelines, intelligent retail systems, and security and safety systems. However, the development of M2M communication involves several challenges to be handled in the future, as follows: connectivity of massive devices, bursty data, the zero latency, scalability in terms of supporting devices, technologies and diverse applications, fast and reliable delivery of messages, and cost of devices. In addition, the development of efficient algorithms for location, time, group, priority, and multi-hop data transmission management for M2M communication is needed [34,111]. Interested readers may refer to [112–116].

Massive MIMO (mMIMO). mMIMO systems are also known as large-scale antenna systems, very large MIMO, hyper-MIMO, and full-dimension MIMO [38,40]. mMIMO systems use antenna arrays with hundreds of antennas at MBS for simultaneously serving many UEs with a single antenna in identical time and frequency. Hence, expensive equipment are mounted on a MBS [39]. A mMIMO system relies on spatial multiplexing, which in turn relies on the channel knowledge at MBS, on both UL and DL channels. A mMIMO system reduces latency and energy, simplifies the MAC layer, shows robustness against intentional jamming, and increases the capacity due to spatial multiplexing.

Visual Light Communication (VLC). VLC is a high-speed data transfer medium for short range LOS optical links in the future cellular networks [117]. The light-emitting diodes (LEDs) provide VLC using amplitude modulation at higher frequencies and achieve higher data rates while keeping the LED’s primary illumination function unaffected. VLC can be used for outdoor applications, where high power laser-based equipments provide transmission

links, and for indoor application, where LEDs provide short distance transmission links. VLC is an energy-efficient technology, works on a wider range of unregulated frequency bands, shows high spatial reuse, and inherits security due to LOS. However, VLC is sensitive to sunlight and not able to work for a long range without LOS, and hence, confined coverage [118,119]. The implementation of VLC has still some unanswered questions, such as: how VLC will work in a long range without LOS and will it work for back-haul data transfer in multihop?

Fast caching. Caching is a way for storing temporary data for reducing data access from slow memory or the network. In a network, content caching⁴ is popular and answers the request while responding in place of the application servers as a proxy, thereby reducing the amount of hits that are directly sent to the ultimate backend server. In caching, three decisions are prominent as: what to cache, where to cache, and how to cache? The authors [55,24] suggested that UEs will have enough memory in the future, and they can work as a cache for any other UE, since a small amount of popular data requires to be cached. Wang et al. [55] provided a caching mechanism based content centric networking (CCN), assuming 5G networks will include CCN-capable gateways, routers, and MBSSs. CCN provides in-network data storage, also known as *universal caching*, at every node in the network. The cached data is uniquely identified at each node. Accordingly, a user can request for a particular content from the content cache of any device within the network, or the request is forwarded to the actual source of content storage.

Interested readers may refer to [120,24] for finding details about some of the above mentioned methodologies and technologies.

7. Applications of 5G networks

The zero latency, high speed data transfer, and ubiquitous connectivity are the salient features of 5G networks that are expected to serve a wide range of applications and services. In this section, we enumerate the most prominent applications of 5G networks, as follows:

Personal usages. This domain of 5G networks would be capable of supporting a wide range of UEs, from scalable to heterogeneous devices. Also the data demands (e.g., multimedia data, voice communication, and Web surfing), would be satisfied while keeping the QoS requirements.

Virtualized homes. Due to C-RAN architectures, users may have only low cost UEs (e.g., set-top box for TVs and residential gateways for accessing the Internet) with services of the physical and data link layers. All the other higher layers' applications may move to the cloud for universal access and outsourced computation services [85].

Smart societies. It is an abstract term for connected virtualized homes, offices, and stores. Accordingly, every digital

and electronic services/appliances, e.g., temperature maintenance, warning alarms, printers, LCDs, air conditioners, physical workout equipment, and door locks, would be interconnected in a way that the collaborative actions would enhance the user experience. Similarly, smart stores would assist in filtering out irrelevant product details, sale advertisements, and item suggestions on the go.

Smart grids. The smart grids would decentralize the energy distribution and better analyze the energy consumption. This would allow the smart grids to improve efficiency and economic benefits. The 5G networks would allow a rapid and frequent statistical data observation, analysis, and fetching from remote sensors and would adjust the energy distribution accordingly.

The tactile Internet. The tactile Internet improves the user experience in a virtual environment to an extent of only milliseconds of interaction latency. The futuristic applications such as automated vehicle platooning, self-organizing transportation, the ability to acquire a virtual sense for physically challenged patients, synchronized remote smart-grids, remote robotics, and image processing with customized/panoramic view would use the tactile Internet protocols [4].

Automation. Self-driving vehicles would take place in the near future, and as a requirement, vehicles would communicate with each other in real-time. Moreover, they would communicate with other devices on the roads, homes, and offices with a requirement of almost zero latency. Hence, an interconnected vehicular environment would provide a safe and efficient integration with other information systems.

Healthcare systems. A reliable, secure, and fast mobile communication can strengthen medical services, e.g., frequent data transfer from patients' body to the cloud or health care centers. Therefore, the relevant and urgent medical services could be predicted and delivered to the patients very fast.

Logistics and tracking. The future mobile communication would also assist in inventory or package tracking using location based information systems. The most popular way would be to embed a radio frequency identification (RFID) tag and to provide a continuous connectivity irrespective of the geographic locations.

Industrial usages. The zero latency property of 5G networks would help robots, sensors, drones, mobile devices, users, and data collector devices to have real-time data without any delay, which would help to manage and operate industrial functions quickly while preserving energy.

8. Real demonstrations and testbeds for 5G networks

In this section, we present some real demonstrations and testbeds for 5G Networks. DOCOMO [10] is developing a real-time simulator for evaluating and simulating mMIMO, small-cell, and mmWave. The simulator has shown 1000-times increase in the system capacity, and 90% users achieved 1 Gbps data rate. DOCOMO also performed a real experiment in the year 2012, where data was uploaded at the speed of 10 Gbps.

⁴ https://developer.akamai.com/stuff/Caching/Content_Caching.html.

Samsung performed data transmission experiment using mmWave at 28 GHz frequency band and achieved the world's first highest data rate of 1.2 Gbps on a vehicle running at the speed of >100 km/h. Further, when the vehicle was nearby a stop, the data transmission speed was achieved up to 7.5 Gbps. In the experiment, the peak data rate was more than 30-times faster as compared to the state-of-the-art 4G technology. Samsung is also developing array antennas that have nearly zero-footprint and reconfigurable antenna modes [13]. Ericsson has achieved the speed of 5 Gbps in a demonstration.⁵ Huawei at University of Surrey in Guildford is developing a testbed, which would be used for developing methodologies, validating them, and verifying a C-RAN for an ultra-dense network.⁶

The European Commission funded METIS [7] has developed more than 140 technical components, including: air interface technologies, new waveforms, multiple access and MAC schemes, multi-antenna and mMIMO technologies, multi-hop communications, interference management, resource allocation schemes, mobility management, robustness enhancements, context aware approaches, D2D communication, dynamic reconfiguration enablers, and spectrum management technologies. They have also implemented some testbeds, as: three D2D communication related testbeds, one massive machine communications related testbed, and one related to waveform design. There are other projects funded by the European Commission working on the development of 5G networks, as: 5GNOW (<http://www.5gnow.eu/>), TROPIC (www.ict-tropic.eu), MCN (www.mobile-cloud-networking.eu), COMBO (www.ict-combo.eu), MOTO (www.fp7-moto.eu), and PHYLAWS (www.phylaws-ict.org).

Small-cells: Sprint, Verizon, and AT&T in the United States, Vodafone in Europe, and Softbank in Japan are developing femtocells [61]. Alcatel-Lucent, Huawei, and Nokia Siemens Networks has been involved in the development of plug-and-play SBSs [44]. **mMIMO:** TitanMIMO by Nutaq is a testbed for 5G mMIMO. The TitanMIMO-4 testbed provides a realistic throughput by aggregating the entire mMIMO channel into a central baseband processing engine.⁷ Lund University and National Instruments, Austin are also involved in the development of testbeds for mMIMO [121]. DOCOMO [10] is developing a real-time simulator for mMIMO.

C-RANs: China Mobile Research Institute has developed C-RANs [47]. The European Commission supported iJoin (<http://www.ict-ijoin.eu/>) is also involved in developing C-RAN architectures, especially, RANaaS [49]. IBM, Intel, Huawei, and ZTE are developing C-RANs [20]. OpenAirInterface (OAI) [122] is an open experimentation and prototyping platform created by the Mobile Communications Department at EURECOM. OAI used for two purposes as: C-RANs and M2M communication. **Full duplex radio:** groups

at Stanford and Rice Universities are focusing on the development of full duplex radios [22]. **CRNs:** Vodafone Chair Mobile Communications Systems at the TU Dresden has created a testbed for studying CRNs in the future cellular networks [123].⁸ **mmWave:** DOCOMO [10], Huawei, the European Commission funded projects, and New York University [110] are developing methodologies and testbeds for mmWave and have successfully performed various experiments.

9. Concluding remarks

In this survey, we discussed salient features, requirements, applications, and challenges involved in the development of the fifth generation (5G) of cellular mobile communication that is expected to provide very high speed data transfer and ubiquitous connectivity among various types of devices. We reviewed some architectures for 5G networks based on the inclusion of small-cells, cognitive radio networks, device-to-device communication, and cloud-based radio access networks. We find out that energy consumption by the infrastructure is going to be a major concern in 5G networks, and hence, reviewed energy-efficient architectures. We figured out several open issues, which may drive the future inventions and research, in all the architectures.

The development of new architectures is not only a concern in 5G networks; there will be a need for handling other implementation issues in the context of users, e.g., interference removal, handoff management, QoS guarantee, channel accessing, and in the context of infrastructures, e.g., load balancing. During our illustration, we included several new techniques, e.g., full duplex radios, dense-deployment techniques, SIC, DUD, mmWave, mMIMO, and VLC. We also discussed the current trends in research industries and academia in the context of 5G networks based on real-testbeds and experiments for 5G networks.

We conclude our discussion with a resonating notion that the design of 5G infrastructure is still under progress. The most prominent issues are enlisted below, and to provide elegant solutions to these issues would contribute in early deployment as well as in long run growth of 5G networks.

- The security and privacy of devices, infrastructures, communication, and data transfer is yet to be explored. We believe that the current solutions based on encryption would not suffice in the future due to a huge number of devices. Intuitively, a solution that would use an authenticated certificate may be feasible [124].
- The development of network devices, infrastructures, and algorithms must be self-healing, self-configuring, and self-optimizing to preform dynamic operations as per the need, for example, dynamic load balancing, QoS guarantee, traffic management, and pooling of residual resources.

⁵ <http://www.ericsson.com/research-blog/5g/ericsson-research-hits-5gbps-5g-labs-demo/>.

⁶ <http://www.fiercewireless.com/tech/story/huawei-invests-5g-testbed-university-surrey/2014-11-04>.

⁷ http://www.kanecomputing.co.uk/pdfs/Nutaq_TitanMIMO-4.pdf.

⁸ <https://mns.ifn.et.tu-dresden.de/Chair/Pages/The-Chair.aspx>.

- The cloud computing is an attractive technology in the current trend for various applications. We have reviewed C-RANs; however, the current solutions do not consider the impact of virtualization for backhaul data transfer, the trust of the cloud, inter-cloud communication, ubiquitous service guarantee, and real-time performance guarantee with zero-latency. Thus, the development of C-RANs must address the big question that how much virtualization is good?
- The multi-RATs are attractive solutions to access different RATs. However, would it be possible for devices to use more than one RAT at an identical time for uplink and downlink channels? Further, the network densification must quantify that how much network density is good?
- The design, development, and usage of user devices, service-application models, and, especially, the network devices must be affordable to cater the needs of overwhelming users, service providers, and network providers.
- The zero latency is a primary concern in most of the real-time applications and services, especially, in the tactile Internet. However, all the existing architectures and implementations of 5G networks are far from achieving the zero latency. Therefore, it is utmost desired that the latest real-time and ultra-reliable network configurations must be improved to a latency free environment.

Acknowledgments

Authors are thankful to Shlomi Dolev and anonymous referees for remarks that improved the quality of the paper.

References

- [1] More than 50 billion Connected Devices, Ericsson, White Paper, 2011. Available at: http://www.akos-rs.si/files/Telekomunikacije/Digitalna_agenda/Internetni_protokol_ipv6/More-than-50-billion-connected-devices.pdf.
- [2] Ericsson Mobility Report: On the Pulse of the Networked Society, Ericsson, 2015. Available at: <http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf>.
- [3] 5G Rouse Cases and Requirements, Nokia Networks, White Paper. Available at: http://networks.nokia.com/sites/default/files/document/5g_requirements_white_paper.pdf.
- [4] G.P. Fettweis, The tactile internet: Applications and challenges, *IEEE Veh. Technol. Mag.* 9 (1) (2014) 64–70.
- [5] The Next Generation of Communication Networks and Services, The 5G Infrastructure Public Private Partnership (5GPPP), European Commission, 2015. Available at: <http://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf>.
- [6] Analysis Understanding 5G: Perspectives on Future Technological Advancements in Mobile, GSMA Intelligence, 2014. Available at: <https://gsmaintelligence.com/research/?file=141208-5g.pdf&download>.
- [7] METIS Final Project Report, Deliverable D8.4, METIS, 2015. Available at: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D8.4_v1.pdf.
- [8] X. Ge, H. Cheng, M. Guizani, T. Han, 5G wireless backhaul networks: challenges and research advances, *IEEE Netw.* 28 (6) (2014) 6–11.
- [9] 5G is Coming: Are you Prepared? Alcatel-Lucent, Strategic White Paper 2015. Available at: <http://www2.alcatel-lucent.com/landing/5g/>.
- [10] 5G Radio Access: Requirements, Concept and Technologies, DOCOMO, 5G Paper. Available at: https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/whitepaper_5g/DOCOMO_5G_White_Paper.pdf.
- [11] 5G: A Technology Vision, Huawei, White Paper, 2013. Available at: <http://www.huawei.com/5gwhitepaper/>.
- [12] 5G: Vision for the Next Generation of Connectivity, Qualcomm, 2015. Available at: <https://www.qualcomm.com/media/documents/files/whitepaper-5g-vision-for-the-next-generation-of-connectivity.pdf>.
- [13] 5G Vision, Samsung Electronics Co., Ltd., 2015. Available at: <http://www.samsung.com/global/business-images/insights/2015/Samsung-5G-Vision-0.pdf>.
- [14] J.G. Andrews, S. Buzzi, W. Choi, S.V. Hanly, A.E. Lozano, A.C.K. Soong, J.C. Zhang, What will 5G be? *IEEE J. Sel. Areas Commun.* 32 (6) (2014) 1065–1082.
- [15] R. Chávez-Santiago, M. Szydelko, A. Kliks, F. Foukalas, Y. Haddad, K.E. Nolan, M.Y. Kelly, M.T. Masonta, I. Balasingham, 5G: The convergence of wireless communications, *Wirel. Pers. Commun.* 83 (3) (2015) 1617–1642.
- [16] L. Gavrilovska, V. Rakovic, V. Atanasovski, Visions towards 5G: Technical requirements and potential enablers, *Wirel. Pers. Commun.* (2015) 1–27.
- [17] X. Zhou, Z. Zhao, R. Li, Y. Zhou, T. Chen, Z. Niu, H. Zhang, Toward 5G: when explosive bursts meet soft cloud, *IEEE Netw.* 28 (6) (2014) 12–17.
- [18] C. I. C. Rowell, S. Han, Z. Xu, G. Li, Z. Pan, Toward green and soft: a 5G perspective, *IEEE Commun. Mag.* 52 (2) (2014) 66–73.
- [19] K. Au, L. Zhang, H. Nikopour, E. Yi, A. Bayesteh, U. Vilaipornsawai, J. Ma, P. Zhu, Uplink contention based SCMA for 5G radio access, in: *IEEE GLOBECOM Workshops*, 2014, pp. 900–905.
- [20] J. Wu, Z. Zhang, Y. Hong, Y. Wen, Cloud radio access network (C-RAN) a primer, *IEEE Netw.* 29 (1) (2015) 35–41.
- [21] E. Oh, K. Son, B. Krishnamachari, Dynamic base station switching-on/off strategies for green cellular networks, *IEEE Trans. Wirel. Commun.* 12 (5) (2013) 2126–2136.
- [22] D. Bharadia, E. McMillin, S. Katti, Full duplex radios, in: *ACM SIGCOMM*, 2013, pp. 375–386.
- [23] W. Nam, D. Bai, J. Lee, I. Kang, Advanced interference management for 5G cellular networks, *IEEE Commun. Mag.* 52 (5) (2014) 52–60.
- [24] F. Boccardi, R.W. Heath, A.E. Lozano, T.L. Marzetta, P. Popovski, Five disruptive technology directions for 5G, *IEEE Commun. Mag.* 52 (2) (2014) 74–80.
- [25] H. Elshaer, F. Boccardi, M. Dohler, R. Irmer, Downlink and uplink decoupling: A disruptive architectural design for 5G networks, in: *IEEE GLOBECOM*, 2014, pp. 1798–1803.
- [26] C. Wang, F. Haider, X. Gao, X. You, Y. Yang, D. Yuan, H.M. Aggoune, H. Haas, S. Fletcher, E. Hepsaydir, Cellular architecture and key technologies for 5G wireless communication networks, *IEEE Commun. Mag.* 52 (2) (2014) 122–130.
- [27] X. Zhang, J. Zhang, W. Wang, Y. Zhang, C. I. Z. Pan, G. Li, Y. Chen, Macro-assisted data-only carrier for 5G green cellular systems, *IEEE Commun. Mag.* 53 (5) (2015) 223–231.
- [28] P.K. Agyapong, M. Iwamura, D. Staehle, W. Kiess, A. Benjebbour, Design considerations for a 5G network architecture, *IEEE Commun. Mag.* 52 (11) (2014) 65–75.
- [29] E. Hossain, M. Rasti, H. Tabassum, A. Abdelnasser, Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective, *IEEE Wirel. Commun.* 21 (3) (2014) 118–127.
- [30] X. Hong, J. Wang, C. Wang, J. Shi, Cognitive radio in 5G: a perspective on energy-spectral efficiency trade-off, *IEEE Commun. Mag.* 52 (7) (2014) 46–53.
- [31] S.S. Hong, J. Brand, J.I. Choi, M. Jain, J. Mehlman, S. Katti, P. Levis, Applications of self-interference cancellation in 5G and beyond, *IEEE Commun. Mag.* 52 (2) (2014) 114–121.
- [32] Available at: http://research.microsoft.com/en-us/projects/spectrum/fcc_dynamic_spectrum_access_noi_comments.pdf.
- [33] S. Chen, J. Zhao, The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication, *IEEE Commun. Mag.* 52 (5) (2014) 36–43.
- [34] Requirements and Vision for NG-Wireless, Outlook: Visions and Research Directions for the Wireless World, Wireless World Research Forum, 2011. Available at: <http://www.wvrf.ch/files/wvrf/content/files/publications/outlook/Outlook7.pdf>.
- [35] G. Fettweis, S.M. Alamouti, 5G: Personal mobile internet beyond what cellular did to telephony, *IEEE Commun. Mag.* 52 (2) (2014) 140–145.
- [36] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhvasi, C. Patel, S. Geirhofer, Network densification: the dominant theme for wireless evolution into 5G, *IEEE Commun. Mag.* 52 (2) (2014) 82–89.
- [37] I.F. Akyildiz, W.-Y. Lee, K.R. Chowdhury, CRAHNS: Cognitive radio ad hoc networks, *Ad Hoc Netw.* 7 (5) (2009) 810–836.

- [38] E.G. Larsson, O. Edfors, F. Tufvesson, T.L. Marzetta, Massive MIMO for next generation wireless systems, *IEEE Commun. Mag.* 52 (2) (2014) 186–195.
- [39] L. Lu, G.Y. Li, A.L. Swindlehurst, A.E. Ashikhmin, R. Zhang, An overview of massive MIMO: benefits and challenges, *J. Sel. Top. Signal Process.* 8 (5) (2014) 742–758.
- [40] Y. Nam, B.L. Ng, K. Sayana, Y. Li, J. Zhang, Y. Kim, J. Lee, Full-dimension MIMO (FD-MIMO) for next generation cellular technology, *IEEE Commun. Mag.* 51 (6) (2013).
- [41] B. Chen, J. Zheng, Y. Zhang, H. Murata, SARA: A service-aware resource allocation scheme for device-to-device communication underlying cellular networks, in: *IEEE GLOBECOM*, 2014, pp. 4916–4921.
- [42] M.N. Tehrani, M. Uysal, H. Yanikomeroglu, Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions, *IEEE Commun. Mag.* 52 (5) (2014) 86–92.
- [43] L. Wei, R.Q. Hu, Y. Qian, G. Wu, Enable device-to-device communications underlying cellular networks: challenges and research aspects, *IEEE Commun. Mag.* 52 (6) (2014) 90–96.
- [44] W. Ni, I.B. Collings, X. Wang, R.P. Liu, Multi-hop point-to-point FDD wireless backhaul for mobile small cells, *IEEE Wirel. Commun.* 21 (4) (2014) 88–96.
- [45] R.Q. Hu, Y. Qian, An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems, *IEEE Commun. Mag.* 52 (5) (2014) 94–101.
- [46] C.X. Mavroumoustakis, A. Bourdena, G. Mastorakis, E. Pallis, G. Kormentzas, An energy-aware scheme for efficient spectrum utilization in a 5G mobile cognitive radio network architecture, *Telecommun. Syst.* 59 (1) (2015) 63–75.
- [47] C-RAN: The Road Towards Green RAN, China Mobile Research Institute, white paper, 2011. Available at: http://labs.chinamobile.com/cran/wp-content/uploads/C-RAN_white_paper_v2_5_EN.pdf.
- [48] J. Liu, T. Zhao, S. Zhou, Y. Cheng, Z. Niu, CONCERT: a cloud-based architecture for next-generation cellular systems, *IEEE Wirel. Commun.* 21 (6) (2014) 14–22.
- [49] P. Rost, C.J. Bernardos, A.D. Domenico, M.D. Girolamo, M. Lalam, A. Mäder, D. Sabella, D. Wübben, Cloud technologies for flexible 5G radio access networks, *IEEE Commun. Mag.* 52 (5) (2014) 68–76.
- [50] V. Yazici, U.C. Kozat, M.O. Sunay, A new control plane for 5G network architecture with a case study on unified handoff, mobility, and routing management, *IEEE Commun. Mag.* 52 (11) (2014) 76–85.
- [51] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H.D. Schotten, T. Hidekazu, H.M. Tullberg, M.A. Uusitalo, B. Timus, M. Fallgren, Scenarios for 5G mobile and wireless communications: the vision of the METIS project, *IEEE Commun. Mag.* 52 (5) (2014) 26–35.
- [52] Z. Zhang, X. Chai, K. Long, A.V. Vasilakos, L. Hanzo, Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection, *IEEE Commun. Mag.* 53 (5) (2015) 128–137.
- [53] S. Han, C.I.Z. Xu, C. Pan, Z. Pan, Full duplex: Coming into reality in 2020? in: *IEEE GLOBECOM*, 2014, pp. 4776–4781.
- [54] G. Wu, C. Yang, S. Li, G.Y. Li, Recent advances in energy-efficient networks and their application in 5G systems, *IEEE Wirel. Commun.* 22 (2) (2015) 145–151.
- [55] X. Wang, M. Chen, T. Taleb, A. Ksentini, V.C.M. Leung, Cache in the air: exploiting content caching and delivery techniques for 5G systems, *IEEE Commun. Mag.* 52 (2) (2014) 131–139.
- [56] X. Duan, X. Wang, Authentication handover and privacy protection in 5G hetnets using software-defined networking, *IEEE Commun. Mag.* 53 (4) (2015) 28–35.
- [57] A. Orsino, G. Araniti, A. Molinaro, A. Iera, Effective RAT selection approach for 5G dense wireless networks, in: *IEEE 81st Vehicular Technology Conference, VTC Spring 2015, Glasgow, United Kingdom*, 11–14 May, 2015, pp. 1–5, 2015.
- [58] H. Song, X. Fang, L. Yan, Handover scheme for 5G C/U plane split heterogeneous network in high-speed railway, *IEEE T. Veh. Technol.* 63 (9) (2014) 4633–4646.
- [59] F. Giust, L. Cominardi, C.J. Bernardos, Distributed mobility management for future 5G networks: overview and analysis of existing approaches, *IEEE Commun. Mag.* 53 (1) (2015) 142–149.
- [60] H. ElSawy, E. Hossain, D.I. Kim, HetNets with cognitive small cells: user offloading and distributed channel access techniques, *IEEE Commun. Mag.* 51 (6) (2013).
- [61] W. Wang, Q. Zhang, Local cooperation architecture for self-healing femtocell networks, *IEEE Wirel. Commun.* 21 (2) (2014) 42–49.
- [62] X. Zhang, W. Cheng, H. Zhang, Heterogeneous statistical QoS provisioning over 5G mobile wireless networks, *IEEE Netw.* 28 (6) (2014) 46–53.
- [63] J. Kim, A.F. Molisch, Quality-aware millimeter-wave device-to-device multi-hop routing for 5G cellular networks, in: *IEEE International Conference on Communications, ICC*, 2014, pp. 5251–5256.
- [64] N. Yang, L. Wang, G. Geraci, M. ElKashlan, J. Yuan, M.D. Renzo, Safeguarding 5G wireless communication networks using physical layer security, *IEEE Commun. Mag.* 53 (4) (2015) 20–27.
- [65] N. Ulltveit-Moe, V.A. Oleshchuk, G.M. Koen, Location-aware mobile intrusion detection with enhanced privacy in a 5G context, *Wirel. Pers. Commun.* 57 (3) (2011) 317–338.
- [66] K. Gai, M. Qiu, L. Tao, Y. Zhu, Intrusion detection techniques for mobile cloud computing in heterogeneous 5G, *Secur. Commun. Netw.* (2015).
- [67] Y. Li, B. Kaur, B. Andersen, Denial of service prevention for 5G, *Wirel. Pers. Commun.* 57 (3) (2011) 365–376.
- [68] J.K. Liu, M.H. Au, W. Susilo, K. Liang, R. Lu, B. Srinivasan, Secure sharing and searching for real-time video data in mobile cloud, *IEEE Netw.* 29 (2) (2015) 46–50.
- [69] R. Kantola, J.L. Santos, N. Beijar, Policy-based communications for 5G mobile with customer edge switching, *Wiley Secur. Comm. Netw.* (2015).
- [70] N. Deng, W. Zhou, M. Haenggi, A heterogeneous cellular network model with inter-tier dependence, in: *IEEE GLOBECOM*, 2014, pp. 1522–1527.
- [71] S. Sharma, A.K. Singh, On detecting termination in cognitive radio networks, *Int. J. Netw. Manag.* 24 (6) (2014) 499–527.
- [72] L. Huang, G. Zhu, X. Du, Cognitive femtocell networks: an opportunistic spectrum access for future indoor wireless coverage, *IEEE Wirel. Commun.* 20 (2) (2013) 44–51.
- [73] W. Wang, G. Yu, A. Huang, Cognitive radio enhanced interference coordination for femtocell networks, *IEEE Commun. Mag.* 51 (6) (2013).
- [74] C. Lee, C. Shih, Coverage analysis of cognitive femtocell networks, *IEEE Wirel. Commun. Lett.* 3 (2) (2014) 177–180.
- [75] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, S. Mohanty, NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey, *Comput. Netw.* 50 (13) (2006) 2127–2159.
- [76] C. Jiang, Y. Chen, K.J.R. Liu, Y. Ren, Cognitive femtocell market: How to price? in: *IEEE GLOBECOM*, 2014, pp. 4514–4519.
- [77] D. Gale, L.S. Shapley, College admissions and the stability of marriage, *Amer. Math. Monthly* (1962) 9–15.
- [78] A. Asadi, Q. Wang, V. Mancuso, A survey on device-to-device communication in cellular networks, *IEEE Commun. Surv. Tutorials* 16 (4) (2014) 1801–1819.
- [79] Y. Li, T. Wu, P. Hui, D. Jin, S. Chen, Social-aware D2D communications: qualitative insights and quantitative analysis, *IEEE Commun. Mag.* 52 (6) (2014) 150–158.
- [80] P. Li, S. Guo, T. Miyazaki, W. Zhuang, Fine-grained resource allocation for cooperative device-to-device communication in cellular networks, *IEEE Wirel. Commun.* 21 (5) (2014) 35–40.
- [81] W. Wang, V.K.N. Lau, Delay-aware cross-layer design for device-to-device communications in future cellular systems, *IEEE Commun. Mag.* 52 (6) (2014) 133–139.
- [82] T.D. Hoang, L.B. Le, T. Le-Ngoc, Resource allocation for D2D communications under proportional fairness, in: *IEEE GLOBECOM*, 2014, pp. 1259–1264.
- [83] P. Mell, T. Grance, The NIST definition of cloud computing, 2011.
- [84] A. Checko, H.L. Christiansen, Y. Yan, L. Scolar, G. Kardaras, M.S. Berger, L. Dittmann, Cloud RAN for mobile networks – A technology overview, *IEEE Commun. Surv. Tutorials* 17 (1) (2015) 405–426.
- [85] B. Han, V. Gopalakrishnan, L. Ji, S. Lee, Network function virtualization: Challenges and opportunities for innovations, *IEEE Commun. Mag.* 53 (2) (2015) 90–97.
- [86] I.F. Akyildiz, P. Wang, S.-C. Lin, SoftAir: A software defined networking architecture for 5G wireless systems, *Comput. Netw.* 85 (0) (2015) 1–18.
- [87] Y. Cai, F.R. Yu, S. Bu, Cloud computing meets mobile wireless communications in next generation cellular networks, *IEEE Netw.* 28 (6) (2014) 54–59.
- [88] S. Barbarossa, S. Sardellitti, P.D. Lorenzo, Communicating while computing: Distributed mobile cloud computing over 5G heterogeneous networks, *IEEE Signal Process. Mag.* 31 (6) (2014) 45–55.
- [89] Y. Liu, Y. Zhang, R. Yu, S. Xie, Integrated energy and spectrum harvesting for 5G wireless communications, *IEEE Netw.* 29 (3) (2015) 75–81.
- [90] A. Leu, B. Mark, S. Tang, Analysis of handoff interference and outage along arbitrary trajectories in cellular networks, *IEEE Trans. Wirel. Commun.* 7 (9) (2008) 3597–3607.

- [91] Y. Zhou, D. Li, H. Wang, A. Yang, S. Guo, QoS-aware energy-efficient optimization for massive MIMO systems in 5G, in: Sixth International Conference on Wireless Communications and Signal Processing, WCSP, 2014, pp. 1–5.
- [92] S. Singh, J.G. Andrews, Joint resource partitioning and offloading in heterogeneous cellular networks, *IEEE Trans. Wirel. Commun.* 13 (2) (2014) 888–901.
- [93] L. Goratti, S. Savazzi, A. Parichehreh, U. Spagnolini, Distributed load balancing for future 5G systems on-board high-speed trains, in: 5GU, 2014, pp. 140–145.
- [94] J.G. Andrews, S. Singh, Q. Ye, X. Lin, H.S. Dhillon, An overview of load balancing in hetnets: old myths and open problems, *IEEE Wirel. Commun.* 21 (2) (2014) 18–25.
- [95] I.K. Son, S. Mao, Y. Li, M. Chen, M.X. Gong, T.T.S. Rappaport, Frame-based medium access control for 5G wireless networks, in: *Mobile Networks and Applications*, 2015, pp. 1–10.
- [96] H. Nikopour, E. Yi, A. Bayesteh, K. Au, M. Hawryluck, H. Baligh, J. Ma, SCMA for downlink multiple access of 5G wireless networks, in: *IEEE GLOBECOM*, 2014, pp. 3940–3945.
- [97] K. Smiljkovic, P. Popovski, L. Gavriloska, Analysis of the decoupled access for downlink and uplink in wireless heterogeneous networks, *IEEE Wirel. Commun. Lett.* 4 (2) (2015) 173–176.
- [98] H. Elshaer, F. Boccardi, M. Dohler, R. Irmer, Load & backhaul aware decoupled downlink/uplink access in 5G systems, in: 2015 IEEE International Conference on Communications, ICC 2015, London, United Kingdom, June 8–12, 2015, pp. 5380–5385, 2015.
- [99] H. Kim, N. Feamster, Improving network management with software defined networking, *IEEE Commun. Mag.* 51 (2) (2013) 114–119.
- [100] H. Farhadi, H. Lee, A. Nakao, Software-defined networking: A survey, *Comput. Netw.* 81 (2015) 79–95.
- [101] N. McKeown, T. Anderson, H. Balakrishnan, G.M. Parulkar, L.L. Peterson, J. Rexford, S. Shenker, J.S. Turner, OpenFlow: enabling innovation in campus networks, *Comput. Commun. Rev.* 38 (2) (2008) 69–74.
- [102] E.J. Hernandez-Valencia, S. Izzo, B. Polonsky, How will NFV/SDN transform service provider opex? *IEEE Netw.* 29 (3) (2015) 79–108.
- [103] B. Cao, F. He, Y. Li, C. Wang, W. Lang, Software defined virtual wireless network: framework and challenges, *IEEE Netw.* 29 (4) (2015) 6–12.
- [104] J. Xie, D. Guo, Z. Hu, T. Qu, P. Lv, Control plane of software defined networks: A survey, *Comput. Commun.* 67 (2015) 1–10.
- [105] D. Kreutz, F.M.V. Ramos, P.J.E. Verissimo, C.E. Rothenberg, S. Azodolmolky, S. Uhlig, Software-defined networking: A comprehensive survey, *Proc. IEEE* 103 (1) (2015) 14–76.
- [106] M.Y. Arslan, K. Sundaresan, S. Rangarajan, Software-defined networking in cellular radio access networks: potential and challenges, *IEEE Commun. Mag.* 53 (1) (2015) 150–156.
- [107] I. Alsmadi, D. Xu, Security of software defined networks: A survey, *Comput. Secur.* 53 (2015) 79–95.
- [108] R. Wang, H. Hu, X. Yang, Potentials and challenges of C-RAN supporting multi-RATs toward 5G mobile networks, *IEEE Access* 2 (2014) 1187–1195.
- [109] Y. Niu, Y. Li, D. Jin, L. Su, A.V. Vasilakos, A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges, *Wirel. Netw.* 21 (8) (2015) 2657–2676.
- [110] T.S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G.N. Wong, J.K. Schulz, M. Samimi, F. G. Jr, Millimeter wave mobile communications for 5G cellular: It will work!, *IEEE Access* 1 (2013) 335–349.
- [111] P.F. Stenumgaard, J. Chilo, J. Ferrer-Coll, P. Ångskog, Challenges and conditions for wireless machine-to-machine communications in industrial environments, *IEEE Commun. Mag.* 51 (6) (2013).
- [112] C. Anton-Haro, M. Dohler, Machine-to-machine (M2M) Communications: Architecture, Performance and Applications, Elsevier, 2014.
- [113] Z. Fan, R.J. Haines, P. Kulkarni, M2M communications for e-health and smart grid: an industry and standard perspective, *IEEE Wirel. Commun.* 21 (1) (2014) 62–69.
- [114] G.C. Madueño, C. Stefanovic, P. Popovski, Reliable reporting for massive M2M communications with periodic resource pooling, *IEEE Wirel. Commun. Lett.* 3 (4) (2014) 429–432.
- [115] M. Condoluci, M. Dohler, G. Araniti, A. Molinaro, K. Zheng, Toward 5G densenets: architectural advances for effective machine-type communications over femtocells, *IEEE Commun. Mag.* 53 (1) (2015) 134–141.
- [116] R. Ratasuk, A. Prasad, Z. Li, A. Ghosh, M.A. Uusitalo, Recent advancements in M2M communications in 4G networks and evolution towards 5G, in: *ICIN 2015*, 2015, pp. 52–57.
- [117] S. Wu, H. Wang, Visible light communications for 5G wireless networking systems: from fixed to mobile communications, *IEEE Netw.* 28 (6) (2014) 41–45.
- [118] L. Grobe, A. Paraskevopoulos, J. Hilt, D. Schulz, F. Lassak, F. Hartlieb, C. Kottke, V. Jungnickel, K. Langer, High-speed visible light communication systems, *IEEE Commun. Mag.* 51 (12) (2013) 60–66.
- [119] A. Jovicic, J. Li, T. Richardson, Visible light communication: opportunities, challenges and the path to market, *IEEE Commun. Mag.* 51 (12) (2013) 26–32.
- [120] Q. Li, H. Niu, A. Papanthassiou, G. Wu, 5G network capacity: Key elements and technologies, *IEEE Veh. Technol. Mag.* 9 (1) (2014) 71–78.
- [121] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, F. Tufvesson, A flexible 100-antenna testbed for massive MIMO, in: *IEEE GLOBECOM Workshops*, 2014, pp. 287–293.
- [122] N. Nikaein, M.K. Marina, S. Manickam, A. Dawson, R. Knopp, C. Bonnet, OpenAirInterface: A flexible platform for 5G research, *Comput. Commun. Rev.* 44 (5) (2014) 33–38.
- [123] M. Danneberg, R. Datta, G. Fettweis, Experimental testbed for dynamic spectrum access and sensing of 5G GFDM waveforms, in: *IEEE 80th Vehicular Technology Conference*, 2014, pp. 1–5.
- [124] S. Dolev, Ł. Krzywiecki, N. Panwar, M. Segal, Vehicle authentication via monolithically certified public key and attributes, in: *Wireless Networks*, 2015, pp. 1–18.



Nisha Panwar is a Ph.D. student at Department of Computer Science, Ben-Gurion the Negev, Israel. She received her Master of Technology (M.Tech.) degree Engineering from National Institute of Technology, Kurukshetra, India in 2011. Her research interests include security and privacy issues in vehicular networks, computer network and communication, and distributed algorithms.



Shantanu Sharma is a Ph.D. student at Ben-Gurion University of the Negev, Israel. He received his Master of Technology (M.Tech.) degree in Computer Science from National Institute of Technology, Kurukshetra, India, in 2011. He was awarded a gold medal for the first position in his M.Tech. degree. His research interests include designing models for MapReduce computations, distributed algorithms, mobile computing, and wireless communication.



Awadhesh Kumar Singh received his Bachelor of Technology (B.Tech.) degree in Computer Science from Madan Mohan Malaviya University of Technology, Gorakhpur, India, in 1988, and his Master of Technology (M.Tech.) and Ph.D. degrees in Computer Science from Jadavpur University, Kolkata, India, in 1998 and 2004, respectively. He joined the Department of Computer Engineering at National Institute of Technology, Kurukshetra, India, in 1991, where he is presently a Professor and serving as the Chairman of the department. Earlier, he also served as the Chairman of the department during 2007–2009. He is currently the Chief Vigilance Officer of the institute. His research excellence is acknowledged by the paper in journals, *IEEE Signal Processing Letters*, *Journal of Computer Science and Technology*, *Wireless Personal Communication*, *Journal of Super Computing*, *Journal of Information Science and Engineering*, *Journal of Communication Networks and Information Security*, *Journal of Applied Evolutionary Computation*, and alike. His research interests include distributed algorithms, mobile computing, fault tolerance, and wireless communication.