Key Technologies in 5G: Network Architecture

Saeed Sheikhzadeh, Mohammad Reza Javan *

***Abstract*—To address the continuously increasing demand for high data rates, beside air interface technologies, there are requirements to introduce new and effective ways to design architecture of networks. This paper proposes several technologies which can be used as promising candidates to change the future of wireless communication architecture especially in Fifth generation (5G) wireless network. In this article, we discuss various promising cellular architectures such as full duplex communication, device-to-device communication, mobile femtocell, visible light communication and visualization in 5G.**

***Index Terms*—5G, network architecture, full duplex, device-to-device, mobile femtocell, visible light communication, NFV, SDN, C-RAN.**

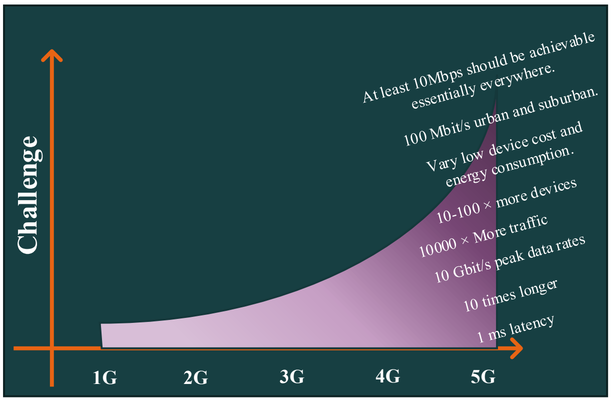


Fig. 1. The challenge of 5G.

I. Introduction

The vendors, operators, and industry partners believe that communications beyond 2020 and 5G technology will be a combination of the existing systems and new revolutionary technologies designed to meet new applications requirements. 5G will be the set of technical components and systems needed to handle these requirements and overcome the limits of current systems [1]. The purpose of 5G is to provide "*Zero latency gigabit experience*" for users/objects of the network [1]. These targets show that an increase of data rates and a reduction of latency are equally important in 5G. The challenge of cellular network and 5G is shown in Fig. 1

The technologies which will be used in 5G cellular network should have some characteristics to overcome the requirements for future demands: [1]:

- Formore capacity, the new 5G system should be designed in a way that enables deployments in new frequency bands.
- 5Gshouldbe designed in such away that is able to handel huge number of devices want to connect to the network.
- The latency and jitter of different services should be considered.
- Forsensors, a battery life of 10 years needs to be achievable. Moreover, for smartphones and tablets, reduced power consumption and increased battery life will be very important. Furthermore, the energy consumption for the operators needs to decrease.
- These technologies should keep equipment cost low as possible.
- Peakdatarates of a 5G system and, more importantly, the data rate of the cell-edge users should meet their targets.
- 5G should provide high data rates to high speed users.
- Sincelocation-basedservices are becoming more important, the accurate positioning shall also be possible with 5G both in indoors and outdoors.
- Securityofsensitive personal data and safety from inserting false information to the system should be ensured in designing of 5G technologies.

The remainder of the paper is organized as follows. In Section II, we describe the Full Duplex (FD) communication. Section III comprises of Device-to-Device (D2D) communications. In Section IV, Mobile Femtocell

Manuscript received December 15, 2016; accepted May 22, 2017
Saeed Sheikhzadeh is with the Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran. E-mail:saeed.sheikhzadeh@gmail.com.
Mohammad R. Javan is with the Department of Electrical and Robotics Engineering, Shahrood University, Shahrood, Iran. E-mail:javan@shahroodut.ac.ir.

(MFemtocell) is presented. Section V discuss about Visible Light Communication (VLC). Section VI gives the detailed description of Software Defined Networking (SDN), Network Function Visualization (NFV) and the Cloud Radio Access Network (C-RAN). Finally, we conclude our paper in Section VII.

II. Full Duplex

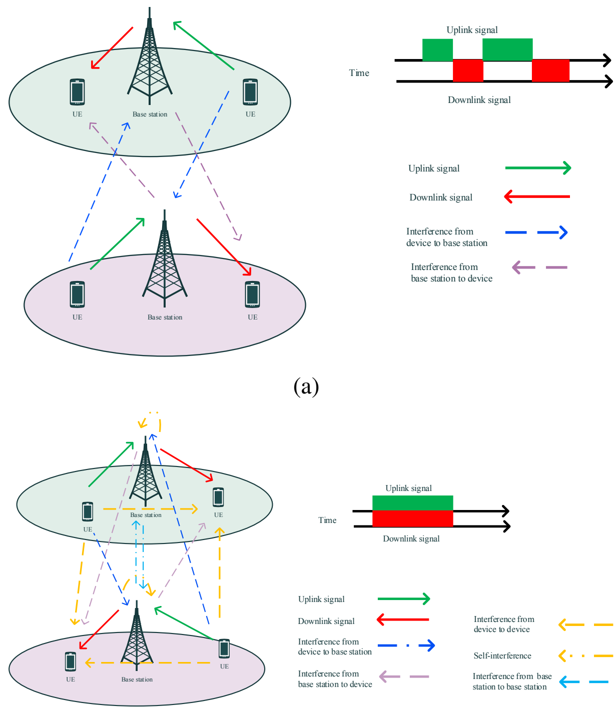
Traditional radio transceivers are generally not able to simultaneously receive and transmit on the same frequency channel because of the crosstalk[1] between the transmitter and the receiver. Most of today's communication systems use time-division duplex (TDD) or frequency-division duplex (FDD) to provide orthogonality between transmit and receive signals, hence they are Half Duplex (HD) communication. The HD scenario is shown in Fig. 2(a).

In FD case, each Base Station (BS) can simultaneously use the same frequency band for scheduling uplink and downlink transmissions [2]. The main driving force behind the increasing interest in FD communications is the promise of nearly doubled channel capacity compared to traditional communication mode, but at the cost of reduced energy efficiency. The difference between FD and HD have been investigated by quantifying the performance gains in terms of either having increased throughput or reduced Outage Probability (OP), albeit achieved at the cost of increased complexity [3].

In some times, FD schemes may not always outperform their HD counterparts, hence in this condition hybrid schemes that switch between HD and FD can also be developed for adaptively using the radio resources and maximizing the spectrum efficiency [4] and [5]. The FD configuration is shown in Fig. 2(b).

Beside the capability of FD to increase spectral efficiency, FD system can decrease the latency within multihop links and avoid the guard period that is necessary when switching the direction of the transmission from downlink to uplink in TDD systems. At the network level, FD system by simultaneously process user plane and control plane signals, can decrease the latency and boost the operation of the overall system [6].

Each transmission potentially experiences higher interference from self-transmission and from neighboring cells compared to the traditional HD cellular systems. In other words, the intended received signal can be more weaker than the transmitted signal and it is very difficult, but not impossible, to detect the received signal under internal interference from the transmission antennas. In



(b)

Fig. 2. HD and FD wireless communications (a) HD scenarios, (b) FD scenarios.

downlink, the additional interference during FD operation comes from all the active uplink User Equipments (UE) on the same channel which we call UE-to-UE interference. In uplink, additional interference comes from: 1) Tx-to-Rx Self-Interference (SI), and 2) neighboring BSs interference due to simultaneous downlink transmission which we call BS-to-BS interference.

From the interference perspective, FD has some differences from HD. We can summarize them as follows [3]:

- BS-to-BS interference in downlink: Usually BSs are installed at higher elevations and have fewer obstructions between each other, and hence, the path loss between BSs is generally much smaller than the path loss between BSs and UEs, especially in an outdoor environment. Therefore, there is very strong interference in FD between BSs. Techniques to mitigate BS-to-BS interference are necessary, especially in outdoor environments. There are some methods to mitigate BS-to-BS interference including the null forming in the elevation angle at the BS antennas and interference management through C-RAN architecture.
- Self interference: There is a major breakthrough in the cancellation circuit design, hence FD operations can be deployment in small cell devices for small

[1]It is also called self interference.

cell deployment, where the smaller coverage area makes it a more suitable environment to deploy FD radios.
- User to user in uplink interference: Since this type of interference depends on the location as well as the transmission power of UEs, an intelligent coordination mechanism is needed. The goal of the coordination is to select those UEs for simultaneous transmission such that their rate/power allocation would create less interference on each other.

Excessive SI may even result in reduced capacity for FD systems that falls below that of HD systems. Compared to the weak received signal of interest, the SI signal can have 60–100 dB higher power. Therefore, the SI signal must be attenuated significantly to allow detection of the actual received signal and enable FD communications in the first place [6]. It is important that, for obtaining the acceptable throughput, the SI level after performing SI suppression be at least 3 dB lower than the noise level. Consensus reached by both industry and academia shows that it is critical to perform efficient SI suppression/cancellation in implementing radical FD communication systems [5] and [7]. There are two methods for omitting the self-interference: 1) passive isolation [8], and 2) active cancellation [9].

To decrease the complexity of active cancellation techniques and circles, the passive isolation between the transmitter and receiver can be used. This leads to the power of the SI leaking to the receiver becomes small and no complicated active cancellation is needed. When the transmit and receive antennas are separated, by increasing the spacing between the antennas or using different polarizations, the electromagnetic isolation between the antennas can be improved in a straightforward manner and resulting in lower SI power. Due to these techniques are passive in nature, they require no tracking of the SI signal and its possible distortion while still can provide a significantly decrease in the amount of SI. We can mention some passive isolation methods such as the use of band-gap structures as high-impedance surfaces and inductive loops, the port-to-port isolation between the antenna feeds, connecting lumped elements between antenna feeds and resonant structures like wavetraps [6].

Active SI cancellation is used to help passive antenna based techniques to mitigate the SI effects. Active cancelation acts in two steps: first at the input of the receiver chain, and then after the Analog-to-Digital Conversion (ADC) [6]. The first step is required in order to prevent the complete saturation of the receiver components and ADC. The second step is performed to attenuate the remaining SI signal below the noise floor.

Theoretically, an FD system has an infinite dynamic range and perfect channel estimation can perfectly eliminate the SI signal. However in practice, due to the hardware limitations such as transmit/receive signal quantization, nonlinearities, and In-phase and Quadrature (I/Q) mismatch, system is unable to reach the theoretical performance [5].

FD is very useful and applicable for cognitive radio networks. The secondary users sense the spectrum and transmit whenever the spectrum holes are detected continuously which is very critical. Cognitive radio networks can use the potential to achieve simultaneous sensing and transmission with use of FD systems. Secondary users with FD radios can sense the target spectrum band in each time slot and determine if the primary users use the spectrum or not, and at the same time, can transmit data or decide to keep silent based on the sensing results [10].

III. DEVICE-TO-DEVICE COMMUNICATIONS

Generally, D2D communication is a type of communications in which devices communicate to each other without any communications infrastructures. D2D communication introduces new usage models based on the adjacency of users including social networking applications, peer-to-peer content sharing, and public safety communications in the absence of network coverage. Since D2D network realize ad hoc mesh network, hence, can be of critical use in natural disasters. In a critical situation, an urgent communication network can be set up using D2D functionality in a short time, replacing the damaged communication network and Internet infrastructure. In the traditional cellular networks, a central entity is responsible for coordinating the communications between the network devices through the BS meaning that no direct communications between devices is allowed. However, emerging context-aware application which needs location discovery and communications between neighboring devices makes D2D communications necessary for future wireless networks as well as 5G. In D2D communications, two neighboring devices communicate with each other over the cellular bandwidth without or with little BS coordination [11].

D2D functionality will increase the coverage, offload backhaul, provide the fallback connectivity, and increase the spectrum utilization, capacity per area, and battery life. One of the advantages of D2D functionality is to reduce some load of the network in a local area such as a stadium or a big mall by allowing direct transmission among cell phones and other devices. The load balancing and load management can be optimized by network and device pro-active caching of common information and offloading the devices to establish direct links. Indeed, D2D communication has been used in the ad-hoc and personal area networking technologies in unlicensed spectrum bands e.g., Industrial, Scientific, and Medical (ISM) bands. Although this type of communication requires less control, it has certain shortage such as limited content sharing, no point-to-multipoint links, synchronization issues, authentication, interference issue, and security concerns [12]. The spectral efficiency, throughput per area, energy efficiency, battery life, and latency can be achieved by reducing the distance between nodes in D2D communication. The coverage can be enhanced by other devices cooperation which can be the only communication in case of poor or no coverage, coverage holes, and emergency situation. Indeed, when the distance between the transmitter and the receiver is large, the channel attenuation will be also high leading to poor received signal. Cooperative communication is a promising technology which can be used to improve the channel quality. In this type of communications, a third node called relay station is used to help the communications quality improves. One solution is to instal relay stations which could be expensive. However, with D2D functionality, some (idle) users could take the responsibility of relaying the information. Hence, full potential of cooperation can be realized only through the implementation of device relaying. Generally, D2D communications are commonly used in:

- Safety applications and disaster scenarios.
- Novel commercial Proximity Services (ProSe) scenarios.
- Network traffic offloading.
- Industrial automation and machine-to-machine communication.

In 5G, D2D communication is considered as another tier where the set of devices cooperate with each other to dramatically increase the network capacity by either reusing the same spectrum as the macro cell or by using unlicensed spectrum. In the D2D communications, the operator might have different levels of control (full, partial, or no control) over the resource allocation among the source, the destination, and the relaying devices. Therefore, the D2D communications is classified into four main types as follows [11]:

- Devices communicate directly with each other and the operator acts as supervisor (Fig. 3(c)).
- Devices communicate with the help of other device as a relay and the operator acts as supervisor to establish and control the links (Fig. 3(a) and Fig. 3(b)).
- Devices communicate directly with each other and the operator is not involved into the process of establishing and controlling the links and the devices are responsible for this task (Fig. 4(b).
- Devices communicate with the help of other device as a relay and the operator is not involved into the process of establishing and controlling the links and devices are responsible for this task (Fig. 4(a)).

In the two first cases, the resource allocation is performed by the BS. Therefore, the BS can handle the problem of the interference management using centralized methods. However, in the two last cases, the resource allocation between devices is done without centralized entity to supervise. Operating in the same licensed band, D2D communication has an impact on other existed users which use licensed band. Hence to reduce impact on the performance of other users, the smart interference management strategies and appropriate resource allocation schemes needs to be considered. The ad-hoc mode of D2D communication in the licensed spectrum (two last cases) offers limited applications similar to the unlicensed counterpart. However, when the network infrastructure acts as the supervisor, D2D communication in the licensed band has many applications such as social networking, video sharing, mobile relaying, and gaming, as well as many benefits such as traffic offloading, capacity enhancement (frequency reuse), extended cellular coverage, improved energy efficient communication [12].

There are two access schemes in D2D communication, closed access, and open access [13]. In closed access, the device which works as relay, has a trust list containing the name of the users which can connect. The list of trusted devices can contain the users in a neighborhood or workplace that know each other, or the users that have been authenticated via a trusted party such as an organization. In open access scheme, each device can act as a relay for other devices without any restrictions. In this case, the security is a challenging issue and must be considered carefully.

There are three main options for the integration of D2D into the cellular networks [14]:

- D2D communication can be with network-controlled or not: In network-controlled approach, infrastructure nodes (BSs and control entities) play a central role in establishing, arbitrating, and managing D2D connections. The infrastructure performs the fundamental tasks such as spectrum management, security, information brokering, and mobility management.
- D2D can work in both the in-band and out-of-band fashions: In-band refers to the D2D communication type whose traffic uses the same licensed frequen-

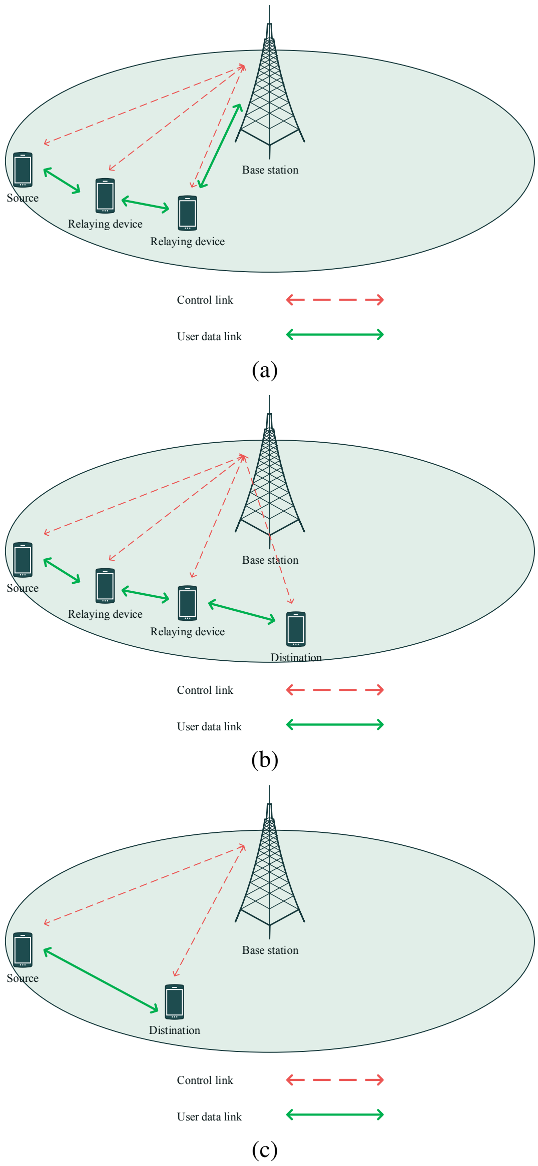


Fig. 3. D2D communications (a) device relaying communication with operator controlled link establishment for transmitting the data from BS, (b) Device relaying communication with operator controlled link establishment for transmitting the data from other user, (c) direct D2D communication with operator controlled link establishment [11].

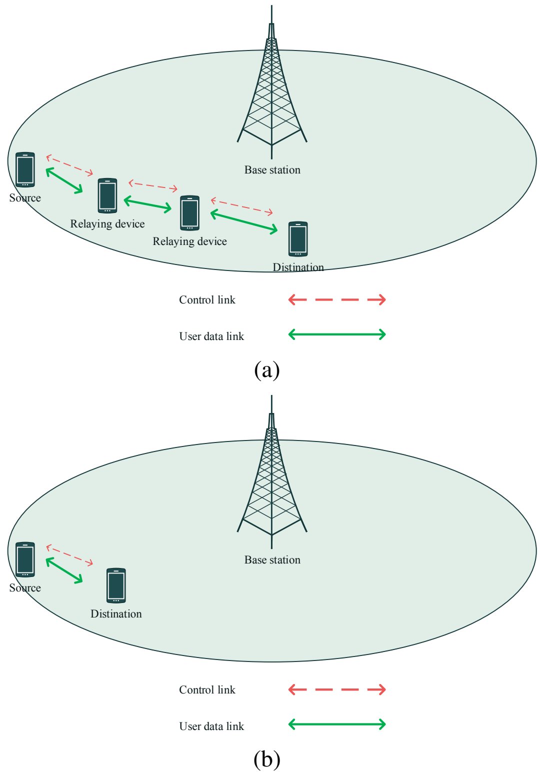


Fig. 4. (a) device relaying communication with device controlled link establishment (b) direct D2D communication with device controlled link establishment.

cies as those used by the ordinary D2D traffic. The main benefit of this approach is the higher degree of control that the operators retain on who transmits and how which could limits the interference. Cooperation among users in this approach becomes more easy to enforce and check. Furthermore, in the view of terminals, there is no need to carry additional radio interfaces. The opposite approach, i.e., the out-of-band D2D, is used to offload cellular networks through other networks.

- D2D can work in overlay or underlay fashion: Overlay fashion refers to the fact there is no part of the spectrum specifically reserved to that D2D communication and it same as primary users in cognitive radio. When networks operate in underlay fashion, D2D transfers and traditional cellular communications share the same radio resources and are scheduled within the cellular bands in an opportunistic fashion same as secondary users in cognitive radio.

IV. Mobile Femtocell

The use of femtocells has become a promising way to improve the throughput of the wireless communication networks due to the ability to absorb the indoor traffic and reduce the originating traffic from the outdoor BSs. The deployment of Femtocell Access Points (FAPs) in macrocells' area has several benefits for operators and subscribers who access to the FAPs for getting service. From the operators' point of view, the advantages of FAPs' deployment are reducing the network deployment cost, increasing the spectrum efficiency, and offloading the data traffic originating from the existing macrocell in the indoor environments. From the subscribers' perspective, they can achieve a high Signal to Interference Plus Noise Ratio (SINR) leading to substantially improved indoor coverage and high throughput [15] and [16].

Traditional BSs such as macrocells, microcells, picocells, and femtocells are designed for low speed users. In this configuration, the only way for the high speed users to use the cellular network is to connect to macrocells. Public vehicles such as trains and buses are moving hotspots with many people potentially requesting diverse data services, e.g. web browsing, video streaming, and gaming. However, these kind of users suffer from higher outage probability, lower throughput, and higher probability of occurrence of call-drops due to the poor signal quality inside the vehicle. The users inside a moving vehicle may execute multiple handovers which, in turn, may cause a significant increase in signaling load and call-drops in the network. Furthermore, due to the high penetration loss through the vehicle's metallic enclosure, the user suffers poor network connection. To overcome these shortages of the cellular networks for high and medium speed users, MFemtocell is introduced [17]–[19]. MFemtocell is a kind of base station which is deployed in the vehicular environment and UE is connected to MFemtocell instead of outside macrocellular. In other words, with combining the concept of moving network and Femtocell network, MFemtocells can move and dynamically change its connection to the operator's core network. The MFemtocell can be deployed on public transport buses, trains, and private cars. To establish connection and transmit/receive data to/from backhaul network, MFemtocells use strong antenna which be installed on the roof or outside of the vehicle. The outside antenna is connected to the inside femtocell by wire. Therefore, the UE which is connected to MFemtocell can archive strong signal and high data rate. Therefore, using the MFemtocell in 5G can improve high speed users throughput and reduce the outage probability. In the MFemtocell, the device and its associated users are all viewed as a single unit to the connected base station. The MFemtocell configuration is shown in Fig. 5.

There are some potentially benefits in implementation of MFemtocell listed as follows [18]:

- Improving the spectral efficiency of the entire network.
- Signaling overhead reduction in the network for improving the system performance.
- Performing a handover as a representative of all its associated users. This will reduce the handover attempts as the users move between cells in the network.
- Due to a relatively shorter range of communication with their serving MFemtocell, the battery life of the associated users can be prolonged.

Moreover, there are several challenges in deployment of mobile femotcells [19]:

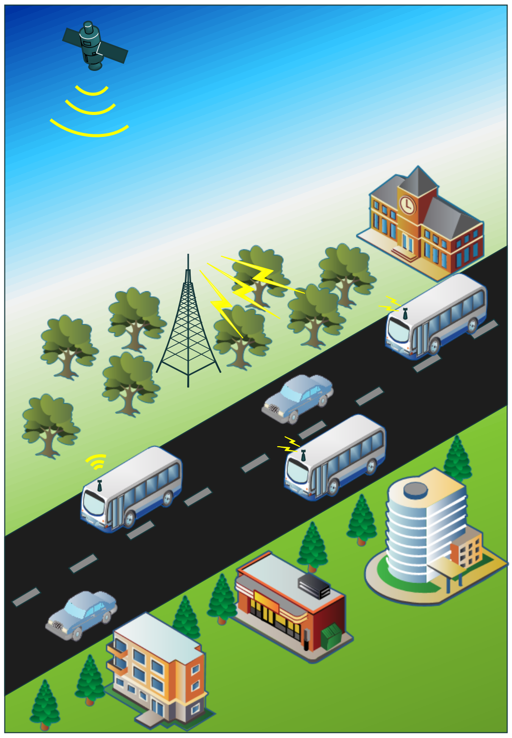


Fig. 5. The scenarios of MFemtocell.

- Finding the most reliable backhaul to carry the traffic between the core network and MFemtocells' users.
- The spectrum sharing problem between the macro BS and MFemtocell.

MFemtocell can connect to different kinds of backhauls depending on the the speed of the vehicles and the availability of the wireless backhaul networks. Various configurations are classified as follows [17]:

- Slow and medium vehicular speeds in the city: In this situation, usually the cellular network is available and the MFemtocell can connect to the cellular network via the macrowave link, the standardized cellular air interface, or using WiFi as radio access technology.
- Slow and medium vehicular speeds out of the city: In this situation, vehicles such as buses, cars, or trains move around the country. In this condition, there is some possibility that network coverage is not available. Hence, the femtocell network uses the satellite networks to connect to the core of the cellular network.

- High vehicular speeds: For this case, due to frequently network handover, connecting to the macrocell is not an effective way for backhauling the MFemtocell traffic. Therefore, the satellite network is the best possible solution.
- In ship and airplane: In this situation the only solution is satellite.

V. Visible Light Communication

The increased wireless data traffic is creating pressure on the dwindling Radio Frequency (RF) spectrum. This pressure drives the need for alternative technologies. The VLC also known as Li-Fi or Optical Wireless Communication (OWC) [20], is a new technology which is introduced to overcome the concern about spectrum scarcity. VLC generally uses fast switching Light Emitting Diodes (LEDs) as its source. On the receiver side, photodiodes is used to convert the optical signals to electrical signals. These sources have the ability to provide illumination and communication for short range indoor links simultaneously [21]. This means that VLC enables systems that illuminate and at the same time provide broadband wireless data connectivity.

VLC systems rely on two disjoint industries: the broadband wireless communications industry and the commercial lighting industry. Since in practice, VLC devices can become marketable, they must have ability to provide high quality lighting and reliable data communication [22]. On the other hand, due to good light quality, low energy consumption, small size, and long lifetime, LED-based illumination devices are a significant substitution to traditional incandescent-based and fluorescent-based illumination devices in the near future and they will be deployed in many countries around the world as indoor illumination, display devices, and traffic lights. At the present time, with huge growth in communication and illumination technology, the notable properties of LED make it particularly proper for wireless communication. On the transmitter side, to achieve high data rate, the signal by on-off keying modulation is modulated. On the receiver side, the photodiodes are used to convert the optical signals to electrical signals at very high rates. In sum, it is appropriate combining illumination devices with wireless communication devices to achieve both green communication and energy-saving illumination. Since LED is safe for eyes and inexpensive, it is more suitable for indoor applications than laser. However the available bandwidth is less than that of laser, which can achieve a data rate up to Gbps. Unlike the highly focused laser, LED is a diffuse source that can provide adequate coverage in a room by placing lots of small LED elements on a panel. Furthermore, the direction and placement of LED elements are determined the efficiency of illumination and communication [20]. An optical attocell covers an area of 1–10m^2 and distances of about 3 m [22].

By combining the data on the visible LED's light way above the human eye's fusion rate, illumination as well as communication can be realized simultaneously. Recently, due to the fact that the visible light spectrum offers a lot of untamed free bandwidth, VLC has engrossed much attention in the wireless communication systems. Moreover, VLC can be used to provide ultra-fast speed wireless communications. Although VLC may not be good solution for large cellular wireless communication, it can be considered as supplementary used for small cell wireless coverage as hot spots under large heterogeneous networks. VLC was originally developed for indoor last mile wireless service delivery. Moreover, it is hard to be used for mobile communications due to the narrow beamwidth in light waves. With using the visible light spectrum, VLC can easily reach up to several Gbps and can have important role in 5G technology [23]. Sophisticated modulation schemes, such as the Optical Orthogonal Frequency Division Modulation (OOFDM) and simple On-Off Keying (OOK) and Pulse Position Modulation (PPM) can be used to increased the data rate [24]. Moreover, Multiple-Input Multiple-Output (MIMO) technology and Red-Green-Blue (RGB)-LED with Wavelength Division Multiplexing (WDM) are exploited to achive ambitiously Gbps transmissions data rate. VLC uses unregulated spectrum, specifically from 428 to 750 THz, which provides huge communication bandwidth to deliver services such as large files and super high definition video transfer [20]. Moreover, the ultra wide bandwidth is also more robust to multipath fading in various environments. It has to be noted that VLC is not subject to fast fading effects as the wavelength is significantly smaller than the detector area [22]. The scenario of VLC is shown in Fig. 6

VLC as wireless communication technology has several benefits such as [21], [24] and [20]:

- Safety, due to the light is restricted to an area surrounded by opaque objects such as walls and luggages which light unable to pass. In other words, the coverage areas of a VLC have applicable security.
- Eco-friendly and energy efficiency.
- Easily deployment of VLC infrastructure by adding a few cheap front-end components to the existing lighting infrastructure.
- Availablity of vast bandwidths.
- The lightening infrastructures are presence everywhere.
- There is no interference between the indoor user

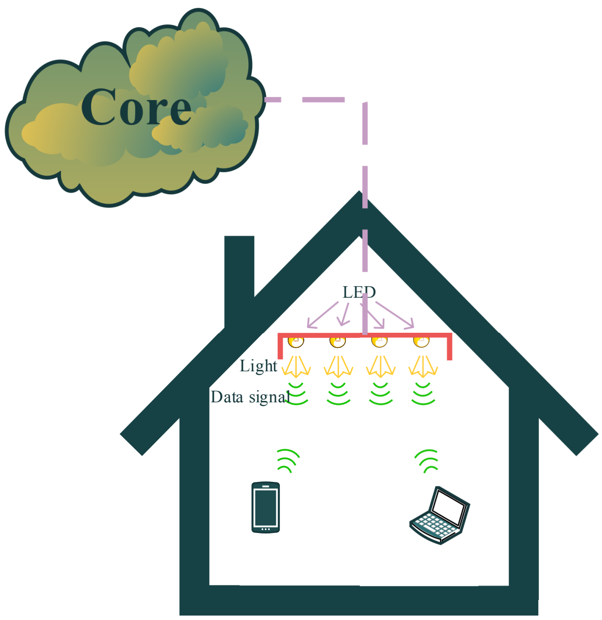


Fig. 6. VLC scenario.

which use VLC and the outdoor user which use RF at all due to different spectra.

- Because there is no interference, RF base station can transmit with low power.
- The resources are used most efficiently.
- High spatial reuse.

Moreover, there are also some disadvantages and chalenges for VLC [24] and [20]:

- A confined coverage,
- Sensitivity to Line-of-Sight (LoS), mobility, blocking and sun-light,
- The lack of uplink support.
- Hazards to human eyes.

VI. Visualization in 5G

Introducing the new service in current networks is becoming extremely difficult. The nature of existing hardware, changing the exist network configuration and integrate the separate parts to each other are the main challenge. However, new emerging technologies, such as NFV, SDN and C-RAN, are promising solutions to this end [25].

A. Software Defined Networking

The paths for data flows are determined, configured, and programmed by the control plane. Data forwarding paths at the hardware level to destination is based on this control information. In traditional networks the control and data planes are combined in a network node [26].

A SDN has been proposed to offer scalable and flexible management with a logical centralized control model [27]. In other words, in SDN, there is a central software program (controller) which dictates the overall network behavior. There are many advantages in decoupling which enables both planes to evolve independently, for example, high flexibility, being vendor-agnostic, programmability, and the possibility of realizing a centralized network view [28]. In this architecture, the network devices change to simple packet forwarding devices (data plane) and the control logic is implemented in the controller (control plane) [29]. The introduction of new ideas in this new paradigm is much easier, due to simplify the change and manipulate of the network compares to last method. Moreover, in SDN, the network configuration is centralized and in this arrangement operators do not have to configure all network devices individually to make changes in network behavior, instead, with global knowledge of the network state, make network traffic forwarding decisions in controller. In other words, SDN is defined in [26] and [30] as: "*In the SDN architecture, the control and data planes are decoupled, network intelligence and state are logically centralized, and the underlying network infrastructure is abstracted from the applications*". SDN provides the global network view which causes to dynamic topology control (i.e., adjusting switch usage depend on load and traffic mapping) [26].

Another feature of SDN is the network programmability which allows seamless communication at all levels, from hardware to software and ultimately to end users. This characteristic makes applications and network aware from each other which causes to efficiently use of resources and opens up the potential for new applications [26]. Moreover, SDNs provide the possibility of control the behavior of the network directly, by configuring the packet forwarding rules installed on each switch for programmers [31].

There are four main features which SDN focus on them [26]:

- Separation the control plane and the data plane from each other,
- A centralized controller and view of the network,
- Open interfaces between the devices in the control plane (controllers) and those in the data plane
- Programmability of the network by external applications

In [26], SDN functional architecture is divided into three part which shown in Fig. 7:

- The physical network equipment such as Ethernet switches and routers are located in the bottom layer which forms the data plane,

- The controllers are located in the central layer which facilitate setting up and tearing down flows and paths in the network,
- The bottom layer or application contains such as energy-efficient networking, security monitoring, and access control for operation and management of the network. An application refers to a service provided by the network operator.

The central layer linked with the bottom layer and upper layer by an Application Programming Interface (API) which called as the southbound API and northbound API, respectively.

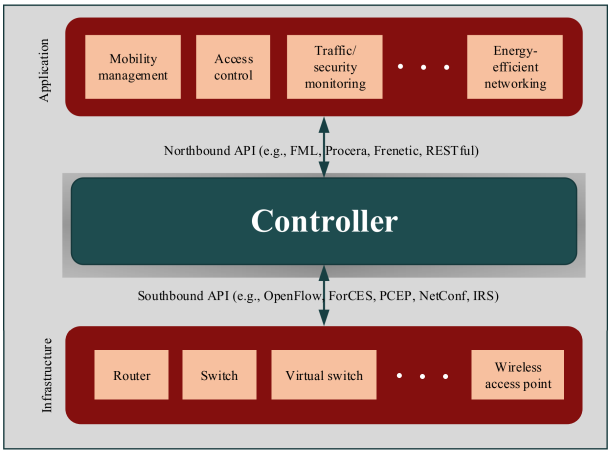


Fig. 7. SDN functional architecture [26].

SDNs have several functionalities which can support multiple tasks at the same time such as [31]:

- Implementation shortest-path routing by calculating the forwarding rules for each switch by running Dijkstra's algorithm on the graph of the network topology by the controller,
- The controller can selectively shut down links or even whole switches after directing traffic along other paths for energy efficiency propose.
- Balancing the load between back-end servers in a data center, by splitting flows over several server replicas and migrate flows to new paths by the controller,
- And etc.

B. *Network function virtualization*

The European Telecommunications Standard Institute (ETSI) proposes the following definition for NFV [32]:

"*Network Functions Virtualisation aims to transform the way that network operators architect networks by evolving standard IT virtualisation technology to consolidate many network equipment types onto industry standard high volume servers, switches and storage, which could be located in Datacentres, Network Nodes and in the end user premises.*"

On the other hands, base on ETSI, we can define the concept of NFV as decouple the network function from the hardware and these functions are implemented using software. NFV uses standard computing virtualization technology to consolidate in commodity hardware. This fuctions previously performed by specific hardware appliances. Moreover, NFV uses virtualization and cloud computing in telecommunication networks. The network functions are virtualized by NFV as Virtualized Network Function (VNFs) and these network functions are used to create networking services by interconnection [33]. VNFs are the basic elements to provide the complete virtualization of service delivery [34]. On other hand, the plan of network operators is migration from legacy standalone hardware appliances to an architecture based on VNFs. NFV is applicable to any part of network (data plane packet processing and control plane function) and any kind of networks technology (mobile and fixed networks). Potential examples include switching elements, mobile network nodes, tunnelling gateway elements, converged, and network-wide functions, application-level optimization, security functions, etc. [32].

Indeed, NFV uses virtialization technology to separate software instance from hardware platform and helps to faster networking service provisioning . In other words, NFV changes implementation of network functions by performing software virtialization techniques and runs them on commodity hardware [25]. On the other hands, NFV combines many different functional modules such as L2 switch, L3 router, application delivery controller to obtain cost effectively and acceptable performance. The most common example of this technology is to run an open source software based firewall in a virtual machine.

There are several benefits which NFV can potentially bring [25] and [32]:

- Reduce capital investment,
- Reduce energy consumption,
- Reduce Time to Market for introduction and deployment of new service,
- Allow network appliance multi-version and multi-tenancy,
- Allow to introduction targeted service based on geography or customer sets,
- Introducing targeted and tailored services based on customer needs,
- and etc.

However, beside these benefits, there are several challenges which network operators face when deploying virtual appliances [25] and [32]:

- It is shown that latency variate abnormally and throughput is instable even when the underlying network is only lightly utilized,
- The another problem is how to migrate from the existing network infrastructure to NFV-based solutions,
- How to efficiently place the virtual appliances and dynamically instantiate them on demand are come from separation of functionality from location,
- Managing and using many virtual network appliances while ensuring security issue,
- Ensuring the appropriate level of resilience to hardware and software failures,
- and etc.

On the implementation of NFV, it is possible that the capacity of VNF be less than the corresponding physical version on dedicated hardware. However, this degradation should keep as small as possible. On the other hand, with holding the latency requirement, the algorithms of splitting of load should be designed [25].

In virtual network structure, NFV architecture should be flexible to be able use VNFs in the proper situations and time, dynamically allocate and scale hardware resources for them, and interconnect them to achieve service chaining [25]. With using NFV, the implementation of network functions is done by software which can run on a range of industry standard server hardware, and can be moved or instantiated in various location in the network without the need for installation of new equipments.

NFV and SDN can work together and complementary each other but not depend to each other. NFV goals can be achieved without requiring SDN, however, the two concepts and solutions can be combined and potentially greater value accrued.

C. Cloud Radio Access Network

Mobile data transmission volume is continuously rising. It is forecasted to grow 13-fold from 2012 until 2017 according to Cisco [35] with smart phones and tablet users driving the growth. Therefore, to satisfy the growing user demands, mobile network operators have to increase the network capacity. C-RAN is a novel mobile network architecture which has the potential to solve most of the telecommunication's challenges. In C-RAN, baseband resources are pooled so that they can be shared between BSs. It is able to adapt to non-uniform traffic and utilizes the resources of the network, such as BSs and frequency, more efficiently. Moreover, C-RAN is seen as a typical realization of the mobile networks supporting soft and green technologies in 5G mobile networks in year 2020 horizon [35].

There are two kinds of functionality associated with BSs: baseband processing and radio functionalities. The main tasks of baseband processing module are coding, modulation, and FFT. Moreover, the radio module is responsible for digital processing, frequency filtering, and power amplification. The base station can be classified into three generations.

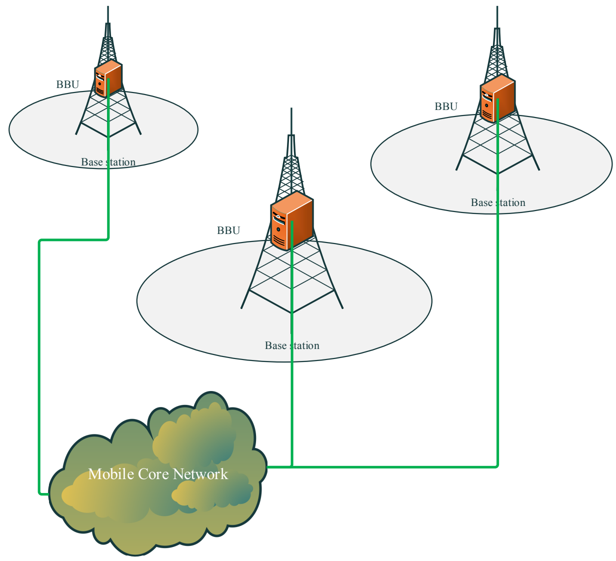


Fig. 8. The first generation of BS.

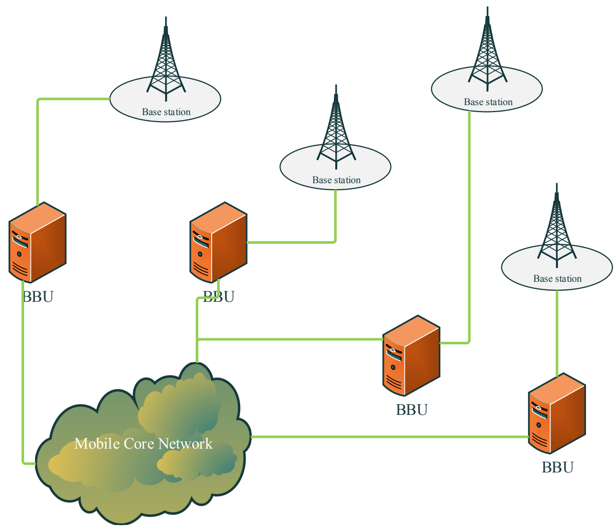


Fig. 9. The second generation of BS.

The first generation is deployed in 1G and 2G mobile networks. In this generation, the radio and baseband processing functionalities are integrated inside the BS and the antenna module is close to the radio module. This kind of architecture is shown in Fig. 8.

3G and 4G networks use the second generation of BSs. In this generation, the radio unit is called a Remote

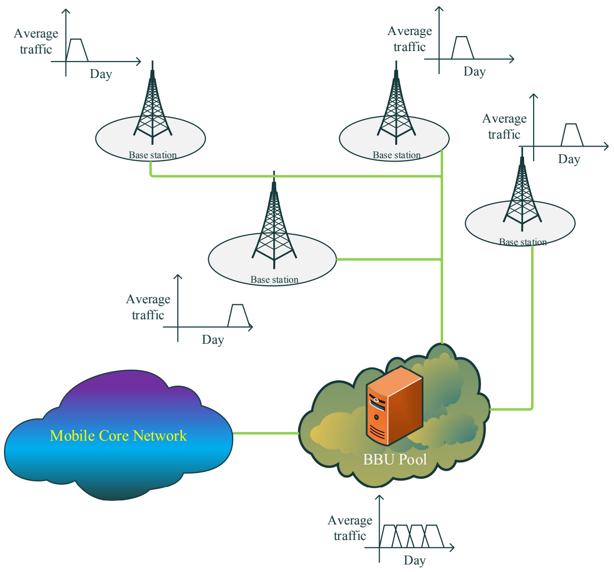


Fig. 10. The third generation of BS.

Radio Head (RRH) or Remote Radio Unit (RRU). The main sub-functions of RRH are as follows:

- Digital processing,
- Digital to analog conversion,
- Analog to digital conversion,
- Power amplification and filtering.

The baseband processing module is called Base Band Unit (BBU) or Data Unit (DU). In this generation, RRH can connect to BBU by optical fiber and microwave which the distance of them can be extended up to 40 km. This solutions have so many benefits to operators of network such as, conveniently and easily access to BBU, reducing and saving the site rental, serving many RRH by only one BBU and etc. Moreover RRH can be installed on the towers or roofs. The architecture of this generation is shown in Fig. 9.

The third generation of BSs is called C-RAN which will be used in 5G. This generation is totally different from the previous generations. In this generation, in order to improve the efficiency of BBU and balance the load between the BSs with heavy and light load, all of the BBUs are gathered into one entity called the BBU-Pool. BBU-Pool is connected to all BSs. A BBU-Pool is a virtualized cluster which can consist of general purpose processors to perform baseband processing. The architecture of this generation is shown in Fig. 10.

The utilization of C-RAN has many advantage such as [35]:

- C-RAN has the potential to decrease the cost, power, and energy consumption of the network operation: Since the number of BBUs which are needed in C-RAN compared to the traditional architecture is reduced, the expense, the power, and the energy consumption are reduced compared to the traditional Radio Access Network (RAN) architecture. Moreover, virtualized BBU-Pool of one operator can be used by different network operators allowing them to rent RAN as a cloud service meaning that the network deployment cost is reduced.
- Improving scalability and easing network maintenance: BBUs which are used in C-RAN can be deployed and upgraded in an easy way. Hence, the network maintenance and deployment become artless.
- Facilitating the deployment of the mechanisms such as enhanced Inter-Cell Interference Coordination (eICIC) and Coordinated Multi-Point (CoMP) to increase the spectral efficiency and throughput: As the BBUs from many sites are co-located in one pool, they can interact with lower delays, and therefore, these mechanisms can be developed easily. By reducing the delay, the network performance is improved, and methods for performing the load balancing between the cells can be easily used. Furthermore, advanced features of 3GPP Long Term Evolution-Advance (LTE-A) such as CoMP and interference mitigation can be efficiently supported by C-RAN which is essential especially for small cell deployments.
- Balancing nonuniform traffic: the network load varies throughout the day. It is because the subscribers are moving between different areas during the day from home to work or vice versa. Hence, the load of the network rises during the day in downtown and rises during the night in suburb.

VII. Conclusion

In this article, the network architectures that can be used in 5G cellular architecture have been discussed. Here, we have focused on full duplex, device-to-device communications, mobile femtocell, visible light communication, and visualization in 5G. These technologies can lead to fundamental changes in the design of cellular networks.

References

[1] Nokia, "5G use cases and requirements," White Paper.

[2] L. Wang, F. Tian, T. Svensson, D. Feng, M. Song, and S. Li, "Exploiting full duplex for device-to-device communications in heterogeneous networks," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 146–152, May 2015.

[3] S. Goyal, P. Liu, S. Panwar, R. Difazio, R. Yang, and E. Bala, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Communications Magazine*, vol. 53, no. 5, pp. 121–127, May 2015.

[4] K. Yamamoto, K. Haneda, H. Murata, and S. Yoshida, "Optimal transmission scheduling for a hybrid of full-and half-duplex relaying," *IEEE Communications Letters*, vol. 15, no. 3, pp. 305–307, March 2011.

[5] Z. Zhang, X. Chai, K. Long, A. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 128–137, May 2015.

[6] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, May 2015.

[7] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept 2014.

[8] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 680–694, February 2014.

[9] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4494–4510, Nov 2013.

[10] Y. Liao, L. Song, Z. Han, and Y. Li, "Full duplex cognitive radio: a new design paradigm for enhancing spectrum usage," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 138–145, May 2015.

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

[12] H. Mustafa, M. Imran, M. Shakir, A. Imran, and R. Tafazolli, "Separation framework: An enabler for cooperative and D2D communication for future 5G networks," *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, pp. 1–1, 2015.

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

[14] F. Malandrino, C. Casetti, and C.-F. Chiasserini, "Toward D2D-enhanced heterogeneous networks," *IEEE Communications Magazine*, vol. 52, no. 11, pp. 94–100, Nov 2014.

[15] H.-S. Jo, C. Mun, J. Moon, and J.-G. Yook, "Self-optimized coverage coordination in femtocell networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 10, pp. 2977–2982, October 2010.

[16] L. Huang, G. Zhu, and X. Du, "Cognitive femtocell networks: an opportunistic spectrum access for future indoor wireless coverage," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 44–51, April 2013.

[17] M. Chowdhury, S. Q. Lee, B. H. Ru, N. Park, and Y. M. Jang, "Service quality improvement of mobile users in vehicular environment by mobile femtocell network deployment," in *2011 International Conference on ICT Convergence (ICTC)*, Sept 2011, pp. 194–198.

[18] F. Haider, H. Wang, H. Haas, D. Yuan, H. Wang, X. Gao, X.-H. You, and E. Hepsaydir, "Spectral efficiency analysis of mobile femtocell based cellular systems," in *2011 IEEE 13th International Conference on Communication Technology (ICCT)*, Sept 2011, pp. 347–351.

[19] F. Haider, C. Wang, B. Ai, H. Haas, and E. Hepsaydir, "Spectral-energy efficiency trade-off of cellular systems with mobile femtocell deployment," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2015.

[20] S. Wu, H. Wang, and C.-H. Youn, "Visible light communications for 5G wireless networking systems: from fixed to mobile communications," *IEEE Network,*, vol. 28, no. 6, pp. 41–45, Nov 2014.

[21] F. Zafar, D. Karunatilaka, and R. Parthiban, "Dimming schemes for visible light communication: the state of research," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 29–35, April 2015.

[22] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, February 2014.

[23] H.-H. Chen, "Feature topic: Visible light communications [message from the editor-in-chief]," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 2–3, April 2015.

[24] R. Zhang, J. Wang, Z. Wang, Z. Xu, C. Zhao, and L. Hanzo, "Visible light communications in heterogeneous networks: Paving the way for user-centric design," *IEEE Wireless Communications*, vol. 22, no. 2, pp. 8–16, April 2015.

[25] B. Han, V. Gopalakrishnan, L. Ji, and S. Lee, "Network function virtualization: Challenges and opportunities for innovations," *IEEE Communications Magazine*, vol. 53, no. 2, pp. 90–97, Feb 2015.

[26] S. Sezer, S. Scott-Hayward, P. Chouhan, B. Fraser, D. Lake, J. Finnegan, N. Viljoen, M. Miller, and N. Rao, "Are we ready for SDN? implementation challenges for software-defined networks," *IEEE Communications Magazine*, vol. 51, no. 7, pp. 36–43, July 2013.

[27] M. Dong, H. Li, K. Ota, and J. Xiao, "Rule caching in SDN-enabled mobile access networks," *IEEE Network*, vol. 29, no. 4, pp. 40–45, July 2015.

[28] S. Yeganeh, A. Tootoonchian, and Y. Ganjali, "On scalability of software-defined networking," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 136–141, February 2013.

[29] H. Kim and N. Feamster, "Improving network management with software defined networking," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 114–119, February 2013.

[30] ONF, "Software-defined networking: The new norm for networks,," White Paper. [Online]. Available: https://www.opennetworking.org

[31] N. Foster, A. Guha, M. Reitblatt, A. Story, M. Freedman, N. Katta, C. Monsanto, J. Reich, J. Rexford, C. Schlesinger, D. Walker, and R. Harrison, "Languages for software-defined networks," *IEEE Communications Magazine*, vol. 51, no. 2, pp. 128–134, February 2013.

[32] ETSI, "Network functions virtualisation: An introduction, benefits, enablers, challenges & call for action," October 2012, White Paper. [Online]. Available: http://portal.etsi.org/NFV/NFV_White_Paper.pdf

[33] A. Bradai, K. Singh, T. Ahmed, and T. Rasheed, "Cellular software defined networking: a framework," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 36–43, June 2015.

[34] J. Matias, J. Garay, N. Toledo, J. Unzilla, and E. Jacob, "Toward an SDN-enabled NFV architecture," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 187–193, April 2015.

[35] A. Checko, H. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. Berger, and L. Dittmann, "Cloud ran for mobile networks - a technology overview," *IEEE Communications Surveys Tutorials,*, vol. 17, no. 1, pp. 405–426, Firstquarter 2015.