ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΨΗΦΙΑΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ

ΚΑΤΕΥΘΥΝΣΗ Τηλεπικοινωνιακά Δίκτυα και Υπηρεσίες Τηλεματικής

Τίτλος Εργασίας

“OPERATOR SURVEILLANCE AND BILLING IN 5G D2D SYSTEMS”

Όνομα φοιτητή

ΜΠΟΥΤΟΣ ΑΝΑΣΤΑΣΙΟΣ

Αθήνα, 2017

ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΣΧΟΛΗ ΨΗΦΙΑΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΜΑΤΙΚΗΣ

ΚΑΤΕΥΘΥΝΣΗ Τηλεπικοινωνιακά Δίκτυα και Υπηρεσίες Τηλεματικής

Τριμελής Εξεταστική Επιτροπή

Δημητρακόπουλος Γ.

Επίκουρος καθηγητής, Τμήμα Πληροφορικής κ Τηλεματικής, Χαροκόπειο Πανεπιστήμιο

Μιχαλακέλης Χ.

Επίκουρος καθηγητής, Τμήμα Πληροφορικής κ Τηλεματικής, Χαροκόπειο Πανεπιστήμιο

Αναγνωστόπουλος Δ.

Καθηγητής, Τμήμα Πληροφορικής κ Τηλεματικής, Χαροκόπειο Πανεπιστήμιο

Ο Μπούτος Αναστάσιος

δηλώνω υπεύθυνα ότι:

1) Είμαι ο κάτοχος των πνευματικών δικαιωμάτων της πρωτότυπης αυτής εργασίας και από όσο γνωρίζω η εργασία μου δε συκοφαντεί πρόσωπα ούτε προσβάλλει τα πνευματικά δικαιώματα τρίτων.

2) Αποδέχομαι ότι η ΒΚΠ μπορεί, χωρίς να αλλάξει το περιεχόμενο της εργασίας μου, να τη διαθέσει σε ηλεκτρονική μορφή μέσα από την ψηφιακή βιβλιοθήκη της, να την αντιγράψει σε οποιοδήποτε μέσο ή/και σε οποιοδήποτε μορφότυπο, καθώς και να κρατά περισσότερα από ένα αντίγραφα για λόγους συντήρησης και ασφάλειας.

ACKNOWLEDGEMENT

I would like to express my gratitude to my supervisor, Mr. George Dimitrakopoulos, for his useful comments, remarks and engagement throughout the learning process of this Master's thesis.

Furthermore, I would like to thank the other two members of the thesis committee, Mr. Christos Michalakelis and Mr. Dimosthenis Anagnostopoulos, for their effective cooperation and contribution to the completion of this Master's thesis. Also, I would like to thank Harokopio University for giving me the opportunity to participate in this postgraduate program.

Finally, I would like to thank my family for supporting me all these years, and also my partner, Anna Maria, for her encouragement during the whole program.

TABLE OF CONTENTS

Περίληψη στα ελληνικά

Η Device-to-Device (D2D) επικοινωνία φαίνεται να είναι μια καίρια, νέα τεχνολογία που θα μας απασχολήσει στο βραχυπρόθεσμο μέλλον με την έλευση της 5ης γενιάς δικτύων. Μέσα από αυτή την τεχνολογία θα μπορέσουμε να αυξήσουμε την απόδοση του δικτύου, να αυξήσουμε τη χωρητικότητα και να έχουμε καλύτερη αξιοποίηση του φάσματος, και τέλος να μειώσουμε την καθυστέρηση της επικοινωνίας (latency) μεταξύ των χρηστών. Επιπλέον, η τεχνολογία αυτή μπορεί να βρει εφαρμογή και σε άλλα πεδία, όπως στην αποφόρτιση των τηλεπικοινωνιακών δικτύων, στη δημόσια ασφάλεια, στην ομαλή λειτουργία του κυκλοφοριακού για την αποφυγή ατυχημάτων, σε στρατιωτικές εφαρμογές και άλλα. Στο πρώτο κεφάλαιο θα έχουμε μια σύντομη περιγραφή σχετικά με τα τεχνικά χαρακτηριστικά των υπαρχόντων δικτύων, όπως είναι τα 2ης, 3ης και 4ης γενιάς, καθώς και τις διαφορές τους όσον αφορά τον τρόπο που γίνεται η καταγραφή και η χρέωση των υπηρεσιών τους. Στο επόμενο κεφάλαιο θα δούμε μια λεπτομερή ανάλυση των δικτύων 5ης γενιάς, όπως καίριες τεχνολογίες που θα πρέπει να εφαρμοστούν με σκοπό την αύξηση της απόδοσης του δικτύου, της χωρητικότητας και της βέλτιστης αξιοποίησης του φάσματος, και τέλος κάποια σημαντικά ζητήματα σε σχέση με τη σχεδίαση των δικτύων αυτών. Το επόμενο κεφάλαιο, που θα μας απασχολήσει και περισσότερο, μιας και αποτελεί τον πυρήνα της διπλωματικής εργασίας, είναι η παρουσίαση δύο διαφορετικών κατηγοριών δικτύων επικοινωνίας, όπου η πρώτη αφορά την «κλασική» επικοινωνία του κεντρικού σταθμού-κυψέλης με το τερματικό του χρήστη (BS-to-device) και η δεύτερη την απευθείας επικοινωνία των χρηστών μεταξύ τους (device-to-device). Σε αυτό το κεφάλαιο θα δώσουμε βάση κυρίως στη δεύτερη κατηγορία, που είναι η επικοινωνία μεταξύ των χρηστών (device-to-device), και θα παρουσιάσουμε τα τέσσερα διαφορετικά δυνατά σενάρια επικοινωνίας μεταξύ τους, καθώς και τα διαφορετικά συστήματα χρέωσης που μπορούν να εφαρμοστούν πάνω σε αυτά.

Λέξεις-κλειδιά: 5^{ης} γενιάς δίκτυα, device-to-device επικοινωνία, αξιοποίηση φάσματος, καθυστέρηση

Abstract

Direct Device-to-Device (D2D) communication seems to be a promising new technology that can increase spectrum utilization and capacity, enhance network performance and throughput, and finally provide low-latency services between end-users in the longer-term 5G networks. It can also be used in a plethora of different fields, such as network traffic offloading, public safety, social services and applications such as gaming and military applications.

The first chapter will make a short reference to the technical specifications of the existing networks such as 2G, 3G and 4G, and the differences in their charging and billing systems. In the next chapter there will be an analysis of 5G networks, such as key technologies that should be used in order to enhance network performance and capacity and increase spectrum utilization, and finally some critical design issues.

The next chapter, which is also the main part of the thesis, is a presentation of a two-tier cellular network that involves a macrocell tier (BS-to-device communications) and a device tier (device-to-device communications) and the different billing systems that can be applied. In this chapter emphasis will be placed on the device tier (D2D communication), where we will present four different ways of communication: a) the device relaying with operator-controlled link establishment (DR-OC), b) the direct D2D communication with operator-controlled link establishment (DC-OC), c) the device relaying with device-controlled link establishment (DR-DC), and d) the direct D2D communication with device-controlled link establishment (DC-DC).

Device terminal relaying makes it possible for devices in a network to function as transmission relays for each other and realize a massive ad hoc mesh network. This is obviously a dramatic departure from the conventional cellular architecture and brings unique technical challenges. In such a two-tier cellular system, since the user data is routed through other users' devices, security must be maintained for privacy. To ensure minimal impact on the performance of existing macrocell BSs, the two-tier network needs to be designed with smart interference management strategies and appropriate resource allocation schemes. Furthermore, novel-billing models should be designed to tempt users (devices) to participate in this type of communication.

The following chapter will include a comparative analysis of the different ways of D2D communication that were presented in the previous chapter, and their different billing systems.

Finally, in the conclusion there will be a summary of the results of the four different types of device-tier communications and their proposed billing options.

Keywords: D2D, 5G, DR (device relaying), DC (direct communication), OC (operator-controlled), DC (device-controlled)

LIST OF FIGURES

LIST OF ABBREVIATIONS

2G	2nd Generation
3G	3rd Generation
4G	4th Generation
APN	Access Point Name
AuC	Authentication Center
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station
CDMA	Code-Division Multiple Access
CDR	Call Data Record
D2D	Device-to-Device
DC-DC	Direct Communication - Device Controlled
DC-OC	Direct Communication - Operator Controlled
DR-DC	Device Relaying - Device Controlled
DR-OC	Device Relaying - Operator Controlled
DVB	Digital Video Broadcasting
EDGE	Enhanced Data rates for GSM Evolution
EIR	Equipment Identity Register
eNB	evolved Node B
EPC	Evolved Packet Core
FDM	Frequency-Division Multiplexing
FDMA	Frequency-Division Multiple Access
FFT/IFFT	Fast Fourier Transform/Inverse Fast Fourier Transform
GGSN	Gateway GPRS Support Node
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HDTV	High Definition Television
HetNet	Heterogeneous Wireless Networks
HLR	Home Location Register
HSS	Home Subscriber Server
ICI	Inter-Carrier Interference
IEEE	Institute of Electrical and Electronics Engineers
IMEI	International Mobile Equipment Identity
IMS	IP Multimedia Subsystem
IPv6	Internet Protocol version 6

ISDN	Integrated Services Digital Network
ISI	Inter-Symbol Interference
LAS-CDMA	Large Area Synchronized Code-Division Multiple Access
LTE	Long-Term Evolution
MAC	Medium Access Control
MIMO	Multiple Input Multiple Output
MME	Mobile Management Entity
MMS	Multimedia Messaging Service
MS	Mobile Station
MSC	Mobile Switching Center
NFV	Network Function Virtualization
NSS	Networking and Switching System
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OMC	Operations and Maintenance Center
OSS	Operation Subsystem
PCEF	Policy Control Enforcement Function
PCRF	Policy Charging and Rules Functionality
PDN	Public Data Network
PHY	Physical
QoS	Quality of Service
RAND	Random Numbers
RAT	Radio Access Technology
RFID	Radio Frequency Identification
RNC	Radio Network Controller
SDN	Software Defined Networking
SFN	Single Frequency Network
SGSN	Serving GPRS Support Node
SGW	Service Gateways
SIM	Subscriber Identify Module
SIR	Signal-to-Interference Ratio
SMS	Short Message Service
SMSC	Short Message Service Centre
SNR	Signal-to-Noise Ratio
SRES	Signed Response
UDR	Usage Data Record
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunications System
VLR	Visitor Location Register

WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access

1. INTRODUCTION

MOTIVATION

5G mobile and wireless communication systems forecast increasing numbers of mobile devices and growing traffic volumes. Traffic volumes in wireless communication have increased in the past few years, and the growth is expected to continue also in the near future. Traffic volumes beyond the year 2020 can be even 1,000 times higher than traffic volumes nowadays. Future cellular networks will become denser with small cells such as pico-cell and femto-cell for better coverage, capacity and spectral efficiency.

Overall targets for 5G systems are higher throughput per area and per user, and lower latency. The 5G systems will support a huge amount of devices, and with lower energy consumption compared to current systems. D2D communication is seen as one answer to growing demands for future mobile and wireless communication systems. D2D might offer higher data rates and lower latency due to the short distance between the D2D link. D2D communication is also energy efficient as no data communication via Base Station (BS) occurs thus the traffic loads of BS decreases.

D2D communication will be divided in two categories. In the first one the communication will be network-controlled as in a conventional cellular system, and in the second one it will be non-network controlled when out-of-coverage, and D2D will (re)use the same licensed spectrum as the cellular links. D2D communication takes place directly between a D2D pair without a base station controlling the communication. D2D communication can also be used in out-of-coverage areas or when cellular network fails, mainly for public safety purposes.

PURPOSE

As we mentioned before, Direct Device-to-Device (D2D) communication seems to be a promising new technology that can increase spectrum utilization and capacity, enhance network performance, increase throughput, and finally achieve low-latency services between end-users in the longer-term 5G networks. The purpose of this thesis is to present the possible ways of device-to-device (D2D) communication in 5G networks and the different levels of control and billing an operator can define.

STRUCTURE OF THE THESIS

The thesis has been structured in the following chapters that are shortly outlined below:

Chapter 2 provides an overview of the architecture of the existing cellular networks, such as, 2G, 3G and 4G and their billing and charging system.

Chapter 3 includes an analysis of 5G networks. More specific, includes the key technologies beneath 5G networks, such as, cell densification and network offload, usage of mmWave spectrum and finally the use of massive MIMO technology. Furthermore, in this chapter we include design issues of 5G technology like waveforming, cloud-based networks and energy efficiency.

Chapter 4 provides an overview of device-to-device (D2D) communication types and their main technical challenges. The chapter also, describes different types of billing/charging scenarios based on the different way of communication, such as, direct relaying with operator-controlled link establishment (DR-OC), direct D2D communication with operator-controlled link establishment (DC-OC), device relaying with device-controlled link establishment (DR-DC) and direct D2D communication with device-controlled link establishment (DC-DC). Finally, chapter 4 compares the revenue for operators and devices while they participate in a conventional network and in a two-tier network, like the above we refer to before.

Chapter 5 includes the conclusion of 5G D2D communication and their different billing/charging scenarios that we discuss in the last chapter.

2. Operator Billing/Charging in 2G, 3G and 4G networks

Before we discuss the billing and charging system in today's networks, let's make a short reference to the technical specifications of each network.

2G emerged in the late 1980s. It uses digital signals for voice transmission and has a speed of 64 kbps. It was the first generation that provided facility of SMS (Short Message Service) and uses a bandwidth of 30 to 200 KHz. Next to 2G, the 2.5G system uses packet-switched and circuit-switched domains and provides data rate up to 144 kbps, e.g. GPRS, CDMA and EDGE [1].
3G uses Wide Band Wireless Network with which clarity is increased. The data is sent through Packet Switching. Voice calls are interpreted through Circuit Switching. Along with verbal communication, it includes data services, access to streaming and new services such as Global Roaming. It operates at a range of 2,100 MHz and has a bandwidth of 15-20 MHz used for high-speed internet services.

4G offers a downloading speed of 100 Mbps. Moreover, it provides the same features as 3G, and additional services such as Multi-Media, to watch T.V. programs with more clarity and send data much faster than previous generations. LTE (Long Term Evolution) is considered 4G technology [1]. 4G is being developed to accommodate the QoS and rate requirements set by forthcoming applications such as wireless broadband access, Multimedia Messaging Service (MMS), video chat, mobile TV, HDTV content, Digital Video Broadcasting (DVB), minimal services like voice and data, and other services that utilize bandwidth.

2.1. Architecture in 2G technology

2.1.1 GSM Mobile Network Infrastructure

The GSM mobile network is divided into four major systems:
- Mobile Station (MS)
- Base Station Subsystem (BSS)
- Networking and Switching System (NSS)
- Operation Subsystem (OSS)

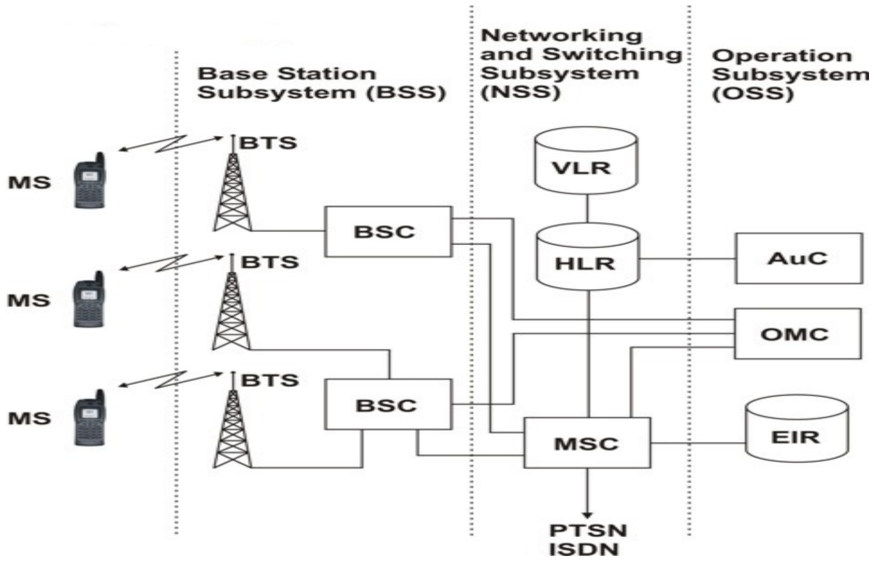

Figure 1 - the major systems of a GSM mobile network

2.1.2 Mobile Station (MS)

It refers to the terminal equipment used by wireless subscribers. It consists of:

- SIM-Subscriber Identity Module
- Mobile Equipment

The SIM is removable and with appropriate SIM, the network can be accessed using various mobile equipment.

The equipment identity is not linked to the subscriber. The equipment is validated separately with IMEI and EIR. An integrated circuit chip with a microprocessor, random-access memory (RAM) and read-only memory (ROM) contained in the SIM. Also, the SIM should be valid and should authenticate the validity of MS while accessing the network.

The SIM also stores subscriber related information such as IMSI, cell location identity, etc.

Functions of Mobile Station (MS):

- Radio transmission and reception
- Radio channel management
- Speech encoding/decoding
- Radio link error protection
- Flow control of data
- Rate adaptation of user data to the radio link
- Mobility management

2.1.3 Base Station Subsystem (BSS)

The BSS connects the MS and the NSS and it is composed of the following:

- Base Transceiver Station (BTS), also called Base Station.
- Base Station Controller (BSC).

The BTS and BSC communicate across the standardized Abis interface. The BTS is controlled by the BSC and one BSC can have many BTS under its control.

Base Transceiver Station (BTS)

The BTS houses the radio transceivers and handles the radio-link protocols with the Mobile Station. Each BTS comprises radio transmission and reception devices including signal processors, antenna, etc. Each BTS is capable to support 1 to 16 RF carriers. The specifications differentiating the BTSs are; antenna type, antenna height, power level, and number of carriers.

Functions of BTS:

- It is in charge for frequency and time synchronization.
- The process of modulation for trans-direction, encryption, channel coding, and multiplexing,
- Arrange in advance for transmission from the mobiles, depending of their distance from BTS (Timing Advance).
- Detect random access requests from mobiles, as well as measure and monitor the radio channels for handover and power control.

Base Station Controller (BSC)

The BSC manages the radio resources for one or a group of BTSs. It handles radio-channel setup, frequency hopping, handovers, and control of the RF power levels. The BSC provides the time and frequency synchronization reference signals broadcast by its BTSs. It establishes connection between the mobile station and the MSC. The BSC is connected via interfaces to the MSC, BTS and OMC.

2.1.4 Networking and Switching Subsystem (NSS)

The NSS is responsible for performing call processing and subscriber-related functions. The network switching subsystem includes the following functional units:

- Mobile Switching Center (MSC)
- Home Location Register (HLR)
- Visitor Location Register (VLR)

Mobile Switching Center (MSC)

The MSC performs all the switching functions for all mobile stations located in the geographic area controlled by its assigned BSSs. Also, it interfaces with other MSCs, and other system entities.

Main functions of the MSC:

- Call handling that copes with the mobile nature of subscribers considering location registration, authentication of subscribers and equipment, handover and prepaid service.
- MSC-BSS signaling protocol management.
- Management of required logical radio link channel during calls.
- Controls inter-BSS and inter-MSC handovers.
- Handling location registration and ensuring interworking between mobile station and VLR.
- Acting as a gateway MSC to interrogate HLR.
- Standard functions of a switch such as charging.

Home Location Register (HLR)

The HLR is a central database that contains detailed information of each mobile phone subscriber that is authorized to use the GSM core network. HLRs store details of every SIM card issued by the mobile phone operator. A unique identifier included in every SIM, which is called IMSI, and is the primary key to each HLR record. One HLR per GSM network is recommended, and it may be a distributed database. Permanent data in HLR is changed by the man-machine interface. Temporary data such as location information changes dynamically in HLR.

Visitor Location Register (VLR)

The VLR is always integrated with the MSC. When a MS roams into a new MSC area, the VLR connected to that MSC would request data about the MS from the HLR. Later, if the mobile station makes a call, the VLR has the information needed in order to setup a call without having to interrogate the HLR every time. VLR contains information such as:

- Mobile identity
- Any temporary mobile sub identity
- HLR's part for the mobiles that are currently located in MSC service area.
- ISDN directory number of the mobile
- A directory number to route the call to the roaming station

2.1.5 Operation Subsystem (OSS)

Operation and Maintenance Center (OMC)

The OMC is a function through which the network operator can control and monitor the system by performing the following functions:

- Software installation
- Billing/Charging management
- Traffic management
- Subscriber administration
- Data analysis performance
- Tracing of subscribers and equipment
- Configuration management
- Mobile equipment management

Equipment Identity Register (EIR)

The EIR consists the identity of mobile station equipment called International Mobile Equipment Identity (IMEI), which may be valid, suspect, and prohibited. When a mobile station accesses the system, the equipment validation procedure is evoked before giving the services.

The information is available in the form of three lists:

- White List - The terminal is allowed to connect to the network.
- Grey List - The terminal is under observation from the network for possible problems.
- Black List - The terminals reported as stolen are not type-approved. They are not allowed to connect to the network. The EIR informs the VLR about the list, where the particular IMEI is in.

Authentication Center (AuC)

The AuC is associated with an HLR. It stores an identity key called authentication key (Ki) for each mobile subscriber. This key is used to generate the authentication triplets.

- RAND (Random Number)
- SRES (Signed Response) - to authenticate IMSI
- Cipher Key - to cipher communication over the radio path between the MS and the network.

2.2. Architecture in 3G technology

2.2.1 3G Network Infrastructure

When it comes to a 3G network, the only difference is the Radio Tower (NodeB) that supports a higher data rate (faster speed) together with a Radio Network Controller (RNC). The backend comprises the usual core components, the HLR, VLR, MSC, with the addition of the SGSN and GGSN, which provide access to the Packet Switching Network.

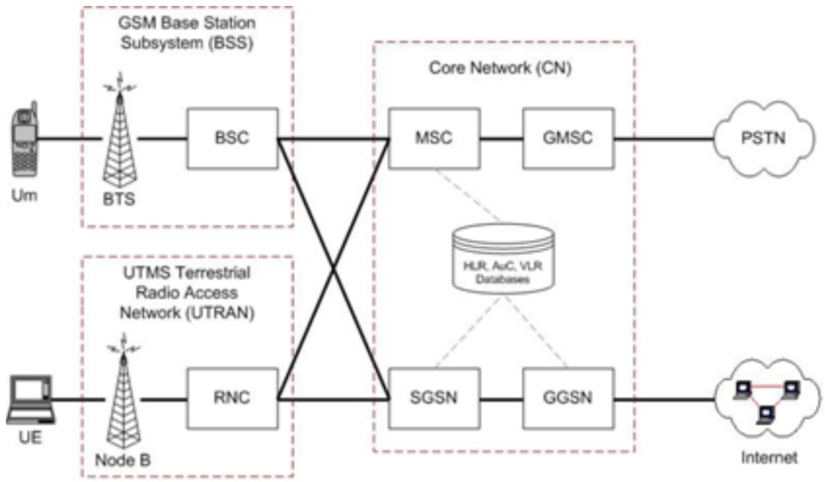

Figure 2 - how 2G/GSM (top) and 3G/UMTS (bottom) networks work side by side

Gateway GPRS Support Node (GGSN)

The GGSN is a main component of the GPRS network. The GGSN is in charge for the internetworking between the GPRS network and external packet-switched networks, such as the Internet. From an external network's perspective, the GGSN is a router to a "sub-network," because the GGSN "hides" the GPRS infrastructure from the external network. When the GGSN receives data addressed to a specific user, it checks if the user is active. If they are, the GGSN forwards the data to the SGSN serving the mobile user, but if the mobile user is inactive, the data is discarded. On the other hand, mobile-originated packets are routed to the right network by the GGSN.

Serving GPRS Support Node (SGSN)

The SGSN is in charge for the delivery of data packets from and to the mobile station within its geographical service area. Its tasks include packet routing and transfer, mobility management (attach/detach and location management), charging functions, logical link management, and authentication. The location register of the SGSN enlists location information such as current VLR and current cell, and user profiles (e.g. IMSI, addresses used in the packet data network) of all GPRS users registered with it.

2.3. Architecture in 4G technology

2.3.1 4G Network Infrastructure

In a 4G network things become a little different. Again the Radio Tower is connected to a device called eNodeB (femtocell), which connects to the EPC. The EPC is an all-in-one solution that contains the following components: Mobile Management Entity (MME) (replacing the MSC), which in turn connects to the Service Gateways SGW and PGW (which replace the GGSN and the SSGN), and the Policy Charging and Rules Functionality (PCRF) performing bandwidth shaping and billing. That connects users to the Public Data Network (Internet).

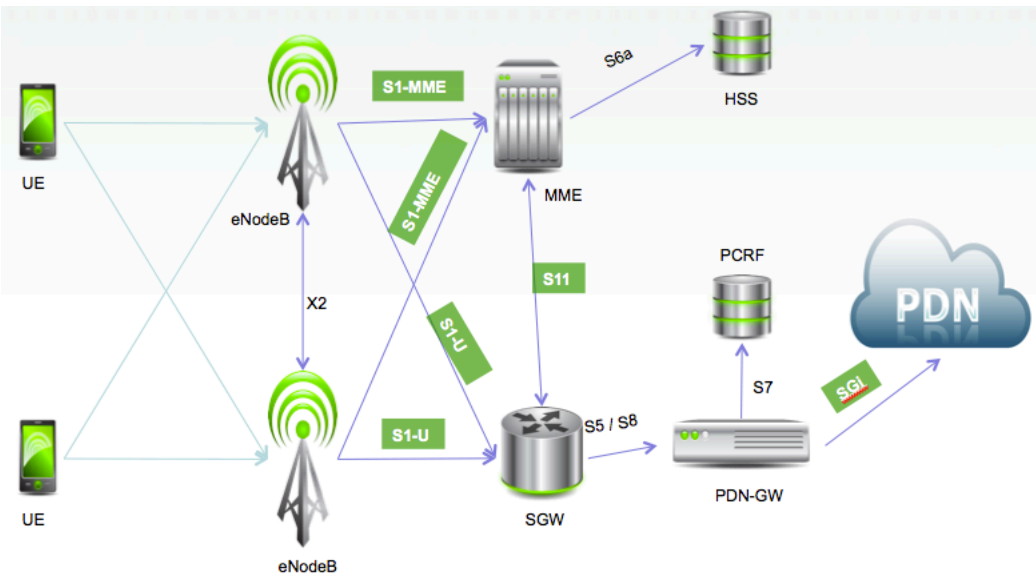

Figure 3 - the functions of a 4G network

- The Home Subscriber Server (HSS) component has been carried forward from the UMTS and GSM and is a central database that contains information about all the network operator's subscribers.
- The Packet Data Network Gateway (PGW) communicates with the outside world, using SGi interface. Each packet data network is identified by an access point name (APN). The PGW has the same role as the gateway GPRS support node (GGSN) in UMTS and GSM.
- The serving gateway (S-GW) acts as a router, and forwards data between the base station (eNodeB) and the PDN gateway and has the same role as the serving GPRS support node (SGSN) in UMTS and GSM.
- The mobility management entity (MME) controls the high-level operation of the mobile by means of signaling messages.
- The Policy Control and Charging Rules Function (PCRF) is in charge of policy control decision-making, as well as for controlling the flow-based charging functionalities in the Policy Control Enforcement Function (PCEF), which resides in the PGW.

2.4. Billing - Usage Capturing

A customer starts generating usage at the network as soon as he/she starts using the products and services sold by the operator. A network element is a combination of software plus hardware and responsible for metering events and overall service control for any type of service.

“Event” refers to a single billable occurrence of product usage, typically captured electronically by a network. For instance, when a mobile phone user makes a telephone call, an event is generated which contains information about that phone call, such as the time of day the call was made, call duration and the number that was called.

An event along with all its attributes is called Call Data Record (CDR). A data collector in the network switch captures the usage in the form of Call Data Record (CDR)/Usage Data Record (UDRs). These raw CDRs/UDRs are in turn converted by the mediation system into a format understandable by the Billing System. The following diagram in Fig. 4 shows the process where the network elements capture usage data and send Raw UDRs to the Mediation System, and finally to the Billing System for rating and billing.

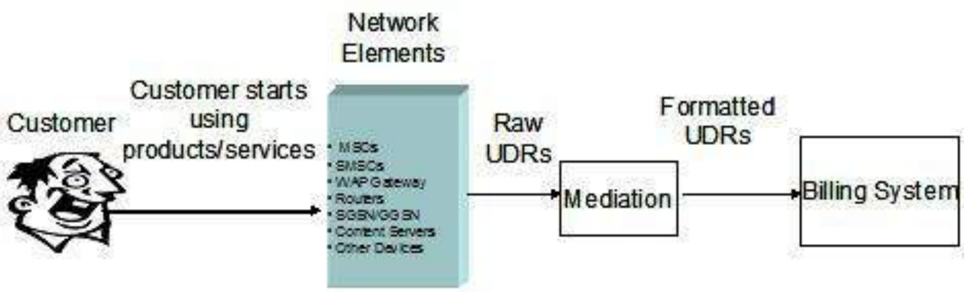

Figure 4 - process diagram of CDR, mediation and billing system

For service controlling and producing different types of CDRs; for example in GSM, the following network elements are presented

- MMS traffic is captured by the MMSC.
- SMS traffic is captured by the SMSC.
- Voice calls are captured by the MSC (Mobile Switching Centre).
- Roaming CDRs are captured by the roaming partner's switching element.
- Data traffic is captured by the GGSN.

GSM service providers are doing billing based on the services they are providing to their customers. All parameters are simple enough to charge a customer for the provided services. This chapter provides an overview of the frequently used billing techniques and parameters applied to charge a GSM subscriber.

Telephony Service

These services can be charged on per call basis. The call initiator has to pay the charges, and the incoming calls nowadays are free. A customer can be charged based on different parameters such as:

- International call or long-distance call.
- Local call.
- Call made during peak hours.
- Call made during night time.
- Discounted call during weekends.
- Call per minute or per second.
- Many more other criteria can be designed by a service provider to charge their customers.

SMS Service

Most of the service providers charge their customers' SMS services based on the number of text messages sent. There are other prime SMS services available where service providers charge more than the normal SMS charge. These services are being availed in collaboration with television networks or radio networks that require SMS from the audience.

Most of the time, the charges are paid by the SMS sender but for some services, such as stocks and share prices, mobile banking facilities, leisure booking services, etc., the recipient of the SMS has to pay for the service.

GPRS Services

Using a GPRS service, you can browse, stream videos or download files from the Internet. So a service provider will charge you based on the data uploaded as well as data downloaded on your mobile phone. These charges will be based on per Kilobyte data downloaded/uploaded.

An additional parameter could be a QoS provided to you. If you want to watch a movie, then a poor QoS may work because some data loss may be tolerable, but if you are downloading for example a zip file, then a single byte loss will corrupt your complete downloaded file.

Another parameter could be peak and off-peak time to download a data file or to browse the Internet.

Supplementary Services

Most of the supplementary services are being provided based on monthly rental or absolutely free. For example, call forwarding, call waiting, call on hold, and calling number identification are offered for free.

Call barring is a service used by service providers just to recover their dues, etc., otherwise this service is not being used by any subscriber.

Call conferencing service is a form of simple telephone call where the customers are charged for multiple calls made at a time. No service provider charges extra for this service.

Advice of Charge (AoC) can be charged based on the number of queries made by a subscriber.

Other Network Elements

Other network elements include components such as SMS Service Centre, Voice Mail Box, and SMS Flow.

SMS Service Centre

It interfaces with the MSC having interworking functionality to provide Short Message Service (SMS) to mobile subscribers. The SMS can be destined to a fax machine, PC on the Internet or another MS. The location of the recipient MS is queried by the MSC and delivered.

Voice Mail Box

When the mobile subscriber is not in a position to answer the incoming calls in case is busy or out of service area, then the call gets diverted to a mail box which has already been activated by the subscriber. For this, a separate connectivity has been established from the MSC. The subscriber will be alerted through SMS later and can retrieve the message.

SMS Flow

- When a user sends an SMS, the request is placed via the MSC.
- The MSC forwards the SMS to the SMSC, where it gets stored.
- The SMSC queries the HLR to find out where the destination mobile is and forwards the message to the destination MSC if the destination mobile is available.
- If the mobile is not available the message gets stored in the current SMSC itself. In most installations, if a mobile is not available for SMS delivery, the SMSC does not retry. Instead, the destination MSC informs the SMSC when the mobile comes back in range. SMS handling is a store-and-forward operation unlike USSD.

An SMS has a validity period for which it will wait for the destination mobile to be available. After that time, the SMSC will delete the message. The validity period can be set by the user. Normal validity is 1 day.

2.5. Limitations of conventional networks

4G networks are not enough essential to support a vast amount of connected devices with significant spectral efficiency and low latency, which will be the main aspect in the near future communication and computing. Below, we can have a view of some crucial aspects in which conventional cellular networks lag behind, thereby motivating the evolution of 5G networks

Lack of bursty data traffic. Not a few mobile applications use to send heartbeat messages to their servers and occasionally request for very high data transfer rate on a short period of time. Those kinds of data transmission, devour much more battery resources of mobile user equipment (UE) with increasing bursty data in the network, and hence, may crash the core network.

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

No distinguish between indoor and outdoor users. In the conventional cellular networks there is a single BS installed close to the center of the cell and interacts with all the UEs independently if the user is indoor or outdoor; while UEs remain indoors about 80% of the day and outdoors the other 20% of the day. In addition, the communication among indoor UE and the BS is not efficient in terms of data transfer rate, spectral efficiency, and energy-efficiency, because of the attenuation of signals passing through walls.

Co-channel interference. In a conventional cellular network, two separate channels are used. The first one is used for the data transmission from UE to the 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 channels operate at an identical frequency, i.e., a full duplex wireless radio, then a high level of co-channel interference in uplink and downlink channels is a major issue in 4G networks. It also prevents the network densification, i.e., the deployment of many BSs in a geographical area.

Inefficient utilization of BS's processing capabilities. The processing power of a base-station (BS) in the conventional cellular networks, can be used only by its correlated MSs, and they are designed to support peak time traffic. However, the processing power of a BS can be distributed across an extensive geographical area when it is lightly loaded. For instance, during daytime, BSs in business areas are overloaded, while BSs in residential areas are almost idle, and backwards. Also BSs in residential areas are over-subscribed in holidays and weekends while BSs in business areas are almost idle. However, the almost idle BSs consume an identical amount of power as over-subscribed BSs; thus, we have an increased overall cost of the network.

Lack of heterogeneous wireless networks support. The heterogeneous wireless networks (HetNets) are consist of wireless networks with diverse access technologies, for example, 3G and 4G networks, WiFi, wireless local area networks (WLAN), and Bluetooth. The HetNets are already standardized in 4G; however, the basic architecture was not intended to support them. Moreover, in the conventional cellular networks a UE is allowed to have a DL channel and a UL channel that have to be associated with a single BS that prevents the maximum utilization of HetNets. In HetNets, a UE is able to choose its UL and DL channels, between two different BSs that belong to two different wireless networks for maximum performance.

3. 5G Analysis

Wireless communication started in the early 1970s. In the four decades that followed, mobile wireless technology evolved from 1G to 5G generations. Fifth generation technology offers very high bandwidth, such as the user has never experienced before. Fifth generation technologies offer various new advanced features, which makes them most powerful and will place them in huge demand in the future.

Nowadays different wireless and mobile technologies exist, such as third generation mobile networks (UMTS - Universal Mobile Telecommunication System, cdma2000), LTE (Long Term Evolution), WiFi (IEEE 802.11 wireless networks), WiMAX (IEEE 802.16 wireless and mobile networks), as well as sensor networks, or personal area networks (e.g. Bluetooth, ZigBee). Mobile terminals include a variety of interfaces such as GSM, which are based on circuit switching. All wireless and mobile networks implement an all-IP principle, which means all data and signaling will be transferred via IP (Internet Protocol) on network layer.

5^{th} generation technology provides facilities such as camera, MP3 recording, video player, high phone memory, audio player, etc., that user never imagined before. 5^{th} generation mobile multimedia Internet networks can be completely wireless communication without limitation, which will be supported by LAS-CDMA (Large Area Synchronized Code-Division Multiple OFDM [Orthogonal frequency-division multiplexing]), IPv6 and few others.

5G technologies offer extreme data capabilities, unrestricted call volumes, and infinite data broadcast together within the latest mobile operating system. Fifth generation should make an important difference and add more services and benefits to the world over 4G. Fifth generation should be a more intelligent technology that interconnects the entire world without limits. This generation is expected to be released around 2020.

3.1. Key technologies for 5G

The need for radically higher data rates, 1000× higher than the previous generation network, will be addressed through combined gains in the following three categories:

1. Extreme densification and offloading to improve the area's spectral efficiency, with more active nodes per unit area and Hz.
2. Increased bandwidth by moving into mmWave spectrum, but also by making better use of WiFi's unlicensed spectrum in the 5-GHz band. In one word, more Hz.
3. Increased spectral efficiency through advances in MIMO, to support more bits/s/Hz per node.

The combination of more nodes per unit area and Hz, more Hz, and more bits/s/Hz per node, will compound into many more bits/s per unit area.

3.1.1 Extreme Densification and Offloading

A simple but extremely effective way to increase network's capacity is by making smaller cells. This technique has been applied over several cellular generations [2]. In the early 80s, when emerged the first generation, cells had sizes on the order of hundreds of square km. Since then, those sizes have been progressively shrinking and nowadays they are often fractions of a square km in urban areas.

For instance, in Tokyo, the spacing between BSs can be as small as two hundred meters, giving a coverage area under a tenth of a square km. Networks have been evolving quickly to include nested small cells, as for example picocells (range 100 meters maximum) and femtocells (WiFi-like range), in addition to distributed antenna systems which resemble in functionality picocells from a capacity and coverage standpoint. All their baseband processing are located at a central site and share cell IDs [3].

By shrinking the cells, we experienced numerous benefits, the most important of which was the reuse of spectrum across a geographic area and the ensuing reduction in the number of users competing for resources at each BS. Thus, in principle, cells can shrink almost indefinitely without a sacrifice in SIR, until nearly every BS serves a single user (or is idle). This allows each BS to devote its resources, as well as its backhaul connection, to an ever-smaller number of users.

Unfortunately, by shrinking the cells and by extreme densification, some challenges arise:

- Preserving the expected cell-splitting gains as each BS becomes more lightly loaded, particularly low-power nodes.
- Determining the appropriate associations between users and BSs across multiple radio access technologies (RATs), crucial for the optimization of the edge rate.
- Supporting mobility through such a highly heterogeneous network.
- Affording the rising costs of installation, maintenance and backhaul.

3.1.2 mmWave Spectrum

Terrestrial wireless systems have to a large degree confined their operation within the relatively small range of microwave frequencies extending from several hundred MHz to a few GHz and corresponding to wavelengths ranging between a few centimeters and about a meter. At this point, this spectral band—often called “beachfront spectrum”—has become nearly fully occupied, particularly during peak times. Regardless of the efficacy of densification and offloading, a lot more bandwidth is required [4].

Although beachfront bandwidth allocations can be made decidedly more efficient by modernizing regulatory and allocation procedures, to put large amounts of new bandwidth into play leaves us with only one way to go: increase frequency. Fortunately, great amounts of relatively idle spectrum do exist in the mmWave range of 30-300 GHz, where wavelengths are 1-10 mm as shown in Fig. 5.

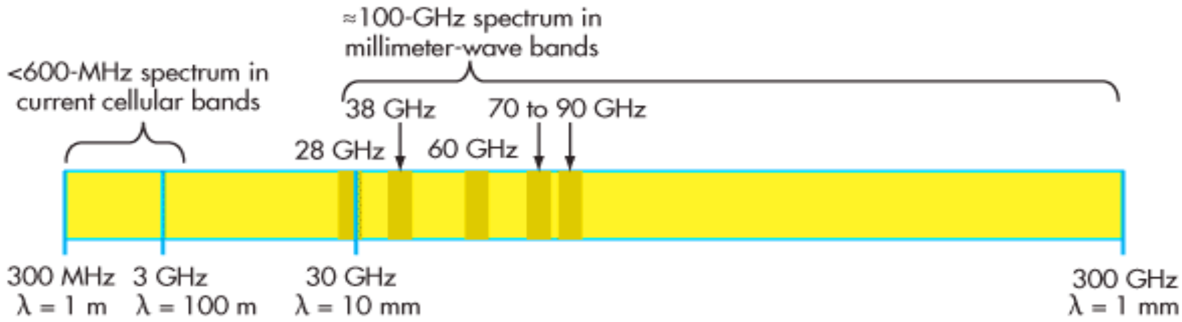

Figure 5 - mmWave spectrum

The main reason that mmWave spectrum was idle is that, until recently, it had been considered unsuitable for mobile communications because of rather hostile propagation qualities, including atmospheric and rain absorption, strong pathloss, penetration through objects, low diffraction around obstacles, and, furthermore, because of strong phase noise and exorbitant equipment costs. Therefore, according to the dominant perception, such frequencies, and particularly the large unlicensed band around 60 GHz, were suitable mainly for very-short-range transmission [5], [6]. Thus, the focus had been on WiFi (with the WiGiG standard in the 60-GHz band), and on fixed wireless in the 28, 38, 71-76 and 81-86 GHz. However, semi-conductors are maturing, their costs and power consumption are rapidly falling and the other obstacles related to propagation are now considered increasingly surmountable, given time and focused effort.

3.1.3 massive MIMO

The first introduction of MIMO communication into WiFi systems was around 2006, and into 3G cellular a short while later. Before going into "massive MIMO," a potential enabler of 5G, we must have a clear understanding of the technology behind the conventional, smaller-scale MIMO (multiple-input/multiple-output). MIMO deployment uses multiple antennas that are located at both the source (transmitter) and destination (receiver) [7]. Those antennas are linked in order to minimize error and increase efficiency of a network. This method's ability to multiply the capacity of the antenna links has made it a crucial component of wireless standards including 802.11n (Wi-Fi), 802.11ac (Wi-Fi), HSPA+, WiMAX and LTE.

Massive MIMO, as you might guess, takes MIMO technology and scales it up to hundreds or even thousands of antennas and terminals. These antennas, attached to a base station, focus the transmission and reception of signal energy into small regions of space, thus offering new levels of throughput and efficiency. The more antennas that are used, the finer the spatial focusing can be. Fig. 6 represents MIMO technology.

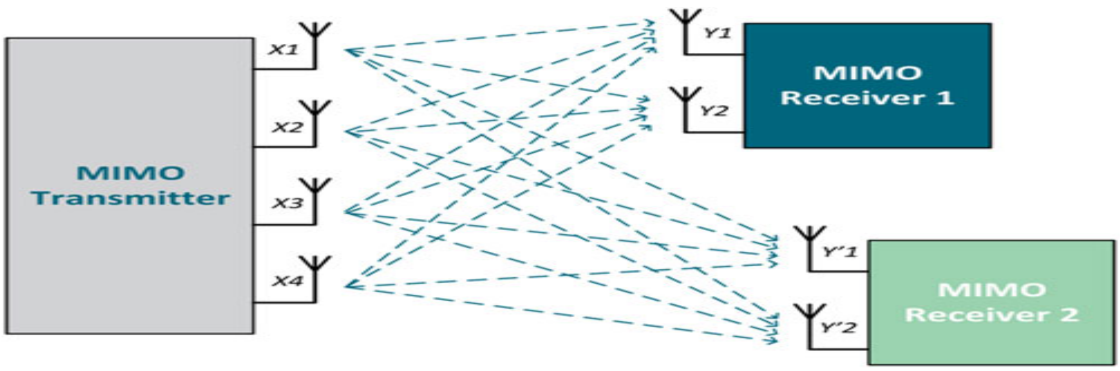

Figure 6 - Illustration of a massive MIMO technology

Using massive MIMO technology has a lot of benefits. Utilizing more antennas in a MIMO system, creates more degrees of freedom in the spatial area, consequently this enables greater improvement in performance to be gained:

- **Increasing data rate:** The increase in the number of antennas allows for a greater number of paths to be used, and hence a much greater level of data to be transferred within a given time.

- **Increasing basic link SNR**: One of the basic advantages of the use of MIMO systems is that they can be used to improve the SNR of the overall system. The use of large or massive MIMO allows it to be taken to a higher level.

- **Channel hardening:** Expanding the quantity of antennas significantly to create a massive MIMO system means that the system becomes more robust to the actual entries of the channel matrix. Thus, this has further advantages in the area of signal processing. It is necessary for linear detectors to perform matrix inversions; this can easily be done as part of the processing. It also allows simple detection methods to achieve very good performance as the dimensions increase. This decrease in the complexity of processing has a significant effect on many areas, allowing it to be used in many new applications.

Deploying a system with hundreds or thousands of antennas and terminals is not exactly plug-and-play. This requires a more advanced processing capability in the nodes. Moreover, each node must be able to determine the data transmitted from one antenna to the data transmitted from another; otherwise network performance will be limited [8].

Some other challenges, according to the Institute of Electrical and Electronics Engineers (IEEE), include:

- Making many low-cost, low-precision components that work effectively together.
- Acquisition and synchronization for newly joined terminals.
- Exploitation of extra degrees of freedom provided by the excess of service antennas.
- Finding new deployment scenarios for reducing internal power consumption to achieve total energy efficiency reductions.

3.2. Design Issues for 5G

To begin with, in case of achieving 1000 times higher data rates, 5G networks must decrease latency, decrease energy consumption, lower costs, and support many low-rate connections. In this chapter we are going to discuss important up-to-date research areas that support the above requirements. We are going to start with the most basic part of PHY (physical layer) the waveform and afterwards consider the progress of a cloud-based and virtualized network architectures, latency and control signaling plus energy efficiency.

3.2.1 Waveform: Signaling and Multiple Access

Signaling and multiple access formats, like waveform design, have essentially changed at each cellular generation, and to a substantial degree they have been each generation's defining technical feature. They have also often been at the epicenter of intense intellectual and industrial disputes, which have played out in the wider media. The first generation approach, based on analog frequency modulation with FDMA, transformed into a digital format for 2G; although it employed both FDMA and TDMA for multiple access, it was generally known as "TDMA" because of the novelty of time multiplexing.

In the meantime, a niche spread spectrum/CDMA standard developed by Qualcomm to compete for 2G became the dominant approach to all global 3G standards. Once the limitations of CDMA for high-speed data became inescapable, there was a subtle but unmistakable retreat toward TDMA, with minimal spectrum spreading retained and with the important addition of channel-aware scheduling [9].

Due to the increasing signal bandwidths needed to support data applications, orthogonal frequency-division multiplexing (OFDM) was unanimously adopted for 4G in conjunction with scheduled FDMA/TDMA, as the virtues of orthogonally were viewed with renewed appreciation.

It is natural to entertain the possibility that the transition to 5G could involve yet another major change in the signaling and multiple access formats.

OFDM and OFDMA: OFDM has become the prevailing signaling format for high-speed wireless communication, creating the basis of all current WiFi standards and LTE, as well as of wire-line technologies such as digital TV, digital subscriber lines, and commercial radio.

Its qualities include:

- Being able to cope naturally with frequency selectivity.
- Computationally efficient implementation via FFT/IFFT blocks and simple frequency-domain equalizers.
- A perfect match for MIMO, since OFDM allows for the spatial interference from multiantenna transmission to be dealt with at a subcarrier level, without the added complication of intersymbol interference.

OFDM (Orthogonal Frequency Division Multiplexing) and OFDMA (Orthogonal Frequency Division Multiple Access) are both wideband digital communication technologies with a small difference between them. However, they are both based on the same concept of bundling evenly spaced multiple sub-carriers with special characteristics into one big chunk, still transmitting separately over the transmission media. When it comes to providing simultaneous multi-user access, however, the two technologies have a significant difference in their channel allocation mechanism.

OFDM is a Frequency Division Multiplexing (FDM) mechanism, which works by dividing a single wideband signal into a large set of narrowband sub-carriers in such a way, that all of the sub-carriers are orthogonal to each other and evenly spaced. In other words, OFDM divides one high-speed signal into numerous slow signals that are more robust at the receiver’s end, so that the sub-channels can then transmit data without being put under the same intensity of multipath distortion faced by single carrier transmission. The numerous sub-carriers are then collected at the receiver and recombined, forming one high-speed transmission [10].

The orthogonally of the sub-carriers provide high spectral efficiency and low Inter-Carrier-Interference (ICI). Since each and every sub-carrier is treated as a different narrowband signal, each one of them modulated individually, it makes it easy to deal with frequency selective fading due to multipath. In other words, simplified channel equalization is required due to the sub-carrier nature of the narrowband. Moreover, low data rate of each sub-carrier reduces Inter Symbol Interference (ISI) greatly, and that results in very high Signal-to-Noise Ratio (SNR) of the system. As a result of all the above advantages, it is made possible to implement Single Frequency Network (SFN) and solve spectrum limitation issues in the commercial implementation of such a system.

In OFDM systems, only a single user can transmit on all of the sub-carriers at any given time. To accommodate multiple users, an OFDM system must employ Time Division Multiple Access (TDMA) (separate time frames) or Frequency Division Multiple Access (FDMA) (separate channels). Neither of these techniques is time or frequency efficient. The major drawback to these static multiple access schemes is the fact that different users seeing the wireless channels (Sub-carriers) differently is not being utilized. OFDM technologies typically have nomadic, fixed and one-way transmission standards, ranging from TV transmission to Wi-Fi, as well as fixed WiMAX.

On the other hand, OFDMA is the multi-user OFDM technology where users can be assigned on both a TDMA and FDMA basis, where a single user does not necessarily need to occupy all the sub-carriers at any given time. In other words, a subset of sub-carriers is allocated to a particular user. This enables several users to transmit low data rate simultaneously, and it can also be dynamically assigned to the best non-fading, low-interference channels for a particular user, avoiding bad sub-carriers to be assigned. Point-to-Multipoint fixed and mobile systems use OFDMA, and most emerging systems use OFDMA, such as Mobile WiMAX and LTE.

To summarize, comparing OFDM and OFDMA, we find the following main differences:

- OFDM supports multiple users (Multiple Access) via TDMA basis only, while OFDMA supports either on a TDMA or FDMA basis or both at the same time.
- OFDMA supports simultaneous low data rate transmission from several users, but OFDM can only support one user at any given moment.
- Further improvement in OFDMA over OFDM robustness to fading and interference since it can assign subset of sub-carrier per user by avoiding bad channels.
- OFDMA supports per channel or sub-carrier power while OFDM needs to maintain the same power for all sub-carriers.

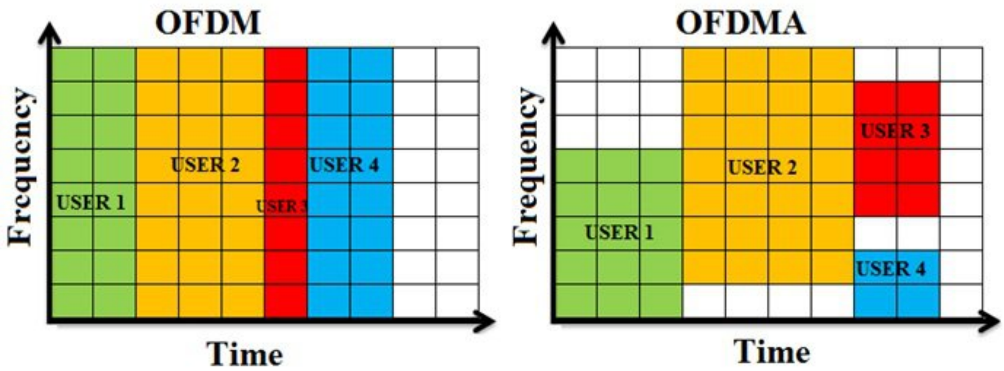

Figure 7 - differences between OFDM and OFDMA

3.2.2 Cloud-Based Networking

One of the most significant event, is the movement of data to the cloud. Thus, they can be accessed from anywhere and through a variety of terminals. This, in a general point of view reclassifies the endpoints and the time frame for which network services are provisioned. It requires that the network be much more nimble, flexible and scalable. All things considered, two innovation patterns will get to be distinctly vital later on, is the NFV (network function virtualization) and the SDN (software defined networking).

Both NFV and SDN are tend to be some of the greatest advance in mobile communication networking in the past two decades, tends to drastically change the way network services are provided. Although the move toward virtualization is thus far taking place only within the core network, this trend might eventually expand toward the edges. In fact, the term cloud-RAN is already being utilized, but for now largely to refer to schemes whereby multiple BSs are allowed to cooperate [11].

1) Network Function Virtualization: NFV enables network functions that were traditionally tied to hardware appliances to run on cloud computing infrastructure in a data center. It should be noted that this does not imply that the NFV infrastructure will be equivalent to commercial cloud or enterprise cloud. What is expected is that there will be a high degree of reuse of what the commercial cloud offers.

It is obvious to expect that some requirements of mobile networks, for example, the separation of the data layer, management layer and finally control layer will not be feasible within the commercial cloud. All things considered, the separation of the network functions from the hardware infrastructure will be the cornerstone of future architectures. The key benefit will be the ability to elastically support network functional demands.

Furthermore, this new architecture will allow for significant nimbleness through the creation of virtual networks and of new types of network services [12]. However, computing technology advances have made huge progress the past few years, where this vision can turn into a reality, with the ensuring architecture having recently been termed Software Defined Networking (SDN).

2) Software Defined Networking: SDN is an emerging architecture that is dynamic, manageable, adaptable and cost-effective making it ideal for the high-bandwidth, dynamic nature of today's applications. This architecture decouples the network control and forwarding functions enabling the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services. [13].

Some of the specific advantages of SDN is the efficiency and lower operating expenses, better and more granular security, enhanced network flexibility and holistic management. Moreover, as a plus of SDN can be considered the cloud abstraction, since it is easier to unify cloud resources. Last but not least is the guaranteed content delivery, in other words, the ability to shape and control data traffic. [14].

3.2.3 Energy Efficiency

In 5G, the energy efficiency of the communication chain will need to improve by about the same amount as the data rate just to maintain the power consumption level or even reduce it. This implies a several-order-of-magnitude increase in energy efficiency, which can be considered as a challenge. Unsurprisingly, in recent years there has been a surge of interest in the topic of energy efficient communications. In addition to laudable environmental concerns, it is simply not viable from a logistical, cost or battery-technology point of view to continually increase power consumption [15].

Due to the rapidly increasing network density, the access network consumes the largest share of the energy. Research has focused on the following areas.

1) Resource Allocation: The literature is rich in contributions dealing with the design of resource allocation strategies aimed at the optimization of the system's energy efficiency.
2) Network Planning: Energy efficient network planning strategies include techniques for minimizing the number of BSs for a coverage target and the design of adaptive BS sleep/wake algorithms for energy savings. Since networks have been designed to meet peak-hour traffic, energy can be saved by switching off BSs when they have no active users or simply very low traffic. Of course, there are different degrees of hibernation available for a BS and attention must be given in order to avoid unpleasant coverage holes (this is usually accomplished through an increase of the transmitted power from nearby BSs) [16].
3) Renewable Energy: Another intriguing possibility is that of BSs powered by renewable energy sources such as solar power. This is of great interest in developing countries lacking a reliable and ubiquitous power grid, but it also generates broader interest as it allows "drop and play" small cell deployment (if wireless backhaul is available) rather than "plug and play." A more stable scenario is where the resource allocation makes efficient use of both renewable and traditional energy sources.
4) Hardware Solutions: Finally, much of the power consumption issues will be dealt with by hardware engineers, with recent work in low-loss antennas, antenna muting, and adaptive sectorization according to traffic requirements.

To summarize, energy efficiency will be a major research theme for 5G, spanning many of the other topics in this chapter:

- True cloud-RAN could provide an additional opportunity for energy efficiency since the centralization of the base-band processing might save energy, especially if advances on green data centers are leveraged.
- The tradeoff between having many small cells or fewer macrocells given their very different power consumptions is also of considerable interest.
- A complete characterization of the energy consumed by the circuitry needed for massive MIMO is currently lacking.
- MmWave energy efficiency will be particularly crucial given the unprecedented bandwidths.

3.3. Applications of 5G Networks

High-speed data transfer, zero latency, and ubiquitous connectivity are some of the main advantages of 5G networks that are expected to serve a wide range of services and applications. In this part, we present some of the most remarkable applications of 5G networks, as follows:

Personal usages. This part of 5G networks would be able to support a wide range of UEs, from scalable to heterogeneous devices. Also the data demands like, voice communication, Web surfing, and multimedia data would be satisfied while maintaining a high level of QoS.

Healthcare systems. Having a secure, reliable, and fast mobile communication the healthcare system can de enhanced, for example, frequent data transfer from patients' body to the cloud or healthcare centers. Accordingly, the suitable and urgent medical services could be delivered to the patients as fast as possible.

Logistics and tracking. The future mobile communication could also benefit package tracking or inventory by 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 regardless of the geographic locations.

Automation. In the near future, self-driving vehicles would take place, and as a requirement, vehicles would communicate with each other in real-time. Furthermore, they would be able to communicate with other infrastructures on the roads, offices and homes with almost zero latency. Thus, an interconnected vehicular environment would provide a safe and efficient integration with other information systems.

Smart societies. With the term "smart societies" we refer to all connected virtualized offices, homes, and stores. For instance, smart stores would assist in filtering out irrelevant product details, sale advertisements, and item suggestions on the go. Therefore, every digital and electronic services/appliances like warning alarms, printers, temperature maintenance, air conditioners, and door locks, would create an interconnected network in a way that the collaborative actions would enhance the user experience.

Smart grids. Energy distribution would be decentralized with the use of smart grids and a better analyze of energy consumption would be achieved. This would allow smart grids to improve efficiency and economic benefits. 5G networks would also allow a rapid and frequent statistical data observation, analysis, adjusting accordingly the energy distribution.

Industrial usages. The zero latency attribute of 5G networks will benefit sensors, drones, mobile devices, and users to have real-time data without any delay, which would help to operate and manage industrial functions fast while maintaining energy efficiency.

4. Billing in 5G D2D Communication

In a conventional cellular system, devices do not have the capability of directly communicating with each other within the licensed cellular bandwidth, and all communications take place through the BSs (base stations). In this chapter, we present a two-tier cellular network that involves a macrocell-tier (BS-to-device communications) and a device-tier (device-to-device communications). Device terminal relaying makes it possible for devices in a network to function as transmission relays for each other and realize a massive ad hoc mesh network.

Obviously, this is a great departure from the conventional cellular architecture and brings unique technical challenges. In such a two-tier cellular system, since the user data is routed through other users' devices, it is important to maintain security for privacy reasons. To ensure minimal impact on the performance of existing macrocell BSs, the two-tier network must be designed with appropriate resource allocation schemes and smart interference management strategies. Furthermore, novel-pricing models should be designed to tempt devices to participate in this type of communication. The chapter provides an overview of these major challenges in two-tier networks and proposes some pricing schemes for different types of device relaying.

With the term “device” here we refers to any portable wireless device with cellular connectivity such as cellphone, tablet, laptop, etc. a user owns. Device relaying makes it possible for devices in a network to function as transmission relays for each other and establish a massive ad hoc mesh network. Of course, this will be possible with device-to-device (D2D) communication functionality, which allows two nearby devices to communicate with each other in the licensed cellular bandwidth with limited BS involvement or even without any base station (BS) involvement.

D2D functionality also has a crucial role in mobile cloud computing and facilitates effective sharing of resources (spectrum, computational power, applications, social contents, etc.) for users in close proximity to each other. Furthermore, service providers can take advantage of D2D functionality to take some load off the network in a local area such as a stadium or a shopping mall by allowing direct transmission among cell phones and other devices. Moreover, D2D communication can be of critical use in natural disasters. In an earthquake or hurricane, an urgent communication network can be set up in a short period of time using D2D functionality, replacing the damaged communication network and Internet infrastructure.

In the following paragraphs, we provide a categorization of D2D communication based on the degree of involvement of the cellular operator.

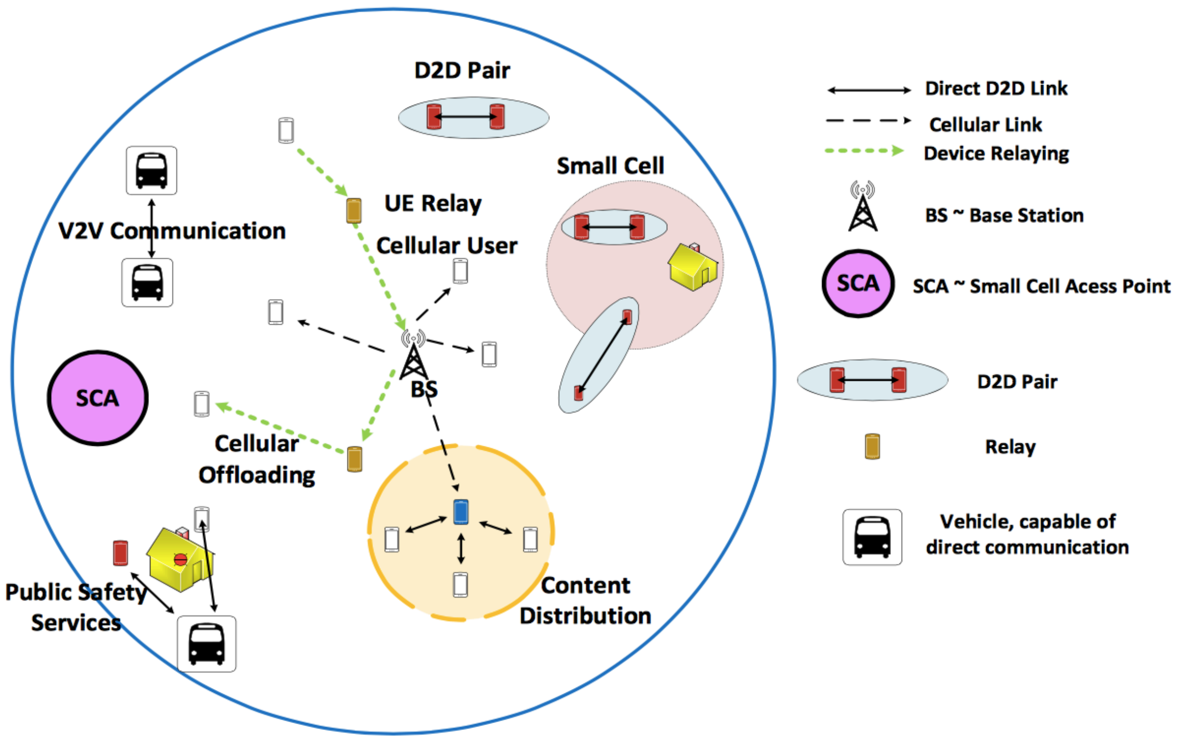

Figure 8 - A general scenario supporting D2D communication

4.1. OVERVIEW OF D2D COMMUNICATION TYPES AND MAIN TECHNICAL CHALLENGES

In this part, we present a two-tier 5G cellular network, the macrocell-tier and the device-tier. The macrocell tier involves base station (BS)-to-device communications as in a conventional cellular system. The device tier involves D2D communications. If a terminal (device) connects the cellular network through a BS, this device is said to be operating in the macrocell tier. If a device connects directly to another device or realizes its transmission through the assistance of other devices, these devices we can easily say that are operating in the device tier. In that kind of system, the BS will continue to serve the devices as in a conventional network. However, at congested areas or cell edges, devices will be allowed to communicate with each other, creating an ad hoc mesh network.

In the device-tier communications, the operator might have different levels of control, based on the business model; it either deploys full/partial control over the resource allocation among destination, source, and relaying devices, or prefers not to have any control at all. Thus, we can define the following four main types of device-tier communications:

1) **Device relaying with operator-controlled link establishment (DR-OC):** This type of communication is applicable for a device that is at the edge of a cell, i.e. in the coverage area, which has poor signal strength. In this type of communication, the devices will connect with the BS by relaying their information through other devices. This type of communication will be helpful for the device to attain a higher quality of service and respective increased battery life. The operator communicates with the relaying devices for partial or full control link establishment.

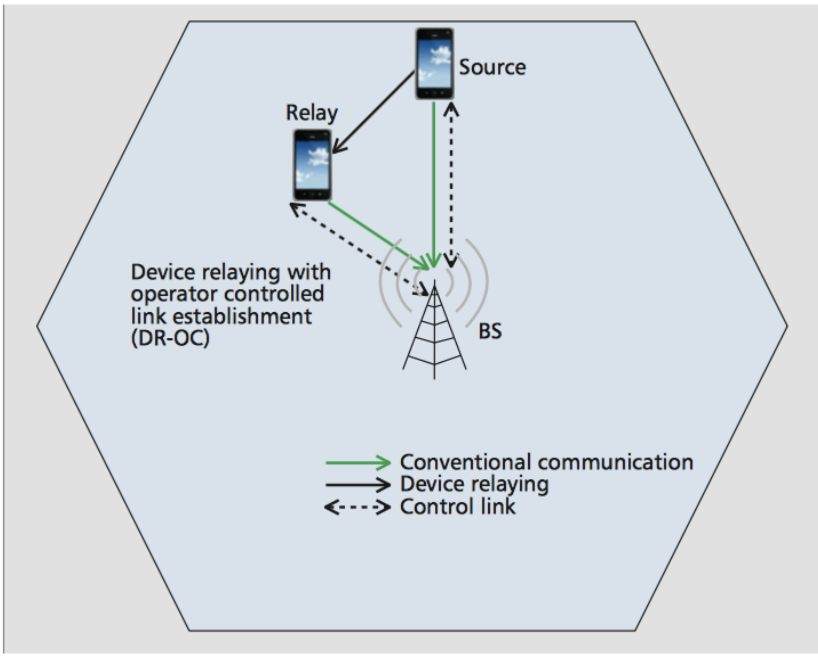

Figure 9 - illustration of device relaying communication with operator-controlled establishment (DR-OC). A device communicates with the BS through relaying its information via other devices.

2) **Direct D2D communication with operator-controlled link establishment (DC-OC):** In this type of communication, the source and destination devices exchange data with each other without the involvement of a base station, but they are supported by the BS for link formation.

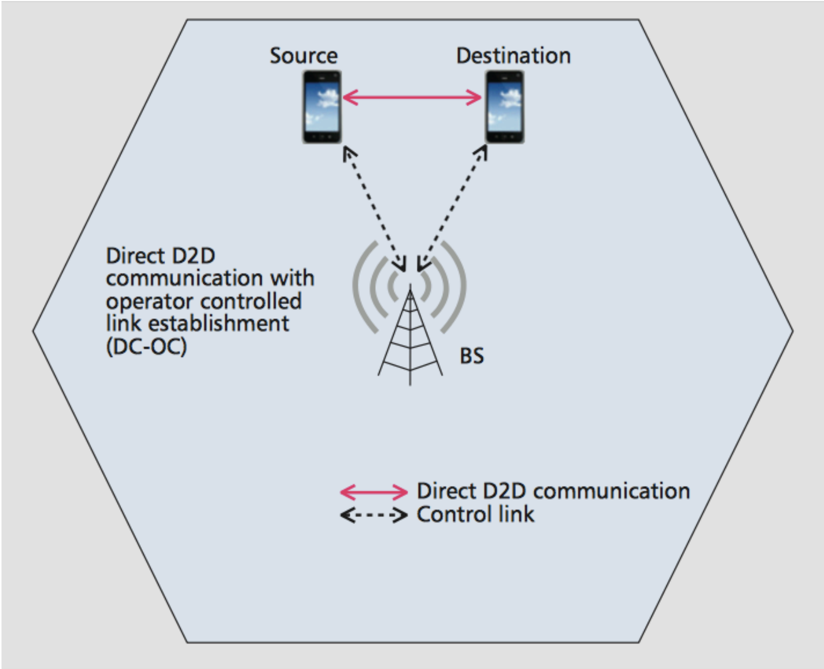

Figure 10 - illustration of direct D2D communication with operator-controlled link establishment (DC-OC).

The source and destination devices can talk and exchange data with each other without the need for a BS, but they are assisted by the BS through a control link.

3) **Device relaying with device-controlled link establishment (DR-DC)**: The operator is not involved in the process of link establishment. Therefore, source and destination devices are responsible for coordinating communication using relays between each other.

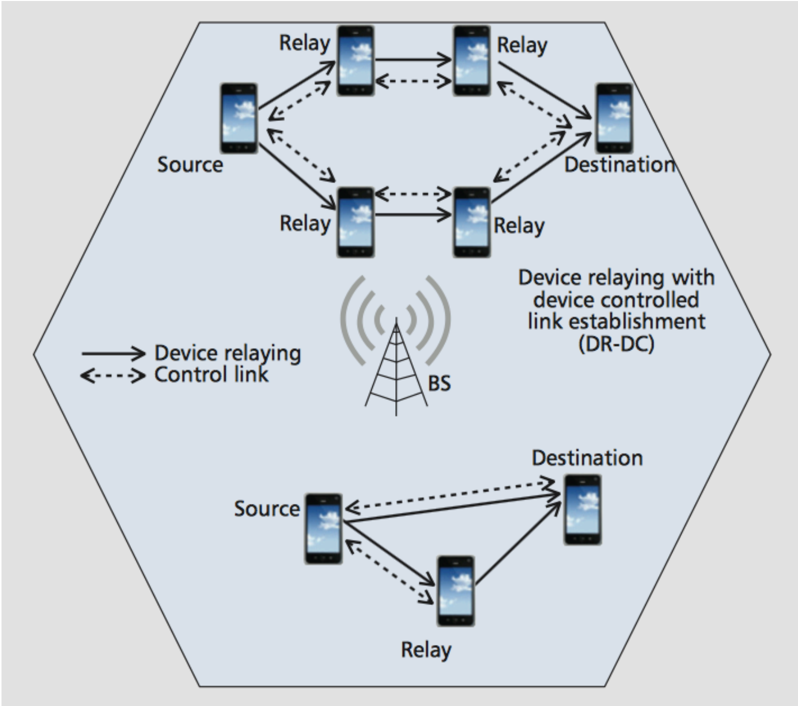

Figure 11 - illustration of device relaying communication with device-controlled link establishment (DR-DC).

Source and destination devices directly talk with each other and do not use any control link from the operator.

4) **Direct D2D communication with device-controlled link establishment (DC-DC)**: In this type of communication, the source and destination devices have direct communication with each other and the link formation is itself controlled by the devices without any assistance from the BS. Thus, source and destination devices should use the resource in such a way as to ensure limited interference with other devices in the same tier and the macrocell tier.

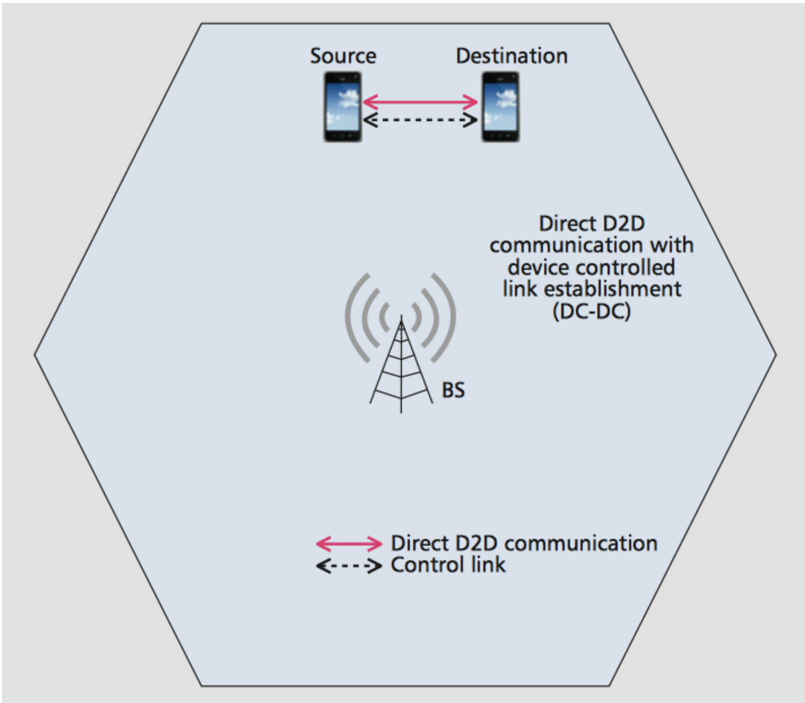

Figure 12 - illustration of direct D2D communication with device-controlled link establishment (DC-DC).

All the above should be carefully designed in order to have a reliable and advanced network. By the word "reliable," we refer to the technical challenges that need to be addressed, such as security and interference management.

In D2D communication, the routing of user data is through the devices of the other users, so the main area of concern is security, because privacy needs to be maintained. Closed access will ensure the security for the devices that want to operate on the device level. In closed access, a device has a list of certain reliable devices, like the users in the close vicinity or office to whom you know each other, otherwise the users that have been legitimated through a reliable party, like an association, can unswervingly communicate with each other, sustaining a level of discretion, whereas the devices not on this list need to use the macro cell level to communicate with it. Also the devices in a group can set a proper encryption between each other to avoid divulging their information to other devices. Since there is no type of supervision, security in such an instance is a challenging problem. Security issues in D2D communication include the identification of potential attacks, threats, and vulnerability of the system.

Concerning interference management, in device relaying communication with operator-controlled link establishment (DR-OC) and direct device-to-device communication with operator-controlled link establishment (DC-OC), the base station can execute the resource allocation and call setup process [17]. So, the base station can alleviate to a certain degree the problem of interference management by using centralized methods. But in device relaying communication with device-controlled link establishment (DR-DC) and direct device-to-device communication with device-controlled link establishment (DC-DC) resource allocation between devices will not be supervised by the centralized unit. Operating in the same licensed band, devices will inevitably impact macrocell users.

To make sure that there is as minimal impact as possible on the performance of existing macrocell BSs, a two-tier network should be designed using smart interference management strategies and appropriate resource allocation schemes. Apart from the interference between the macrocell and device tiers, interference can also be seen among users in the device tier.

To address resource allocation for this type of communication, different approaches can be employed such as admission control and power allocation, cluster partitioning, relay selection and other.

In DR-OC, as illustrated in Fig. 9, since the BS is one of the communicating units, the aforementioned challenges can be addressed with the help of the base station like authenticating the relaying devices through encryption for maintaining adequate privacy of the information of the devices. The challenge of spectrum allocation amid the relaying devices to prevent them from interfering with other devices will also be managed by the BS.

In DC-OC, as illustrated in Fig. 10, the devices communicate directly with each other, but the BS controls the formation of links between them. Basically the BS has complete control over the D2D connections, like connection setup and maintenance, and resource allocation. So for assigning resources to every D2D connection, the network can either assign resources in an identical manner as a regular cellular connection or in the form of a dedicated resource pool to all D2D connections.

On the other hand, in DR-DC and DC-DC, as shown in Figs. 11 and 12, there is no BS to control the communication between the devices, but several devices are communicating with each other by using cooperative or non-cooperative communication by playing the role of relays for the other devices. Since there is no centralized supervision of the relaying, distributed methods will be used for processes like connection setup, interference management, and resource allocation. Before the data transmission phase, two devices need to find each other and the adjacent relays. The devices can periodically broadcast identity information so that other devices may be aware of their existence, and then they will decide whether or not to start a device-to-device direct or device relaying communication [18].

4.2. Billing in D2D Communication

Apart from the technical issues we summarized in the previous section, another challenge that operators need to address is how they will control and charge for their services in D2D communication. It is critical for operators to have a convenient and tempting way of charging, otherwise users may turn in to traditional D2D technologies, which are free of charge but have lower speed and less security.

Devices that act as relays for other users use their own resources such as battery, data storage, and bandwidth. Therefore, tempting billing models should also be designed in order to force devices to be part of this type of communication. Furthermore, a secure environment should be exist, since devices in direct D2D communication want to use the process of selling and buying resources among themselves. In this case the operator is the one who can control and create this secure environment for this kind of process. Therefore, it can expect some payment from the devices for the security and QoS in D2D communication.

Below we present some billing scenarios for each case of the direct relaying (DR) communication and the device-to-device (D2D) communication.

4.2.1 Billing for DR-OC

The main challenge in device relaying with operator-controlled link establishment (DR-OC) is to motivate the users for participating. Since they use their own resources (e.g. battery, bandwidth) to relay the information of other devices, they need tempting monetary or other kinds of incentives to become involved in this type of communication. One possible scenario is that the operator can offer some discounts on monthly bills based on the amount of data they relay through their devices. From the operator's point of view giving such discounts is reasonable, because the operator benefits from providing service to devices with poor coverage or demanding data rates higher than available in the macrocell tier.

Another possible incentive for the relaying device is that the operator can offer some free services in exchange for the amount of data they have relayed, instead of a monthly discount on the bill. To implement this in real time, the base station can send some control signals to users and give them the option instead of using the conventional data transmission to use the award-data transmission.

In Fig. 13 we illustrate the revenues of the operator and the devices in DR-OC and compare them with the conventional (non-cooperative) communication.

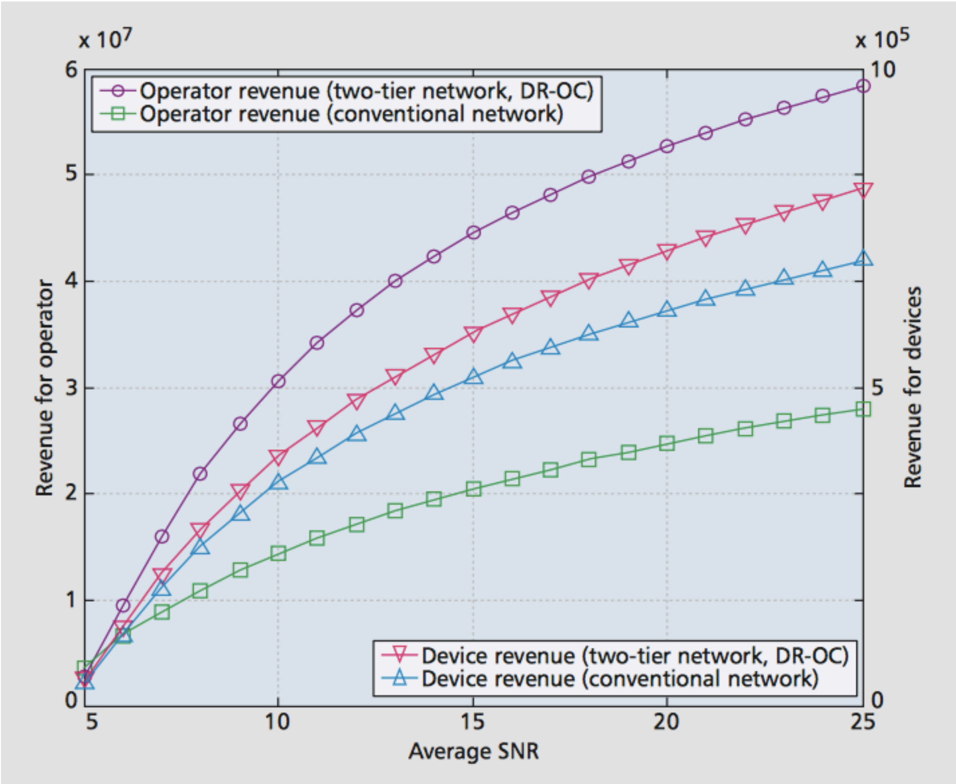

Figure 13 - Comparison of device and operator revenues in conventional single-tier cellular system and proposed two-tier system with DR-OC.

In DR-OC, fixed pricing is used by the operator for the spectrum and free spectrum is offered to the devices in exchange for data relayed by them. This is equal to less throughput for the "reward." In cooperative communication the devices experience better quality of service; therefore it can be expected that they will be charged more by the operator compared to the conventional cellular system. It is observed in Fig. 13 that the revenues of the operator and devices in a two-tier network with DR-OC increase compared to those in a conventional single-tier cellular system. Specifically, for the numerical values given, the operator revenue for high SNRs doubled, while the revenue of the devices increases by about 25 percent in the best-case scenario. Therefore, we can conclude that DR-OC is more beneficial for the operator.

4.2.2 Billing for DC-OC

An important challenge for operators in DC-OC is to have tempting pricing schemes in order for the users to be satisfied by using this service rather than free WiFi or Bluetooth. Spectrum trading is a solution that can satisfy both the operator and the user. In spectrum trading, the seller's (operator) main priority is to maximize the profit, while the buyer's (users) is to maximize the utility of spectrum usage. However, since these objectives are in contrast with each other, an optimal and sustainable solution for spectrum trading would be required so the profit and utility are maximized, while both operator and user are satisfied.

An optimal and tempting solution for spectrum trading seems to be the first-price sealed-bid auction (FPSBA), also known as blind auction. In this type of auction, all bidders (users) simultaneously submit sealed bids, so that no bidder knows the bid of any other participant, then the highest bidder pays the price they submitted [19]. In DC-OC, when there is only one operator (monopoly market) the pair of devices that want to communicate with each other submit their bids to the BS. Then the BS selects the bids and allocates the channels. The BS can set a threshold price for a certain channel for which devices should place a bid higher than this price to acquire it.

Furthermore, the BS can set to auction the same channel to more than two users, assuming sufficient distance to avoid interference and gain more revenue from a channel. This can be achieved easily, when a pair that wants to submit their bid to the BS can include their positions using GPS. Also, there are other methods for the positioning of devices, such as handset-based, and SIM-based, by which the BS can locate the position of the devices.

As an example to demonstrate the revenues of operators and devices in a two-tier cellular system with DC-OC, we should consider the following scenario: Assume there is only one operator in the market that has 50 available channels, each with a bandwidth of 5 MHz, and the number of devices varies between 100 and 800. In a conventional single-tier cellular system, at any one time only 100 devices can talk with each other under the assumption of 50 available channels.

Let's imagine a two-tier cellular system with DC-OC and set a reserved price "i" for the submitted bids. Therefore, devices can submit any bid higher than the reserved price based on their valuation of spectrum. Let pc=1 denotes the unit price of spectrum in a single-tier cellular system. We further assume that the submitted bid for spectrum by the devices in a two-tier system can take values in the range of [i, Pmax], where Pmax is the maximum value of bids. The operator collects all the bids and allocates the available channels to the highest bidders. The revenue of the operator is the summation of all payments from the winning devices, while the revenue of the devices is the difference between their valuation and their payment to the operator.

In Fig. 14 we illustrate the revenues for both operator and devices. It is observed that the revenue of the operator in DC-OC increases with the increasing number of devices, since their value of bids increases. In the single-tier cellular system, however, the revenue of the operator remains constant since the operator is already using the full network capacity. The revenue of the operator in the two-tier network with DC-OC is 6 percent higher than that of the single-tier cellular system.

The revenue of the devices in both cases decreases as the number of devices increases. However, the amount of revenue that the devices can generate in a two-tier network is higher than the single-tier counterpart by 270 percent when the number of devices is 100. This gain is reduced to 27 percent when the number of devices increases to 800. Since the value of bids increases with the increasing number of devices, the revenue of devices decreases. It should also be noted that since there are more channels available in DC-OC, the devices could enjoy more throughput than those in the conventional network.

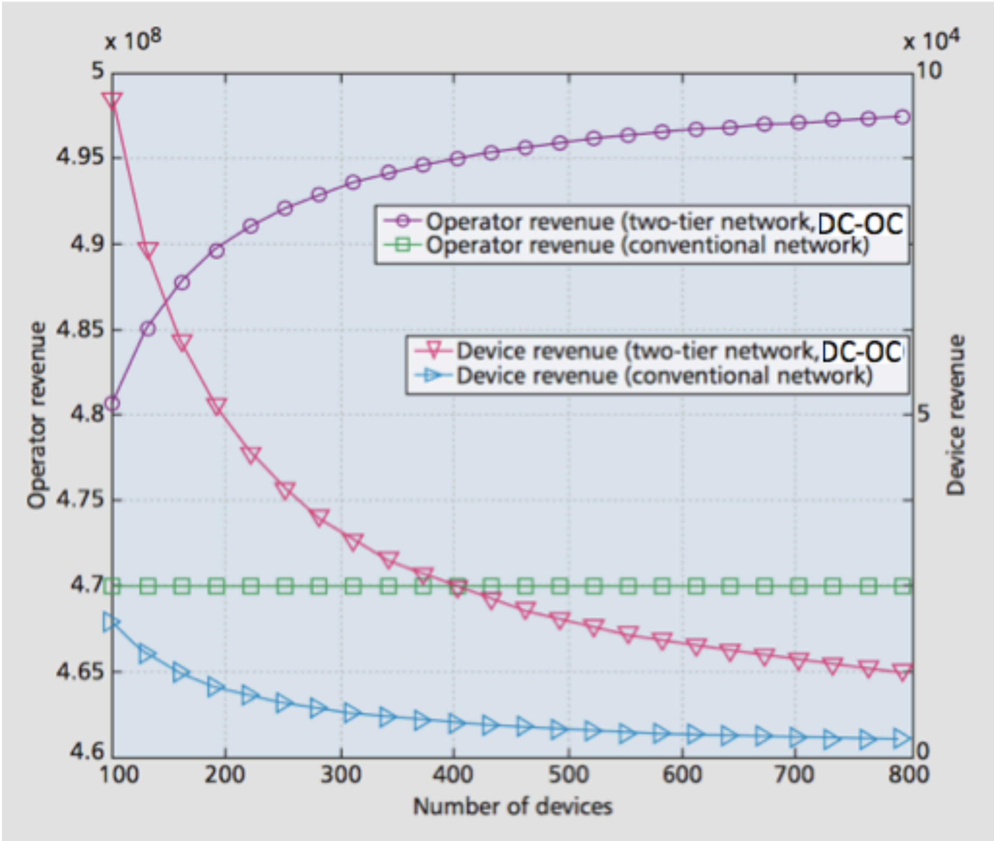

Figure 14 - Single operator and a first-price sealed-bid auction (FPSBA) are used. Comparison of device and operator revenues in conventional single-tier cellular system and the proposed two-tier system with DC-OC.

4.2.3 Billing for DR-DC and DC-DC

As long as, the operator does not have any control over this relaying communication, it is obvious that it should not suppose to make any profit from this. However, in closed-access DR-DC, devices know each other via different ways; such as being a member of a group or association or having some kind of a relationship, they can agree on a pricing scheme or simply relay the information without expecting any payment from other devices, similar to when a user uses Bluetooth or WiFi to transfer a file. In open-access DR-DC, users can adopt different pricing approaches such as cooperative game, bargaining game, or double auction [19].

For the pricing between devices in DC-DC, we can consider that the users are selfish and have a set of social relationships, and hence will not display the same behavior toward all other devices. Since the user who is considered selfish will also need to send messages using other devices at times, it is in the selfish user’s interest not to ask for monetary payment from the devices in its work group/community. Whenever a selfish device receives a message from another device, it refers to these records and calculates the percentage of meeting days, and the standard deviation of the daily meetings. Based on this, the selfish devices display altruistic behavior toward devices they meet often, and these devices appeared that are better off than the devices that are always selfish [20].

5. CONCLUSION

Device-to-device (D2D) communications is seen as a new technology that will be implemented in 5G mobile networks to provide high performance in cellular network, improve coverage, provide spectral efficiency and high data rates, as well as offer new peer-to-peer services with QoS guarantees.

In this thesis, we presented some of the most important technologies in 5G networks, such as cell densification and channel offloading, the use of mmWave spectrum and the use of massive MIMO’s technology benefits. Moreover, we present the design issues that need to be applied such us signaling, cloud-based networking and energy efficiency. Then we presented a two-tier 5G cellular network, the macrocell-tier and the device-tier. We focused on the device-tier, where we provided a categorization of four different D2D communication types, considering the degree of involvement of the operator.

Following that, we discussed some crucial technical challenges that should be addressed for each type, such as interference management, resource allocation issues and security. In the end, we focused on billing issues for each type of communication and we compared the device and operator revenues in a conventional single-tier network and a two-tier network. From the graphs of figure 13, it is obvious that in the two-tier network, both operators and users obtain more revenue than in a conventional cellular network and also users experience better QoS.

Finally, in the graphs of figure 14 it is clearly illustrated that the revenue of the operator in the two-tier network is 6 percent higher than that of the conventional network. From the device’s side the amount of revenue that can be obtained in a two-tier network is higher than the single-tier by 270 percent when the number of devices is 100 and it is reduced by 27 percent when the number of devices increases to 800 since the value of bids increases with the increasing number of devices, the revenue of devices decreases.

REFERENCES

[1].“5G Technology – Redefining wireless Communication in upcoming years” by Akhilesh Kumar Pachauri 1 and Ompal Singh published in International Journal of Computer Science and Management Research Vol 1 Issue 1 Aug 2012 ISSN 2278 – 733X.

[2]. V. Chandrasekhar, J. G. Andrews, and A. Gatherer, “Femtocell networks: A survey,” IEEE Commun. Mag., vol. 46, no. 9, pp. 59–67, Sep. 2008.

[3]. J. G. Andrews, “Seven ways that HetNets are a cellular paradigm shift,” IEEE Commun. Mag., vol. 51, no. 3, pp. 136–144, Mar. 2013.

[4]. “Estimated spectrum bandwidth requirements for the future development of IMT-2000 and IMT-Advanced,” ITU-R, Geneva, Switzerland, Rep. M.2078, 2006.

[5]. T. S. Rappaport, J. N. Murdock, and F. Gutierrez, “State of the art in 60-GHz integrated circuits and systems for wireless communications,” Proc. IEEE, vol. 99, no. 8, pp. 1390–1436, Aug. 2011.

[6]. T. Baykas, C.-S. Sum, Z. Lan, and J. Wang, “IEEE 802.15.3c: The first IEEE wireless standard for data rates over 1 Gb/s,” IEEE Commun. Mag., vol. 49, no. 7, pp. 114–121, Jul. 2011.

[7]. T. L. Marzetta, “The case for MANY (greater than 16) antennas as the base station,” in Proc. ITA, San Diego, CA, USA, Jan. 2007.

[8]. F. Rusek et al., “Scaling up MIMO: Opportunities and challenges with very large arrays,” IEEE Signal Process. Mag., vol. 30, no. 1, pp. 40–60, Jan. 2013.

[9]. P. Bender et al., “CDMA/HDR: A bandwidth efficient high-speed wireless data service for nomadic users,” IEEE Commun. Mag., vol. 38, no. 7, pp. 70–77, Jul. 2000.

[10]. D. Lopez-Perez, A. Valcarce, G. De La Roche, E. Liu, and J. Zhang “OFDMA femtocells: A roadmap on interference avoidance,” IEEE Commun. Mag., vol. 47, Sep. 2009.

[11]. N. Lee, R. W. Heath, Jr., D. Morales-Jimenez, and A. Lozano, “Base station cooperation with dynamic clustering in super-dense cloud-RAN,” in Proc. IEEE GLOBECOM, Dec. 2013, pp. 784–788.

[12]. Network Function Virtualization—An Introduction, Benefit, Enablers, Challenges and Call for Action, white paper, Oct. 2012. [Online]. Available: http://portal.etsi.org/NFV/NFV_White_Paper.pdf

[13]. Software-Defined Networking, the New Norm for Networks, Apr. 2012. [Online]. Available: http://www.opennetworking.org

[14]. S. Sezer et al., "Are we ready for SDN? Implementation challenges for software-defined networks," IEEE Commun. Mag., vol. 51, no. 7, pp. 36– 43, Jul. 2013

[15]. "Special issue on energy-efficient wireless communications," IEEE J. Sel. Areas Commun., vol. 29, no. 8, Sep. 2011.

[16]. E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in Proc. IEEE GLOBECOM, Dec. 2010, pp. 1–5.

[17]. L. Lei et al., "Operator Controlled Device-to-Device Communications in LTE-Advanced Networks," IEEE Wireless Commun., vol. 19, no. 3, June 2012, pp. 96–104.

[18]. G. Fodor *et al.*, "Design Aspects of Network Assisted Device-to-Device Communications," IEEE Commun. Mag., vol. 50, no. 3, Mar. 2012, pp. 170–77.

[19]. E. Hossain, D. Niyato, and Z. Han, Dynamic Spectrum Access and Management in Cognitive Radio Networks, 1st ed., Cambridge Univ. Press, 2009.

[20]. M. N. Tehrani, M. Uvsal, H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: challenges, solutions, and future directions" IEE Communication magazine, Volume: 52, Issue: 5, May 2014