Table of Contents

SYLLABUS

Course Catalog

3 Credit hours (3 h lectures). Introduction; Next Generation Network Architectures; Physical Layer Design; MAC Layer Design; Applications: Device-to-Device Communications, Machine-to-Machine Communications, Internet of Things, Vehicular Communications, Wearable Technologies, etc.; Quality of Service (QoS) and Quality of Experience (QoE); Energy Considerations; Practical Considerations; Open Issues.

Textbook

No one textbook. Students will be referred to several recent books and IEEE articles.

References

BOOKS

1. Wei Xiang, Kan Zheng and Xuemin (Sherman) Shen, Editors, (2017). *5G Mobile Communications*. Springer.

JOURNAL ARTICLES

2. Agiwal, Mamta, Abhishek Roy, and Navrati Saxena. "Next generation 5G wireless networks: A comprehensive survey." *IEEE Communications Surveys & Tutorials* 18.3 (2016): 1617-1655.
3. Chin, Woon Hau, Zhong Fan, and Russell Haines. "Emerging technologies and research challenges for 5G wireless networks." *IEEE Wireless Communications* 21.2 (2014): 106-112.
4. Demestichas, Panagiotis, et al. "5G on the horizon: key challenges for the radio-access network." *IEEE Vehicular Technology Magazine* 8.3 (2013): 47-53.

Instructor

Instructor: Dr. Mohammad M. Banat

Email Address: banat@just.edu.jo

Prerequisites

Prerequisites by topic Communication Systems, Digital Communications

Prerequisites by course EE 450, EE 551

Topics Covered

Week	Topics
1	Introduction
2-4	Next Generation Network Architectures
5-6	Modulation, Coding, and Multiple Access
7-8	Millimeter Wave Communications
9-10	Massive MIMO Systems
11-12	Device-to-Device Communications
13-14	Machine-to-Machine Communications
15-16	Internet of Things

Evaluation

Assessment Tool	**Expected Due Date**	**Weight**
Mid-Term Exam		25%
Term Project Report		15%
Presentations		10%
Final Exam	According to the university final examination schedule	50%

I. INTRODUCTION

I.1. Brief Overview

The continuing current revolution in wireless and mobile communications started in the late 1970s. Initial systems provided analog voice calls at acceptable quality. Modern technologies provide high quality mobile broadband services with end-user data rates of several megabits per second over wide areas and tens, or even hundreds, of megabits per second locally. Introduction of new digital modulation techniques, frequency reuse, penetration of packet-based communication, and rapid advancement in physical layer technologies, have significantly contributed in this evolution.

Initiation of new types of mobile devices such as smart phones and tablets (and many other devices), has produced a great variety of new applications which are used for mobile connectivity. Currently all-IP based fourth generation LTE networks have become a part of everyday life. As a result, a set of new, user-oriented mobile multimedia applications, like mobile video conferencing, streaming video, e-healthcare and other online services are coming up. An exponential growth in network traffic has resulted as a consequence. These new applications do not only satisfy user requirements, but also open up new business horizons for wireless operators to increase their revenues.

The vision of next generation wireless communications lies in providing very high data rates (typically in Gbps), extremely low latency, manifold increase in base station capacity, and significant improvement in quality of service (QoS), compared to current networks. The future will witness a networked society that features unbounded access to information and allows – everyone everywhere every time everything – sharing of data. New technology components need to be developed to facilitate the evolution of existing wireless based technologies. Wireless technologies, like long-term evolution (LTE), High Speed Packet Access (HSPA) and Wireless Fidelity (Wi-Fi), will be incorporating new technology components that will help them to meet the needs of the future. Needless to say, totally new technologies are almost certain to emerge.

Exercise I.1

Prepare a concise 5-page report on each one of the following technologies:

- Wi-Fi
- LTE
- HSPA

I.2. Evolution of Wireless Technologies

Wireless communications is an important part of present day society. Wireless communications started when Marconi used electromagnetic waves to transmit the letter “S” along a distance of 3 km in the form of three dot Morse code. Since then, several generations of wireless communications have appeared. With the fast growth of wireless technologies, the data rates, mobility support, coverage and spectral efficiency are all on the rise. Along the path from earlier generations to current ones, networks have moved from exclusive use of circuit switching to exclusive use of packet switching, with mixed forms of switching in between.

Exercise I.2

Prepare a concise 5-page report on:

- Mobility support, coverage and spectral efficiency in wireless networks
- Circuit-switched connections
- Packet- switched connections

I.2.A. First Generation Wireless (1G)

The 1st generation started in the late 1970s and early 1980s. It had a data rate up to 2.4 kbps. It used analog signaling and FDMA to allocate channels. Its disadvantages include: very low capacity, unreliable handoff, and no security procedures.

I.2.B. Second Generation Wireless (2G)

The 2nd generation was introduced in mid to late 1990s. Digital technology is used in 2nd generation mobile telephones. Global System for Mobile communications (GSM) was the first 2nd generation system, primarily used for voice communication, with a data rate of up to 64 kbps. GSM used TDMA for channel allocation. It provided services like Short Message Service (SMS) and e-mail. 2G mobile handset battery lasts longer because of the radio signals having low power. Some 2G systems (especially in north America) used CDMA for channel allocation.

I.2.C. 2.5G

A 2.5G system generally uses 2G system frameworks, but it applies packet switching along with circuit switching. It can support data rate up to 144 kbps. Packet transmission for data communication in 2.5G is called General Packet Radio Services (GPRS). The main 2.5G technologies were GPRS, Enhanced Data Rate for GSM Evolution (EDGE), and Code Division Multiple Access (CDMA).

I.2.D. 3G and 3.5G

The 3rd generation (3G) was established in late 2000. It supports transmission rate up to 2 Mbps. 3G systems provide high speed mobile access to services based on Internet Protocol (IP). Aside from transmission rate, unconventional improvement was made by maintaining quality of service (QoS). Additional features like global roaming and improved voice quality made 3G a remarkable generation. The major disadvantage for 3G handsets is that, they require more power than most 2G models.

Since 3G involves the introduction and utilization of Wideband Code Division Multiple Access (WCDMA), Universal Mobile Telecommunications Systems (UMTS) and Code Division Multiple Access (CDMA) 2000 technologies, the evolving technologies like High Speed Uplink/Downlink Packet Access (HSUPA/HSDPA) and Evolution-Data Optimized (EVDO) has made an intermediate wireless generation between 3G and 4G named as 3.5G with improved data rates of 5-30 Mbps.

I.2.E. 3.75G

Fixed Worldwide Interoperability for Microwave Access (WIMAX) and Long-Term Evolution technology (LTE) are seen as the future of mobile data services. Fixed WIMAX and LTE have the potential to supplement the capacity of the network and provide a substantial number of users the facility to access a broad range of high speed services like on demand video, peer to peer file sharing and composite web services.

Along with this, a supplementary spectrum is accessible which allows operators to manage their network very compliantly, and offers better coverage with improved performance for less cost.

I.2.F. 4G

4G is generally referred as the descendant of the 3G and 2G standards. 3rd Generation Partnership Project (3GPP) has standardized LTE Advanced (LTE-A) as a 4G standard along with Mobile WIMAX. A 4G system improves communication networks by providing a complete and reliable solution based on IP. Services like voice, data and multimedia are provided to subscribers, at quite higher data rates as related to earlier generations, everywhere and at every time.

I.2.G. 5G

With an exponential increase in the demand of the users, 4G will be easily replaced with 5G with an advanced access technology named Beam Division Multiple Access (BDMA) and Non- and quasi-orthogonal Filter Bank multi carrier (FBMC) multiple access.

Project I.1-M

- Prepare a complete study of the use of multicarrier modulation in wireless communications.

It is commonly assumed that 5G cellular networks must address several challenges that are not effectively addressed by 4G. These challenges include:

- higher capacity,
- higher data rate,
- lower end to end latency,
- massive device connectivity,
- reduced cost,
- quality of experience (QoE) provisioning

Ever increasing popularity of smart devices, introduction of new emerging multimedia applications, together with an exponential rise in wireless data (multimedia) demand and usage is already creating a significant burden on existing cellular networks. 5G Wireless systems, with improved data rates, capacity, latency, and QoS are expected to provide the solution of most of the current cellular network problems. Potential facilitators to address the 5G main challenges, and main 5G design fundamentals are summarized in Figure I.1.

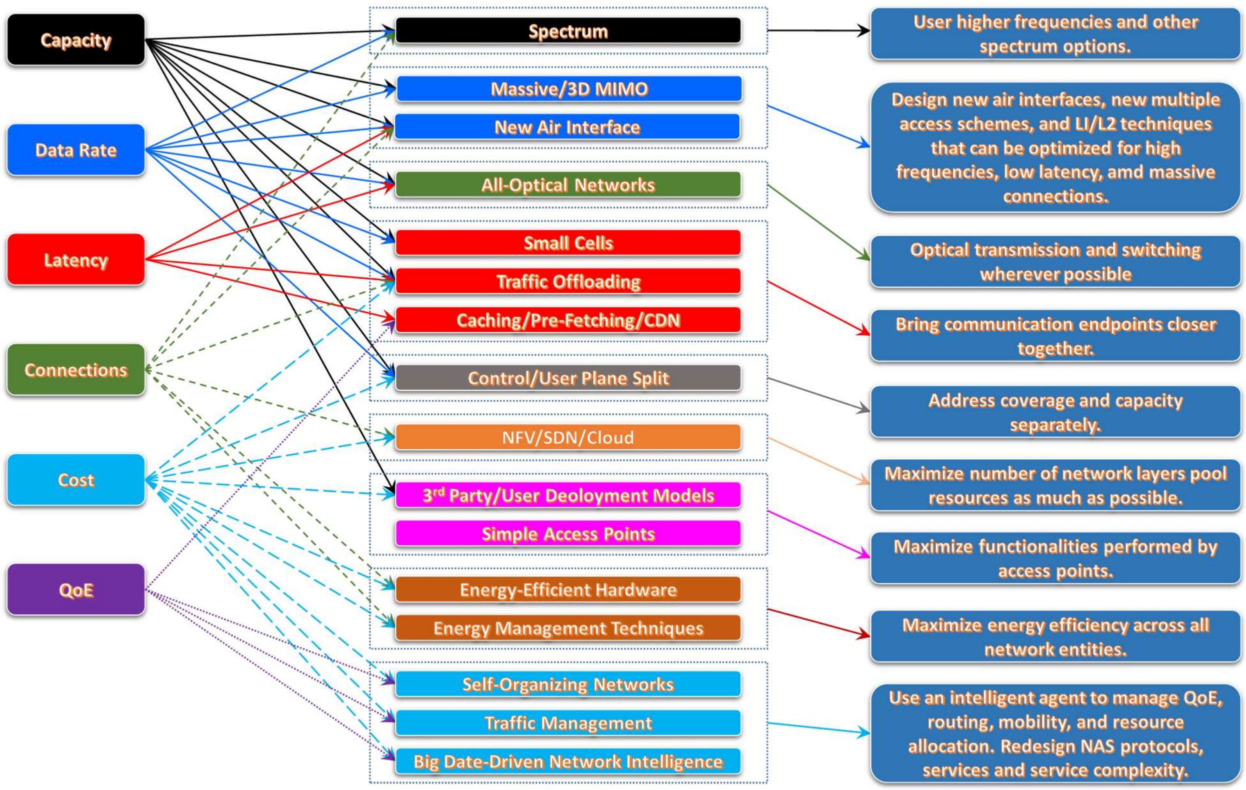

Figure I.1: 5G challenge, facilitators, and design fundamentals [1]

Gen.	Access Technology	Data Rate	Band	Bandwidth	FEC	Switching	Applications
1G	AMPS FDMA	2.4 kbps	800 MHz	30 kHz	NA	Circuit	Voice
2G	GSM TDMA	10 kbps	850/900/ 1800/1900 MHz	200 kHz	NA	Circuit	Voice +Data
	CDMA			1.25 MHz			
2.5G	GPRS	50 kbps		200 kHz		Circuit/ Packet	
	EDGE	200 kbps					
3G	WCDMA UMTS			5 MHz	Turbo Coding		

I.2-Evolution of Wireless Technologies

<table>
<tr><td></td><td colspan="2">CDMA2000</td><td>384 kbps</td><td rowspan="3">800/850/ 900/1800/ 1900/2100 MHz</td><td>1.25 MHz</td><td rowspan="3"></td><td></td><td rowspan="3">Voice +Data +Video Calling</td></tr>
<tr><td rowspan="2">3.5G</td><td colspan="2">HSUPA/HSDPA</td><td rowspan="2">5-30 Mbps</td><td>5 MHz</td><td rowspan="7">Packet</td></tr>
<tr><td colspan="2">EVDO</td><td>1.25 MHz</td></tr>
<tr><td rowspan="2">3.75G</td><td colspan="2">LTE OFDMA/SC-FDMA</td><td rowspan="2">100-200 Mbps</td><td>1.8, 2.6 GHz</td><td>1.4-20 MHz</td><td rowspan="2">Concatenated Codes</td><td rowspan="4">Online gaming+ HDTV</td></tr>
<tr><td>WiMAX SOFDMA</td><td>Fixed WiMAX</td><td>3.5, 5.8 GHz</td><td>3.5, 7 MHz (3.5 GHz) 10 MHz (5.8 GHz)</td></tr>
<tr><td rowspan="2">4G</td><td colspan="2">LTE-A OFDMA/SC-FDMA</td><td>DL 3Gbps
UL 15 Gbps</td><td>1.8, 2.6 GHz</td><td>1.4-20 MHz</td><td rowspan="2">Turbo Coding</td></tr>
<tr><td>WiMAX SOFDMA</td><td>Mobile WiMAX</td><td>100-200 Mbps</td><td>2.3, 2.5, 3.5 GHz</td><td>3.5, 5, 7, 8.75, 10 MHz</td></tr>
<tr><td>5G</td><td colspan="2">BDMA, FBMC</td><td>10-50 Gbps</td><td>1.8, 2.6, 30-300 GHz</td><td>60 GHz</td><td>LDPC</td><td>UHD Video, VR</td></tr>
</table>

I.3. Pre-5G Cellular Networks

It is well known that mobile data consumption is exploding, driven by increased penetration of smart devices (smartphones and tablets), better hardware (e.g., better screens), better user interface design, compelling services (e.g., video streaming), and the desire for anywhere, anytime high-speed connectivity.

Global mobile traffic experienced around 70% growth in 2014. Smartphones (26% of the total number of global mobile devices) are responsible for 88% of total mobile data traffic. The percentage of smart devices is expected to rise to more than 50% in 2019. Since 2012 video and multimedia traffic has comprised more than half of the global mobile traffic. An average mobile user is expected to download around 1 terabyte of data annually by 2020.

Exercise I.3

Calculate the maximum number of concurrent full HD streaming (1080p) downloads on an LTE connection with a theoretical maximum download date rate of 150 Mbps.

More end users are using multiple devices with different capabilities to access a mix of best effort services (e.g., instant messaging and email) and services with quality of experience (QoE) expectations (e.g., voice and video streaming). Over-the-top (OTT) players provide services and apps, some of which compete directly with core operator services (e.g., voice, SMS, and MMS). Connectivity is increasingly evaluated by end users in terms of how well their apps work as expected, regardless of time or location (in a crowd or on a highway), and they tend to be unforgiving toward the mobile operator when these expectations are not met.

Current research is heavily directed at applications of the Internet of Things (IoT), the Internet of Vehicles (IoV), Device to Device (D2D) communications, Machine to Machine (M2M) communications, among other technologies. IoT, which adds "anything" as an additional dimension to connectivity (in addition to anywhere and anytime), is becoming a reality.

Smart wearable devices (e.g., bracelets, watches, glasses), smart home appliances (e.g., televisions, fridges, thermostats), sensors, autonomous cars, and cognitive mobile objects (e.g., robots) promise a hyper-connected smart world that could create many interesting opportunities in many sectors of life such as healthcare, agriculture, transportation, manufacturing, logistics, safety, education, and many more.

Project I.2

- Prepare a complete study of the capacity of a cellular network.

Presently, almost all wireless communications use spectrum in 300 MHz to 3 GHz band, often termed as "sweet spot" or "beachfront spectrum". This band derives benefits from its reliable propagation characteristics over several kilometers in different radio environments. However, this band cannot accommodate the exploding mobile traffic and connectivity. The key essence of next generation wireless networks lies in exploring the unused, high frequency mm-wave band, ranging from 3 to 300 GHz. The spectrum from 59 to 64 GHz is open for unlicensed wireless communications. The spectra from 71 to 76 and 81 to 86 GHz are open for licensed peer to peer communications. Radio astronomy, radars, airport communications and many military applications have already been using the mm-wave bands over the last few decades.

Of the huge 3-300 GHz mm-wave spectrum, only ranges 57-64 GHz and 164-200 GHz are un-suitable for communications. The 57-64 GHz band is suitable for very short range point-to-point and point-to-multipoint applications. Of course, Even a small fraction of available mm-wave spectrum can support hundreds of times of more data rate and capacity over the current cellular spectrum. The availability of a big chunk of mm-wave spectrum is opening up a new horizon for spectrum future wireless communications.

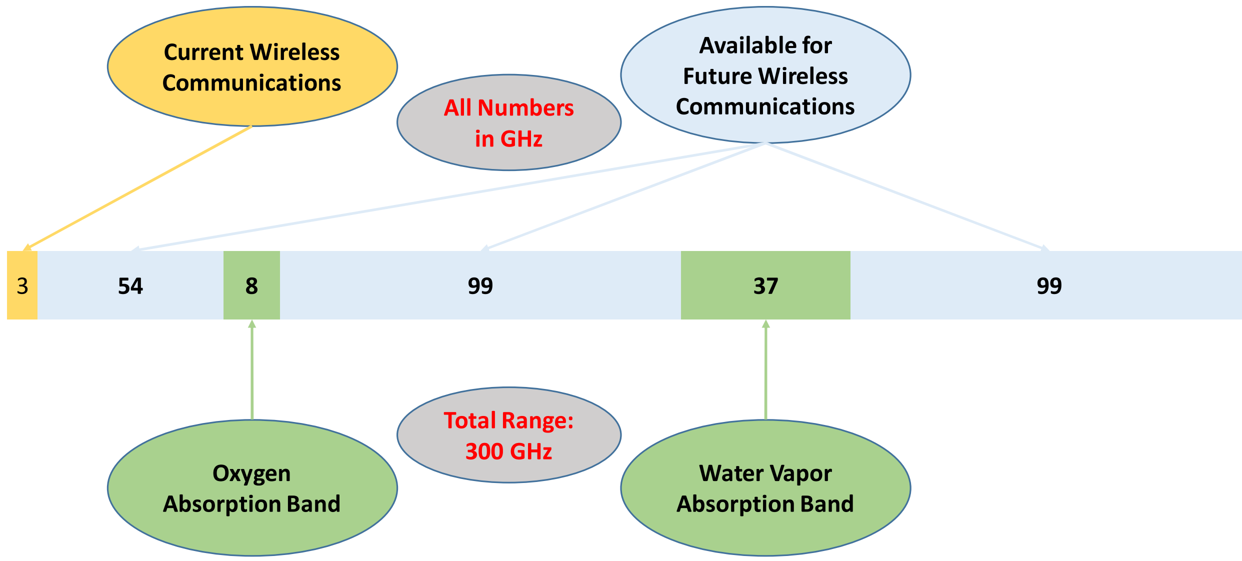

Figure I.2: mm-wave Spectrum [2]

Exercise I.4

- Calculate the range of wavelengths used in current wireless communications (frequency range: 0.3-3 GHz).
- Calculate the range of wavelengths used in next generation wireless communications (expected frequency range: 3-300 GHz).

Exercise I.5

How does the value of carrier frequency affect the allowable transmitted signal bandwidth?

Project I.3

- Prepare a study of what mm-wave bands are suitable for which wireless applications.

I.4. 5G Vision and Motivation

5G wireless communications envision magnitudes of increase in wireless data rates, bandwidth, coverage and connectivity. This is sought to be achieved with substantial reduction in latency and energy consumption.

I.4.A. Requirements on 5G Networks

- 1-10 Gbps data rates in real networks. This is almost 10 times increase from traditional LTE network's theoretical peak data rate of 150 Mbps.
- 1 ms round trip latency. Almost 10 times reduction from 10 ms round trip time in 4G.
- High bandwidth in unit area. It is needed to enable large numbers of connected devices with higher bandwidths for longer durations in a specific area.

- Enormous number of connected devices. In order to realize the vision of IoT, emerging 5G networks need to provide connectivity to thousands of devices.
- Perceived availability of 99.999%. 5G envisions that network should practically be always available.
- Almost 100% coverage for 'anytime anywhere' connectivity. 5G wireless networks need to ensure complete coverage irrespective of users' locations.
- Reduction in energy usage by almost 90%. Development of green technology is already being considered by standards bodies. This is going to be even more crucial with high data rates and massive connectivity of 5G wireless.
- High battery life. Reduction in power consumption by devices is fundamentally important in emerging 5G networks

Project I.4

Prepare a detailed study of the mutual independence of the transmission bit and symbol rates, bit and symbol energies, power, number of bits per symbol, distances in the signal space constellation, signal to noise ratio, and error probability.

Some wireless vendors and operators assume that 5G development should start in a backward compatible way with existing 4G LTE networks. This would help in continuing services using the same carrier frequency to traditional devices. 5G networks have been demonstrated at the PyeongChang 2018 winter Olympic games, where Samsung, Intel, and Korea Telecom (KT) have collaborated to bring 5G service experience to the public. The partners provided a 4K streaming video service via a 5G network to highlight the massive high-volume and high-speed data transmission potential of 5G. KT provided the 5G data network through a collaboration led by Intel with partners including Ericsson, Nokia and Alibaba, while Samsung unveiled its 5G mobile tablet device to deliver a 4K streaming video via Intel's base stations.

Future networks are expected to connect diverse nodes in different proximities. Small, micro, pico and femto cell deployment is already underway. Thus, dense 5G networks will have high co-channel interference, which will gradually render the current air interface obsolete. This pushes in the concept of sectorized and directional (energy focused) antennas, as opposed to the age-old omnidirectional antennas. Therefore, Space Division Multiple Access (SDMA) and efficient antenna design are utmost necessary.

Overall layout of 5G wireless networks breaks the rules of BS centric cellular concept and moves towards a device centric topology. The propagation and penetration of mm-wave signal in outdoor environment is quite limited. Thus, node layout cannot follow traditional cellular design or even any definite pattern. Ultra dense deployment is necessary in areas requiring high data rates, like subway stations, malls and offices.

Line of Sight (LOS) communication is undisputed preference over Non Line of Sight (NLOS) communication. Reflected, scattered and diffracted signals still might have sufficient energy, which needs to be explored when LOS is completely blocked. 5G cellular technology needs to work with an enormous number of users, variety of devices and diverse services. The primary concern therefore, is the integration of 5G BSs with the legacy cellular networks (e.g. 4G, 3G and 2G).

I.4.B. Use Cases

5G will penetrate into every single element of our future society and will create an all-dimensional, user-centered information ecosystem. 5G will break the limitation of time and space to enable an immersive and interactive user experience. 5G will also shorten the distances between humans and things, and implement a seamless integration to achieve easy and smart interconnection between people and all things. 5G will enable us to realize the vision – “Information is a finger away, and everything will be kept in touch”. There are a great many use cases proposed by different organizations. Mobile Internet and the Internet of Things (IoT) are the two main market drivers in the future development of mobile communications, and they will trigger a large range of use cases.

I.4.C. Use Cases of Mobile Internet

Mobile Internet is disrupting the traditional business model of mobile communications, enabling unprecedented user experiences and making a profound impact on every aspect of people’s work and life. Mobile Internet will promote the continued evolution of the way humans interact with information, and provide users with ultimate experience through more immersive services including but not limited to:

- Video services, such as immersive Ultra High Definition (UHD) and three-dimensional (3D) video
- Augmented reality
- Virtual reality
- Online gaming applications
- Tactile Internet
- Remote computing
- 3D connectivity: aircrafts and drones

In 2020 and beyond, there will be an explosive growth in mobile data traffic. It is estimated by International Mobile Telecommunication (IMT-2020) 5G Promotion Group that the global mobile data traffic will grow by more than 200 times from 2010 to 2020, and by nearly 20,000 times from 2010 to 2030.

It is observed that the current average traffic asymmetry ratio of mobile broadband is in favor of the downlink, and this is expected to increase due to a growing demand for audiovisual content. Mobile Internet is aiming at people-oriented communications with a focus on the user experience. Towards 2020 and beyond, the increasing popularity of ultra-high definition (UHD), 3D and video immersion will significantly drive up the data rates. At the same time, consumers will continue to demand better experiences on mobile communications wherever they are. A consistent service experience is expected in all scenarios, including ultra-dense scenarios such as stadiums, open-air gatherings, and high-speed moving scenarios such as high-speed trains, vehicles and subways. The total number of devices connected by global mobile communications networks will reach 100 billion in the future.

I.4.D. Use Cases of Internet of Things

IoT is focused on communications between things and between things and people, involving not only individual users, but also a large number of various vertical industrial customers. IoT services types and relevant requirements of IoT services are very diverse.

The IoT has extended the scope of mobile communications services from interpersonal communications to interconnection between things (smart devices), and between people and things, allowing mobile communications technologies to penetrate into broader industries and fields. The potential IoT use cases include:

- Smart grid and critical infrastructure monitoring
- Environmental monitoring
- eHealth services
- Remote object manipulation like remote surgery
- Internet of vehicles
- Smart wearables
- Sensor networks
- Smart cities

Some IoT services like mobile health will have strict requirements on transmission rates, while services such as IoV will demand millisecond-level latency and nearly 100% reliability. In addition, many IoT devices may be deployed in remote locations, or in areas where transmission losses can be a problem, such as indoor corners, basements and tunnels. Therefore, the coverage of mobile communications networks need to be further enhanced. In order to penetrate into more IoT services, 5G should be more flexible and more scalable, to support massive device connections and meet diverse user requirements.

5G will also need to meet extremely high security requirements, particularly for services such as e-banking, security monitoring, safe driving, and mobile health. 5G will also be able to support lower power consumption to build greener mobile communication networks and to enable much longer terminal battery life, especially for some IoT devices.

I.4.E. Classification of 5G Use Cases

It is helpful to group countless emerging use cases into several use case families. Use cases in each family share similar characteristic and requirements.

Usage Scenarios

- Enhanced mobile broadband: Mobile broadband addresses human-centric use cases for access to multimedia content, services and data. The demand for mobile broadband will continue to increase, leading to enhanced mobile broadband. The enhanced mobile broadband usage scenario will come with new application areas and requirements in addition to existing mobile broadband applications for improved performance and an increasingly seamless user experience.

- Seamless wide-area coverage scenario: Seamless coverage for high mobility is desired, with much improved user data rate compared to existing data rates. However, the data rate requirement may be relaxed compared to the hotspot scenario.
- High-capacity hot-spot scenario: For an area with high user density, very high traffic capacity is needed, while the requirement for mobility is low and the user data rate is higher than that of wide area coverage.

- Ultra-reliable and low latency communications: This use case has stringent requirements for capabilities such as throughput, latency and availability. Some examples include wireless control of industrial manufacturing or production processes, remote medical surgery, distribution automation in a smart grid, transportation safety, etc.
- Massive machine type communications: This use case is characterized by a very large number of connected devices typically transmitting a relatively low volume of non-delay-sensitive data. Devices are required to be of low cost, and have a very long battery life.

I.5. High-Level Key Capabilities

Key performance indicators (KPIs) for 5G include the user experienced data rate, connection density, end-to-end delay, traffic volume density, mobility, and peak date rate.

Table I.1: 5G Performance indicators

Performance Indicator	Definition
User Experienced Data Rate (bps)	Minimum achievable data rate for a user in real network environment
Connection Density (/km^2)	Total number of connected devices per unit area
End-To-End Latency (ms)	Duration between the transmission of a data packet from the source node and the successful reception at destination node
Traffic Volume Density (bps/km^2)	Total data rate of all users per unit area
Mobility (km/h)	Relative speed between receiver and transmitter under certain performance requirement
Peak Date Rate (bps)	Maximum achievable data rate per user

Several problems are anticipated if today's networks are used to handle the explosive development of mobile Internet and IoT:

- The energy efficiency level, overall cost per bit and complexity of network deployment and maintenance cannot effectively handle 1000 times traffic growth and the massive number of connected devices.
- Co-existence of multiple radio access technologies (RAT) causes increased complexity and degraded user experience.
- Existing networks can not realize accurate monitoring of network resources and effective awareness of services, and therefore they cannot intelligently fulfill the diversified requirements of future users and services.
- Widely distributed and fragmented spectrum will cause interference and coexistence complexity.

To solve these problems, 5G should have the following capabilities to achieve sustainability. In terms of network construction and deployment, 5G networks need to:

- Provide higher network capacity and better coverage, while decreasing the complexity and cost of network deployment, especially the deployment of ultra dense networks.
- Have a flexible and scalable architecture to adapt to the diverse needs of users and services.
- Make flexible and efficient use of various spectrum resources, including paired and unpaired spectrum[1], re-farmed spectrum[2] and new spectrum, low-frequency and high-frequency bands, and licensed and unlicensed bands.
- Have stronger device-connection capabilities to deal with the access requirements of huge amounts of IoT devices.

In terms of operation and maintenance (O&M), 5G needs to:

- Improve network energy efficiency and the O&M cost-per-bit to cope with data traffic growth and the diverse needs of various services and applications.
- Reduce the complexity caused by the co-existence of multiple radio access technologies, network upgrades, and the introduction of new features and functions, to improve users' experience.
- Make intelligent optimization based on awareness of users behaviors and services contents.

---

[1] A block of spectrum in a lower frequency band and an associated block of spectrum in an upper frequency band. A number of frequency channels are paired with frequency channels in the other block. One frequency is transmitting from a base station to a fixed subscriber. A second frequency is used from the subscriber to the base station. The frequency pair is separated by a "duplex distance" to provide isolation of the two signals. If the two paired frequencies are too close to each other, the transmitter can overload the receiver. (http://www.telecomabc.com/p/paired-spectrum.html).

[2] A service can use portions of the spectrum initially allocated to a different service.

- Provide a variety of network security solutions to meet the needs of all types of devices and services of mobile internet and IoT.

Table I.2: 5G Efficiency indicators

Efficiency Indicator	Definition
Spectrum Efficiency (bps/Hz/cell or bps/Hz/ km^2)	Data throughput per unit spectrum per cell, or per unit area
Energy Efficiency (bit/J)	Number of bits than can be transmitted per unit energy
Cost Efficiency (bit per unit money)	Number of bits than can be transmitted per unit money

Spectrum utilization, energy consumption and cost are the three key factors which must be addressed in sustainable mobile communication networks. In order to achieve sustainability, 5G needs to make significant improvements in the following aspects:

- Spectrum efficiency: 3–5 times.
- Energy efficiency: 100+ times.
- Cost efficiency: 100+ times.

5G systems must dramatically outperform previous generation systems. 5G should support

- Peak data rate: tens of Gbps.
- Traffic volume density: tens of Gbps per square kilometer.
- Mobility: higher than 500 km per hour.
- **User experienced data rate: 0.1–1 Gbps.**
- **Connection density: one million connections per square kilometer.**
- **End-to-end latency: millisecond level.**

Table I.3: 5G Key capabilities and values

Key Capability	Value
Peak Data Rate	20 Gbps
User-Experienced Data Rate	0.1-1 Gbps
Latency	1 ms

Mobility	500 km/h
Connection Density	10^6/km^2
Energy Efficiency	100 times compared to IMT-Advanced
Spectrum Efficiency	3-5 time compared to IMT-Advanced
Area Traffic Capacity	10 Mbps/km^2

All key capabilities may to some extent be important for most use cases, and the relevance of certain key capabilities may be significantly different, depending on the use cases/scenario.

In the enhanced mobile broadband scenario, the user experienced data rate, area traffic capacity, peak data rate, mobility, energy efficiency and spectrum efficiency all have high importance, but mobility and the user experienced data rate would not have equal importance simultaneously in all use cases. For example, in hotspots, a higher user experienced data rate, but a lower mobility, would be required than in the wide area coverage case.

In the ultra-reliable and low latency communications scenarios, low latency is of highest importance, e.g., safety critical applications. Such capability would be required in some high mobility cases as well, e.g., transportation safety, while high data rates could be less important.

In the massive machine type communication scenario, high connection density is needed to support a tremendous number of devices in the network that may transmit only occasionally, at low bit rate and with zero/very low mobility. A low cost device with long operational lifetime is vital for this usage scenario.

I.6. Deployments Scenarios

Use cases will be delivered across a wide range of environments. To facilitate the study of 5G requirements and provide guidance to 5G technical design, several typical deployment scenarios need to be specified.

I.6.A. Indoor Hotspots

The indoor hotspot deployment scenario focuses on small coverage per cell and high user throughput or user density in buildings. The key characteristics of this deployment scenario are high capacity, high user density and consistent user experience indoor.

This scenario represents indoor offices with a total area of 120 m x 50 m. 12 small cells are deployed with an inter-site-distance (ISD) of 20 m. The BS antenna height is 3 m. The carrier frequency options include 4 GHz, around 30 and 70 GHz. The bandwidth for 4 GHz is up to 200 MHz. The bandwidth for around both 30 and 70 GHz is up to 1 GHz. 10 users per cell are distributed uniformly and all users are indoors with 3 km per hour velocity. Full buffer and/or burst traffic model is assumed.

I.6.B. Dense Urban

The dense urban heterogeneous deployment scenario focuses on macro cells with micro cells and high user densities and traffic loads in city centers and dense urban areas. The key characteristics of this deployment scenario are high traffic loads, outdoor and outdoor-to-indoor coverage. This scenario will be interference-limited.

The ISD for the macro cells is 200 m. There are 3 micro cells per macro cell. The macro BS antenna height is 25 m and micro BS antenna height is 10 m. The carrier frequency for macro cell is 4 GHz. The carrier frequencies for micro cell include 4 GHz, around 30 GHz and around 70 GHz. The bandwidth for 4 GHz is up to 200 MHz and bandwidth for both around 30 GHz and around 70 GHz is up to 1 GHz. 10 UEs are distributed per micro sector. 80% users are indoor with a moving speed of 3 km per hour and 20% are in cars with a velocity of 30 km per hour.

I.6.C. Urban Macro

The urban macro homogeneous deployment scenario focuses on large cells and continuous coverage. The key characteristics of this scenario are continuous and ubiquitous coverage in urban areas. This scenario will be interference-limited. The ISD in this scenario is 500 m. The BS antenna height is 35 m. The carrier frequency is 4 GHz, 2 GHz, and around 30 GHz. The bandwidth for 4 GHz is up to 200 MHz and bandwidth for 2 GHz is up to 100 MHz. Full buffer and/or burst traffic model is assumed. 10 UEs are distributed per cell. 80% users are indoor with 3 km per hour velocity and 20% are in cars with 30 km per hour velocity.

I.6.D. Rural

The rural deployment scenario focuses on larger and continuous coverage. The key characteristics of this scenario are continuous wide area coverage supporting high speed vehicles. This scenario will be noise-limited and/or interference-limited. The ISD in this scenario is 1732 m or 5000 m. The BS antenna height is 35 m. The carrier frequency is 700 MHz. The bandwidth is up to 20 MHz. Full buffer and/or burst traffic model is assumed. 10 UEs are distributed per cell. 50% users are indoors with 3 km per hour velocity and 50% are in cars with 120 km per hour velocity. 4 GHz and 2 GHz frequency are also considered in this scenario.

I.6.E. High Speed

The high speed deployment scenario focuses on continuous coverage along track in high speed trains. The key characteristics of this scenario are consistent user experience with very high mobility. In this deployment scenario, dedicated linear deployment along railway line is considered and UEs are located in train carriages.

I.6.F. Other Deployment Scenarios

I.7. Detailed Technical Requirements

Peak Data Rate

Peak data rate is the highest theoretical data rate, which is the received data rate assuming error-free conditions assignable to a single mobile station, when all assignable radio resources for the corresponding link direction are utilized (i.e., excluding radio resources that are used for physical layer synchronization, reference signals or pilots, guard bands and guard times).

The target for peak data rates are 20 Gbps for downlink and 10 Gbps for uplink.

Peak Spectral Efficiency

Peak spectral efficiency is the highest theoretical data rate (normalized by bandwidth), which is the received data rate assuming error-free conditions assignable to a single mobile station, when all assignable radio resources for the corresponding link direction are utilized (i.e., excluding radio resources that are used for physical layer synchronization, reference signals or pilots, guard bands and guard times).

The targets for peak spectral efficiency are 30 bps/Hz for downlink and 15 bps/Hz for uplink. Higher frequency bands imply higher bandwidths but lower spectral efficiency, while lower frequency bands results in lower bandwidths but higher spectral efficiency. Thus, the peak data rate cannot be directly derived from the peak spectral efficiency and bandwidth multiplication.

Bandwidth

Bandwidth means the maximal aggregated total system bandwidth. It may be supported by a single or multiple RF carriers. The value for this KPI is for further study.

Control Plane Latency

Control plane latency refers to the time to move from a battery efficient state (e.g., IDLE) to the start of continuous data transfer (e.g., ACTIVE). The target for control plane latency should be 10 ms.

User Plane Latency

Defined as the time it takes to successfully deliver an application layer packet from the layer 2/3 service data unit (SDU) ingress point[3] to the radio protocol layer 2/3 SDU egress point[4] via the radio interface in both uplink and downlink.

For ultra-reliable and low latency communications (URLLC), the target for user plane latency should be 0.5 ms for UL and 0.5 ms for DL. Furthermore, if possible, the latency should also be low enough to support the use of next-generation access technologies as a wireless transport technology that can be used within the next-generation access architecture. For other cases, the target for user plane latency should be 4 ms for UL and 4 ms for DL.

Latency for Infrequent Small Packets

For infrequent application layer small packet/message transfer, the time it takes to successfully deliver an application layer packet/message from the radio protocol layer 2/3 SDU ingress point at the mobile device to the radio protocol layer 2/3 SDU egress point in the RAN, when the mobile device starts from its most "battery efficient" state.

Mobility Interruption Time

Mobility interruption time means the shortest time duration supported by the system during which a user terminal cannot exchange user plane packets with any base station during transition. The target for mobility interruption time should be zero. This KPI is for intra-system mobility.

Inter-System Mobility

Inter-system mobility refers to the ability to support mobility between the IMT-2020 system and at least one IMT system.

Reliability

Reliability can be evaluated by the success probability of transmitting e.g., 20 bytes within 1 ms, which is the time it takes to deliver a small data packet from the radio protocol layer 2/3 SDU

[3] Ingress traffic is composed of all the data communications and network traffic originating from external networks and destined for a node in the host network. Ingress traffic can be any form of traffic whose source lies in an external network and whose destination resides inside the host network. Ingress traffic can be from all applications accessed via a remote server or over the Internet. Today, almost all internetworks are connected to some external network, the Internet or an autonomous system. The diverse connectivity and communication among them creates data packets flowing in and out of the network. Ingress traffic is all traffic initiated at a remote and external location or within a network not under the host network. Ingress traffic must also be directed toward a segment or node installed in the host network. Whenever users access a website, application or a utility over the Internet, the ingress traffic flows towards that user's system because that entity the user accessed is hosted on an external network. (https://www.techopedia.com/definition/2415/ingress-traffic).

[4] Egress traffic is any data or traffic bound for an external entity and passing through the edge router of the host network to reach its destination node. Egress filtering is a popular network management technique. It scans all egress traffic for any sign of abnormality or malicious activity and then discards any infected data packets. Egress traffic is a term used to define the volume and substance of traffic transmitted from the host network to an external network destination. Egress traffic also contains request packets. These are destined for an application hosted on a remote server, jeopardizing the integrity and availability of the network. To counteract this problem, all egress traffic is filtered. In the case of a security or throughput breach, the traffic is capped. Egress traffic filtering not only ensures that malicious packets do not leave the network, but also manages the flow of information by restricting useless traffic. (https://www.techopedia.com/definition/2398/egress-traffic).

ingress point to the radio protocol layer 2/3 SDU egress point of the radio interface, at a certain channel quality (e.g., coverage-edge).

Coverage

Maximum coupling loss (MCL) on the uplink and downlink between device and base station site for a data rate of 160 bps, where the data rate is observed at the egress/ingress point of the radio protocol stack in uplink and downlink. MCL includes in practice antenna gains, path loss, shadowing, etc. The target MCL is 164 dB.

UE Battery Life

The UE battery life is determined by the battery life of the UE without recharge. The UE battery life in extreme coverage depends on the activity of mobile originated data transfer consisting of 200 bytes UL per day followed by 20 bytes DL from MCL of 164 dB, assuming a stored energy capacity of 5 Wh. The target for UE battery life is beyond 10 years.

UE Energy Efficiency

UE energy efficiency means the capability of a UE to sustain much better mobile broadband data rate while minimizing the UE modem energy consumption.

Cell and Transmission/Reception Point (TRP) Spectral Efficiency

TRP spectral efficiency is defined as the aggregate throughput of all users (the number of correctly received bits, over a certain period of time) divided by the channel bandwidth divided by the number of TRPs. A three-sector site consists of 3 TRPs. In case of multiple discontinuous "carriers" (one carrier refers to a continuous block of spectrum), this KPI should be calculated per carrier. In this case, the aggregate throughput, channel bandwidth, and the number of TRPs on the specific carrier are used.

Area Traffic Capacity

Area traffic capacity means total traffic throughput served per geographic area (in Mbit/s/m^2). This metric can be evaluated by the full buffer model – all user equipment (UEs) always have downlink and uplink traffic loads – or the non-full buffer model. The area traffic capacity is a measure of how much traffic a network can carry per unit area. It depends on the site density, bandwidth and spectrum efficiency.

User Experienced Data Rate

User experienced data rate can be evaluated for non-full buffer traffic and for full buffer traffic. Non-full buffer simulations are preferred for the evaluation of this KPI. For non-full buffer traffic, the user experienced data rate is 5% of the user throughput. User throughput (during active time) is defined as the size of a burst divided by the time between the arrival of the first packet of the burst and the reception of the last packet of the burst.

Connection Density

Connection density refers to the total number of devices fulfilling specific QoS per unit area (per km^2). The target for connection density should be 1,000,000 device/km^2 in the urban environment.

Mobility

Mobility means the maximum user speed (km/h) at which a defined QoS can be achieved. The target for mobility is 500 km/h.

***

II. 5G NETWORK ARCHITECTURES

II.1. Requirements for Radio Access Network (RAN) Architecture

The RAN design for the Next Generation Radio Access Technologies has to fulfill the following requirements:

- The RAN architecture shall support tight interworking between the new radio access technology (RAT) and LTE.
- Considering high performing inter-RAT mobility and aggregation of data flows via at least dual connectivity between LTE and new RAT.
- The RAN architecture shall allow deployments using Network Function Virtualization.
- The design of the RAN architecture shall allow for rapid and efficient deployment of new services.

II.2. Ultra-Dense Networks

To meet user demands and to overcome the challenges of 5G systems, significant change in the strategy of designing 5G wireless networks is needed. New use cases such as high-resolution video streaming, tactile Internet, road safety, remote monitoring, and real-time control place new requirements related to throughput, end-to-end (E2E) latency, reliability, and robustness on the network. Diverse services such as wearable devices, connected cars, connected homes, moving robots, and sensors that must be supported in an efficient and scalable manner.

There are usually several ways to improve throughput in a wireless system:

- Increasing allocated bandwidth.
- Improving spectral efficiency through new coding and modulation technologies (air interface design).
- Cell splitting to improve spectrum reuse. This means increasing of the wireless access point (AP) density with smaller coverage per AP. An example of high AP density can be found in hotspot scenarios such as ultra dense networks (UDNs). Macro cellular and local small cell coexistence architecture in 5G will replace the existing macro cellular-dominated architecture.
- Multiple-input multiple-output (MIMO) transmission.

Ultra-dense small cell deployments and new technologies such as massive multiple-input multiple-output (mMIMO), software defined networking (SDN), and network function virtualization (NFV) are expected to provide an impetus to establish the fundamental design principles of 5G. In line with this, operators need to continuously enhance network capability and optimize its usage. Operators need to deploy more localized capacity, in the form of small cells (e.g., pico and femto cells and remote radio units (RRUs) that are connected to centralized baseband units by optical fiber) to improve capacity. In addition, traffic offloading to fixed networks through local area technologies such as Wi-Fi in unlicensed frequency bands is essential to meet the capacity and data rate challenges of 5G.

Wireless users access the network from indoors (homes, offices, malls, train stations, etc.) for approximately 80 percent of time, and from outdoors for approximately 20 percent of the time. In the current wireless cellular architecture, for a mobile user to communicate (whether from indoors or from outdoors), an outdoors base station is needed perform the communication. So, for indoors users to communicate with the outdoors base station, the signals will have to travel through the walls of buildings. This causes very high penetration losses. This, consequently, results in reduced spectral efficiency, lower data rates, and loss of energy efficiency of wireless communications.

A new design technique that has come into existence for scheming the 5G cellular architecture is to distinct outdoors and indoors setups. One of the key ideas of designing the 5G cellular architecture is to separate outdoor and indoor scenarios so that penetration loss through building walls can somehow be avoided. This idea will be supported with the help of distributed antenna system (DAS) and massive MIMO technology. Using massive MIMO, a geographically dispersed antenna array is deployed which has tens or even hundreds of antenna units. Present MIMO systems use either two or four antennas, because the antennas are installed within transmitters and receivers. MIMO architectures allow for huge capacity gains.

Outdoor BSs will be equipped with large antenna arrays with some antenna elements (also large antenna arrays) distributed around the cell and connected to the BS via optical fibers, benefiting from both DAS and massive MIMO technologies. Outdoor mobile users are normally equipped with limited numbers of antenna elements, but they can collaborate with each other to form a virtual large antenna array, which together with BS antenna arrays will construct virtual massive MIMO links.

Large antenna arrays will also be installed outside of every building to communicate with outdoor BSs or distributed antenna elements of BSs, possibly with line of sight (LoS) components. Large antenna arrays have cables connected to the wireless access points inside the building communicating with indoor users. This will certainly increase the infrastructure cost in the short term while significantly improving the cell average throughput, spectral efficiency, energy efficiency, and data rate of the cellular system in the long run.

Using such a cellular architecture, as indoor users only need to communicate with indoor wireless access points (not outdoor BSs) with large antenna arrays installed outside buildings, many technologies can be utilized that are suitable for short-range communications with high data rates. Some examples include WiFi, ultra wideband (UWB), mm-wave communications (3–300 GHz), and visible light communications (VLC) (400–490 THz). Note that mm-wave and VLC technologies use higher frequencies not traditionally used for cellular communications. These high-frequency waves do not penetrate solid materials very well and can readily be absorbed or scattered by gases, rain, and foliage. Therefore, it is hard to use these waves for outdoor and long distance applications. However, with large bandwidths available, mm-wave and VLC technologies can greatly increase the transmission data rate for indoor scenarios.

II.2-Ultra-Dense Networks

To solve the spectrum scarcity problem, besides finding new spectrum not traditionally used for wireless services (e.g., mm-wave communications and VLC), we can also try to improve the spectrum utilization of existing radio spectra, for example, via cognitive radio (CR) networks.

Project II.1

Prepare a 15-page report on cognitive radio communications.

The 5G cellular architecture should also be a heterogeneous one, with macrocells, microcells, small cells, and relays. To accommodate high mobility users such as users in vehicles and high speed trains, the mobile femtocell (MFemtocell) concept has been proposed, which combines the concepts of mobile relay and femtocell. MFemtocells are located inside vehicles to communicate with users within the vehicle, while large antenna arrays are located outside the vehicle to communicate with outdoor BSs. An MFemtocell and its associated users are all viewed as a single unit to the BS. From the user point of view, an MFemtocell is seen as a regular BS. This is very similar to the above idea of separating indoor (inside the vehicle) and outdoor scenarios. Users using MFemtocells can enjoy high data rate services with reduced signaling overhead.

With the requirements of sub-millisecond latency and bandwidth limitation in traditional wireless spectrum, cellular networks are now poised to break the Base Station (BS) centric network paradigm.

II.3. Ultra-Dense Network Concept

In general, UDN is a new wireless network solution for hotspot scenarios, to provide higher throughput and better user experience. In UDN, the AP density is comparable to or even higher than the user density. Different types of APs will tightly cooperate to achieve higher spectrum efficiency, lower power consumption and seamless mobility.

The main difference between UDN and traditional cellular network is the AP density. In UDNs, the radius of AP coverage is about 10 m and there are thousands of APs in 1 squared km. In traditional cellular networks, cell range is more than 500 m and there are usually less than 3-5 base stations (BSs) in 1 squared km. One or several terminals are usually connected to one UDN AP, whereas hundreds or even thousands of active users can be resident in one macro cell.

Another key point is that the type of APs in UDN is diversified. Small cell station, relay station, distributed Remote Radio Head (RRH) and user equipment (UE) itself can act as AP in UDN. However the macro BS in traditional cellular network is the dominant access point type.

Table II.1: Comparison between UDN and the traditional cellular network

Item	UDN	Traditional Cellular Network
Deployment Scenario	Indoor, Hotspot	Wide coverage

AP Density	More than 1000/km^2	3-5/km^2
AP Coverage	Around 10 m	Several hundreds of meters or more
AP Types	Small Cell, Pico, Femto, Relay, UE	Macro/Micro BS
User Density	High	Low/Medium
User Mobility	Low	High
Traffic Density	High	Low/Medium
Deployment	Heterogeneous, irregular	Single layer, regular
System Bandwidth	Hundreds of MHz	Tens of MHz
Spectrum	> 3 GHz	< 3 GHz

Typical scenarios of UDN include: office, apartment, open-air gathering, stadium, subway and railway station. The common requirements in these scenarios are huge numbers of connections, high density network traffic and high data rate. To meet these requirements, the APs need to be deployed densely, with minimum inter-site distances (ISDs) of ten meters or below, i.e. one or more APs per room for indoor scenarios, and one AP on each lamp post for outdoor scenarios.

In order to provide very high traffic density and better user experience in UDN, there are many new challenges including network architecture, mobility management, interference management and others.

II.3.A. Challenges of Network Architecture

Traditional cellular network architectures are designed for wide area seamless coverage. There are many problems to using traditional cellular network architecture (e.g. 4G architecture) in UDN.

Signaling overhead and lengthy data path

Too many functions such as service control and mobility control are centralized in the Core Network (CN), which include Mobility Management Entity (MME) and Packet Data Network Gateway (PGW)/Serving Gateway (SGW). This is not efficient for UDN with high data traffic and ultra-dense AP deployment, and it leads to signaling overhead and lengthy data transmission path between APs and CN.

Frequent handover

In a communication network, the control plane provides the connectivity and mobility while the user plane provides the data transport. The tight coupling of user plane and control plane over one air interface will result in frequent handover when the AP coverage is very small. It is not efficient and flexible in heterogeneous networking within both macro and UDN AP coverage.

All distributed functions

Higher layer process, Radio Resource Management (RRM), mobility management functions are distributed on each AP independently. In order to better support advanced interference management and resource management for UDN, the functions on each distributed AP needs to be centralized.

Better user experience

UDN is targeted to provide smooth handover and very high data rate for each user with ultra dense APs. The simple data gathering and transmission function of Local GateWay (LGW) cannot support better user experience. More functions are needed in LGW.

Therefore, a new architecture of UDN is needed to support high density AP deployment and flexible network management. In this new architecture, a local centralized user service center is necessary for knowing and measuring user's radio environment. Besides, the RRM and user service control center much closer to the user are required to provide better joint processing and Quality of Service (QoS) control, and a lower mobility anchor is also needed. Meanwhile, the CN functions should be simplified to provide only high level service to user.

II.3.B. Challenges of Mobility Management

With very small coverage and irregular network topology, mobility management in UDN is quite different from traditional cellular network. In the scenarios that ultra-dense APs are deployed, the coverage radius of these APs is extremely small, i.e. only a few meters to tens of meters. According to traditional handover judgment method, frequent handover may occur and the interruption probability of the user's experience of high data rate may increase. Moreover, from the perspective of the network side, it also means a high handover signaling overhead. How to solve the problem is essential for UDN.

II.3.C. Challenges of Interference Management

Huge number of APs may bring much higher throughput and better user experience, but may also lead to problems. Interference management directly impact the system performance. Along with the multiplexing of resources for more access opportunities, interference increases as well and becomes more complex related to traditional cellular network. We need to solve the following problems:

- The ultra-density environment results in more interference sources. For example, in the crowded subway trains, lots of terminals and APs exist, and signals may have more reflected and scattering paths. Then transmission power should increase, which turns out more interference than before, however.

- Decreasing interference and increasing the resource utility are in contradiction and we need to find the proper tradeoff point.
- The existing parameters to evaluate interference impact such as interference threshold may not reflect the overall performance of the network, and then more suitable parameters should be used to give a better indication between interference managing results and throughput, energy efficiency and other system level parameters.

Therefore, we should set up suitable interference models, analyze the typical wireless transmission scenario, and then propose effective interference control approach for UDN.

II.3.D. Challenges of Flexible Networking

With extreme densification and complex heterogeneous deployment, network planning and optimization will become difficult for UDN. It is critical to further enhance Self Organization Network (SON) to support flexible networking. Huge numbers of APs in UDN make it more complex to realize self configuration, self-optimization and self-healing. Ultra high throughput, ultra low latency, ultra high reliability, massive connections are required to be provided in UDN. On another side, UDN is a very complex network to cover both indoor and outdoor scenarios. From view of the access technologies, there are 5G new access technologies, LTE access technologies, WLAN access technology working together. So, flexible network architecture with intelligent network sensing and automotive network optimization is very important to flexible networking and increasing spectrum efficiency in UDN. Based on the above analysis, new network architecture for UDN is an important direction. Key technologies towards flexible networking, self-backhauling, multi-RAT coordination, advanced interference management, advanced mobility management and radio resource management are also essential for UDN.

Project II.2

Prepare a 15-page report on heterogeneous wireless access in 5G.

In addition, due to different radio scenarios and higher spectrum and wider bandwidth, new transmission technologies and new air interface designs are also needed to further enhance the system performance for UDN. Millimeter wave communication is a very attractive direction for UDN, which can provide Gbps user experience and tens of Gbps throughput per AP. And it’s also very suitable for wireless backhauling for UDN APs.

Project II.3

Prepare a 15-page report on sustainable green HetNets.

Project II.4

Prepare a 15-page report on cooperative communication networks.

Project II.5

Prepare a 15-page report on cloud and fog computing.

Project II.6

Prepare a 15-page report on offloading in 5G.

Project II.7

Prepare a 15-page report on network architecture of UDNs.

Project II.8

Prepare a 15-page report on 5G RAN architectures.

Project II.9

Prepare a 15-page report on user-centric wireless networks in 5G.

Project II.10

Prepare a 15-page report on energy harvesting in 5G networks.

Project II.11

Prepare a 15-page report on resource management in 5G networks.

***

III. 5G Spectrum Strategies

III.1. Introduction

In 5G vision, the spectrum issue is one of the most important parts. Governments, agencies, standardization organizations and research institutions from many countries pay high attention to the 5G spectrum strategies.

According to current research, generally, spectrum below 6 GHz is the best resource in the near future. However, due to its scarcity and increasing difficulty to realize international harmonization, it is time to seek spectrum above 6 GHz. The 5G spectrum solution will be comprehensive, indicating combinations of different frequency ranges for different scenarios.

Currently allocated spectrum for IMT is an important part of 5G candidate spectrum. Most existing frequency ranges for mobile communications are below 3 GHz. The amount of spectrum that is currently used for mobile communications is several hundred megahertz.

III.2. 5G Spectrum Demand

How would one calculate IMT spectrum demands? Generally, one would need to

- analyze (and forecast) future market and traffic volume,
- calculate and distribute the traffic on different RATs,
- calculate the required capacity,
- estimate the demanded spectrum.

The actual calculation process can be very complicated, when there are a variety of traffic types, different environments and multiple cell types of different RATs. For example, imagine how to estimate the data rates of a high quality video streaming user located in indoor offices, connecting with future 5G small cells, in the year 2025.

There are differences in the markets and deployments and timings of the mobile data growth in different countries. Therefore, two settings are developed to characterize lower and higher user density settings. These two sets of market study input parameter values have to be considered in the calculations to characterize differences in the user densities in different countries. In some countries, national spectrum requirement can be lower than the estimate derived by lower user density settings and in some other countries, national spectrum requirement can be higher than the estimate derived by higher user density settings.

The areas which have the largest spectrum demand are always in metropolises. When requirements of such a scenario are satisfied, the solution might be applicable to others as well. For example, the total IMT spectrum demand in China in 2020 is 1350-1810 MHz.

The obtained estimation results are huge. Compared with current identified spectrum for IMT in different countries, there are still several hundred to even 1000 MHz deficit. Moreover, in the period of 5G commercialization after 2020, the total spectrum demand may continue to increase.

III.3. Candidate Frequency Bands

III.3.A. BELOW 6 GHZ

Spectrum re-farming

In order to realize the benefits of new generation of mobile communication systems, regulators need to deliver efforts to enable re-farming of spectrum in the frequency band occupied by the old ones.

III.3.B. ABOVE 6 GHZ

Traditionally, spectrum above 6 GHz such as millimeter wave (mmWave) is often used for point-to-point communications in large-power systems such as satellite systems and microwave systems. From the perspective of public mobile communication, mmWave communication technology also attracts attention all over the world as one of the potentially essential technologies of the next generation of mobile communication systems.

Due to the characteristics of high frequency, the key technologies applied for low frequency are hard to directly put into use. How to utilize the advantage and overcome the disadvantage are both opportunities and challenges. Preliminary studies show that a lot of potential suitable frequency ranges could be found from 6 to 100 GHz.

III.3.C. SPECTRUM FOR IOT

Compared with previous generations of mobile communications, 5G needs to meet extremely high performance requirements in more diverse scenarios. Besides scenarios such as seamless wide-area coverage and high-capacity hot-spot, there is also a need of low-power massive-connections including machine-to-machine and man-to-machine, driven by internet of things.

Low-power massive-connection scenario mainly targets sensing and data collecting use cases, such as environmental monitoring and intelligent agriculture. This scenario is characterized by small data packets, low power consumption, low cost, and massive connections. Specifically, probably at least one million connections per squared kilometer need to be supported. The suitable frequency ranges of such 5G usage scenarios will be wide and rather different from those for communication between people.

III.4. Spectrum Sharing

Spectrum sharing for 5G is motivated by the fact that significantly more spectrum and much wider bandwidths than what is available today will be needed in order to realize the performance targets of 5G. So far, 2G, 3G and 4G cellular networks have been designed for dedicated, licensed spectrum. With that, cellular technology can only operate in a subset of the spectrum that is in principle available.

Significantly more spectrum and much wider bandwidths than what is available today will be needed in order to realize the performance targets of 5G. One way is to try the best to fill the spectrum gap by finding new spectrum, especially going to higher frequency band which is less crowded than what is being used today. The other way is to improve total spectrum efficiency by spectrum sharing. Generally speaking, it is expected that the spectrum for 5G will, at least partly, be available under spectrum sharing regimes.

III.4.A. Spectrum Sharing Scenarios

- Vertical sharing refers to spectrum sharing between users of different priority (e.g. primary and secondary), i.e., unequal rights of spectrum access.
- Horizontal sharing is sharing between systems that have the same priorities in the spectrum, i.e. different users have fair access rights to the spectrum. If the sharing users in the spectrum adopt the same technology, it is called homogenous horizontal sharing; otherwise heterogeneous horizontal sharing.

Licensed Mode

In licensed mode, spectrum is allotted to wireless systems as primary use. The only relevant spectrum sharing scenario is homogenous horizontal sharing, i.e. sharing spectrum with other operators using the same RAT, which is also called co-primary or inter-operator spectrum sharing. There are two different ways to achieve horizontal sharing, i.e. mutual renting and limited spectrum pool.

- In the mutual renting sub-scenario the spectrum resources in a band are subdivided into several blocks and each block is licensed to one operator. Operators mutually allow other operators to "rent" parts of their licensed resources. An operator can rent resources from multiple other operators simultaneously. The actual "owner" of a resource has always strict priority in accessing its licensed part of the spectrum, including the possibility of preemption at any time. This approach could be applied to bands that have initially been used in a dedicated licensed way where there is a need to increase peak data rate beyond what is possible within one licensed block.
- The limited spectrum pool scenario allows an operator to obtain an authorization, usually a license, to use up to the whole band on a shared basis with a limited number of other known authorized users. This setup does not provide guarantee for instantaneous access to a minimum amount of spectrum, but it is envisioned that mutual agreements between licensees are such that the long term share of an individual operator has a predictable minimum value. This is similar to the sharing situation in unlicensed bands but it is a priori known how many authorizations a regulator will give out and what the rules for the band will be, hence providing the necessary certainty for investments into large networks.

Licensed Shared Access (LSA) Mode

In LSA mode, a licensee has the authorization to access spectrum that is unused by its owner at certain locations and/or times. This vertical sharing is based on well-defined conditions which are parts of a sharing license. Standardization activities on LSA technical conditions and geo-location databases are ongoing.

Unlicensed Mode

In unlicensed mode, a wireless system has to share spectrum with other unlicensed systems. For heterogeneous horizontal sharing in an unlicensed band a system must be prepared for coexistence with any other technology that may be present in the band. License-exempt spectrum creates a low barrier to market entry.

Unrestricted access to the bands creates unpredictable interference conditions that typically lead to a decrease in accessible capacity when traffic levels are high.

III.4.B. Spectrum Sharing Techniques

There are several different ways to technically realize the above spectrum sharing and to implement corresponding regulatory requirements. Different kinds of spectrum sharing techniques are mapping to the most possible application scenarios.

Coordination Protocol

Coordination protocol means explicit information exchange between the sharing systems via some well-defined interface to avoid interference in sharing spectrum. The protocol defines the behavior of the nodes when receiving certain messages or when certain events take place. This includes the possibility for defining behavior depending on what relation nodes have with each other, e.g. interference relation. Both centralized and distributed structure can be applied to coordination.

The connections between networks or between network and central manager can be wired or over the air. The advantage of this approach is that it can achieve more efficient spectrum sharing with enough information on the interference between different networks by employing smart resource allocation. One obvious drawback is the mandatory need of connection between different networks or to central manager for multiple networks, which may not be easy to have in most scenarios.

Geolocation Database (GLDB) Support

In this solution the system queries a database to acquire information on what resources are available (i.e., unused by other systems of typically higher priority) for operation in its current location. The geolocation database provides rules for operating in the geographical region of the system. A rudimentary geolocation database would provide a list of available channels and possibly associated transmits power limitations to a querying system for a specified geographical region.

The GLDB approach has been considered as an appropriate solution to manage unlicensed access to TV white spaces while ensuring that the bands may, in the future, be vacated of unlicensed transmissions. In this case this solution does not provide exclusive use of the spectrum to any system. The presence of a centrally controlled resource manager opens up for the possibility of future repurposing of the bands, since the behavior of the radios the database controls can be influenced, including the possibility to prevent transmissions completely. In this way a band could be opened up for usage not compatible with the database-controlled devices. For example, 802.22 uses such approach to enable coexistence with TV systems. It is assumed that all devices must be equipped with satellite technology (e.g. GPS) which provide positioning and also gives time reference for inter-BS synchronization. Before communication, UE needs to inform BS the location and possibly antenna pattern, height and etc. Then BS queries database for allowed channels and powers. Finally BS will communicate with corresponding UE with received configuration limit.

III.4.C. Cognitive/Dynamic Spectrum Access (DSA)

Cognitive radio is an automatic machine that uses software defined radio to change frequency band to adapt and sense the surrounding environment. They recognize radio spectrum by sensing when it is unused by the incumbent radio system and use this spectrum in an intelligent way.

There are two different types of systems present in situations in which cognitive radios are used to dynamically access the spectrum:

- A primary user on a portion of the spectrum is a user who owns a license that legally permits the user to use the spectrum for a purpose specified by the license.
- A secondary user is a user that accesses spectrum that is licensed by a primary user. This access has typically to be regulated in such a way that the secondary user does not cause harmful interference to the license holder.

There are two important functions to achieve good spectrum sharing between primary user and secondary user:

- Spectrum sensing function available in secondary systems to sense status of primary systems to find spectrum opportunity.
- Dynamic spectrum access (DSA) function available in secondary system to control spectrum access dynamically according to varying environment of primary systems.

Spectrum Sensing

The main challenge of spectrum sensing is the strict requirement of sensitivity. In order to incorporate the noise uncertainty and multi-path fading/shadowing uncertainty into the specification, the sensing node requires identifying the presence of primary signal in the low signal to noise ratio (SNR) region. Typical spectrum sensing algorithms are as follows:

- ***Energy detection***: Measure the radio frequency energy or the received signal strength indicator over certain observation time and compare the accumulated metric to a predefined threshold.

Advantages	Easily implemented without prior knowledge.
Disadvantages	Sensitive to the noise and interference uncertainty.
	Can’t distinguish primary signal from other interference.

- ***Coherent detection***: Detect known pilot signal of primary systems with matched filter and compare obtained metric to a predefined threshold. This technique requires knowledge of the primary signal.

Advantages	Can distinguish primary signal from interference and noise.
	Works well in low SNR.
	Shorter observation time.
Disadvantages	More implementation complexity by performing timing, carrier sync and equalization.

- ***Covariance-based detection***: Estimate sample covariance matrix of the received signal and make spectrum whiteness tests from eigenvalue-based or covariance statistic-based statistics. Primary signal correlated due to dispersive channel, correlation of multiple receiver antennas or over-sampling effect.

Advantages	Can exploit more general feature of primary signal without signal-specific knowledge.
Disadvantages	Performance highly depends on correlation of primary signal.

- ***Cyclostationary detection***: Analyze the cyclic autocorrelation function or its equivalent Fourier transformation spectrum correlation to seek cyclostationary feature. Cyclostationary feature in primary signal caused by modulation and coding, hopping sequence, cyclic prefix.

Advantages	Capability of differentiating the primary signal from the interference and noise.
Disadvantages	More implementation complexity with wide scan of the cyclic frequencies.

In general, the above schemes are oriented for local spectrum sensing in one node. However, if the primary signal is in deep fading or is obstructed by a large obstacle, the received power at secondary node is too weak to detect. Thus hidden node problem occurs in this situation. This uncertainty can be effectively mitigated by spatial diversity in the cooperative sensing. Additionally, cooperative sensing can relax the sensitivity requirement of local sensing and increase the agility of making sensing decision.

III.4-Spectrum Sharing

The procedure of cooperative sensing is as follows: all cognitive radios that have agreed to participate in the cooperative sensing process sense the spectrum and communicate their results to other cognitive radios. Based on the information type sent by the individual CRs, cooperative sensing can be categorized by soft decision fusion and hard decision fusion.

Hard decision fusion is the merge of the information to produce a common final decision for the whole system if the information that it has received consists of individual preliminary decisions;

Soft decision fusion means what it receives is channel statistics or raw data to perform the procedure of creating the final decision.

Soft decision fusion outperforms hard decision fusion at the expense of large signaling overhead to forward the sensing data. As an alternative, the quantized decision fusion is desirable to achieve the tradeoff between these two issues. It is shown that even a simple cooperative scheme, such as the OR-rule, may improve the performance of the cognitive secondary system significantly. The probability of false detection can be made arbitrarily small while maintaining the level of interference protection of the primary system. This is achieved by adding more cooperating cognitive users that are experiencing uncorrelated fading or shadowing. Considering only few cooperating uncorrelated users give considerable improvement in the sensitivity, compared to the case without cooperation.

***

IV. Massive MIMO

IV.1. Introduction

Massive MIMO can bring at least ten-fold improvements in area throughput by increasing the spectral efficiency (bit/s/Hz/cell), while using the same bandwidth and density of base stations as in current networks. A key goal of the 5G technologies is to improve the area throughput by orders of magnitude; 100x and even 1000x higher throughput are regularly mentioned as 5G design goals.

The area throughput of a wireless network is measured in bit/s/km^2 and can be modeled as follows:

$$\begin{aligned}\text{Area throughput (bit/s/km}^2) = &\text{ Bandwidth (Hz)}\\ &\text{x Cell density (cells/km}^2)\\ &\text{x Spectral efficiency (bit/s/Hz/cell)}\end{aligned} \tag{IV.1}$$

There are three main components that can be improved to yield higher area throughput:

(1) more bandwidth can be allocated for 5G services
(2) the network can be densified by adding more cells with independently operating access points
(3) the efficiency of the data transmissions (per cell and for a given amount of bandwidth) can be improved.

IV.2. MIMO Communications

The spectral efficiency (SE) of a single-input single-output (SISO) communication channel, from a single-antenna transmitter to a single-antenna receiver, is upper bounded by the Shannon capacity. Shannon capacity has the form

$$\begin{aligned}C &= W\log_2\left(1+\frac{P}{N_0 W}\right)\\ &= W\log_2\left(1+\bar{\gamma}\right)\end{aligned} \tag{IV.2}$$

where C is the channel capacity in bits/s, W is the transmission bandwidth in Hz, P is the signal power in Watts and $N_0/2$ is the AWGN power spectral density. Note that the maximum possible SE is given by

$$\frac{C}{W} = \log_2\left(1+\bar{\gamma}\right) \tag{IV.3}$$

The SISO capacity is a logarithmic function of the average signal-to-noise ratio (SNR). To improve the SE we need to increase the SNR, which corresponds to increasing the power of the transmitted signal.

Example IV.1

Suppose we have a system that operates at an SE of 2 bit/s/Hz and we would like to double its SE to 4 bit/s/Hz.

From (IV.3), $\bar{\gamma} = 3$.

To obtain an SE of 4 bit/s/Hz, $\bar{\gamma}$ has to be increased to 15. This corresponds to improving the SNR by a factor 5.

(**verify this**): The next doubling of the SE, from 4 to 8 bit/s/Hz, requires another 17 times more power.

In other words, the logarithm of the SE expression forces us to increase the transmit power exponentially fast to achieve a linear increase in the SE of the SISO channel.

This is clearly a very inefficient and non-scalable way to improve the SE, and the approach also breaks down when there are interfering transmissions in other cells that scale their transmit powers in the same manner. We therefore need to identify another way to improve the SE of cellular networks.

Each base station (BS) in a cellular network serves a multitude of user terminals. Traditionally, the time/frequency resources have been divided into resource blocks and only one of the user terminals was active per block. This terminal can then receive a single data stream with an SE quantified as $\log_2\left(1+\bar{\gamma}\right)$. The efficient way to increase the SE of a cellular network is to have multiple parallel transmissions. If there are G parallel and independent transmissions, the sum SE becomes $G\log_2\left(1+\bar{\gamma}\right)$. Parallel transmissions can be realized by having multiple transmit antennas and multiple receive antennas. There are two distinct cases:

- Point-to-point MIMO, where a BS with multiple antennas communicates with a single user terminal having multiple antennas.
- Multi-user MIMO, where a BS with multiple antennas communicates with multiple user terminals, each having one or multiple antennas.

There are many reasons why multi-user MIMO is the most scalable and attractive solution:

- The wavelength is 5–30 cm in the frequency range of cellular communication (1–6 GHz). This limits the number of antennas that can be deployed in a compact user terminal for point-to-point MIMO, while one can have almost any number of spatially separated single-antenna terminals in multi-user MIMO.
- The wireless propagation channel to a user terminal is likely to have only a few dominating paths, which limits the ability to convey multiple parallel data streams to a terminal in point-to-point MIMO. The corresponding restriction on multi-user MIMO is that the users need to be, say, a few meters apart to have sufficiently different channel characteristics, which is a very loose restriction that is true in most practical scenarios.

- Advanced signal processing is needed at the terminals in point-to-point MIMO to detect the multiple data streams, while each terminal in multi-user MIMO only needs to detect a single data stream.

The canonical multi-user MIMO system consists of a BS with M antennas that serves K single-antenna terminals. The BS multiplexes one data stream per user in the downlink and receives one stream per user in the uplink.

The BS uses its antennas to direct each signal towards its desired receiver in the downlink, and to separate the multiple signals received in the uplink. If the terminal is equipped with multiple antennas, it is often beneficial to use these extra antennas to mitigate interference and improve the SNR rather than sending multiple data streams.

Note that $\min(M,K)$ represents the maximal number of data streams that can be simultaneously transmitted in the cell, while still being separable in the spatial domain. The number $\min(M,K)$ is referred to as the multiplexing gain of a multi-user MIMO system.

Project IV.1

Prepare a 10-page report, explaining the structures of differences between the following MIMO configurations:

- spatial multiplexing
- spatial diversity

IV.3. Massive MIMO Concept

Massive MIMO takes multi-user MIMO communications to a new level by designing a highly scalable communication protocol. Massive MIMO is a multi-user MIMO system with M antennas and K users per BS. The system is characterized by $M >> K$ and operates in time-division duplex (TDD) mode using linear uplink and downlink processing.

There are two ways of implementing the downlink and uplink transmission over a given frequency band.

- In frequency division duplex (FDD) mode the bandwidth is split into two separate parts: one for the uplink and one for the downlink.
- In time-division duplex (TDD) mode the whole bandwidth is used for both downlink and uplink transmission, but separated in time.

If the system switches between downlink and uplink faster than the channels are changing, then it is sufficient to estimate the channels in only one of the directions. When $M >> K$, we note that TDD systems should send pilots only in the uplink and the pilot length becomes independent of the number of BS antennas. We conclude that TDD is the preferable mode.

Project IV.2

Prepare a 10-page report on TDD and FDD systems.

Example IV.2

- $M = 100 - 200$, $K = 1 - 40$.
- $M = 100$, $K = 10$.

The BS antenna array typically consists of M dipole antennas, each having an effective size $\lambda/2 \times \lambda/2$, where λ is the wavelength.

Example IV.3

Consider a square antenna array area of 1 m^2.

If the carrier frequency is 1.5 GHz, then the wavelength is 0.2 m $\Rightarrow$ $\lambda/2 = 0.1$ m.

The total number of accommodated antennas is $10 \times 10 = 100$.

If the carrier frequency becomes 3 GHz, then the total number of accommodated antennas is $20 \times 20 = 400$.

Each antenna is attached to a separate transceiver chain, so that the system can access the individual received signal at each antenna and select the individual signals to be transmitted from each antenna. The array can have any geometry; such as linear, rectangular, cylindrical, and distributed arrays.

The basic massive MIMO transmission protocol is illustrated in Figure IV.1. The time–frequency resources are divided into blocks of size B_c Hz and T_c s, where B_c is the channel coherence bandwidth, and T_c is the channel coherence time. Within the coherence bandwidth, the channel is flat (frequency non-selective), with basically no multipath. Within the coherence time, the channel is slowly fading (time non-selective), with basically no Doppler frequency shift.

Each block is referred to as a coherence interval. Let the coherence interval be used to transmit N_c symbols. The symbol rate is then equal to $R_c = N_c/T_c$. According to the Nyquist-Shannon sampling theorem (such that intersymbol interference can be avoided), we should have

$$R_c \le B_c \tag{IV.4}$$

Therefore, the number of transmission symbols that fit into a coherence interval is given by bounded according to

$$\frac{N_c}{T_c} \le B_c \Rightarrow N_c \le B_c T_c \tag{IV.5}$$

0:

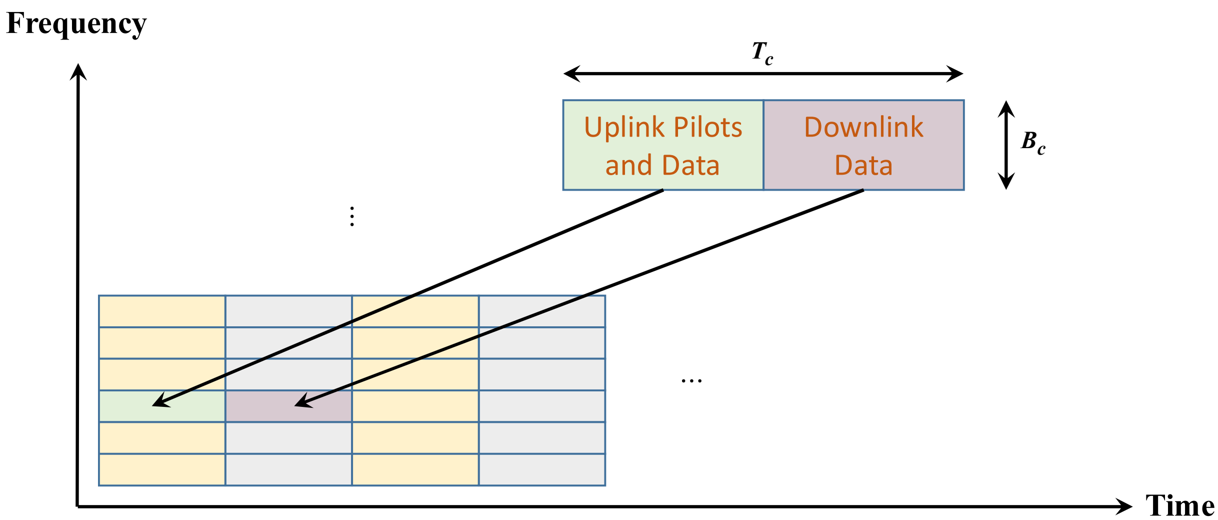

Figure IV.1: Basic massive MIMO transmission protocol

Example IV.4

Let $B_c = 200$ kHz and $T_c = 1$ ms . Then

$$N_c \le 200$$

Such a situation can arise in cases of highway user velocities in urban environments at 2 GHz carrier frequencies.

Note that increasing the coherence bandwidth means increasing the flatness of the channel, while increasing the coherence time makes the channel more static.

Example IV.5

Let $B_c = 2000$ kHz and $T_c = 10$ ms . Then

$$N_c \le 2 \times 10^4$$

Such a situation can arise in cases of low user mobility and short delay spread.

Note that the coherence bandwidth is inversely proportional to the delay spread. The coherence time is inversely proportional to the Doppler spread.

***

V. MILLIMETER-WAVE MOBILE COMMUNICATIONS

V.1. Introduction

Recently, mmWave communication has been investigated because of its ultra broad spectrum. Typical mmWave frequencies range from 6 GHz to 100GHz including 6 GHz, 15 GHz, 28 GHz, 38GHz, 60 GHz, and E-band (71–76GHz, 81–86GHz). From channel measurements, the propagation loss for mmWave transmission is quite large. Its transmission range is thus limited. Therefore, mmWave transmission is more suitable for small cells and dense user scenarios. Owing to these features, mmWave transmission is tailored for a hotspot scenario and has become a promising candidate for the fifth generation.

Channel propagation property is a fundamental topic of mmWave communications. The channel model will affect the spectrum allocation as well as system design and performance evaluation standards.

Since the wavelengths of millimeter wave frequencies are very small, it can utilize polarization and different spatial processing techniques like massive MIMO and adaptive beamforming. With the significant increase in bandwidth, the data links in densely populated areas will now handle greater capacity than present 4G networks. Likewise the base stations are constantly reducing the coverage areas of the cell for spatial reuse, cooperative MIMO, relays and interference mitigation between base stations. Since the base stations are abundant and more densely dispersed in urban areas, this will reduce the cost per base station. Spectrum distributions of over 1 GHz of bandwidth are currently being utilized in the 28 GHz and 38 GHz bands.

As far as building a prototype is concerned, the antenna is essentially being positioned in very close vicinity to the 28 GHz Radio Frequency Integrated Circuit and the front end module; because there will be high signal attenuation at 28 GHz. Realizing the antenna array directly on the printed circuit board of the 5G cellular device will minimize the loss between the antenna and Radio Frequency Integrated Circuit. This implies that an employment of the Radio Frequency blocks in the 5G architecture before the intermediate frequency stage will be reliant on the placement of the 28 GHz antenna array in the cellular phone. Taking this concept into a thought, a minimum set of two 28 GHz antenna arrays can be used for millimeter wave 5G cellular applications, with the two antenna arrays placed in the top and bottom part of the cellular device.

The millimeter wave spectrum is under-utilized and has been left mostly idle until present years. The main reason behind the underutilization is its unsuitability for cellular communications. This is mainly because of unfriendly channel conditions like path loss effect, absorption due to atmosphere and rain, small diffraction and penetration through objects. Others reasons of unsuitability include strong phase noise and excessive apparatus costs.

The large unlicensed band around 60 GHz is appropriate primarily for very short range transmission. So, the emphasis had been given to both fixed wireless applications in the 28, 38, 71-76 and 81-86 GHz and WiFi with the 802.11ad standard in the 60 GHz band. Semiconductors are evolving, as their costs and power consumption values are decreasing rapidly due to the growth of the above-mentioned short range standards.

V.2. Millimeter Wave Propagation

The main propagation issues regarding millimeter wave propagation for 5G cellular communication are:

Path Loss

The free space path loss is dependent on the carrier frequency, as the size of the antennas is kept constant. The product of carrier frequency f_c and wavelength λ is equal to the speed of light c, i.e.,

$$f_c\lambda = c \tag{V.1}$$

When non-isotropic antennas are used with a transmit antenna gain of G_t and a receive antenna gain of G_r, the free space path loss is usually given in the form

$$\frac{P_r(d)}{P_t} = \frac{G_t G_r \lambda^2}{(4\pi)^2 d^2 L} \tag{V.2}$$

where P_t represents the transmit power (watts) and L is the system loss factor which is independent of the propagation environment. The system loss factor represents overall attenuation or loss in the actual system hardware, including transmission line, filter, and antennas. In general, $L > 1$, but $L = 1$ if we assume that there is no loss in the system hardware.

Note that in (V.2), it is assumed that apertures of both transmit and receive antennas are kept fixed regardless of λ. Note that in this case the free-space path loss decreases with f_c^2. If the apertures of both antennas are not to be kept fixed and the carrier frequency increases, the size of the antenna gets reduced and the effective aperture increases by a factor of $\lambda^2/4\pi$. The free space path loss between a transmitter and a receiver antenna grows with f_c^2. So, if we increase the carrier frequency from 3 to 30 GHz, this will add 20 dB of power loss irrespective of the transmitter-receiver distance. If the antenna aperture at one end of the link is kept constant, then the free-space path loss remains unchanged.

Blocking

Microwave signals are less prone to blockages but it deteriorates due to diffraction. On the contrary, millimeter wave signals suffer less diffraction than the microwave signals and exhibit specular propagation (one ray, as opposed to diffuse propagation which means a multitude of rays). Note that millimeter wave signals have shorter wavelengths than microwave signals. This makes mm-waves much more vulnerable to blockages. Recent studies reveal that, with the increase in the transmitter and receiver distance the path loss increases by 20 dB/decade under Line of sight propagation, but decreases by 40 dB/decade plus an added blocking loss of 15-40 dB for non-line of sight propagation. So due to the presence of blockages, the connection will promptly shift from usable to unusable which will results in large scale impediments that cannot be avoided with typical small scale diversity countermeasures.

Atmospheric and Rain Absorption

Within the unlicensed 60-GHz band, the absorption due to rain and air particularly the 15 dB/km oxygen absorption are more perceptible. But these absorptions are insignificant for the urban cellular deployments, where base station spacings might be on the order of 200 m. But actually, these types of absorption are useful as they efficiently increase the segregation of each cell by further attenuating the background interference from more distant base stations.

From the above explanation, it is inferred that the propagation losses for millimeter wave frequencies are resolvable, but only by steering the beam energy with the help of large antenna arrays and then collecting it coherently. But for practical viability, the concept of narrow beam communication is fresh for cellular communications and poses problems like:

- **Link Acquisition** - The main problem that the narrow beams are facing is in establishing links amid users and base stations for both initial access and handoff. The user and base stations will have to locate each other by scanning lots of angular positions where the possibility of a narrow beam is high. This problem poses an important research challenge predominantly in the perspective of high mobility.
- **Need of New Transceiver Architectures** - Wireless millimeter wave systems have gone through significant improvement but still there are some hardware issues which will affect the designing of the communication systems. The analog to digital and digital to analog converters needed for large bandwidths are the prime cause of power consumption. A prime reason of power consumption is because of the use of large antenna arrays. Along with these, high receiver sensitivities are needed to deal with the path loss because it is not feasible that each antenna will be provided with normal fully digital beam formers.

Project V.1

Prepare a 10-page report on channel impairments in mm-Wave transmission systems.

***

VI. INTERNET OF THINGS (IOT)

VI.1. Overview

The idea of an Internet of Things as a network of smart devices dates far back in the past, with the first applications for automated inventory systems coming as early as 1983. However, only from 1999 it took momentum, becoming part of a shared vision for the future of Internet. Today, the growing pervasiveness and ubiquity, in almost any context, of small and cheap computing devices, endowed with sensing and communication capabilities, is paving the way to the realization of the IoT vision.

A large variety of communication technologies has gradually emerged, reflecting a large diversity of application domains and of communication requirements. Some of these technologies are prevalent in a specific application domain such as Bluetooth Low Energy in Personal Area Networks, and Zigbee in Home Automation systems. In addition, such landscape is constantly and rapidly evolving, with new technologies being regularly proposed, and with existing ones moving into new application domains.

A rough distinction is emerging between consumer IoT (cIoT) and industrial IoT (iIoT), with clear implications on underlying technologies and business models. Consumer IoT aims at improving the quality of people's life by saving time and money. It involves the interconnection of consumer electronic devices, as well as of (virtually) anything belonging to user environments such as homes, offices, and cities.

Conversely, industrial IoT focuses on the integration between Operational Technology (OT) and Information Technology (IT) and on how smart machines, networked sensors, and data analytics can improve business-to-business services across a wide variety of market sectors and activities, from manufacturing to public services. It generally implies machine-to-machine interactions, either for application monitoring (e.g., process monitoring in chemical production plants, vehicle fleet tracking, among others), or as part of a self organized system, with a distributed control which does not require human intervention (i.e., autonomic industrial plants).

Despite their evident differences, these two service domains share some general communication requirements, such as scalability, need for lean protocol stack implementations in constrained devices, and friendliness to the IP ecosystem.

Nonetheless, the specific communication requirements of iIoT and cIoT can be very different, in terms of reliability, QoS (latency, throughput, etc), and privacy. cIoT communications are typically machine-to-user, and usually in the form of client-server interactions. In cIoT, desirable features of networked things are low power consumption, ease of installation, integration and maintenance. Indeed, the quantified self paradigm which is currently unfolding with the advent of fitness and health tracking systems, smart watches and sensor rich smartphones requires a high power efficiency, in order to enable long term monitoring by small, portable devices, as part of a ”smart” environment or integrated in our daily wearings. At the same time, such applications need to minimize the risk of exposing such sensitive data as someone's health status or life habits.

Increasing the number of nodes and of exchanged information clearly multiplies the potential vulnerabilities of the system to attacks and to privacy leaks.

Differently from cIoT, iIoT evolves from a large base of systems employing machine to machine communications for control process automation and/or monitoring. In such domains, iIoT is the result of the integration, through the Internet, of hardwired and often disconnected islands, usually based on semi-proprietary protocols and architectures. Such integration magnifies the potential of isolated industrial plants by augmenting their flexibility and manageability, and disclosing the opportunity to deploy new services.

Many of iIoT communications, together with some of cIoT communications, have often to satisfy stringent requirements in terms of timeliness and reliability. Typically, the information exchanged is critical for ensuring a correct and safe behavior of the processes under control. Hence, the communication network must be engineered in order to:

- meet stringent delay deadlines;
- be robust to packet losses;
- be safe and resilient to damages, and more generally, strike the desired balance between capital expenditure / operational expenditure (CAPEX/OPEX) costs and system / service availability.

The advent of 5G communications presents great opportunities for IoT. The increased data rate, reduced end-to-end latency, and improved coverage with respect to 4G hold the potential to cater for even the most demanding of IoT applications in terms of communication requirements. Its support for large amounts of devices enables the vision of a truly global Internet of Things. In addition, for its focus on the integration of heterogeneous access technologies, 5G may play the role of a unified interconnection framework, facilitating a seamless connectivity of ”things” with the Internet.

Nowadays, the IoT landscape includes an extreme diversity of available connectivity solutions which need first to be harmonized across multiple industries, and then properly combined together in order to meet the IoT technical Key Performance Indicators (KPIs).

First forms of IoT connectivity can be dated back to the 80s, with the legacy Radio Frequency Identification (RFID) technologies, and to the 90s, with the Wireless Sensor Networks (WSNs). Due to their attractive application scenarios, both in business and consumer markets, they gained a lot of momentum. Therefore, for the first decade of the 21st century, industrial alliances and Standards Developing Organizations (SDOs) put a lot of effort in developing standardized low power IoT solutions.

No technology has emerged as a market leader, mainly because of technology shortcoming, and business model uncertainties. Now, the IoT connectivity field is at a turning point with many promising radio technologies emerging as true M2M connectivity contenders: Low-Power WiFi, Low-Power Wide Area (LPWA) networks and several improvements for cellular M2M systems. These solutions are very attractive for IoT deployments, being able to fulfill availability and reliability requirements. With the final aim of helping the understanding of this rich and variegated context the reminder of this section overviews the modern IoT connectivity landscape and characterizes in more details the technologies which would potentially have a decisive impact in enabling a global IoT in the upcoming future.

Project VI.1

Prepare a 10-page report on internet of things security.

Project VI.2

Prepare a 10-page report on use of internet of things in smart cities.

Project VI.3

Prepare a 10-page report on use of internet of things in virtual reality applications.

Project VI.4

Prepare a 10-page report on ubiquitous and pervasive computing.

Project VI.5

Prepare a 10-page report on internet of things electromagnetic radiation issues.

Project VI.6

Prepare a 10-page report on Bluetooth networks.

Project VI.7

Prepare a 10-page report on ZigBee networks.

Project VI.8

Prepare a 10-page report on body area networks.

VI.2. IoT Enablers

In order to enable the ubiquitous connectivity required for many of the IoT applications, many more features and functionalities will need to be added to the currently predominantly broadband approach. This inherently leads to a strong heterogeneous networking (HetNet) paradigm with multiple types of wireless access nodes (with different MAC/PHY, coverage, backhaul connectivity, QoS design parameters, among others).

HetNets will offer the required seamless connectivity for the emerging IoT through a complex set of mechanisms for coordination and management. Emerging 5G networks will thus be characterized by interoperability and integration between multiple radio access networks, including unlicensed frequencies.

VI.2.A. 4G-Evolution Feature Enhancements

Narrowband Operation

Work is ongoing on to specify a narrowband version of LTE named narrowband IoT (NB-IoT) especially targeting machine-type communication (MTC) applications with low data rate and need for low module cost and long battery life time. The details of the air interface design are still to be settled and the higher layer protocols starting from Layer 2 and upwards will be common between NB-IoT and LTE. By making the solution similar to LTE the large eco-system for LTE modules can be mobilized to secure availability of device chipsets and a fast rollout. The performance requirements for NB-IoT are similar as for the wideband LTE MTC solution, the main benefit lies in that the narrowband operation leads to flexibility in the deployment.

Low Cost and Enhanced Coverage

Device Power Saving

The main source of the energy consumption for earlier MTC devices is the periodical listening for possible paging messages. For typical MTC traffic patterns, the energy required for transmitting messages is only a small fraction of the total energy consumed. A mechanism to reduce the power consumption for MTC devices is called the Power Saving Mode (PSM). In PSM, the device can turn off all functionality requiring the device to listen while in idle mode. Consequently the device does not receive paging messages or perform link quality measurements in PSM. Since the device remains attached to the network, less signaling is required compared to the approach where the device would be completely shut off when not transmitting. The device can transmit uplink data at any time but is only reachable in downlink when the device has been active in uplink, which happens at configurable time instances or when the device has transmitted uplink data.

VI.2.B. 4G-Evolution and 5G RAT Enablers

Relaying for Increased Coverage

Relaying is a key technology for 5G systems. In its classic formulation, it is meant for extending the communication range of a Base Station (BS) and improve throughput by means of Relay Stations (RSs).

Millimeter Wave Technologies

Device-to-Device Communications

Device to Device (D2D) communications represents a turning point in cellular systems. They entail the possibility that two devices can exchange data without the involvement of the BS or with just a partial aid from the BS. In contrast to WiFi and Bluetooth technologies, which provide D2D capabilities in the unlicensed band, with D2D communications the Quality of Service (QoS) is controllable because of the use of the licensed spectrum.

***

VII. Device to Device (D2D) communications

VII.1. Introduction

D2D communication in cellular networks is defined as direct communication between two mobile users without traversing the Base Station (BS) or core network. In a traditional cellular network, all communications must go through the BS even if both communicating parties are in range for D2D communication. This architecture suits the conventional low data rate mobile services such as voice call and text message in which users are not usually close enough to have direct communication. The intelligence at the user equipment (UE) side will be exploited in 5G networks to support D2D connectivity, which is mainly driven by the inherent and strong motivation for operators to offload traffic from the core network.

However, mobile users in today's cellular networks use high data rate services (e.g., video sharing, gaming, proximity-aware social networking) in which they could potentially be in range for direct communications. Hence, D2D communications in such scenarios can highly increase the spectral efficiency of the network. Nevertheless, the advantages of D2D communications is not only limited to enhanced spectral efficiency. In addition to improving spectral efficiency, D2D communications can potentially improve throughput, energy efficiency, delay, and fairness. The realization of D2D connectivity represents a real step forward for operators to reduce transmission delay and improve energy efficiency, particularly for networks (e.g., social networking) supporting proximity-based services.

The majority of the literature on D2D communications propose to use the cellular spectrum for both D2D and cellular communications (i.e., underlay inband D2D). These works usually study the problem of interference mitigation between D2D and cellular communication. In order to avoid the aforementioned interference issue, some propose to dedicate part of the cellular resources only to D2D communications (i.e., overlay inband D2D). Here resource allocation gains utmost importance so that dedicated cellular resources be not wasted. Other researchers propose to adopt outband rather than inband D2D communications in cellular networks so that the precious cellular spectrum be not affected by D2D communications. In outband communications, the coordination between radio interfaces is either controlled by the BS (i.e, controlled) or the users themselves (i.e., autonomous). Outband D2D communication faces a few challenges in coordinating the communication over two different bands because usually D2D communication happens on a second radio interface (e.g., WiFi Direct and Bluetooth).

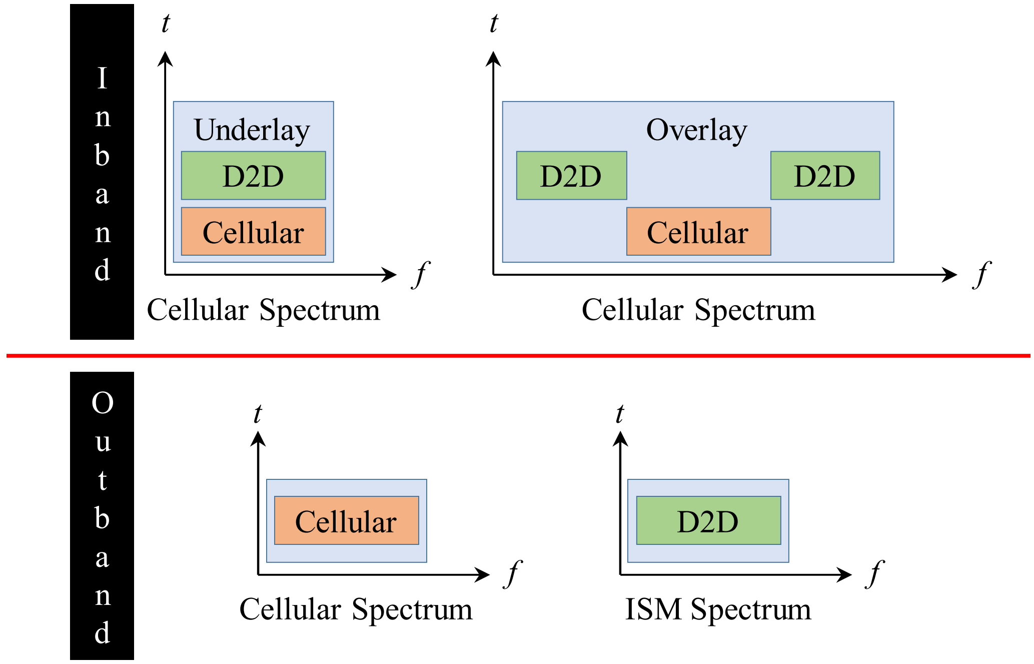

Figure VII.1: Spectral allocation for inband and outband D2D

VII.2. Classifications of D2D Communications

D2D users can communicate with each other in three manners: D2D direct link, relay assisted D2D communications, and clustering-based communications.

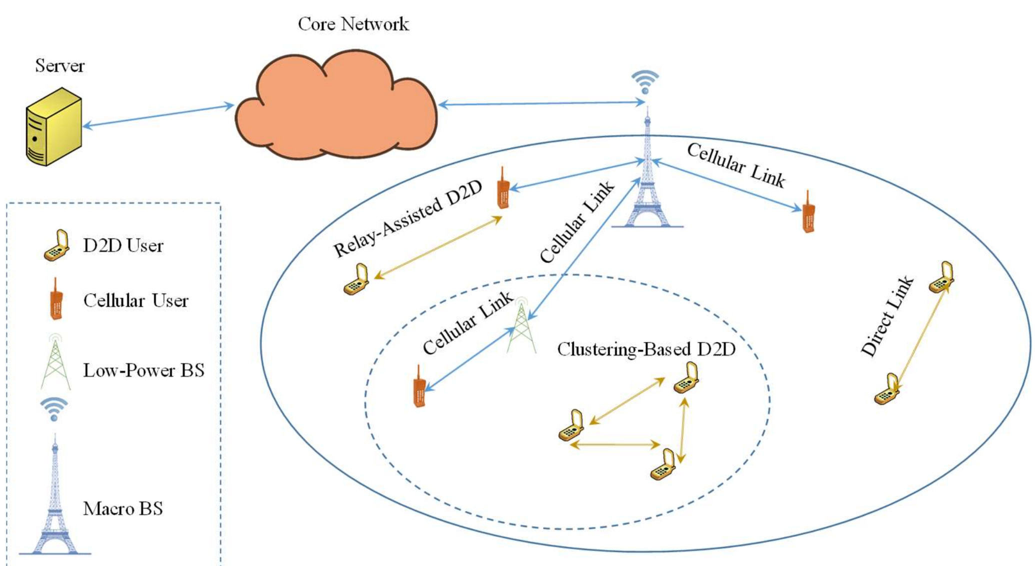

Figure VII.2: Classification of D2D

- ***D2D Direct Link:*** The simplest case of D2D communication occurs when transmitters and receivers exchange data directly with each other without intermediate nodes.
- ***Relay Assisted D2D Communications:*** Given a scenario where a mobile device wants to connect to another node which is out of its communication coverage or is in a poor channel state with the destination node, cellular users may be employed as relays to improve the data transmission between transmitters and receivers.
- ***Clustering-Based Communications:*** In a content sharing or information diffusion scenario, users requesting the same file in a short range can potentially form a cluster to allow the desired file to be multicast within the cluster to save both bandwidth and time delay. Moreover, users with similar interests or with tight social relationship (as illustrated by the tendency to "follow" one another) can also form a cluster to share contents with one other via D2D communications.

VII.3. Inband D2D

The literature under this category, which contains the majority of the available work, proposes to use the cellular spectrum for both D2D and cellular links. The motivation for choosing inband communication is usually the high control over cellular (i.e., licensed) spectrum. The interference in the unlicensed spectrum is uncontrollable which imposes constraints for QoS provisioning. Inband communication can be further divided into underlay and overlay categories.

In underlay D2D communication, cellular and D2D communications share the same radio resources. In contrast, D2D links in overlay communication are given dedicated cellular resources. Inband D2D can improve the spectrum efficiency of cellular networks by reusing spectrum resources (i.e., underlay) or allocating dedicated cellular resources to D2D users that accommodates direct connection between the transmitter and the receiver (i.e., overlay). The key disadvantage of inband D2D is the interference caused by D2D users to cellular communications and vice versa. This interference can be mitigated by introducing high complexity resource allocation methods, which increase the computational overhead of the BS or D2D users.

VII.3.A. Underlaying Inband D2D

Early works on D2D in cellular networks propose to reuse cellular spectrum for D2D communications. To date, the majority of available literature is also dedicated to inband D2D, especially D2D communications underlaying cellular networks.

Spectrum Efficiency

By exploiting the spatial diversity, underlaying inband D2D is able to increase the cellular spectrum efficiency. Interference between the cellular and D2D communications is the most important issue in underlaying D2D communications. Good interference management algorithms can increase the system capacity, and have attracted a lot of attention.

Usually, uplink cellular resources are used for D2D communications. Since reusing uplink resources for D2D users can cause interference to cellular uplink transmissions at the BS, D2D users monitor the received power of downlink control signals to estimate the path loss between

D2D transmitter and the BS. This helps the D2D users to maintain the transmission power below a threshold to avoid high interference to cellular users. If the required transmission power for a D2D link is higher than the minimal interference threshold, the D2D transmission is not allowed.

Power Efficiency

Power efficiency enhancement techniques for D2D-enabled cellular networks is also a very interesting research topic. Sometimes, a heuristic algorithm for power allocation in OFDMA-based cellular networks is used. This is done by performing power allocation along with subcarrier and bit allocation algorithms. The heuristic first allocates the resources for the cellular users and then performs resource allocation for D2D users. If the required power level for D2D transmission is higher than a certain threshold, the D2D pair communicates through the BS. Simulations of this method show that the integration of the proposed heuristic with subcarrier and bit allocation improves the downlink power consumption of the network around 20% in comparison to the traditional OFDMA system without D2D.

QoS/Power Constraints

There are many works which focus on improving the system performance while maintaining certain QoS/power constraints. Resource allocation is sometimes used for D2D communication underlaying cellular network, to guarantee QoS requirements for both D2D and cellular users. The resource allocation problem formulation is a non-linear constraint optimization problem. The problem is split into three subproblems. First, the BS checks the feasibility of the D2D connection based on the SINR requirements (admission control). Next, optimal power control is formulated for the D2D pair. Finally, a maximum-weight scheme is used for resource allocation for cellular and D2D users.

VII.4. Outband D2D

Here the D2D links exploit unlicensed spectrum. The motivation behind using outband D2D communication is to eliminate the interference issue between D2D and cellular link. Using unlicensed spectrum requires an extra interface and usually adopts other wireless technologies such as WiFi Direct, ZigBee or bluetooth. Some of the works on outband D2D suggest to give the control of the second interface/technology to the cellular network (i.e., controlled). In contrast, others propose to keep cellular communications controlled and leave the D2D communications to the users (i.e., autonomous). Outband D2D uses unlicensed spectrum which makes the interference issue between D2D and cellular users irrelevant. On the other hand, outband D2D may suffer from the uncontrolled nature of unlicensed spectrum. It should be noted that only cellular devices with two wireless interfaces (e.g., LTE and WiFi) can use outband D2D, and thus users can have simultaneous D2D and cellular communications.

VII.5. D2D-Assisted Cellular Communication

D2D communications using cellular spectrum under the control of cellular infrastructure enable devices to communicate directly without intermediate nodes. As mobile applications exploiting the proximity of mobile devices continue to gain popularity, D2D communications in cellular networks also attract more and more attention owing to their ability to share data at higher speed via local links. The emergence of D2D communications in cellular networks promises multiple performance benefits. Generally, D2D can support communications at higher bit rates, lower delays, while consuming less energy. For example, two adjacent devices can usually achieve better performances when communicating directly by setting up a D2D link rather than seeking relay help from the BS.

In particular, it is more resource-efficient for adjacent devices to communicate directly with one other as opposed to routing data through a BS or even the core network. In other words, compared to the traditional cellular communications which occupy both uplink and downlink resource, D2D communications save energy and improve spectrum efficiency. Furthermore, switching from an infrastructure path to a direct path may offload the cellular and backhaul traffic, alleviate network congestion, thereby benefiting other users in the network.

VII.6. Research Challenges in D2D-Assisted Networks

To facilitate local area services, D2D communications present a promising technology in future wireless networks to improve network capacity and user experience. Mobile network operators can derive substantial benefit by leveraging the coordinated and network-assisted D2D technologies.

VII.6.A. Synchronization

In an LTE network-assisted D2D scenario, two user equipment (UEs) of a D2D pair are synchronized with the BS, implying that slot and frame timing as well as frequency synchronization are acquired and aligned. Synchronized D2D transmissions are appealing because continuous searching for discovery signals is required in asynchronous discovery schemes. However, UEs can be active only during predetermined time slots for receiving discovery related signals in the process of time synchronized device discovery, which consumes significantly less energy.

However, it is challenging to realize synchronization between two UEs because of at least two reasons

- Two UEs that form a D2D link pair may be associated with different BSs that are not synchronized.
- Even for UEs that are located in the same cell, the distance from different UEs to the BS may be different, and the application of different timing advance adjustments may be required.

Therefore, further study is required to investigate the impact of timing misalignment on system performance, and additional synchronization methods are needed if the impact turns out to be non-negligible.