UNIVERSITÀ DEGLI STUDI DI PADOVA

Master Thesis in
ICT For Internet and Multimedia, Cybersystems

Narrowband IoT: from the end device to the cloud. An experimental end-to-end study.

7th October 2019
Accademic Year 2018/2019

Author: Giulia Calzavara, 1163079
Advisor: Prof. Lorenzo Vangelista
Coadvisor: Ing. Mauro Franchin

"The future belongs to those who believe in the beauty of their dreams."

E. Roosevelt

To Federico and my family,

my Advisor Lorenzo and Mida Solutions' people.

Contents

Introduction

This thesis is about a novel study and experimentation of a Cloud IoT application, communicating over a NB-IoT Italian network.

So far there no been presented studies, which are about the interactions between the NB-IoT network and the cloud. This thesis not only fill this gap but also shows the use of Cognitive Services to interact, through the human voice, with the IoT application. Moreover, in this thesis other types of mobile networks, like 2G, 3G and 4G are considered and analyzed. However, in terms of coverage and support of battery constrained devices, the NB-IoT is confirmed to be the best.

The thesis is organized as follow:

- in Chapter 1 is presented an overview of the IoT world, from the network protocols to the devices;
- in Chapter 2 all the existing Cellular IoT technologies are reviewed;
- in Chapter 3 the NB-IoT protocol is described;
- in Chapter 4 a the description of the Cloud Computing infrastructure, used to add a form of artificial intelligence to the IoT application developed for this thesis, is provided;
- in Chapter 5 a detailed description of the IoT application is provided;
- in Chapter 6 conclusions are drawn.

Chapter 1

Overview

The Internet of Things, referred as IoT, is a branch of Information and Communication Technology (ICT) industry, that aims to virtually connect anything with everything. This process has started when the first cellular phone data connection service was created: the connection between the phone and the Internet, has put a milestone in the telecommunication scenario. Nowadays the frontier of the cellular connectivity includes anything that provides an environment measurement, e.g. smart metering. A recent example is the Smart Metering Implementation Program developed by the British government, which has decided to replace the British meters with ones with near-real-time information on their energy consumption. Another example is the category of wearables, like smartwatches or clothing-sensors; for the second type introduced, the Wiliot semiconductor company has developed in January 2019 the first battery-free Bluetooth® sensor tag, that is able to send beacon via Bluetooth from wallets, bicycles and so on to other ones. To do this, they provide General Data Protection Regulation policies [19] to improve the confidentiality of such a data exchange.

1.1 3GPP and its organization

A key enabler of the connection and the communication between things is the global standards development organization called Third Generation Partnership Project, 3GPP, which is responsible for the standardization of Global System for Mobile Communications (GSM, 2G), Universal Mobile Telecommunications System (UMTS, 3G), Long Term Evolution (LTE, 4G) and 5G radio access technologies. It was founded in 1998 with the aim of creating a single organization coordinating seven regional telecommunications Standard Development Organizations (SDOs) such as ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC. It provides their members with a stable environment to produce the Reports and Specifications that define the 3GPP technologies related to the cellular telecommunications network, like radio access, core transport network and service capabilities, with the aim of providing complete system specifications.

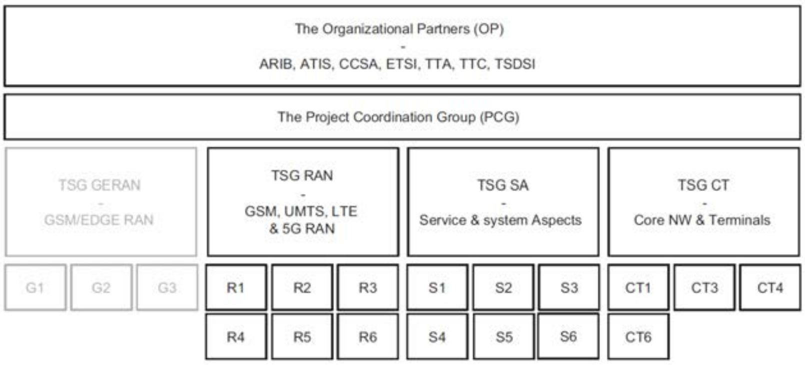


Figure 1.1: Organizational structure of 3GPP.

3GPP organizes its work in releases cycles, as explained in [17]. A release contains a set of work items where each typically delivers a feature that is made available to the cellular industry at the end of the release cycle through a set of Technical Specification, TSs. A feature, different for each version of a release, is specified in four stages:

- Stage 1: for service requirements;
- Stage 2: high-level feature description;
- Stage 3: detailed description;
- Stage 4: development of the performance requirements and conformance testing procedures to ensure proper implementation of the feature.

When the release cycle is finished, a specific version of the specifications used for feature development is published. In each release it is important to ensure that GSM, UMTS, and LTE can coexist in the same geographical area, without interference.

The technical work is well distributed across Technical Specification Groups, TSGs; each group is supported by a set of Working Groups, WGs with technical expertise representing different stakeholders (usually companies, manufacturing or service providers and administration) in the industry. In the figure 1.1 the organizational structure of 3GPP is shown, where, for LTE release 13, the structure was built around four TSGs:

- TSG Service and System Aspects (SA);
- TSG Core network and Terminals (CT);
- TSG GSM/EDGE Radio Access Network (GERAN);
- TSG Radio Access Network (RAN).

The project is then handled by the Project Coordination Group, PCG, and, above it, there are the seven SDOs already cited: ARIB and TTC for Japan, CCSA for China, ETSI in Europe, ATIS in the US, TTA for Korea and TSDSI in India. Within 3GPP, these seven SDOs are called Organizational Partners, OPs.

1.2 Internet of Things

Internet of Things (IoT) is the neologism which means that device becomes smarter, processing becomes intelligent and communication becomes informative. The IoT refers to the interconnections between billion of devices in various contexts, such as: transportation, smart cities, smart domotic, smart health, e-governance, assisted living, e-education, retail, logistics, agriculture, automation, industrial manufacturing, business management and process management. Those connections facilitate the analysis of the data to be processed, the management and the intelligent decision-making in an fashion autonomous.

Referring to [1], IoT architecture can be treated as a system that can be physical, virtual or a hybrid of the two, and consists of a collection of active physical things, actuators, sensors, cloud sensors, specific IoT protocols, communication layers, users, developers and an enterprise layer.

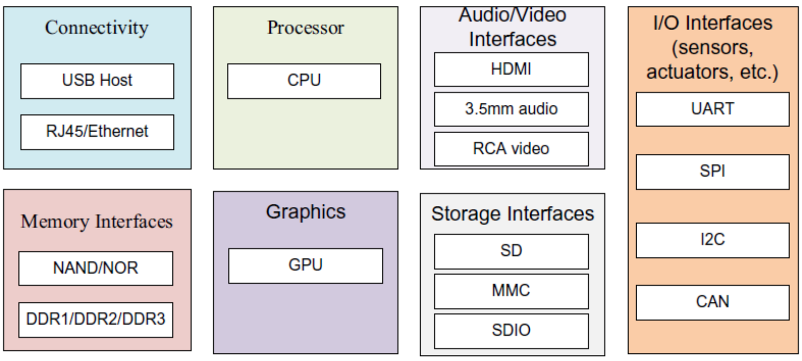


Figure 1.2: IoT device components.

In figure 1.2, the functional blocks for an IoT system are shown:

- **Devices**: they provide sensing, actuation, control and monitoring activities. They exchange and collect data with other connected devices and process the results locally or they send them to centralized servers or cloud. Devices have some interfaces to communicate with other devices, such as I/O interfaces for sensors, for Internet connectivity, for memory and storage, for audio and video.

- **Communication**: it performs the communication between devices and remote servers. It is regulated by a protocol, which usually encompass in data link layer, network layer, transport layer and application layer.

- **Services**: to serve various type of functions such as device modeling, device control, data publishing, data analytics and device discovery.
- **Management**: it provides functions to govern an IoT system.
- **Security**: it secures the IoT system with authentication, authorization, privacy, message integrity, content integrity and data security.
- **Application**: this is the most important functional block, because it acts as an interface that provides the necessary modules to control and monitor a lot of aspects and it also provides the analysis of the system status and it can predict future prospects.

The **Communication Protocol** is the backbone of the IoT systems, because it enables network connectivity, and of course data exchange, for the application. Such protocols define the format of the data exchange, the encoding of the data, the addressing schemes for devices and routing packets from source to destination. Also they provide sequence and flow control and re-transmission of packets that are lost [1]. A comparison between the existing communication protocol is explained in the table 1.1.

Parameters	**WiFi**	**WiMAX**	**LR-WPAN**	**Mobile communication**	**Bluetooth**	**LoRa**
Standards	IEEE 802.11 a/c/b/d/g/n	IEEE 802.16	IEEE 802.15.4 (ZigBee)	2G-GSM, CDMA, 3G-UMTS, CDMA2000, 4G-LTE (NB-IoT)	IEEE 802.15.1	LoRaWAN R1.0
Frequency band	5-60 GHz	2-66 GHz	868/915 MHz 2.4 GHz	from 865 MHz	2.4 GHz	868/900 MHz
Data rate	1 Mb/s-6.75 Gb/s	1 Mb/s-1Gb/s (fixed) 50-100 Mb/s (mobile)	40-250 Kb/s	2G: 50-100 kb/s 3G: 200 kb/s 4G: 0.1-1 Gb/s	1-24 Mb/s	0.3-50 Kb/s
Transmission range	20-100 m	<50 Km	10-20 m	Entire cellular area	8-10 m	<30 Km
Energy consumption	High	Medium	Low	Medium	Medium	Very Low
Cost	High	High	Low	Medium	Low	High

Table 1.1: Comparison between communication technologies

802.11 Its commercial name is Wi-Fi. It includes Wireless Local Area Network (WLAN) communication standards. There are various types of standards, which provide data rates from 1 Mb/s to 6.75 Gb/s, with communication range in the order of 20 m indoor and 100 m outdoor. Some standards are explained in the table 1.2.

802.16 - WiMax It is part of the wireless broadband standards. The name WiMax means Worldwide Interoperability for Microwave Access, and provides data rate from 1.5 Mb/s to 1 Gb/s.

802.15.4 - LR-WPAN It is a collection of Low-Rate Wireless Personal Area Network standards. These forms the basis for specification of high level communications protocols (e.g. Zigbee). The 802.15.4 provides data rates from 40 Kb/s to 250 Kb/s and also ensures low-cost and low-speed communication to power constrained devices.

Standard	**Band**
802.11a	5 GHz
802.11b	2.4 GHz
802.11g	2.4 GHz
802.11n	2.4/5 GHz
802.11ac	5 GHz
802.11ad	60 GHz

Table 1.2: Comparison between WiFi standards

The frequency in which it operates are 868/915 MHz for low data rates and 2.4 GHz for high data rates.

2G/3G/4G - mobile communication The second generation of mobile communication standards includes GSM and CDMA, the third generation includes UMTS and CDMA2000. Regarding the fourth generation, it includes LTE. The IoT devices, that relies on such standards, can communicate over cellular networks.

802.15.1 - Bluetooth™ Bluetooth is a low cost wireless communication technology that is used for data transmission between mobile devices over short range (of 8-10 m) and it operates in 2.4 GHz band, with a data rate of 1 Mb/s up to 24 Mb/s. This standard defines a Personal Area Network (PAN) communication.

LoRaWAN® - LoRa™ LoRaWAN is a long range communication protocol designed by the LoRa Alliance. LoRaWAN is a Low Power Wide Area (LPWA) networking protocol, designed to connect battery operated devices to the internet in regional, national or global wireless networks. It targets the key requirements of the IoT such as bi-directional communication, end-to-end security, mobility and localization services. The LoRaWAN data rates range from 0.3 kb/s up to 50 kb/s and it operates in 868 European bands and 915 MHz USA bands. In unobstructed environment the communication between nodes via LoRaWAN takes place within 20 miles range and the battery lifetime of such nodes is guaranteed up to 10 years.

Cloud Regarding the application part, the Cloud helps to capture data in real-time, to visualize them, to analyze such data in order to make future decisions and manage devices, while implying a "pay-as-you-go" formula. The 10 application domains on which the IoT cloud platforms are currently evolving in the IT market are: application development, device management, system management, heterogeneity management, data management, analytics, deployment management, monitoring management, visualization and research.

Communication Protocols There are a lot of communication protocols that are been studies to build IoT networks. For example, Near Field Communication (NFC) is the technology used to establish near-communication (e.g. devices must be at distance of at most 4 centimeters to communicate). An example of NFC is the contactless payment, which allow mobile payment. Another example is LTE-A, which is the Long Term Evolution Advanced and it indicates various technologies which increment the 4G mobile network performance. In table 1.3 it is also cited uCode, that is an identification number system that can be used to identify IoT devices in the real world.

In figure 1.3 there is a representation of real-application that can be done in the IoT vision.

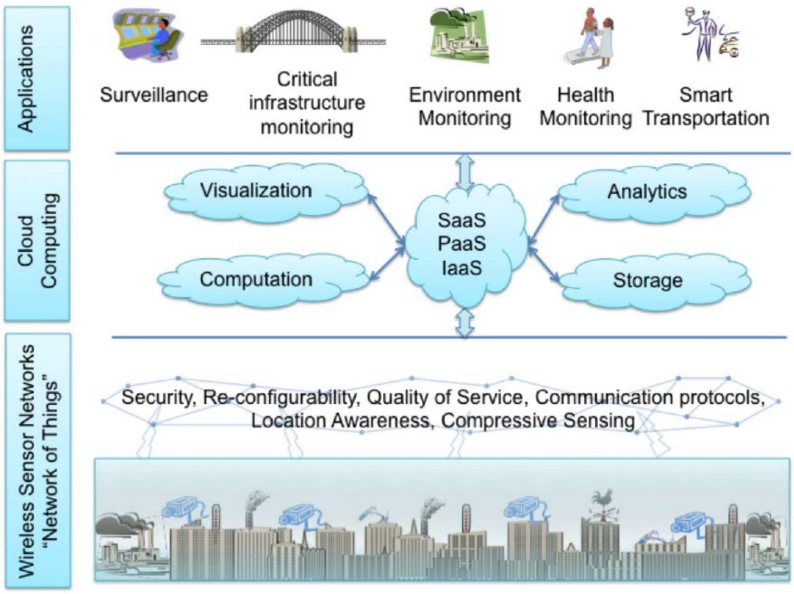


Figure 1.3: The IoT applications

In summary, referring to [3], the "Internet of Things" is composed by six elements, that are represented in figure 1.4 and their examples are described in table 1.3.

Figure 1.4: The IoT elements.

IoT Elements		*Samples*
Identification	*Naming*	EPC, uCode
	Addressing	IPv4, IPv6
Sensing		Smart Sensors, Wearable, sensing devices, Embedded sensors, Actuators, RFID tag
Communication		RFID, NFC Bluetooth, IEEE 802.15.4, WiFi, LTE-A
Computation	*Hardware*	SmartThings, Arduino, Phidgets, Intel Galileo, Raspberry Pi, Gadgeteer, BeagleBone, Cubieboard, Smart Phones
	Software	OS (Contiki, TinyOS, LiteOS, Android) Cloud (Azure, Hadoop, Amazon,..)
Service		Identity-related (shipping), Information Aggregation (smart grid), Collaborative-Aware (smart home), Ubiquitous (smart city)
Semantic		XML language and JSON

Table 1.3: IoT elements and samples.

According to Gubbi et al. [2], the above classification is useful to provide a global definition of Internet of Things for smart environment:

> "Interconnection of sensing and actuating devices providing the ability to share information across platforms through a unified framework, developing a common operating picture for enabling innovative applications. This is achieved by seamless large scale sensing, data analytics and information representation using cutting edge ubiquitous sensing and cloud computing."

Internet of thing connectivity is composed by two families: **critical IoT** and **massive IoT**. The differences are that the first one provides low latency, ultra reliability, very high number of devices and relative low cost, instead of the second one, which does not guarantee low latency and high reliability, but it supports massive number of devices and it provides solutions with very low cost. The NB-IoT, which is the focus of this thesis, is a part of the critical IoT family.

1.3 Wireless Sensor Networks

A set of sensor devices and actuators communicating, deployed in a given area, constitutes a network with no pre-established architecture, the Wireless Sensor Network

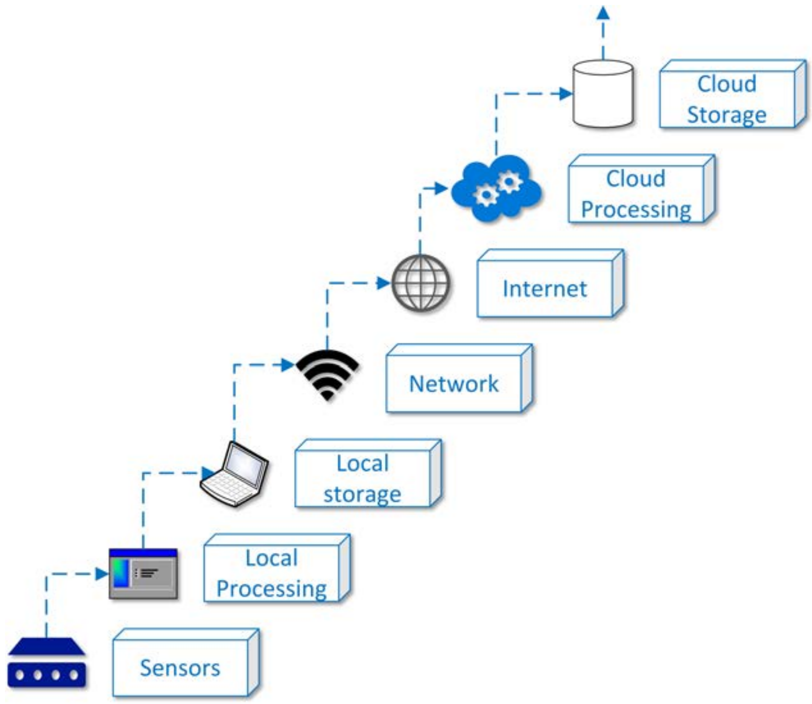


Figure 1.5: The composition of the IoT.

(a representation is provided in figure 1.7). The most important aspect relies in the collaboration of a large number of nodes, because, if we focus on the single node, it is very limited in terms capabilities.

WSNs can be applied in the studies related to environment, like tracking the movements of small animals, monitoring environmental conditions that affect crops and livestock and also Earth monitoring, which also includes bio-complexity mapping of the environments and the study of pollution.

Referring to [9], the technology for the Wireless Sensor Networks is composed by sensors, equipment for the communication and the operating system. Sensors include four basic components, which are represented in figure 1.6: a sensing unit, a processing unit, a transceiver unit and a power unit. The sensing unit is composed by: sensors and Analog to Digital Converters (ADCs); the analog signals produced by the sensors are converted into digital signals by the ADC and then fed into the processing unit. This last unit, which is associated with a small storage unit, manages the collaboration between other nodes. The transceiver unit connects the node to the network. A lot of sensor nodes require their geographical positions be available to them to work with high accuracy; to this end, some sensors can have a location finding system. Due to the limitations imposed by the size of these sensors, power is a very scarce resource; for this reason, power units may be supported also by solar cells or other systems.

The communication technologies include the Bluetooth, that allows data transfers between units over distances up to 10m, and ZigBee, for which the communication range is 75m.

As regarding the operating systems, the embedded wireless communication must satisfy power and real time performance, on heterogeneous software and hardware architectures. A solution is proposed by TinyOS, that is well suited for sensor nodes equipped with a minimum level of hardware. Other alternatives are provided by Microsoft Windows CE, Palm OS and Contiki, which differ in specifications.

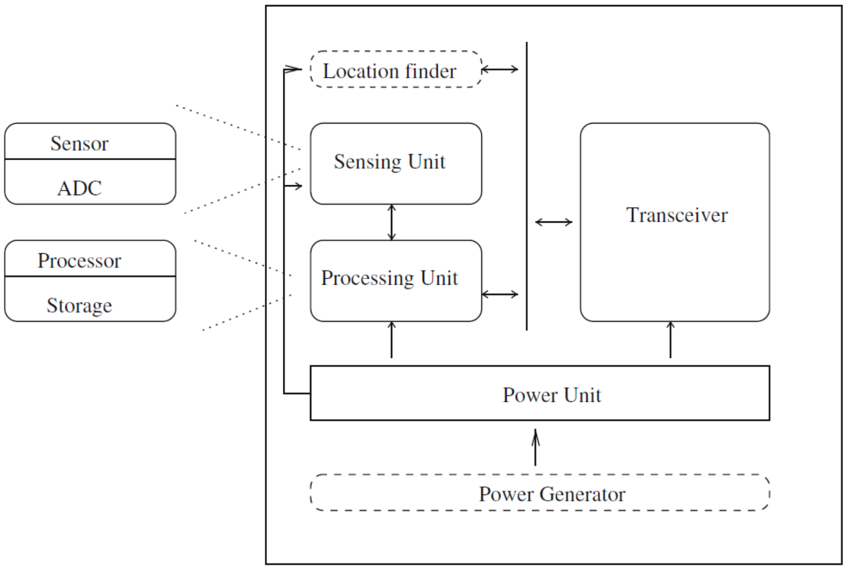


Figure 1.6: Logical component of a sensor.

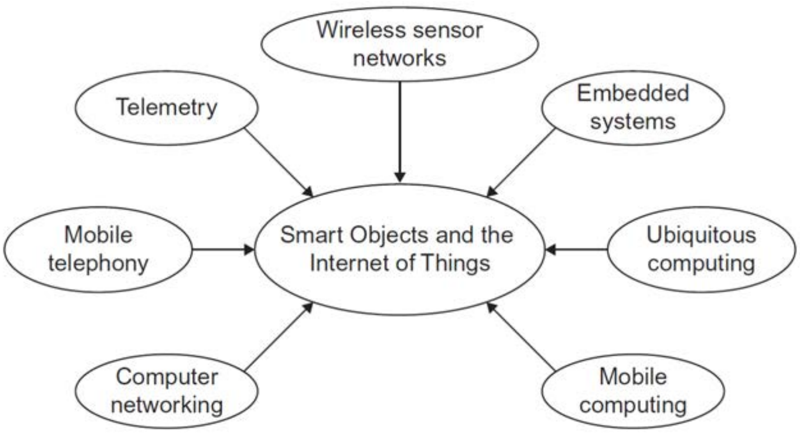


Figure 1.7: Smart object are an intersection between WSN, computer networking, and other things.

1.4 Radio access technologies

To develop the IoT, like authors in [17] have described, new dedicated radio access technologies have been developed for the licensed spectrum, and this make possible to have three new technologies:

- **EC-GSM-IoT**, that is the fully backward compatible solution that can exist over GSM networks, and it has been designed to provide connectivity to IoT devices in the frequency band below 600 MHz;
- **NB-IoT**, which re-uses technical components from LTE to facilitate operation within an LTE carrier. The entire system operates in a narrow-spectrum, starting from only 200 kHz bandwidth and supports extreme coverage and high uplink capacity;
- **LTE-M**, based on LTE, it provides ubiquitous coverage and highly power efficient operation, with a flexible bandwidth of 1.4 MHz, to serve higher-end applications with stringent requirements on throughput and latency.

Focusing on the requirements, the IoT market is divided in at least two categories: massive Machine Type Communication, mMTC and Ultra Reliable and Low Latency Communication, URLLC. Smart Wearables and Sensor Networks are two industrial vertical applications belonging to the mMTC category, while automated driving, industrial automation and eHealth are part of URLLC.

In figure 1.8 there is a graphical representation of the requirements for mMTC and URLLC categories, in terms of coverage, number of supported connections, latency, throughput, mobility, device complexity and device battery life. The center of this chart represents the relaxed requirements, while the edges are the stringent ones.

Parameters	**LTE / E-UTRA 2008**	**LTE Advance 2012**	**5G 2018**
Peak data rate	300 Mbps DL 75 Mbps UL	1 Gbps DL 500 Mbps UL	10 Gbps - 20 Gbps
Channel bandwidth	up to 20 MHz	up to 20 MHz	up to 500 MHz
Over The Air latency	50 ms	10 ms	1 ms
Carriers	1	5	16
Antennas (MIMO)	4	8	64 - 256

Table 1.4: Characteristics of mobile networks.

In table 1.4 a comparison is made between the existing mobile networks.

The 4G cellular technology, known as LTE or E-UTRA, has been introduced in 3GPP in 2008. In LTE the device categories are referred as User Equipment Categories (Cat), and they indicate the capability and the performance of different types of equipment. Cat-1 is the first UE category and it was originally designed to

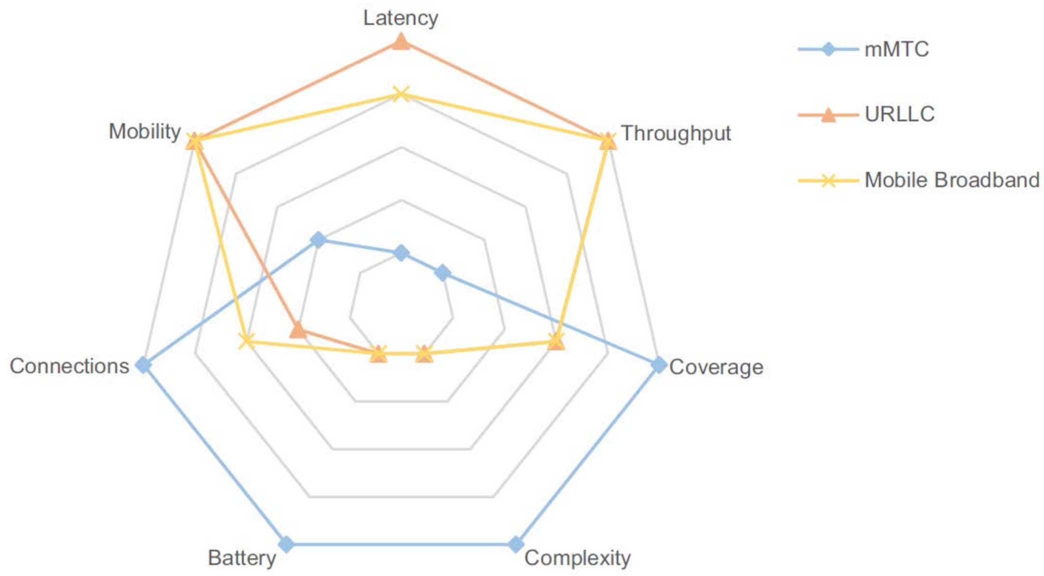


Figure 1.8: eMTC, URLLC and mobile broadband requirements.

support mobile broadband services with data rate of 10 Mbps for the downlink and 5 Mbps for the uplink. In the LTE-Advanced, for what concerns this thesis, there are significantly improvements with respect to the simple 4G. The LTE Narrowband IoT is one category of the MTC introduced in LTE from Release 13. It offers low power consumption, low data rate, no mobility support, limited bandwidth of 180 kHz, extended coverage and low hardware cost. The newest 3GPP LTE NB-IoT standard introduces also the minimal power consumption to extend battery lifetime, [18]. NB-IoT devices are characterized by a non-time critical data transfer and can be used in a range of devices, such as simple wearables, smartwatches or sensors embedded in clothing.

The 5G has been introduced in release 15, which is the starting point for future implementation of 5G activities. The commercial deployment will be in 2020, when the first phase of technical specifications are expected to be ready. The 5G Narrowband IoT is the new radio interface designed to connect, using the cellular infrastructure, a large number of devices in a wide range of application domains. Different data rates have also been introduced ranging from 10s of Kbps in 180 KHz bandwidth (LTE Cat-NB1) up to few hundreds of Kbps (LTE Cat-NB2). In 5G it is also introduced the Resource Spread Multiple Access (RSMA) technology, to have advanced features for massive IoT communication over asynchronous and grant-less access, performing the power saving mode (PSM) and extended discontinuous reception (eDRX) for longer battery lifetime. In figure 1.9 is depicted the future of International Mobile Telecommunication (IMT[1]) from the 2020.

[1]**International Telecommunication Union (ITU)**: it is a specialized agency of the United Nations that is responsible for issues that concern information and communication technologies.

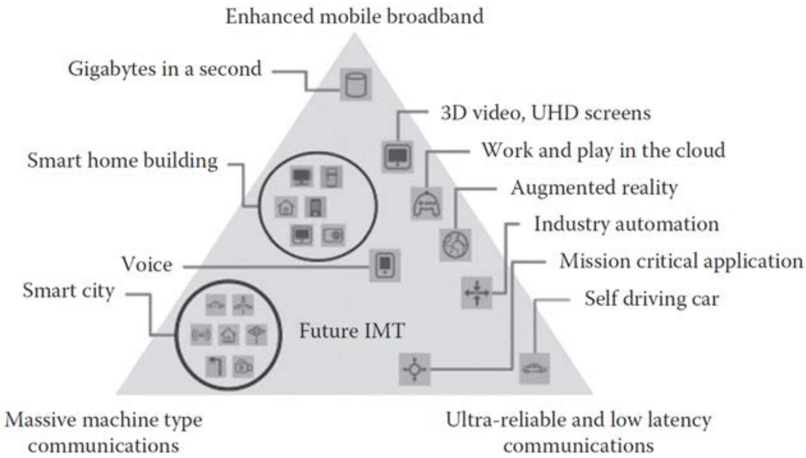


Figure 1.9: Usage scenario for IMT-2020 and beyond.

5G NB-IoT devices are designed to meet the following requirements:

- massive number of low-throughput devices;
- low power consumption;
- longer battery lifetime;
- improved indoor and outdoor coverage;
- low complexity;
- low latency;
- low cost.

On the other hand, cellular network that support NB-IoT are designed in order to have:

- re-use of existing power saving procedures in core network;
- support the sharing of the core network between mobile operators;
- control the UE access for each Public Land Mobile Network (e.g. mobile operators);
- support for SMS;
- support IP header compression for IP-based services;

IMT is a standard and system created by the ITU, for the creation, operation and management of next generation mobile networks and Internet communications.

- support cell selection and re-selection procedures;
- support multicast traffic.

The 3GPP cellular technologies are not the only solutions for IoT traffic. Also Bluetooth and Wi-Fi can serve as bearers for MTC traffic. The cellular technologies are referred as operating in licensed spectrum, the other technologies, instead, operate in unlicensed spectrum. The licensed spectrum is the part of the public frequency space that has been licensed by national or regional authorities to a private company, the Public Land Mobile Network (PLMN) or mobile operator. The Unlicensed spectrum, on the other hand, is the portion of the public frequency space that remain public and, by definition, free of licensing costs. The use of this spectrum is regulated by a set of national/regional technical requirements. An example of this spectrum is the industrial, scientific and medical (ISM) bands. Technologies like Bluetooth, Wi-Fi, Bluetooth Low Energy, ZigBee and Wi-Fi Halow, commonly use such ISM bands to provide short-range communication. There are also systems that can be deployed within both bands and they are regulated by national or regional regulations.

The European Telecommunication Standards Institute (ETSI) is the standardization organization in the European telecommunications industry and provides the Harmonized Standards for the licensed exempt spectrum as the 863-870 MHz and 2400-2483.5 MHz bands. On the other hand, the Federal Communications Commission (FCC), which is the standardization organization for telecommunications in the USA, established as license exempt bands the 902-928 MHz and 2400-2483.5 MHz bands.

Another distinction is related to the area network, and there are: Wireless Personal Area Network (WPAN), Wireless Local Area Network (WLAN), Low-Power WAN (LPWAN). The first comprises the Bluetooth technology, the second is the Wi-Fi and the last is designed to provide long-range connectivity and to support wireless devices in location where WPAN and WLAN cannot provide efficient coverage. The use of low power allows low device cost, compact device design and support flexible usage of the device.

1.5 Application

An IoT application involves sensors and actuators, which are endpoints for IoT networks that collects data and information, like location, images, weather conditions and they perform some actions. They stores such information into the cloud or other storage location, then big data analysis is made to process data, e.g. to obtain prediction of the behaviour of systems, environment, etc. IoT devices that are used as sensors can be used for applications like measuring consumption of gas, water and electricity or to measure weather conditions, pollution levels and environmental activities (e.g. solar activity, noise, and dust). The figure 1.10 describes all the sectors that an IoT application requires to be developed [4].

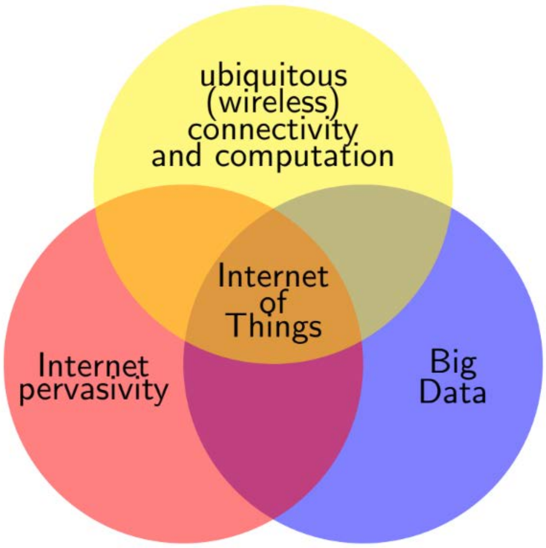


Figure 1.10: How IoT is composed.

Also IoT devices can be used as actuators to control devices (e.g. traffic lights, traffic lanes or home appliance). According to [18], NB-IoT devices are more used as sensors than acting as actuators. This because sensors collect information and route them to a control center, where decision are made and then a corresponding command is sent back to an actuator; to do this there are at least one sensor and at most one actuator (see figure 1.11 for an example).

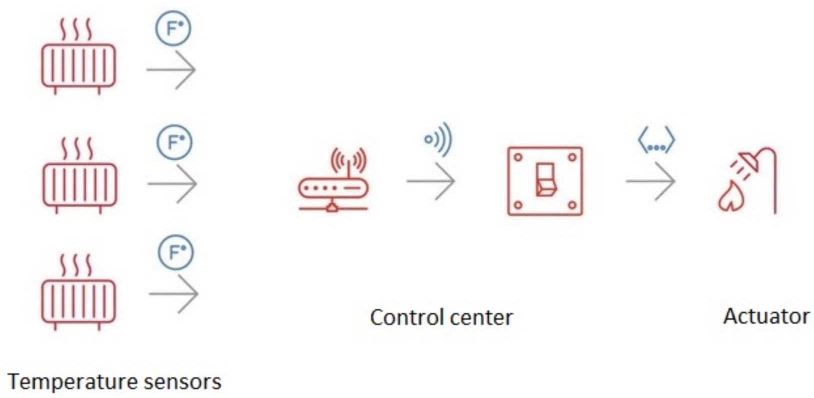


Figure 1.11: Network of temperature sensors, in a house, which detects heat. The control center processes the signal and sends the command of turn on to a sprinkler (the actuator).

The goal of the NB-IoT is to create an highly connected world using sensors and data analytics for sensing, monitoring and controlling all events in homes, cars, agricultural, industrial and environmental fields. It is possible to classify such events in three categories, that are already developed within the 5G NB-IoT specifications:

- Smart City: to monitor traffic, to control infrastructure grids (like electricity, gas) and to manage public safety;
- Smart Home: from home sensors to wearables;
- Smart Transportation: it involves communication between vehicles and pedestrians for traffic warning and traffic safety, for transportation information, and for parking meter communication.

Smart objects used for specific applications, constituting a so called "vertical market", while their computing and analytical services form an "horizontal market". In figure 1.12 the IoT application are shown, in such a way to emphasize the vertical and horizontal markets: here, every domain specific application interacts with domain independent services, whereas each domain sensors and actuators communicate directly with each other [3].

Figure 1.12: This picture is emphasizing the vertical markets and the horizontal integration between IoT application.

1.6 An End-to-End Study

The originality and novelty of this thesis concerns on the fact that here it is developed an experimental end-to-end study by using the Narrowband IoT connections in Italy. When the work for this thesis was done, the NB-IoT implementation in Italy was still in their early days. Also, commercial devices were basically, not simply available and we had to fall back to Development kits. Being the NB-IoT in the early phase of implementation, the interoperability between the infrastructure of different operators (or the same operator but in different geographical regions) and the NB-IoT devices was not yet established. So one specific contribution of this thesis is the selection of the correct combination of the end-device and the network as well as its correct setup.

Another contribution of this thesis is to integrate the Microsoft cloud and the Power BI application with a network of temperature sensors, to view all the interaction of devices, given the NB-IoT connection. Moreover, it is used Microsoft Cognitive Services to interact with the application through the human voice.

Such study help us to view a limitation in the first implementation of NB-IoT antennas: not all antennas have NB-IoT deployments. This incurs in the fact that you cannot develop a NB-IoT solution everywhere.

Also, another important aspect is that in NB-IoT there isn't the possibility of doing a VoLTE call. To overcome such limitation, Microsoft Azure Cognitive Services help us: a string is send to the cloud and the cloud returns a human-like voice message. A little message with a string is, of course, less heavy than a voice call, because the band is not saturated at all.

Since in Microsoft Azure there isn't a way to view a graph of temperature send by sensors, Power BI, a powerful tool of Microsoft, help us to produce and analyze graphs, by start Azure Stream Analytics tool, which provide a query from cloud data (taken from temperature sensors) to Power BI.

Chapter 2

Cellular Internet of Things

2.1 Transmission characteristics

Wireless channels are generally dynamic and unpredictable with respect to wired ones, and this makes difficult an exact analysis of the wireless communication system; for this reason and due to the rapid growth of mobile communication services, the optimization of such communication has become a critical point. In fact, the variety of recent studies highlights the need to develop high performance and bandwidth-efficient wireless transmission technology.

According to [6], radio propagation, in wireless communication, refers to the behavior of radio waves in the propagation from transmitter to receiver, as shown in figure 2.1.

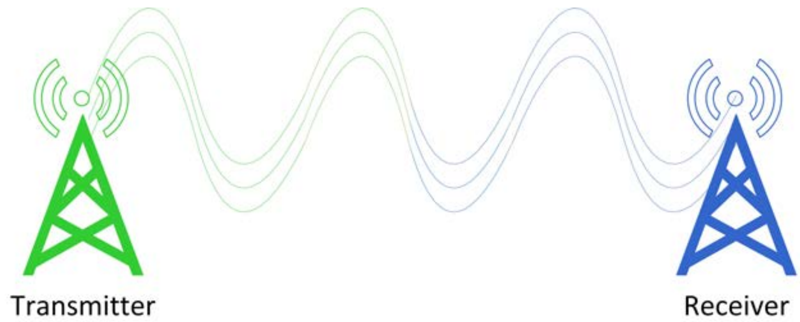


Figure 2.1: Transmitter and Receiver antennas.

Radio waves, when propagated, may be affected by three different types of physical phenomena, whose intensity varies with different conditions and environments:

- *reflection*: occurs when a propagating electromagnetic wave bump into a very big object, with respect to the wavelength, and this forces the signal power to be reflected back to its origin;

- *diffraction*: occurs when the radio path covered from transmitter to receiver is obstructed by a surface with sharp irregularities or small openings;

- *scattering*: it forces the radiation to deviate from a straight path, due to some obstacles with small dimension with respect to the wavelength.

An important characteristic in a wireless channel is the phenomenon called *fading*: it is a variation (degradation) of the signal amplitude over frequency and time; in summary, it is a non-additive signal disturbance in that wireless channel. A channel affected by fading is called "fading channel". Fading phenomenon can be classified into two categories:

- large-scale fading: which occurs when the mobile antenna moves through a large distance, and the signal can be affected by path loss and shadowing.

  - *Path loss* means the reduction in power density of an electromagnetic signal as it propagates through the environment in which it is travelling. It is a deterministic factor and this implies that it can be predicted by considering the distance between the two involved antennas.
  - *Shadowing* is a slow process characterized by variations of median path loss from transmitter to receiver in fixed location. It is a random phenomena, in the sense that its effect can only be predicted by its probabilistic distribution.

- small-scale fading: due by rapid variation of signal levels when the mobile antenna covers short distance, because constructive and destructive interference of multiple signal paths occur.

In figure 2.2 there is a scheme that summarize all phenomenons that can affect the signal in large and small scale fading.

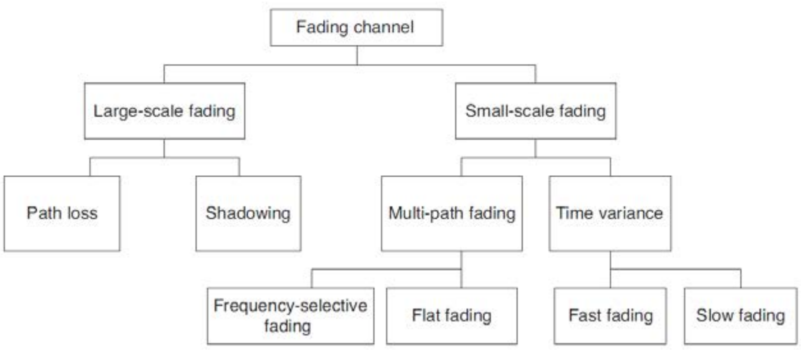
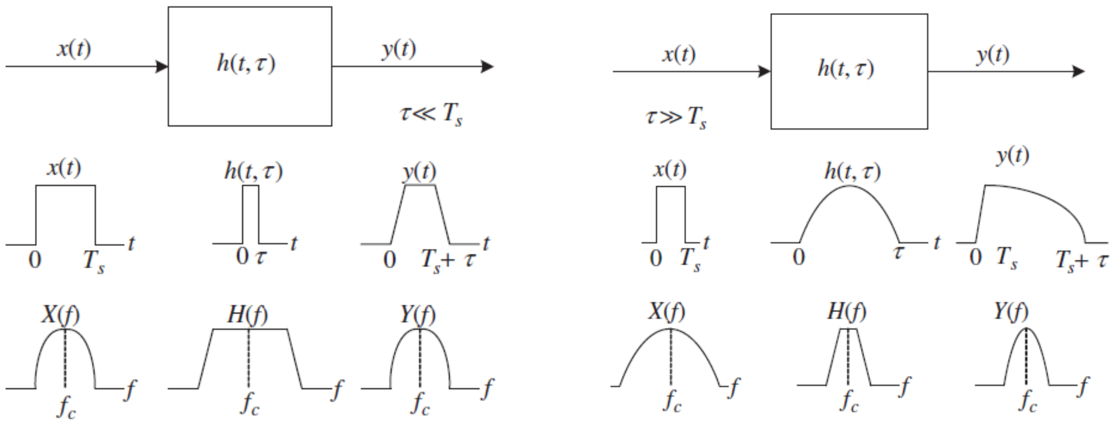


Figure 2.2: Classification of fading channels.

With the movement of the mobile antenna, the type of fading for the corresponding receiver depends on both the transmission scheme, that is specified by signal parameters like bandwidth and symbol period, and the channel characteristics, as multipath delay spread and Doppler spread, which causes, respectively, time dispersion and frequency dispersion; such dispersion induce time-selective fading or frequency-selective fading. a transmit signal can be fading in the frequency domain either in a *selective* or *non-selective* manner, due to the time dispersion phenomenon. The frequency selectivity is governed by signal bandwidth and by how the channel response varies with the frequency.



(a) Frequency non-selective fading channel (b) Frequency selective fading channel

Figure 2.3: Fading characteristics due to time dispersion over multipath channel $(h(t, \tau))$.

In the figure 2.3 there is given an intuitive representation of fading characteristics due to time dispersion. As illustrated there, the transmitted signal is affected by frequency non-selective fading when the signal bandwidth is narrow enough and, in such way, it can be transmitted over the flat response; this is the contrary to what happens in frequency selective fading, where the signal, with wider bandwidth, is filtered out by the finite channel bandwidth.

2.2 Transmission types

2.2.1 Single-Carrier transmission

In figure 2.4 there is a typical end-to-end configuration for a single-carrier baseband communication system: here, there is $g_T(t)$ that is the transmit filter with transmit symbols $\{a_n\}$ in input (with a symbol period of T seconds, and data rate of $R = 1/T$), a band-limited channel $h(t)$ with a bandwidth W, the receive filter $g_R(t)$ that, with the equalizer and the detector, processes the received symbols. The following equation

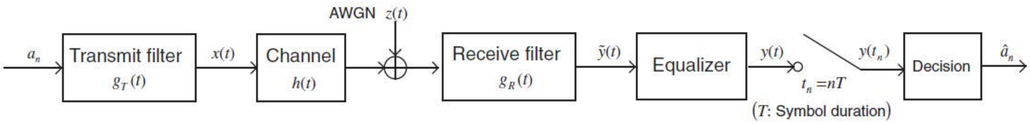


Figure 2.4: Single-carrier baseband communication system model.

expresses the output of the equalizer:

$$y(t) = \sum_{m=-\infty}^{\infty} a_m g(t - mT) + z(t) \tag{2.2.1}$$

where $z(t)$ is the additive Gaussian noise and $g(t)$ is the impulse response of the overall end-to-end system:

$$g(t) = g_T(t) * h(t) * g_R(t) * h^{-1}(t) \tag{2.2.2}$$

Note that $g(t)$ isn't time limited due to the fact that it refers to a finite channel bandwidth. In case the noise term is ignored:

$$y(t_n) = a_n g(0) + \sum_{m=-\infty, m\neq n}^{\infty} a_m g((n-m)T) \tag{2.2.3}$$

Note that the term a_n contains an Inter-Symbol Interference (ISI) component, that is caused by a trail of the impulse response, that could degrade the performance of the digital communication system; this is a cause related to the overlapping of subsequent symbols.

As the symbol rate R increases, the signal bandwidth becomes larger and this implies that the transmitted symbol is affected by multipath fading, and this, as it is already explained, leads to frequency selectivity.

2.2.2 Multi-Carrier transmission

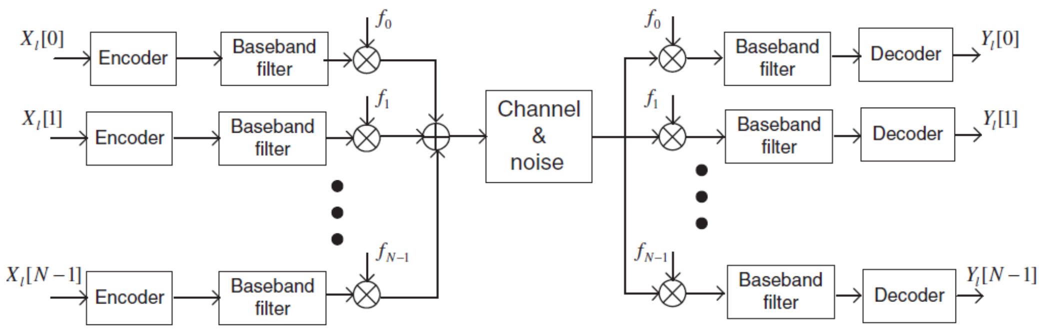


Figure 2.5: Multi-carrier communication system model.

Multi-carrier transmission is the proposed solution to overcome the problem of frequency selectivity of the single-carrier with wideband channel, for high rate data transmission. The figure 2.5 there is a representation of such transmission model, where the wideband channel is dived into N narrowband channels: here the wideband signal is splitted in several narrowband signals which are analyzed through the multiple narrowband filter $H_k(f)$ and then synthesized with the multiple narrowband receiving filter $G_k(f)$. In this way the frequency-selective wideband channel can be approximated by multiple frequency-flat narrowband channels, and this is shown in figure 2.6.

In figure 2.7 there is illustrated a transmitted signal spectrum in the multicarrier transmission system, which occupies multiple sub-bands of equal bandwidth, each centered at different carrier frequency.

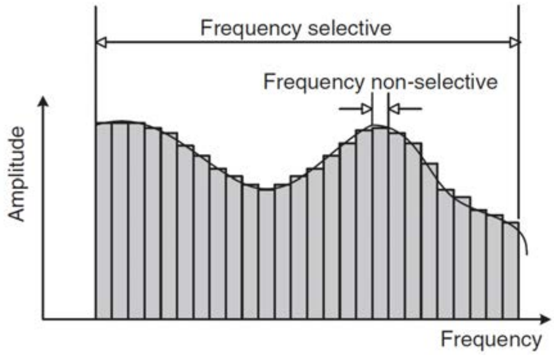


Figure 2.6: Frequency Response of a multichannel transmission system.

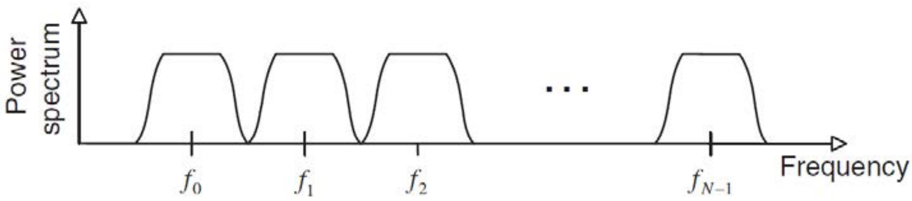


Figure 2.7: Signal spectrum of a multicarrier transmission signal.

2.3 Transmission modulations

2.3.1 Orthogonal Frequency Division Multiplexing

According with authors of [6], Orthogonal Frequency Division Multiplexing (OFDM) transmission scheme is a type of multichannel system which employs multiple subcarriers. In figure 2.8 there is a graphical representation of such communication scheme.

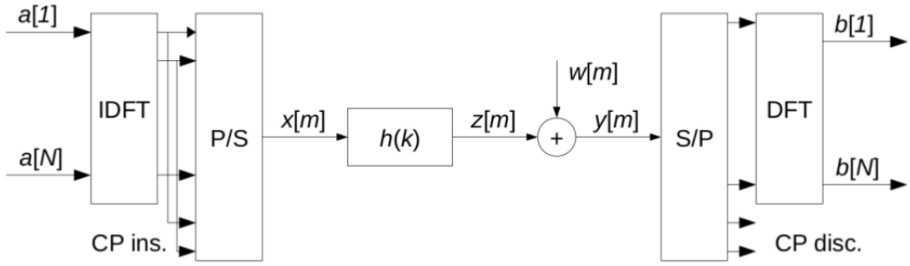


Figure 2.8: OFDM transmission scheme implemented using IDFT/DFT.

The use of multiple orthogonal subcarriers, with finite duration T, implies signal overlapping in the spectrum, as shown on figure 2.9; to overcome such problem, discrete Fourier Transform (DFT) and inverse DFT (IDFT) are used to implement such orthogonal signals. The fast Fourier transform (FFT) and the inverse fast Fourier transform (IFFT) are the implementation of DFT and IDFT, respectively. In the OFDM transmission scheme, $\{X_l[k]\}_{k=0}^{N-1}$ are the transmitted symbols which generate $\{x[n]\}_{n=0}^{N-1}$, that are the samples for the sum of N orthogonal subcarrier signals, after the IDFT calculation. By adding the additive noise $w[n]$, the received sample is calculated as: $y[n] = x[n] + w[n]$, and this forms the chain of received samples $\{y[n]\}_{n=0}^{N-1}$ (and its noisy version is $\{Y_l[k]\}_{k=0}^{N-1}$).

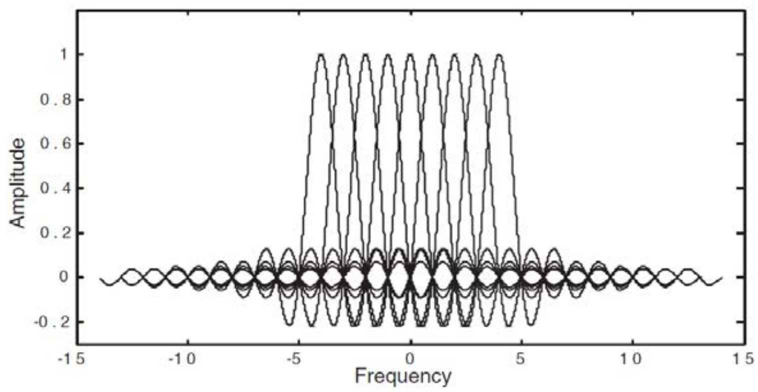
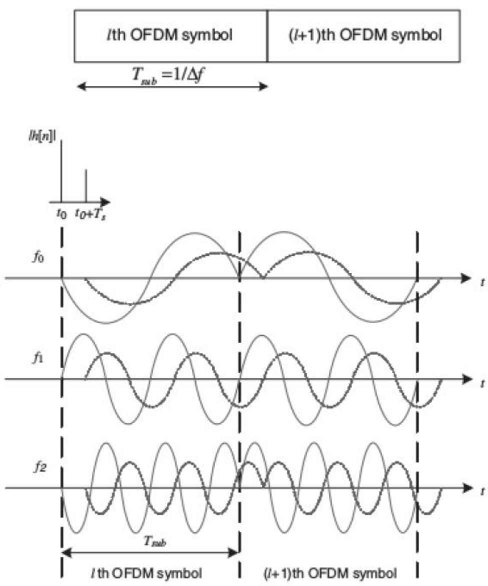
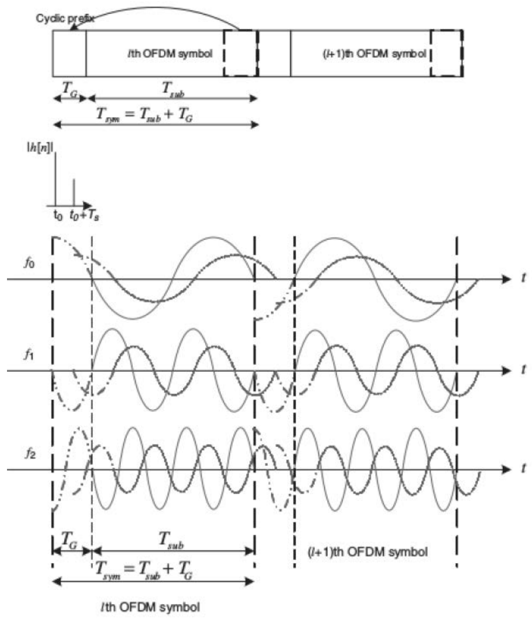


Figure 2.9: Spectrum of OFDM signal in linear scale. It is a sum of shifted sinc functions, where the overlapped neighbors are spaced by 1/T.

The OFDM scheme places a virtual carriers (VCs) around the frequency band in order to reduce the out-of-band radiation, caused by the fact that each subcarrier signal is time-limited for each symbol, which produces non-negligible adjacent channel interference (ACI). Another guard interval, called cyclic prefix (CP), is inserted in the time domain, in order to mitigate the inter-symbol interference (ISI) that can be obtained between OFDM symbols. In figures 2.10 there is depicted the ISI effect without and with the guard interval.



(a) OFDM symbols without guard interval and the ISI effect on the received signal



(b) OFDM symbols with guard interval and the ISI effect on the received signal, that is mitigated by the CP

Figure 2.10: ISI effect with and without the cyclic prefix.

2.3.2 Orthogonal Frequency Division Multiple Access

OFDM is the transmission technique in which all the subcarriers are used to transmit symbols of a single user; it is not a multiple access technique, but it can be combined with existing multiple access techniques, such as Time Division Multiple Access (TDMA), Frequency Division Multiple Access (FDMA) and Code Division Multiple Access (CDMA). The most useful approach in mobile cellular system is to combine OFDM with FDMA, and obtained the so called OFDMA. According with authors in [6], users that make use of the same cell, may have different interference: an efficient way to overcome this problem is to allow that users can select their own subset of subcarriers with better channel conditions, instead of having a single user that uses all the subcarriers at the same time (as what happens in OFDM-CDMA). In figure 2.11 there is a comparison between OFDMA and OFDM-CDMA access techniques.

The amount of physical resources used to develop OFDMA depends on the required data rate of each user and on the multi-user diversity gain among users, that relies on the nature of the channel usage among users in FDMA.

2.3.3 Single Carrier FDMA

As cited in [10], Single Carrier Frequency Division Multiple Access is a modified form of OFDMA. It is a technique for high data rate uplink communications used in cellular

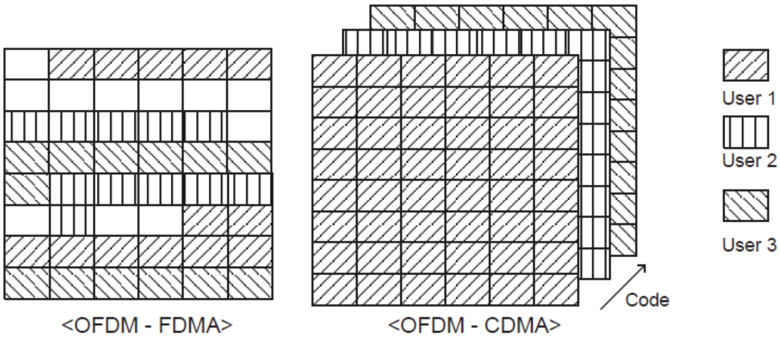


Figure 2.11: Multiple access techniques used in OFDM system.

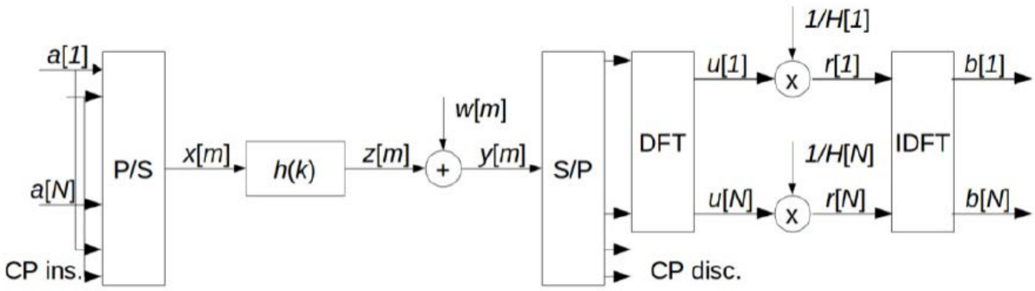


Figure 2.12: Single-Carrier FDMA scheme.

systems.

An advantage of using SC-FDMA is the peak-to-average power ratio (PAPR), that is lower than that is achieved in OFDMA. PAPR is the ratio between the maximum power and the average power of the complex signal $s(t)$; it is due by the fact that linear amplifiers impose a non-linear distortion on their outputs caused by their saturation characteristics. Such non-linear characteristic of High Power Amplifier, when excited by a large input, causes the out-of-band radiation that affects signals in adjacent bands and in-band distortions that result in received signal rotation, attenuation and offset. To reduce the PAPR, SC-FDMA is adopted for uplink transmission in 3GPP LTE standards, instead of OFDMA, because it spreads the input signal with DFT, subsequently taken this into IFFT, and then reduce the PAPR to the level of single-carrier transmission, as shown in figure 2.12.

2.4 Telecommunication Standards

2.4.1 Licensed and License exempt band regulations

Cellular technologies can operate in two types of spectrum, [17]:

- *licensed spectrum*: it corresponds to a part of the public frequency space which is licensed by national or regional authorities to a private company (e.g. a mobile network operator), with the condition of providing a certain service to the public (e.g. cellular connectivity);

- *unlicensed spectrum*: also called license exempt frequency bands, it corresponds to the portions of the public frequency space which are free of licensing costs; each equipment manufacturers that use such spectrum must meet a set of national or regional technical regulations. Industrial, Scientific and Medical (ISM) bands are the most popular license exempt frequency bands, and they are specified in the International Telecommunication Union (ITU) Radio Regulations.

In the United States, License exempt bands are defined by the Federal Communications Commission (FCC) in the Electronic Code of Federal Regulations; it states that unlicensed spectrum are 902-928 MHz and 2400-2483.5 MHz. In the ancient continent, the Europe, is the European Telecommunication Standards Institute (ETSI) that defines the unlicensed spectrum in the Harmonized standards, and states such spectrum in 863-870 MHz and 2400-2483.5 MHz.

The 3GPP requirements, mentioned in section 1.4, are defined to regulate the secure coexistence between 3GPP systems operating in adjacently within the same band or in different bands; also they define in-band requirements to guarantee systems' minimum level of performance on the whole and on a per link basis. Requirements for license exempt bands are defined to support the coexistence between systems outside and inside the band. Although inside the band systems in the frequency domain are overlapping, because they may use the same frequency resource at any point in time, then to limit such interference, regulations are defined to limit the used of output power. In this way the *duty cycle* and the *dwell time* on a specific frequency resource may be defined. Duty cycle defines the ratio by which a transmitting device uses a radio resource; dwell time defines the maximum contiguous time by which a transmitting device uses a radio resource. To provide high and robust coverage for a lot of devices in unlicensed bands, such requirements establish a strict design boundaries to follow, by meeting service requirements.

2.4.2 Narrowband Cellular IoT: NB-CIoT

According with 3GPP technical specification [13], Narrowband Cellular IoT is an optimized standard for the IoT communications, which takes account of the following considerations:

- compatibility with stand-alone deployments in a low minimum system bandwidth;
- compatibility with the battery technologies that are used in low cost IoT products;
- constraints in the mobile station maximum transmitting power.

In uplink and downlink, NB-CIoT solution uses a minimum system bandwidth of 200 kHz and, in each directions, channels are divided into narrow-bandwidth subcarriers: 48 on downlink and 36 on uplink. This is done to enable the communication with mobile stations that have lower peak transmitting power, but with worst coverage.

On downlink, OFDMA is adopted, to provide:

- high bandwidth efficiency and flexible time/frequency resource provisioning;
- straightforward support for frequency hopping to easy manage interference;
- multi-subcarrier broadcast channel that reduces power consumption on reception;
- reduction of interference to co-located LTE in adjacent channels.

Also, on downlink, PSK and 16-QAM modulations are used.

On the other hand, on uplink, FDMA is used, because each subcarrier is modulated and pulse-shaped (to ensure that users which operates on adjacent subcarriers are separated in frequency) as a single carrier. Also, on uplink, GMSK and PSK modulations are used.

Classical modulation techniques are more in detail in the appendix *Digital Modulation*.

2.5 Communication Technologies: Antennas

It is important to find techniques to achieve high-speed data transmission or high reliability. To do this, it is possible to classify multiple antenna techniques into two main categories: *diversity* and *spatial-multiplexing* techniques. Diversity techniques refer to the fact that the multiple antennas receive the same information-bearing signals by improving the transmission reliability; this is done when a channel is converted in a more stable AWGN-like channel without any catastrophic signal fading [6]. On the other hand, in spatial-multiplexing techniques the independent data streams are simultaneously transmitted by multiple antennas, and this achieves an higher transmission speed. The choice of the use of one category with respect to the other is due by how much speed in transmission do you want: with spatial-multiplexing it is possible to achieve the maximum speed (that is the same as the capacity of the MIMO channel), otherwise with diversity techniques the achievable transmission speed is lower with respect to the capacity of the MIMO channel.

Antennas can achieve diversity gain by following some ways, like:

- **Space diversity**: in the sense that antennas must be sufficiently separated, because they are used to implement independent wireless channels;

- **Polarization diversity**: because independent wireless channels are implemented by taking care to the fact that vertically and horizontally polarized paths are independent;
- **Time diversity**: it can be achieved by considering the fact that the same information is transmitted repeatedly at sufficient separated time instances;
- **Frequency diversity**: the same as before, but considering the separation of frequency bands;
- **Angle diversity**: to receive the same information-bearing signal at different angles, it is possible to use multiple receive antennas with different directions.

Antennas are the way for a network operator to spread cellular signals to users. In Italy there are five main network operators: TIM s.p.a., Vodafone Italia, Wind Tre, Iliad and Rete Ferroviaria Italiana (this last is only for private use of the railway company RFI). Also there are a lot of virtual operators, like Fastweb Mobile, PosteMobile and so on, which use the existing network infrastructure of the main operators. To do this thesis' work, Vodafone was our choice. In figures 2.13 and 2.14 there are shown where are placed antennas near Mida Solutions s.r.l., where my internship was done and the most of the work was done, and near DEI, where I have done my studies in telecommunication engineer and the antenna-test part of the thesis was done. Every cell has a Physical Cell Identifier (PCI) which allows to the device, that is connecting to the network, to identify the specific cell in an unambiguous manner. In figures aforementioned, it is written PCIs for the two antennas used in my thesis.

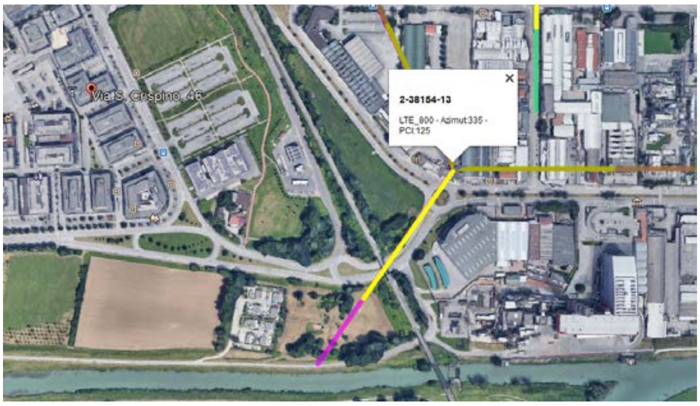


Figure 2.13: Antenna near Mida Solutions' office

LTE has five frequencies, and they are shown in the table 2.1.

The choice of the MIMO is because, at least in theory, a 4x4 MIMO can send twice the data of a 2x2 MIMO. In practice this isn't really what happens and the difference is negligible. Carriers supported in LTE is divided in three categories:

- 4G has only one carrier;

Figure 2.14: Antenna near DEI building

LTE frequencies	***Bandwidth***	***MIMO***
800 MHz	10 MHz	2x2
1400 MHz	20 MHz (only downlink)	2x0
1800 MHz	20 MHz	2x2 4x4
2100 MHz	10 MHz(if sharing with UMTS) 15 MHz	2x2
2600 MHz	15 MHz	4x4

Table 2.1: LTE frequencies

- 4G+ has from 2 up to 3 carriers;
- 4.5G supports from 4 carriers.

LTE 800 is chosen for the NB-IoT because the signal can go very far and can pass through "signal-resistant" materials, like reinforced concrete walls, with the help of the additional 6 dB of power boost. Moreover LTE 800 is installed via software in physical cells, it has little bandwidth of 180 kHz and it is added two guard intervals, one to the left and one to the right, of 45 kHz (in total there are 270 kHz of bandwidth) and it is placed in the in-band. The Voice over LTE (VoLTE) technology is avoided in LTE 800 because if one device does a VoLTE call, it saturates the bandwidth [14].

By following the angle diversity implementation, in LTE every antenna must have a certain degree of regulation, of more or equals to 60 degrees, to avoid interference with other cells; this concept is referred as "lobe opening", [15]. The signal lobe is shown in figure 2.15. The lobe is opening of more or less 5 degrees above and below the theoretical line that is perpendicular to the ground, and then it propagates, making, of course, the entire lobe. However, in figure 2.16 there is a representation of polar pattern of a lobe, represented as an array of lobes (credits: [12]). Following the z axis, in the right, at 0 degree, there is the *main lobe*, that corresponds to the signal propagation, but in the opposite side (at 180 degrees) there is the so called *back lobe*; all the other lobes are called *side lobes* and they represent all the undesired radiation in all the directions. The lobes radiation pattern grows with the radio wavelength. In the transmitting antenna, if side lobe has excessive radiation, this may cause interference to other cells; on the other hand, in receiving antenna, side lobes can capture interfering signals and then increase the level of the noise.

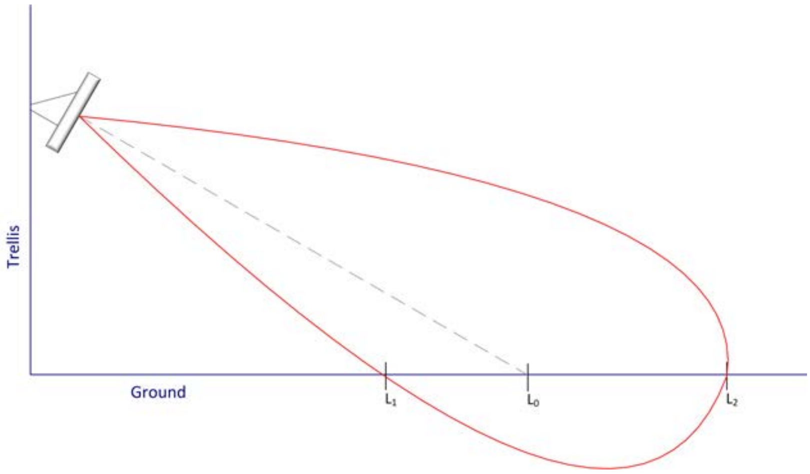


Figure 2.15: Practical antenna lobe.

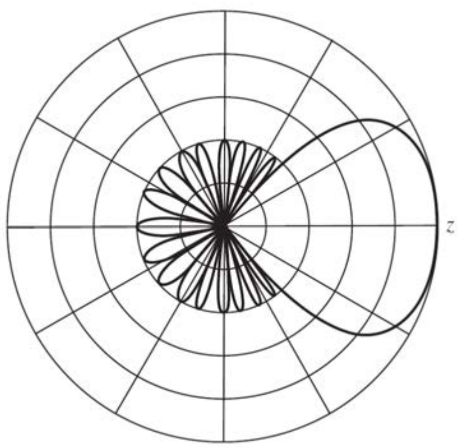


Figure 2.16: Theoretical antenna lobe.

In figure 2.17 there is shown the orientation of LTE 800 antennas. LTE 800 antennas are usually tilted lower than the others, because for narrowband modulation it is important to not have losses in the quality of the signal also if there are "signal-resistant" materials, as cited before. In figures 2.13 and 2.14 there are antennas that are represented with a beam of three axes: this is also explained in figure 2.17. Here "physical" means that they are physical cells, instead of "software" which means that for NB-IoT there are not built dedicated physical antennas, but NB-IoT itself is configured, via software, to operate "in-band", as it will be explained in section 3.3.1.

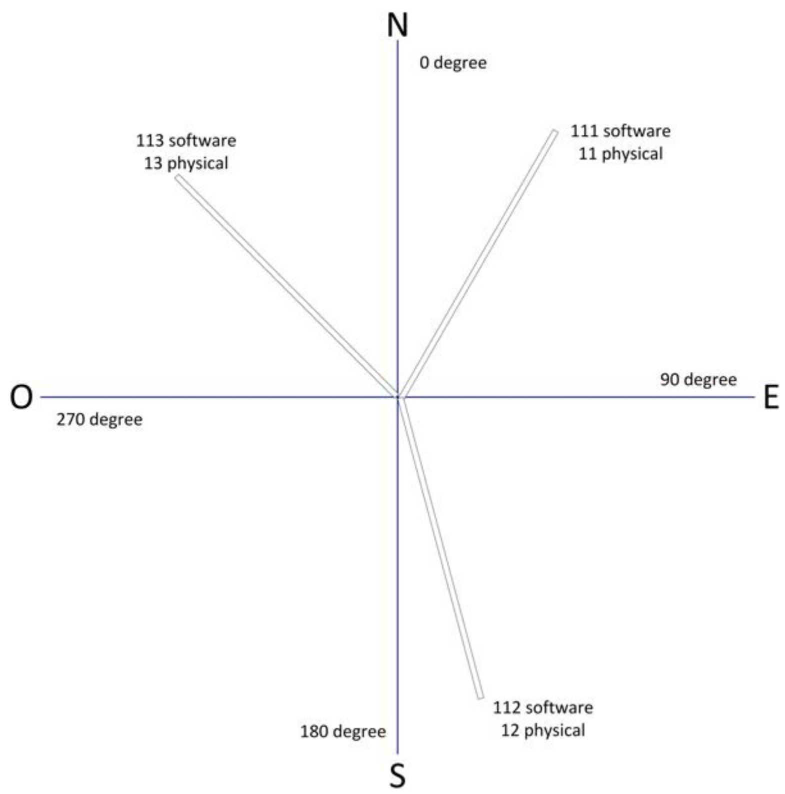


Figure 2.17: Orientation of LTE 800 antennas.

Chapter 3

NarrowBand IoT

In Release 13 of the 3GPP 5G specifications a new radio interface is introduced [16]: the Narrowband IoT (NB-IoT). According with [18], NB-IoT is designed to connect, through cellular infrastructure, a large number of devices in a wide range of application domains, forming the so-called Internet of Things. The 5G NB-IoT is also planned to introduce advanced features for massive IoT, like Resource Spread Multiple Access (RSMA), which requires asynchronous and grant-less access, power saving mode schemes and extended discontinuous reception for longer battery lifetime, [24].

NB-IoT is a part of the Low Power Wide Are Network (LPWAN), and operates in licensed spectrum bands. 3GPP includes such technology as a part of the LTE standards, because it can be based on the big ecosystem offered by the LTE technology and mobile operators. 5G NB-IoT promises to enhance existing cellular-base use cases, like: smart homes, smart cities, smart transportation, smart grids, smart industry, mobile virtual reality, object tracking and mission critical control. In coming few years, it has been calculated that there will be an explosion in the number of IoT connected devices; for example, in 2025 to be precise, 5 billion devices will be connected through 5G NB-IoT network. To make such expansion possible, User Equipment and networks have the following requirements and goals:

- a cell sector can support at least 52,547 connected devices;
- devices have low power consumption and longer battery lifetime, in order to enable a single battery charge for multiple number of years;
- the target for 5G NB-IoT devices is to achieve an extended coverage compared to legacy GPRS devices;
- devices will be ultra-low complexity to achieve very low cost;
- devices will have low latency, in the order of 10 ms or less.
- re-use of existing power saving procedures of the core network level;

- share of the core network by multiple mobile operators;
- control of the UE access for each Public Land Mobile Network (PLMN [1]);
- support of Short Message Service (SMS)
- support IP header compression and multicast traffic for such services that are based on IP;
- support of cell selection and re-selection procedures in both idle and connected modes;

The only thing that is not mentioned is the voice call: VoLTE call through NB-IoT, would saturate the bandwidth and other devices could not use the same NB-IoT cell.

3.1 NB-IoT Protocol Stack

NB-IoT follows a network protocol stack which is designed into a layered architecture for both transmitting and receiving side. Each layer has a protocol by which it can communicate with the peer node at the same level; such protocol enable the exchange of messages, packets or Protocol Data Units (PDUs) to provide services or functions to the upper layer or to use functions and services of the lower layer.

The structure of the layer is based on the model of Open Systems Interconnection (OSI), which is developed as an international standard for computer networks by the International Standards Organization (ISO) [18]. Typically, the OSI architecture is divided vertically into *data plane* and *control plane* and horizontally into *Access Stratum* (AS) and *Non Access Stratum* (NAS). Data plane is such plane in which the user data flows between the two nodes, while control plane is the one where the control information is exchanged. Access Stratum is responsible for the handling and processing the physical transmission or reception on the media, based on what type of network is considered which can be done, for example, via Ethernet cable or in a wireless channel (like the one of the NB-IoT). Non Access Stratum includes the five upper layer (Network, Transport, Presentation, Session and Application) that are almost the same across different type of physical media, because they are independent of the physical media.

In figure 3.1 there is the description of the OSI data plane protocol stack [18]. Every layer exchanges data units as Service Data Unit (SDU) or PDU: the first one refers to the exchange of data unit within the layer (intra-layer), while the second means the exchange of data unit between layers (inter-layer). It follows that each layer has its own SDU wich is appended with a header in the packet unit: on the transmitting path a layer appends its SDU, then encapsulates the packet with an

[1]**PLMN** is the combination of wireless communication services offered by a specific operator in a specific country.

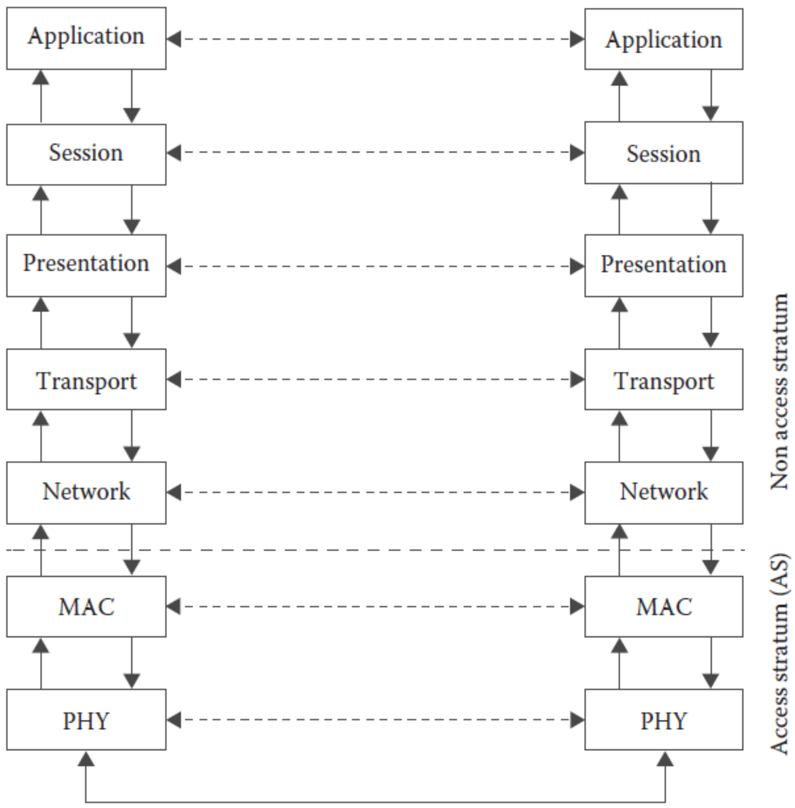


Figure 3.1: OSI data plane protocol stack.

header and sends the packet to the layer below, while on the receiving path a layer sends its SDU to the upper layer. Each layer knows the size of its header and then can strip the PDU received from the lower layer to obtain the SDU. Of course, each layer may add a checksum or an integrity-protection trailer to its SDU.

Considering the correspondence between the NB-IoT and the OSI protocol stack, in NB-IoT the OSI MAC layer is composed by Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC) and Medium Access Control (MAC), which, with Physical Layer, constitute the Data Plane. In Control Plane there are two additional elements: Non-Access Stratum (NAS) and Radio Resource Control (RRC).

3.2 LTE NB-IoT Protocol Architecture

LTE is an evolution of UMTS and it has inherited some terms by its predecessor, and they are used by 3GPP: E-UTRA is the Enhanced UMTS Terrestrial Radio Access and represents the User Equipment, while E-UTRAN is the Enhanced UMTS Terrestrial Radio Access Network and it is the core network. This last one is composed by the eNodeB, which is the base station (the central controller), that is connected to a large

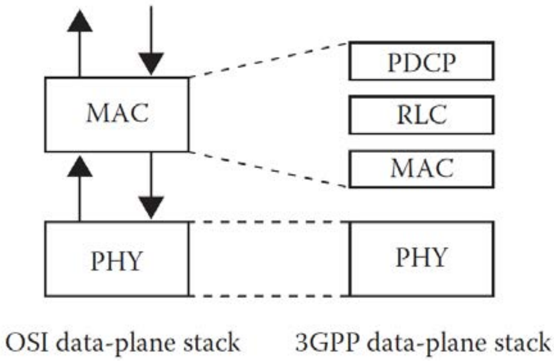


Figure 3.2: NB-IoT Data Plane.

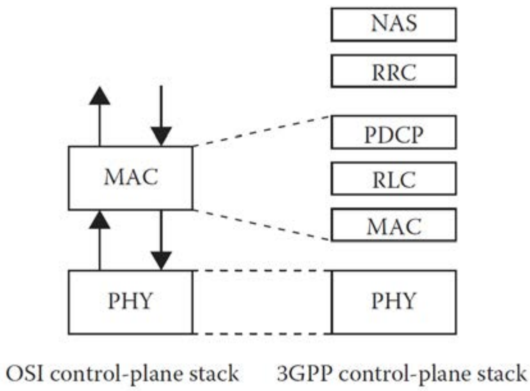


Figure 3.3: NB-IoT Control Plane.

number of NB-IoT devices. Also, the eNodeB can be connected with other eNodeBs and with the core network (EPC, Evolved Packet Core) via of some protocols, like X2 and S1, respectively.

As we can see in figure 3.4, the eNodeB is connected to the Mobility Management Entity (MME), to carry control-plane messages and signalling, but eNodeB is also connected to the Serving Gateway (S-GW) and Packet Gateway (P-GW) in order to carry data-plane messages. Since the network is composed by thousands of devices, there can be multiple MMEs communicating with the same eNodeB which perform load-balancing among themselves, by providing:

- NAS signalling (e.g. attach and tracking area update procedures);
- authorization and authentication;
- selection of S-GW and P-GW;
- interception of piggybacked signalling messages or data-plane messages.

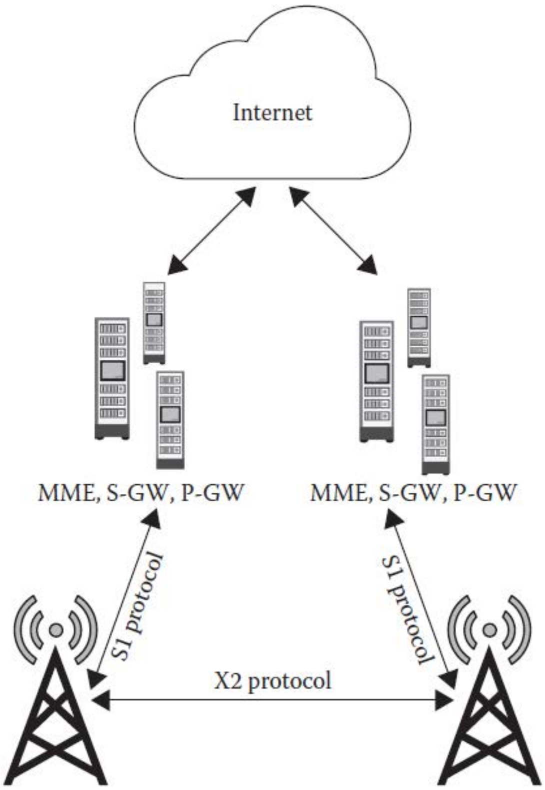


Figure 3.4: LTE NB-IoT network architecture.

Near MME there is S-GW, which provides:

- packet forwarding and re-routing to P-GW;
- accounting for User Equipment traffic;
- local mobility anchor (e.g. if the device moves to different EPC, its traffic is re-routed through its home S-GW (HS-GW));
- interception of data plane packets.

Then there is P-GW, which is the access point that provides connectivity to the User Equipment, and its main functions are:

- support of IP traffic;
- mapping of Evolved Packet System (EPS) bearer parameters, to route traffic to the User Equipment;
- packet filtering and inspection;

- data rate enforcement and accounting for User Equipment traffic volume in uplink and downlink;
- interception of data plane packets.

The other component of the EPC is the Home Subscriber Server (HSS), which is used for storing and updating User Equipment subscription information, and it has the following functions:

- it provides UE identification and addressing: it contains the International Mobile Subscriber Identity (IMSI) of the UE;
- it also provides authentication between the MME and the UE;
- it provides security keys used for ciphering and integrity protection for the signalling and data plane messages exchanged between the UE and the eNodeB.

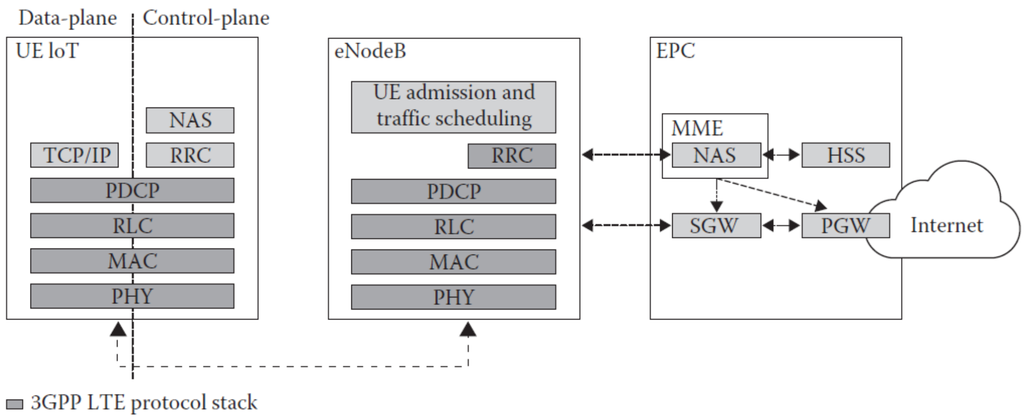


Figure 3.5: LTE NB-IoT protocol stack for both UE and eNodeB.

In figure 3.5 there is another draw of the NB-IoT protocol stack, where there are more emphasis in the UE and the eNodeB communication scheme.

3.2.1 NB-IoT Modes of Operation

The radio interface of the NB-IoT can support three modes of operations, which are described in figure 3.6 and they are listed below [17]:

- **Inband**: that is done by using the LTE carrier bandwidth, reserving resource blocks; each block occupies 180 KHz of bandwidth.
- **Guardband**: by using guard resource blocks within an LTE carrier's guardband;

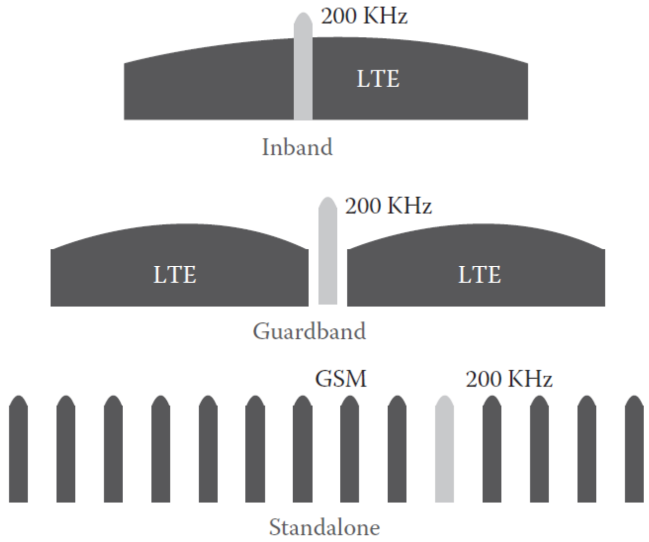


Figure 3.6: LTE NB-IoT modes of operations.

- **Standalone**: by using a dedicated carrier other that LTE (e.g. GSM), and it occupies one GSM channel of 200 KHz.

For NB-IoT standalone operation there are 200 kHz of channel bandwidth and 180 kHz of transmission bandwidth [18], that are shown in figure 3.8(a). Similar mode is the inband, where there is the same transmission bandwidth. As regarding the guardband operation, it uses the same channel bandwidth as the inband mode, except that the 3 MHz one is note used. Inband and guardband mode of operations are represented in figures 3.8(b) and 3.8(c).

In figure 3.7 there is a representation of the frequency domain waveform of the 200 kHz resource block channel bandwidth. This is related to the first block that appears in the inband mode of operation. This figure is very important to view and explain why guard bands are used in this type of operation: they are used because without them there is a overlap of signals and the receiver is not able to recover the original signal. This mode of operation is used to deploy NB-IoT in physical antennas.

3.3 Radio Resource Control Sublayer

Radio Resource Control, RRC, is the control plane sublayer used to exchange data during transmission and reception with the eNodeB, and it involves two actors: Signaling Radio Bearer (SRB) and Data Radio Bearer (DRB); such messages are encapsulated in PDU packets. Before exchanging packets with the eNodeB, a cell selection procedure must be performed: this is done when the UE is powered on or the USIM is inserted and then it selects a cell of the target PLMN for the first time. Here, UE

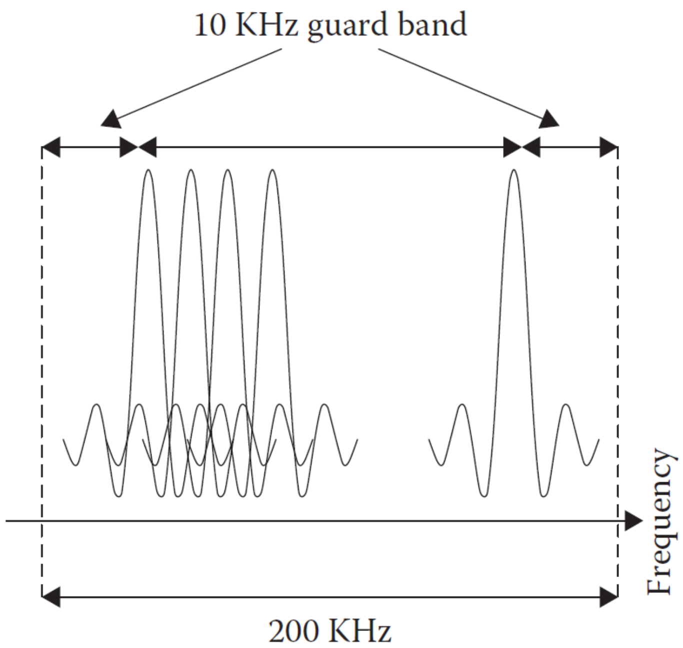
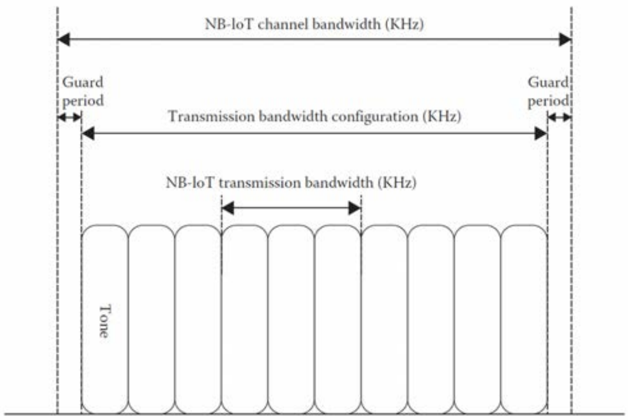


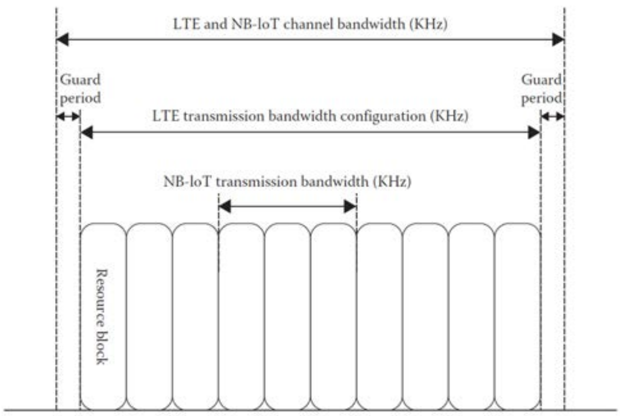
Figure 3.7: Frequency domain of 200 kHz channel bandwidth.

needs to tune to each broadcast channel frequency and acquires Master Information Block (MIB) and System Information Block (SIB) messages. With the help of MIB and SIB, the UE is able to identify that the observed cell belongs to the target PLMN and then it camps on it; if the UE camps on a cell, this last becomes a *serving cell* of the UE. If multiple cells exist, the UE chooses the strongest one. After the camping, the UE periodically does a cell reselection to find if there is any other strongest cell, and this is called, of course, cell re-selection. On the other hand, the eNodeB transmits in a repetitive way the MIB and SIB messages, in order to increase the probability of a UE being able to acquire them reliably [18]. In figure 3.13 it is shown the procedure followed by the UE to select or re-select the eNodeB. In figure 3.9(a) there is a representation of what happens when the UE is connected to the eNodeB: here, Non-Access Stratum messages are sent from eNodeB to the UE in the Downlink way or vice versa in the Uplink procedure (figure 3.9(b)).

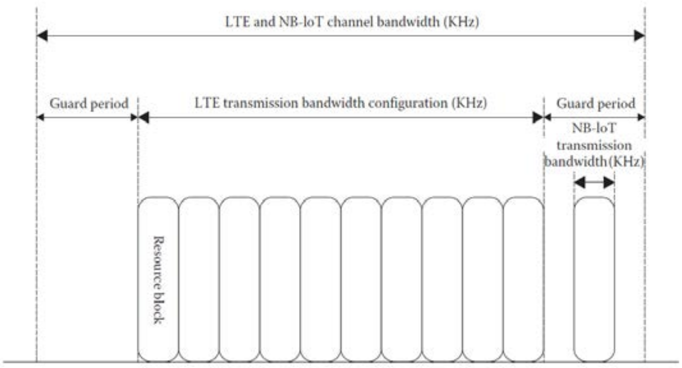
In figure 3.10 it is shown an example of how a device is able to choose a cell: here the strongest cell is an LTE antenna, but the device is served by the nearest NB-IoT cell; the detected problem, named "partial NB-IoT deployment", is the fact that the path loss from the serving cell is huge and also there is interference with the LTE cell. As regarding in-band mode of operation, link-level coverage performance is taken into consideration [20] to choose the best cell. In guard-band mode of operation, instead, it is done repeatedly the cell re-selection to compute the best parameters.



(a) Standalone operation.

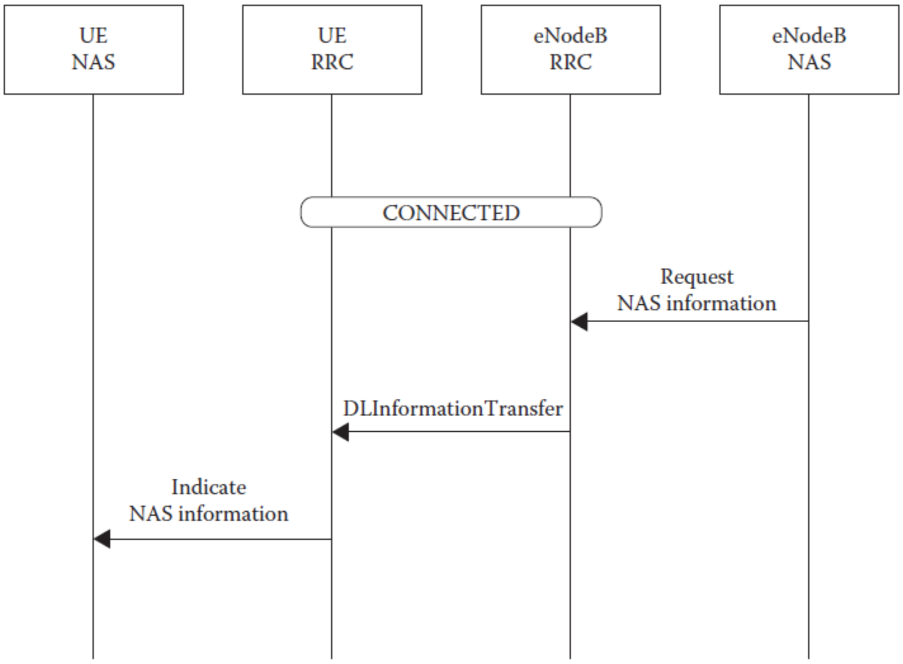


(b) Inband operation.

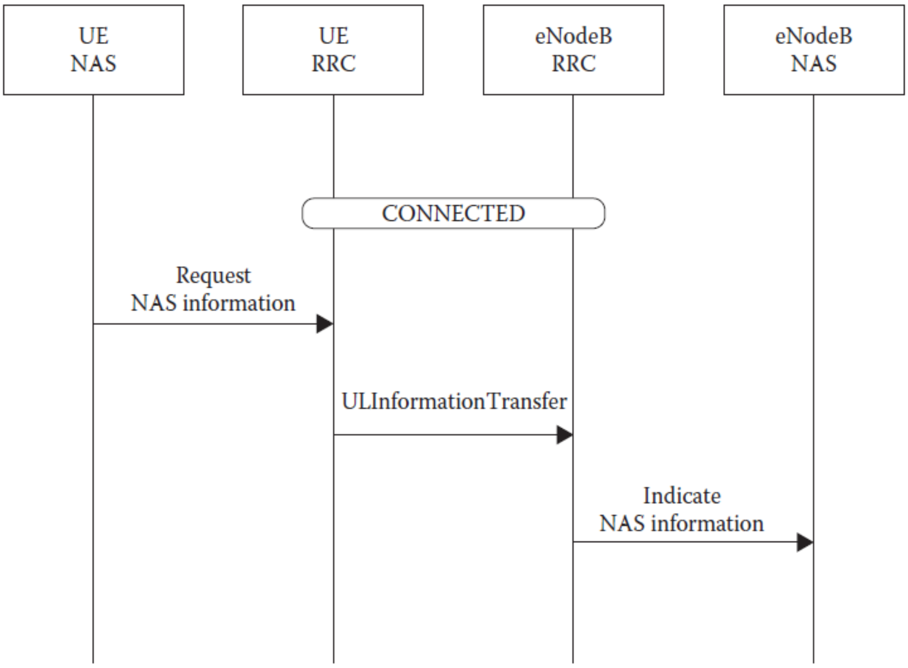


(c) Guardband operation.

Figure 3.8: Channel and transmission bandwidth for the three types of operations.



(a) Downlink connected mode.



(b) Uplink connected mode.

Figure 3.9: Downlink and Uplink connected mode.

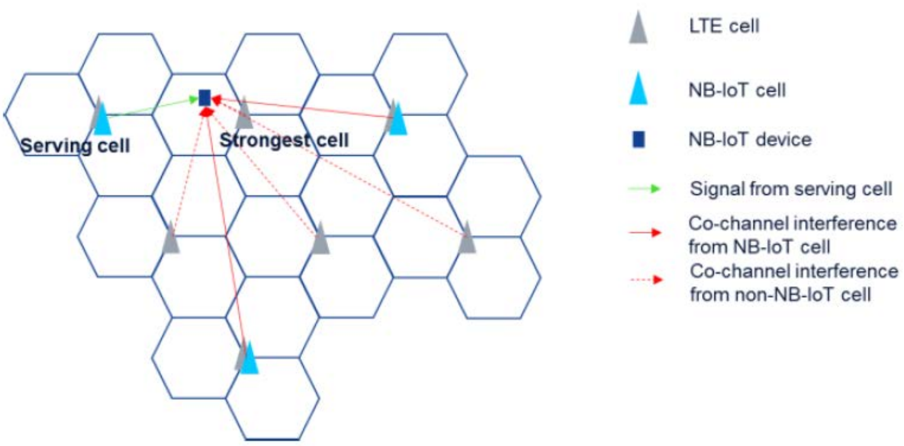


Figure 3.10: Choice of a cell for a NB-IoT device.

3.3.1 Modes of Operation

NB-IoT can transfer data in two modes: idle and connected. Idle mode occurs before the device establishes the connection with the network and, at the end of this mode, there is the connection mode where the device can transfer and exchange data. In the list below there is the description of this RRC states:

- **IDLE mode**: this includes initial cell selection of re-selection of a eNodeB, acquisition of MIB and SIB, paging and mobility;
- **CONNECTED mode**: this includes the data transfer between the UE and the eNodeB, and also the detection of any resource assigned to the UE for transmission and reception by the Narrowband Physical Downlink Control Channel (NPDCCH). This procedure is called "Random Access Procedure and Access Control" [17].

The **Random Access Procedure** follows some steps, that are drawn in figure 3.12: after synchronization with the non barred cell, the device can send a *random access preamble* to the eNodeB; if the base station detects this preamble, it sends back a message (called *Message 2*) that contains scheduling information about device's radio resources, and then it is able to calculate device transmission capability. After the reception of this message, the device includes here its identity and its scheduling request, also with its buffer and power status, and sends it back to the cell (*Message 3*). Then the eNodeB sends to UE the *Message 4* with the connection setup or resume message, after the resolution of contentions, caused by multiple devices that transmit the same random access preamble in the first step. Finally, the device replies with an RRC connection setup, or resume, complete message which means that it finishes the transition to connected state.

The UE, on the other hand, can optimize the connection setup with two procedures: the User Plane CIoT Evolved Packet System (EPS) and the Control Plane

CIoT EPS optimization procedures; both aims to reduce the message exchange betweeen eNodeB and UE but the first one aims to resume configurations established in a previous connection and so it minimizes the signalling needed to setup data radio bearer, while the second one allows transmission of data before setting up a data radio bearer.

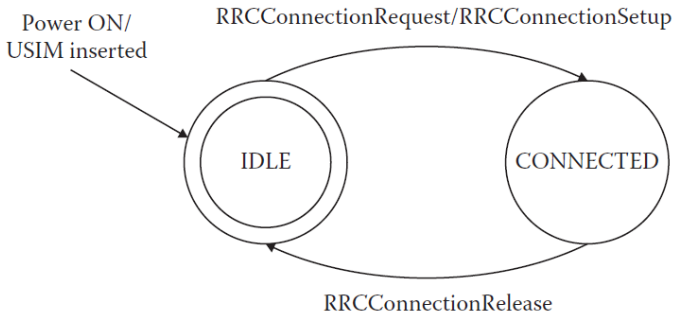


Figure 3.11: Radio Resource Control modes of operation.

In figure 3.11 there is a schema that describes the transition from IDLE mode to CONNECTED one and vice versa. The IDLE mode occurs when the UE is first powered on and the USIM is inserted. The UE then moves from the IDLE to the CONNECTED mode when the connection is established and moves back to the the IDLE mode when a connection is released. Another mode can be used and it is called **power-saving mode** and it is verified when the UE is powered-off while it remains registered with the network.

The selection and re-selection behavior can be observed when the NB-IoT evaluation kit (that will be explained at the end of this chapter, in section 3.8), used to connected to the NB-IoT infrastructure for this thesis, is powered on and the USIM is inserted: the led-light blinks very fast for 200ms and goes slowly for 1800 ms. When the connection is established the light remains constant, while when the power saving is reached and the UE camps on the cell, the light flickers slowly. When data transfer is active, the light blinking becomes very quick. The following procedure describes the AT-Commands to search for the network and to open a socket service.

- AT+QICSGP = context ID, protocol type (IPV4, IPv6), APN, username, password. This command is used to configure parameters of a TCP/IP context.

- AT+QIOPEN = context ID, socket service index, socket service type (TCP, UDP), IP address, domain name address, remote port, local port, access mode (buffer, direct push, transparent). This is used to open the socket service. An example is the socket open to the "www.google.com" website: AT+QIOPEN = 1, 1, "TCP", "172.217.23.110", 80, 1234, 1.

- AT+QISEND = connected ID. This is used to send data in access or direct push mode. The network will reply with "SEND OK" when the connection has been

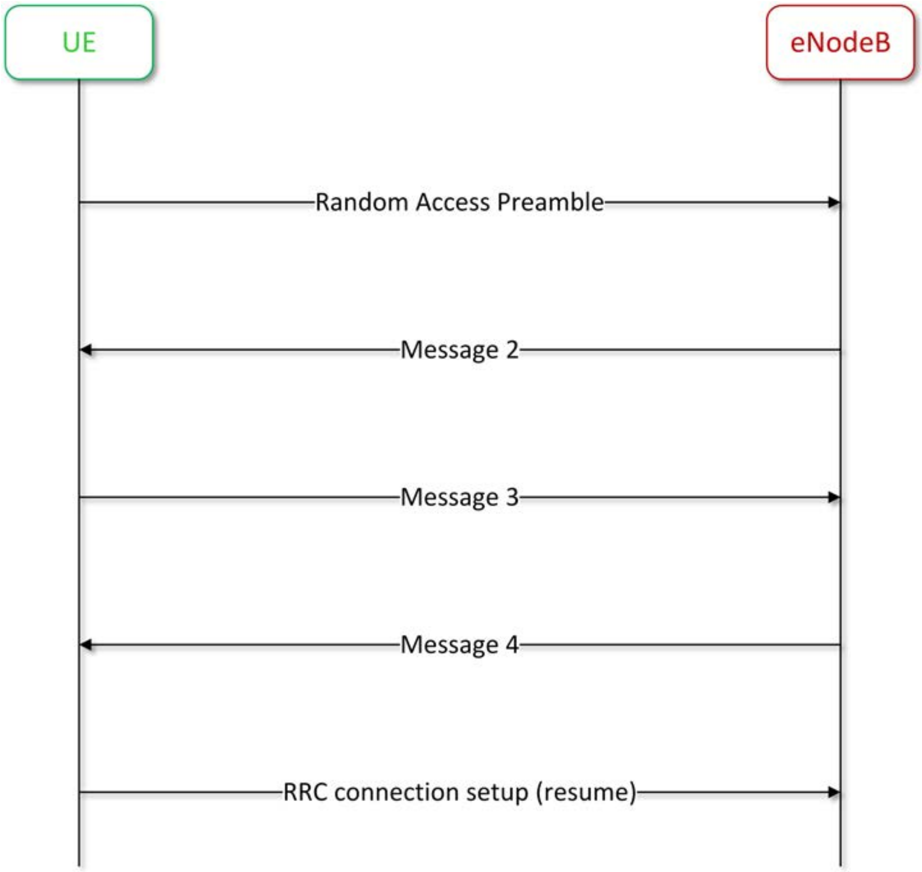


Figure 3.12: NB-IoT Random Access Procedure.

established and the sending is successful. The context ID is the same for all the commands and is chosen by the AT+QICSGP.

As shown before, a device can support multiple downlink or uplink carrier frequencies to the cell: this is called **Multi-Carrier Support**. This behavior is used to provide a load-balancing between the large number of NB-IoT devices on different carriers and then to avoid contention between NB-IoT devices and achieving higher throughput [18].

There is a distinction between two types of carriers: *anchor carrier* and *non anchor carrier*. The first one is used to do the synchronization and to carry system information, while the second one can convey transmitted data and it is used during the Random Access Procedure, that has been mentioned at the beginning of this paragraph. When a device is in idle mode it camps on an anchor carrier and when it needs to be switched to the connected mode, it will camp on a non-anchor carrier. After the device complete the data session in the connected mode, it comes back to the anchor carrier to be put in idle mode.

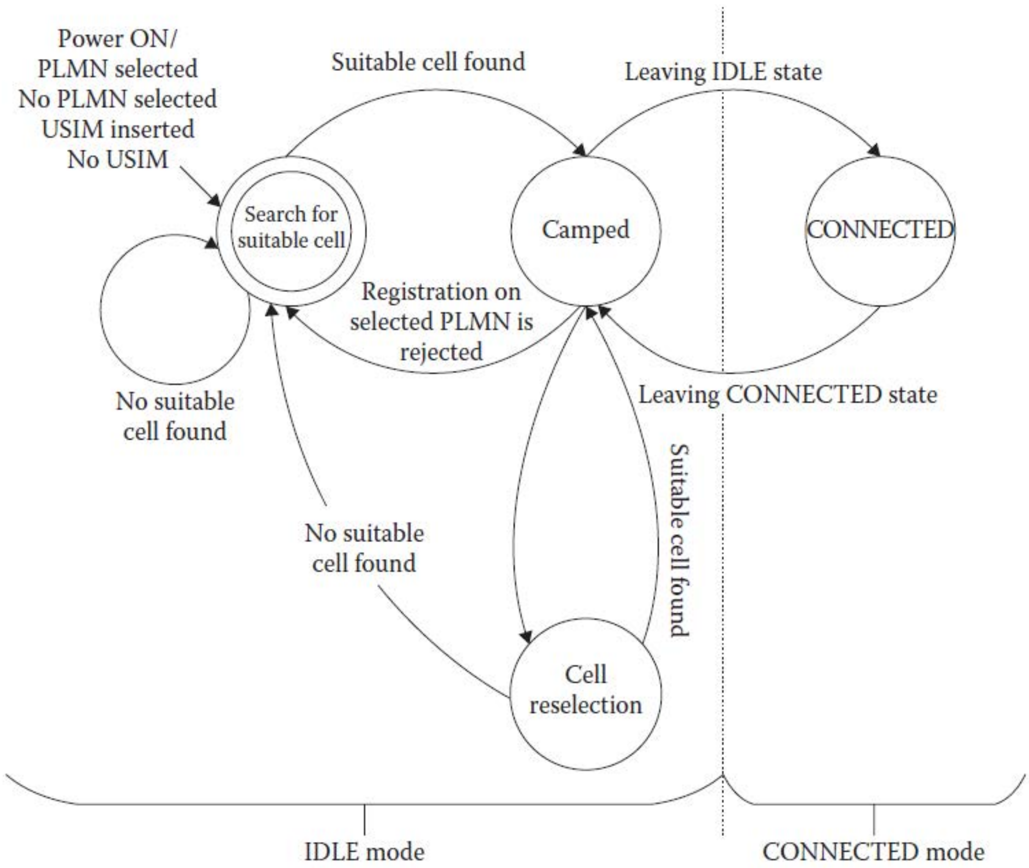


Figure 3.13: Cell Selection and Re-Selection.

3.3.2 Logical Channels

In RRC are defined control-plane logical channels that are used to transmit and receive RRC messages to and from the eNodeB, respectively. They are listed below:

- Broadcast Common Control Channel (BCCH) for receiving broadcast MIB and SIBs;
- Common Control Channel (CCCH) and Dedicated Control Channel (DCCH) to exchange RRC messages;
- Dedicated Traffic Channel (DTCH) for exchanging data-plane traffic carried on DRBs.

Logical channel are mapped to transport channel in the MAC sublayer and then are mapped to physical channel at the Physical sublayer.

3.4 Packet Data Convergence Protocol Sublayer

The PDCP sublayer is the more thin layer in the NB-IoT architecture and it is used to control-plane and data-plane. Its main functionality is to provide integrity and security protection to exchanged packets [18]. It has the following functionalities:

- assignment of a sequence number to the transmitted SDU and handled the sequence number of received SDUs.
- packets header compression and decompression;
- ciphering (encryption) and deciphering (decryption) of PDUs;
- integrity protection and verification of control plane PDUs only;
- ordering received PDUs and detection of duplicated packets.

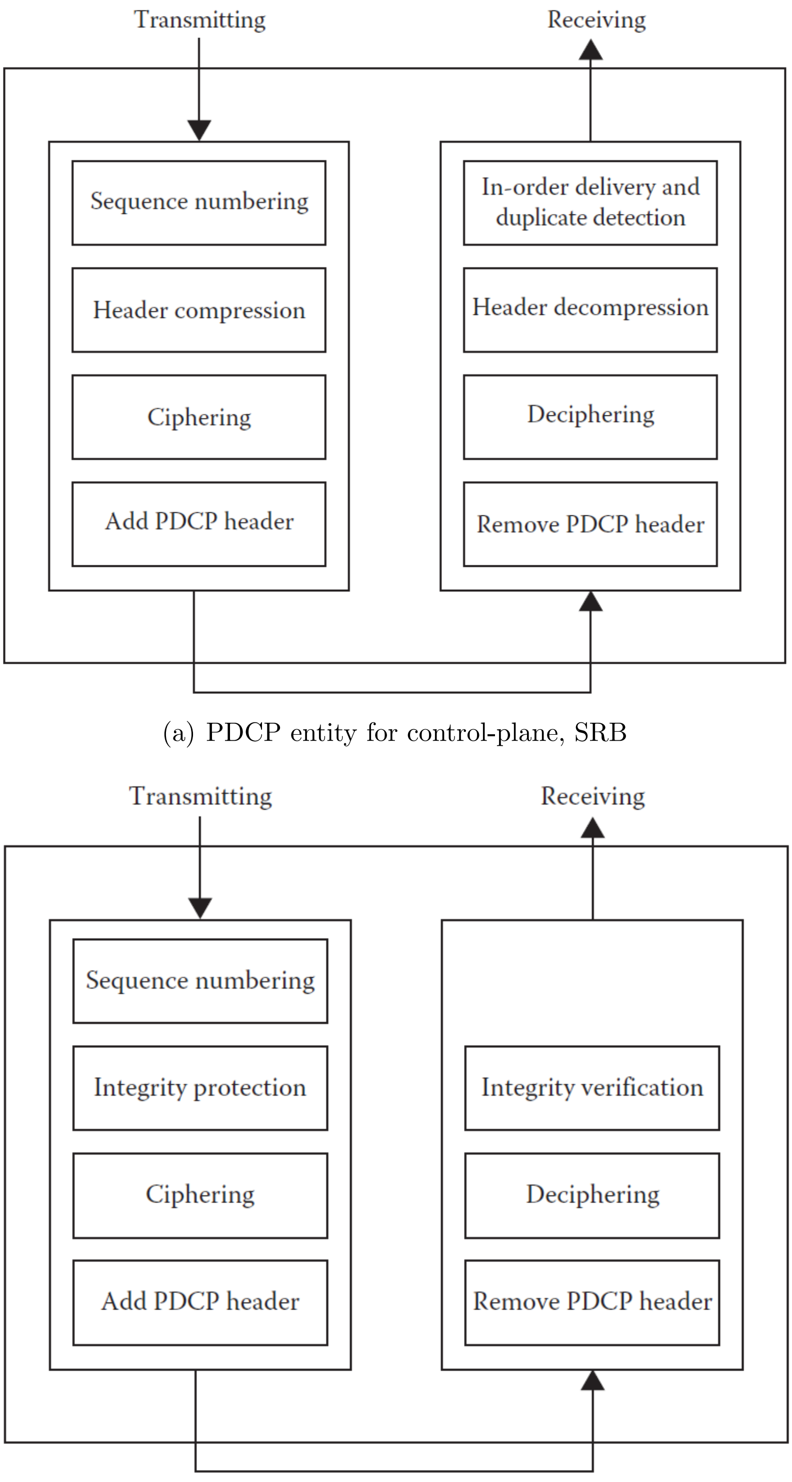
In figures 3.14(a) and 3.14(b) there are a description of what happens in this sublayer in Signalling Radio Bearer and Data Radio Bearer, respectively.

3.5 Radio Link Control Sublayer

RLC is the most important sublayer in the NB-IoT protocol stack, because it is responsible for the reliable transfer and to guarantee that PDUs can be received to the intended receiver. It operates in three modes: transparent mode, unacknowledgement mode and acknowledgment mode. In the first one a received PDU from RRC sublayer is queued for transmission into a buffer until an uplink opportunity is indicated by the MAC sublayer. Into the unacknowledgement mode, the UE detects if the RLC PDU received is duplicated and then re-order all the received PDUs; the difference between unacknowledgement mode and the acknowledgement one is that in this last mode an ACK is sent back from receiver to transmitter in case of errors or missing PDUs, when the receiver is in phase of re-ordering. These three types of transmissions are drawn in figures 3.15(a), 3.15(b) and 3.15(c), where at the bottom there are logical channels and the dotted line represents the radio interface.

3.5.1 RLC Services

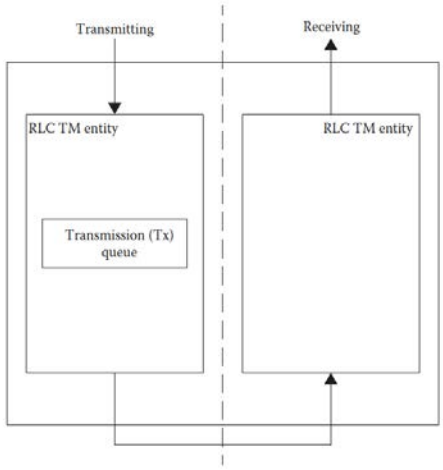
RLC provides services to the upper layers, like data transfer in Transparent Mode and Unacknowledgement one, while in Acknowledgement Mode there is also an indication of the successful delivery status of PDUs. Lower layers, on the other hands, provides notification of a transmission opportunity and, of course, data transfer services [22].



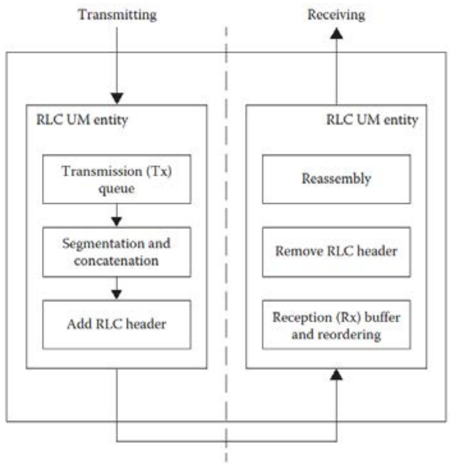
(a) PDCP entity for control-plane, SRB

(b) PDCP entity for data-plane, DRB

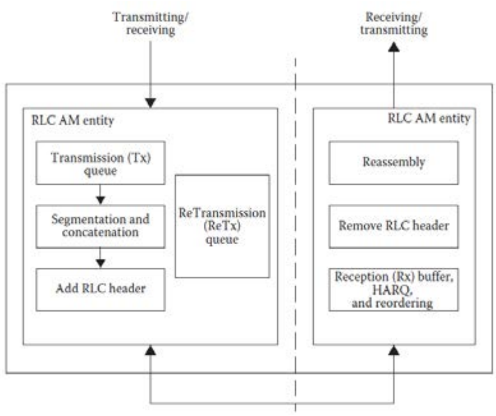
Figure 3.14: PDCP entities.



(a) Transparent mode.



(b) Unacknowledgement mode.



(c) Acknowledgement mode.

Figure 3.15: Transmission and reception modes in RLC.

3.6 Medium Access Control Sublayer

The MAC sublayer is the lowest sublayer that interfaces directly with the physical one. Its functionalities are:

- random access and contention resolution procedure;
- hybrid ARQ operations;
- priority scheduling for data and signalling RBs;
- mapping of logical channel to and from transport ones (figure 3.16);
- buffer status, data volume and scheduling requests reporting;
- discontinuous reception procedure to achieve power-saving mode.

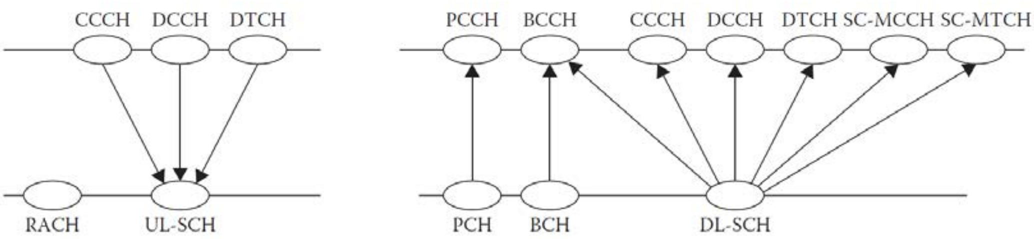


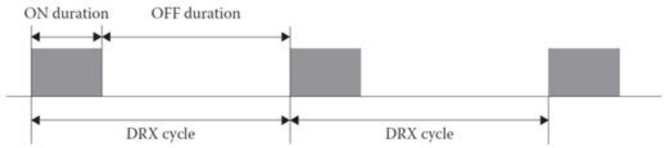
Figure 3.16: Mapping of logical and transport channels.

In figure 3.16 there is a representation of the mapping between logical and transport channels fro Uplink and Downlink procedure, respectively. All of the names written here, are described in section 3.7.2. The only difference with PHY channels is that here there is Random Access Channel which has no corresponding logical channel, since the random access message and procedure are originating and received by the MAC sublayer only [18].

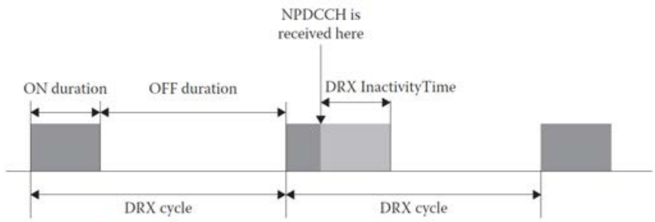
The most important procedure done in the MAC sublayer is the *Random Access Procedure*: this is already explained in section 3.3.1, but here is reported since it is the first operation that a UE initiates to connect to an eNodeB. The main purpose of this procedure is to achieve uplink synchronization and to obtain the grant to start the RRC connection and NAS Attach procedure.

3.6.1 Discontinuous Reception

The most important aspect of MAC sublayer it the Discontinuous Reception (DRX), which grants the preservation of energy and UE's battery. The UE sleeps during the DRX and it occasionally wakes up to monitor NB-PHY DL Control Channel



(a) Sleeping phase in DRX.



(b) Wake up phase in DRX.

Figure 3.17: DRX procedure when there is no NPDCCH received and when one NPDCHH is received.

(NPDCCH), respecting a specific time interval; on the other hand the eNodeB configures the DRX parameters to make UE able to establish connection with RRC or to instantiates re-configuration procedures.

A DRX cycle starts in a subframe that has this offset:

$$[(SFN \times 10) + subframe - number]mod(drxCycle) = drxStartOffset \quad (3.6.1)$$

here, SFN means System Frame Number (that is a number between 0 and 1023), the drxCycle indicates the length of the DRX cycle in subframes (and includes ON time followed by OFF time interval) and drxStartOffset is, obviously, the number used to calculate the starting subframe number for a DRX cycle. When the DRX cyle begins, the OnDuration timer starts and during this interval, UE monitors NPDCCH for any grants but still sleeping, as shown in figure 3.17(a). Otherwise, during the OnDuration timer, if an NPDCCH indicates a downlink or uplink transmission, the UE wakes up and starts the drx-InactivityTimer, as shown in figure 3.17(b). An important aspect is that DRX cycle starts if and only if all the timers are stopped.

3.7 Physical Sublayer

The PHY sublayer is responsible for physical channels, transmissions and reception of MAC PDUs. It is the bottom layer in the NB-IoT protocol stack and it is really what is used in this thesis, because with the help of this layer it is possible the communication from the NB-IoT Development Kit to the eNodeB (reference in 3.8).

3.7.1 Transmission schemes

To analyze signals in NB-IoT a technique called *framing* is used the : a specific portion of the signal is taken into consideration and processed separately. In this way the receiver is able to quickly decode the signal and reduce possible errors. The coding and decoding of the signal in NB-IoT is done in the access stratum: here time frames are used and their structure is illustrated in figure 3.18. Into the highest level there is one hyperframe cycle, of more or less 2 hours, that contains 1024 frames. One frame, with duration of 10 ms, consists of 10 subframes that are splitted into two slots of 0.5 ms, as shown in figure 3.18. Each slot, in DL, is then divided in OFDM symbols in time domain and 12 subcarriers in frequency domain, as shown in figure 3.19(b)), and one symbol is composed by a signal with its Cyclic Prefix (CP), positioned at the beginning of the signal, and, generally, with duration, for example, of 4.7μ.

Each hyperframe and subframe are labeled with a Hyper System Frame Number (HSFN) and a System Frame Number (SFN), respectively. NB-IoT supports a subcarrier spacing of 15 kHz, for which each frame contains 20 slots, [17], but in UL, the protocol supports an additional subcarrier spacing of 2.75 kHz and each frame is divided then into five slots, each of 2 ms (this is shown in figure 3.19(b)).

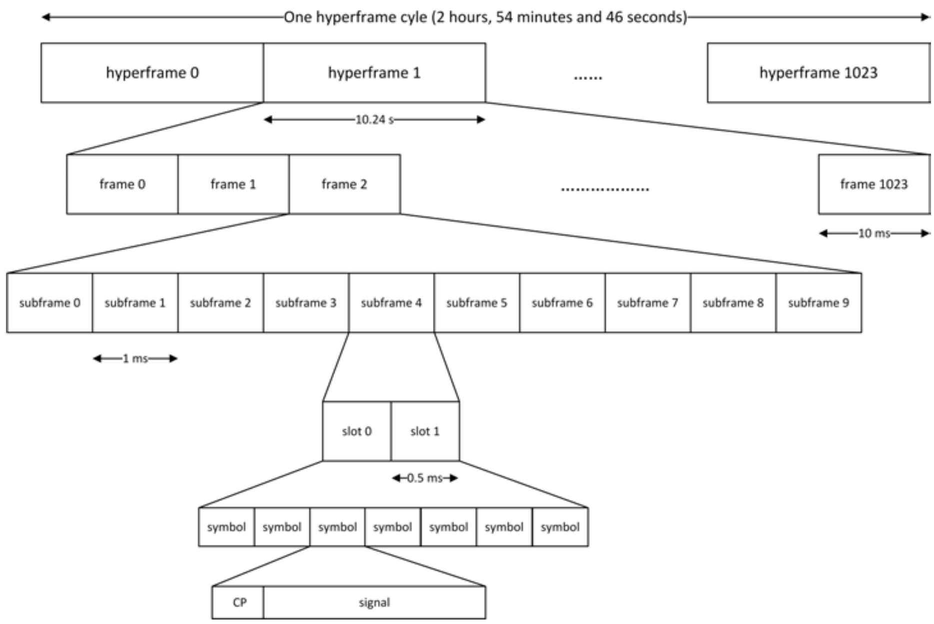
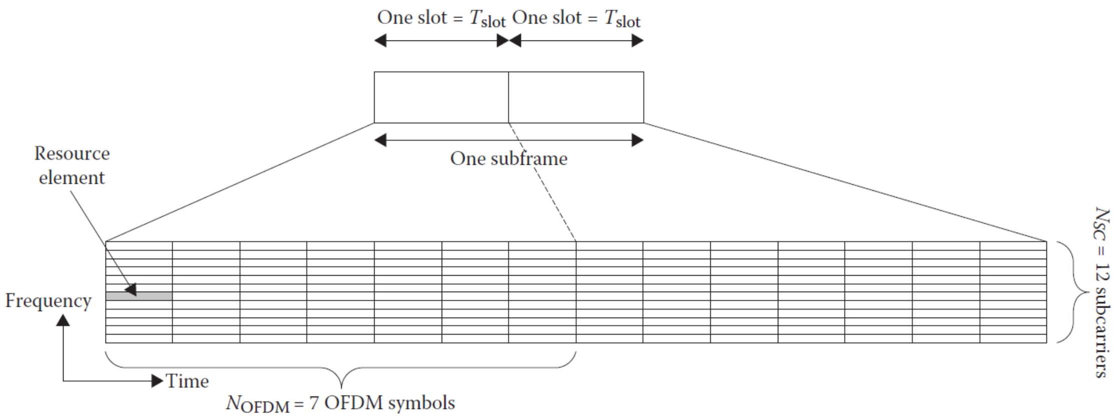
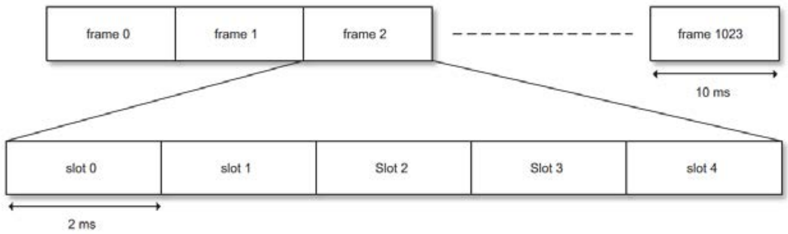


Figure 3.18: LTE NB-IoT frame structure for 15 kHz subcarrier spacing.

In the specific case of downlink transmission, one slot can contains symbols and the overall matrix, made by such symbols, is called Physical Resource Block (PRB). In figure 3.20 there is shown a PRB, which occupies two consecutive slots. The PRB is used to specify the mapping of physical channels and signal into Resource Elements (REs), which are the smallest physical channel unit and is identifiable by its subcarrier



(a) Downlink resource grid.



(b) Additional subcarrier spacing in NB-IoT UL signal transmission.

Figure 3.19: Channel and transmission bandwidth for the three types of operations.

index k and the symbol index l within the PRB. As an example, a PRB spans 12 subcarriers over 7 OFDM symbols: this means that there are $12 \times 7 = 84$ REs in this PRB.

For the UL case, the resource unit (RU) is the one that is used to mapping the UL physical channels to the REs; they depends on the subcarrier that is configured and on the number of subcarriers that are allocated in the UL transmission.

Downlink operation

To do downlink operations, NB-IoT employs Orthogonal Frequency-Division Multiple Access (OFDMA), which is been explained in section 2.3.2. The DL waveform has 12 subcarriers and it is identical for stand-alone, in-band and guard-band modes of operation, which are explained in section 3.2.1. NB-IoT supports one or two logical antenna ports to do the DL, and when there are two logical antenna ports, they are based on transmit diversity using Space-Frequency Block Coding: here, logical antenna port 0 transmits symbol pair (s_{2i}, s_{2i+1}), while the port 1 transmits symbol pair (s^*_{2i}, s^*_{2i+1}). Each pair is then mapped to consecutively available two REs within the OFDM symbol. The equation 3.7.1 show how such modulated symbol pairs are

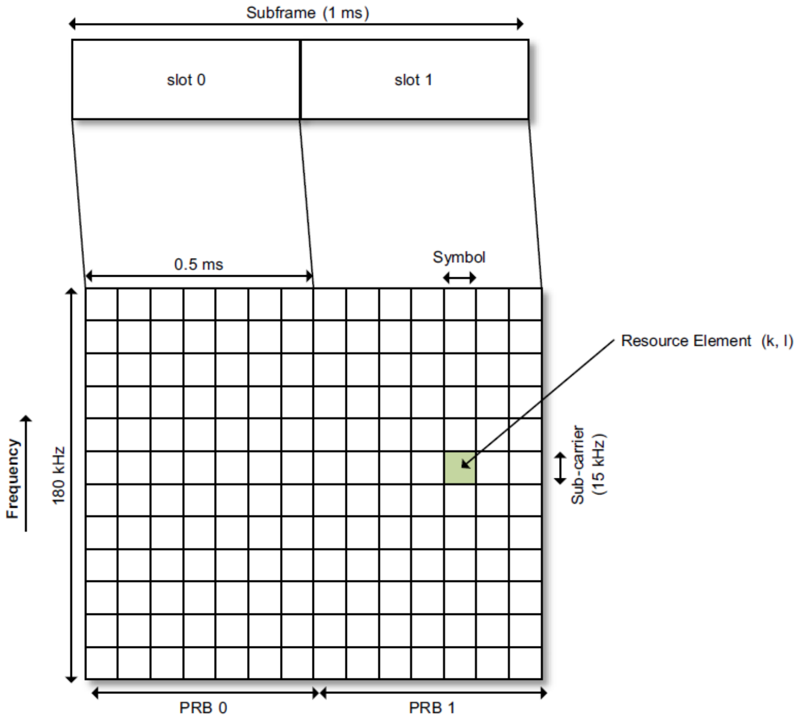


Figure 3.20: Additional subcarrier spacing in NB-IoT UL signal transmission.

calculated:

$$\begin{bmatrix} y_{2i}^{0} \\ y_{2i}^{1} \\ y_{2i+1}^{0} \\ y_{2i+1}^{1} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & j & 0 \\ 0 & -1 & 0 & j \\ 0 & 1 & 0 & j \\ 1 & 0 & -j & 0 \end{bmatrix} \begin{bmatrix} Re(s_{2i}) \\ Re(s_{2i+1}) \\ Im(s_{2i}) \\ Im(s_{2i+1}) \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} s_{2i} \\ -s_{2i+1}^{*} \\ s_{2i+1} \\ s_{2i}^{*} \end{bmatrix} \tag{3.7.1}$$

where $Re(x)$ is the real part of symbol x, while $Im(x)$ represents the complex part of the symbol and y_x^p is the precoded symbol.

The mapping operation from logical antenna ports to the base station physical antenna ports is important for the NB-IoT in-band operation, because many devices are equipped with single narrow bandwidth antenna and this offers low spatial diversity and limited frequency diversity; moreover to achieve extended coverage range, a repetition-based link adaptation is required. To allow coherent combining to be done by the device and optimize its receiving performance, DL waveform must be transmitted with a continuous and stable phase trajectory: in such way the device is able to detect the received signal far below the thermal noise floor [17].

Uplink operation

In NB-IoT communication SC-FDMA modulation is used to perform uplink signal transmission, as already explained in 2.3.3, and it uses 15 kHz subcarrier spacing. In NB-IoT are supported two types of transmission: multicarrier and single carrier and this affects the waveform. So, generally speaking, it can be said that for NB-IoT multicarrier UL transmission it is used SC-FDMA modulation, while in the single carrier UL transmission it is used OFDM. The description of the DL can be used to talk about UL, as they use the same waveform transmission and the same mapping operation.

3.7.2 Physical Channels

As in LTE, in NB-IoT there are three categories of channels, which are explained below [23]:

- **physical channels**: that carry user data and control messages;
- **logical channels**: which provide services for the MAC layer;
- **transport channels**: which offer information transfer to MAC and higher layers.

To carry information among layers, NB-IoT performs a mapping between logical and physical channels, and this is shown in figure 3.21.

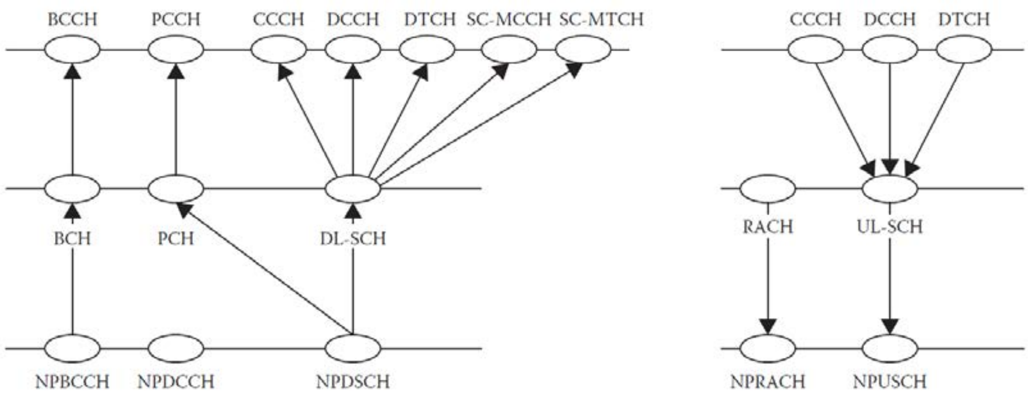


Figure 3.21: Mapping between transport channels and physical ones for UL and DL operations.

Logical channels can be divided in two groups:

- control channels, which carry the control plane information, include:
  - BCCH Broadcast Control Channel, for system information to all devices connected to the eNodeB;

  - PCCH Paging Control Channel, used for paging information when searching a unit on a network;
  - CCCH Common Control Channel, for random access information like the connection setup;
  - DCCH Dedicated Control Channel, to carry user-specific control information, like power control and handover;
  - SC-MCCH Single Cell Multicast Control channel, that is used for multicast reception information in the single cell case;

- traffic channels
  - DTCH Dedicated Traffic Channel, that is used for user data transmission;
  - SC-MTCH Single Cell Multicast Channel, used for multicast data transmission.

In transport channels there are some types of channels that are specific of the DL or UL operations, which do the mapping from transport channel to the logical ones:

- Downlink:
  - BCH Broadcast channel;
  - PCH Paging channel;
  - DL-SCH Downlink Shared Channel;
- Uplink:
  - RACH Random Access Channel;
  - UL-SCH Uplink Shared Channel.

The last channels that are written in figure 3.21 are the physical channels, and they are:

- Downlink:
  - NPBCCH Narrowband Physical Broadcast Control Channel: this is used for receiving broadcast control PDU, called MIB-NB, from the eNodeB.
  - NPDCCH Narrowband Physical Downlink Control Channel: it carries only control information;
  - NPDSCH Narrowband Physical Downlink Shared Channel: this is used to carry traffic for DL-SCH and PCH, by carrying only one resource block for a user equipment per subframe;
- Uplink:

– NPRACH Narrowband Physical Random Access Channel: it transmits random access preamble, which is composed by four symbol groups transmitted without gaps on a single subcarrier, to achieve transmit diversity;

– NPUSCH Narrowband Physical Uplink Shared Channel: it is used to transmit uplink transport block (at maximum one per carrier).

3.8 Use Cases and Deployment

During the initial state of this thesis, a real problem was to find a development kit compatible with the networks available in Padova to test NB-IoT connection; those networks are nowadays at early stage of deployment and little information about them was available. Different people have been contacted and finally it was possible to select a candidate kit, for which however there was no guarantee that it would work. Checking and setting parameters have been one of the personal contribution to this thesis. Another important contribution is to test such network with some temperature sensors that communicate also with the Azure Cloud. In figure 3.22 there is shown the preferable candidate for the evaluation kit, also used in this thesis.

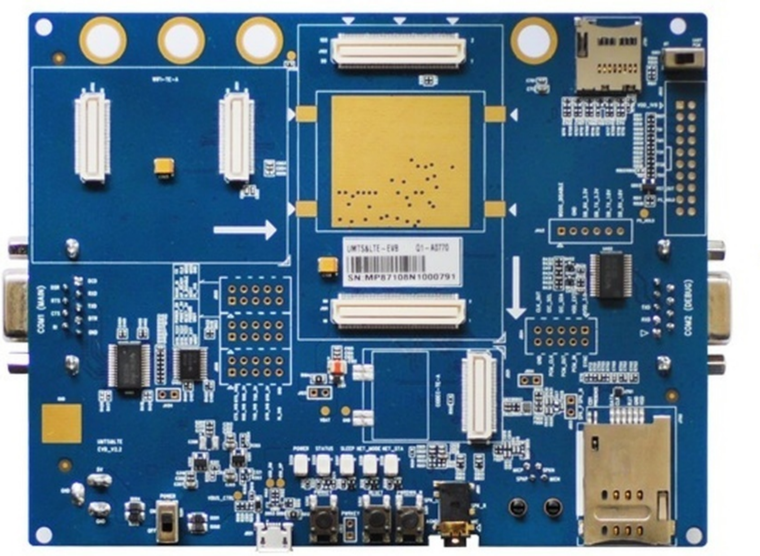


Figure 3.22: NB-IoT development evaluation kit.

Chapter 4

IoT Cloud Systems

Internet of Things is a network of objects in which devices, machines, vehicles and other physical devices are interconnected and equipped with smart systems for sensing, computing and communicate. When such smart devices are connected to the cloud, large amount of data can be processed and analyzed to create useful information for the end users. Moreover, the intrinsic limitations of mobile devices, like the battery life, are lighted by taking advantage of powerful resources of the Cloud. Traditional centralized cloud architectures are not useful to provide such advantages to the IoT world, because they have excessive network load, high end-to-end service latency and unbearable energy costs. To overcome these problems, an edge-computing approach has been adopted in recent years: cloud architectures become increasingly distributed as they are splitted in small edge cloud nodes, which made the cloud infrastructure well organized into hierarchical layers, with different computing and storage capabilities; all of these things make the *IoT Cloud Networks* very near to the end user. All the advantages of the IoT Cloud Networks are listed below [25]:

- *low latency*: cloud services can be placed at the edge of the network to achieve close proximity with the final user, in order to have real-time interactions;

- *high reliability*: services can be replicated across the highly distributed platform to achieve an increment of the fault tolerance;

- *reduced operational cost*: the large input/output requirements that some functions request, can be placed close to sources/destinations to reduce network load and associated operational costs;

- *high flexibility*: this is the most important advantage, because the virtualization allows sharing heterogeneous physical infrastructure among multiple services, without the need of dedicated deployments;

- *location awareness and mobility support*: the location of mobile users can be used to personalized contextual services;

- *scalability*: the storage and computing distributions allow scaling the number of connected devices and services without saturating the internet network.

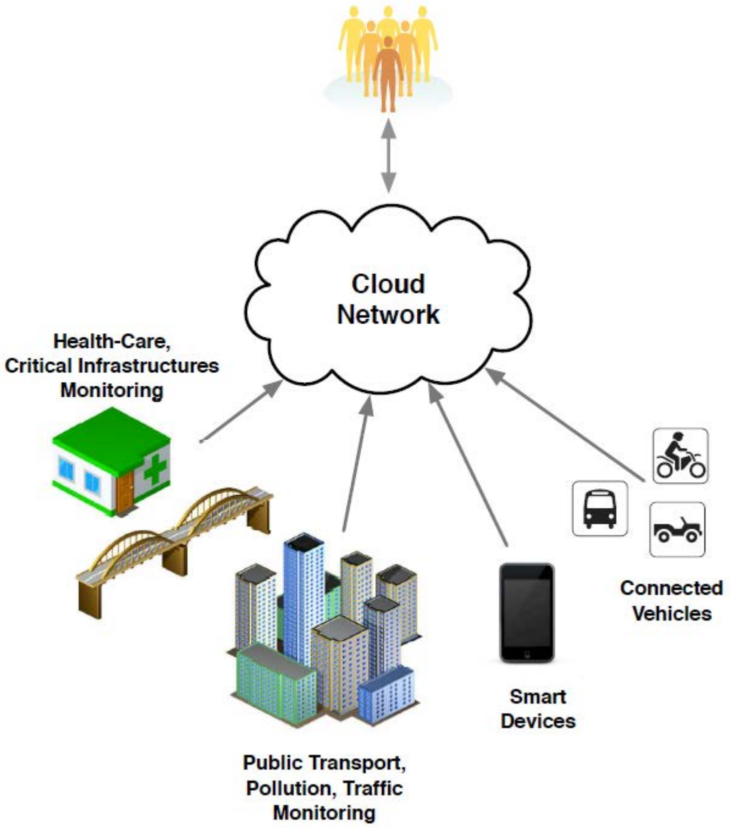


Figure 4.1: IoT cloud network.

In this thesis this type of edge-cloud is used to develop a Java application that is able to interface with an infrastructure of Temperature sensors controlled by the human voice. This is possible to reach by the use of some small cloud nodes, or **Foggy Nodes**, that are part of the Microsoft Azure cloud, like IoT Hub, Cognitive Services and Stream Analytics. The application is then deployed in a Windows Virtual Machine to be used as a Windows Service. The Foggy nodes will be explained in section 4.3.

The two perspective to view the Internet of Things is data-centric and Internet-centric. The first one is referred to the fact that all the things composing an IoT network can send data, which are then sent to the Internet, and this is why we can have two way to deal with IoT. Data-centric, in combination with Internet-centric, makes a *cloud system*.

Cloud computing is the strongest mechanism to interface with the huge amount of data that IoT sensors in a network send through the internet. With the help of services that a cloud provider supplies, data are instantaneously collected and prepared to be analyzed.

To realize the full potential of the cloud computing and of the smart sensors, it is necessary to have a combined framework and a cloud that is at the center of the architecture. In figure 4.2 there is a conceptual framework which integrates the ubiquitous sensing devices and the cloud applications. The advantage of using cloud computing is that it provides flexibility of dividing associated costs and also it is highly scalable. Also, sensors can join the network and offer their data to the cloud storage for processing. Some services offered by the cloud are based on artificial intelligence, which offers data mining and machine learning tools to convert information to knowledge. By the use of all the services provided by the cloud computing, the data generated, tools used and the visualization created “disappears” in the background, emphasizing the full potential of the IoT in various application domain [2]. Cloud computing can offers such services as Infrastructures, Platforms or Software; in this way there is three types of application that the cloud can offers:

- Software-as-a-Service (SaaS): uses Internet to deliver applications, which are managed by a third-party vendor and then offered to the final user. SaaS applications run directly through the web browser, and this means that they do not require any downloads or installations on the client side. Examples of SaaS are: Google Apps, Dropbox, Cisco WebEx;
- Platform-as-a-Service (PaaS): it provides cloud components to certain software and it delivers a framework for developers so that they can create customized applications. Examples of PaaS are some services of AWS and of Microsoft Azure;
- Infrastructure-as-a-Service (IaaS): it provides the cloud infrastructure service and it also includes servers, network, operating systems, storage, based on virtualization technology. Some examples are: AWS, Microsoft Azure and Google Compute Engine.

There isn’t a real choice between services: the only advantage is that to manage SaaS applications on large scale, it can be used PaaS in order to coordinate the cloud. Cloud computing also offers “provisioning” of resources: this concept is related to the allocation of resources and services to a customer; this means how a customer can obtain cloud services and resources from a cloud provider. This concept of provisioning can be distinct in:

- advanced provisioning: the supply can be done by a contract;
- dynamic provisioning: the user pay only resources that he uses;
- self-provisioning: it is a combination of the previous ones. For example, Microsoft Azure uses this last approach.

In figure 4.2 there is a block that is named Dynamic Provisioning: it extends the concept of provisioning in the sense that it implements the logic for cloud provisioning and managing of the virtualized resources in the private and public cloud environments.

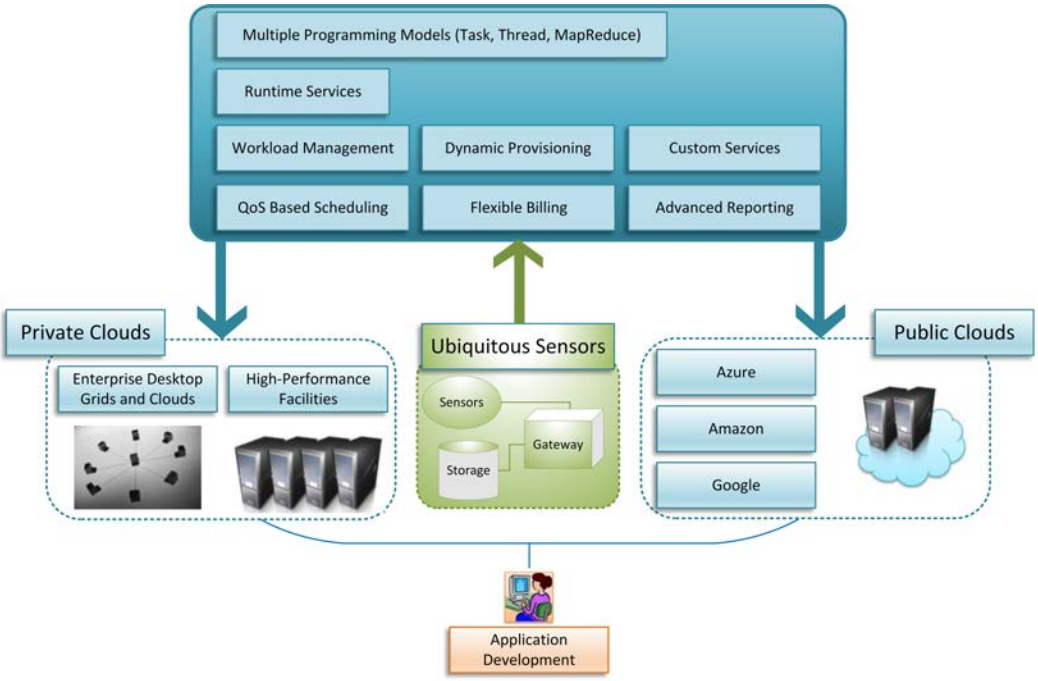


Figure 4.2: Overview of the role of the cloud in an IoT application.

4.1 IoT Cloud System Architecture

Generally, the architecture of an IoT Cloud System is defined by the composition of software units and APIs, as shown in figure 4.3.

Software-defined IoT units are the abstraction of real IoT resources with their runtime environments and capabilities. These units are used to encapsulate IoT resources and, to this end, they use well defined APIs to create a virtual runtime topology on which it is possible to deploy and execute the virtual IoT network. The main advantages of using software units are: flexible customization, reliable management of large-scale IoT cloud systems and central point management.

The background cloud engine is based on **well-defined APIs**, **unit prototypes** and **provisioning APIs**. The first one enables provisioning and controlling the units at runtime, for example, for smart metering to start the execution of sensing the environment. Moreover, well-defined APIs are able to interconnecting software-defined IoT units to dynamically deliver IoT resources to the applications. Of course, software-defined IoT units, thanks to these APIs, have a mechanism to map the the

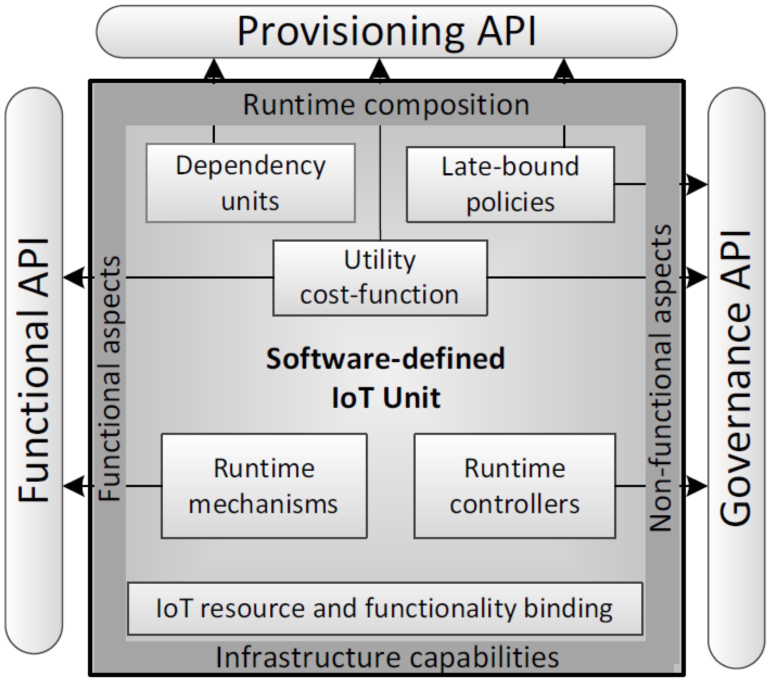


Figure 4.3: Architecture of the IoT Cloud System.

virtual resources with the underlying real physical infrastructure. An example of well-defined APIs are packages of *Microsoft Azure IoT* or *Microsoft Cognitive Services*. On the other hand, unit prototype is the resource container hosted in the cloud and enriched with functional, provisioning and governance capabilities; they are based on OS-level virtualization, for example Virtual Machines, and this concept do not poses limitations, because by utilizing the well-defined APIs, the unit prototypes are dynamically configured, provisioned, interconnected, deployed and controlled at runtime [26]. An example of a unit prototype is the Virtual Machine, called "Mida Virtual Machine", uses to deploy the application that has been developed for this thesis, which provides the control of an IoT temperature sensors network, controlled by the human voice. The last block of the cloud architecture is the provisioning APIs: they are software components that can be configured by the final user to be used within the well-defined APIs, to provide more services. Google Gson API is an example of provisioning API: even if it is not a property of Windows, it is used to give a formal model (e.g. JSON) to the telemetry packets sent by temperature sensors to the Azure Cloud.

The other two types of APIs that are in figure 4.3 are **functional APIs** and **governance APIs**. The last one is used to perform runtime control operations, like adding or removing units in the cloud topology, while the second one is the responsible for communication protocol (e.g. AMQPS, that is the one chosen for this thesis for its point-to-point communication and message queuing) and cloud storage.

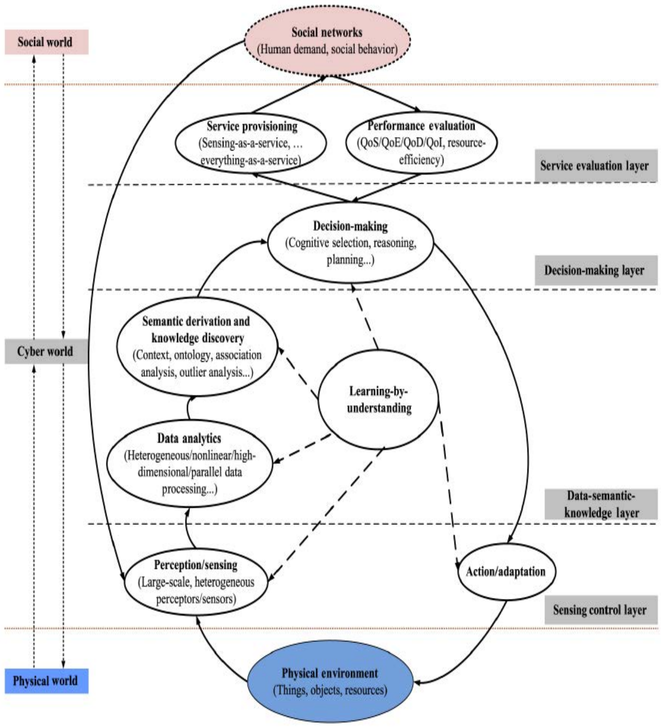


Figure 4.4: Cognitive Internet of Things framework.

4.2 Cognitive IoT

In current years a lot of research has been done to allow objects to see, hear and smell the physical world for themselves. The most of the attention when an IoT infrastructure is developed, is based on communication, connectivity, computing and other related things, but the more innovative aspect that comes from these recent research is that a lot of IoT devices are developed to have comprehensive cognitive capabilities.

A formal definition of Cognitive Internet of Things is given by authors in [28]: it is a network paradigm that interconnects physical objects that behave as *agents*, with minimum human interactions, following the methodology of *understanding-by-building* to learn from both the physical environment and social networks. The two targets of Cognitive IoT are:

- bridging the physical world and the social world together to form and **intelligent physical-cyber-social system** (iCPS);
- enabling smart resource allocation, automatic network operation and intelligent service provisioning.

In figure 4.4 there is a representation of the Cognitive IoT framework.

Referring to [29], Cognitive IoT is based in IoT paradigm, but enriched with cooperative mechanisms to achieve better performance and intelligence. Here, sensors can perceive the network conditions by analyzing the knowledge and making intelligent decisions and the Cloud provider gives them the "cognitive capacities". This is the case of this thesis: here the physical object is not able to senses the world by itself, but uses the potentiality of the Cognitive Services of the cloud; this because smart temperature sensors have limited battery life and also use NB-IoT connectivity, which is very limited in transmission and reception power, to communicate with the cloud. Cognitive services here play a central role to control temperature sensors, because with the human voice it is possible to start the environment monitoring and to stop this communication; the cloud cognitive service is then able to convert the voice in a command. Furthermore the sensor is able to detect that the temperature is more than 30 Celsius degree: it sends to the cloud a string with the details and, on the other hand, the cloud sends back a voice alarm.

4.3 NB-IoT in the Edge Cloud

One of the targets of NB-IoT development is to simplify the Radio Access Network Protocol (RAN) by removing functionality not needed for applications in the IoT: in this way the IoT system is more cost-efficient to deploy. NB-IoT, in recent years, is the most attractive network for implementation in cloud computing platforms [30]. Although the implementation of applications in RAN protocols can be done in cloud platform called **C-RAN**, NB-IoT applications, instead, can be also deployed in cloud, but which is called **Fog Cloud**. The difference between C-RAN and Fog Cloud is that the last one has reduced capability than the first one, because while it uses the same resources, such as networking, compute and storage, the difference is based on demand of the applications when low latency, massive connections, geo-distributed sensor networks and fast mobile applications: the fog node reply much faster and more dynamically that the C-RAN to these requests. So, Fog Computing is a good solution to address the demand for massive connections and low-latency applications, because it pushes some processing tasks from the cloud servers to the Fog Nodes [31].

4.3.1 Fog Node Architecture

The main characteristic of the Fog node is the distribution of virtual server extended from the cloud center, which are geo-located close to the mobile devices, sensors or IoT devices. The server chosen to deploy such Fog Computing architecture is the *Green Server*, which is a server based on Arm 64-bit. The Operating System chosen for this architecture is the Ubuntu latest version (19.04), in which authors in [31] had proposed to add, on top of it, a lightweight IaaS cloud system called OpenStack, to unify software execution environment. Also it is used a cloud service middleware, which provides interfaces to the various applications or service providers. In [31] NB-

IoT devices are viewed as fog nodes: using such devices, it is achieved extremely-wide-coverage locations of devices, since the major propriety of NB-IoT is precisely achieve wide coverage for massive connections at low cost. In [32] there is a mathematical model to locate NB-IoT devices in the random environment: authors use probabilistic distance tools, that are in contrast with the most of the existing models for analyzing positioning, because they are based on deterministic scenarios. In this study some system factors are used to analyze the positioning: required processing gain, path-loss exponent, number of participating eNodeBs, coordination among participating eNodeBs and network traffic load from non-participating eNodeBs; also are taken into consideration how directional antennas and power control impact on the location awareness.

In figure 4.5 there is shown the fog node computing concept in the NB-IoT system architecture.

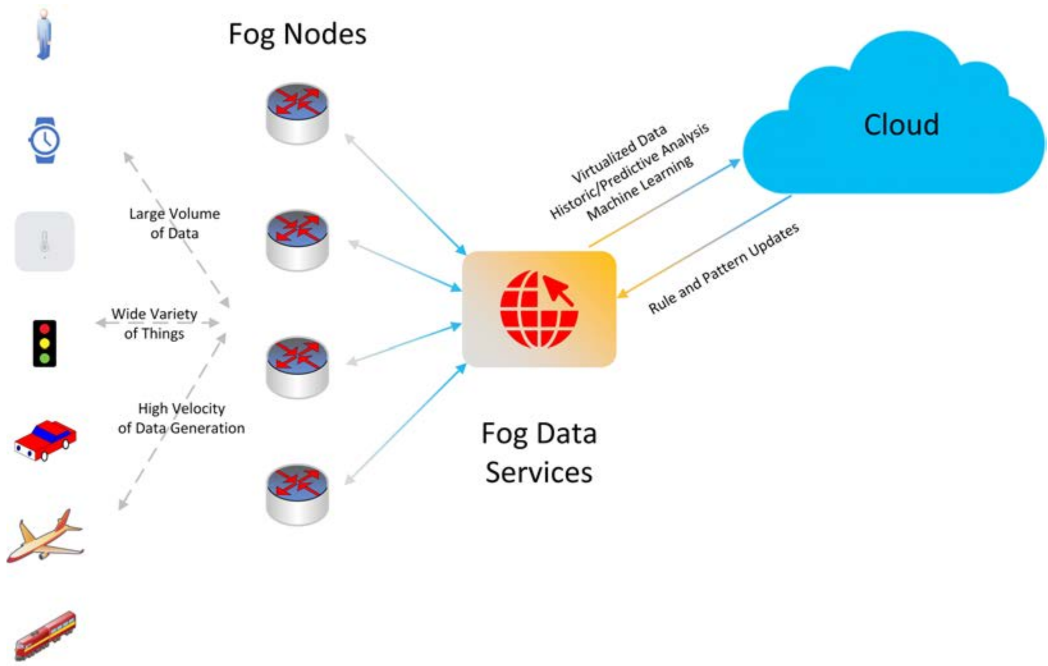


Figure 4.5: Fog node computing architecture.

4.4 5G Cloud-Based IoT

Recent researches have been predicted that in 2020 there are more than 20 billion of IoT connections in the world. In the 5G IoT will play a central role because it has the advantage of making new profits for mobile operators.

In existing IoT network communications a dedicated Core Network is used to build IoT communications, which includes Packet Gateway (P-GW), Serving Gateway (S-GW) and Home Subscriber Server (HSS) (that are already explained in section 3.2): this dedicated Core Network is necessary to have IoT Connection Management Platform (CMP), which is built for user management, where users are authenticated with SIM cards, and Device Management Platform (DMP), that is only for device management, as its name said. The existing 2G, 3G and 4G networks can not satisfy all requirements that IoT communications needs. Also, a dedicated centralized Core Network and the platforms for connection and device management are not applicable for IoT business development, especially due to the rapidly growth of IoT business. A solution is proposed by authors in [33], which have been proposed a four-layered architecture for 5G massive IoT connections; layers are *sensor layer*, *access layer*, *data center layer* and *cloud layer* and they are represented, with all the related details, in figure 4.6. Sensor layer is the responsible of devices which can execute low-power computing and communication; devices can use multiple wireless communication technologies, like LoRa, Zigbee, Bluetooth, WiFi, etc, to be connected with the access layer. The access layer contains gateways, 2G/3G/4G and 5G cellular base stations and NB-IoT. In this way different IoT use case can use different bandwidth, data rate, latency and cost: according to these difference in use case, devices are distribute to different wireless interfaces, using different transmission technologies. Data center layer includes Edge Data Center and Center Data Center and, finally, devices are connected to the cloud layer where business management platforms and application servers are deployed. In the cloud layer CMP and DMP platform are deployed here and also it does a centralized control, a management of global resources and also it allows data storage. This new network architecture is very interesting because it takes into consideration the fact that IoT devices are powered by batteries and then a lightweight IP protocol is used, and also it consider different communication technologies for different use cases.

4.5 Comparing platforms: Azure VS AWS

Until now Microsoft Azure was discussed more than other platform. This is the moment to compare it to other architectures and the most famous counterpart is Amazon AWS.

The first question that can be asked when someone talks about cloud computing is if Data Center are better than Cloud. And the answer is: it depends! It depends on what is the business of the company, but if we talk about IoT, it is better to have a Cloud System. Because data center is the ideal solution for companies that need customization and dedicated system that give them a full control of their data and equipment, but in massive sensor connections, the amount of data is huge and it is prohibitive to control all the equipment. Furthermore, data center has limited storage capacity, because once it is installed, nobody is able to change the amount of memory it has without purchasing and installing more equipment. Cloud computing is the best choice because it is scalable to a company needs and it has potentially

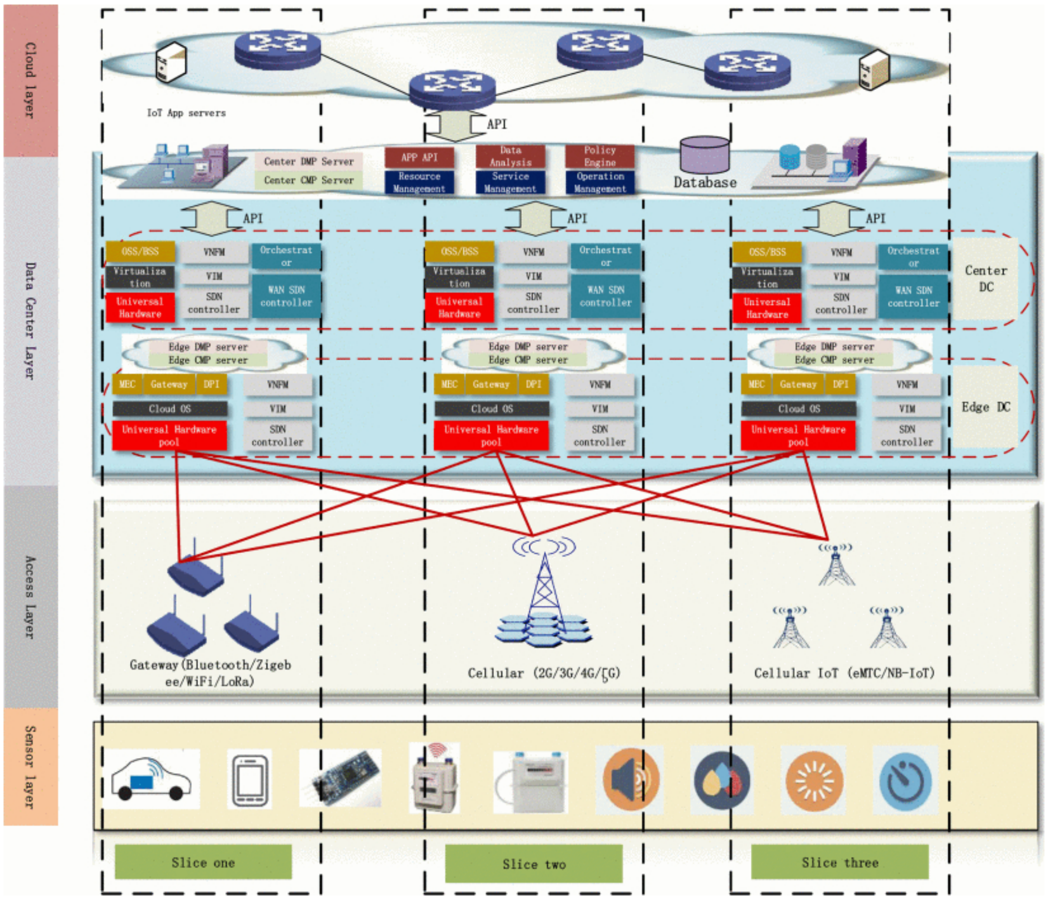


Figure 4.6: Massive 5G IoT Cloud-based network architecture.

unlimited capacity, but the disadvantage is that a third party (the cloud vendor) manage company's data.

In the cloud world there are some cloud vendors that offer various services, like Google Cloud Platform, Amazon AWS, Microsoft Azure. Since Google Cloud Platform is the least used, in this thesis there is a comparison between the most fierce opponents: Microsoft Azure and Amazon AWS.

Microsoft Azure was developed in 2010 and it is a flexible cloud platform that allows fast development, debugging and iteration of the applications, as well as their further management through a network of Microsoft data centers [34]. Applications can be developed with any programming languages and any tools, but the recommended tool to use is Visual Studio Code, because it has a lot of packages to develop and test the developed application. Also Azure allows integration between public cloud applications and existing IT environments. To use Azure a subscription to a Microsoft Live account with a credit card is needed. In the phase of registration, and also after this, it can be chosen all the resources that are needed. When a resource is created, like a Virtual Machine, there are a lot of options to choose, like the preferable

Operating System, and this permits a customization of the resource; an example of this is shown in figure 4.7.

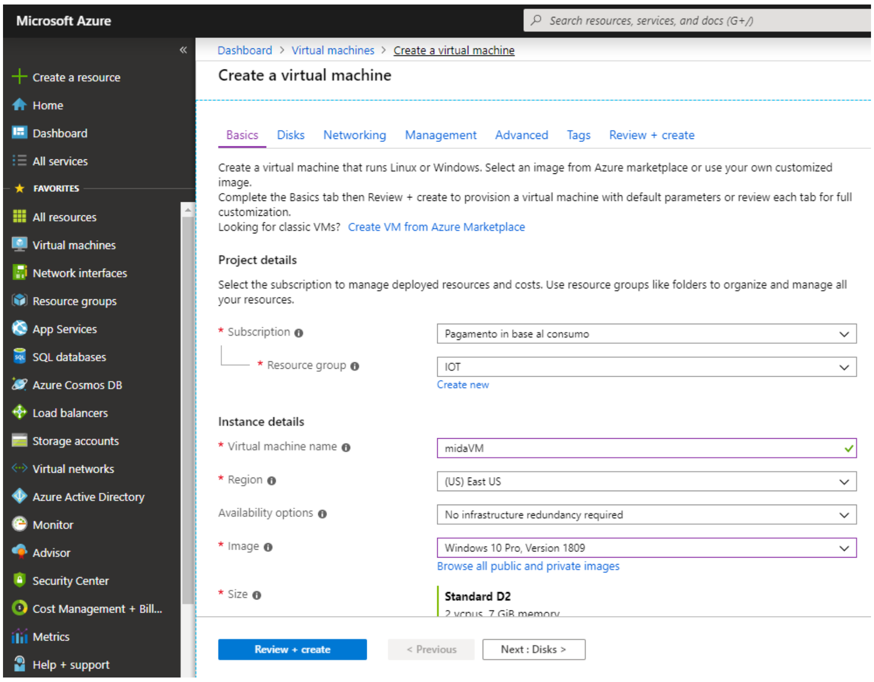


Figure 4.7: Example of how to create a VM on Azure.

On the other hand, Amazon Web Services is a cloud computing platform whose main features are low price, high speed, scalability, adaptability and security (this last one is certified by industrial security certificates). It was developed in 2006 and it has two options: Elastic cloud computing (EC2), that is the central part of Amazon's platform and it contains all the user's virtual machines, and Amazon Simple Storage Service (S3), which is an online web service of storage. Amazon doesn't provide a portal like Azure, because it provides services to other web sites or client operations for developers and not all of these services are available to the end users. Also here, it is possible to develop applications with any programming languages and any tools.

The study conducted in [34] has highlight strengths and weaknesses of two of the major cloud platforms. Microsoft Azure has an easy and intuitive interface useful to manage virtual resources, very easy also for the creation of Virtual Machines; on the other hand, Amazon AWS offers more options oriented to managing Linux VMs instead of Windows ones.

In this thesis Microsoft Azure has been chosen for the easy and more appealing portal and also because it allows the integration with Power BI, which is a powerful external tool from Microsoft, to create graphs of all the type starting from a dataset taken from IoT Hub, that is a resource of Azure.

Chapter 5

The experimental end-to-end study

In this thesis it has been developed a Java application that controls the data flow of some temperature sensors through the human voice. During the execution of the application, if the temperature is more than 30 Celsius degree, a vocal alarm is sent from the cloud to the device where the application is running. This application is located in a Windows Virtual Machine and it is running as a Windows service; the VM used to deploy such application is stored in the Microsoft Azure Cloud. Data flows are collected in the cloud by **IoT Hub** service and then they are processed and analyzed by a powerful tool of Microsoft which is called **Power BI**: to connect the IoT Hub with Power BI it is used a service of the Azure cloud, which is called **Stream Analytics**. After the data analysis, a graph is drawn and it shows all the temperatures registered by one sensor at a time.

5.1 Data analysis environment

Millions of devices, today, are connected to the internet and generate an enormous amount of data. A lot of these devices are part of a larger IoT solution, where devices send their data to the cloud for storage, processing and further analysis [35]. IoT solutions can be viewed at high level and be broken down into two essential parts: device connectivity and data processing and analysis, as shown in figure 5.2; more in detail, device connectivity is made by local devices that are connected to the cloud, while data processing and data analytics part is connected to the cloud.

The left part of the figure 5.2 shown the device connectivity with the cloud; the cloud gateway is the mediator that gathers the incoming data and makes it available for processing by other services. But the real challenges are how to create a secure and efficient connection between devices and the back-end solution and how to take care of the most used IoT devices that generally have some specific characteristics, such as: limited power resources, lack of physical access to the device, human device interaction and need of application protocols customization. Moreover, there is another key characteristic: the arrow between devices and the cloud, in figure 5.2 is bidirectional;

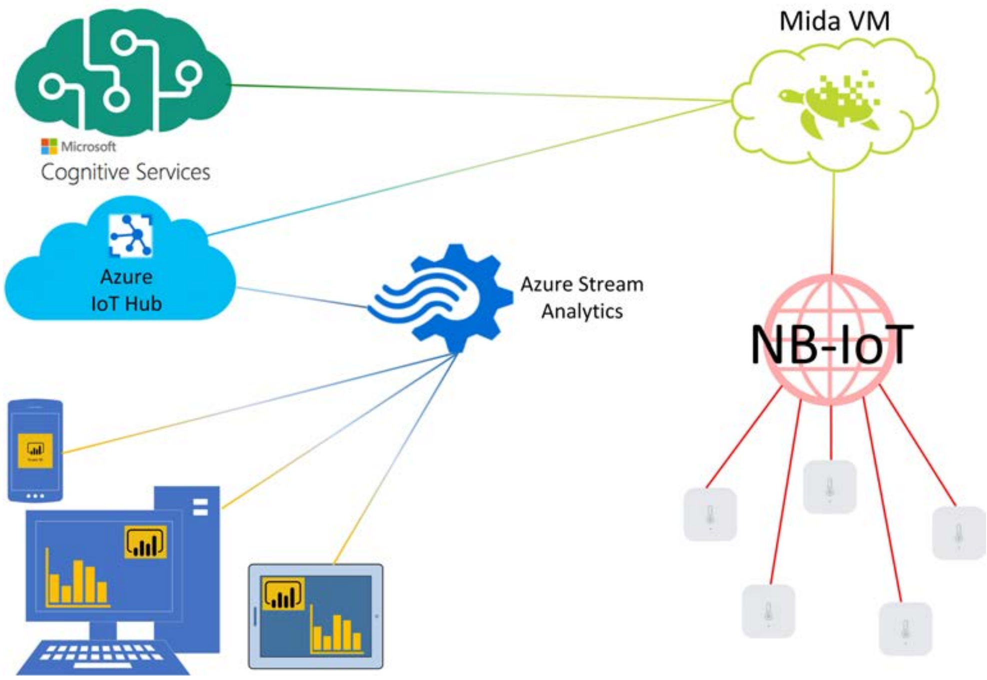


Figure 5.1: Description of the end-to-end study.

this means that IoT solutions require also *bidirectional* communication between, of course, devices and the cloud, in order to allow not only devices to send data to the cloud (device-to-cloud communication), but also the reception of processed messages and information (cloud-to-device communication). The Azure service that permits these communication is Azure IoT Hub.

Azure IoT Hub acts as an high-scale gateway that enables and manages all bidirectional communication to and from devices. It is possible to use the IoT Hub to build IoT solutions with reliable and secure communication between the billion devices that interact with the cloud, by virtually connecting any device to it.

IoT Hub is used because it provides some features that are necessary to build an IoT solutions, and they are listed below:

- **scalability**, to support hardware and software scenarios, like the huge collection of devices that are simultaneous connected;
- **security** offered to have data protection, device and user authentication;
- device monitoring to identify issues in devices connectivity and also to provide operation logs;
- **extension** of various component is allowed.

This powerful Azure resource is also able to connect through IP devices, with supported protocols like HTTP, MQTT and AMQPS. Devices which do not support

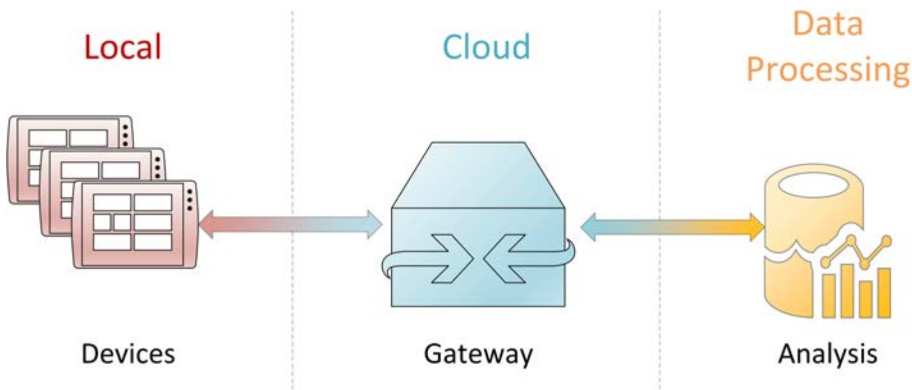


Figure 5.2: Device connectivity with the cloud and data processing.

these protocols or that cannot access directly the Internet, because they use industry specific protocols through internal networks, can still connect to Azure IoT Hub via custom cloud or field gateway, respectively [36]. In figure 5.3 there is the graph of the telemetry sent in about one hour of temperature's sensor activity.

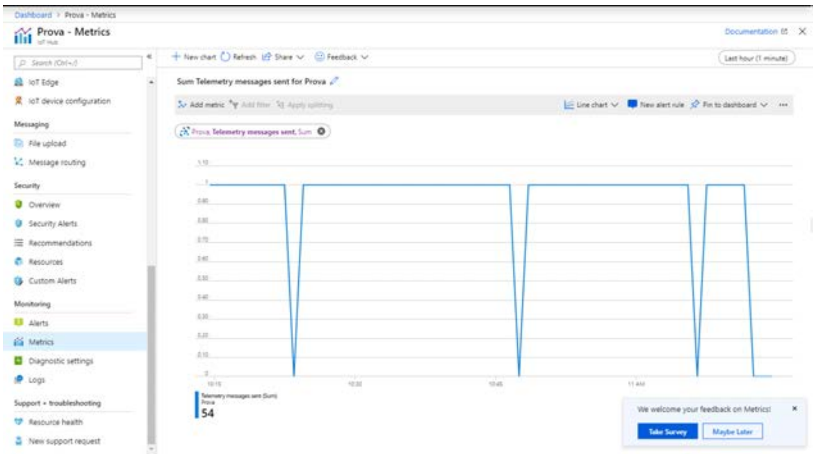


Figure 5.3: IoT Hub metrics tool.

All data send by devices are stored in IoT Hub, waiting to be picked up for further analysis; this means that IoT Hub is only a temporary holding place for incoming data [35]: data are consumed as soon as they come, but to do this there is an Azure resource that help us, and it is Azure Stream Analytics, which takes data from IoT Hub and drop them where user wants, maybe doing some data transformations. Stream Analytics is a real-time event processing which allows analytical capabilities on streaming data from devices, by focusing on making access to data. It is designed to create solution for a low cost, because it is based on the model "pay-as-you-go", which means that costs are based on the number of units and the amount of data processed. Furthermore it is very easy to use, because data transformations are done via Stream

Analytics Query Language, which is a sort of SQL language that supports IntelliSense and auto-complete directly in the query editor. Moreover, Stream Analytics is used for real-time insights via Power BI and also for long term storage for downstream batch analysis. In figure 5.4 the Stream Analytics used in this thesis is shown, where “temperaturedeviceiothub” is the name of the IoT Hub and “temperatureBI” is the name of the Power BI dashboard used.

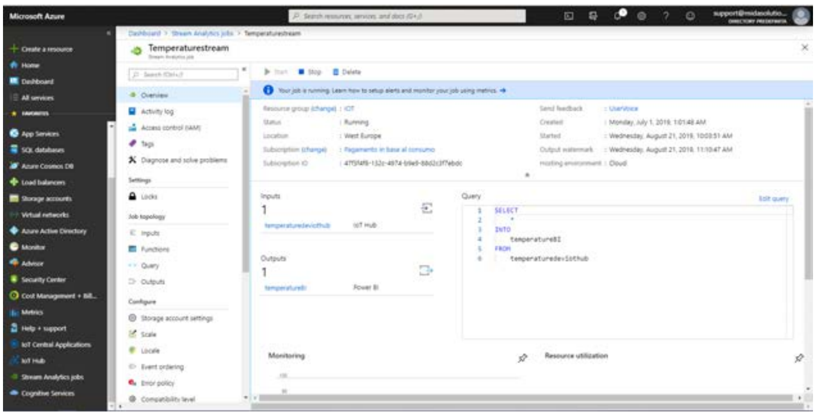


Figure 5.4: The Azure Stream Analytics used in this thesis.

Azure Stream Analytics, with the help of Azure IoT Hub, offers the capabilities to stream millions of events per second, by providing real-time analytics, However, the mechanism for visualizing the real time analytics is missed in this cooperation: Microsoft Power BI is the solution, that has been used also in this thesis, to overcome such problem.

Power BI is a business analytics service offered by Microsoft; its target is to provide interactive data visualizations with a simple interface, where users can create their own reports and dashboards for every IoT applications [37]. In figure 5.5 there is shown how the cooperation between the IoT Hub, the Stream Analytics and Power BI can be done: the first two services are part of the cloud, while the last one is an external tool.

Figure 5.5: Connection between Azure IoT Hub, Azure Stream Analytics and Microsoft Power BI.

When data arrive at Power BI, the user must creates a dashboards, which is a graphical interface to collect all the dataset. Into the dashboard there is the possibility of creates a report that contains the analysis of the processed data and it is possible to view graphs. An example of the graph resulting from the IoT application developed for this thesis is in figure 5.6; in this figure the red line points out the 30 degrees threshold, beyond which the alarm is started. The x-axis is the time in which temperature are taken, while the y-axis represents the degrees.

Figure 5.6: Power BI temperature graph.

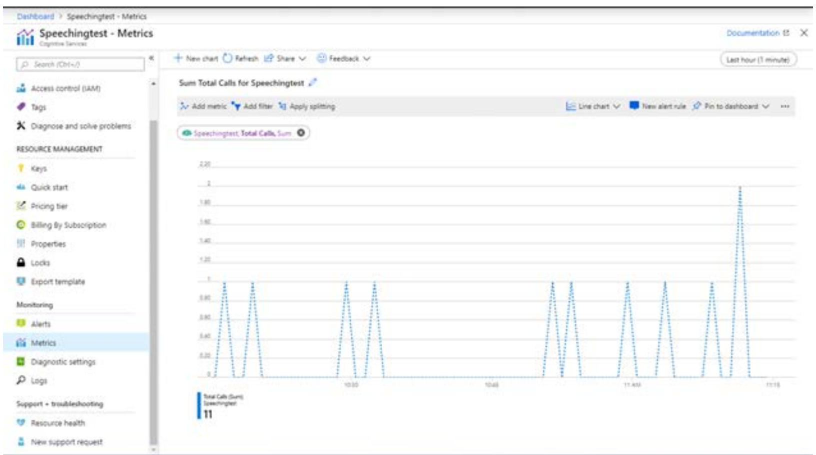


Figure 5.7: Graph of all the call done by device to Text-to-Speech cognitive service.

5.2 Cognitive Services

Azure Cognitive Services give to an IoT developed application the AI capabilities of recognize the human voice, reply with an human-like voice, analyze data to find anomalies or to do predictions, and so on. In this thesis it has been used **Text-to-Speech** (TTS) [39] and **Speech-to-Text** [38] services. The first one has been used to send a message text from the temperature sensor which measures more than 30 Celsius degrees: the text string is received in the TTS Azure cognitive service and the cloud is able to send back a human-like voice alarm, which can be heard from the final user's computer, tablet or smartphone speakers [40]. The human-like voice is very fluid and natural-sounding, because Azure uses a well trained neural networks, whose results can be adjusted with **Speech Synthesis Markup Language** (SSML). SSML is an XML-based markup language that can be used to specifity the input text to be converted into synthesized speech by the TTS service, and developers can fine-tune the pronunciation, the volume and the speaking rate. Speech-to-Text, instead, is used to start or stop the collection of temperature sensors' data by the final user's voice. In the figure 5.7 there is shown how many calls are done by the application to the TTS cognitive service to receive the alarm voice.

Chapter 6

Conclusions

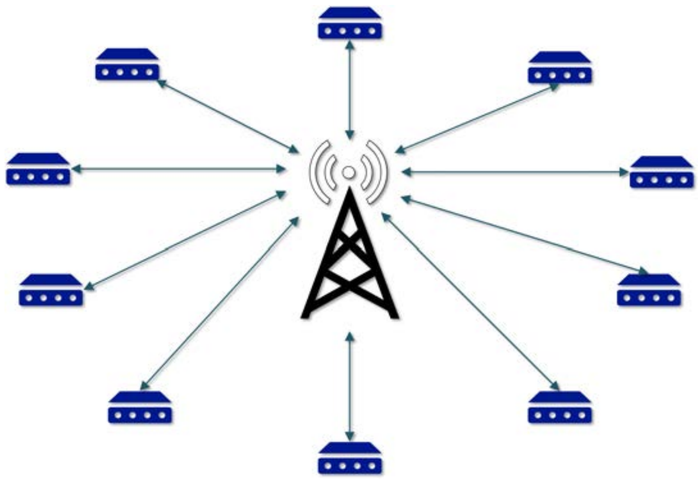


Figure 6.1: Star topology of NB-IoT sensors network.

When one wants to deploy an IoT network, a variety of wireless technologies are available. Some machine-type communications (MTC) are based on Local Area Networks, like Wi-Fi 802.11ah, or Personal Area Network, like ZigBee or Bluetooth. NB-IoT, instead, is the more favorable candidate, because it offers power efficiency, cost savings and wider deployment. Moreover, NB-IoT does not need gateways using another wireless technology to connect to devices and systems, because it can directly connect sensors to base station, with a star topology (figure 6.1). This direct connection dramatically cuts costs and provides more flexibility. Another improvement is that NB-IoT is deployed to be compatible with existing LTE networks and this is an advantage over LoRa, Weightless and SigFox protocols. 3GPP, the organization that specified the NB-IoT, moreover, has been able to provide NB-IoT with all the security measures present in LTE networks, including secure authentication, signalling protection and data encryption [21].

In this thesis it has been developed:

- a suitable connection from development kit over a NB-IoT network in Italy to the Azure Cloud;
- a test application in the Cloud collecting temperature data coming from temperature sensors via the NB-IoT network, through the development kit;
- an artificial intelligence sample application using Microsoft Business Intelligence to try the stream processing of the data and producing suitable alerts;
- a text-to-speech, speech-to-text and natural language voice recognition application using Microsoft Azure services to reproduce suitable sounds from the alarms and instructing the application to perform some actions.

Acknowledgments

The author would like to thanks Engineer M. Franchin and Engineer A. Licciardello to have given her the opportunity to take part of the Microsoft Inspire Event that has been held at Las Vegas in July 2019.

Appendix A

Digital Modulations

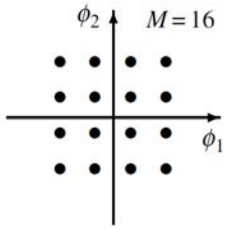
A.1 QAM

Quadrature amplitude modulation (QAM) it the digital modulation that conveys two analog message signals, or two digital bit streams, by modulating the amplitudes of two carrier waves, using the amplitude modulation (AM) analog scheme. The two carrier waves of the same frequency are out of phase with each other by 90 degrees, a condition known as orthogonality or quadrature. The transmitted signal is created by adding the two carrier waves together. At the receiver, the two waves can be coherently demodulated because of their orthogonality property. Another key property is that the modulations are low-frequency/low-bandwidth waveforms compared to the carrier frequency, which is known as the "narrowband assumption".

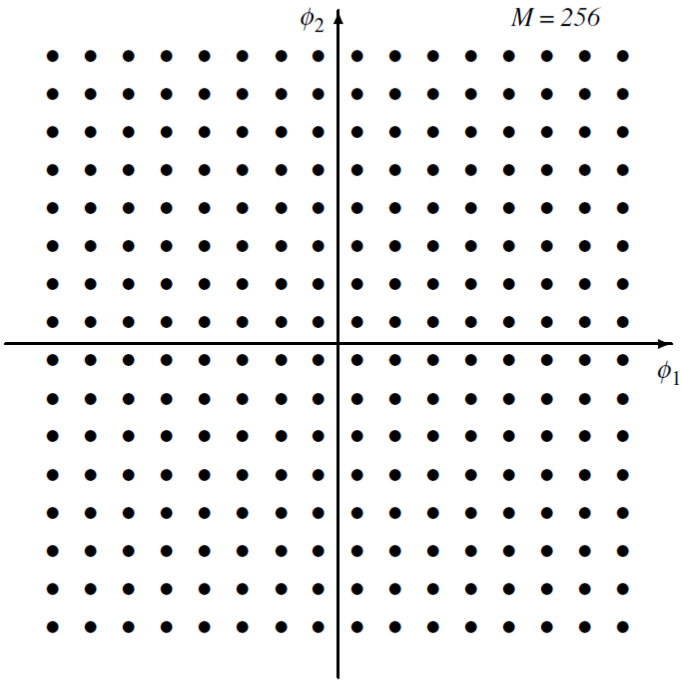
QAM is used extensively as a modulation scheme for digital telecommunication systems, such as in 802.11 WiFi standards. Arbitrarily high spectral efficiencies can be achieved with QAM by setting a suitable constellation size, limited only by the noise level and linearity of the communications channel.

Generally, it is used in the mobile communication the 256-QAM for the downlink and the 64-QAM for the uplink.

In QAM the constellation diagram is very important, because it is the graphical representation of constellation points, where they are usually drawn in a square grid with equal-spaced vertical and horizontal lines; in figure A.1 there are not drawn the lines, because they are not necessary. Since in digital telecommunication the data is usually binary, the number of points in the grid is usually a power of 2, and the QAM can be classified as: 4-QAM, 16-QAM, 64-QAM and so on. By moving to an higher-order constellation, it is possible to transmit more bits per symbol. However, if the target when a constellation is developed is to have the same energy among signals, the points must be closer together and they are thus more susceptible to noise and other corruptions; this results in a higher bit error rate and so higher-order QAM can deliver more data with less reliably than lower-order QAM, and this is only for constant mean constellation energy. Using higher-order QAM without increasing the



(a) 16-QAM.



(b) 256-QAM.

Figure A.1: Example of constellations for QAM.

bit error rate requires a higher SNR (signal-to-noise ratio) by increasing signal energy, reducing noise, or both. Communication systems designed to achieve very high levels of spectral efficiency usually employ very dense QAM constellations [41].

In application environments, such as the ones for mobile telecommunication, need a higher order QAM constellation but the multipath interference increases. There is a spreading of the points in the constellation, decreasing the separation between adjacent states, making it difficult for the receiver to decode the signal properly. So, there is the so called reduced noise immunity.

A.2 PSK

Phase-shift keying PSK modulation can be viewed as a special case of QAM, where the amplitude of the transmitted signal is constant, but it varies in phase [42]. It is a digital modulation process which transmits data by modulating the phase of a constant frequency reference signal, that is called the *carrier wave*. The modulation is done by varying the sine and cosine inputs signal in time. It is widely used for wireless LANs, RFID and Bluetooth communications.

Any digital modulation scheme uses a finite number of distinct signals to represent digital data. PSK uses a finite number of phases, where each of them has assigned a unique pattern of binary digits. Usually, each phase encodes an equal number of bits. Each pattern of bits forms the symbol that is represented by the particular phase. The demodulator, which is designed specifically for the symbol-set used by the modulator, determines the phase of the received signal and maps it back to the symbol it represents, thus recovering the original data.

There are three main classes of digital modulation techniques used for the transmission of digitally represented data:

- Amplitude-shift keying (ASK);
- Frequency-shift keying (FSK);
- Phase-shift keying (PSK).

All of them transmit data by changing some aspect of the base signal, like the carrier wave, in response to the data signal. In the case of PSK, the phase is changed to represent the data signal.

PSK schemes are represented by a constellation diagram: it shows the points in the complex plane in which the real and imaginary axes are defined as the axes in phase and quadrature, respectively, due to their 90 degree separation. The amplitude of each point, along with the in-phase axis, is used to modulate a cosine wave (or sine) and the amplitude, along the quadrature axis, to modulate a sine wave (or cosine).

In PSK, the constellation points chosen are usually positioned with uniform angular spacing around a circle. This gives maximum phase-separation between adjacent points and thus the best immunity to corruption. Since they are positioned on a circle, than they can all be transmitted with the same energy. In this way, the modulation of the complex numbers, that they represent, will be the same and thus it will be the amplitude needed for the cosine and sine waves. An example of a constellation of PSK is shown in figure A.2.

Two common examples of PSK modulation are Binary Phase Shift Keying (BPSK), which uses two phases, and Quadrature Phase Shift Keying (QPSK), which uses four phases.

A.2.1 BPSK

BPSK is the more simple form of Phase Shift Keying modulation. It uses two phases which are separated by 180 degrees and so can also be called 2-PSK.

This system handles the highest noise level, or distortion, before the demodulator reaches an incorrect decision. This makes it the most robust modulation of all the PSKs. However, it is only able to modulate at 1 bit/symbol and so it is unsuitable for high data-rate applications.

A.2.2 QPSK

QPSK uses four points on the constellation diagram, equi-spaced around a circle. With its four phases, QPSK can encode two bits per symbol in order to minimize the Bit Error Rate (BER).

The mathematical analysis shows that QPSK can be used either to double the data rate compared with a BPSK system while maintaining the same bandwidth of the signal, or to maintain the data-rate of BPSK but with an half of the bandwidth needed. In this latter case, the BER of QPSK is exactly the same as the BER of BPSK.

Given that radio communication channels are allocated and it have a maximum bandwidth, the advantage of QPSK over BPSK becomes evident: QPSK transmits twice the data rate in a given bandwidth compared to BPSK, considering the same BER.

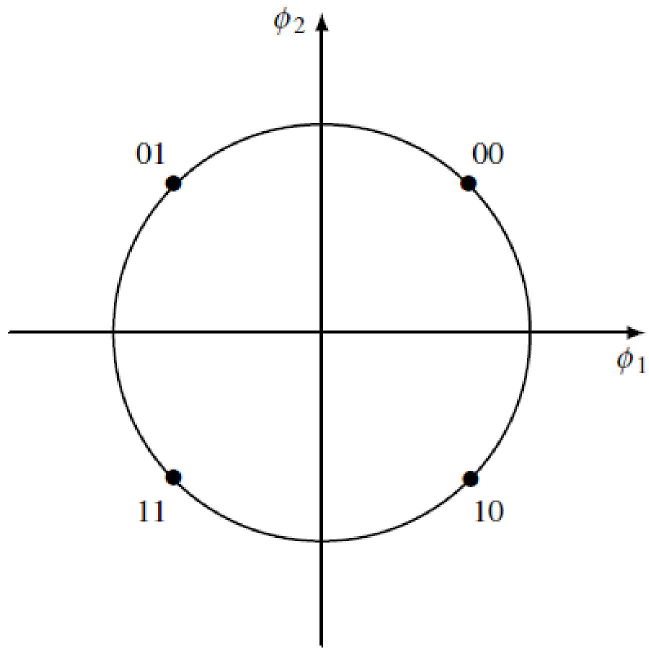


Figure A.2: Mapping of a PSK constellation.

Appendix B

Abbreviations

1G	1st Generation
2G	2nd Generation
3G	3rd Generation
3GPP™	Third Generation Partnership Project
4G	4th Generation
5G	K 5th Generation
ACB	Access Class Barring
ACK	Acknowledgement
AES	Advanced Encryption Standard
AM	Acknowledgement Mode
AMBR	Acknowledgement Maximum Bit Rate
AMD	Acknowledgement Mode Data
AMPS	Acknowledgement Mobile Phone System
ARP	Allocation and Retention Priority
ARQ	Automatic Repeat Request
AS	Access Stratum
BCCH	Broadcast Common Control Channel
BCH	Broadcast Control Channel
BSR	Buffer Status Report
C-RNTI	Cell Random Network Temporary Identifier
CCCH	Common Control Channel
CDMA	Code Division Multiple Access
CDMA2000-EV-DO	CDMA2000 Evolution Data Optimized
CE	Control Element
CEL	Coverage Enhancement Level
CIoT	Cellular Internet of Things
CMAS	Commercial Mobile Alert Service
CP	Cyclic Prefix
CRC	Cyclic Redundancy Check

CRS	Cell-specific Reference Signal
CSFB	Circuit Switched FallBack
CSG	Closed Subscriber Group
DCCH	Dedicated Control Channel
DFT	Discrete Fourier Transform
DL-SCH	Downlink Shared Channel
DRB	Data Radio Bearer
DRX	Discontinuous Reception
DTCH	Dedicated Traffic Control Channel
E-RAB	EPS Radio Access Bearer
E-UTRA	Evolved UMTS Terrestrial Radio Access
E-UTRAN	Evolved UMT Terrestrial Radio Access Network
EAB	Extended Access Barring
EARFCN	E-UTRA Absolute Radio Frequency Channel Number
EDGE	Enhanced Data GSM Environment
EGPRS	Enhanced General Packet Radio Service
eMBB	Enhanced Mobile Broadband
EPC	Evolved Packet Core
EPRE	Energy Per Resource Element
ETSI	European Telecommunications Standards Institute
ETWS	Earthquake and Tsunami Warning System
FCC	Federal Communications Commission
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FTT	Fast Fourier Transform
GBR	Guaranteed Bit Rate
GPRS	General Packet Radio Service
GSM®	Global System for Mobile Communication
H-SFN	Hyper System Frame Number
HARQ	Hybrid Automatic Repeat Request
HFN	Hyper Frame Number
HSS	Home Subscriber Service
IFFT	Inverse Fast Fourier Transform
IMEI	International Mobile Equipment Identity
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IMT	International Mobile Telecommunication
IoT	Internet of Things

IS-95	Interim Standard 95
ITU	International Telecommunication Union
LCG	Logical Channel Group
LSB	Least Significant Bit
LTE™	Long Term Evolution
M2M	Machine-to-Machine
MAC	Medium Access Control Sublayer
MAC	Message Authentication Code
MBMS	Multimedia Broadcast Multicast Service
MBR	Maximum Bit Rate
MCL	Maximum Coupling Loss
MME	Mobility Management Entity
MNC	Mobile Network Code
MPL	Maximum Path Loss
MSB	Most Significant Bit
MTC	Machine Type Communication
NACK	Negative Acknowledgment
NAS	Non Access Stratum
NPBCCH	Narrowband Physical Broadcast Control Channel
NPDCCH	Narrowband Physical Downlink Control Channel
NPDSCH	Narrowband Physical Downlink Shared Channel
NPUSCH	Narrowband Physical Uplink Shared Channel
NR	New Radio
NRS	Narrowband Reference Signal
OFDM	Orthogonal Frequency Division Multiplexing
P-GW	Packet GateWay
P-RNTI	Paging Random Network Temporary Identifier
PCCH	Paging Common Control Channel
PCH	Paging Control Channel
PCI	Physical Cell ID
PDCP	Packet Data Convergente Protocol Sublayer
PDN	Packet Data Network
PDU	Protocol Data Unit
PLMN	Public Land Mobile Network (i.e., Mobile Operator)
PRB	Physical Resource Block
QCI	QoS Class Identifier
RA	Random Access

RACH	Random Access Channel
RAT	Radio Access Technology
RB	Radio Bearer
RF	Radio Frame
RLC	Radio Link Control Sublayer
RoHC	Robust Header Compression
RRC	Radio Resource Control Sublayer
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RTT	Round Trip Time
RU	Resource Unit
S-GW	Serving Gateway
S-TMSI	SAE-Temporary Mobile Station Identifier
SC-FDMA	Single Carrier Frequency Division Multiple Access
SC-MCCH	Single Cell Multicast Control Channel
SC-MTCH	Single Cell Multicast Transport Channel
SC-PTM	Single Cell Point To Multipoint
SDU	Service Data Unit
SF	SubFrame
SFN	System Frame Number
SI	System Information
SIB	System Information Block
SINR	Signal-to-Interference-plus-Noise Ratio
SMS	Short Message Service
SRB	Signalling Radio Bearer
TDD	Time Division Duplex
TDMA	Time Frequency Division Multiplexing
TM	Transparent Mode
TMD	Transparent Mode Data
TR	Technical Report
TS	Technical Specification
TTI	Transmission Time Interval
UE	User Equipment
UL-SCH	Uplink Shared Channel
UMD	Unacknowledged Mode Data
UMTS™	Universal Mobile Telecommunications System
URLLC	Ultra Reliable and Low Latency Communication
USIM	UMTS Subscriber Identity Module
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network

List of Figures

Bibliography

[1] P.P. Ray, *"A survey on Internet of Things architectures"*, Journal of King Saud University - Computer and Information Sciences, 2018.

[2] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami *"Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions"*, Future Generation Computer Systems, Volume 29 Issue 7, 2013.

[3] A. Al-Fuqaha. M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, *"Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications"*, IEEE Communications Surveys and Tutorials, Volume 17, Issue 4, 2015.

[4] L. Vangelista, Slides of the course of Internet of Things and Smart Cities, 2018.

[5] A. Goldsmith *"Wireless Communications"*, Stanford University, 2004.

[6] Y.S. Cho, J. Kim, W.Y. Yang, C.G. Kang, *"MIMO-OFDM Wireless Communications With Matlab®"*, IEEE Press, 2010.

[7] L. Katzela, M. Naghshineh *"Channel assignment schemes for cellular mobile telecommunication systems: A comprehensive survey"*, IEEE Communications Surveys and Tutorials, Volume 3, Issue 2, 2000.

[8] J.P. Vasseur, A. Dunkels *"Interconnecting Smart Objects with IP: The Next Internet"*, Elsevier Inc., 2010.

[9] M. Conti, *"Secure Wireless Sensor Networks. Threats and Solutions"*, Springer, 2015.

[10] H.G. Myung, J. Lim, D.J. Goodman, *"Single Carrier FDMA for Uplink Wireless Transmission"*, IEEE Vehicular Technology Magazine, 2006.

[11] C.X. Mavromoustakis, G. Mastorakis, J.M. Batalla, *"Internet of Things (IoT) in 5G Mobile Technologies"*, Springer International Publishing, 2016.

[12] W.L. Stutzman, G.A. Thiele, *"Antenna Theory and Design, 3rd edition"*, Wiley, 2012.

[13] 3GPP Technical Specification Group GSM/EDGE Radio Access Network, *"Cellular system support for ultra-low complexity and low throughput Internet of Things (CIoT) (Release 13)"*, 3GPP Technical Report, 2015.

[14] L. Diaz, T. Milligan, *"Antenna Engineering Using Physical Optics: Practical CAD Techniques and Software"*, Har/Dskt edition, 1996.

[15] W.L. Stutzman, G.A. Thiele, *"Antenna Theory and Design"*, John Wiley and Sons, 2012.

[16] R. Ratasuk, N. Mangalvedhe, Y. Zhang, M. Robert, J-P. Koskinen, *"Overview of narrowband IoT in LTE Rel-13"*, IEEE Conference on Standards for Communications and Networking (CSCN), 2016.

[17] O. Liberg, M. Sundberg, Y.P.E. Wang, J. Bergman, J. Sachs, *"Cellular Internet of things, Technologies, Standards and Performance"*, Elsevier Academic Press, 2018.

[18] H. Fattah, *"5G LTE Narrowband Internet of Things (NB-IoT)"*, CRC Press, 2019.

[19] Wiliot company, *"Battery-Free Bluetooth® Technology"*, `https://www.wiliot.com/blog/demo`

[20] N. Mangalvedhe, R. Ratasuk, A. Ghosh, *"NB-IoT Deployment Study for Low Power Wide Area Cellular IoT"*, IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2016.

[21] Anritsu Company, *"NB-IoT: Characteristics and Considerations for Design and Verification"*, `https://www.anritsu.com/en-US`, 2018.

[22] ETSI TS 136 322, *"LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Link Control (RLC) protocol specification"*, 3GPP TS 36.322 v.14.0.0 Release 14, 2017.

[23] R. Ratasuk, N. Mangalvedhe, J. Kaikkonen, M. Robert, *"Data Channel Design and Performance for LTE Narrowband IoT"*, IEEE 84th Vehicular Technology Conference, 2016.

[24] P.A. Maldonado, P. Ameigeriras, J.P. Garzon, J.N. Ortiz, J.M.L. Soler, *Narrow-Band IoT Data Transmission Procedures for Massive Machine Type Communications*, IEEE Network Magazine, 2017.

[25] M. Barcelo, A. Correa, J. Llorca, A.M. Tulino, J.L. Vicario, A. Morell, *"IoT-Cloud Service Optimization in Next Generation Smart Environments"*, IEEE Journal on Selected Areas in Communications (Volume: 34, Issue: 12), 2016.

[26] S. Nastic, S. Sehic, D-H. Le, H-L. Truong, S. Dustdar, *"Provisioning Software-defined IoT Cloud Systems"*, International Conference on Future Internet of Things and Cloud, 2014.

[27] P.P. Ray, *"A survey of IoT cloud platforms"*, Future Computing and Informatics Journal, 2016.

[28] Q. Wu, G Ding, Y. Xu, S. Feng, Z. Du, J. Wang, *"Cognitive Internet of Things: A New paradigm Beyond Connection"*, IEEE Internet of Things Journal, 2014.

[29] M. Zhang, H. Zhao, R. Zheng, Q. Wu, W. Wei, *"Cognitive Internet of Things: Concepts and Application Example"*, IJCSI International Journal of Computer Science Issues (Volume: 9, Issue: 6), 2012.

[30] Y. D. Beyene, R. Jantti, O. Tirkkonen, K. Ruttik, S. Iraji, A. Larmo, T. Tirronen, J. Torsner, *"NB-IoT Technology Overview and Experience from Cloud-RAN Implementation"*, IEEE Wireless Communications Journal (Volume: 24, Issue: 3), 2017.

[31] Y.C.P. Chang, S. Chen, T.J. Wang, Y. Lee, *"Fog Computing Node System Software Architecture and potential applications for NB-IoT industry"*, IEEE International Computer Symposium, 2016.

[32] F.Tong, S. He, *"On Positioning Performance for the Narrow-Band Internet of Things: How Participating eNBs Impact?"*, IEEE Transactions on Industrial Informatics, 2019.

[33] Y. Wang, Z. Zhang, P. Zhang, Z. Ma, G. Liu, *"A New-Cloud-Based Network Framework in 5G Massive Internet of Things Connections"*, IEEE International Conference on Communication Technology, 2017.

[34] B.S. Dordevic, S.P. Jovanovic, V.V. Timcenko, *"Cloud Computing in Amazon and Microsoft Azure platforms: performance and service comparison"*, 22nd Telecommunications Forum Telfor (TELFOR), 2014.

[35] S. Klein, *"IoT Solutions in Microsoft's Azure IoT Suite. Data Acquisition and Analysis in the Real World"*, Apress publication, 2017.

[36] Microsoft documentation, *"What is Azure IoT Hub?"*, `https://docs.microsoft.com/en-us/azure/iot-hub/about-iot-hub`

[37] Microsoft documentation, *"Visualize real-time sensor data from Azure IoT Hub using Power BI"*, `https://docs.microsoft.com/en-us/azure/iot-hub/iot-hub-live-data-visualization-in-power-bi`

[38] Microsoft documentation, *"What is speech-to-text?"*, `https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/speech-to-text`

[39] Microsoft documentation, *"What is text-to-speech?"*, https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/text-to-speech

[40] L. Larsen, *"Learning Azure Cognitive Services. Use Cognitive Services APIs to add AI capabilities to your applications"*, Packt Publishing, 2018.

[41] N. Benvenuto, M. Zorzi, *"Principles of Communications Networks and Systems"*, Wiley, 2011.

[42] N. Nevio, G. Cherubini, *"Algorithms for Communications Systems and their Applications"*, Wiley, 2002.