A Survey of Cellular IoT Connectivity Solutions

Saeed Alsabbagh*, Amine Adouane†, Pengwenlong Gu‡, Cedric Adjih‡, Nadjib Aitsaadi*

*Paris-Saclay University, UVSQ, DAVID, F-78035, Versailles, France

†Algiers University, A-16000, Algiers, Algeria

‡Inria Saclay, F-91120, Palaiseau, France

firstname.lastname@*uvsq.fr, †univ-alger.dz, ‡inria.fr

Abstract—A wide range of telecommunication systems have been designed to meet the increasing demand for Internet of Things (IoT) connectivity solutions. This includes technologies operating in both licensed and unlicensed frequency spectra. The support of IoT use cases has been one of the main objectives of the fifth-generation (5G) mobile cellular networks since its beginning in 3GPP Release 15 (Rel-15). Both i) Ultra-Reliable and Low-Latency communication (URLLC) and ii) massive Machine-Type Communication (mMTC) usage scenarios are designed to accommodate IoT applications. However, new mid-range IoT applications, such as wearables, industrial wireless sensor networks, and video surveillance has requirements that can not fit into any of the 5G categories. Therefore, 3GPP in its Rel-17 has proposed a new category of 5G devices, named Reduced Capability (RedCap) devices aiming to have lower cost/complexity, and extended battery lifetime compared to the reference 5G devices. In this article, we provide a deep overview of the key IoT communication protocols. After that, we give a comprehensive insight into RedCap Rel-17, including its main features: i) UE cost/complexity reduction and ii) UE power-saving enhancement. Moreover, we discuss its impact on the overall 5G system performance, namely coverage losses and system capacity.

Index Terms—5G-mMTC, Cellular IoT, RedCap, NR-Light, LTE-M, NB-IoT, LPWAN, mid-range IoT, UE complexity Reduction, UE Power-Saving, eDRX, RRM Relaxation.

Fig. 1. M2M/IoT system components [**?**]

I. INTRODUCTION

The Internet of Things (IoT) is an essential element of the technological and societal transformation witnessed in the last decade. The effect of IoT systems spans a wide range of applications and services, including security, healthcare, transport, energy, logistics, etc [**?**]. IoT provides the capability to create smart objects that can communicate with each other and enable more efficient operations and accurate decision-making [**?**]. This would pave the way toward a far-reaching vision of an intelligent, autonomous world. By the end of 2023, the worldwide IoT market will reach around 300 billion U.S. dollars with more than 15 billion connected devices, and these numbers are expected to be doubled in the following years [**?**]. The fast growth of IoT systems is led by the recent evolution in different engineering domains, especially wireless communication systems [**?**].

Machine-To-Machine communication (M2M) usually refers to communication scenarios between a device and a server or directly between end terminals without human interaction [**?**]. A strong similarity between IoT systems and M2M exists, where IoT is considered the successor of M2M. Both technologies aim t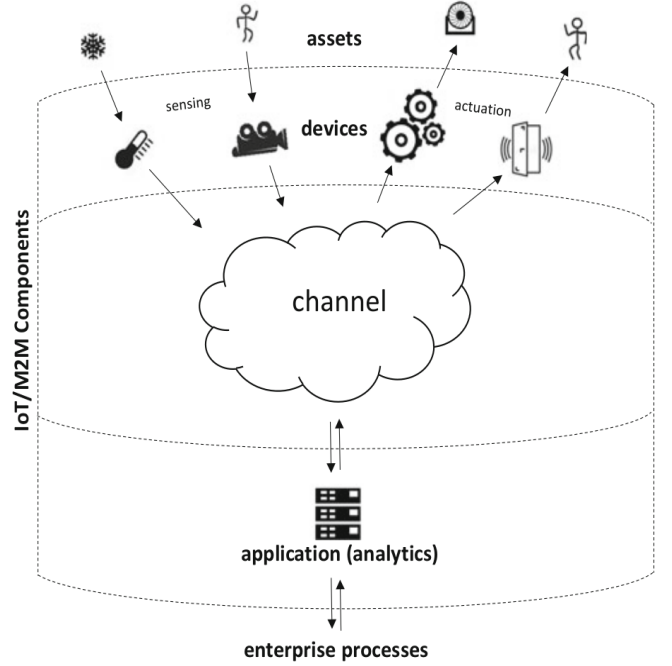o control/access physical assets through some application. Fig. **??** depicts a typical IoT/M2M system that usually consists of three main elements: 1) Devices that interact with the physical assets; 2) Communication Medium which enables the exchange of signals and; 3) Applications that can perform analytics and take actions.

Most M2M systems are built and designed in a vertical way [**?**]. This means that they deal with one type of parameter (e.g., position, humidity, pressure, temperature, etc.) and perform a particular task (e.g., turn on/off a switch). The heterogeneous nature of M2M applications has led to a wide range of proprietary networking protocols that serve a single-purpose application. This lack of standardization has limited the interaction between different applications and slowed the full automation transition. A horizontal expansion is required to accommodate the need for interoperability between different applications. Existing for many years and showing its efficiency, the Internet Protocol (IP) was the logical candidate to enable this move. An IP-based connectivity framework, linking the devices with servers running the applications, is the typical form of today's IoT systems. Fig. **??** illustrates the main difference between M2M communications and IoT
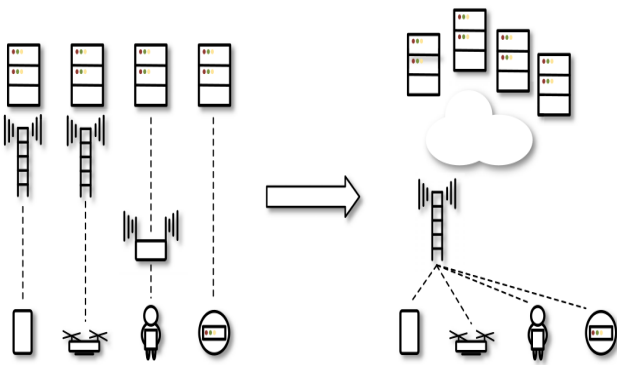
Fig. 2. M2M vs IoT Systems [?]

systems.

No unified definition of IoT systems was agreed on worldwide, where every company adopts one aligned with its market strategy. However, in 2012, the International Telecommunication Union (ITU) defined IoT systems in a more general way as follows [?]:

> "A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable Information and Communication Technologies (ICT)."

According to ITU, things can be perceived as *physical* or *virtual* objects with minimum communication capabilities, allowing them to be identified and integrated into information networks. Physical objects are types of equipment in the real world that are usually able to collect, store, and process data (e.g., sensors, controllers, actuators, etc.). On the other hand, virtual things are elements of the information world that can be stored, processed, and accessed, such as software applications and multimedia content. It is important to note that the advancement of IoT is based on the foundations of existing and evolving information and telecommunication technologies, such as Artificial Intelligence (AI), big data, cloud infrastructure, and wireless networks.

From a networking perspective, IoT is a network of networks connected through some communications and information infrastructure. IoT networks share some defining characteristics that can differ from one use case to another. Here is a list of features that are generally associated with IoT networks [?], [?]:

- Enormous scale: the number of connected devices in the network that need to be managed is magnitudes larger than traditional human-centered communication.
- Interconnectivity: in the foreseen vision of IoT, all devices must be interconnected to a global communication infrastructure.
- Heterogeneity: devices in IoT are based on different hardware and connected through variable information and communication platforms.
- Sporadic nature: the network parameters (e.g., the number of connected devices, their location, and traffic) vary in a very dynamic and unpredictable way.

- Low-capability devices: in most IoT applications, devices are battery-powered and should autonomously work for several years. This imposes that they have a low processing capacity to minimize power consumption.
- Small infrequent data transmission: the network should be optimized to support the exchange of a small amount of data with minimum overhead.
- Simple design: it is essential to have a light communication protocol and network architecture to meet the limited capabilities of IoT devices.
- Extended Coverage: most IoT devices are located inside buildings and have an undermined coverage due to their limitation in size and power. Therefore, providing good coverage is one of the main challenges for IoT networks.
- Other requirements: some additional high-level requirements, such as identification, authentication, charging, and security, are relevant to the IoT systems. These functions can be addressed either using services provided by an intermediary network or using some proprietary solutions.

Based on the various presented characteristics, IoT applications can be grouped into several categories. It is worth noting that IoT applications can be classified into four segments: *Massive IoT*, *Broadband IoT*, *Critical IoT*, and *Industrial Automation IoT* as described and illustrated in Fig. ??. A brief summary of these classes is provided in the following [?]:

1) Massive IoT: describes applications where a very large number of low-capability end-point devices sporadically transmit a small volume of data. These applications in most cases are data collection applications like sensors that read the temperature, humidity, pressure, etc. Massive IoT devices are expected to work for many years (> 10 years) without the need to replace or change their battery.

2) Broadband IoT: includes applications that require lower latency and higher throughput than the massive IoT. Wearables, drones, augmented reality, and virtual reality devices are typical examples of Broadband IoT.

3) Critical IoT: unlike massive IoT, a fewer number of devices are expected to be connected to the network with relaxed constraints on the requested throughput. However, the main performance requirement needed is to guarantee a real-time reliable data transmission. One use case of this category is autonomous vehicles.

4) Industrial Automation IoT: this segment focuses on the integration of wireless IoT connectivity to the industrial wired infrastructure. IIoT is useful to achieve higher productivity and optimize the production process.

The different segments of IoT networks have different QoS performance demands, and one connectivity technology can not fit them all. Several connectivity solutions have been adopted including Cellular technologies and unlicensed ones. In this article, we provide a comprehensive insight into the work done by 3GPP in Rel-17 to deliver the new cellular technology to support IoT applications in 5G, so-called Reduced Capability devices (NR Red Cap). This includes a presentation of its main characteristics and a discussion of the recent developments observed in 3GPPP 5G IoT standardization.
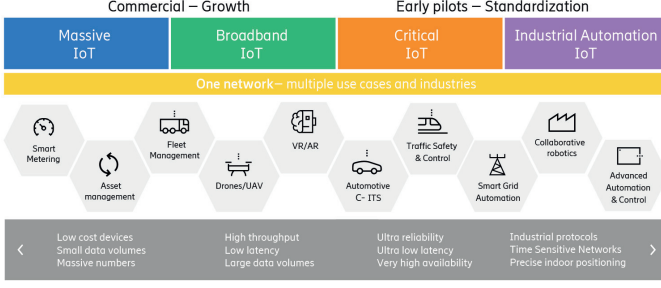
Fig. 3. Cellular IoT segments [**?**]

Fig. 4. IoT standards

The remainder of this article is organized as follows. First, we start in Section **??** by analyzing and comparing the various existing IoT connectivity solutions in both the licensed and unlicensed spectra. After that, Section **??** introduces the main motivations of the work on the new 5G IoT technology (RedCap) and the standardization process in 3GPP. Section **??** and **??** study the key features of RedCap Rel-17, namely, UE complexity reduction and UE power-saving enhancement. The analysis of the potential impact of RedCap devices on the overall system performance is given in Section **??**. We conclude in Section **??** by presenting the recent studies in academia published after 3GPP RedCap Rel-17 and the newly started work item (RedCap phase 2) in 3GPP Rel-18.

II. STATE-OF-THE-ART OF IoT CONNECTIVITY

Providing large-scale stable connectivity is considered the main challenge in order to achieve the ultimate goal of the IoT world. As presented in the abive section, IoT applications extremely vary in terms of Quality of Service (QoS) and network requirements. Such a wide diversity makes it almost impossible to have single communication technology that fits all IoT segments [**?**]. Wired networks are able to provide extremely reliable connectivity with high data rates and low latency. However, the cost of installation and rigid architecture has limited their exploitation in IoT use cases. On the other hand, the re-configurable and dynamic nature of wireless communications and the recent advancement in achievable capacity and latency make them the perfect candidates for IoT connectivity.

Wireless technologies can be classified based on the QoS they offer, i.e., coverage range (short, medium, long), data rate (high or low), latency, cost/complexity, etc. Another common way to group wireless protocols is based on the used frequency spectrum where one can find licensed and unlicensed technologies [**?**]. The choice of an adequate connectivity solution depends on several factors, such as deployment cost, coverage, and power consumption. Most of today's indoor IoT applications rely on short/medium-range unlicensed wireless technologies. These technologies are designed to achieve coverage of tens to hundreds of meters, which is perfect for providing the connectivity of personal and local devices. Medium-range solutions (e.g., WiFi, Zigbee, etc.) are a perfect
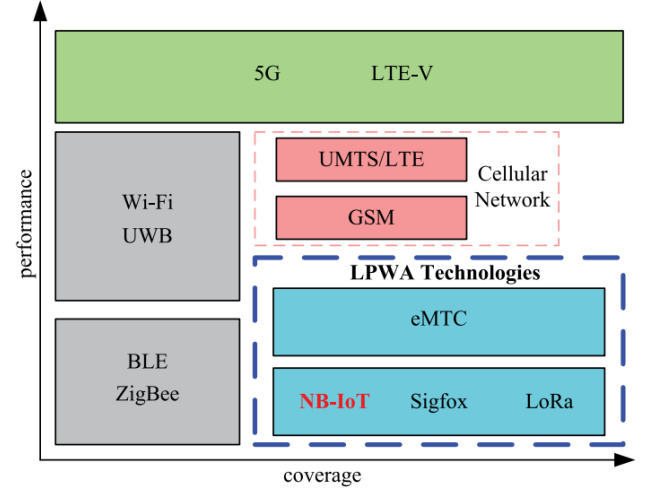
fit for home automation scenarios. Furthermore, Bluetooth and Bluetooth Low-Energy (BLE) are used more for fitness and medical wearables [**?**]. Finally, Radio-Frequency IDentification (RFID), with its limited data transmission and short coverage range, is widely used in most retailers as a solution for stock monitoring and self-checkout [**?**].

In fact, the larger scope of IoT uses cases share common characteristics of having a large number of devices, low data rate, low power consumption, and enhanced wide coverage. Therefore, traditional short-range technologies will fail to meet the coverage requirements. Similarly, the well-known cellular networks, namely, Global System for Mobile Communication (2G/GSM), Universal Mobile Telecommunications System (3G/UMTS), and Long Term Evolution (4G/LTE) are *inappropriate* for supporting most IoT applications due to their *high cost*, *complex design*, and *excessive power consumption*. Thus, different industries have been motivated to invest and develop new revolutionary solutions that are well-suited for serving IoT communications. The main interest was developing a new Low Power Wide Area (LPWA) network class [**?**]. Unlike conventional human-centred networks, which are designed to provide high data rates and mobility, LPWA networks focus on achieving wide coverage for a large number of low-cost battery-powered nodes with relaxed requirements in terms of throughput and latency. A lot of IoT businesses were motivated to use LPWAs as a connectivity solution. This has accelerated the development and the rise of LPWA in the last decade [**?**]. LPWA networks can now be classified into unlicensed spectrum technologies and licensed spectrum. Fig. **??** illustrate the trade-off between coverage and performance provided by different IoT standards. In the following, we further detail existing LPWA technologies since they are solely designed for IoT applications.

A. Standards over Unlicensed Spectrum

LPWA networks over the unlicensed spectrum are usually proprietary patented wireless technologies. These technologies rely on frequency resources over the Industrial, Scientific, and
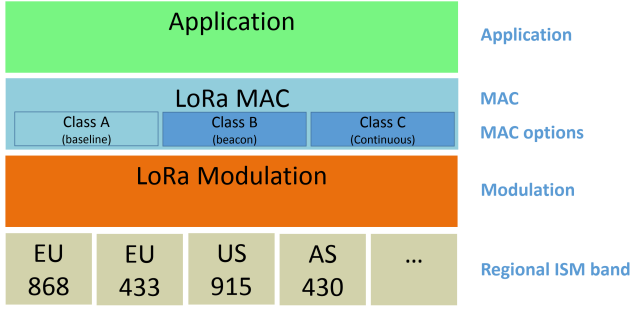
Fig. 5. LoRaWAN Protocol Stack

mMedical (ISM) bands, which will help the providers cut deployment costs. Many LPWA protocols have been developed in the last decade, including, LoRaWAN, Sigfox [?], INGENU [?], TELENSA [?], and Dash7 [?]. However, the two most competitive ones that share a big part of the IoT market today are LoRaWAN and Sigfox [?]. More details on both of these technologies are given below.

1) LoRaWAN: stands for Long Range Wireless Access Network, and it is a communication protocol and system architecture developed and maintained by LoRa Alliance in 2015 [?]. LoRaWAN is an open-source Medium Access Control (MAC) that is based on LoRa RF modulation technology patented for Semtech in 2014 [?] as illustrated in Fig. ?? [?]. LoRa is a spread-spectrum modulation technique that is derived from Chirp Spread Spectrum technology (CSS) [?]. CSS spreads a narrow-band signal over a wider channel bandwidth that makes the received signal more robust against interference (i.e., higher Signal-to-Interference-and-Noise Ratio (SINR)). The mathematical explanation of LoRa modulation is given in [?].

Typical spreading spectrum techniques like Direct Sequence Spread Spectrum (DSSS) tend to multiply the low-rate data signal by a spreading code (so-called chip sequence) of a higher rate. This will spread the bandwidth of the original signal on a wider bandwidth corresponding to the chip rate. The spreading of the signal's bandwidth will make it more resilient to channel attenuation, like multi-path fading. This gain in the link budget is called a processing gain, and it is proportional to the $log10$ ratio of the chip rate and the data signal rate [?]. Spread Spectrum techniques are used in a lot of applications like military communication since it provides a more robust and secure connection.

A similar technique is used in LoRa modulation, where the data signal is spread over a wider bandwidth using a higher rate chirp-like signal. A chirp signal is a tone that changes its instantaneous frequency over time between two frequencies, f0, and f1 [?]. The Spreading Factor (SF) is defined as the amount of spreading code applied to the original signal. The symbol rate R depends on SF following the formula [?]:

$$R = SF \times \frac{BW}{2^{SF}} \tag{1}$$

Having a high SF will reduce the achievable data rate. However, increasing SF will increase signal robustness, and that

means a longer coverage range. LoRa modulation has a total of six spreading factors (SF7 to SF12). The LoRa physical layer operates in different ISM bands depending on region 433/868-MHz in Europe, 915-MHz in the US, etc. Within these bands, multiple fixed channels of 125 KHz, 200 KHz, and 500 KHz are available for UL/DL transmissions [?]. Depending on the SF and channel bandwidth, LoRa data rate varies between 50kbps and 300kbps [?]. Choosing the appropriate SF is a trade-off between data rate and coverage [?].

LoRaWAN networks have three main components: End Devices, Gateways, and Network Servers. The architecture of LoRaWAN networks is organized in a star-of-stars topology where the gateways transparently relay the messages between end devices and the central network server that runs applications at the backend [?]. In LoRaWAN, there are three categories of end devices with different capabilities ranging from class A, B, to C [?]. Class A represents end devices with the minimum functionality to work in LoRaWAN networks. Class A devices provide the lowest power consumption and are used for applications where the communication is triggered by the end device. Downlink transmissions from the server are performed only after a received message from an end device. Class B devices are designed to allow extra downlink transmissions. End devices have a receive window at scheduled times where they should listen for incoming messages. Gateway will communicate to end devices a time-synchronized Beacon where it should be active. Finally, class C devices have nearly continuously received windows, which reduces the latency for server-to-end-device communication [?]. Class C consumes the most power, and it is suitable for end devices that are connected to a power supply network, like traffic lights. All the devices must be compatible with Class A [?]. More details on the different classes of LoRaWAN and different MAC layer structures can be found in ITU-T standard Y.4480 [?].

2) Sigfox: Another well-known LPWAs technology that runs over unlicensed spectrums is Sigfox network. Sigfox is a LPWA networking protocol that was founded by a French start-up that holds the same name in 2009. Later, in 2022, the ownership was transferred to UnaBiz [?]. Sigfox network can be considered as a type of cellular network. However, instead of targeting human-centeric applications, the main goal is to serve massive IoT use cases that need ubiquitous coverage combined with lower power consumption and lower cost [?]. The network is based on one-hop star topology where end devices are connected to gateways/base stations that convey their messages to the sigfox cloud platform backend managed by the Sigfox company itself [?]. Sigfox Network Operators (SNOs) install proprietary base stations and connect them to the backend servers through an IP-based network [?]. Each gateway is able to serve up to a million end-devices with a coverage of 30–50 km in rural areas and 3–10 km in urban areas [?].

The initial version of Sigfox was designed to support only UL communication, but later it became a bidirectional communication technology. Even though downlink transmissions are still only triggered by end devices through an uplink transmission [?]. The messages between base stations and end devices are modulated using Differential Binary Phase

Shift Keying (DBPSK) for UL and Gaussian Frequency Shift Keying (GFSK) for the DL [**?**]. Sigfox signal is carried over an Ultra-Narrow Band (UNB) of 100 Hz in sub-GHZ ISM bands [**?**]. The main advantage of using UNB is to reduce the noise contribution on the occupied bandwidth, and therefore with the same targeted error probability; the signal can reach much longer distances [**?**]. Each message is typically transmitted by default three times on different sub-channels to increase the reliability since no acknowledgment functionality is applied [**?**]. In this regime and given the regulation on the emission over ISM bands, the end-device can send up to 140 messages per day, with a payload size of 12 octets, at a data rate up to 100 bps [**?**]. Using the 3-repetitions scheme, the message takes around 6 seconds to be transmitted [**?**]. It is worth noting that Sigfox does not provide any kind of security functionality, and it is up to the customer to adopt some technique in the application layer [**?**]. Sigfox is considered a complementary solution to other communication systems and can be used together with any other type of network such as: Wi-Fi, Bluetooth, GPRS, NB-IoT, LoraWAN, etc. [**?**]. The fact that Sigfox DL traffic is triggered by UL polling makes it more suitable for data acquisition applications rather than command-and-control ones [**?**]. No detailed specification on the standard is available for the public, but extra details are found in the patent published by sigfox in 2013 [**?**].

B. Standards over Licensed Spectrum

Traditional cellular networks (i.e., 2G/3G/4G) have historically been designed to provide high-quality mobile voice and data services [**?**]. They have proven their ability to deal with a large amount of traffic over a wide coverage area with a better performance in terms of data rate and latency than LoRaWAN or SigFox [**?**]. However, IoT applications cannot simply be served by existing cellular systems. Cellular systems have a limited capacity, and supporting an enormous number of terminals would be challenging, especially with the existing resource allocation approaches. For example, in LTE, for each transmission, a minimum of an entire Resource Block (RB) of 180 kHz is used [**?**]. This would be beneficial to reduce the signaling overhead for some applications, but for IoT traffic, it is extreme and will waste the network resources and subsequently deteriorate the overall performance. Another aspect of the problem is that IoT devices are meant to be low-end and battery-powered modules that should last for years, and supporting complex systems like cellular networks is unfeasible with such limited capabilities. On the other hand, the aforementioned unlicensed spectrum technologies have some drawbacks that make them not suitable for all IoT use cases. Being dependent on the unlicensed spectrum will make the connection less reliable and susceptible to interference [**?**]. Therefore, following the emergence of the IoT market, 3GPP since its Rel-13 has started the work on developing additional IoT connectivity solutions based on existing cellular technologies.

The work in 3GPP to support MTC communication over cellular networks goes back as far as Rel-8. A technical report TR 22.868 [**?**] was published under the title "Study

on Facilitating Machine-to-Machine Communication in 3GPP Systems". The early work from 3GPP Rel-8 to Rel-11 was focused on the overload and congestion problem on data and signaling planes, and the shortage of resources to guarantee access for a large number of terminals [**?**]. Several features were developed under the umbrella of Cellular IoT (CIoT) to meet the demanding IoT requirements, especially in terms of coverage enhancement, connection density, and energy efficiency. By the time of 3GPP Rel-12, most cellular IoT deployments were relying on GSM/GPRS networks, due to their widespread coverage and low-cost UE. However, the non-optimal spectrum usage of GSM/GPRS compared to LTE has encouraged 3GPP to start the work on developing MTC UEs for LTE networks. A Study on the Provision of Low-Cost MTC UEs Based on LTE was published in the Technical Report (TR) 36.888 [**?**]. it was the beginning of "LTE for MTC" technology, the so-called eMTC or LTE-M. Later in 3GPP Rel-13, a study on Extended Coverage GSM (EC-GSM) and Narrow-Band Internet of Things (NB-IoT) technologies was initiated by 3GPP's technical specification group GSM EGPRS RAN (GERAN) [**?**]. This study aimed to investigate the feasibility of supporting ultra-low complexity and low-throughput IoT applications over cellular systems. The findings of this study were later documented in 3GPP TR 45.820 [**?**]. After Rel-13 TSG GERAN was closed and its responsibilities were transferred to TSG RAN who took charge of developing both EC-GSM and NB-IoT along with LTE-M [**?**]. These technologies are the representatives of LPWA standards operating on the licensed spectrum [**?**]. They are widely deployed around the world with more than 250 active mobile IoT network [**?**]. A more detailed comparison between Cellular IoT technologies is presented in the following.

1) LTE-M: is a cellular IoT standard based on the legacy LTE network. LTE-M is considered a simplified version of 4G networks optimized to support scenarios that require medium to low rates, wide coverage, a high number of connections, and extended battery lifetime [**?**]. In Rel-12, 3GPP initiated the work of LTE-M by introducing a new category of MTC devices (so-called, Cat-0) as low-cost MTC devices [**?**]. LTE-M devices use typical 4G multiple access techniques of orthogonal frequency division multiple access (OFDMA) in the downlink and single-carrier FDMA (SC-FDMA) in the uplink. It shares the same spectrum resources as other LTE UEs [**?**]. However, compared to its counterpart Cat-1, the lowest LTE UE category in Rel-11, Cat-0 can provide a 50% of complexity/cost reduction [**?**]. This gain comes from the fact that Cat-0 devices are equipped with a single RF chain (i.e., no MIMO support) and can optionally operate in Half-Duplex mode [**?**]. The cut in the device capabilities will limit UL/DL throughput to 1 Mbps [**?**]. In addition to cost gain, LTE-M devices will enjoy a longer battery lifetime ($\sim$10 years [**?**]) due to the newly introduced UE power-saving techniques, extended Discontinuous Reception (eDRx) and Power Saving Mode (PSM) [**?**]. In Rel-13, 3GPP launched a new category of LTE-M devices, known as enhanced MTC (eMTC) or equivalently (Cat-M/Cat-M1). This category aims to further reduce device capabilities and cater to lower IoT applications. Cat-M devices have a restricted bandwidth of 1.4 MHz instead
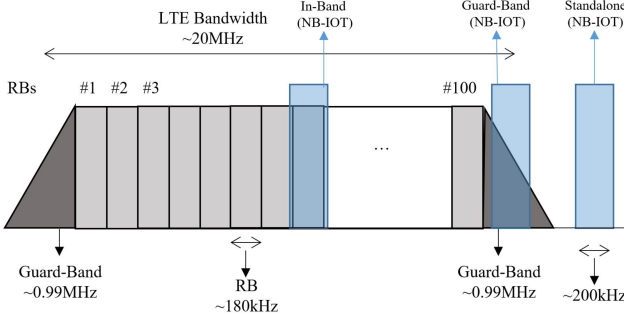
Fig. 6. Deployment modes for NB-IoT

of 20 MHz. However, it provides extended coverage of 15 dB more than the conventional LTE coverage [?] using more repetition in the time domain. Devices of this class are 25 % cheaper than original Cat-0 LTE-M devices and still provide relatively the same throughput of 1 Mbps [?]. It is worth noting that, even with their limited capabilities, Cat-M devices are able to support real-time functionalities like voice over LTE (VoLTE) [?]. The deployment of LTE-M networks is spread all over the world [?]. Many telco operators were able to support LTE-M devices using their legacy LTE network. LTE-M operates over the same spectrum resources as LTE, and only a software update to the network is required to support it without any hardware modification [?].

2) Narrow-Band IoT (NB-IoT): is a narrow-band cellular LPWAN technology that was standardized by 3GPP in Rel-13 in 2016 [?]. It aims to provide wider coverage, lower device cost, longer battery life, and higher connection density compared to LTE-M [?]. NB-IoT is another deviation of LTE and is based on its functionalities and network architecture. NB-IoT was introduced as a new category of MTC devices, so-called Cat-NB1/Cat-NB2 [?]. It can coexist with GSM and LTE in sub-GHz deployments [?], where three modes of operations are adopted [?], [?]:

- Independent (stand-alone) operation: It is based on utilizing and re-framing any available spectrum of GSM carriers in the form of 200 KHz channels.
- Guard-band operation: this mode makes use of the available resource blocks in the unexploited guard band of LTE carrier frequency.
- In-band operation: it utilizes the same frequency resource blocks available within LTE carriers.

An illustration of NB-IoT deployment modes is shown in Fig. **??**. Deploying NB-IoT in sub-GHz bands (i.e., 700MHz, 800MHz, and 900MHz licensed frequencies) makes it an excellent choice for providing extensive coverage. The NB-IoT network is able to support a link budget of 164dB which is 20 dB stronger than the traditional GSM and LTE network [?] while serving up to 50,000 UEs per cell with the possibility to scale up the capacity by allocating more carriers for NB-IoT [?]. It is worth noting that, independent and guard-band modes are able to provide better indoor coverage since the received signal will not suffer from the same interference experienced in the In-band mode [?]. NB-IoT supports bi-directional communication using OFDM for the downlink and

SC-FDMA for the uplink, and unlike LTE-M only half-duplex operations are supported [?]. The small bandwidth of NB-IoT (only 180 kHz) helps cut the cost and energy consumption at the expense of reducing the data rate (> 50 kHz [?]) and increasing the latency. In addition, even though NB-IoT was built on top of the legacy LTE network, it can be considered as a new interface from the protocol stack point of view [?]. Many optimizations have been adopted to both the control plane and user plane in order to support the transmission of small data packets [?] (e.g., multiple transmissions are carried on one RB [?]). A simplified mobility function is provided with no handover support technology [?], and due to the narrow bandwidth of NB-IoT, new physical channels were introduced for synchronization, cell search, as well as new reference signals for channel estimation, tracking, and demodulation [?]. Regarding UE power-saving, the same techniques of PSM and eDRX are re-used in NB-IoT with more extended sleeping periods [?]. For more detailed information on NB-IoT specifications and features we refer to the following resources [?], [?].

3) EC-GSM: is a standard-based technology developed by 3GPP as one of the promising candidates for LPWA CIoT. The work on EC-GSM started in 2015 as part of Rel-13 [?] along with NB-IoT. It is based on eGPRS technology and can be deployed to existing GSM networks through a software upgrade [?]. The traffic of legacy GSM devices and EC-GSM-IoT is multiplexed on the same physical channels [?] and operates over GSM 850–900 and 1800–1900 MHz bands [?]. EC-GSM channels have a bandwidth of 200 kHz, whereas conventional tools of combining FDMA with TDMA are used for multiple access [?]. Two modulation techniques are used, namely Gaussian Minimum Shift Keying (GMSK) and 8PSK, providing a peak data rate of 70 and 240 kbps, respectively [?]. It offers an improvement of 20 dB in the coverage compared to the typical GSM coverage [?] allowing the signal to reach locations with challenging conditions such as deep indoor basements [?]. EC-GSM is also designed to support low-complexity and low-energy devices with a battery life of up to 10 years [?]. Even though this GSM-based technology provides a performance equivalent to NB-IoT, and can be extremely useful for supporting low-end IoT applications, it does not get the same attention as its two counterparts CIoTs [?]. This is because a lot of mobile operators are planning to shut down their 2G/3G deployments in order to free up some resources for the upcoming 5G network [?].

TABLE **??** summarizes the different features of the main LPWA technologies presented in this section. It can be seen from TABLE **??** that all LPWAN technologies (Licensed and Unlicensed) support a wide-area coverage with a long battery lifetime at the expense of a low data rate.

III. REDUCED CAPABILITY DEVICES (REDCAP)

ITU has identified the minimum technical performance requirements for 5G services known as the IMT-2020 standard (see the report M.2410-0 [?]). 5G system in Rel-15 was originally designed to support three main usage scenarios, namely, enhanced Mobile BroadBand (eMBB), Ultra-Reliable

TABLE I
COMPARISON OF DIFFERENT LPWA TECHNOLOGIES

Feature	non-3GPP		3GPP		
	LoRaWAN	Sigfox	NB-IoT	LTE-M	EC-GSM
Devices per cell	10K	10K	50K	18K	20K
Maximum Coupling Loss	157 dB	153 dB	164 dB	154 dB	164 dB
Coverage Range	$< 15km$	$< 50km$	$< 15km$	$< 11km$	$< 15km$
MAC Solution	Spread Spectrum	UNB	OFDM/SC-FDMA	OFDM/SC-FDMA	TDMA/FDMA
Channel Bandwidth	125-500 kHz	100 Hz	180 kHz(LTE carrier) 200 kHz(GSM carrier)	20 MHz (LTE Cat-0) 1.4 MHz (LTE Cat-M1)	200 kHz
Modulation type	CSS	GFSK/DBPSK	BPSK/QPSK	QPSK/QAM	GMSK/8PSK
Peak Rate	50-300 kbps	100 bps	250 kbps	1 Mbps	70-240 kbps
Duplex Mode	Half-duplex	Half-duplex	Half-duplex	Full/Half-Duplex	Half-Duplex
Voice Support	No	No	No	Yes (VoLTE)	No
Mobility Support	Yes	No	Limited	Yes	Limited
Module Cost	9–12 dollars	5–10 dollars	7–12 dollars	10–15 dollars	~ 5 dollars
Battery lifetime	> 10 years				
Localization Support	Yes				

and Low-Latency Communications (URLLC), and massive Machine Type Communications (mMTC). These three classes have significantly heterogeneous requirements in terms of data rate, latency, connection density, energy efficiency, reliability, and mobility. Both URLLC and mMTC are intended to target IoT applications. URLLC which was primarily introduced in Rel-15 addresses time-sensitive applications such as industrial operations and remote surgery. A further enhanced version of URLLC was introduced in 3GPP Rel-16 within the enhanced URLLC (eURLLC) work item [?]. On the other hand, mMTC is meant to serve low-end IoT applications that have relaxed requirements in terms of data rate and latency.

One of the important goals of 5G systems is to enable the automotive industry. The framework of Industry 4.0 aims to create more interconnected smart factories. This industrial transformation can help improve efficiency, reduce operational costs, increase safety, and enhance productivity. It is important to note that developing a massive industrial wireless sensor network is the base to achieve this foreseen transformation.

Alongside smart factories, the future information society contains an important element so-called smart cities. The latter include the optimization of a variety of city activities such as lighting, transport, waste management, parking, video surveillance, etc. The collected data from all these applications can be used to efficiently manage and control city resources and improve provided services to residents. 5G is considered a vital enabler in this envisioned shift toward smart cities.

Another use case that will play an important role in the future 5G community is Wearables. Devices like smartwatches, extended reality (XR) devices, and medical devices will have a huge impact on everybody's life. The spread of 5G network will provide the connectivity needed for these new applications.

The forenamed mid-range IoT use cases have requirements that fall in between those of eMBB, URLLC, and mMTC. They are demanding a performance higher than what is offered by mMTC but lower than URLLC and eMBB as shown in Fig. **??**. One can see that existing 5G usage scenarios either fail to meet the minimum requirements of mid-range IoT applications or they are too powerful which can lead to a waste of network resources.
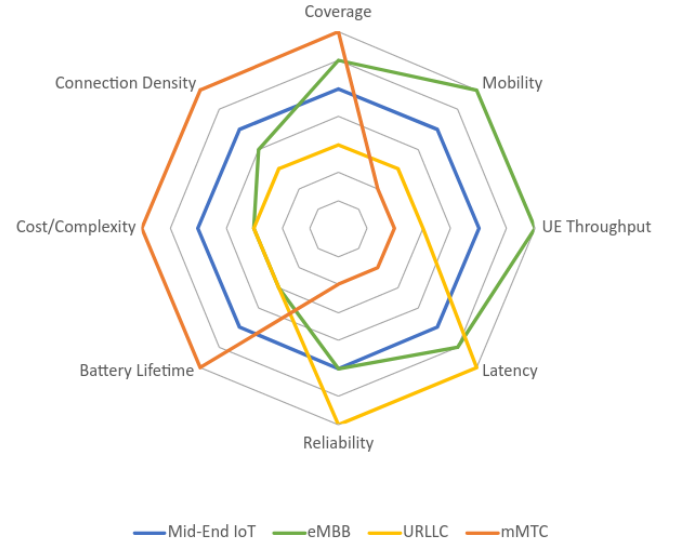

Fig. 7. RedCap Radar chart

In this regard, to further optimize the usage of network resources, 3GPP has initiated in Rel-17 the work to support a new category of devices with a reduced cost/complexity compared to the high-end Rel-16 eMBB and URLLC NR devices. This new category was presented as a lightweight version of NR devices under the name of NR-Light or NR-lite. However, now it is officially called Reduced Capability devices (RedCap). The main objective of RedCap UE category is to serve three reference use cases: Industrial wireless sensors (e.g., various kinds of sensors and actuators), Video Surveillance (e.g., surveillance cameras that are essential for smart city), Wearables (e.g., AR/VR devices, Human-assistive devices, smart watches, glasses, etc.). TABLE **??** presents the main requirements specified by 3GPP in the study item on support of reduced capability NR devices TR 38.875 [?]. From TABLE **??**, it is noted that these three representative use cases have slightly different requirements. However, they can be categorized into one class of mid-range IoT services to avoid market fragmentation [?]. RedCap devices share

common generic constraints of being low cost/complexity and compact in size. They are also meant to support 5G frequency bands in both frequency range 1 (FR1), and frequency range 2 (FR2), and support the two modes of operation: Frequency Division Duplex (FDD) and Time Division Duplex (TDD).

RedCap devices approximately provide a performance similar to low-end LTE UEs such as Cat-1. Nevertheless, it was agreed to develop RedCap as a 5G standalone technology to benefit from the efficiency of the 5G core network and facilitate the transition to a full 5G society. Also, it is important to note that RedCap is not designed to serve Low-Power Wide-Area (LPWA) use cases. 3GPP confirmed in its study on "self-evaluation towards IMT-2020 submission" [?] that together LTE-M and NB-IoT fulfill IMT-2020 requirements for mMTC class. In Rel-17 the study phase on the support of RedCap devices was initiated by multiple RAN work groups. The study aimed to find technical solutions to support the new category while having a minimal impact on existing network scenarios. Three main objectives were announced for this work item: UE complexity Reduction, UE battery lifetime enhancement, and coexistence with other technologies. The following sections explain in detail the proposed features and their impact on the performance. The collected findings during the study phase are documented in the technical report TR 38.875 [?]. The study phase was followed by standardization work to specify the required updates of NR standard, specifically, UE capabilities in TS 38.306 [?] and RRC parameters in TS 38.331 [?] (see section 4.2.21).

IV. UE COMPLEXITY REDUCTION

The design of low-cost/complex devices was the main priority of the RedCap work item. RedCap devices are built on the foundation of reference Rel-15/16 NR devices. This means that RedCap will try to optimize NR configurations in a way that suits the use cases of mid-range IoT applications. Many technical solutions have been suggested to reduce UE cost/complexity. Most of these modifications are related to features in the physical layer (e.g., bandwidth, no. of antennas, etc.). In the following, we present the main complexity reduction features discussed in RedCap Rel-17 [?].

A. Reduced the number of UE Rx antennas

The support of multi-input multi-output communication is essential for most modern-day wireless systems. Increasing the number of receiving antennas (Rx) may dramatically enhance the experienced coverage and throughput at the user terminal. However, these desired advantages come at the cost of unpleasant complex RF chains and overhead processing. The reference NR device supports the following minimum antenna configuration:

- In FR1: 2 Rx for FDD mode and 4 Rx for TDD mode
- In FR2: 2 Rx

While for RedCap UEs only 1 Rx, and optionally 2 Rx, was considered as the minimum requirement in the study. Reducing the number of Rx branches simplifies both UE RF components (including Antenna array, Filters, LNAs, mixer, and local oscillator) and Base-Band operations (ADC/DAC,

FFT/IFFT, Synchronization/cell search block, MIMO processing). The overall device cost/complexity reduction can reach $30-50\%$. More information about the evaluation methodology and the estimation of cost reduction can be found here R1-2009293 [?]. This cut in the device cost/complexity comes at the expense of degrading network performance. Having one Rx branch limits RedCap devices from performing MIMO operations and benefiting from spatial diversity. This may impact both coverage and spectral efficiency. Nevertheless, RedCap's requirements are still met. Another aspect concerning energy consumption, where in theory it should be reduced due to the usage of fewer Rx branches. However, the time needed to receive a certain payload is increased as a result of the degradation in the device data rate.

B. UE bandwidth reduction

Legacy NR devices must support two ranges of carrier frequencies, namely, FR1 and FR2. Within these two possible ranges, UE will transmit/receive signals over a carrier bandwidth. In 5G NR, The maximum allowed carrier bandwidth is 100 MHz in FR1 and 200 MHz (optionally 400 MHz) in FR2 (see [?], [?]). Higher carrier bandwidth can be reached using Carrier Aggregation (CA). Given the less demanding requirement of RedCap use cases in terms of data rate, it is possible to decrease the maximum supported bandwidth. For RedCap UEs, the maximum allowed bandwidth is 20 MHz in FR1 and 50 MHz (optionally 100 MHz) in FR2. It was also suggested to abandon the support of CA for RedCap devices.

This reduction in UE bandwidth results in a substantial decrease in device cost/complexity, estimated to be on the order of $\sim 30\%$ in FR1 and $\sim 20\%$ in FR2. Most of the reported gain is associated with less processing in base-band parts (ADC/DAC, LDPC decoding, etc.). It must be noted that this cut has a minor impact on coverage, spectral efficiency, and latency. The main affected metric is the achievable peak data rate. However, it is still enough to meet the predefined requirements of RedCap use cases.

C. Half-duplex FDD Operation

5G UEs are capable of supporting duplex communication, i.e., communicate in both directions uplink and downlink. Two duplexing approaches exist: Frequency Division Duplex (FDD), and Time Division Duplex (TDD). FDD implies that UL and DL transmission occurs on different carrier frequencies. On the other hand, in TDD, the separation between the two flows takes place in the time domain. While FDD offers better coverage due to frequency separation, TDD provides higher throughput because of the simplified processing. In general, a reference 5G device supports both TDD and FDD modes. Furthermore, FDD is divided into Full-Duplex (FD-FDD) and Half-Duplex (HD-FDD) modes. Full-Duplex FDD means that the transmission on both DL and UL happens at the same time. On the contrary, HD-FDD operation obligates that UE is only allowed to receive or transmit at a given time. This approach of communication gained a lot of attention in LTE for use cases that do not require high throughput but have some constrain on energy consumption such as IoT applications. In

TABLE II
REDCAP REFERENCE USE CASES AND REQUIREMENTS [?]

Requirement	Industrial wireless sensors	Video Surveillance	Wearables
Data Rate	2 Mbps	7.5-25 Mbps	5-50 Mbps in DL (peak 150 Mbps) 2-5 Mbps in UL (peak 50 Mbps)
Latency	100 ms (5-10 ms in safety applications)	500 ms	-
Availability	99.99%	99% − 99.9%	-
Battery lifetime	few years	-	Multiple days (up to 1-2 weeks)
Traffic pattern	Heavy UL	High-end video (Heavy UL)	-
Mobility	Stationary	Stationary	Low-Mobility
Coverage	$MCL = 143dB$		
Cost/Complexity	Low cost/complexity devices compared to eMBB and URLLC		
Size	a design with compact form factor		

TS 36.211 [?], two types of HD-FDD were defined: Type A and Type B. The main difference between the two types is the guard period between receiving and transmitting, where Type B allows a wider guard. In RedCap study [?], it was suggested to use both types of HD-FDD types, where Type A is prioritized.

Compared to a reference NR device, a RedCap device with HD-FDD system can save $\sim 7\%$ for Type A and $\sim 10\%$ for Type B. Most of the cost saving comes from replacing the duplexer with a switch and a low-pass filter and having a simplified RF transceiver. However, operating in half-duplex mode may have a negative impact on the latency. Some additional delay can be experienced especially in case of simultaneous DL and UL traffic. Nevertheless, the latency requirement for RedCap use cases is still fulfilled. Also, it is worth noting that introducing HD-FDD operation will have some impacts on the specification, e.g., determining UL/DL switching time, applicable bands, etc [?].

D. Relaxed maximum number of MIMO layers

One way of using multiple antennas in wireless systems is spatial multiplexing. Spatial multiplexing is based on dividing the transmitted data into several traffic flows, so-called MIMO layers, and sending them simultaneously on different available antennas. This very powerful technique helps enhance the system's spectral efficiency and multiply the achievable data rate. The reference NR device can support, depending on antenna configuration, up to 4 MIMO layers in FR1 and 2 layers in FR2. For RedCap the maximum number of MIMO layers was reduced to one layer FR1/FR2 and optionally 2 layers for FR1 TDD.

Reducing the maximum number of MIMO layers helps in simplifying the signal processing and eventually the overall cost/complexity. The gain in cost is $\sim 17\%$ when reducing from 4 to 1 layer in FR1 TDD and $\sim 11\%$ when reducing to 1 layer for the case of FR2. The main impact of cutting the number of supported MIMO layers is decreasing the spectral efficiency and subsequently the achievable data rate. The peak rate may decrease by $50-75\%$. Despite this reduction, RedCap devices can still be able to fulfill the defined requirements.

E. Relaxed maximum modulation order

Modulation order is an important factor that affects the experienced throughput. Having a higher modulation order can dramatically increase the spectral efficiency (e.g., 256-QAM =8 bits per symbol).

In the RedCap study and in order to cut the device cost, it was suggested to change the maximum mandatory modulation orders as follows: 64-QAM instead of 256-QAM for DL and 16-QAM instead of 64-QAM for UL.

The proposed relaxation of modulation orders decreases UE's cost/complexity by reducing the required processing in both RF and base-band. The estimated reduction in the cost is $\sim 6\%$ for the downlink case and $\sim 2\%$ for the uplink.

Similar to the reduction of maximum MIMO layers, the main impacted performance indicator here is the spectral efficiency. The peak rate will decrease by $25-33\%$, depending on the considered relaxation. Even with this degradation in the achievable throughput, the predefined requirements RedCap device are met.

F. Combinations of UE complexity reduction features

Alongside studying individual techniques listed in subsections A-E, it was suggested to mix some of them together. Different combinations of complexity reduction features and their impact on the device's overall material cost are presented in TR 38.875 (See TABLES 7.8.2-1, 7.8.2-2, and 7.8.2-3). Cost decrease in the order of $50-70\%$ was gained. For example, using the following setting of 20 MHz, 1 layer, 1 Rx, DL 64-QAM, and UL 16-QAM together instead of the reference NR device in FR1 TDD can achieve an estimated cost reduction of $\sim 70.7\%$.

Other features related to higher layers were discussed in the study phase of RedCap. This includes a reduction of the maximum number of Data Radio Bearers (DRBs), a reduction of Sequence Number (SN) in PDCP and RLC, and a reduction of L2 buffer size (see clause 7.1 in [?]). However, they were not part of the final recommendations for the design of RedCap devices because they needed additional changes to the specification and their incorporation in reducing the device cost was marginal.

At the end of this section, we summarize in TABLE ?? the list of recommendations published by 3GPP for designing RedCap devices. It is important to note that these techniques do not require a deep change in the specification of 5G NR nor deteriorate the network performance. However, they can bring a substantial cut in cost, complexity, and size of the user terminal. In the following section, a review of the possible

methods proposed to enhance UE battery lifetime is presented, which was the second objective of RedCap.

V. UE POWER-SAVING ENHANCEMENT

Another important goal of the work on reduced capability devices is to enhance the UE energy consumption and extend its battery lifetime. In 3GPP Rel 16, enhancing UE power-saving was a major topic covered by several technical reports, such as in TR 38.840 [?]. Multiple techniques (including, Wake-Up Signal (WUS), Paging Early Indication (PEI), and dynamic search space set switching) were adopted by UEs in order to consume energy more efficiently. All UE power-saving features available for a reference NR device will be by default accessible for RedCap devices [?]. In fact, due to the low-mobility/stationary nature of the RedCap device along with the bursty intermittent traffic, additional enhancements were proposed to improve UE power-saving further. Three main features were discussed during the RedCap study phase:

1) Suggest less persistent monitoring scheme of Physical Downlink Control Channel (PDCCH). This includes decreasing the number of blind decodes of PDCCH [?];
2) Increase UE sleep time by extending the duration of discontinuous reception (DRX) in both RRC_Idle and RRC_Inactivity states;
3) Relax the mobility function by reducing the frequency of Radio Resource Management (RRM) measurements.

Only the last two were included later in the final recommendations for the design of Rel-17 RedCap devices. The first feature was abandoned due to its marginal power-saving gain $(0 - 12\%)$, compared to the big impacts on the blocking rate of PDCCHs that causes higher latency [?]. More details of the two included in Rel 17, and the potential improvements they can bring will be given in the following subsections.

A. Extended DRX for RRC Inactive/Idle

In 5G systems, gNBs transmit Downlink Control Information (DCI) messages over PDCCHs. These DCI messages have multiple formats based on the information they want to convey to UE (e.g., allocated time/frequency resources on the uplink and downlink, UL power control, etc.). UEs monitor several PDCCHs to decode the control messages from gNB, typically one per slot [?]. However, in some low-latency use cases such as URLLC, UEs can be configured to listen to PDCCHs more frequently.

In IoT applications, most of the time UEs do not have data to send or receive. Thus, continuous channel sensing leads to unnecessary energy waste. To solve this problem, in 5G NR, UEs are allowed to discontinuously monitor the PDCCHs and turn off their transceiver chains in order to reduce their power consumption. Before 5G, Discontinuous Reception (DRX) was considered one of the main power-saving measures that have been adopted by many 3GPP cellular technologies [?], [?]. When UE has a DRX option configured, it will periodically alternate between active (ON) and sleep (OFF) periods in a duty cycle scheme as shown in Fig. ??. Depending on the state of the UE, i.e., RRC_CONNECTED, RRC_idle/inactive, there exist two modes of DRX: Connected mode DRX (CDRX) and
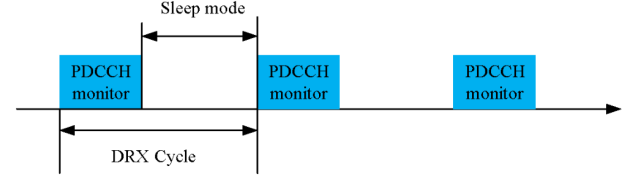

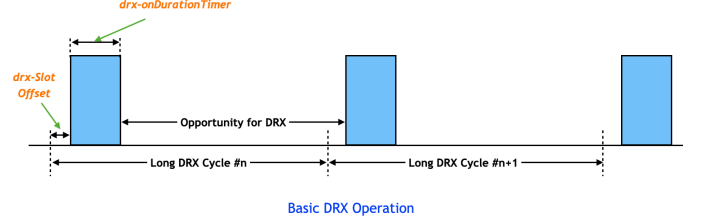Fig. 8. Schematic diagram of DRX

Fig. 9. Basic DRX Operation

Idle/Inactive mode DRX (IDRX). It is worth noting that the power-saving gain achieved by DRX has a negative impact on latency [?]. The type of traffic transmitted and the choice of system parameter will define the tradeoff between the two objectives [?]. In the following, we discuss the operational details and the parameters related to both CDRX and IDRX.

In CDRX mode, the Radio Resource Control (RRC) entity controls DRX operation by configuring the following parameters [?] in the Media Access Control (MAC) layer [?]:

- *drx-LongCycleStartOffset*: contains two parameters *drx-LongCycle* and *drx-StartOffset*. *drx-LongCycle* sets the DRX cycle duration as a value between 10 ms and 10.24 s. And the other parameter, *drx-StartOffset*, defines the starting point of the ON duration as illustrated in Fig. ??. It also takes values in multiples of 1 ms. Note that the basic unit used in both DRX cycle and start offset is 1 ms which guarantees that UEs can start the ON period at the beginning of a radio subframe.
- *drx-onDurationTimer*: a fixed timer at the beginning of every DRX cycle where UE is awake and listening for the control channel. When this timer expires, UE returns to sleep mode if there is no message transmitted in PDCCH as shown in Fig. ??. This timer takes values in multiples of 1/32 ms (up to 1600 ms). The choice of the values of drx-onDurationTimer is consistent with the slot clock that varies depending on sub-carrier spacing.
- *drx-InactivityTimer*: a fixed timer starts after a PDCCH occasion which indicates a new UL or DL transmission for the MAC entity. This timer overlaps the drx-onDurationTimer and cancels its effect. In other words, the UE will continue monitoring the PDCCH and listen to the control information related to the upcoming messages even after the expiration of the drx-onDurationTimer. Once the InactivityTimer times out and there is no message to receive, UE will return to sleep mode as shown in Fig. ??. This timer takes values in multiples of ms (up to 2560 ms).

TABLE III
REDCAP MAIN FEATURES

|	FR1		FR2	
	Ref. NR device	RedCap device	Ref. NR device	RedCap device
Bandwidth	100 MHz	20 MHz	200 MHz	100 MHz
Min number of Rx branches	FDD: 2 Rx TDD: 4 Rx	1 or 2 Rx	2 Rx	1 or 2 Rx
Duplexing Mode	FDD and TDD	Half-duplex FDD	TDD	Half-duplex FDD
Max modulation order	DL: 256QAM UL: 64QAM	DL: 64QAM UL: 16QAM	DL: 64QAM UL: 64QAM	DL: 16QAM UL: 16QAM
Max MIMO Layers	4 or 2	1 or 2	2	1 or 2
Cost Reduction	0%	70%	0%	50%
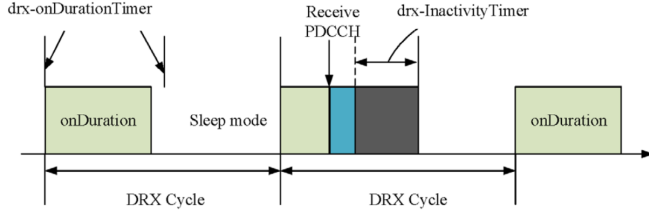
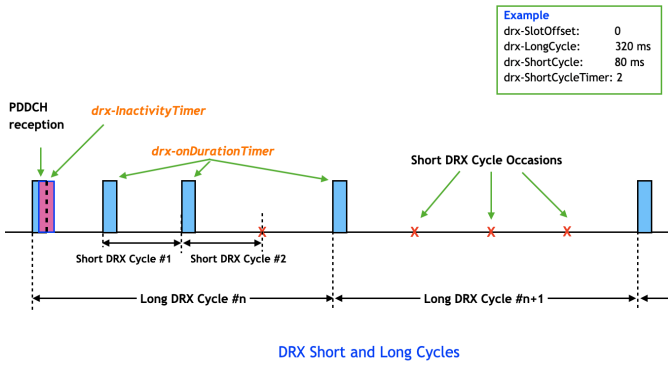Fig. 10. Schematic diagram of drx-InactivityTimer

Fig. 11. Schematic diagram of drx-ShortCycleTimer

- *drx-ShortCycle*: This parameter sets an optional (shorter) DRX cycle duration. The duration of short DRX cycles can have values in the range between 2 and 640 ms.
- *drx-ShortCycleTimer*: This parameter is defined when UE is configured with both Short and Long DRX Cycles. UE switches to operate in Long DRX cycles after a defined number (drx-ShortCycleTimer) of Short Cycles with no activity, as illustrated in Fig. **??**. This parameter takes integer values up to 16.

A CDRX cycle starts with an ON period in which UE checks for control messages, i.e., DCI, on PDCCHs [**?**]. If UE has detected a packet arrival, it starts a drx-InactivityTimer (this timer is restarted as long as UE still has an ongoing connection) [**?**]. Upon the expiry of the drx-InactivityTimer, UE enters a short DRX sleep cycle and later switches to long DRX after several iterations of short cycles [**?**]. After a configured number of long DRX cycles, the network releases UE's radio access context [**?**] and the UE enters to RRC_Inactivity and later RRC_Idle modes where it can benefit from an extended IDRX cycle [**?**]. Fig. **??** summarizes the different steps of DRX operation.

In RRC_Inactive/Idle states, UEs do not have any active sessions. However, while sleeping, they are required to be woken up to listen to paging messages, and short messages, i.e., System Information (SI) update notifications and emergency notifications (Earthquake/Tsunami Warnings (ETWS), Commercial Mobile Alerts (CMAS)) [**?**], [**?**]. UEs in RRC_Idle/Inactivity states are reachable using the paging procedure [**?**]. In RRC_IDLE mode, the UE receives paging messages initiated by the core network. On the other hand, in RRC_INACTIVITY mode, the paging control channel (PCCH) carries RAN-initiated paging messages. The same paging messages are sent over all beams, and it is up to the UE to select the suitable beam by sensing the Synchronization Signal Blocks (SSBs) [**?**]. Since mobile networks are synchronized communication systems, UEs need to be woken up only at specific time instants, so-called Paging Occasion (PO) [**?**]. A PO consists of a set of PDCCH monitoring occasions located in the pre-configured pagingSearchSpace as specified in TS 38.213 [**?**]. Multiple POs form a Paging Frame (PF) that has the same length as one Radio Frame (i.e., 10 ms) [**?**]. In practice, a PO is transmitted over two channels, a PDCCH, and an associated PDSCH [**?**]. UEs scan the configured PDCCH search space for a paging indication whose CRC is scrambled by the Paging-Radio Network Temporary Identifier (P-RNTI) [**?**]. Upon the detection of a paging indication, UE will decode the associated PDSCH looking for the pagingRecordList [**?**]. If its UE_ID is among the paged UEs in pagingRecordList it will initiate the connection with the base station [**?**]. Each UE monitors one PO per DRX cycle and is able to precisely determine the System Frame Number (SFN) that carries its PF/PO using the following formulae [**?**]:

$$(\text{SFN}+\text{PF\_offset}) \, mod \, \text{T} = (\text{T} \, div \, \text{N}) \times (\text{UE\_ID} \, mod \, \text{N}) \quad (2)$$

where PF_offset is the offset from the beginning of the current SFN. T is the UE IDRX cycle or also known as the paging cycle (i.e., PagingCycle). The value of T comes from two different sources. The first is defaultPagingCycle which is broadcasted by the gNB to all UEs through system information messages (SIB). The second source is UE-specific cycle defined by RRC and/or upper layers. The UE will compare the two values and set the IDRX cycle, T, to the minimum between them [**?**]. It is worth noting that PagingCycle is multiples of radio frames and has a maximum value of $\approx 2.56s$ [**?**]. This upper bound was later extended and specified in Rel-17 to $10.24s$ [**?**]. N represents the number of total paging frames

in T. UE_ID is the UE identification, i.e., 5G-S-TMSI stored in USIM. The full description of these parameters is given in Clause 7 in 3GPP TS 38.304 [**?**].

In the RedCap studies, an IDRX cycle beyond $10.24s$ was discussed [**?**]. The two main objectives were to study the possible mechanisms to achieve this IDRX extension and examine the potential power-saving gain. The proposed approach is to reuse the applicable parts of the eDRX mechanism in LTE RAN, including the Hyper-SFN (H-SFN) and the Paging Time Window (PTW), see clause 7.4 in 3GPP TS 36.304 [**?**] for more details. This mechanism was initially designed for NB-IoT and LTE-M to extend the long DRX cycles up to a few hours, which is upper bounded by the upper value of the H-SFN (1048576 radio frame $\approx$2.91 hrs).

The evaluation results obtained by several vendors are summarized in appendix E.1 in 3GPP TR 38.875 [**?**]. A clear power-saving enhancement by extending eDRX beyond $10.24s$ can be observed. The results showed an $80-90\%$ power saving with a DRX cycle of $10485.76s$ compared to the normal IDRX setting of $2.56s$. The study also suggested a clear power-saving gain when the DRX cycle is extended from $10.24s$ up to a few minutes. This reduction in energy consumption may substantially extend the battery lifetime of UEs which is important for some use cases like industrial wireless sensors [**?**]. Based on the advantages mentioned above, RedCap devices are recommended to support an extended DRX cycle of $10485.76s$ in both RRC_IDLE and in RRC_INACTIVE states [**?**]. Such a long eDRX cycle is already supported by the 5G core only for UEs in RRC_Idle state and not RC_Inactivity [**?**]. Therefore, a work item started by 3GPP group SA WG2 in 2022 to study the mechanism and feasibility of this extension [**?**]

B. RRM relaxation for stationary devices

Mobility management plays an important role in cellular networks, which consists of a list of operations, including measurements, cell selection/(re)selection, and handover. Both gNB and UE cooperate to achieve these Radio Resource Management (RRM) operations. gNBs configure UEs to achieve intra-frequency, inter-frequency, and inter-RAT cell selection/re-selection following the requirements specified in TS 38.133 [**?**]. The configuration includes defining parameters like the frequency points, the periodicity, and the number of measured samples. UEs periodically report to the gNB the values of Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), and Signal-to-Interference-and-Noise Rate (SINR) measured from both the serving and neighboring cells. These RRM measurements can guarantee that each UE connects to the best available cell and ensure a good quality of communication. In 5G networks, mobility management is a challenging task, especially, with the introduction of new technologies such as Beamforming [**?**]. In order to connect to the best available Beam, UE needs a very precise evaluation of all possible beams. Even though these measurements are essential to achieve the best performance, they are in some scenarios either unnecessary or not optimized for the mobility profile of the UE. This can lead to excessive power consumption that can drain the UE battery. In

some use cases such as industrial wireless sensors and video surveillance, UEs are more stationary, and the required RRM measurements can be relaxed. In Rel-16, 3GPP has defined a new power-saving technique to mitigate the aforementioned problem of unnecessary measurements. It was a suggested feature in the study on UE power-saving in NR [**?**].

RRM measurements relaxation aimed to specify less frequent UE-specific measurements that adapt the mobility state of the terminal. This feature can be initiated once the UE fulfills some criteria, so-called triggers. RRM measurement relaxation triggers are configured at UEs by the serving cell through SIB messages. In Rel-16, two triggers were primarily specified: *UE with low mobility* and *UE not at cell edge*. According to TS 38.804 [**?**], they are defined as followed:

Criterion for UE with low mobility: This trigger is for UEs in a low-mobility state, which is based on the fact that the level of RSRP does not dramatically change when UE is moving slowly. UEs are meant to monitor the variation of RSRP level and compare it to a prefixed threshold, which can be formulated as follows:

$$\text{Srxlev}_{\text{Ref}} - \text{Srxlev} < \text{S}_{\text{SearchDeltaP}} \tag{3}$$

where Srxlev and $\text{Srxlev}_{\text{Ref}}$ are the current and reference "Cell selection Rx level" in dB. These values are directly derived from the measured RSRP level (see clause 5.2.3.2 in TS 38.804 [**?**]). $\text{S}_{\text{SearchDeltaP}}$ specifies the threshold of the maximum allowed variation of RSRP. The value of $\text{Sexlev}_{\text{Ref}}$ is updated in three conditions: 1) After selecting or re-selecting a new cell; 2) if the UE is moving toward the center of the cell (i.e., Srxlev - $\text{Srxlev}_{\text{Ref}} < 0$), and 3) if the condition in equ. **??** has not been met for a period of $\text{T}_{\text{SearchDeltaP}}$. We see that the level of "low mobility" is not strict and is defined by choosing the value of $\text{T}_{\text{SearchDeltaP}}$ and $\text{S}_{\text{SearchDeltaP}}$ and it is up to the operator to set these parameters [**?**].

Criterion for UE not at cell edge: This trigger is for UEs that are not located at the edge of the serving cell. It allows UEs to compare their current level of RSRP, i.e., Srxlev, with a prefixed threshold $\text{S}_{\text{SearchDeltaP}}$ as follows:

$$\text{Srxlev} > \text{S}_{\text{SearchThresholdP}} \tag{4}$$

The value of $\text{S}_{\text{SearchDeltaP}}$ is transmitted to the UE via SIB messages. If Srxlev is greater than ($\text{S}_{\text{SearchThresholdP}}$), the UE is considered close to the cell center and less likely to attach to another neighbor cell. This permits the UE to relax the mobility management setting by having less frequent RRM measurements and saving more power.

Due to the low mobility property of some RedCap use cases, new triggers for RRM measurement relaxation can be introduced. Several enhancements were discussed in the study phase including:

1) Introduction of stationary UE state,
2) Use beam measurements to determine UE mobility state,
3) Definition of the mobility state in the subscription information, i.e., in USIM.

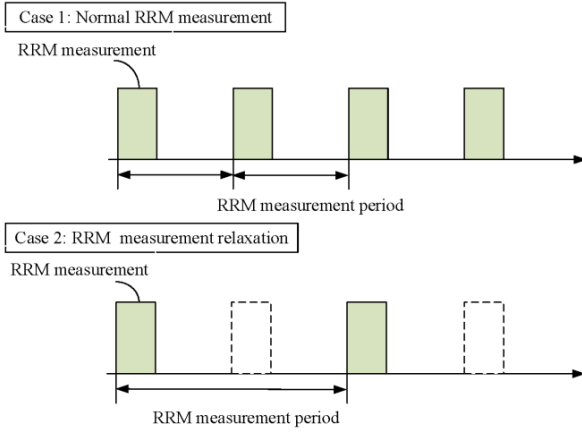However, only two triggering criteria were later specified for RedCap devices [**?**]:

Fig. 12. Example of RRM measurement relaxation

Criterion for a stationary RedCap UE*:* This trigger defines RedCap UEs that are in a stationary state. It follows the same approach used in the low-mobility criterion by sensing the variation of the level of RSRP:

$$\text{Srxlev}_{\text{RefStationary}} - \text{Srxlev} < \text{S}_{\text{SearchDeltaP-Stationary}} \quad (5)$$

UE must sense the level of the received signal, and according to the parameters (duration and threshold) defined by the network, RRM relaxation mode will be triggered.

Criterion for a stationary RedCap UE not at cell edge*:* This trigger combines the two previous criteria and triggers an intensive RRM measurement relaxation once the UE fulfills both "Stationary" and "no at cell edge" conditions.

Once the UE fulfills any of the aforementioned criteria it can start the RRM relaxation operation. Many approaches to perform RRM measurements relaxations have been discussed during the study phase of RedCap and in previous work items of 3GPP [?], [?], [?], including:

1) Increase the interval between RRM measurement;
2) Reduce the number of neighboring cells to be measured;
3) Reduce the number of measured carrier frequencies; and
4) Reduce the number of monitored reference signals.

The methodology that has been adopted so far in NR is based on increasing the RRM measurements cycle time. This is done by skipping some measurement occasions as illustrated in Fig. **??**. UE has to perform RRM measurements every N DRX cycle where N is a scaling factor that is communicated to UE as defined in TS 38.133 [?].

Different companies have contributed to the study item of RedCap [?] by evaluating and analyzing the gain and impact of using RRM measurement relaxation. One source [?] has shown a power-saving gain of $3.6\% - 26.6\%$ when increasing the RRM measurement cycle 4 times with different UE states of idle/inactive and connected modes. This gain comes at the expense of a slight increase in the handover failure rate of 0.26%. Another contribution from Huawei [?] showed an enhancement of 25% in UE power saving by increasing RRM measurement intervals from a few seconds to $1hour$. This study also showed a potential gain of $13.54 - 16.25\%$ in UE

power-saving if the number of measurement of SSBs blocks is reduced by $62.5 - 75\%$. Ericsson on the other hand, claimed that the power saving gained by increasing the measurement intervals beyond one hour (the maximum possible in NR Rel-16) is insignificant [?]. Their conclusion was based on the evaluation of UEs in both RRC_IDLE or RRC_INACTIVE states.

VI. REDCAP IMPACT ON NETWORK PERFORMANCE

An important objective of the RedCap study is to evaluate the potential impact on the overall system performance arising from introducing this technology. The presented features, in Sections **??** and **??**, have great potential in terms of lowering the device cost/complexity and enhancing the UE battery lifetime. However, they might deteriorate other performance aspects, such as latency, and throughput, etc. In addition, the existence of RedCap devices might negatively affect other users in the network, namely, eMMB devices. Therefore, a more detailed and quantitative evaluation is required to estimate this performance degradation. Along with enabling some functionalities to mitigate and limit the potential impact. Two main problems were largely discussed in RedCap study [?]: Coverage Loss and Network Capacity/Spectral-Efficiency. We present in the following subsections the evaluation methodologies agreed on and the final results.

A. Coverage Loss

RedCap devices are expected to experience additional coverage losses [?]. The main impact on coverage is caused by the following considerations:

- Reduced the number of Rx antennas: This causes a loss in spatial diversity which can affect the quality of the received signal and subsequently causes coverage losses.
- UE bandwidth Reduction: RedCap device will operate on a smaller bandwidth than reference NR device. Smaller Bandwidth means less number of frequencies to carry the signal. This loss in frequency diversity will limit the options for frequency hopping which consequently leads to a coverage loss in some propagation scenarios [?].
- Antenna efficiency loss: the smaller size of RedCap device in some use cases such as wearables will affect the antenna gain. The estimated coverage loss is around 3 dB [?].

The link budget evaluation is a widely used method to analyze coverage performance. Several formulas are used to calculate the link budget such as Maximum Path Loss (MPL), Maximum Coupling Loss (MCL), and Maximum Isotropic Loss (MIL). 3GPP decided to use MIL in its work on NR coverage enhancements [?]. MIL can calculate the relative difference in dB between physical channels. This can help determine the bottleneck channel (i.e., the channel with the worst link budget). More details on the link budget template can be found in appendix A.3 of [?].

Any of the physical channels can be the bottleneck that limits the coverage of UE [?]. Therefore, different types of control and data channels along with the random-access channels were included in the link budget evaluation. A list

TABLE IV
PHYSICAL CHANNELS USED IN COVERAGE EVALUATION

Name	Description
PRACH	Physical Random Access Channel is used by UE to send a preamble in order to initiate the connection
PBCH	Physical Broadcast Channel carries system information messages and the primary/secondary synchronization signals PSS/SSS
PDCCH	Physical DL Control Channel conveys DCI to UEs e.g., scheduling information
PDSCH	Physical DL Shared Channel is used to carry UE data on the DL
PUCCH	Physical UL Control Channel is used to send feedback to gNB e.g., acknowledgments and channel-state reports
PUSCH	Physical UL Shared Channel is used to carry UE data on the UL
Msg2	Message 2 or Random Access Response (RAR) is an indication from the gNB about the reception of UE preamble and carries the uplink grants for Msg.3
Msg.3	Message 3 is used by UE to transmit RRC message(e.g, RrcRequest) and includes information about the device identity and capabilities
Msg4	Message 4 or Contention Resolution message confirms that the gNB has correctly identified the UE, and resolved the possible contention and it transfers the UE to RRC_connected state

TABLE V
SIMULATION ASSUMPTIONS USED IN COVERAGE EVALUATION

Scenario Parameters		Rural	Urban	Indoor
Carrier frequency		0.7 GHz	2.6 GHz and 4 GHz	28 GHz
Sub-carrier spacing		15 kHz	30 kHz	60 kHz
Duplexing		FDD	TDD	
Delay Spread		300 ns		30 ns
gNB TX chains		2	4	2
gNB RX chains		2	4	2
UE TX chains		1		
UE RX chains	Ref. NR	2	4	2
	RedCap	1 or 2		
UE bandwidth	Ref. NR	20 MHz	100 MHz	
	RedCap	20 MHz		50 MHz or 100 MHz
UE speed profile		3 km/h		
UE antenna correlation		low		

of the main channels and the associated messages are given in TABLE **??** [**?**]

Three propagation scenarios have been studied: Rural and Urban in FR1 and Indoor in FR2 [**?**]. A list of the main parameters of each of these scenarios is summarized in TABLE **??**, which also includes the common assumptions and the specifications used to differentiate between RedCap and reference NR devices.

For both RedCap and reference UEs, the target performance requirements of physical channels/signals are defined, i.e., min throughput, Block Error Rate (BLER), no. of HARQ re-transmission, etc. After that, the coverage evaluation can be carried out by following the three main steps [**?**]:

1) Perform a Link-level Simulation (LLS): This step aims to define the required SNIR level to meet the performance target for each of the examined physical channels. Note that the LLS simulations is done using the parameters defined in TABLE **??**.

2) Calculate the link budgets: MILs of different channels are calculated using the results of the LLS (i.e., SINR levels).

3) Determine the required coverage recovery: Once the MIL level of all channels is calculated, coverage recovery for RedCap device can be estimated. The amount coverage recovery is identified by the bottleneck channel for the reference NR device within the same deployment scenario. The bottleneck channel is the physical channel that has the lowest MIL(see clause 6.3 in [**?**]).

Different companies have contributed to the RedCap study by performing the link budget analysis. Simulation results of different sourcing companies can be found in appendix C of 3GPP TR 38.875 [**?**] and in [**?**]. As a conclusion of the coverage evaluation of RedCap devices, the following outcomes are made:

- In FR1, for both Rural and Urban scenarios, due to the limited size of the RedCap device, a reduced antenna efficiency of 3 dB is expected. As a result, both PUSCH and Msg3 have a MIL level that is less than the bottleneck channel of the reference NR UE and a coverage recovery is needed. However, other UL physical channels do not require coverage compensation since they have a MIL level higher than the bottleneck channel. For the DL physical channels, the coverage is dependent on the frequency bands (i.e., 0.7 GHz, 2.4 GHz, or 4 GHz) and the assumptions of the DL Power Spectral Density (PSD). The most affected signals are Msg2 and Msg4 with a potential degradation of 6 dB, especially for RedCap UEs with 1 Rx. This coverage degradation can be mitigated by using the Transport Block Size (TBS) scaling technique.

- In FR2 with mmWaves, the antennas have a smaller form. Thus, the size limitation of RedCap devices will not have a negative effect on antenna efficiency. Therefore, the MIL levels of UL channels are equivalent to the ones of the reference NR UE. Thus, no coverage recovery is needed. However, for the DL channels, a coverage recovery in the range [2-3 dB] is needed for PDSCH, for RedCap UE with 50 MHz BW and 1Rx.

- No new coverage recovery solutions are needed for Red-Cap UEs. RedCap devices can use the existing solutions (e.g., PUSCH repetition) as specified in the Rel-17 [**?**] to compensate the additional coverage losses on the UL. Other techniques from Rel-15/16, such as TBS scaling, are also available for RedCap UEs.

B. Capacity Analysis

Similar to the coverage performance, the introduced complexity reduction features of RedCap devices are expected to affect the overall system capacity negatively. Simplifications, such as reducing the number of antennas and the modulation order, have a direct impact on spectral efficiency and consequently may reduce the system's capacity. Therefore, further examinations need to be done before approving these features. In particular, a System-Level Simulation (SLS) must be conducted to evaluate the impact on the system's capacity. The simulation assumptions defined in 3GPP TR 38.802 [**?**] were re-used as a baseline. A list of the main simulation parameters for both RedCap and Reference NR devices is given in TABLE **??**. The rest of the simulation parameters

TABLE VI
ASSUMPTIONS FOR SLSs

Parameters	RedCap UE	Reference UE
Scenarios	• Dense Urban: 2.6 GHz and 4 GHz (TDD) • Indoor: 28 GHz (TDD)	
Traffic Model	• Full buffer (optional) • FTP model 3 • IM traffic model	• Full buffer (optional) • FTP model 3
Traffic load	• Full buffer (optional): 10 users per cell • Other traffic models: low ($< 30\%$) and medium ($30 - 50\%$) resource utilization	
Scheduled BW	• 20 MHz in FR1 • 50 or100 MHz in FR2	• 100 MHz for FR1/FR2
Max modulation order	• 64QAM for DL • 16QAM for UL	• 254QAM for DL • 64QAM for UL
UE Rx / MIMO layers	• 1Rx or 2Rx	• 4Rx in FR1 • 2Rx in FR2
RedCap UE ratios	• 0%, 25%, 50%, or 100% of the UEs in the cell are RedCap devices	

can be found in Tables 6.4-1 and D-1 in [**?**]. Several traffic models, including full buffer traffic and bursty traffic such as File Transfer Protocol Model 3 (FTP3) and Instant Messaging (IM), are used in simulations. Each of them is characterized by different packet sizes and inter-arrival time. The overall Resource Utilization (RU) in the cell is related to the number of UEs and the traffic model used. The complexity reduction features of RedCap devices are consistent with the one presented in Section **??**. The metric used to evaluate the system performance is the average Spectral Efficiency (SE) and it is defined as follows:

$$SE = \frac{\text{cell average throughput (Mbps)}}{\text{cell bandwidth (MHz)} \times RU} \quad (6)$$

*check the storyline later

The results of the SLS from different companies are gathered in Tables D-2 to D-25 in appendix D of [**?**]. The evaluations covered different combinations of the parameters presented in TABLE **??**. The outcome of the capacity and spectral efficiency evaluation can be summarized as follows:

• In the case where both types of UEs have bursty traffic, FTP3 is assumed for eMBB users and IM traffic for RedCap devices. The results show that introducing RedCap devices will have a marginal impact on the throughput of other eMBB UEs and the overall system capacity. It is worth noting that given the considered traffic models, the RedCap devices generate 0.1 MB payload every 2 s (400 kbps) while eMBB devices have 0.5 MB payload every 200 ms (20 Mbps). It means that the generated load of RedCap devices is very limited even when the number of RedCap devices connected to the cell is much

larger than eMMB devices. The IM traffic model is only valid for the use cases of video surveillance and industrial wireless sensors where the exchanged traffic is dominated by the U.

• In wearable use cases, both types of devices have the FTP3 traffic model and the amount of data in the DL is much higher than in the UL. The results between companies are quite different. On one hand, Nokia stated that the throughput of eMBB device is not affected by the introduction of RedCap device. On the other hand, Huawei/HiSilicon claimed that the spectral efficiency degradation could reach 30% and 50% when RedCap devices have respectively 1Rx and 2Rx.

• Similar conflicted results among companies can be observed when both eMBB UEs and RedCap devices are modeled with full buffer traffic. Nokia claims a minor degradation in the spectral efficiency while Huawei/HiSilicon showed that the impact can reach 54-70% depending on the number of Rx chains used for RedCap devices.

VII. BEYOND REDCAP REL-17

In March 2021, 3GPP published a full technical report TR 38.875 [**?**]. It also includes a summary of all the recommended features to reduce UE cost/complexity and enhance UE power-saving. After this, several research efforts were also made to investigate the proposed solutions further.

Given the coverage loss reported on the UL, especially with PUSCH channel, [**?**] highlighted that such degradation might have an impact on certain categories of RedCap devices. Services like healthcare monitoring and require reliable real-time UL transmission. Several possible 5G techniques to enhance UL performance were discussed, including, Dual Connectivity (DC), Carrier Aggregation (CA), and Supplementary Up Link (SUL). DC and CA were excluded since they cannot meet the RedCap complexity reduction requirements of operating in a single connectivity mode and supporting one band at a time. On the other hand, in SUL, UE operates at a single frequency band and switches to another prefixed one when the quality of the received signal drops (e.g., from 700 MHz band to 3.5 GHz band). The evaluation of this technique was done through link-level simulations. The results showed a gain in the link budget of 8.33 dB which was reflected in improving the probability of detection and BLER. This gain comes at the expense of some additional delay due to the switching cost.

Another work conducted by Nokia [**?**] investigated some modifications to the proposed RRM measurement relaxation feature. It aims to quantify and analyze the power-saving gain while using different thresholds of the "Cell Edge" criterion. The study was done for RedCap UEs with two mobility profiles: Stationary and Pedestrian (3 km/h). For stationary UE, the results show that reducing the threshold of "Cell edge" (i.e., increase cell border) can significantly improve UE power saving (up to 96% better than the baseline "No RRM relaxation" case) and without any additional impact in terms of handover failures ratio (HOFR). The pedestrian UEs on the other hand, experienced a substantial increase in the HOFR

while reducing the "Cell-Edge" which affected the service continuity.

Another work on UE power-saving for RedCap devices was conducted in [**?**]. The paper reviewed the existing power-saving techniques in 5G NR, including discontinuous reception (DRX), Wake-Up Signal (WUS), RRM relaxation, and Paging enhancement. The study proposed an approach to enhance the paging for RedCap UEs that are geographically located close to each other , which is similar to Paging Sub-grouping specified in Rel-16 (see clause 7.3 in [**?**]). However, no power-saving gain analysis based on simulation results is given in this work.

A. RedCap in Rel-18

Based on the foundations put in Rel-17 RedCap WI, 3GPP continued in Rel-18 the work on enhanced support of reduced capability NR devices in the work item (NR_redcap_enh) [**?**]. The main objective of the work on Rel-18 RedCap is to expand the 5G market with use cases that have relatively lower cost/complexity and lower power consumption [**?**]. Following the same logic in Rel-17, a further reduction of the Bandwidth to 5 MHz is suggested to cut the device cost/complexity [**?**]. Another simplification regarding the restriction of the UE peak data rate to 10 Mbps was discussed. The proposed features and the analysis of their impact are documented in TR 38.865 [**?**]. This extension of RedCap Rel-18 aims to enlarge the range of RedCap applications with use cases that need relatively lower cost/complexity, and more power-saving devices. However, it is important to note that this expansion is not meant to overlap with the existing LPWA applications [**?**].

Another work item related to RedCap later started in December 2022 entitled "5GS support of NR RedCap UE with long eDRX for RRC_INACTIVE State" [**?**]. In RRC_idle state, the 5G system already supports paging cycles beyond 10.24s since both NB-IoT and LTE-M are connected to it. However, in RRC_INACTIVE state, handling of UEs data/signals is not defined when the UE is unreachable for a period of more than 10.24s. Therefore, this work aims to specify the required mechanism to support long eDRX in RRC_INACTIVE state.

VIII. Conclusion

This paper gives a comprehensive overview of the recent work of 3GPP on supporting a new cellular IoT device known as RedCap. It starts by shedding light on the different classifications and characteristics of IoT networks. Traditional IoT connectivity protocols are discussed from both the unlicensed and cellular spectrum. For the unlicensed band, both LoRaWAN and Sigfox are widely used for data acquisition application that requires long coverage with very limited payload. Cellular IoT technologies, namely LTE-M, EC-GSM, and NB-IoT, provide more reliable connections with higher data rates, benefiting from working in the licensed spectrum. The emergence of the mid-range IoT market has encouraged 3GPP to expand the 5G ecosystem and introduce a lighter version of 5G UEs under the name RedCap devices. It aims to cater to the new mid-range IoT applications, whose requirements fall in between other 5G classes (i.e., eMBB, URLLC, mMTC).

The design of RedCap devices should further optimize the radio interface to have devices with lower cost/complexity, smaller physical size, and longer battery lifetime compared to reference 5G UEs. Different features have been proposed to reduce UE complexity including, limiting UE Bandwidth, reducing number of RX, etc. The total cost/complexity cut can reach a $50 - 70\%$ lower than the reference 5G UEs. To meet the other objective of reducing UE power consumption, both eDRX and RRM measurements relaxation were adopted. Simulation results estimated a power saving gain of $80 - 90\%$ when increasing the IDRX cycle to 10485.76 s (~ 2.91hrs). Reducing the frequency of RRM measurements four times can decrease the UE power consumption by 25%. Furthermore, the article also highlights the potential coverage and capacity degradation linked to the introduced features of RedCap devices. Simulation results showed some additional coverage losses of 3 dB on the UL channels. This limitation can be compensated using coverage recovery techniques such as PUSCH repetition, TBS scaling, etc. The work to further enhance and simplify RedCap devices will continue in Rel-18 where 3GPP has launched a new work item on RedCap phase 2 [**?**] [**?**].