Abstract—The Internet of Things (IoT) ecosystem emerges rapidly and paves the way towards collecting huge amounts of data that can feed innovative information-based systems. The connectivity infrastructure that will support the IoT ecosystem encompasses deployments that define the so-called Low-Power Wide Area Networks (LPWANs). LPWANs are an attractive solution for IoT service provisioning, since they incorporate energy-efficient and cost-effective technologies. In this context, a comprehensive study is provided to assist in the currently increasing research and development interest on LPWANs. First, recent cloud-based and open-source approaches for data management are discussed, while the major key wireless technologies for network access are presented. Second, the potential of Long Range (LoRa) modulation and Long Range Wide Area Network (LoRaWAN) protocol, as key LPWAN wireless technologies in unlicensed spectrum bands, is further analyzed. The fundamental principles of these technologies are presented, and a thorough study on the relevant research activities is conducted.

Index Terms—Internet of Things (IoT), long range (LoRa), long range wide area network (LoRaWAN), low-power wide area network (LPWAN), wireless sensor and actuator network (WSAN), wireless sensor network (WSN).

I. INTRODUCTION

ONE OF the key targets of beyond 2020 networks is the interconnection of numerous devices to support innovative applications with heterogeneous requirements. This target has been described in the International Mobile Telecommunications 2020 (IMT-2020)1 report [1] as one of the three most demanding scenarios for the 5G era, under the term massive Machine Type Communications (mMTC). In this context, the Internet of Things (IoT) technology [2] serves as a key enabler for the realization of the mMTC scenario.

IoT has been defined as 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 [3]. Based on this definition, various terms have been coined to highlight i) the application perspective (e.g., Broadband IoT and Critical IoT), and ii) the network perspective (e.g., Massive IoT [4] and Network of Things (NoT) [5]). In all cases, the message is one: in the coming years, the massiveness of IoT devices is expected to be a major challenge. Recent estimations foresee that, by the end of 2030, about 500 billion IoT devices will be deployed [6].

In this context, two main approaches are developed. On the one hand, mobile networks (having high coverage from multiple Base Stations (BSs) already in the field) and local area networks (covering also indoor scenarios) expand their capabilities towards including IoT communications in the pallet of their services. An example of such capabilities is included in the recent work around the 5G New Radio for Reduced Capability (NR RedCap) devices [7]. On the other hand, the deployment of networks dedicated to support IoT services is evolving notably, having as a main representative the Long-Power Wide Area Networks (LPWANs) [8].

LPWANs target the deployment of Gateways (GWs) operating as access points, to support End Devices (EDs) operating as sensors (for gathering data of interest) and/or as actuators (for performing an action). As such, LPWANs can be considered as a type of Wireless Sensor and Actuator Networks (WSANs) [9]. From the application perspective, LPWANs support applications with low-rate requirements, which implies low energy demands for the EDs (they can remain operational in the field, with no battery charging or replacement, for long periods -a few years-) [10]. Also, from the deployment perspective, wide areas are covered with a small amount of GWs (due to the long-range type of wireless links). These characteristics (low energy demands and long-range support) make LPWANs a cost-effective solution for IoT.

The scope of this paper is to present LPWANs as major candidates for the support of the IoT ecosystem, by decomposing various related technologies and revealing research challenges and directions. To this end, extra focus is given to the Long Range (LoRa)/LoRa Wide Area Network (LoRaWAN) systems, which appear to be a prominent solution for wireless access over unlicensed spectrum. Taking into account the existing surveys in the area ([11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]), we strive for an up-to-date comprehensive review on LPWANs, by providing valuable insights on the underlying technologies and state-of-the-art directions for research and development.
A Comprehensive Study on LPWANs With a Focus on the Potential of LoRa/LoRaWAN Systems

Fig. 1 reflects the adopted approach; with the solid-colored boxes, a chain of core sections is depicted, defining a main reading path for a reader familiar with the topic. The auxiliary tabular boxes, which complement the main reading path, represent self-contained sections, targeting at building the necessary background for a non-expert reader.

Specifically, the contributions of this paper are the following:

• Understanding the importance of adopting a holistic end-to-end approach for a network study, we shed light on both the access-side and the server-side of LPWANs. For the former, we discuss the main considerations in the communication among the EDs and the GWs, whereas for the latter, we provide an indicative list of implementation options in terms of cloud support. We also dedicate a separate section on data management; a key concept of IoT networking and consequently of LPWAN networking.

• Hoping to serve as a reference point for LPWAN researchers, we have gathered and examined, under the same set of criteria, a complete list of LPWAN wireless technologies that appear in the literature. Also, to the best of our knowledge, this survey is one of the few to present in such detail the specifications of the LoRa and LoRaWAN technologies, and the first one to focus on the latest LoRaWAN release, i.e., v1.0.4.

• Given the rapid growth of the research in the area of LoRa/LoRaWAN, we provide a comprehensive state-of-the-art analysis, where the related research work is classified into three categories: analytical approaches, simulation-based studies, and real deployments. To assist the future researcher in the field, we also provide an extensive survey on simulators that can be used for LoRa/LoRaWAN evaluations.

The rest of this paper is organized as follows. In Section II, we describe the fundamental characteristics of LPWANs. In Section III, we present the main issues in the access part of LPWAN systems, whereas in Section IV we discuss server-side considerations in terms of cloud implementations. In Section V, we examine the topic of data management in LPWANs. We describe numerous LPWAN solutions in Section VI, and compare them in Section VII. Afterwards, we turn our attention to the LoRa/LoRaWAN networking stack. We provide a high-level view both in architectural terms and in protocol stack terms in Section VIII. We delve into LoRaWAN in Section IX and LoRa in Section X. Having achieved familiarity with this LPWAN solution, we identify its important research challenges, thanks to the published state-of-the-art work in Section XI, and the key research methodology of simulations in Section XII. Finally, Section XIII contains some final remarks.

II. LPWAN BACKGROUND

LPWANs define a wireless network family, which can be considered as part of the third generation of IoT networking, where the focus is the provisioning of cloud services to connected things [28]. Fig. 2 compares the key wireless technology families in terms of data rate, range, and power consumption. Clearly, the LPWAN technologies are the winners in terms of range and power consumption.

The following are well-established LPWAN technologies: Extended Coverage-Global System for Mobile communications-IoT (EC-GSM-IoT), Long Term Evolution-Machine type communication (LTE-M), Narrowband-IoT (NB-IoT), MY IoT (MIOTY), Random Phase Multiple Access (RPMA), Weightless, Wize, Sigfox, and LoRa. The
first three technologies (i.e., EC-GSM-IoT, LTE-M, and NB-IoT) have been defined by the 3rd Generation Partnership Project (3GPP) and are referred to as Cellular IoT (CIoT) technologies, whereas the rest have not been standardized by 3GPP. This criterion also splits the technologies into those that use the licensed part of the spectrum (3GPP technologies) and those that operate in the license-free spectrum (non-3GPP technologies).

From an architectural perspective, LPWANs are composed of five types of nodes: the EDs and the GWs, which compose the access-side, and the Network Gateway, the LPWAN Authentication, Authorization, and Accounting (AAA) Server, and the Application Server, which compose the server-side. EDs are deployed in massive numbers and tend to broadcast their messages in the hope of being received by a GW that covers the area. GWs forward these messages to the server-side, where numerous functions at the network and application layers take place. At the server-side, the Network Gateway receives this transmitted information and either performs internal processing or re-distributes the information to the rest of the interested parties, i.e., the LPWAN AAA Server and the Application Server. Existence of the LPWAN AAA Server and the Application Server is optional, because the Network Gateway may be able to handle all the necessary server-side functionality. Nevertheless, common practices dictate the use of dedicated servers for the different functionalities and types of messages, e.g., an LPWAN AAA Server may authorize specific EDs to participate in the network, whereas an Application Server may be the bridging point between internal server-side application data and external application end user. A growing trend in server-side implementations is the utilization of a cloud infrastructure. Although the nomenclature and the full node deployment might be different from one technology to another, an LPWAN system architecture can be usually described by a structure similar to the one depicted in Fig. 3.

III. ACCESS-SIDE CONSIDERATIONS

Focusing on their access-side, LPWANs need to meet certain demands, such as energy efficiency, long range communication, low-cost operation, and support of Uplink (UL) dominated connections.

A. Energy Efficiency

For all LPWAN technologies the main energy draining factor is the transmission process. Hence, an appropriate radio interface and a careful transmission planning are essential for the system’s longevity. Furthermore, lightweight protocols are needed to achieve not only the minimization of the number of transmissions, but also the minimization of each transmission’s length. For specific scenarios, EDs are powered by solar panels, moving towards an energy harvesting approach, off-boarding an important factor of the network management functionality.

B. Long Range

As expected by the networking paradigm’s name, long range support is a prerequisite for an LPWAN technology. Practically, LPWANs are systems that establish links in the order of 10^3 m and 10^4 m. Furthermore, experimental setups have seen communication in the order of 10^5 m. Such distances are achieved thanks to robust Physical layer (PHY) protocols. More specifically, the 10^4 m is the main LPWAN operational range and refers to deployments in urban/suburban environments; the 10^5 m is an achievable range in areas that are mostly obstacle-free, Line of Sight (LOS) links, and rural environments [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]. Also, the 10^5 m range has been shown feasible in experimental setups. Indeed, a number of such experiments have been conducted worldwide and a LoRaWAN record based on weather balloons transmissions was achieved at 766,000 meters.\(^2\)

C. Low Cost

In the access-side, GWs are deployed in low numbers (due to the long range) and operate as packet forwarders, pushing the main functionality back to the server-side; this minimizes the GW cost. On the other hand, EDs are deployed in massive numbers and, thus, the minimization of their development cost is performed by constraining the functionality and limiting the hardware/software needs. Indeed, EDs can be built from scratch, using the appropriate radio module and a common development board, such as Arduino, Raspberry Pi, or BeagleBoard. Cost reduction can also be achieved in the server-side, as described in the next section.

D. UL Dominance

The LPWAN access part supports the Wireless Sensor Networks (WSNs) [29], and, as such, it is characterized as UL dominant; that is, most of the transmissions are performed from EDs to GWs. Typically, EDs are deployed on the field and are expected to provide the server-side with information obtained by their sensors, based on a periodical or event-driven scheme. It should be noted that, contrary to the typical mesh or ad-hoc network topology used in WSNs, LPWANs follow a star system development [30].

IV. CLOUD SUPPORT FOR THE SERVER-SIDE

In the server/control-side of LPWANs, various software components are expected (e.g., Network Gateway, Application Server, and the Application Server). Existence of the LPWAN AAA Server is mandatory, as it handles the authentication of EDs before they can transmit to the network. The AAA Server is the entry point to the network and is responsible for authorizing specific EDs to participate in the network. From an architectural perspective, the AAA Server is the main functionality back to the server-side; this minimizes the GW cost. On the other hand, EDs are deployed in massive numbers and, thus, the minimization of their development cost is performed by constraining the functionality and limiting the hardware/software needs. Indeed, EDs can be built from scratch, using the appropriate radio module and a common development board, such as Arduino, Raspberry Pi, or BeagleBoard. Cost reduction can also be achieved in the server-side, as described in the next section.

2https://www.thethingsnetwork.org/article/lorawan-distance-world-record, Accessed: Jun. 6, 2022
Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Technology (NIST), the three main representatives are computing. According to the National Institute of Standards established itself as a growing field in the area of cloud computing. Recently, the Everything-as-a-Service (XaaS) approach has established itself as a growing field in the area of cloud computing. According to the National Institute of Standards and Technology (NIST), the three main representatives are Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS). NIST definitions for all three of them are given in [39]; however, we provide our own definitions here:

- **SaaS**: The end user’s ability to use existing applications developed by the cloud service provider, via an Application Programming Interface (API) or a client request. The end user might be granted authorization to configure a few application parameters, but neither user-developed applications are allowed nor access to the underlying platform or hardware infrastructure is permitted.

- **PaaS**: The end user’s ability to deploy its own software, assuming that the software technology can be supported by the cloud environment (e.g., Java applications will require a certain environment to run in the cloud infrastructure). The end user might be granted authorization to configure a few environment parameters, but has no control over the underlying platform or the hardware infrastructure.

- **IaaS**: The end user’s ability to deploy entire platforms in the cloud environment, assuming the platform’s components can be supported by the cloud infrastructure (e.g., a system component requiring a minimum amount of memory to run). The end user is given the option to configure the platform as needed, e.g., select Operating System (OS) and storage size, but has no control over the underlying hardware infrastructure.

More on the concept of cloud computing can be found in [40], in which the Database-as-a-Service (DBaaS) model -an emerging service model in the area of cloud support to IoT systems- is presented. An informative resource is also [41], in which the topic of IoT forensics (Forensics-as-a-Service (FaaS)) is discussed, giving valuable insights in IoT networking security aspects based on cloud computing.

RPMA is characterized as a PaaS LPWAN solution, since Ingenu (the company behind the technology) offers the required tools to the end users for developing their own applications on top of the RPMA platform. Apart from RPMA, a number of companies offer XaaS solutions in the LPWAN domain, by allowing access to backend servers, storage space, or a deployed infrastructure.

Numerous key industry entities operating in this area could be mentioned [42]. Amazon’s Amazon Web Services (AWS) is a well-established platform offering cloud solutions, and, in the context of IoT, IoT Core has great interoperability with LoRaWAN systems. Microsoft provides similar capabilities with its Azure IoT platform. Google also invests heavily in such a platform with its Google Cloud Platform (GCP),3 Alibaba shows substantial momentum in this field with its Alibaba Cloud and the corresponding IoT platform. IBM serves IoT Cloud needs with its IBM Watson IoT Platform. Tencent runs the namesake cloud platform and offers solutions via its IoT Hub. Finally, Oracle is another option to be considered in this area, with its Oracle Cloud Infrastructure (OCI) and the IoT Intelligent Applications Cloud.

Moreover, IoT-specialized cloud service providers have emerged, showcasing expertise on IoT functions and data management. The ecosystem of such IoT-specialized cloud providers is already large and keeps expanding at a rapid pace.

V. DATA MANAGEMENT IN LPWANs

LPWANs are part of the third IoT generation [28], which focuses on cloud storage and processing, by following data-centric management [29] and data-driven networking [43]. The entirety of data-driven networking, in the context of LPWANs, faces the key data challenges of the so-called five Vs of big data: data generated in high Volumes, with extreme Velocity, having intense Variety, being Valuable, and needing to check their Veracity. LPWAN architectures that address these challenges have already emerged [44]. For the definition of such architectures, common reference models from the broader Industry 4.0 sector are used [45]; namely, the Reference Architectural Model Industrie 4.0 (RAMI4.0) [46], the Industrial Internet Reference Architecture (IIRA) [47], the International Data Spaces Reference Architecture Model (IDSRAM) [48], and the IoT World Forum’s (IoTWF) Reference Model [49].

While in the following sections we present LPWANs from the point of view of networking between the EDs and the GWs, in this section we focus on the result of the basic function that the EDs perform: through sensors they receive data and transmit it to some aggregation point - the data is the

3GCP includes in its cloud pallet the IoT Core service (not to be confused with AWS’s IoT Core). However, during August 2022, Google announced that will retire IoT Core by mid-August 2023, prompting the users to migrate to a specialized IoT cloud service provider.
result. In this section, we briefly study all the basic operations on data in LPWANs: acquisition, storage, analysis, and visualization. In later sections, networking issues are studied, either for each technology separately (Sections VI–VII), or focusing on LoRa/LoRaWAN networks (Sections VIII–X).

### A. Fundamental Data Processes for LPWANs

Data management in LPWANs includes all the fundamental processes that are applied on data when it comes to communications systems: namely, data acquisition, data storage, data visualization, and data analysis.

**Data acquisition** is at the heart of sense and actuation systems, since EDs’ main operation is collecting and/or acting on data. Two distinct approaches exist: the pull model and the push model. The pull model expects from the user (usually sitting beyond the server-side) to ask for the data. A key representative of the pull model is the HyperText Transfer Protocol (HTTP), which permits a user to apply the GET method to retrieve information.4 On the other hand, the push model supports the dispatch of data to the interested user by the time the data are transmitted by EDs. The Message Queuing Telemetry Transport (MQTT) protocol and Webhooks are well-established and widely used push methods. Due to the increasing interest in WSN, the Organization for the Advancement of Structured Information Standards (OASIS), which develops MQTT, is working towards the standardization of MQTT for Sensor Networks (MQTT-SN).

**Data storage** is also of vital importance for the successful setup of an LPWAN system. An initial approach is to store data in plain files with .json, .csv, or .txt (less frequently used) extensions. JavaScript Object Notation (JSON) has gained a lot of attention in recent years, and along with Comma Separated Values (CSV), is considered the main human-readable data representation format today. eXtensible Markup Language (XML) is another option and its applicability in semantic Web and Web of Things (WoT) [50] context can be proven useful to system developers focusing on such data interchanges. Efficient XML Interchange (EXI) is another step towards this direction [51], having been developed for computationally constrained devices. However, a more sophisticated approach would dictate the utilization of a database paradigm.

Several options exist, with the main considerations being: a document-oriented database (e.g., DynamoDB, MongoDB), which may also export data in the form of popular representations, such as .json, .csv, and .xml extensions files; a timeseries database optimized for the flow of IoT data (e.g., InfluxDB, Prometheus); or a graph-based database, in which each graph node could be a UL transmission distinguished by the reception time from the GW, with the sequence of transmissions creating a chain (e.g., Dgraph, Neo4j).

**Data visualization and analysis** is applied on stored data to transform them into information. Grafana is an open platform that is utilized in numerous established projects. Mathworks has developed a solution that offers both visualization capabilities and support of other operations upon data: the ThingSpeak platform. In the same field, platforms like TagoIO are also widely used. MATLAB is used in the ThingSpeak platform for data analysis and processing, whereas TagoIO supports JavaScript (Node.js) and Python. Naturally, LPWANs support the current evolution in IoT data analysis (and especially in industrial data management), towards exploiting Artificial Intelligence (AI) toolkits. In this context, Machine Learning (ML) models can be utilized, based on well-established toolboxes or frameworks, such as PyTorch and TensorFlow. ML models can be deployed not only in system components with high computing capabilities, such as a GW, an NS, or a monitoring platform, but also in EDs themselves, thanks to TinyML [52].

### B. Advanced Data-Driven Services on Top of LPWANs

Beyond the fundamental data processes, innovative data-driven concepts have emerged (especially in the industrial sector). Some of the most critical ones are the following: i) common frameworks for representation of the nodes of a network, referring mainly to the concept of Asset Administration Shell (AAS) [53], including common data formats and communication protocols, referring mainly to the work conducted by the Open Platform Communications (OPC) Foundation and especially the OPC-Unified Architecture (OPC-UA) group; and ii) the potential of exploiting the common representation and data modeling schemes towards the digital twinning of LPWAN systems.

**Representation and Data modeling:** The term “assets” (as defined within the manufacturing sector) refers to any interconnected device, sensor, actuator, machine, or even processes and systems, inside a controlled communication environment (e.g., inside a manufacturing plant). Considering this, AAS is the basis for the discovery, access, and description of an asset [54]. The specifications of the AAS are defined in IEC 62378 [55], while its implementation is a subject of many open-source approaches, such as the AASX [56] and the Eclipse Basyx [57]. **Discovery** is a fundamental mechanism in AAS and a prerequisite for any communication establishment. There are different approaches for that, including the use of directories-style systems, search engines, crawlers, networking protocols, etc. **Access** refers to a set of APIs used for accessing the assets. The use of APIs ensures unified access, at the cost of requesting every resource to implement them. Asset description refers to data models for describing the asset procedures and status, such as the OPC-UA models, defined by the OPC Foundation. The communication model of OPC-UA fits well in the LPWAN networks, since it is a pub/sub protocol, as described in open62541 [58].

One step beyond the default data models of OPC-UA, new information models are created, which are called Companion Specifications, and they are specific to vertical industries (as for instance those defined by the Mechanical Engineering Industry Association VDMA). In addition, the Digital Twin (DT) community has built related tools and models, such as the DT Definition Language (DTDL) [59], which specifies an Industry 4.0 language used for inter-AAS communication.

*The digital twin concept* [54], [60]: DT refers to the digital presence of any “asset” enriched with full administration
flows, the Node-RED tool is also considered as a DT-enabler. Moreover, dedicated cloud-based tools in digital twinning have already emerged, as for instance the PaaS solution of Azure DT. The realization of a DT could be based on emulation-oriented systems, such as those provided by the Emulated IoT (ELIoT) [63], [64]. In ELIoT, the EDs can be abstracted as clients and the Network Gateway as server, and an LPWAN system can be emulated as a client-server model. For the communication part of a system and its data flows, the Node-RED tool is also considered as a DT-enabler.

VI. LPWAN TECHNOLOGIES

Having laid the foundation of LPWANs, we proceed with a discussion on wireless technologies that can support such type of networks.

A. EC-GSM-IoT

The EC-GSM-IoT technology was first introduced in Release 13 by 3GPP, while further improvements have been added in recent Releases. The main target of the technology was to use the existing mobile network infrastructure and functionality (mainly 2G/GSM) for establishing long range IoT communications.

A key point when developing this technology was to maintain the advanced features of GSM and of General Packet Radio Service (GPRS), while extending the coverage area of these systems and optimizing the protocol for IoT devices [66]. More specifically, EC-GSM-IoT is based on enhanced GPRS (eGPRS)/Enhanced Data rates for Global Evolution (EDGE), a 2.75G wireless cellular technology, which serves as an improvement of GSM in terms of data rate. GSM offers extended coverage compared to other cellular technologies. Its improvement, eGPRS/EDGE, maintains this advantage, while supporting higher data rates (384kb/s). EC-GSM-IoT combines the coverage advantage of GSM with the higher data rates of eGPRS/EDGE; it attains a maximum data rate of 240kb/s.

Contrary to GSM and eGPRS/EDGE, EC-GSM-IoT uses enhanced Discontinuous Reception (eDRX) of 52min, compared to GSM’s DRX of 11min, performs less operations during inactive periods and utilizes a Power Saving Mode (PSM), leading to an extended ED lifetime [67]. Other features of EC-GSM-IoT are the optimization of the utilized system information and the absence of redundant re-transmissions; factors that are considered critical in the effort to reduce device complexity and further improve battery usage [67]. Also, EC-GSM-IoT is equipped with additional channels to achieve wider coverage, with an overlaid Code Division Multiple Access (CDMA) scheme being used for increased capacity [68]. Security improvements in the areas of integrity, authentication of both ends, and the use of ciphering algorithms, classify EC-GSM-IoT as LTE-level secure [68].

In the deployment field, EC-GSM-IoT is implemented using low-cost EDs that can operate for up to ten years with an average battery usage of 5W/h. These EDs transmit in multiple GSM frequency bands depending on the deployment area, ranging from 395MHz to 1960MHz, using 200kHz channels. The deployment of EC-GSM-IoT’s channels is performed in a GSM in-band scheme. Messages are transmitted in two power classes: either at 23dBm or at 33dBm. These two classes define the two coverage areas cases; the former having a Maximum Coupling Loss (MCL) of 154dB and the latter of 164dB. The data rate depends on the modulation type. The use of Gaussian Minimum Shift Keying (GMSK) offers data rates between 0.35kb/s and 70kb/s, while the use of 8-Phase Shift Keying (8PSK) offers a maximum of 240kb/s. These data rates correspond to the UL as well as the Downlink (DL). Apart from CDMA, Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) are used as multiple access schemes in both the DL and the UL.

B. LTE-M

Another CIoT technology that can satisfy the needs of low power and extended coverage is LTE-M. We note that in the context of this paper, with the term LTE-M we refer to the common characteristics of LTE-MTC, LTE Cat M1, and LTE-M technologies. LTE-M, as a CIoT technology, utilizes the LTE spectrum for establishing wireless communication links, allowing for coexistence with other cellular technologies, such as 2G, 3G, 4G, and 5G.

The LTE-M deployment costs and time-to-market are reduced, since LTE’s BSs can be used for LTE-M links with a software upgrade. LTE-M adopts security and privacy mechanisms from LTE to cope with confidentiality, end user (i.e., ED) authentication, and message integrity. Moreover, with the improvements added in 3GPP Releases 14 and 15, LTE-M EDs with mobility capabilities can be supported, since in-band LTE operation permits handovers from one cell to the other. Other improvements added in Releases 14, 15, and 16 include the standardization of a new device category, namely Cat. 2, which offers substantial enhancements in ED’s performance. Furthermore, Voice Over LTE (VoLTE) features were added in Release 14 [69]. Release 15 built upon these improvements and presented new use cases, supporting higher ED mobility, adding a new power class of 14dBm, reducing latency and ED’s power consumption, increasing spectral efficiency, and improving access control [66]. The evolutionary approach of the technology continued in Release 16. Worth mentioning enhancements include new power saving techniques for EDs, extended support of EDs’ mobility, and improvements that targeted coexistence with New Radio (NR), i.e., the 5G radio interface [70].

Overall, LTE-M is based on a simplified version of LTE protocols and functionality towards reducing ED complexity and cost, and improving power consumption. Specifically, in LTE-M, there is no control channel for wideband operation, Transmission Modes (TM) support is reduced; Hybrid
Automatic Repeat Request (HARQ) is also reduced [68], since the message size is already limited and the addition of a large number of redundant bits for error control purposes should be restricted.

In the deployment field, LTE-M EDs are able to operate in low power modes and utilize power saving techniques, such as PSM, eDRX, and Connected mode DRX (C-DRX), to live up to ten years with a 5W/h battery consumption. The frequency bands used for message transmissions are cellular, ranging from 462.5MHz to 2655MHz, and their selection is based on each area’s spectrum regulation and operators’ decisions. The Bandwidth (BW) is 1.08MHz and the in-band LTE operation is applied. LTE-M defines two power classes: the class of 23dBm and the class of 20dBm. Both classes result in the same coverage area, with an MCL of 155.7dB. Despite the extended area that the system can cover, data rates are the highest among other LPWAN solutions, reaching 1Mb/s for both streams. Especially for Cat. 2 devices, data rates of 7Mb/s in UL and 4Mb/s in DL using 5MHz BW can be reached. For both UL and DL, Quadrature Phase Shift Keying (QPSK) and 16-Quadrature Amplitude Modulation (16-QAM) are utilized for modulation purposes. In UL, Single Carrier-FDMA (SC-FDMA) is used as multiple access scheme, whereas in DL the use of Orthogonal FDMA (OFDMA) is preferred; same as in LTE networks. It is worth mentioning that techniques such as 15kHz tone spacing and Turbo codes are utilized to increase system performance in both the UL and the DL [68].

C. NB-IoT

NB-IoT is one of the most prominent LPWAN CIoT technologies. NB-IoT defines a radio interface for IoT communications, moving one step further from EC-GSM-IoT and LTE-M technologies, which are solely based on GSM and LTE radio interfaces, respectively.

NB-IoT is specified in 3GPP Release 13; much like EC-GSM-IoT and LTE-M, future improvements of the technology are included in Releases 14, 15, and 16. Specifically, in Release 14, a new device power class was established at 14dBm, mobility improvements were introduced, and enhancements in power consumption and latency were proposed among other upgrades of the technology [69]. Data rate was also improved to reach the scale of 250kb/s for both UL (for multi-tone deployment) and DL. Release 15 brought further improvements, including increase of the system’s range to 120km, support of micro-cells, pico-cells, and femto-cells, and wake-up enhancements for EDs that operate in DRX or eDRX [66]. Release 16 expanded the technology even further, with key improvements among others being power saving techniques in the wake-up procedure of EDs, directions towards Self Organizing Networks (SONs), greater mobility support, and coexistence of the technology with NR [70].

As a 3GPP CIoT technology, NB-IoT utilizes bands in the licensed part of the spectrum, both in the sub-GHz region, such as B8 or B26, and above 1GHz, such as B66 or B74. The BW of the system is 180kHz. Three modes of operation are defined, as depicted in Fig. 5 [71]. In the stand alone mode, NB-IoT utilizes a portion of the spectrum that is used for GSM EDGE Radio Access Network (GERAN) by operating in a GSM carrier. In the guard-band mode, unused resource blocks inside an LTE carrier are utilized. These blocks are at the edges of the carrier and are typically used for tackling interference issues. Finally, in the in-band operation, NB-IoT uses an LTE Physical Resource Block (PRB); hence the size of 180kHz BW, since LTE uses twelve subcarriers of 15kHz each. The mode of operation is unknown to the ED when it connects to the network [72].

In the deployment field, NB-IoT uses the same power saving techniques with LTE-M; namely, PSM, eDRX, and C-DRX. Battery life is calculated to be approximately ten years with the typical 5Wh battery usage, while solutions are designed to permit EDs to last even longer. NB-IoT’s key advantage is its MCL value, which can reach 164dB for standalone deployment, allowing extremely extended coverage. Supported modulation schemes include QPSK in both UL (multi-tone) and DL, and $\frac{\pi}{2}$-QPSK or $\frac{\pi}{2}$-Binary PSK (BPSK) in UL (single-tone) [23]. Being able to coexist with and utilize LTE bands, NB-IoT uses SC-FDMA in UL and OFDMA in DL for multiple access; however, the area of access mechanisms in NB-IoT is undergoing extensive research [73]. Again, to improve system performance, techniques such as 15kHz tone spacing in both streams (or even 3.75kHz tone spacing in single-tone UL operation) and Turbo coding in UL are used. A complete analysis of the PHY of NB-IoT is provided in [74].

D. MIOTY

Moving to non-3GPP technologies that operate in unlicensed bands, the first examined candidate is MIOTY. MIOTY is an LPWAN standard that follows one of the 3 European Telecommunications Standards Institute (ETSI) specifications for the so-called Low Throughput Networks (LTN). In [75], ETSI defines 3 LTN protocol families: Lfour, Telegram Splitting Ultra Narrow Band (TS-UNB), and Dynamic Downlink Ultra Narrow Band (DD-UNB). MIOTY is a realization of the second protocol family, TS-UNB, and is developed by the Fraunhofer Institute.\footnote{\textsuperscript{5}As of 2020, MIOTY is promoted by MIOTY Alliance.}

MIOTY can offer communication services to EDs with high mobility of up to 120km/h. The underlying LTN protocol, i.e., TS-UNB, offers great advantages in the area of robustness against co-channel interference; a factor highly critical for links operating in the unlicensed part of the spectrum. This robustness is achieved thanks to GW’s ability to decode the entire message using a portion of the received packets. Adding to Telegram Splitting (TS) benefits, the split messages require less time to transmit, resulting in reduced Time on Air (ToA), and, hence, an improvement in battery usage. The Fraunhofer Institute promotes the usage of innovative energy harvesting techniques to achieve greater ED lifetime.

\textsuperscript{5}As of 2020, MIOTY is promoted by MIOTY Alliance.
As a TS-UNB technology, MIOTY uses the Telegram Splitting Multiple Access (TSMA) for allowing more than one EDs to access the communication channel. In TS, the packets to be transmitted are first divided into several smaller sub-packets and then they are sent in a non-continuous fashion, with each transmission being separated by the other in the time domain. The sub-packets are distributed pseudorandomly in both time and frequency within a certain radio frame. More on TS can be found in [75], [76], and [77].

In the deployment field, two power classes are supported by MIOTY, based on the frequency band used for the communication: 868MHz in Europe and 915MHz in North America. This leads to the 14dBm power class for the European band and the 23.3dBm class for the North American band. With these power classes and proper configuration, EDs can operate in the field for about twenty years. The link budget of a MIOTY link is calculated at 154dB, with a receiver sensitivity of −139dBm [75]. Data rate, by official statement, is 0.514kb/s, but estimations set it at 0.4kb/s, by using a BW of 200kHz.

A newer feature of MIOTY is the differentiation of EDs in communication classes based on the application scenarios. For monitoring-only purposes, MIOTY’s Class Z EDs transmit only UL packets without DL reception capabilities. Class A EDs offer additional (limited) DL options, in the form of configuration messages sent by a BS to an ED. Finally, the recently introduced Class B supports the transmission of DL commands in both multicast and broadcast mode.

E. RPMA

RPMA was invented and patented [78] by Myers with On-Ramp Wireless Inc. assignee the company that is now called Ingenu. RPMA is explained as a concept in [15] and [79]. It is actually a Direct Sequence Spread Spectrum (DSSS) technique that pseudorandomly delays the transmissions of EDs to minimize overlapping occurrences, resulting in increased Signal to Interference and Noise Ratio (SINR) value for every ED.

Contrary to other popular LPWAN technologies, and especially the unlicensed ones, RPMA does not make use of the sub-GHz frequency bands for link establishment, but it uses the globally available 2.4GHz band for wireless transmission. Despite not having the added advantage of high penetration due to low frequency that other LPWAN solutions have, RPMA behaves extremely well in terms of coverage. For tackling the issue of interference in a band that also operates one of the most widely used wireless technologies, i.e., WiFi (IEEE 802.11), RPMA utilizes its high robustness and a frequency plan that assigns small channels in-between the channels of WiFi [80]. For access control over the wireless channel, RPMA uses the namesake multiple access scheme.

Another key feature of RPMA is its solid security and privacy support. EDs communicate with RPMA BSs having ensured authentication for both end points of the link; messages are sent encrypted by the EDs, and the backhaul communication of the system is achieved via Secure Sockets Layer (SSL)/Transport Layer Security (TLS) connections. In the deployment field, RPMA uses a significantly higher BW than the typical case of LPWANs, reaching 1MHz [81]. Regarding transmission power, EDs transmit at a maximum of 21dBm in both Europe and the U.S., while RPMA BSs can transmit at 21dBm in Europe and at 30dBm in the U.S. [82]. In this power class, and by optimizing the usage of EDs for longevity, EDs can last in the field for more than ten years; with rare message transmissions and the use of small packets, it is estimated that EDs can last for about twenty years.

Another advantage of RPMA is its very high link budget. It can reach 168dB, with a receiver sensitivity of −142dBm for BSs and −133dBm for EDs. The data rate is estimated to be 78kb/s for UL and 19.5kb/s for DL, at maximum [15], [25].

Interestingly, as of 2018, Ingenu has decided to offer RPMA as a PaaS solution [83].

F. Weightless

The Weightless Special Interest Group (SIG)6 started developing three solutions for LPWANs, namely Weightless -W, -N, -P. Weightless-W operates in the TV White Space (TVWS) region, making it a worth examining solution, but the different regulations in different parts of the world regarding TVWS spectrum complicate its deployment. Weightless-N operates in the sub-GHz unlicensed region and seems like a promising solution due to its numerous similarities with Sigfox, one of the most popular LPWANs. The technology is used by NWave; however, Weightless-N’s link budget is considered to be floating [84], leading Weightless SIG to another alternative. The suggested solution is Weightless-P [85], a technology that this paper focuses on.

Weightless-P supports fully acknowledged (ACK) bi-directional links; various techniques can be used to offer Quality of Service (QoS), such as Forward Error Correction (FEC) in block interleaving mode, Cyclic Redundancy Code (CRC), and ARQ. It also offers several advanced security and privacy features, with mutual authentication via Advanced Encryption Standard (AES) 128/256, message integrity, confidentiality via AES 128/256, and secure Over The Air (OTA) key exchange being supported [86].

In the deployment field, the BSs transmit at a typical power of 27dBm, while being able to transmit also at 30dBm. EDs have a lower transmit power, at 14dBm, being also capable for 30dBm transmissions. Importantly, the power consumption of EDs at idle state is 100µW. With this power usage, an ED can last for three to eight years [87]. Even though the technology can theoretically operate at both licensed and unlicensed spectrum, for the time being, it operates in the following unlicensed bands in the sub-GHz region: 138MHz, 433MHz, 470MHz, 780MHz, 868MHz, 915MHz, and 923MHz. The choice of band is subjected to each area’s frequency regulation. BW is set at 12.5kHz or 100kHz for UL, while at DL broadcast, unicast, and multicast channels exist, using 100kHz, 50kHz, or optionally 12.5kHz.

The link budget of Weightless-P can reach at most 153dB, assuming an Additive White Gaussian Noise (AWGN) channel. Its value depends on transmission parameters, with the

6Now called Weightless Alliance.
highest figure being obtained for UL transmissions when the receiver sensitivity is set at $-135.5\text{dBm}$, with a required Signal to Noise Ratio (SNR) of $-17.5\text{dB}$, for a typical Bit Error Rate (BER) of $10^4$, while Coding Rate (CR) is set at $\frac{3}{4}$, and Spreading Factor (SF) at 8. The $153\text{dB}$ value is true, if Offset QPSK (OQPSK) is used as a modulation method. The other modulation method that the technology utilizes is GMSK with a Bandwidth-symbol Time (BT) product of 0.3. These modulations are used in both UL and DL. For multiple access, UL messages use TDMA and FDMA combined, whereas in DL only FDMA is utilized. Considering the data rates, the maximum value is $100\text{kb/s}$ in both streams for $100\text{kHz}$ BW, GMSK modulation, and proper CR selection, but in UL falls to $10\text{kb/s}$ at maximum for $12.5\text{kHz}$ BW. OQPSK performs worse in terms of data rates; a penalty that has to be paid for the great system range it offers.

G. Wize

Different from the techniques presented so far, Wize provides, in addition to the PHY functionality, a full protocol stack. It also adopts the conventional LPWAN architecture; however, the server-side is a data collection system called Head-End, whereas LAN modems, placed between the GWs and the EDs, could be used to serve multiple EDs. The Wize protocol stack includes four layers: PHY, Data Link, Presentation, and Application (APP). The Wize specification is supported by the Wize Alliance and can be found in [88]. Based on this specification, we describe below (following a bottom-up approach) the main aspects of each layer.

In the deployment field, the related PHY parameters are described in [89]. Practically, the $169\text{MHz}$ Industrial, Scientific, Medical (ISM) band is utilized and targets deployments in the European region. Six channels are defined, each $12.5\text{kHz}$ wide, permitting transmissions with a power of up to $27\text{dBm}$, having the obligation to respect a duty cycle of $10\%$. The modulation used is either Gaussian Frequency Shift Keying (GFSK) or 4GFSK, both with a BT product of $0.5$. The former modulation can give data rates of either up to $2.4\text{kb/s}$ or $4.8\text{kb/s}$, while the latter can go as high as $6.4\text{kb/s}$. Optimal sensitivity by popular market products is achieved with the $2.4\text{kb/s}$ PHY, giving a value of $-119\text{dBm}$ [90], resulting in a maximum link budget of $146\text{dB}$.

On top of PHY resides the Wize Data Link layer. Two different packet formats are defined: the DL broadcasting format and the default UL/DL format. In the UL, multiple transmissions of the same message (repetitions) are allowed, to increase the probability of message reception. Possible duplicates that may appear due to the redundancy or the broadcasting nature of UL messages are resolved in the Head-End. In the DL, unicasting transmissions for device-specific commands or broadcasting software update commands are sent to EDs. As in many LPWANs, the UL-initiated communication leads to low power consumption. In Wize, the EDs remain idle during non-UL periods, except for a small DL window, which opens shortly after the UL.

Security aspects and metadata management are performed at the Wize Presentation layer. Regarding system security, Wize supports both access authentication using AES Cipher-based Message Authentication Code (CMAC) for ED - GW communication and end-to-end authentication using AES CMAC. Encryption using AES Counter (CTR) mode is also provided.

The data plane of the APP is related to the use case supported. Wize targets mainly smart city applications, with gas metering and water management verticals being the most adopted ones. Regardless of the use case, application administration options (control plane) are offered as well.

H. Sigfox

Sigfox is one of the widely adopted non-3GPP technologies for LPWANs and owes its name to the homonymous company, Sigfox, which first introduced the technology.\(^7\) A comprehensive review of the technology can be found in [91]. Based on the Sigfox set of specifications, EDs can be developed, while the server-side is also provided by Sigfox Cloud. EDs communicate with BSs via Sigfox links and the backhaul communication is performed within a Virtual Private Network (VPN), via fixed cable links (Digital Subscriber Line (DSL)) as a primary solution and cellular technologies as a backup. Satellite support can be used as an alternative solution as well [92].

In the deployment field, Sigfox defines three power classes for EDs, according to the region that the technology is used. Currently, there are seven region zones, named RC1, RC2, …, RC7. RC1 consists mostly of European countries and parts of Africa and Middle East; RC2 includes North America and parts of South America; RC3 is dedicated to Japan; RC4 covers some other countries of South America and the Asia-Pacific area; and zones RC5, RC6, and RC7 are exclusive for South Korea, India, and Russia, respectively. For RC1, RC3, RC6, and RC7 the maximum Transmission Power (TP) is set at $16\text{dBm}$, for RC5 at $14\text{dBm}$, whereas in RC2 and RC4 at $24\text{dBm}$. However, TP also depends on the device class. In Sigfox, four device classes are defined. $U_i$, $i=0 \ldots 3$, with $U0$ being the strongest and $U3$ the weakest. For $U1$ devices, TP is in the range of $[TP_{U0-9dB} \, TP_{U0-4dB}]$, for $U2$ the range is $[TP_{U0-14dB} \, TP_{U0-9dB}]$, and for $U3$ we have a TP below $TP_{U0-14dB}$. This TP, along with other technology parameters, allows battery powered EDs to have an autonomy at the scale of a decade.

Sigfox’s frequency plan is also region-based and utilizes ISM parts of the spectrum. Devices that operate in RC1 areas transmit at $868.130\text{MHz}$, in RC2 at $902.200\text{MHz}$, in RC3 at $923.200\text{MHz}$, in RC4 at $920.800\text{MHz}$, in RC5 at $923.300\text{MHz}$, in RC6 at $865.200\text{MHz}$, and in RC7 at $868.800\text{MHz}$. The total BW used in a region is $192\text{kHz}$ and the transmission BW is only $100\text{Hz}$; i.e., this is an implementation of an Ultra Narrow Band (UNB) scheme.

The transmission rate for Sigfox is $600\text{b/s}$ in the DL and either $100\text{b/s}$ or $600\text{b/s}$ in the UL, depending again on the region. RC2 and RC4 support $600\text{b/s}$, whereas the rest of the regions allow UL transmissions of $100\text{b/s}$. These data rates allow battery powered EDs to have an autonomy at the scale of a decade.

\(^7\)Sigfox was acquired by UnaBiz on April 21st, 2022, (https://www.sigfox.com/en/news/una-biz-appointed-new-owner-sigfox-sa-and-sigfox-france-sas-under-receivership-proceeding, Accessed: Jun. 6, 2022).
might seem extremely low, but in practice they are sufficient, considering the amount of information transmitted and the size of the packets. In Sigfox, packets have a payload size of 0-12Bytes and a total size of 26Bytes for the UL and a fixed payload size of 8Bytes for the DL. As a result, the transmission of 26Bytes with a 100b/s rate needs only 2.08s. To transmit their messages, devices utilize the available channels in a Random-FDMA (RFDMA) scheme. EDs choose three channels in a pseudo-random fashion to transmit their message three times in three different frequencies, in an effort to increase the probability of successful message reception by at least one BS, thus decreasing the amount of lost information and improving the QoS.

The total number of messages an ED can send in one day depends on several parameters including the region, the payload size, and the bit rate. In RC1, RC6, and RC7, a duty cycle of 1% is imposed due to regulations; in RC2 and RC4, a Frequency Hopping (FH) scheme is adopted as described earlier with the three times transmission approach; and in RC3 and RC5, Listen Before Talk (LBT) is preferred. In the European region case, for a 12Bytes payload (26Bytes total size) message transmissions with rate 100b/s in a 1% duty cycle, an ED can transmit six messages an hour and one hundred forty messages a day. In DL, messages are less frequent and are initiated by a UL message; with a DL duty cycle being of 10%, the maximum number of messages per day and per device is equal to four. The sensitivity of a Sigfox BS is also region dependent, and for a typical RC1 transmission can reach $-142\text{dBm}$, resulting in a theoretical link budget of 163.3dB. Nevertheless, a typical link budget of 156dB is considered reasonable.

### I. LoRa/LoRaWAN

LoRa refers to the PHY of the protocol stack and, more precisely, to a proprietary Chirp Spread Spectrum (CSS) modulation, developed by Semtech [93]. Despite being tightly connected with LoRaWAN, it should be noted that LoRaWAN is an open Medium Access Control sublayer (MAC) protocol residing on top of LoRa and maintained by LoRa Alliance.

LoRa EDs may transmit at a maximum power of 16dBm, whereas GWs have two power classes: one defined by the same maximum TP with the EDs, and the other at a maximum value of 27dBm. Chipsets built by Semtech support transmission of packets in ISM bands, on a region dependent fashion or by utilizing universally available bands, such as 2.4GHz. Transmission frequencies are set by the protocol used in MAC. LoRaWAN, which is the main proposal in this area, defines eight ISM regions [94], as described later. Bandwidth is also non-fixed, and can be either at 125kHz, 250kHz, or 500kHz, in several widely used chipset implementations, while other realizations can go down to 7.8kHz [95].

Several LoRa parameters, such as BW, may differ from one link to another, resulting in a link budget that can be adapted to each use case, with its maximum value reaching 154dB. Other parameters that can be configured in a LoRa link include Carrier Frequency (CF), SF, CR, and TP [96]. CF is typically adjustable and depends on the deployment region. For SF, the options are six: from SF7 to SF12. As for CR, it should be selected from the set $\{\frac{1}{4}, \frac{1}{5}, \frac{1}{7}, \frac{1}{8}\}$. TP can be set at a lower value than the maximum allowed if the channel quality permits it. Data rate is variable and can take values in the interval of [0.162kb/s, 50kb/s]; however, LoRaWAN-based systems that implement LoRa in PHY may achieve data rates of up to 5.47kb/s, 11kb/s, or 21.9kb/s depending on the deployment region (despite Semtech’s chipsets having greater capabilities), whereas the maximum value of 50kb/s is achieved via Frequency Shift Keying (FSK) modulation. LoRaWAN supports also Long Range - Frequency Hopping Spread Spectrum (LR-FHSS) in some regions, with a maximum bit rate of 0.162kb/s or 0.325kb/s, again in a region-based fashion [94].

Since LoRa operates in ISM bands, transmissions should respect the use of medium by other technologies. Common regulations regarding LoRa include the modulation of duty cycle, i.e., the percentage of time that an ED is active transmitting, with values such as 0.1%, 1%, and 10%. Practically, the duty cycle regulates the ToA of a LoRa transmitted packet for several regions. The ToA regulator role may also be performed by dwell time, whereas some regions are subjected to both regulations. Instead of duty cycle, LBT is also an option in a few cases.

With the LoRaWAN MAC protocol, EDs are accessing the transmission medium using an ALOHA-type scheme with variable frame sizes [97]; its performance has been measured to be better than ALOHA owing to SF orthogonality, other transmission characteristics (e.g., capture effect), and protocol improvements [98], [99]. Nevertheless, multiple access performance is affected by the imperfect nature of SF orthogonality [100], [101]. The topic of multiple access has attracted several research efforts, and a number of different approaches have been proposed [102], [103], [104], [105], [106]. In LoRaWAN, DL transmissions are either unicast for every ED class specified in the protocol (Classes A, B, and C) or multicast for ED Classes B and C.

### J. Other LPWAN Technologies

The LPWAN area is developing at a rapid pace, and more technologies are emerging. As discussed earlier, one of the Weightless SIG proposals is Weightless-N, which is implemented by a smart cities oriented system, called Nwave [107]. Major industry players, such as Sony and Amazon, offer their solutions, namely Eltres [108] and Sidewalk [109], respectively (Sidewalk operates with LoRa and/or FSK and/or Bluetooth Low Energy (BLE)). Sensor Networks Over White spaces (SNOW) [110] is another worth considering option for LPWAN deployment, while Wi-Fi HaLow [17], [111] is gaining more and more attention. DASH7 [112] is a well-established protocol, being supported by DASH7 Alliance. The interested reader can also search for alternatives, such as Telensa, NB-Fi, iQRF, Thingstream, Helium, SAT4M2M, Hiberband, and Wi-SUN.

---

3RC3 also supports 1% duty cycling, [https://build.sigfox.com/sigfox-radio-configurations-rc3](https://build.sigfox.com/sigfox-radio-configurations-rc3), Accessed: Aug. 31, 2022.

4LoRaWAN has been approved as an International Telecommunication Union (ITU) Recommendation as of November 29th, 2021, [https://www.itu.int/rec/T-REC-Y.4480-202111-I/en](https://www.itu.int/rec/T-REC-Y.4480-202111-I/en), Accessed: Jun. 6, 2022.
TABLE I
OPERATIONAL CHARACTERISTICS OF VARIOUS LPWAN TECHNOLOGIES

| Technology     | Multiple Access [UL / DL] | Modulation [UL / DL] | Bandwidth (kHz) [UL / DL] | Frequency (MHz) | Maximum Data Rate (kb/s) [UL / DL] | Maximum Transmission Power (dBm) [UL / DL] | Maximum Link Budget (dB) |
|----------------|---------------------------|----------------------|---------------------------|----------------|------------------------------------|---------------------------------------------|-------------------------|
| EC-GSM-IoT     | TDMA, FDMA, or CDMA       | GMSK or 8PSK         | 200                       | cellular bands | GMSK: 70 / 8PSK: 240              | 23 or 33                                    | 164 [MCL]               |
| LTE-M          | SC-FDMA / OFDMA           | QPSK or 16-QAM       | 1040 or 5000              | cellular bands | 1000 or 7000 / 1000 or 4000       | 14, 20, or 23 (3GPP Release dependent)      | 155.7 [MCL]             |
| NB-IoT         | SC-FDMA / OFDMA           | 4-QPSK, 4-DPSK, or QPSK/QPSK | 180               | cellular bands | 250                                | 14 or 23 (3GPP Release dependent)           | 164 [MCL]               |
| MIOTY          | TSMA                      | TS-UNB               | 200                       | ISM region dependent | 0.4                                | 14 or 23.3 (region dependent)              | 154                     |
| RPMA           | RPMA                      | DSSS (RPMA)          | 1000                      | ISM 2400        | 78 / 19.5                          | 21 / 21 / 21 / 30 (region dependent)        | 168                     |
| Weightless-P   | TDMA and FDMA / FDMA      | OQPSK or GMSK        | 12.5 or 100 / 12.5, 50, or 100 | ISM region dependent | 10 or 100 / 100 (modulation and BW dependent) | 14 or 30 / 27 or 30 (region dependent)      | 153                     |
| Wize           | ALOHA-like (unslotted)    | GFSK or 4GFSK        | 12.5                      | ISM 169         | GFSK: 2.4 or 4.8 / 4GFSK: 6.4     | 27                                          | 146                     |
| Sigfox         | ALOHA-like (unslotted), Random-FDMA | DBPSK / GFSK | 0.1                  | ISM region dependent | 0.1 or 0.6 / 0.6 (region dependent) | 14, 16, or 24 (region dependent)           | 156                     |
| LoRa/LoRaWAN   | ALOHA-like (unslotted)    | LoRa (CSS), FSK, or LR-FHSS | 125, 250, or 500         | ISM region dependent | LoRa: 5.47, 11, or 21.9 (region dependent) | FSK: 50 / LR-FHSS: 0.162 or 0.325 (region dependent) | 154                     |

VII. COMPARISON OF TECHNOLOGIES

In this section, we compare the LPWAN technologies presented in Section VI, in terms of operational characteristics, application perspectives, and deployment aspects.

A. Operational Characteristics

Based on the presentation of the main LPWAN technologies, provided in the previous section, Table I summarizes their key characteristics. Multiple access schemes utilize time, frequency, and code division, as well as ALOHA-like random access. Both narrowband and wideband modulations are used. The frequency bands used in each technology are largely dictated by the standardization body, with all non-3GPP solutions but RPMA operating in the sub-GHz region. On the other hand, RPMA uses the 2.4GHz band, whereas 3GPP technologies operate in cellular bands decided by the operator. Data rates are mainly of the order of 10^1kb/s, with only a few exceptions. All technologies have similar maximum transmission power, offering a low power class in 14dBm or 16dBm, and a high power class between 20dBm and 30dBm.

As it can be observed in Table I, all the technologies have a link budget of over 146dB, justifying their capability of covering the needs of LPWAN systems. Nominal values for the achievable ranges are given below:

- The existing literature agrees on the following values for the maximum achievable range for the following technologies: LTE-M, 10km regardless of the environment; Weightless-P, 2km in urban environments; RPMA, 15km in rural environments; LoRa/LoRaWAN, 15km in rural environments and 5km in urban environments.
- For NB-IoT, most publications place the maximum achievable range in rural environments at 15km; however, there are references that put it either higher, at 35km [16], [25], or lower, at 10km [22]; for urban environments, the value is 1km. For Sigfox, in rural environments most publications mention the 50km threshold, with one reference giving a lower value at 40km [22], whereas for urban environments the values of 5km and 10km are mentioned in the literature.
- For EC-GSM-IoT, MIOTY, and Wize no nominal values are mentioned in the set of the referenced surveys. Nevertheless, for all three of them there exist references by organizations or vendors, that place their maximum range at 15km [67], 20km [113], and 50km [114], respectively.

The achievable range is heavily affected by several parameters. Firstly, it is clear from the range values presented above that the communication environment (urban vs. rural) plays a major role. Secondly, the path between transceivers is of crucial importance. LOS links perform orders of magnitude better than non-LOS ones (e.g., for RPMA, it is reported that LOS links can reach up to 500km [17], [21]); also, significant differences may exist between different non-LOS links, depending on the propagation properties of the specific environment (e.g., residential buildings vs. harbor-like premises [30]). Thirdly, the range depends on hardware implementation characteristics, and specifically on the radio modules of both end points.
Table I includes the maximum transmission power for both communication end points and the maximum link budget, for every technology. The interested reader is encouraged to use these values to calculate an estimate of the maximum range supported by each technology for each deployment environment. Furthermore, in Section XI we refer to numerous research publications that examine the impact of environmental parameters in LoRa links. Despite being focused in LoRa/LoRaWAN, Section XI sheds light on the topic of the maximum achievable range.

Summarizing, all the technologies share common operational characteristics with some small variations. These variations can lead to technologies that fit better to specific application scenarios.

B. Application Perspectives

All LPWAN technologies can satisfy the requirement of long range connectivity under the constraint of low power consumption; however, a particular solution may be preferred for a specific vertical, due to its inherent characteristics.

Focusing on the set of 3GPP technologies, all three solutions can be considered as general-purpose, although they tend to be the favorable options for use cases with the following characteristics: i) QoS differentiation is an essential part of the system, ii) applications are characterized as mission-critical, iii) DL messages are expected frequently or they are lengthy, and iv) the network is rolled out by an operator. Among the 3GPP technologies, NB-IoT serves the need of long range connectivity extremely well and has acceptable performance regarding power usage. As a result, NB-IoT can cover a wide range of applications: from industrial verticals to smart buildings and smart cities applications, and from health monitoring to disaster reporting. LTE-M has lower MCL than NB-IoT; however, its range capability is still high. It can be applied to the same set of applications as NB-IoT, but, due to its greater BW and the higher offered data rates, it can also serve more demanding streams, such as image or video streams [21]. Based on this, examples of applications, for which LTE-M would be an excellent fit, are the monitoring/surveillance applications for authorized access in residential buildings, offices, and industrial areas, which are based on more detailed image sensor data instead of occupancy sensor data. Nevertheless, such applications are energy costly. EC-GSM-IoT has similar capabilities to NB-IoT; however, due to the impending switch-off of several GSM networks across the globe by the end of this decade, it is questionable whether this is a future-proof choice.

Regarding non-3GPP technologies, MIOTY and Sigfox have many similarities in terms of achievable Key Performance Indicators (KPIs). They have similar frequency plan, data rate support, maximum transmission power allowed, and link budget. Both technologies can be characterized as general-purpose solutions, since they perform well in diverse application contexts characterized by short UL messages and sparse DL commands. Such applications include the following: smart cities, smart buildings, smart factories, smart agriculture, fleet and asset tracking, and the utilities sector (i.e., all types of smart metering). However, their low data rate can be a limiting factor in some scenarios. RPMA has a substantial BW and link budget, hence it can be used for a wide range of applications. Nevertheless, it might not be considered as a general-purpose solution, due to the utilization of the 2.4GHz band, which puts building penetration and outdoor-to-indoor or indoor-to-outdoor coverage in question. Weightless-P and the combination of LoRa and LoRaWAN have a lot in common, and hence their application field is similar. The six Working Groups (WGs) of the LoRa Alliance [115] reveal that the same verticals being served by MIOTY and Sigfox can also be served by Weightless-P and LoRaWAN; as a result, both are considered as general-purpose solutions. Finally, Wize applications are part of the protocol specification, and include gas metering and water management.

C. Deployment Aspects

Although the existing LPWAN technologies share many common operational characteristics, presently, four of them enjoy greater acceptance considering the number of deployments: LTE-M, NB-IoT, Sigfox, and LoRa. Based on the estimates provided by a market research conducted by IoT Analytics [116], in 2019, more than nine out of ten (92%) LPWAN systems were deployments based on these four technologies. Fig. 6 reveals how this percentage of 92% is distributed among the four dominant technologies.

As it can be observed in Fig. 6, the LoRaWAN technology is the most adopted among the four, both in terms of the number of network operators and the number of countries with established networks. This domination of LoRaWAN should be partly attributed to cost reasons: NB-IoT and LTE-M utilize the licensed spectrum and require more demanding BSs in terms of computing power and energy; Sigfox networks operate on a subscription-based scheme. Another reason could be the public availability of the LoRaWAN protocol, which contradicts the approach of many non-3GPP technologies.

Specifically, all 3GPP technologies have publicly available specifications in the form of 3GPP Releases, whereas non-3GPP technologies rely on their Standards Development Organization (SDO) for the openness of their protocol: MIOTY, RPMA, and Sigfox are proprietary (although, MIOTY is based on an open ETSI specification), whereas Weightless and Wize are open. The combination of LoRa/LoRaWAN is an
interesting one: the MAC protocol is fully open, whereas LoRa PHY is proprietary. Nevertheless, thanks to Semtech specifications (e.g., [93], [95]) and experimentation by academia and industry that includes reverse engineering (e.g., [117]), numerous aspects of the PHY have come to light. The openness of a standard appears to be one of the crucial factors for its adoption.

In Fig. 6, note that for the LoRaWAN technology the number of countries covered exceeds the number of network operators; this interesting pattern reveals the higher applicability of licence-free LPWAN technologies in cross-country deployment scenarios. It should be mentioned that the LoRaWAN numbers in Fig. 6 refer to deployments using the LoRaWAN MAC on top of the LoRa PHY, meaning that there might be even more LoRa systems deployed with other protocol stacks on top of them; this highlights the lead of LoRa among LPWAN PHY technologies, in terms of adoption.

D. Complementarity of LPWANs

To cap the discussion on LPWAN technologies comparison off, we proceed with a comparative examination of two of the most widely adopted solutions: LoRa and NB-IoT. This comparison can serve as an example on how the different LPWAN solutions may complement each other, rather than compete. By looking at Fig. 7 [16], [22], we can observe that where LoRa lacks performance, NB-IoT excels and vice versa. LoRa systems offer great coverage, are easy to deploy and cost efficient, and support EDs that can operate for several years while being powered by simple batteries. On the other hand, NB-IoT deployments can reach great distances, scale easily, perform superbly in terms of latency and QoS, and have a larger payload length. These indicators lead to the deduction that LoRa and NB-IoT are more complementary technologies than competitors.

We conclude that for each technology there are many use case scenarios in which it can perform better, but there is no one solution which is universally better that the other. Weightless Alliance has expressed this eloquently, by stating that there is no single network that can serve every IoT need, with each LPWAN technology offering different trade-offs in IoT KPIs, thus being better suited in different application scenarios. For the rest of this paper, we focus on LoRa/LoRaWAN, by using this networking stack as a complete LPWAN example. Even though Sections VIII–XII refer to LoRa/LoRaWAN, only Sections VIII–X are LoRa/LoRaWAN specific; i.e., Sections XI and XII can also be applied to research endeavors in other LPWAN solutions, with careful mapping of each technology’s characteristics onto the targeted use case.

VIII. LoRa/LoRaWAN Systems

The combination of LoRa PHY and LoRaWAN MAC is a highly efficient solution for the establishment of LPWAN systems. This is due to the key characteristics of the underlying combined technologies, i.e., the LoRa modulation and the LoRaWAN protocol. In this section, the LoRa/LoRaWAN systems are further analysed.

A. Architectural View

From the architectural perspective, the LoRa/LoRaWAN system deployment follows a star-of-stars topology, as shown in Fig. 8.

EDs are deployed in a star topology around GWs and broadcast their messages by using single-hop LoRa links, forming the network access-side of the architecture. GWs are deployed in a star topology around a central node, called Network Server (NS). The main role of the GWs is to forward UL messages to an NS and distribute DL messages to the EDs. Once the messages arrive to an NS, packet deduplication is applied and then the data are forwarded to an Application Server for application-layer processing of the payload. For the ED to NS connection, a separate authentication server can be used, called Join Server. All the entities of the server-side (Network, Application, and Join servers) can be hosted either in a single location (single point of presence) or in a distributed cloud platform.

10https://www.weightless-alliance.org/background. Accessed: Jun. 6, 2022
1) Server-Side: The NS is the core component of the architecture, containing the intelligence of the deployed network. The NS usually resides in a computer system with full computing capabilities, although the advancements in single board computers (e.g., Raspberry Pi) over the recent years have led to NS implementations in such systems. The current trend dictates the deployment of NS in a cloud infrastructure along with the rest of the server-side architecture.

The NS may be either public or private. A number of industry partners and network operators have allowed access to their public NS, with the best-known solution in this area being the NS offered by The Things Network (TTN). However, on many occasions, it is preferred to develop a private NS for managing devices and applications. Solutions by TTN, Chirpstack, Loriot, ResIoT, Actility, and others are widely used. These solutions might offer additional components of the LoRaWAN system, such as an application server, a security or authentication server, a geolocation server, and a managing platform.

The NS, as the central point of the star-of-stars topology (Fig. 8), is responsible for a number of operations, including the following: ED activation; message deduplication; direct communication with applications or with the application server (should it exists); ED update shipping via Firmware Update Over The Air (FUOTA) messages; sending MAC commands; and implementation of the Adaptive Data Rate (ADR) algorithm. ED activation can be performed either using the Over The Air Activation (OTAA) method or the Activation By Personalization (ABP) method. Both methods are explained later.

Packet deduplication is a critical operation, sincearrival of duplicate packets at the NS is common. EDs transmit their messages via broadcasting. With GWs’ deployment becoming denser, a UL message can be received by more than one GWs. GWs are designed with no intelligence by default, meaning that they act as plain packet forwarders regarding the upstream traffic. As a result, duplicate packets appear in the NS, adding to its responsibilities the identification of such packets and the handling of the identified duplicates (drop, mark, or do something else) in accordance with the deduplication policy at hand. Another cause of duplicate packets appearance is the loss of ACK packets from NS to ED, which causes re-transmissions of the same UL packet by the ED. Duplicate packets can also emerge due to the QoS policy, i.e., an NS may instruct an ED to broadcast the same frame up to 15 times, in an effort to increase its reception probability; a scheme reminiscent of Wize’s repetitions or Sigfox’s triple UL. A Received Signal Strength Indication (RSSI) deduplication policy is described in [118].

The data received by the NS must be forwarded to the applications built on top of the network infrastructure. Usually, an application server is deployed for handling such operations, which offers significant advantages in terms of i) NS performance and complexity, since the NS has to communicate only with the application server and not with each application separately, and ii) system scalability, since the addition of a new application requiring a dedicated application server will have no impact on the configuration of the NS other than establishing a new communication channel between it and the new application server.

Apart from UL traffic, LoRaWAN supports message transmission from the NS to the EDs. An example of such DL traffic is FUOTA messages [119]. FUOTA, a process running in the APP layer, allows EDs to update their firmware without the need of physical access to the device. Usually, it is supported by a firmware update server and the application server; however, the distribution of the update file to the EDs is performed by the NS. Another example of DL traffic are the MAC commands initiated by the NS for Class A and Class B EDs, as described later. These commands may ask an ED for its state or modify its transmission parameters. For Class B EDs, there are additional beacon-related commands.

Another operation performed by the NS is the ADR algorithm [120]: an effort to offer the best trade-off between data rate and range for EDs, maximize the network capacity, and minimize the necessity of UL ACK messages. When an ED sends UL packets to the GW, the GW forwards them to the NS using the Internet Protocol (IP). The GW does not act on the packets (apart from data integrity checking [118]), but it does attach some additional information, such as timestamp, RSSI value, and SNR level. Based on these metadata for all UL packets by all EDs, the NS can have a holistic view of the network, and thus decide what is the best SF configuration for each ED.

The SF selection follows a simple approach: for high (low) link budget use smaller (larger) SFs; thus, increasing (decreasing) the data rate and decreasing (increasing) the link range, the message airtime, and the energy consumption. The impact of the SF configuration in the discussed trade-off will become evident in the sequel, when the performance metrics of a LoRa link will be presented. The underlying logic of ADR is based on the position of each ED. An ED that sends messages via a high link budget connection is either positioned in close proximity of a GW or has an obstacle-free path, allowing higher data rates and better energy management, without the need of increasing the link’s range. On the other hand, distant EDs or EDs with “bad” links with a GW have to sacrifice data rate for a more robust SF configuration.

It should be noted that the ADR scheme with fixed positioned EDs is implemented by the NS (network-managed ADR); if an ED is in motion, the notion of Blind ADR is utilized [121]. In general, EDs in motion decide by themselves whether to implement ADR, since they have greater knowledge of the current link state compared to the NS. For highly mobile EDs, ADR is preferably disabled, due to the highly variable communication channel. Valuable insights in ADR are given in [122] and [123], whereas an extensive discussion on the topic and the different ADR implementations can be found in [124].

2) Network Access-Side: EDs and GWs are the components of the access part of the architecture. EDs are usually small, energy-constrained devices with low computing capabilities. These characteristics are imposed by the need to be cost-effective, since they are deployed in massive numbers. EDs can be stationary or perform operations while on the move.
Typical building blocks of an ED are: the power unit, the radio module, the memory block, the processor, and the sensor(s). Part of the radio module is also the antenna, which, in market devices, can be printed. A memory unit in the order of few MBs or even KBs is considered adequate for satisfying the needs and the constraints of an ED. The processor of such a device has rather limited capabilities compared to modern processors used in home or enterprise computing systems, but is sufficient for an ED. Finally, an ED may have multiple sensors, being able to perform several measurements, varying from temperature sensing to position awareness. Sensing is the main operation of EDs, hence the WSN characterization of the access part of IoT systems. The main modules of an ED are depicted in Fig. 9. More sophisticated views of a smart sensor’s architecture can be found in [29] and [125]. An alternative view of the ED structure is given in [26], depicting all the sub-modules and encompassing the communication of the ED with its physical environment. Recent advancements in microelectronics have led to the development of chips that can incorporate both the radio transceiver and the microprocessor for LoRa EDs.

As for the GWs, they rely on a fixed power source for their operation. They also have a radio module for LoRa communications; additionally, they may have several other network interfaces, which can be either wireless (e.g., cellular, IEEE 802.11) or wired (e.g., Ethernet). Of course, an antenna is present, which is usually significantly more powerful than the ED antennas. Computing capabilities are higher than those of the EDs. A GW can receive and decode messages by more than one EDs, and more than one GWs may receive the messages broadcast by an ED. GWs can receive simultaneously data in different frequencies and different SFs.

B. Protocol Stack View

The LoRa/LoRaWAN protocol stack fits into the main elements of the architecture, as shown in Fig. 10. Fig. 10 also depicts the layers of the Open Systems Interconnection (OSI) stack, which are used both for the IP communication between the GW and the NS and the server-side internal communications.

1) Server-Side: It is apparent that the NS participates in both data and control flows. The NS receives from a GW an APP payload encapsulated in a LoRaWAN frame, which is encapsulated in appropriate types of packets of the rest of the protocol stack (L4 to L1). The NS performs the necessary decapsulation and, depending on its type, the message either remains in the NS (control messages that are not related to authentication actions or application messages in an architecture that does not include an application server) or is further forwarded to the server-side: i) to an application server (data flow), or ii) to a join server (authentication-related control flow). The reverse order of encapsulation and decapsulation takes place in DL messages.

A key observation here is that LoRaWAN resides above the transport layer in the NS; this is as expected since GW-to-NS and vice versa communication is achieved via links that transfer APP messages encapsulated in transport layer protocol segments, IP datagrams, and data link frames. A complete definition of the system’s backend is given in [126].

2) Network Access-Side: The stack of an ED is simpler than that of the NS. In the case of a data flow, an APP message is encapsulated in a LoRaWAN MAC frame, which is then encapsulated in a LoRa PHY packet to be broadcast via the LoRa radio interface. The broadcast LoRa packets are received by a GW, which performs decapsulation to extract the APP payload. The extracted payload is placed in a typical transport layer segment, then it is encapsulated in an IP datagram, and finally, it is encapsulated in the desired data link frame (e.g., cellular, IEEE 802.11, Ethernet) to be transmitted to the NS. The reverse operation takes place in the DL. For control flow information, the APP-related encapsulation and decapsulation may be skipped.

Fig. 11 shows the protocol stack of an ED [97], which consists of four layers: the APP, the LoRaWAN, the Regional Parameters, and the PHY; practically, this stack can be narrowed down to three layers since the Regional Parameters (Table II) simply define which PHY can be supported in which bands. In the PHY, the LoRa modulation is the main choice; other options include the FSK modulation [94] and LR-FHSS [127]. On top of the PHY, the LoRaWAN MAC specifies three device classes, which define the available DL windows and ultimately determine the power consumption scheme of the ED.

The full frequency plan documentation can be found in [94]. In Europe, no plan for the use of the EU433 band exists. On the other hand, EU868 is used extensively for LoRa transmissions. In all regions other than US915 and AU915, sixteen communication channels may be used with three of them being mandatory to implement, since they can be used for ED join-request messages. Table III depicts the mandatory channels of EU868 for UL transmission.

### Table II: LoRaWAN Frequency Planning Based on Deployment Region

| Region                  | Frequency Band (MHz) | Abbreviation |
|-------------------------|----------------------|--------------|
| some Asian countries    | 923                  | AS923        |
| Australia               | 915 - 928            | AU915        |
| China                   | 779 - 787, 470 - 510 | CN779, CN470 |
| Europe & some African countries | 863 - 870, 433 | EU868, EU433 |
| India                   | 865 - 867            | IN865        |
| North America           | 902 - 928            | US915        |
| Russia                  | 864 - 870            | RU915        |
| South Korea             | 920 - 923            | KR920        |

---

11 https://www.thethingsnetwork.org/article/breaking-down-stm32wl-the-central-element-of-generic-node-sensor-edition-and-first-dual-core-lorawan-system-on-chip-soc?mc_cid=57d7f74b8f&mc_eid=8698751288, Accessed: Jun. 6, 2022.
TABLE III

EU868 MANDATORY CHANNELS FOR UL TRANSMISSIONS

| Channel | Modulation | BW (kHz) | Center Frequency (MHz) | SF |
|---------|------------|----------|------------------------|----|
| 1       | LoRa       | 125      | 868.1                  | 7 - 12 |
| 2       | LoRa       | 125      | 868.3                  | 7 - 12 |
| 3       | LoRa       | 250      |                        | 7   |
| 4       | LoRa       | 125      | 868.5                  | 7 - 12 |

Regarding DL transmissions, the same channels as in UL transmissions are available; however, the default DL channel is set at 869.525MHz (LoRa, SF9BW125), which permits transmission for GWs operating with a TP as high as 27dBm (0.5W).

On top of the LoRaWAN MAC stands the APP layer. The LoRa Alliance has established six WGs that are responsible for the further development of the different vertical sectors; these include all the familiar “smart” applications, such as agriculture, buildings, cities, industry, logistics, and utilities [115]. As discussed in Section V, IoT-related APP protocols can be utilized for data acquisition, whereas LoRa Alliance promotes the standardization efforts for specific verticals, e.g., the Open Metering System (OMS) over LoRaWAN for smart metering applications [128].

IX. OVERVIEW OF LORAWAN MAC

LoRaWAN is an open protocol being developed by LoRa Alliance. The protocol is officially defined as:

“LoRaWAN® specification is an LPWA networking protocol designed to wirelessly connect battery operated ‘things’ to the internet in regional, national or global networks.”

LoRa Alliance is an association that was established in 2015, having Semtech as a founding member and being supported by numerous industry players. More on LoRa Alliance can be found in [129].

A. LoRaWAN Device Classes

The main characteristic of the LoRaWAN protocol is the establishment of three device classes; namely Class A, Class B, and Class C. These three device classes define the ED’s transmission scheme.

Class A stands for All, meaning that all EDs have to be able to implement at least this class to perform LoRaWAN transmissions. Class A defines two open DL windows with a maximum of 15s and 16s delay after a UL transmission, respectively (typical values are 1s and 2s or, in case a join-request UL message has been preceded, 5s and 6s). During these windows, an ED may receive data or control messages from the NS. Regarding the UL, Class A EDs can transmit either a data message, a MAC command, an answer to a previous NS-initiated MAC command, or a combination of them. The transmission scheme can be seen in Fig. 12.
Class B stands for Beacon, meaning that EDs open additional receive windows, based on synchronization beacon frames that are sent either from the NS to an ED via the GW or directly from a GW to an ED. Differentiating from Class A, messages received in Class B’s supplementary DL slots can be only data-related. However, additional Class B-specific MAC commands can be sent by the NS to an ED during Class A’s receive windows. For UL, again, there is a set of MAC commands available either as an ED-originated request or as an answer to an NS-originated MAC request. Class A to Class B switching decision originates from EDs’ APP. The transmission scheme can be seen in Fig. 13.

Class C stands for Continuous, meaning that an ED implementing the Class C approach is constantly waiting for DL transmissions between its UL transmissions. Class C adds no extra MAC commands, but any of Class A MAC commands have to be transmitted in the related available windows and not in the continuous reception window. Should an ambiguity occur with a Class A’s receive window and a Class C’s continuous reception in a Class C-enabled ED, Class A’s DL messages are prioritized. If a Class A successful reception takes place in the first window, no second window is opened, as mentioned earlier, meaning that the continuous DL window will take the place of the second window. The transmission scheme can be seen in Fig. 14.

Naturally, Class A is the best option power-wise, but forces the NS to wait for the next UL transmission if it is not served by the two open windows. On the other hand, Class C reduces latency for NS-to-ED communication and practically communication ceases being UL-initiated (except for the first transmission, which must be a UL), but the ED has to be ready to receive DL messages all the time, resulting in greater energy consumption. Class B stands in the middle point of Class A and Class C.

### B. LoRaWAN Frames & ED Activation

Using the functionality offered by the lower LoRa PHY, LoRaWAN MAC formats its frames to support message exchange between the EDs and the GW. The format of the LoRa PHY packet is discussed on the corresponding section. The format of the LoRaWAN MAC frame can be seen in Fig. 15.

The frame has three main fields. The MAC Header (MHDR) field is the header field for MAC, the payload of MAC resides inside the MAC Payload field, and the Message Integrity Code (MIC) field assists in message integrity checking. In the latter field, the MIC is calculated using the fields of the frame and the AES CMAC algorithm. Considering the MAC Payload field, it may either include the payload of a message, Join-Request/Join-Accept messages for EDs trying to join the network, or other MAC commands. In Fig. 16, we provide a deeper view of the packet encapsulation process in LoRaWAN, for the data flow.  

An extensive analysis of each field presented in Fig. 16 is not in the scope of this paper. The interested reader can find more in official documentation [97]. Here, we focus on the 3-bit Frame Type (FType) field, in an effort to present the different types of frames supported by the protocol. In LoRaWAN, eight types of frames are supported, as shown in Table IV.

Frames with an FType value of 010 and 011 represent UL and DL traffic with no ACK, respectively. An FType value of 100 and 101 is used for UL and DL messages that require ACK. 111 value is proprietary, in a sense that it can be used to describe messages that do not follow the standard format. Frames 000, 001, and 110 are used for ED’s activation via the OTAA process.

In LoRaWAN, two activation processes are distinguished: OTAA and ABP. In OTAA, the ED has to first be “personalized”, meaning that a device Identifier (ID), called End Device Identifier (DevEUI), a join ID, called Join Server Identifier (JoinEUI), and an application key, called Application Key (AppKey), have to be defined. With these values set, the ED can send a Join-Request message with DevEUI and JoinEUI values.

12In DL messages, the field of Payload CRC (PL CRC) does not exist.
13In Specification version 1.0.x, 110 value is described as RFU, meaning Reserved for Future Use. In Specification version 1.1, 110 value is used for Rejoin-Request messages by the ED.
14To avoid confusion, in older versions of the protocol, JoinEUI is mentioned as Application Identifier (AppEUI). As of v1.0.4 and v1.1, the JoinEUI nomenclature is adopted.
Fig. 16. Packet encapsulation in LoRaWAN for the data flow. Control flow may correspond to join-related processes (orange coloring in Join Server in Fig. 10, instead of green coloring) or network commands (red coloring in Network Server in Fig. 10, instead of green coloring).

Fig. 17. OTAA sequence diagram (ED: End Device, GW: Gateway, NS: Network Server, JS: Join Server, AS: Application Server).

The required keys and the DevAddr. This method might seem simpler than OTAA, but has two serious drawbacks. Firstly, the ED is associated with the specific network through which was activated. Secondly, the keys are derived by the user, meaning that a careless selection may lead to risks in the security of the ED or even of the entire WSN.16

After successful activation of the device with either of these two methods, DevAddr, NwkSKey, and AppSKey are set to the ED.

C. Security in LoRaWAN

In LoRaWAN, security is provided via the use of two of the keys described in the previous subsection; namely NwkSKey and AppSKey. NwkSKey stands for Network Session Key, although it refers to security aspects in LoRaWAN MAC. NwkSKey is used by the ED and the NS for the encryption and decryption processes of the LoRaWAN MAC payload in case of control messages, as well as for ensuring integrity (MIC field in frame format). On the other hand, AppSKey, an abbreviation of Application Session Key, is used for the encryption and decryption processes of the MAC payload of data messages, i.e., the fields of the APP message.

It can be inferred that encryption is guaranteed end-to-end, but that is not the case with integrity. Encryption is vital for network’s sustainability, due to the broadcasting nature of UL transmissions that can be captured by any GW in proximity; even GWs that are not part of the system. On the other hand, the NS is considered as a trusted entity in general; however, a user may implement additional confidentiality and integrity mechanisms for the applications.

Despite the built-in security of the protocol, numerous open issues need to be addressed for both distant and on-site attacks. Regarding the activation procedure, even though OTAA is a more secure activation process than ABP, malevolent users can utilize common methods, such as replay attacks (in case of improper DevNonce generation) and jamming,17 to undermine

16Due to the importance of that risk, [97] emphasizes the necessity of proper key derivation.

17Jamming is typically considered as a PHY attack; however: i) there exist data link layer implementations of it, and ii) it can be combined with other techniques to perform attacks in higher layers of the stack.
the join procedure [130]. ABP itself includes a difficult key generation, especially in deployments with plenty of EDs. The techniques of replay attack and jamming can also be combined to perform known attacks, such as a wormhole attack [131], for either a control or a data flow message.

Apart from remote software attacks, malevolent users can exploit EDs by accessing them physically, since EDs are deployed and left to operate on the field. This is a common consideration for IoT systems, especially for those that are deployed in public places, e.g., smart cities, and cover large areas, e.g., LPWANs. Physical attacks appear to have the greatest likelihood to occur in LoRaWAN deployments [132], and, as a result, system operators must handle different types of breaches by utilizing mitigation techniques. By gaining physical access to an ED, an attacker can damage the ED, obtain the session keys to perform an active attack, or sniff UL or DL messages for a passive exploitation. Countermeasures to unauthorized access include physically impenetrable EDs’ cases, alert mechanisms, secure communication among the modules of an ED (e.g., secure communication between the Central Processing Unit (CPU) and the radio module), and ED’s self-destruct mechanisms once its security has been breached, which will down the violated ED but ensure the sustainability of the rest of the system.

A reference on the topic of LoRaWAN security can be found in [132]. Even broader discussions encompassing the LPWAN field can be found in [86], the entirety of WSN field in [133], and the entirety of IoT field in [41], [134], and [135].

TABLE V

| Chipset | Feature | Frequency (MHz) | Maximum Transmission Power (dBm) | Rx Sensitivity (dBm) | Maximum Link Budget (dB) | Supported Modulation | Bit Rate (kb/s) |
|---------|---------|-----------------|----------------------------------|---------------------|-------------------------|---------------------|----------------|
| LLCC68  |         | 150 - 960       | 22                               | -129                | 151                     | FSK, GPSK, MSK, GMSK, LoRa | Programmable bit rate from 1/76 to 26.5 LoRa and 300 FSK |
| SX1261/62/68 | 150 - 960 | 22 | -148 | 170 | Programmable bit rate up to 26.5 LoRa and 300 FSK |
| SX1272/73 | 860 - 1000 | 20 | -137 | 157 | Programmable bit rate up to 300 |
| SX1276/77/78/79 | 137 - 1020 | 20 | -108 | 168 | Programmable bit rate up to 300 |
| SX1280/81 | 2400 | 12.5 | -132 | 144.5 | LoRa, FLRC, (G)FSK | Programmable |

X. OVERVIEW OF LoRA PHY

In contrast to LoRaWAN, LoRa is a proprietary PHY modulation being developed by Semtech. Its definition follows:

“LoRa® is a modulation technique, suitable for interconnecting things with energy constraints in long range wireless environments, with low needs of data rate.”

LoRa development started from Cycleo (acquired by Semtech in 2012), and eventually the wireless communication method [136] and both the modulation [137] and demodulation [138] processes were patented by Seller and Sornin with Semtech assignee. LoRa is based on the CSS patent [139] by Sforza with Nanoscale Labs assignee (Semtech is the current assignee). A brief history of LoRa can be found in [140].

A. Implementation of LoRa in Chipsets

Semtech is responsible for the development of chipsets that are capable of interconnecting EDs with GWs via an established LoRa link. Numerous chipsets are in use today inside the EDs; families of SX1261/62/68, SX1272/73, and SX1276/77/78/79 are considered the most common, LLCC68 is the preferable option in smart home applications due to lower maximum link budget and worse receiver sensitivity, and SX1280/81 is a chipset that can offer LoRa connectivity in the 2.4GHz band. These chipsets can be found inside LoRa modules built by other hardware manufacturers, such as Microchip Technology (RN2483 is one of the most widely used modules and is based on SX1276/77/78/79 family), Murata Manufacturing (its CMWX1ZZABZ-078/-093 module uses SX1276 chipset), STMicroelectronics, HopeRF, Gemtek, AcSiP, and others. A comparison among Semtech’s chipsets can be found in Table V. Values of Table V are nominal, as provided by Semtech.

Regarding the GWs, numerous chipsets are also available. SX1250 which works along with SX1302, SX1255/57/58 (frequency: 400MHz-510MHz / 860MHz-1000MHz / 779MHz-787MHz), SX1301, SX1302, and SX1308 are common options. Major companies develop their own GWs based on these chipsets, such as Cisco (GW based on SX1301), Kerlink, Multitech, and others.

B. Performance Metrics of a LoRa Link

The main metrics regarding a LoRa link are the maximum achievable bit rate and the maximum supported range. The bit rate can be calculated by considering three of the key performance factors (BW, SF, CR), while the range is depicted in receiver’s sensitivity or link budget equations and is related to the fourth factor (TP). For the interested reader, Appendix A provides a further mathematical definition of BW, SF, chip, and chirp in the context of LoRa.

The bit rate of a LoRa link is given by the equation:

$$R_b = SF \times \frac{BW}{2^{SF}} \times \text{RateCode}$$  (1)

Because $R_b$ is proportional to $SF$ linearly and inversely proportional to it in an exponential way, an increase of $SF$ will have a negative impact on bit rate. On the other hand, regarding
We can also calculate the link budget with respect to the fading margin in dB (practically referring to all the phenomena affecting gains), SystemLosses ends in dB (practically loss due to connection of antennas to radio modules via, e.g., wiring), ChannelLosses in dB (practically referring to all the phenomena affecting wireless signal transmission, i.e., reflection, refraction, diffraction, scattering, and absorption), and $M$ is the fading margin in dB. We can also calculate the link budget with respect to the transmitted power and the receiver’s sensitivity, as follows:

$$P_{Rx(dBm)} = P_{Tx(dBm)} - S(dBm) \quad (5)$$

Fig. 18 provides a visualization of the inversely proportional relationship between $SF$ and bit rate and the proportional relationship between $SF$ and range. Naturally, an increase of $SF$ results in exponential decrease of supported bit rate and linear increase of link’s range.

Ending this quantitative section, we examine the regulations of ISM transmission by exploring the concepts of ToA and duty cycle. LoRa utilizes the ISM bands in a region-based scheme, although chipsets that support globally free bands, e.g., 2.4GHz, exist. Due to the usage of the ISM bands, LoRa links have to comply with a number of regulations. Specifically, they have to:

- change transmission channel in a pseudorandom fashion in every transmission, and
- respect the maximum duty cycle (or dwell time) and the maximum transmitted power (Effective Isotropic Radiated Power (EIRP) and Effective Radiated Power (ERP)).

Compliance to these regulations is necessary, since a plethora of technologies reside in ISM bands. Focusing on the European region, the duty cycle imposes the maximum time allowed to transmit, with common maximum values being 0.1%, 1%, and 10%. In the North American region, dwell time is the regulated value. The available transmission time can be calculated as follows.

The structure of a LoRa PHY packet can be viewed in Fig. 19.

A Preamble being transmitted in the form of a number of up-chirps followed by two and a quarter down-chirps indicates the start of the packet. After that, the Header and CRC fields are included if the explicit mode is used. The size of these fields depends on CR, and a typical rate code of $\frac{4}{8}$ with an 8Bytes Payload plus Payload CRC fields size, would result in a combined size of Header and CRC of 2Bytes. In the example of Fig. 19, we observe that the size of Header and CRC fields is half the size of Payload and Payload CRC fields, thus indicating a CR of $\frac{4}{8}$. The Payload field includes the payload information that the ED broadcasts, assuming a UL transmission.

Based on Fig. 19, the transmission time of a PHY packet can be split into the time needed to transmit the preamble and the time needed to transmit the payload (i.e., the header, the CRC, the payload, and the payload CRC). As a result, the ToA

$$P_{Rx(dBm)} = P_{Tx(dBm)} + SystemGains(dB) - SystemLosses(dB) - ChannelLosses(dB) - M(dB) \quad (4)$$

Fig. 19. LoRa PHY packet format [95].
will be:

\[ T_{\text{payload}}(s) = T_{\text{preamble}}(s) + T_{\text{payload}}(s) \]  

The \( T_{\text{preamble}} \) is given by equation (7):

\[ T_{\text{preamble}}(s) = (n_{\text{ preamble}} + 4.25) \times T_s(s) \]  

where \( n_{\text{ preamble}} \) is fixed for EU868 band in 8 symbols and \( T_s \) is the transmission time for one symbol in s, as expressed in equation (8) and derived in Appendix A (equation (A4)):

\[ T_s(s) = \frac{2^SF}{BW(\text{Hz})} \]  

On the other hand, the \( T_{\text{payload}} \) is given by equation (9), shown at the bottom of the page, where \( T_s \) is the transmission time for one symbol in s, \( BW \) is the synchronization word and has a fixed value of 8, \( PL \) is the number of payload bytes, \( SF \) is the spreading factor, \( CRC \) is a one-bit flag that can be set at 1 if the used protocol above is LoRaWAN, \( H \) is the header type as discussed in Fig. 19 (0: explicit, 1: implicit), \( DE \) is a one-bit flag that notifies for low data rate optimization when SF11 or SF12 is used (0: high data rate, 1: low data rate), and \( CR \) is the coding rate.\(^{18}\) Finally, equation (10), shown at the bottom of the page, provides the total \( ToA \) in s, taking into consideration equations 6-9.

Having regulated the percentage of the time that an ED can transmit, a pattern emerges for the UL transmissions. An ED can transmit, then it has to wait for at least a minimum amount of time during which it might receive DL messages, and after that it can re-transmit. We refer to this waiting time as \( T_{\text{interval}} \). By examining the relationship between \( ToA \) and \( T_{\text{interval}} \), while taking into account the regulation about duty cycle, we have:

\[ T_{\text{interval}}(s) = \frac{ToA(s)}{DutyCycle} - ToA(s) \]  

This results in \( T_{\text{interval}} \) being orders of magnitude greater than \( ToA \).

XI. RESEARCH ON LoRA/LoRAWAN

There are various open research topics related to the implementation, usage, and optimization of the LoRa/LoRaWAN networks (\cite{18}, \cite{25}, and \cite{27}). The major ones are listed below:

- Network performance validation and evaluation metrics (e.g., \cite{11})
- Range evaluation and link-level behavior (e.g., \cite{13})

\(^{18}\)In calculations, the \( CR + 4 \) quantity refers to the denominator of rate code, and its total value should be 5 \( \cdots \) 8; hence, \( CR \) takes values 1 \( \cdots \) 4.

- Power consumption, energy management, and lifetime of the network (e.g., \cite{81})
- Multiple Access (e.g., \cite{99})
- ADR algorithm (e.g., \cite{122})
- ED mobility support and roaming (e.g., \cite{123})
- Security in PHY, MAC, and Network layer (NET) (e.g., \cite{130})
- Optimal GW placement and localization challenges (e.g., \cite{141})
- Data-based decision making, by incorporating ML (e.g., \cite{142})
- Meeting recent advancements from Information Technology (IT) and Telco domains, e.g., Software-Defined Networking (SDN), Blockchain, and network slicing (e.g., \cite{143})

The research work on addressing challenges related to the above-listed topics includes approaches for all the well-known steps in research methodology \cite{19}, \cite{144}; namely: mathematical analysis, simulation study, and real-life system implementation.

A set of surveys has also been published examining IoT, WSN, LPWAN, and LoRa/LoRaWAN from various perspectives, including open research challenges of those networks. This set of publications is summarized in Table VI. Indicative research contributions that target specific problems and challenges are studied in the following subsections.

A. Analytical Approaches

The first step towards the study of LoRa/LoRaWAN is to evaluate the performance and the behavior of the network using mathematical tools. Such tools include but are not limited to methods such as: calculation of performance metrics in probabilistic manner and comparison with well-known results (e.g., performance of LoRaWAN throughput in comparison with protocols, such as ALOHA, Carrier Sense Multiple Access (CSMA), and variations of them), modeling of the system with stochastic tools under various scenarios and assumptions (e.g., Poisson process, Markov process, stochastic geometry), and others.

In \cite{101}, the authors provide an analysis for the probability of successful UL transmissions when taking into account co-SF interference, inter-SF interference, or both of them. The imperfect orthogonality of SF is proven to affect UL LoRa links. In \cite{123}, a statistical approach to the effect of SF selection and the different levels of mobility in ADR is presented. The authors gather UL data based on a real testbed and with these data they perform a Kruskal-Wallis non-parametric test. Their analysis exposes a high correlation between SF selection dictated by Blind ADR and performance, especially in lower
TABLE VI
LIST OF REFERENCED SURVEYS

| Target of the Survey and Tutorial | References |
|-----------------------------------|------------|
| IoT                               | [3] (full overview), [9] (IoT full overview), [79] (full overview with focus on specific vertical), [41] (security aspects), [145] (Blockchain applicability), [146] (Blockchain applicability), [54] (DT full overview), [60] (DT full overview), [61] (IoT edge computing), [62] (IoT security aspects), [134] (security aspects), [135] (security aspects) |
| WSN/WSAN                         | [29] (full overview), [133] (security aspects), [144] (evaluation methodology), [147] (ML applicability in WSN challenges) |
| LPWAN                            | [12] (LoRa and Sigfox comparison), [15] (design considerations, comparison, research challenges), [16] (LoRa and NB-IoT comparison), [17] (overview, comparison, research challenges), [21] (full overview), [22] (Sigfox, LoRaWAN, and NB-IoT comparison), [23] (LoRaWAN, DASH, and NB-IoT comparison and mobility support), [26] (ED longevity), [86] (ISM LPWAN security aspects) |
| LoRa/LoRaWAN                     | [11] (full overview), [13] (full overview), [14] (full overview), [18] (LoRaWAN full overview), [19] (LoRa research challenges), [20] (full overview), [24] (LoRaWAN full overview), [25] (LoRa research challenges), [27] (full overview), [124] (ADR algorithm), [127] (LR-PHSS), [132] (LoRaWAN v1.1 security aspects), [148] (LoRaWAN simulators) |

mobility experiments (p-value is always at least 2 significant digits between increment pairs of SFs).

The modeling of interference in LoRa networks either in the form of inter-SF or in the form of intra-SF is of great interest. For the latter case, the authors in [149] derive an interference model based on their proposal for a coherent LoRa transceiver. They calculate the necessary expressions for Symbol Error Rate (SER) and Frame Error Rate (FER), by utilizing the Discrete Fourier Transform (DFT). Due to heavy computational complexity, they also use Q-functions for approximation purposes and later they compare their theoretical model with simulation results for the typical non-coherent transceiver and their proposed coherent one. In [150], the performance of CSS is examined in regard to SF orthogonality, via the derivation of closed form expressions. In [151], the authors give insights in LoRa modulation, examining it from a quantitative signal processing perspective, and he also proposes an optimal receiver based on Fast Fourier Transform (FFT). Results are drawn by comparing the optimal receiver with the typical FSK modulation, in terms of BER vs. SNR.

In [152], the authors propose their own MAC solution in order to alleviate the low Packet Delivery Ratio (PDR) problem of pure ALOHA-based approach in LoRaWAN. They name their protocol Lightweight Carrier Sensing (LCS) and compare it with pure ALOHA in a threefold way: mathematical approach, simulation, and small scale implementation. LCS has two main features: i) a hardware-level “hearing” mechanism, called Channel Activity Detection (CAD), and ii) a packet drop mechanism when the channel is sensed busy, instead of a random backoff. For the initial evaluation part (i.e., the mathematical analysis), the authors derive a probability model for their mechanism’s PDR and compare it with the known PDR of pure ALOHA. The results showcase a considerable improvement in PDR for the proposed mechanism over pure ALOHA, especially for large numbers of EDs. Another MAC suggestion can be found in [106], in which the alternative of slotted ALOHA is discussed. The proposed protocol is based on FM-RDS broadcasting for time dissemination, with the authors exploring the applicability of it in all three possible research steps. For the mathematical approach, they derive closed form expressions for Type-A and Type-B uncertainty of the transmission time of packets, which is caused due to a packet moving across the levels of the entire protocol stack inside an ED.

In [153], the authors present an analytical approach based on stochastic geometry to evaluate LoRa networks’ scalability. Their efforts target the derivation of closed form expressions for outage probability regarding two events: i) the event of a packet having SNR below the required threshold in the receiver, and ii) the event of less than 6dB difference between the interested signal and the interferers (practically capture effect). For these two events, the probabilities are calculated and the results are further validated with Monte Carlo simulations. In [154], the OTAA mechanism of Class A EDs in LoRaWAN is evaluated. The authors formulate the problem as a discrete time Markov Chain (MC) and derive the transition probabilities for all the possible transitions between states, in an effort to estimate the energy and time that an ED needs during activation.

Reference [155] examines the topic of energy harvesting in the more general context of WSNs, providing analytical expressions for power consumption and throughput of the system when realistic assumptions are taken, such as battery inefficiencies. To provide these analytical expressions, the authors formulate the problem as a Markov Decision Process (MDP). In [156], the same problem of energy harvesting is discussed, but this time specifically for LoRaWAN systems. An MC model is presented and results are drawn regarding PDR for Class A EDs. These results are later compared to simulation experiments. Energy issues are also discussed in [157], in which the authors derive analytical expressions for energy consumption, ED lifetime, and energy cost for different types of UL transmissions, and propose their own energy models. Their analysis is based on measurements from an experimental testbed, as described later. The same approach can be found in [125]; the authors derive analytical expressions for energy consumption. These expressions are validated through simulations and the results are discussed further in the corresponding simulation studies subsection that follows.

A resource allocation framework based on a Network Slicing scheme orchestrated by a Deep Federated Reinforcement Learning (DFRL) system architecture is presented in [142]. The authors focus their efforts on providing different network slices for IIoT use cases. The slices are allocated thanks to GWs acting as agents, and TP and SF acting as environmental parameters. QoS is the reward that needs to be maximized, according to the typical agent–environment interaction in reinforcement learning problems [158]. The QoS in this work is interpreted as slice.
delay, energy consumption, throughput, and Packet Loss Rate (PLR).

Network slicing for resource allocation purposes is also considered in [143]. The authors distinguish application scenarios in three slices according to their QoS needs and formulate the problem of network slicing for each ED with a game theoretic approach. The slicing is performed in a distributed way at the GW level. Three games are designed: the first one between the GWs for cell edge EDs assignment, the second one for physical channel reservation for each slice in each GW, and the final for channel assignment in each ED of each slice for each GW. The first and third games are designed as matching theory problems (one-to-many and one-to-one, respectively), whereas the second one is formulated as a bankruptcy game. The model is further validated via simulations as explained in the next subsection.

B. Simulation Studies

Simulation-based research dominates the study of WSNs and consequently of LPWANs, due to simulators’ capability to support experimentation with crowded networks. To this end, numerous simulation environments can be utilized, with main considerations being the ones presented in a comprehensive overview in Section XII, later. This subsection presents indicative research work conducted with the aid of LoRa/LoRaWAN dedicated simulators, as well as other, more generic network simulation tools, such as OMNeT++ and ns-3.

In [11], the authors develop a simulator to compare LoRaWAN capacity with pure ALOHA, with the results depicting an equivalent performance of the two protocols. They perform further tests, that are based on a testbed as described in the corresponding section that follows. In [14], the authors conduct a simulation study having Packet Error Rate (PER) and PLR as targeted metrics and examining the impact of these two in the capacity of a LoRaWAN system with respect to the load of all EDs. Their findings reveal that LoRaWAN performs satisfactorily when the load is low, but, as load increases, both PER and PLR deteriorate drastically due to the ALOHA nature of the protocol. In [81], a comparative study is provided for several LPWAN solutions, with LoRa being among them. The study includes simulations conducted in MATLAB with realistic assumptions, in an effort to estimate the performance of all of these solutions in terms of coverage and energy consumption and, consequently, the impact on EDs’ lifetime.

In [99], the authors quantify the throughput of pure ALOHA, which is supposed to be the multiple access mechanism adopted by LoRaWAN. They also propose a slotted ALOHA mechanism based on a synchronization procedure (built on top of the LoRaWAN protocol), in an effort to improve throughput. To validate their results, they conduct MATLAB simulations and real testbed experiments, as described later. In [100], the authors develop the two endpoints of a LoRa transmission (a modulator and a demodulator) in MATLAB, in an attempt to estimate the impact of SF’s imperfect orthogonality on BER and PER with respect to Signal to Interference Ratio (SIR), for various combinations of SF for transmissions and interferers.

The analytical results provided in [101] are further validated with simulations that extract the system throughput under different conditions. Simulation details are given, but no further elaboration on the simulation environment is available. In [125], the authors propose an energy model for LoRaWAN EDs and validate their model via simulations. Again, simulation details regarding the environment are not given; however, the parameters of the tested scenarios are provided. Results follow intuition regarding the effects of SF and CR selection on the energy consumption, the relationship between range and energy consumption, and the most demanding tasks energy-wise for an ED. Regarding multiple access, a proposal can be found in [103], in which a variant of CSMA/Collision Avoidance (CSMA/CA) is examined as an option, instead of the pure ALOHA that LoRaWAN mainly utilizes. Experiments in crowded OMNeT++ simulations are performed, which reveal a sizeable improvement in successful packet reception and, therefore, in system throughput.

In [96], a discrete event simulator based on Python is described, called LoRaSim. The simulator is used to obtain results that are related to the scalability potential of LoRa and, also, the Data Extraction Rate (DER); a metric that compares the total number of received messages to the total number of sent messages. The discrete event simulator of LoRaSim is used and modified extensively, with one such example being [104], in which a new MAC approach is examined, called MAC on Time (MoT). The authors compare their proposed protocol, a combination of TDMA and contention-based schemes, with the ALOHA-based LoRaWAN protocol. The results reveal a superiority of MoT both in throughput and in latency. Moreover, in [159], the same simulator is used for measuring DER and interference scaling, as a result of the use of directional antennas on EDs and the increase of GWs. In [160], the LoRaSim simulator is extended with MAC considerations, as described in the corresponding subsection of the simulator (Section XII-F). Results include evaluation of successful transmissions on goodput and energy consumption; whereas insights on the impact of ACK packets in LoRaWAN systems’ scaling are also given.

The same extended simulator is used to evaluate the ADR algorithm, as described in LoRaWAN 1.1. The study in [122] reveals a slow convergence of the algorithm, leading to the deterioration of the energy consumption and the PDR. A further extension, called LoRaEnergySim, is provided in [161], in which realistic assumptions are taken regarding the energy consumption, and results are drawn for both energy and DER. The same simulator (LoRaEnergySim) with the necessary modifications [162] is used to evaluate the proposed slotted ALOHA MAC described in [106]. Results depict a clear improvement in both PDR and energy efficiency, especially when high SFs are used and/or the payload length is long. For the combination of SF7 and small packets, pure ALOHA seems the preferable option, at least energy-wise.

LoRa-MAB is another SimPy-based discrete event simulator. The simulator examines the energy consumption per node
per successful sent packet and the successful packet transmissions for a proposed algorithm discussed in [163]. The algorithm focuses on an intelligent distributed resource allocation mechanism that is designed via utilization of Exponential-weight for Exploration and Exploitation (EXP3) algorithm to solve the Multi-Armed Bandit (MAB) problem. LoRa-FREE completes the presentation of SimPy-based simulators and is described in [164], including a mechanism of transmitting buffered data in scheduled bulk UL, contrary to transmitting data at the time of generation. Key metrics include Data Delivery Ratio (DDR) - a metric similar to PDR- and energy, but also collision-related metrics and the study of ACK-related transmissions. The same simulator is used to evaluate Time-slotted LoRaWAN (TS-LoRa) protocol, as discussed in [105].

In [165], the authors present their LoRaWAN simulator, as implemented in MATLAB. After validating the accuracy of their simulator against results of analytical models, they perform several experiments to examine the relation among UL delivery rate (practically PDR), energy consumption, range, and EDs’ number. In all occasions, an increment in range or EDs’ number deteriorates the PDR and the energy consumption. Similarly, a stricter duty cycle lowers the DL response rate; however, it benefits the energy management. The parameters of the simulations are well-provided by the authors.

Another simulator development contribution is described in [166], in which a simulator using the Java programming language is implemented. With this simulator, the authors perform numerous scalability analysis tests by calculating the PDR with respect to the number of EDs. Their study reveals an unfavorable effect of SF increase, inter-arrival time decrease, and (in the case of high numbers of EDs) payload size increase. CR seems to have no impact on PDR, whereas GWs with multiple channels tend to provide higher PDR, as expected. The capture effect that is present in LoRa transmissions improves the PDR of LoRaWAN when compared to the pure ALOHA protocol. QoS requirements are also examined in terms of required PDR for vastly populated LoRa networks. More work with the same simulator includes the further evaluation of the LCS protocol proposed in [152]. The authors perform comparative experiments between LCS and pure ALOHA, in terms of energy consumption (based on their proposed energy models) and supported EDs; LCS appears to be superior in both metrics.

OMNeT++ is one of the de facto simulation environments, being used extensively in academic research. The framework supports experimentation with a vast type of networks, with FLoRa framework supporting LoRaWAN simulations. In [167], the authors elaborate on the development of the framework and showcase the potential of the implemented simulator with ADR-related simulations. Three types of experiments are distinguished with ideal path loss, moderate variability of path loss, and highly variable path loss for urban and sub-urban environments. ADR clearly improves the performance of the network in terms of PDR and energy consumption when path loss is not taken into consideration. However, with path loss present, the results depend on the environment and the level of path loss variability. To this end, an improved ADR+ algorithm is proposed, which excels both no-ADR implementation and ADR implementation, in the two aforementioned metrics.

ns-3 is one of the most widely used simulators in networking research, with several references regarding the topic of LoRa/LoRaWAN connectivity. In [98], the required modules for simulating such networks in ns-3 are presented and simulations are developed in order to assess throughput and coverage. References [168] and [169] further discuss these modules and provide results regarding PDR. One more research team has extended these ns-3 modules for simulations focusing on energy consumption [170]. In [171], another effort on developing the necessary modules for LoRa/LoRaWAN simulations in ns-3 is described, providing insights in PDR. A different research group discusses the development of necessary ns-3 modules for LoRa connectivity and examines energy consumption, PDR, and the feasibility of the CSMA protocol on top of LoRa, when taking into account these two metrics (i.e., energy consumption and PDR) plus the collision rate [172]. Another research topic is the optimization of GW placement. The authors in [141] utilize the Gap statistics, the K-means clustering, and the Fuzzy C-means methods to derive their own model; namely DPLACE. The model is tested and compared with other GW placement methods in ns-3 simulations, showing promising results in terms of Capital Expenditure (CapEx), Operational Expenditure (OpEx), number of required GWs, and PDR.

Publication [149] has already been mentioned in the mathematical approach subsection; however, the authors go a step further by putting their proposed interference model for coherent LoRa transmitter to the test, via the Monte Carlo method. The approximation from Q-function, and the coherent and non-coherent transceivers are examined, with a twofold result: i) the approximations are proven to be extremely close to the simulation results (on many occasions identical), and ii) terms of SER and FER, the coherent transceiver outperforms the non-coherent one. Metrics for SNR and SIR are also provided. Same validation approach for the mathematical analysis of a problem can be found in [153]. The authors confirm their results via Monte Carlo simulations, with the outcomes of their experiments being almost identical to the ones of the numerically calculated outage probabilities.

The mathematical analysis of [154] is also tested via simulations. The proposed MC for the OTAA procedure in LoRaWAN is tested via Scilab simulations with the parameters of the simulations being well-described. Results reveal that high channel quality and the existence of multiple available channels have a positive impact on delay and energy metrics, whereas crowded networks are unfavourable to these performance factors. Again, the mathematical formulation of a problem is validated through simulations in [150], in which the authors set up experiments in MATLAB, to estimate the performance of CSS in terms of BER and throughput, and compare it with the BPSK performance; a modulation that is used extensively in narrowband solutions.

In [156], simulations are performed based on a discrete event simulator developed by the authors using C++, in an effort to compare results extracted from an MC formulation of
PDR estimation problem in energy harvesting Class A EDs, as described in the previous subsection. Extending the lifetime of an LPWAN system is an area of extensive research, and [173] proposes the innovative Long-Lived LoRa MAC protocol to be used in LoRa networks, composed of energy harvesting Class A EDs. The longevity of the system can be further improved via offloading the messages of EDs with drained energy to EDs with higher energy load and via a parameter selection based on heuristic methods. The proposed model is evaluated in ns-3, in terms of network lifetime and throughput. Results are also compared to a real testbed.

Blockchains are the most widely researched realization of distributed ledgers and their applicability in several networking paradigms is investigated constantly. The intersection of Blockchains protocols with IoT solutions is summarized in [145] and [146]. Related work focusing on LoRa/LoRaWAN connectivity exists both in simulation environments and in real testbeds. In [130], the authors propose an architecture with a Blockchain network running in parallel with the typical LoRaWAN system. The Blockchain network is deployed to improve OTAA process and secure the join messages against replay attacks and jamming, thanks to a proposed two-factor authentication mechanism. The mechanism is examined in a virtualized environment with Ethereum client, and showcases great results in terms of throughput; however, the increase in latency is considerable as the number of EDs rises.

IoT is one of the main verticals of LoRa/LoRaWAN, with a key research area in this frame being the resource allocation as mentioned in [142], in which the authors developed a DFRL framework to handle such needs. Their model is tested in Tensorflow (Python), with results being drawn regarding slice delay, energy consumption, throughput, and PLR; compared with another network slicing approach; and showing considerable improvements in all four factors. Network slicing in the context of LoRa/LoRaWAN connectivity is also examined in [143], in which the authors perform ns-3 simulations to validate a multi-game theoretic model that they developed for resource allocation. Results are drawn from experiments with different network configurations and slicing strategies, and reveal great performance in terms of delay, energy consumption, throughput, and PLR.

C. Real System Implementations

Implementations of IoT systems with the support of LoRa/LoRaWAN include: i) small-scale research-oriented testbeds with a few EDs, and ii) large scale deployments serving verticals of the “smart” paradigm.

One of the pioneering works on the topic of system evaluation by means of an experimental testbed is [30], in which a LoRa-specific path loss model based on real measurements is derived. The testbed consists of one Kerlink GW and two SX1272-based EDs: one mounted on top of a moving car and one on a boat. The results are promising in terms of range for both cases, assuming lenient requirements for PDR. Regarding the path loss model, it is coined to fit the performed measurements, with different formula parameters for each mobility scenario. Recent publications address coverage, with [174] examining indoor scenarios using at most ten SX1272-based EDs and at least two devices of this type to communicate with one GW. Results are drawn for path loss, RSSI, energy consumption, and Packet Reception Rate (PRR) (practically PDR).

References [11] and [13] are two more of the earliest publications in LPWAN systems. In [11], the authors focus on LoRa connectivity and evaluate some of the specified nominal values. One SX1276-based ED and a Cisco GW in an urban environment are used to evaluate: i) the RSSI metric, which is calculated slightly higher than the specified, ii) the coverage in terms of PDR, where the system performs satisfactory, and iii) the throughput, which is close to nominal values at the time of the reference. In [13], the authors provide a quick overview of Sigfox and RPMA, whereas they sketch in greater detail LoRa/LoRaWAN connectivity. They conclude their work with some coverage tests for two use cases. Considering the smart buildings vertical, a successful coverage test is conducted in a nineteen-story building, placing the GW in the middle point of it and thirty-two EDs scattered in different places inside of it. Smart city application scenarios are also examined with a Kerlink GW providing connectivity to IMST iM880A-L-based EDs as far as 2 km.

More on the topic of measurement campaigns can be found in [175]. The authors perform extensive LoRa measurements in urban and rural scenarios, thanks to a combination of both a well-established deployment and an experimental one. The entire system consists of three Kerlink GWs and multiple EDs, but for the needs of the experiments two EDs are utilized. Range and coverage are proven feasible even in the 10^3 m scale, as depicted by RSSI and SNR values, having lenient PER requirements. Part of the authors’ work is also the proposal of a reference architecture that is close to the typical LoRaWAN case, but with slightly different terminology, and a protocol stack that relies on IEEE 802.15.4 above LoRa and Constrained Application Protocol (CoAP) in APP.

As described in the previous subsection, [99] proposes a slotted ALOHA access method on top of LoRaWAN. Apart from MATLAB simulations, the authors deploy a real system to perform several tests. The network is consisted of a maximum of twenty-four RFM95W EDs and a GW from Multitech based on a SX1301 and two SX1357, in an indoor environment. The experiments verify the improvement of their slotted ALOHA proposal compared to pure ALOHA. Reference [100] has a similar approach, with the authors validating their MATLAB simulation results via the establishment of a link between a USRP and an SX1276-based transceiver. Another MAC consideration is given in [106], in which a slotted ALOHA scheme is examined. The mathematical approach and the simulation study have already been discussed for the evaluation of this scheme. On top of them, a testbed consisted of two SX1272-based ED and two Si4703 FM-RDS decoder for time dissemination purposes is also developed.

An innovative approach to the topic of multiple access is given in [102]. The authors propose a new architecture for the access part of the system containing an extra component; namely the Cluster Head (CH). EDs’ UL messages can be received by the distant GW or by the near CH and then
forwarding DL messages to the EDs of its cluster. To ameliorate
the energy consumption of EDs, a TDMA-based scheme is
also proposed. Findings from testbed experiments containing
one GW, one CH, and nine SX1276-based EDs showcase
enhancements in PDR, latency, and EDs’ lifetime.

Authors in [131] explore the feasibility of different jam-
mimg techniques in LoRa communications. They implement
an Arduino-based jammer and an Arduino-based sniffer using a
HopeRF module for each, and perform two successful jam-
ming attacks in two RN2483-based EDs, being served by one
GW. Firstly, they conduct a selective jamming attack, i.e., jam-
mimg the communications of one ED, while leaving intact the
communications of the other ED, with significant outcomes:
98% of the jammed packets do not reach the GW, regard-
less the SF configuration. Secondly, they perform a wormhole
attack, i.e., jamming the communications of an ED and later
replaying them, with the results showcasing great resistance
of low SFs, but also great vulnerability of high SFs. All the
outcomes are statistically significant, since the experiments
incorporate the sending of $10^5$ UL messages for each con-
figuration of the first experiment and $10^5$ UL messages for each
configuration of the second experiment. The authors cap their
work off by discussing the applicability of their experiments
to commercial applications and the possible countermeasures.

The deployment presented in [166] is consisted of three
GWs in urban placement and four RFM95W-based (or
SX1276-based) EDs with different boards. Experiments are
conducted for indoor fixed position EDs, outdoor fixed posi-
tion EDs, and moving EDs, resulting in throughput charac-
terization in PHY (similar to theoretical) and APP (about a
tenfold decrease). Also, coverage estimation based on PDR
is provided, resulting in an optimal of 7.5km range with a
greater than 95% PDR, a poor behavior in indoor scenarios,
and a minor role of velocity in mobility cases (at least com-
pared to other propagation factors, such as LOS). Part of the
discussed testbed is also utilized to evaluate the LCS mecha-
nism presented in [152]. Two of the RFM95W-based EDs
and one GW were used in an indoor environment to test the
transmissions detection precision of the protocol, which was
proved to be highly accurate.

In [172], the results of the energy-oriented simulation study
conducted in [170] are validated thanks to a measurement cam-
paign that includes a Kerlink GW and four SX1272 EDs in
an outdoor environment with several distances being exam-
ined. Energy considerations appear to be a topic of intense
research in LoRa/LoRaWAN deployments, and in [176] the
authors deploy a real network in an indoor environment to
control a Heating, Ventilation, and Air Conditioning (HVAC)
system, according to room’s occupation. The network cons-
ists of one GW and one RFM95W-based ED, along with
one ISM radio ED for comparison reasons. The decision for
on/off switching of HVAC is left to a Random Neural Network
(RNN) controller, residing in the GW (in the context of that
reference the GW is the BS), with the model being trained
on a cloud infrastructure. Results are promising in terms of
ED’s and building’s energy consumption, range, and path loss,
especially when compared to the ISM transceiver.

In [157], the authors provide a rigorous analysis of the
energy consumption in the context of LoRaWAN. The utilized
testbed is consisted of an SX1272-based ED and a Kerlink
GW, being deployed in an indoor environment. Thanks to
the measurements and analytical formulas, numerous results
regarding energy consumption, ED’s lifetime, and UL trans-
missions’ energy cost are drawn. Most surprisingly, the ACK
ULs are more energy effective than the non-ACK ones; an out-
come that can be attributed to the fact that Class A EDs that
received an ACK in the first receive window go to idle state
without opening the second receive window. The proposed
Long-Lived LoRa protocol is evaluated via simulations, as dis-
cussed in [173]; however, a real testbed is also deployed to
further investigate the model’s extended lifetime and system
throughput. The testbed is consisted of one Raspberry Pi 3
GW based on the well-known RAK2245 HAT solution, fifteen
EDs operating with the SX1276 chipset and the open-source
solution of the ChirpStack NS. Both the GW and the EDs are
placed indoors.

Blockchain applicability in LoRa/LoRaWAN is examined
not only via simulations but also in testbeds. In [177], the
authors propose the HyperLoRa framework, an innovative
approach that ameliorates NS’s workload by exploiting GWs
capabilities. A central cloud service that processes all the data
is prone to data falsification or loss. With the proposed solu-
tion, a Blockchain network at the edge with different ledgers
for different types of data (application data at the server-side,
OTAA join-request messages at the GW’s cluster) is built.
The network is tested with Hyperledger Fabric (HyperLoRa
is the proposed system), four GWs (with two SX1255 and
one SX1301 in each), and one ED SX1276, whereas the pro-
liferation is achieved with Locust framework for up to two
thousand EDs. Results are promising: the distributed message
processing between server-side and GW leads to amelioration
of CPU utilization in NS, the BW utilization between GW and
NS is improved, whereas throughput is practically the same
with or without HyperLoRa. In short, HyperLoRa has close to
no HyperLoRa implementation performance, but with greater
resource utilization.

Further investigation of Blockchains’ integration in
LoRaWAN systems can be found in [178]. The authors explore
alternative approaches to the topic of data storage and acces-
sibility. To this end, they propose a LoRaWAN-based system,
in which the data are stored to a Swarm storage space and
the access to them is provided via Ethereum smart con-
tacts. Validation of their proposed system is achieved with
a testbed that includes one Raspberry Pi 2-based ED with the
Dragino LoRa/GPS Hat, one Raspberry Pi 3-based GW with
the iC880A module, a local proxy server, the Swarm stor-
age system, and an Ethereum client. Results reveal a viable
memory utilization for all entities (cloud server-side, GWs,
and EDs), but throughput appears to be an issue due to EDs pro-
liferation. Nevertheless, Blockchain’s transactions are focused
on GWs, which are fewer in numbers in real systems and can
aggregate the traffic of multiple EDs.

Finally, in [179], a Blockchain-based architecture for IIoT
is proposed. Authors address the problem of data manipu-
lation, and the integration of Blockchain protocols to IIoT
as a solution to it. The suggested Blockchain IIoT architecture is called BIIT, and targets data reliability and improved energy efficiency with little computational overhead. To validate their architecture, the authors set up a testing platform, that includes one Arduino-based ED with the Dragino LoRa shield for the LoRa-related experiment and the same ED with the Sixfab Cellular IoT Application Shield for the LTE-M-related experiment, a Raspberry Pi acting as a local TTN GW, and a Raspberry Pi 3 having a threefold purpose: a Blockchain plugin for TTN, a Blockchain client, and a Blockchain miner. The results on packet overhead are very promising when the number of signed packets is high.

D. Summary of Research Points

Published research in the area of LoRa/LoRaWAN connectivity showcases:

1) The contribution approach may be categorized as analytical, simulation, or field deployment. Mathematical analysis is needed to optimize system deployment and operation, simulations are needed to examine the behavior of the system in terms of ED proliferation, whereas field tests are permitted thanks to low CapEx and OpEx.

2) Substantial effort from the research community is allocated to two directions:
   - Examine well-known wireless connectivity challenges in the context of LoRa/LoRaWAN systems, that affect mainly signal propagation in PHY and network access in MAC.
   - Explore the applicability of next generation networks’ key features and technologies, such as ML, SDN, Blockchain, and network slicing.

In Table VII, the recent research work in the area of LoRa/LoRaWAN is classified based on the approach of each contribution. While the references depicted in Table VII are a meaningful set that can provide the reader with a complete view of the state-of-the-art, the interested reader is encouraged to check for more works by the authors of the cited publications.

XII. OVERVIEW OF SIMULATION TOOLS

The proliferation of EDs in WSNs dictates the simulation of the interested system to be rolled out, in an effort to estimate its performance, locate any flaws, and try to optimize its development. In this context, the available simulators can assist network engineers’ job, since they allow them to test the system to be deployed, by using virtual system components and by examining several use cases and scenarios. Among the great advantages of simulation environments are: i) the lack of physically deployed nodes, which could be proven costly in the case of sub-optimal deployment, and ii) the ability to easily scale up the system, by just adding virtual nodes.

Focusing on IoT networks and especially on the area of LPWANs and one of the key representatives, LoRa/LoRaWAN, the target of these simulations is threefold:

1) evaluate the number of devices in the network, especially in terms of EDs, to estimate the capacity of the network and the margin for upscale,
2) position EDs and GWs in a way that the network’s coverage is optimized, and
3) use appropriate radio parameters in terms of BW, SF, CR, and TP, to estimate the performance of the network. Other common simulation use cases in the area of IoT include the estimation of EDs’ longevity, the evaluation of EDs’ mobility, and the examination of applicability of common WSN algorithms.

A number of commercially available simulation environments, free and/or open-source, are able to perform the required simulations for a wireless network. Some indicative simulators are [180]:

- The RF design system Planet by Infovista, which allows the design and optimization of wireless cellular networks that have been standardized by 3GPP.
- The software suit Signal Pro by EDX, which includes several modules of widely used networks (e.g., LTE, WiMAX) and provides the necessary tools for RF analysis with the use of propagation models for both indoor and outdoor environments.
- The cellular radiocommunication design tool ASSET Radio by TEOCO, which comes with capabilities of analyzing the coverage of cellular networks, their capacity, and the configuration of several parameters.
- The software HTZ Communications by ATDI, which can be used in designing several wireless systems, such as cellular networks, IoT deployments, satellite systems, and others.
- A number of software suites by CelPlan, which cover a number of simulation use cases, such as design and optimization of an outdoor wireless system, design of a network in an indoor environment, analysis of Point-to-Point (P2P) links, and others.
- Several network design tools from Siradel, which include tools for cellular system design, WiFi networks, specialized tools for modeling the signal propagation in 5G networks, and the tool S_IoT, which can be used for LPWAN design.
- The Pathloss platform, which can assist the design of radio links that operate in a wide area of spectrum; from 30MHz up to 100GHz.

Despite the usefulness of these tools, in the rest of this section we will focus on simulation environments that are free to use and support the LoRa modulation and/or the LoRaWAN protocol, either in their core codebase or via the use of extensions. Along with the discussion of these simulators, references of published work utilizing them are given. An overview of the available open-source LoRaWAN simulators can be found in [148]. Additional information on the broader topic of WSN simulators and on the evaluation methodology for such networks is available in [144].

A. COOJA

Contiki OS Java (COOJA) [181], [182] is a network simulator developed for WSNs. The simulator is used for the study of the behavior of EDs that are called motes and run the Contiki OS [183]. COOJA is an open-source simulator and can be found in [184].
TABLE VII
A CLASSIFICATION OF THE LoRa/LoRaWAN RESEARCH WORK

| Ref. | Algorithm or Process Design | Performance Analysis | Hardware, Software, or Protocol Development | Mathematical Approach | Evaluation Process | Real System Implementation |
|------|-----------------------------|----------------------|-------------------------------------------|-----------------------|--------------------|---------------------------|
| [11] | RSSI, Capacity, PDR, Throughput | √ (no info) | Contiki | Urban, 1 ED, 1 GW |
| [13] | Coverage | | Coverage | Indoor & Urban, 32 EDs, 1 GW |
| [14] | PDR, PLR, Capacity | √ (no info) | Path loss model | Urban, 2 EDs, 1 GW |
| [10] | RSSI, Path loss | | Coverage, Energy, Lifetime | MATLAB |
| [81] | Scalability, DER | Simulator development | LoRaSim | D2D link |
| [98] | Throughput, Coverage | Simulator development | MATLAB | Indoor, 24 EDs, 1 GW |
| [99] | Synchronization procedure | Throughput | Slotted Aloha on top of LoRaWAN | MATLAB | USRP to transceiver |
| [100] | BER to SIR, PER to SIR | | Throughput | Successful UL probabilities | MATLAB | 2 EDs, 1 GW |
| [101] | Throughput, Successful UL | | PDR, Latency, Lifetime | TDMA-based MAC | Indoor, 9 EDs, 1 CH, 1 GW |
| [102] | PDR, Latency, Lifetime | Network architecture | Contiki NG | OMNet++ |
| [103] | Throughput | CSMA/CA | Contiki NG | MoToSim |
| [104] | Throughput, Latency | MAC protocol | | |
| [105] | PDR, Energy | TS-LoRa | LoRaFRE | | |
| [106] | PDR, Energy | Slotted Aloha, Simulator extension | Type-A, B uncertainty | LoRaSimEnergySim (LoRaMACSim) | 2 EDs, 2 PM- R&S decoders |
| [122] | ADR | ADR | Kruskal-Wallis | LoRaWANSim | Urban |
| [123] | Energy model | Energy | Analytical models | LoRaWANSim | Urban |
| [130] | OTAA, two-factor authentication | Throughput, Latency | Blockchain-based architecture | DliberSim | 2 EDs, 1 GW |
| [131] | Jamming percentage | Sniffer/Jammer development | | MATLAB | Indoor, 2 EDs, 1 GW |
| [141] | GW placement model | CapTel, OpTel | PDR vs. SNR | Urban |
| [142] | Resource allocation | Energy, Delay, Throughput, PLR | DPRL architecture | TensorFlow (Python) |
| [143] | Resource allocation | Energy, Delay, Throughput, PLR | Network slicing | Game theory | ns-3 |
| [149] | Interference model | SER, PER, SNR, SIR | Coherent LoRa transceiver | DFT, Q-function | Monte Carlo |
| [150] | SF orthogonality, BER, Throughput | Closed form expressions | | MATLAB |
| [151] | BER vs. SNR | Optimal receiver | | MATLAB |
| [152] | Two energy models | PDR, Energy, Scalability | LCS MAC protocol | PDR probabilities | LoRaSim (Java) | Indoor, 2 EDs, 1 GW |
| [133] | Scalability | Stochastic geometry | √ (no info) | Monte Carlo |
| [134] | Energy, Delay | | | Sclab |
| [135] | Energy, Throughput | | | MDP |
| [136] | PDR | | | C++ |
| [157] | Energy models | Energy, ED lifetime | Analytical models | Indoor, 1 ED, 1 GW |
| [159] | DER, Scalability | Goodput, Energy | Simulator extension | LoRaSim | Indoor, 1 ED, 1 GW |
| [160] | Goodput, Energy | Simulator extension | LoRaSim EnergySim | Indoor, 1 ED, 1 GW |
| [161] | Energy, DER | Simulator extension | LoRaSim EnergySim | Indoor, 1 ED, 1 GW |
| [163] | Resource allocation | PDR, Energy | Simulator development | LoRaMAB | |
| [164] | UL scheduling | DRR, Energy | Simulator development | LoRaFRE | |
| [165] | PDR, Energy, Delay response rate | Simulator development | LoRaSimEnergySim (MATLAB) | Indoor |
| [166] | Throughput, QoS, PDR, Coverage | Simulator development | LoRaSim (Java) | Urban & Indoor, 4 EDs, 1 GW |
| [167] | ADR+ | PDR, Energy | Simulator development | OMNeT++ | Indoor, 1 ED, 1 GW |
| [168] | | PDR | Simulator extension | ns-3 |
| [169] | | PDR | Simulator extension | ns-3 |
| [170] | | Energy | Simulator extension | D2D link |
| [171] | | Error model | PDR | Simulator development | ns-3 |
| [172] | | PDR, Energy | Simulator development, CSMA enhancement | 4 EDs, 1 GW |
| [173] | | Heuristic parameter selection | Throughput, System Lifetime | Long_Lived LoRa | Indoor, 5 EDs, 1 GW |
| [174] | | Path Loss, PRR, Energy, RSSI | Long_Lived LoRa | Indoor, 5 EDs, 1 GW |
| [175] | | RSSI, SNR, PDR | Architecture, Protocol stack | Urban & Rural, 2 EDs, 3 GWs |
| [176] | | Range, Energy, Packet loss | Blockchain-based architecture, HyperLoRa | Indoor | Indoor, 1 ED, 1 GW |
| [177] | | Edge computing, Two ledgers | Throughput, CPU, BW utilization | Blockchain-based architecture, HyperLoRa | HyperLedge Fabric | Indoor, 1 ED, 1 GW |
| [178] | | Data storage and retrieval | Throughput, Memory | Blockchain-based architecture | Swarm storage, 1 ED, 1 GW |
| [179] | | Management architecture | Overhead | Blockchain-based architecture | 1 Edge Node | Urban, 1 ED, 1 GW |

Contiki is an OS that finds usage in devices operating under power restrictions. It is an open-source software and additional information can be found in the official software repository [185]. Recently, a new branch has been developed, called Contiki NG [186], that is gaining increasing attention and can support several protocols frequently examined in
WSNs, such as CoAP, Simple Network Management Protocol (SNMP), and Routing Protocol for Low-Power and Lossy Networks (RPL).

Thanks to these advantages, Contiki is used by numerous EDs of the IoT ecosystem and, consequently, of the LPWAN domain. Networks that are consisted of such EDs, can be simulated with the aid of the COOJA simulator. The simulator environment supports the analysis of the message exchange between EDs and the aggregator, the obtainment of information about the state of the nodes and the communication, and the adjustment of the communication parameters. A graphical interface is included, in which the topology of the network and the EDs’ measurements can be visualized. Key asset of COOJA is its high extension capability, which can provide insights in energy consumption and coverage (e.g., WSN-Maintain [187]), apart from the built-in results on network behavior and IoT-related protocols performance. Plugins that support ED’s mobility also exist [188]. Motes in COOJA are programmed using either C or Java.

B. TOSSIM

A similar approach to COOJA is TinyOS Simulator (TOSSIM) [189], a discrete event simulator for EDs that run TinyOS [190]. TOSSIM is also open-source and its repository can be found in [191].

TinyOS can be found in devices that are operating under power and computing constraints, in a concurrent way, while being easily modular. These four aspects are the main motivation points behind TinyOS development. TinyOS is open-source itself and its main repository can be found in [192]. TinyOS, and consequently TOSSIM, are based on the network embedded systems C (nesC) programming language [193], [194], a limited dialect of C. TOSSIM supports protocols that are of interest in the area of WSNs, such as Destination-Sequenced Distance-Vector Routing (DSDV) and Directed Diffusion.

In TOSSIM, the user can simulate the behavior of virtual EDs or even real motes (TOSSIM uses the same notation for EDs as COOJA). Since the deployed EDs run nesC, the loading of a real deployment’s code to the simulator is supported. Visualization is supported by TinyViz, with JTOSSIM being an alternative. TOSSIM is highly extensible, with an example of this being the modeling of energy consumption via PowerTOSSIM [195]. Mobility support also exists [196], whereas numerical results about the performance of the network can be obtained. Regarding the coverage estimation of the network, to the best of our knowledge, there are no options for importing real map data into the simulator.

C. CupCarbon

A different approach to the previous simulators is CupCarbon [197], [198], [199]. The simulator focuses on simulating WSNs in the context of smart cities, by supporting three of the most widely used technologies: ZigBee, LoRa, and WiFi. CupCarbon can be found in [200].

CupCarbon comes with a number of advantages. Its Graphical User Interface (GUI) provides visualizations with the use of multiple mapping methods, such as Google Maps and OpenStreetMap. Simulator’s documentation is fully explanatory [201] and includes all the mapping options, a description of the GUI, a presentation of the scripting language used in the simulator, and a vast amount of examples that can be used in an effort to get familiar with the simulator environment or start examining WSNs and their related algorithms, such as Low-Energy Adaptive Clustering Hierarchy (LEACH) and Distributed Least Polar-angle Connected Node (D-LPCN). The scripting language that is used is called SenScript and is consisted of a small number of commands, suitable for creating simple and more complex scenarios. In the latest version of the simulator, EDs’ programming using Python is also supported.

Another important aspect of CupCarbon is the utility of event creation and the integration of events in a simulation. These events can be observed and reported by the EDs. Regarding the storage, extraction, and visualization of the results, the user can rely on CSV files, log files in the form of text, and the console of the simulator. Furthermore, graphs can be drawn inside the simulator, focusing on the EDs’ power consumption.

D. LoRaWANSim (MATLAB)

A recently proposed simulator called LoRaWANSim [165] has been implemented in MATLAB, with the source code available in [202]. Due to a name conflict with a simulator discussed in Section XII-F, we will refer to the current simulator as LoRaWANSim (MATLAB).

LoRaWANSim (MATLAB) focuses solely on LoRa/LoRaWAN systems. It can simulate both LoRa PHY and LoRaWAN MAC, while considering EU868 limitations. The simulator offers many advantages in terms of accurate result extraction; most notably, the support of numerous GWs and the account of all types of interference (among EDs, among GWs, and a mixture of these types). Since the simulator is implemented in MATLAB, this is the programming language that is used to set up the simulations.

There is no GUI implemented, but a map can be drawn with no real terrain, serving mainly as a visualization of the EDs’ and the GWs’ position relative to each other. Results include performance of the network in terms of PDR and energy consumption, and are accurate when compared to analytical models that are discussed in [165]. Being a LoRaWAN-focused simulator, it supports the key algorithmic process of the protocol: ADR performed by NS. Mobility is not supported.

E. LoRaSim (Java)

Another simulator dedicated to the study of LoRaWAN systems is presented in [166], developed using the Java programming language. In Section XII-F, a simulator with the same name is described, so we refer to the current simulator as LoRaSim (Java).

LoRaSim (Java) is a discrete event simulator. It has been developed to simulate devices operating in US915 band. Configuration files for GW, EDs, and the network can be used
as an input to the simulator, including typical parameters, such as BW, SF, CR, and TP. Results about network performance can be obtained, such as PDR. Algorithmic issues can be examined, as the one described in [152], in which a proposed carrier sensing mechanism is compared to pure ALOHA in terms of power consumption and system capacity, using the developed simulator.

No GUI details are provided; however, LoRaSim (Java) supports the placement of all LoRa access network’s nodes (GW and EDs) in a coordinated system, but not with the use of a real map. EDs’ mobility is not examined in the corresponding references.

F. LoRaSim/MoTSim/LoRaWANSim/LoRaEnergySim/LoRaMACSim (Python)

LoRaSim is a discrete event simulation environment for LoRa networks. The simulator has been developed using the Python programming language and the simulation framework SimPy, and thus it should not be confused with the namesake Java-based simulator discussed earlier. Its main target is to examine the collisions in LoRa systems and provide insights on the topic of scaling such networks.

Relevant published works for LoRaSim are [96] and [159]. In these references, the capabilities of the simulator are discussed, while the various simulations along with their setups are presented. The simulator can support the creation of multiple LoRa EDs and their communication with one or more GWs. LoRa EDs are simulating the behavior of an element that is based on SX1272 chipset, while LoRaSim’s GWs are based on SX1301. The adjustable parameters in LoRaSim are BW, CF, SF, CR, and TP. Results can be obtained about the network performance via metrics, such as DER. A map with the relative positions of GWs and EDs can be drawn; however, it is not based on a real environment. There is no further visualization and mobility is not supported. The simulator can be found in [203].

LoRaSim has proved to be a big success with several research teams working on the simulator, in an effort to modify or extend the environment to suit their needs. One such research contribution can be found in [104], in which the authors base their MoTSim simulator in LoRaSim, trying to examine their proposed MAC protocol. Subsequently, another research team has extended the LoRaSim simulator, via implementing the necessary support for the LoRaWAN protocol [160]. The extended version of the simulator is called LoRaWANSim (not to be confused with the previously discussed MATLAB-based simulator). Extensions are focused on MAC, to support: DL data and ACKs, duty cycle limitations, collision models, re-transmissions, and data rate adaptation based on unsuccessful packet transmissions. Results include energy consumption and network performance indicators.

A further improvement to the original simulator, named LoRaEnergySim, came from a third research team and is described in [161]. Simulations are organized in a framework that contains three main entities: ED, air interface, and GW. An important aspect of the air interface class is the consideration of known propagation models, such as the log-distance model and the COST-231 model, with the latter fitting nicely the European deployments. For the ED, a realistic energy model is considered based on the LoRaWAN_EFM32, whereas the GW is based on the iC880A module. Algorithmic-wise, the ADR algorithm as implemented by TTN is also included. The simulator is available in [204].

Based on LoRaEnergySim, another research contribution is described in [106], with additional capabilities for the simulator. The LoRaMACSim takes into account two more well-known propagation models; the Rayleigh fading and the ETSI TR 136 942. Moreover, the topic of multiple access can be studied thoroughly, thanks to the consideration of the slotted ALOHA and CSMA/CA protocols, apart from the usual approach of pure ALOHA. LoRaMACSim is also available, in [162].

G. LoRa-MAB

One more discrete event simulator built with the Python programming language and the SimPy framework is LoRa-Multi-Armed Bandit (LoRa-MAB) [163]. The goal of the simulator is to provide an environment for estimating resource allocation in LoRa networks in an intelligent distributed way and analyze the scalability potential of a LoRa network.

To this end, the authors have developed a simulator that has multiple input parameters, supporting a plethora of EDs and GWs, while assuming real radio propagation models and other realistic impact factors, such as capture effect. EDs are distinguished to normal EDs and smart EDs, with the latter being able to perform the proposed intelligent distributed algorithm for resource allocation. Results can be drawn for energy and other performance metrics, such as packet reception rate (which can be thought of as PDR), whereas coverage can be examined, but not with the use of a real map.

EDs’ mobility is not discussed and a GUI is not provided. LoRa-MAB is available in [205].

H. LoRaFREE

The final discrete event simulator built with the Python programming language and the SimPy framework to be discussed in this paper is LoRa Fine-Grained Scheduling for Reliable and Energy-Efficient Data Collection in LoRaWAN (LoRaFREE) [164]. This simulator targets the validation of the namesake FREE algorithm.

The authors in [164] have developed a highly configurable simulator able to support multiple EDs, with several input parameters, including the usual BW, SF, CR, and TP. LoRaFREE uses realistic assumptions in a number of areas, such as a packet error model, the imperfect orthogonality of SFs, fading events, and a duty cycle limitation for EDs and GWs (European deployments). Results can be drawn for numerous performance metrics, with the most important of them being the DDR, the collisions, and several ACK-related metrics. Energy consumption can be measured thoroughly in the simulator environment. EDs and GWs can be drawn in a non-real map. Algorithmic issues can be examined, as depicted by the study of the proposed FREE mechanism.
No additional GUI is provided apart from the projection of EDs and GWs in a coordinated system. Mobility is not taken into consideration. LoRaFREE can be found in [206].

I. OMNeT++

The Objective Modular Network Testbed in C++ (OMNeT++) is a simulation environment that is based on modules written in C++. With these modules, simulations for various networking systems, both wireless and wired, may be created. It follows the discrete event simulation pattern, as many other simulators in the field.

Its popularity among other simulation environments is justified by the numerous available frameworks. The widely used INET framework includes models for examining several well-known protocols from all the layers of the TCP/IP protocol suite, for nodes that can be either at fixed position or in motion. Other commonly used frameworks are: SimuLTE, which is based on INET and is used for evaluating the behavior of LTE and LTE-A networks; Simu5G, which extends the simulation capabilities to 5G networks; Artery, which targets Vehicle to Everything (V2X) simulations; Castalia [207], which targets the simulation of networks consisted of energy-constrained devices, such as WSNs; PAWiS [208], which also strives for WSN simulations; and others.

With the increasing adoption of LoRa networks, it was expected that a framework for such systems would be developed. The Framework for LoRa (FLoRa) [167] simulates the behavior of EDs, GWs, and NS in a LoRa network. It supports adjustment of parameters, including a dynamic scheme of parameter modification reminiscent to ADR. Details about the framework can be found in the corresponding Web page [209] and its repository [210]. The capabilities of gathering statistics regarding the transmissions (PDR) and the power consumption are considered valuable. FLoRa supports an energy model based on the power consumption of the SX1272/73 chipset. It is worth to mention that OMNeT++ (and as a result FLoRa) provides an intuitive GUI that offers visualization of the network and even simulations based on realistic physical environment, while tools to log events and analyze them are also available.

J. ns-3

The final simulator to be discussed is network simulator 3 (ns-3) [211], which has an extensive usage in networking research. It is part of the simulator series ns, with its core codebase developed with the C++ programming language, while also supporting the setup of simulations using Python. ns-3 utilizes built-in libraries, which can be combined for the development of various networking systems, IP-based or not.

Simulations in ns-3 are performed with the use of five main Abstract Data Types (ADTs): the node, the application, the channel, the network device (netdevice), and the topology helper. More information can be found in the official documentation [212]. The vast amount of libraries assists the development and simulation of different kinds of networks: from 3GPP networks to IEEE 802.11 systems, and from networks that are based on modern paradigms such as SDN to systems that experience intense heterogeneity of new and old technologies. In this context, LPWANs can be considered, and more specifically LoRaWAN systems, which are a subject of intense research.

A module able to simulate LoRaWAN networks is available and can be found in [213], while published works [98], [168], and [169] explore and greatly extend the LoRaWAN capabilities of ns-3. In the same context, work published in [170] extends these modules for realistic energy consumption estimation. Other efforts on ns-3 LoRaWAN modules development also exist, with [171] focusing on PDR using ns-3 applications implemented by a different research group, whereas another research team investigates the feasibility of CSMA protocol on top of LoRa in [172]. Moreover, several recording commands and log file creation commands can aid the analysis of network’s behavior, whereas the input of a simulation is highly configurable. Simulations to log any network performance metric of interest can be developed. Algorithmic processes are regularly simulated in ns-3 and in the context of the LoRaWAN module such efforts do exist, ranging from the common ADR algorithm to typical WSN-related algorithms, such as the Pruning algorithm [213].

Considering the graphical interface, ns-3 is a terminal-based simulator. Nevertheless, visualizations can be achieved with the use of an animator, such as NetAnim or PyViz. ns-3 provides several mobility models, supports coverage evaluation, while the use of reminiscent to real environments further improves the validity of the extracted results. ns-3 is an open-source project [214].

K. Summary of Simulation Environments

Table VIII provides a comprehensive summary of the presented simulation environments. The simulators are examined in terms of the supported programming languages, the existence of a GUI, and the capabilities of them in terms of EDs energy consumption, network coverage, network performance (throughput, data rate, PDR, or related metrics), ED mobility, and the support of algorithmic processes or protocols. A performance factor that could have also been discussed is scalability, i.e., the number of EDs that can be simulated; however, there is no apparent distinction in this area: all these simulators can provide results in simulations crowded with EDs.

XIII. Final Remarks

In this final section, we position our work against existing surveys in the literature. We also provide conclusions and directions for future work.

A. LPWAN and LoRa/LoRaWAN Surveys

The literature on LPWAN and LoRa/LoRaWAN is already vast and valuable. Table IX includes recent and high impact surveys that examine different aspects and challenges of the LPWANs and the LoRa/LoRaWAN technologies. In the two main columns of the table, we indicate: i) the LPWAN technology that is extensively (E) or briefly (B) examined in each one
TABLE VIII
OVERVIEW OF SIMULATION ENVIRONMENTS WITH LoRa/LoRaWAN SUPPORT

| Simulation Environment | Ref. | Target | Programming | GUI | Energy | Coverage (Map) | Network Performance | ED Mobility | Algorithmic Processes |
|------------------------|------|--------|-------------|-----|--------|----------------|---------------------|-------------|----------------------|
| CooJA                  | [181]| IoT generic (Confluent EDs) | C, Java | Built-in | ✔ | ✔ (real) | ✔ | ✔ | ✔ |
| TOSSIM                 | [182]| IoT generic (TinyOS EDs) | nesC, C, C++ | TinyViz, TOSSIM | ✔ | - | ✔ | ✔ | ✔ |
| CupCarbon              | [197]| EDs with Zigbee, WiFi, LoRa radio | SenScript, Python | Built-in | ✔ | ✔ (real) | ✔ | ✔ | ✔ |
| LoRaWANSim             | [165]| LoRa/LoRaWAN specific | MATLAB | - | ✔ (not real) | ✔ | - | ✔ |
| LoRaSim                | [152]| LoRa/LoRaWAN specific | Java | - | ✔ (not real) | ✔ | - | ✔ |
| LoRaSim, LoRaWANSim, MoTSim, LoRaEnergySim, LoRaMACSim | [163]| LoRa/LoRaWAN specific | Python | - | ✔ (not real) | ✔ | - | ✔ |
| LoRa-MAB               | [160]| LoRa/LoRaWAN specific | Python | - | ✔ (not real) | ✔ | - | ✔ |
| LoRaFREE               | [164]| LoRa/LoRaWAN specific | Python | - | ✔ (not real) | ✔ | - | ✔ |
| OMNeT++ (PLoRa)        | [167]| Network generic (LoRa/LoRaWAN specific) | C++ | Built-in | ✔ | ✔ (real) | ✔ | ✔ | ✔ |
| ns-3                   | [98]| Network generic (LoRa/LoRaWAN specific) | C++, Python | NetAnim, PyViz | ✔ | ✔ (partly real) | ✔ | ✔ | ✔ |

In the following list, we highlight the most indicative aspect of each one of the surveys presented in Table IX.

- Augustin et al. [11], Nolan et al. [12], and Centenaro et al. [13] share a similar approach for their studies; they provide some experimental results (based on real testbeds), considering, among others, the spatial coverage of the network.
- Bankov et al. [14] provide a rigorous overview of LoRaWAN, by indicating important performance aspects regarding channel access, overhead, back-off policy, retransmissions, data rate selection, and network load.
- Raza et al. [15] follow a holistic approach, touching each LPWAN topic presented in the second column of Table IX.
- Callebaut et al. [26] provide a complete study of energy considerations for the three most adopted LPWAN technologies: NB-IoT, Sigfox, and LoRa/LoRaWAN.
- Gkotsiopoulos et al. [27] perform a complete survey on LoRa research challenges, mainly in terms of system performance and capacity support, through an extensive examination of the published state-of-the-art work.
- Sinha et al. [16] and Mekki et al. [22] identify the most prominent LPWAN technologies and focus exclusively on their comparison in terms of several IoT-related KPIs. They also examine application scenarios for the discussed solutions.
- Wang and Fapojuto [17] provide a comprehensive study where the reader can find research directions on LPWANs.
- Adelantado et al. [18] provide intuition regarding LoRaWAN MAC capabilities, through quantitative examples on ToA, throughput, and PDR. Furthermore, applications of LoRaWAN systems are given.
- Saari et al. [19] provide a broad classification of research areas in the context of LoRa/LoRaWAN.
- Haxhibeqiri et al. [20] give an extensive overview on the topic of LoRa/LoRaWAN research and present works that have used all three evaluation methods: mathematical modeling, simulation, and field experimentation.
- Qadir et al. [21] depict the PHY and the MAC of RMPA as IEEE 802.15.4k implementations. Moreover, they discuss research challenges and potential for many LPWANs, examining also less-investigated aspects, such as security.
- Sundaram et al. [25] provide a valuable taxonomy of LoRa/LoRaWAN research efforts. They classify the studies based on the evaluation methodology (mathematical modeling, simulation, or field experimentation). Their insights for each research class are of great interest.
- Ayoub et al. [23] heavily focus on mobility management capabilities and they also provide a well-documented comparison in terms of IoT KPIs and application performance.
- Ertürk et al. [24] provide a complete study on LoRaWAN connectivity, discussing, among others, LoRaWAN localization, energy considerations, and applications.
Overall, the newer surveys (including ours) provide greater reference pool, and can be specialized to emerging research endeavors, e.g., LoRaWAN ADR mechanisms [124], LoRaWAN security analysis [132], etc. In this framework, this paper, following an approach similar to [15], provides a fresh look to each of the most prominent technologies of LPWANs. In addition, the most appealing license-free technology, i.e., the LoRa/LoRaWAN, is examined in depth. A wide range of aspects is discussed, similarly to the holistic approaches in [24] and [27]. Beyond this, we bring to the surface some new/less-studied aspects, referring to the network deployment as well as to cloud-based network and resource management.

**B. Conclusion and Directions for Future Study**

LPWAN is an emerging connectivity solution looking to embrace the challenge of satisfying the needs of the IoT’s third generation. Its main target is to serve connectivity demands in terms of long range, low power and low data rate, with low setup and operational cost. Moreover, LPWANs tend to be supported by cloud infrastructures; an observation that confirms the current trend on virtualizing as much as possible of the LPWANS’ functionality. As such, LPWANs can be considered as key features of the broader evolution towards a data-driven IoT world. In this framework, they are considered as a potential candidate for industrial solutions, where advanced data-driven services are designed, based on recent advancements, such as the AAS and the digital twinning.

Irrefutably, the LPWAN service paradigm (i.e., mainly data collection from highly distributed nodes) is now ready to inherit ML solutions from the field of big data [147]. Such solutions can be deployed for: i) efficient monitoring needed in various vertical industries, ii) add-on services based on data-extracted knowledge, regarding the habits of individuals, groups, and even entire industries, and iii) prediction models for efficient public protection and disaster relief. Through the ML toolkit, open challenges at the network deployment-level can be addressed, since decision-making based on a well-known environment/context can be applied.

Digging into the LPWAN connectivity/access technologies, they have already been established in several (3GPP and non-3GPP) standards, covering various application demands and network performance requirements. As revealed from the technology comparison study presented in this paper, a standout solution for LPWANs in unlicensed bands is the LoRa/LoRaWAN networking approach, referring to systems that use LoRa PHY and LoRaWAN MAC technologies. Overall, at access level, the integration of LPWANs with other
modern networks is an open challenge. Especially, for the integration with mobile networks, features such as the SDN and the network slicing, can be adopted at the core/management part of a network where LPWAN technologies address the access part. Indeed, 5G and Beyond-5G (B5G) networks target a Multiple-Radio Access Technology (Multi-RAT) scheme. Towards this, there are numerous propositions on how to incorporate different technologies of the access part into a unified core network. 3GPP LPWAN solutions can work inbred with cellular protocols, but for non-3GPP solutions, such as the LoRa/LoRaWAN, we can resort to concepts like Non-3GPP Inter-Working Function (N3IWF) [215].

Regarding the IoT/LPWANs EDs, their proliferation escalates rapidly. IoT connections have already surpassed connections of human-operated devices [4]. It is this massiveness of IoT’s EDs, that requests large scale research studies to be conducted through simulations (where performance and behavior of highly dense and distributed networks can be examined easier). That’s the motivation behind the study on the available simulators provided in this paper.

In the same context, towards providing IoT/LPWAN EDs (i.e., sensor-like devices) with the same set of connectivity capabilities with the IP network devices, IPv6 adoption is a major step. From early 2010s, IPv6 has started to be deployed in network entities, trying to solve the issues of the limited address space of IPv4, the slower routing due to checksum operations, and the fragmentation of datagrams. While the latter two are not expected to be an issue for most LPWANs (at least not a greater issue than to the rest of existing networking families), the former is a serious concern for the entirety of IoT space. IPv6 will eventually be the solution to these problems. LoRa Alliance works towards this end, trying to establish an IPv6 Adaptation layer [216]. Still though, IPv6’s roll-out is not expected to be finalized before the next decade.

Considering all the aspects presented above (cloud infrastructures, ML-based decision making, integration with mobile networks, etc.) a fertile field for further and more intent research work is defined.

APPENDIX A
UNDERSTANDING THE LoRa MODULATION

LoRa is an implementation of CSS,\textsuperscript{19} which is a member of the Spread Spectrum (SS) family [217]. SS modulations spread the signal over a wider BW, achieving better performance in terms of intentional interference (jamming), unintentional interference, and intentional detection of signals intended for others (eavesdropping) [218]. Several SS approaches have been proposed over the years, with Direct Sequence SS (DSSS) and Frequency Hopping SS (FHSS) being the main representatives. Other methods include Time Hopping SS (THSS) and Hybrid SS (HSS). LoRa CSS implementation is considered as a DSSS alternative [118].

\textsuperscript{19}LoRa networks is not the only WSN approach that relies on CSS for its PHY. IEEE 802.15.4a is another example of CSS utilization and can be used for ZigBee connectivity.

In this context, the desired signal (which is a sequence of bits) is transformed to a sequence of pulses using a code. The code is a pseudorandom sequence composed of pulses, which are also called chips. Hence, the chip rate of a code is the number of pulses per second (chips per second) at which the signal is transmitted (or received). Practically, the chip rate, \( R_c \), equals the spread BW, \( BW \), so:

\[ R_c = BW \]  

For the actual transmission, a sequence of symbols is sent. The symbols are defined as the different ways to apply a sweep of the signal (a.k.a. chirp) within the available BW (Fig. 20).

The number of chips that are used per symbol is a quantity derived by the SF and equals to \( 2^{SF} \). Given that, the symbol or chirp rate, \( R_s \), is calculated by the division of the chip rate by the number of chips used in a chirp, thus:

\[ R_s = \frac{R_c}{2^{SF}} \]  

The reciprocal of rate is duration, so we can also derive the duration of the chip and the chirp (\( T_c \) and \( T_s \), respectively) as:

\[ T_c = \frac{1}{R_c} = \frac{1}{BW} \]  

\[ T_s = \frac{1}{R_s} = \frac{2^{SF}}{BW} \]  

For the calculation of the bit rate, we multiply the number of bits that are represented in a symbol (or chirp) with the symbol rate. Again, the number of bits that a symbol represents is derived by the SF and actually equals the value of the SF that is used in the link.

\[ R_b = SF \times R_s \]  

Nevertheless, the bit rate in LoRa is slightly lower than that, due to the inclusion of a FEC scheme. LoRa defines a four-value Rate Code, picking an element from the set \( \{ \frac{4}{5}, \frac{4}{6}, \frac{4}{7}, \frac{4}{8} \} \). Rate Code is a fraction that has as a nominator the amount of useful information of a message and as a denominator the total number of bits, i.e., information bits plus redundant bits for error correction. Given the FEC scheme, the bit rate in LoRa is as follows (which is actually the equation (1) provided in Section X):

\[ R_b = SF \times \frac{BW}{2^{SF}} \times \text{RateCode} \]
# Appendix B

## Acronyms

| Acronym | Description |
|---------|-------------|
| 16-QAM  | 16-Quadrature Amplitude Modulation |
| 3GPP    | 3rd Generation Partnership Project |
| 8PSK    | 8-Phase Shift Keying |
| AAA     | Authentication, Authorization, and Accounting |
| AAS     | Asset Administration Shell |
| ABP     | Activation By Personalization |
| ACK     | Acknowledgment |
| ADR     | Adaptive Data Rate |
| ADT     | Abstract Data Type |
| AES     | Advanced Encryption Standard |
| AI      | Artificial Intelligent |
| API     | Application Programming Interface |
| APP     | Application layer |
| AWGN    | Additive White Gaussian Noise |
| AWS     | Amazon Web Services |
| AppEUI  | Application Identifier |
| AppKey  | Application Key |
| AppNonce| Application Nonce |
| AppSKey | Application Session Key |
| B5G     | Beyond-5G |
| BER     | Bit Error Rate |
| BLE     | Bluetooth Low Energy |
| BPSK    | Binary Phase Shift Keying |
| BS      | Base Station |
| BT      | Bandwidth-symbol Time |
| BW      | Bandwidth |
| C-DRX   | Connected mode DRX |
| CAD     | Channel Activity Detection |
| CDMA    | Code Division Multiple Access |
| CF      | Carrier Frequency |
| CH      | Cluster Head |
| CloT    | Cellular IoT |
| CMAC    | Cipher-based Message Authentication Code |
| CPU     | Central Processing Unit |
| CR      | Coding Rate |
| CRC     | Cyclic Redundancy Code |
| CSMA    | Carrier Sense Multiple Access |
| CSMA/CA | Carrier Sense Multiple Access / Collision Avoidance |
| CSS     | Chirp Spread Spectrum |
| CSV     | Comma Separated Values |
| CTR     | Counter (AES mode) |
| CapEx   | Capital Expenditure |
| CoAP    | Constraint Application Protocol |
| D-LPCN  | Distributed Least Polar-angle Connected Node |
| DBaaS   | Database-as-a-Service |
| DD-UNB  | Dynamic Downlink Ultra Narrow Band |
| DDR     | Data Delivery Ratio |
| DER     | Data Extraction Rate |
| DFRL    | Deep Federated Reinforcement Learning |
| DFT     | Discrete Fourier Transform |
| DL      | Downlink |
| DSDV    | Destination-Sequential Distance-Vector Routing |
| DSL     | Digital Subscriber Line |
| DSSS    | Direct Sequence Spread Spectrum |
| DT      | Digital Twin |
| DTDL    | DT Description Language |
| DevAddr | End Device Address |
| DevEUI  | End Device Identifier |
| DevNonce| End Device Nonce |
| EC-GSM-IoT| Extended Coverage-GSM-IoT |
| ED      | End Device |
| EDGE    | Enhanced Data rates for Global Evolution |
| ELIoT   | Emulated IoT |
| ETSI    | European Telecommunications Standards Institute |
| EXI     | Efficient XML Interchange |
| EXP3    | Exponential-weight for Exploration and Exploitation |
| FDMA    | Frequency Division Multiple Access |
| FEC     | Forward Error Correction |
| FER     | Frame Error Rate |
| FFT     | Fast Fourier Transform |
| FH      | Frequency Hopping |
| FSK     | Frequency Shift Keying |
| FType   | Frame Type |
| FUOTA   | Firmware Update Over The Air |
| FaaS    | Forensics-as-a-Service |
| GCP     | Google Cloud Platform |
| GERAN   | GSM EDGE Radio Access Network |
| GFSK    | Gaussian Frequency Shift Keying |
| GMSK    | Gaussian Minimum Shift Keying |
| GPRS    | General Packet Radio Service |
| GUI     | Graphical User Interface |
| GW      | Gateway |
| HARQ    | Hybrid Automatic Repeat Request |
| HTTP    | HyperText Transfer Protocol |
| HVAC    | Heating, Ventilation, and Air Conditioning |
| ID      | Identifier |
| IDS-RAM | International Data Spaces Reference Architecture Model |
| IIoT    | Industrial IoT |
| IMT-2020| International Mobile Telecommunications 2020 |
| IP      | Internet Protocol |
| ISM     | Industrial, Scientific, Medical |
| IT      | Information Technology |
| ITU     | International Telecommunication Union |
| IaaS    | Infrastructure-as-a-Service |
| IoT     | Internet of Things |
| IoTWF   | IoT World Forum |
| JSON    | JavaScript Object Notation |
| JoinEUI | Join Server Identifier |
| JoinNonce| Join Server Nonce |
| KPI     | Key Performance Indicator |
| LBT     | Listen Before Talk |
| LCS     | Lightweight Carrier Sensing |
| LEACH   | Low-Energy Adaptive Clustering Hierarchy |
| LOS     | Line of Sight |
| LPWAN   | Low-Power Wide Area Network |
| Abbreviation | Description |
|--------------|-------------|
| LR-FHSS      | Long Range - Frequency Hopping Spread Spectrum |
| LTE-M        | Long Term Evolution-Machine type communication |
| LTN          | Low Throughput Networks |
| LoRa         | Long Range |
| LoRaWAN      | LoRa Wide Area Network |
| MAB          | Multi-Armed Bandit |
| MAC          | Medium Access Control sublayer |
| MC           | Markov Chain |
| MCL          | Maximum Coupling Loss |
| MDP          | Markov Decision Process |
| MHDR         | MAC Header |
| MIC          | Message Integrity Code |
| MIOTY        | My IoT |
| ML           | Machine Learning |
| MQTT         | Message Queuing Telemetry Transport |
| MQTT-SN      | MQTT for Sensor Networks |
| MoT          | MAC on Time |
| Multi-RAT    | Multiple-Radio Access Technology |
| N3IWF        | Non-3GPP Inter-Working Function |
| NB-IoT       | Narrowband-IoT |
| NET          | Network layer |
| NIST         | National Institute of Standards and Technology |
| NR           | New Radio |
| NS           | Network Server |
| NetID        | Network Identifier |
| NoT          | Network of Things |
| NwkSKey      | Network Session Key |
| OASIS        | Organization for the Advancement of Structured Information Standards |
| OCI          | Oracle Cloud Infrastructure |
| OFDMA        | Orthogonal-FDMA |
| OMS          | Open Metering System |
| OPC          | Open Platform Communications |
| OPC-UA       | OPC-Unified Architecture |
| OQPSK        | Offset Quadrature Phase Shift Keying |
| OS           | Operating System |
| OSI          | Open Systems Interconnection |
| OTA          | Over The Air |
| OTAA         | Over The Air Activation |
| OpEx         | Operational Expenditure |
| P2P          | Point-to-Point |
| PDR          | Packet Delivery Ratio |
| PER          | Packet Error Rate |
| PHY          | Physical layer |
| PLR          | Packet Loss Rate |
| PRB          | Physical Resource Block |
| PRR          | Packet Reception Rate |
| PSM          | Power Saving Mode |
| PaaS         | Platform-as-a-Service |
| QPSK         | Quadrature Phase Shift Keying |
| QoS          | Quality of Service |
| RAMI4.0      | Reference Architectural Model Industrie 4.0 |
| RFDMA        | Random-FDMA |
| RNN          | Random Neural Network |
| RPL          | Routing Protocol for Low-Power and Lossy Networks |
| RPMA         | Random Phase Multiple Access |
| RSSI         | Received Signal Strength Indication |
| RedCap       | Reduced Capability |
| Rx           | Reception |
| SC-FDMA      | Single Carrier-FDMA |
| SDN          | Software-Defined Networking |
| SDO          | Standards Development Organization |
| SER          | Symbol Error Rate |
| SF           | Spreading Factor |
| SINR         | Signal to Interference and Noise Ratio |
| SIR          | Signal to Interference Ratio |
| SNMP         | Simple Network Management Protocol |
| SNR          | Signal to Noise Ratio |
| SON          | Self Organizing Network |
| SS           | Spread Spectrum |
| SSL          | Secure Sockets Layer |
| SaaS         | Software-as-a-Service |
| TDMA         | Time Division Multiple Access |
| TLS          | Transport Layer Security |
| TM           | Transmission Modes |
| TOSSIM       | TinyOS Simulator |
| TP           | Transmission Power |
| TS           | Telegram Splitting |
| TS-UNB       | Telegram Splitting Ultra Narrow Band |
| TSMA         | Telegram Splitting Multiple Access |
| TTN          | The Things Network |
| TVWS         | TV White Space |
| ToA          | Time on Air |
| Tx           | Transmission |
| UL           | Uplink |
| UNB          | Ultra Narrow Band |
| V2X          | Vehicle to Everything |
| VPN          | Virtual Private Network |
| VoLTE        | Voice Over LTE |
| WG           | Working Group |
| WSAN         | Wireless Sensor and Actuator Network |
| WSN          | Wireless Sensor Network |
| WoT          | Web of Things |
| XML          | eXtensible Markup Language |
| XaaS         | Everything-as-a-Service |
| eDRX         | enhanced Discontinuous Reception |
| eGPRS        | enhanced GPRS |

**ACKNOWLEDGMENT**

The authors would like to thank the anonymous reviewers for their valuable comments and the editor for organizing the review that has greatly improved this paper.

**REFERENCES**

[1] “Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond,” Int. Telecommun. Union, ITU Recommendation ITU-R M.2083-0, Sep. 2015. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf
[2] A. Al-Faqiha, M. Guizani, M. Mohammadi, M. Aleldhari, and M. Ayyash, “Internet of Things: A survey on enabling technologies, protocols, and applications,” *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015, doi: 10.1109/COMST.2015.2444095.

[3] “Overview of the Internet of Things, V1.0,” Int. Telecommun. Union, ITU/Int. Telecommun. Union ITU-T Y.2060, Jan. 2012. [Online]. Available: https://www.itu.int/rec/T-REC-Y.2060-201201-I

[4] (Ericsson, Stockholm, Sweden). *Ericsson Mobility Report*. Jun. 2021. Accessed: Jun. 6, 2022. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports

[5] *Networks of Things*, document SP 800-183, Nat. Inst. Stand. Technol., Gaithersburg, MD, USA, Jul. 2016. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/sp/nist.sp.800-183.pdf

[6] “Internet of Things,” San Jose, CA, USA, Cisco, White Paper, 2016. Accessed: Jun. 6, 2022. [Online]. Available: https://emaronsonindia.com/wp-content/uploads/2020/02/Internet-of-Things.pdf

[7] “Study on support of reduced capability NR devices (release 17)” v17.0.0, 3rd Gener. Partnership Project, Sophia Antipolis, France, 2017. doi: 10.1109/COMST.2017.2721379. [Online]. Available: https://ieeexplore.ieee.org/document/8502812

[8] K. E. Nolan, W. Guibene, and M. Y. Kelly, “An evaluation of low power wide area networks (LPWAN) for large-scale IoT deployment,” *ICT Exp.*, vol. 5, no. 1, pp. 1–17, Mar. 2019. doi: 10.1016/j.ictexp.2017.12.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405959517302953

[9] H. Xu, W. Yu, D. Griffith, and N. Golmie, “A survey of LoRaWAN technologies for large-scale IoT deployment,” *ICT Exp.*, vol. 5, no. 1, pp. 1561–1581, 2nd Quart., 2019. doi: 10.1109/COMST.2018.2877382. [Online]. Available: https://ieeexplore.ieee.org/document/8502812

[10] M. A. Ertürk, M. A. Aydin, M. T. Büyükkaklašlar, and H. Evirgen, “A survey on LoRaWAN architecture, protocol and technologies,” *Future Internet*, vol. 11, no. 10, p. 216, Oct. 2020. doi: 10.3390/fi11100116. [Online]. Available: https://www.mdpi.com/1999-5093/11/10/216

[11] J. P. S. Sundaram, W. Du, and Z. Zhao, “A survey on LoRa networking: Research problems, current solutions, and open issues,” *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 371–388, 1st Quart., 2020. doi: 10.1109/COMST.2019.2949598. [Online]. Available: https://ieeexplore.ieee.org/document/883217

[12] D. C. Marinescu, Cloud Computing: Theory and Practice, 2nd ed. Cambridge, MA, USA: Morgan Kaufmann, 2017.

[13] M. A. Ertürk, M. A. Aydin, M. T. Büyükkaklašlar, and H. Evirgen, “A survey on LoRaWAN architecture, protocol and technologies,” *Future Internet*, vol. 11, no. 10, p. 216, Oct. 2020. doi: 10.3390/fi11100116. [Online]. Available: https://www.mdpi.com/1999-5093/11/10/216

[14] J. P. S. Sundaram, W. Du, and Z. Zhao, “A survey on LoRa networking: Research problems, current solutions, and open issues,” *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 371–388, 1st Quart., 2020. doi: 10.1109/COMST.2019.2949598. [Online]. Available: https://ieeexplore.ieee.org/document/883217

[15] D. C. Marinescu, Cloud Computing: Theory and Practice, 2nd ed. Cambridge, MA, USA: Morgan Kaufmann, 2017.
Nikos Passas received the Diploma degree (Hones.) from the Department of Computer Engineering, University of Paras, Greece, in 1992, and the Ph.D. degree from the Department of Informatics and Telecommunications, University of Athens, Greece, in 1997, where he is currently a Member of the Teaching Staff and was a Group Leader of the Green, Adaptive and Intelligent Networking Research Group. Over the years, he has participated and coordinated a large number of national and European research projects. He has published more than 120 papers in peer-reviewed journals and international conferences and has also published one book and 11 book chapters. He has served as a Guest Editor and Technical Program Committee Member for prestigious magazines and conferences, such as IEEE Wireless Communications Magazine, Wireless Communications and Mobile Computing, IEEE Vehicular Technology Conference, IEEE PIMPS, and IEEE GLOBEFISH.

Lazaros Merakos received the Diploma degree in electrical and mechanical engineering from the National Technical University of Athens, Greece, in 1978, and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York, Buffalo, in 1981 and 1984, respectively. He was with the faculty of the University of Connecticut from 1983 to 1986, and Northeastern University, Boston from 1986 to 1994. Since 1994, he has been a Professor with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens (NUDA), where he has been the Scientific Director of the Network Operations and Management Center. Since June 2020, he has been serving as the Chair for the newly founded Department of Digital Industry Technologies, NUDA. He has been the Principal Investigator for NUDA in many EU funded research and development projects in the area of mobile/wireless networks. His research interests include network technologies, services, and applications, where he has authored more than 300 publications in international journals and conferences. He was a recipient of the Glabella Award for the Best Paper presented at the 1994 International Zurich Conference on Mobile Communications and the Best Paper Award in the 2008 IEEE International Symposium on Communications, Control and Signal Processing. He is the Chairman of the Board of the Greek Universities Network, a non-profit organization for the design and development of advanced ICE services for the 25 universities in Greece.