UNIVERSIDAD DE JAÉN

Escuela Politécnica Superior de Linares
DEPARTAMENTO de Ingeniería de Telecomunicación

TESIS DOCTORAL

•

LOW POWER WIDE AREA NETWORKS BUNDLING

PRESENTADA POR:
Andreas Philipp Matz

DIRIGIDA POR:
Dr. José-Angel Fernández-Prieto
Dr. Ulrich Birkel

JAÉN, 5 DE NOVIEMBRE DE 2022

SUPERVISORS:
Prof. Dr-Ing. Ulrich Birkel
Prof. Dr. José-Angel Fernández-Prieto

TUTOR:
Prof. Dr. Joaquín Cañada-Bago

LOCATION:
Jaén

Ohana means family.
Family means nobody gets left behind, or forgotten.

— Lilo & Stitch

This thesis is dedicated to my family, who have been a constant source of support and encouragement throughout the challenges of my research and the process of writing this document. Without you, this work would have been impossible. To Jenny, thank you for being on my side in happy and sad times – you mean the world to me. To all of my friends, thanks for the words of encouragement, the fruitful discussions and the laughs whenever we meet.

In loving memory of Reinhild Matz,
who taught me to never give up.

1959 – 2022

ABSTRACT

Low Power Wide Area Networks (LPWANs) are a novel group of wireless access technologies that provide an energy-efficient and long-range network service to Internet of Things (IoT) devices. There is a wide range of technologies that are significantly different in terms of network performance, which is called the Quality of Service (QoS). These differences are especially pronounced between LPWANs operated by Mobile Network Operators (MNOs) in the licensed frequency spectrum and those installed in the unlicensed Industrial, Scientific, and Medical (ISM) and Short Range Devices (SRD) bands. Therefore, extensive analysis is required to find the appropriate LPWAN for a specific application. One of the most popular licensed LPWANs is Narrowband Internet of Things (NB-IoT), which was introduced in Long Term Evolution (LTE) Release 13 and is designated as a Massive Machine Type Communication (mMTC) enabler in Fifth Generation (5G) Release 15 and onwards. In this work, a systematic QoS analysis of NB-IoT is performed to identify the central mechanisms and conditions that influence the QoS in different scenarios. The results indicate that NB-IoT meets most theoretical design goals in a commercial deployment, but exhibits high latency at very low signal levels. Afterwards, its applicability to various smart metering use cases is discussed; according to the results of the evaluation, NB-IoT is not suitable for latency sensitive applications.

As a comparison, the same analysis is performed on Long Range Wide Area Network (LoRaWAN), which is a popular unlicensed LPWAN technology. It provides a low-cost, energy-efficient and long-range network service to low-end IoT endpoints. Measurements are performed both in a deployed network, as well as in a shielded, interference-free laboratory environment. The measurements indicate that LoRaWAN meets most theoretical specifications in a deployment built from commercial hardware, but experiences a degradation in coverage due to interference and is limited by the duty cycle regulations of the unlicensed spectrum. Therefore, it is mostly appropriate for low-end use cases without hard QoS requirements. A comparison of NB-IoT with LoRaWAN illustrates that the two technologies compliment each other; combining them in a single device could provide a versatile solution for a wide range of use cases.

In traditional networks, multipath protocols provide such functionality: they bundle multiple network links together and improve the QoS. However, existing multipath protocols introduce too much overhead for narrowband and constrained network links, and while there are devices with multiple LPWAN interfaces, there is no universal LPWAN-specific bundling protocol. Therefore, the novel *Narrowband Bundling Protocol (NBP)* is proposed, which aims to improve the latency, reliability, and coverage of the aggregated connection. It provides an efficient multipath transport service to IoT devices that works over both Internet Protocol (IP) and non-IP paths. Furthermore, a proxy architecture is proposed that provides backwards compatibility by intercepting User Datagram Protocol (UDP) packets. NBP is evaluated in simulations and a field test, and its QoS improvements are verified. For the future, many protocol extensions are planned that will improve its versatility.

We are like dwarfs sitting on the shoulders of giants.
We see more, and things that are more distant, than they did,
not because our sight is superior or because we are taller than they,
but because they raise us up, and by their great stature add to ours.

— Bernard of Chartres, by John of Salisbury

ACKNOWLEDGMENTS

This thesis would not have been possible without the support and inspiration from others. I am immensely thankful to my supervisor Prof. Dr.-Ing. Ulrich Birkel, who encouraged me to start a PhD thesis and provided the resources to complete it. His guidance in identifying promising research topics and organizational efforts in obtaining the necessary equipment were invaluable to the success of this work. His advice in professional and personal matters, helpful criticism and words of encouragement helped to overcome the challenges of this work.

I am very grateful to my supervisor Prof. Dr. José-Angel Fernández-Prieto for the extensive feedback on my papers and thesis document, the fruitful discussions in defining the research topics and the kind and understanding words in personal matters. His organizational efforts and focus on the goal of completing this thesis were essential for its success.

I am also very thankful to my tutor Prof. Dr. Joaquín Cañada-Bago, who supported the creation of the papers and this thesis document with his immense experience and improved the outcome considerably.

At Technische Hochschule Mittelhessen, my special thanks goes to my former colleges at the Department of Electrical and Information Engineering: Mahmoud Mansour, who shared his deep knowledge on cloud infrastructures with me; Mark Weber, who provided help in matters regarding the laboratory and IT infrastructure; Lukas Metzger, for his support during the LoRaWAN measurement campaign; Marco Künkeler, for providing support with virtualization and network issues. I am also thankful to the administrative staff for their restless work to keep the department running.

I would like to thank the entire FlexQuartier Gießen team for the fruitful discussions and furthering my knowledge, especially: Stefan Lechner, Thomas Stetz, David Rühl, Martin Hofmann, Sebastian Wolf, Falco Klaus, Janna Walter, Moritz Hofmann, Milena Potpara and Svende Voss.

Last but not least, I am especially thankful to my family and friends for their support over all these years. My special thanks goes to Sebastian Büttner, who patiently answered all my programming questions, provided constructive feedback and gave advice on overcoming challenges during the implementation of the NBP prototype. Finally, I would like to thank Elke and Jennifer Halefeldt for ensuring readability and grammatical correctness of my papers and this thesis.

Andreas Matz
November 2022

CONTENTS

LIST OF FIGURES

0-RTT Zero Round-trip Time

3GPP Third Generation Partnership Project

4G Fourth Generation

5G Fifth Generation

A Antenna Gain

ACK Acknowledgment

ADR Automatic Data Rate

AF Assured Forwarding

A_{NBP} Coverage Area of NBP

A_{PN} Coverage Area of Path N

API Application Programming Interface

ATSSS Access Traffic Steering, Switching, and Splitting

B Bandwidth

BER Bit Error Rate

BLER Block Error Rate

BPSK Binary Phase Shift Keying

CE Coverage Extension

CID Connection Identifier

cMTC Critical Machine Type Communication

CL Coupling Loss

CR Code Rate

CRC Cyclic Redundancy Check

CS Class Selector

CSS Chirp Spread Spectrum

DCCP Datagram Congestion Control Protocol

DevAddr Device Address

DL downlink

DS Differentiated Services

DSCP Differentiated Services Code Point

DSL Digital Subscriber Lines

DSSS Direct Sequence Spread Spectrum

E_b Energy per Bit

ECL Enhanced Coverage Level

ECN Explicit Congestion Notification

EF Expedited Forwarding

EMI Electromagnetic Interference

eNodeB Evolved Node B

ETSI European Telecommunications Standards Institute

EU-433 433 MHz ISM Band

EU-868 868 MHz SRD Band

FCnt Frame Counter

FCtrl Frame Control

FEC Forward Error Correction

FHDR Frame Header

FHSS Frequency Hopping Spread Spectrum

f_{max} Maximum Frequency

f_{min} Minimum Frequency

FOpts Frame Options

FPort Frame Port

FRMPayload Frame Payload

FSK Frequency Shift Keying

GRE Generic Routing Encapsulation

GSM Global System for Mobile Communications

HARQ Hybrid Automatic Repeat Request

HoL Head-of-Line

HTTP Hypertext Transfer Protocol

IANA Internet Assigned Numbers Authority

IETF Internet Engineering Task Force

INIT Initialize

IM Interference Margin

IMT-2020 International Mobile Telecommunications 2020 Standard

IoT Internet of Things

IP Internet Protocol

IPv4 Internet Protocol Version 4

IPv6 Internet Protocol Version 6

ISM Industrial, Scientific, and Medical

k Boltzmann's Constant

LACP Link Aggregation Control Protocol

LAG Link Aggregation Group

LoRa Long Range

LoRaWAN Long Range Wide Area Network

LOS Line-of-Sight

LPWAN Low Power Wide Area Network

LTE Long Term Evolution

μ Mean Value

MAB Multi-Armed Bandit

MAC Medium Access Control

MACPayload Medium Access Control Payload

MCL Maximum Coupling Loss

MCS Modulation and Coding Scheme

MHDR Medium Access Control Header

MIC Message Integrity Code

ML Machine Learning

mMTC Massive Machine Type Communication

MNO Mobile Network Operator

MP-DCCP Multipath Datagram Congestion Control Protocol

MP-QUIC Multipath QUIC

MPTCP Multipath Transmission Control Protocol

MQTT Message Queuing Telemetry Transport

MTC Machine Type Communication

MTU Maximum Transmission Unit

RAT Radio Access Technology

N Number of Measurement Samples

N_0 Noise Spectral Density

N_{Rep} Number of ECL Repetitions

N_{RU} Number of Resource Units

N_{SF} Number of Subframes (in NB-IoT)

NB-IoT Narrowband Internet of Things

NBP Narrowband Bundling Protocol

NF Noise Figure

NLOS Non Line of Sight

NPDCCH Narrowband Physical Downlink Control Channel

NPDSCH Narrowband Physical Downlink Shared Channel

NPUSCH Narrowband Physical Uplink Shared Channel

NRS Narrowband Reference Symbol

NTP Network Time Protocol

NwkSKey Network Session Key

OFDM Orthogonal Frequency Division Multiplexing

OSI Open Systems Interconnection

P_B Bit Error Probability

$P_{CMAX,c(i)}$ Cell-specific Maximum Transmit Power on Slot i

P_I Interference Power

P_N Thermal Noise Power

$P_{N,eff}$ Effective Noise Power

$P_{RX,min}$ Sensitivity

P_S Signal Power

P_{TX} Transmission Power at the Antenna Connector

$P_{TX,ISM}$ Maximum Permissible Transmission Power in the ISM bands

PCB Printed Circuit Board

PDCP Packet Data Convergence Protocol

PDN Packet Data Network

PDU Protocol Data Unit

PAN Personal Area Network

PHB Per-Hop Behavior

PHY Physical Layer

PHYPayload Physical Layer Payload

PHDR Physical Header

PHDR-CRC Physical Header Cyclic Redundancy Check

PLR Packet Loss Ratio

PRB Physical Resource Block

QoS Quality of Service

QPSK Quadrature Phase Shift Keying

R_b Bit Rate

$R_{b,eff}$ Effective Bit Rate

R_C Chip Rate

R_{max} Maximum Data Rate

RB Resource Block

R_{Code} Rate Code

RE Resource Element

RF Radio Frequency

RLC Radio Link Control

RobE Robust Establishment

ROHC Robust Header Compression

RSRP Reference Signal Received Power

RSRQ Reference Signal Received Quality

RSSI Received Signal Strength Indicator

RU Resource Unit

RX receiver

σ Standard Deviation

SC-FDMA Single Carrier Frequency Division Multiple Access

SCTP Stream Control Transmission Protocol

SDN Software Defined Networking

SF Spreading Factor

SFD Start of Frame Delimiter

SID Subflow Identifier

SINR Signal to Interference and Noise Power

SNR Signal to Noise Ratio

SRD Short Range Devices

SYN Synchronize

T Temperature

T_b Time per bit

T_C Time per Chip

$T_{interval}$ Packet Interval

$T_{L,buf}$ Buffer latency

$T_{L,NBP}$ Latency of an NBP Connection

$T_{L,PN}$ Latency of Path N

T_{min} Minimum Transmission Time

T_{RU} Transmission Time of one Resource Unit

$T_{RXDelay}$ Delay Before LoRaWAN RX Window

T_{SF} Transmission Time of one Subframe (in NB-IoT)

TBS Transport Block Size

tc Traffic Control

TCP Transmission Control Protocol

ToA Time on Air

ToS Type of Service

TX transmitter

UDP User Datagram Protocol

UE User Equipment

UL uplink

UMTS Universal Mobile Telecommunications System

URLLC Ultra-reliable and Low Latency Communication

USB Universal Serial Bus

VINF Virtual Network Interface

VM Virtual Machine

VPN Virtual Private Network

WAN Wide Area Network

WSN Wireless Sensor Network

x Subcarrier Activity Factor

Part I

STATE OF THE ART AND RESEARCH APPROACH

1 INTRODUCTION

Over the last decade, the IoT has become an integral part of everyday life. Countless devices sense their surroundings and interact with each other and the cloud, producing novel traffic patterns known as Machine Type Communication (MTC). There is a great variety of IoT use cases, ranging from high-bandwidth applications such as video surveillance to small battery powered devices such as temperature sensors. In general, two main groups of communication patterns can be identified within the MTC category: Critical Machine Type Communication (cMTC), which is also referred to as Ultra-reliable and Low Latency Communication (URLLC), and Massive Machine Type Communication (mMTC) [44]. One of the most common mMTC use cases are Wireless Sensor Networks (WSNs), which consist of large fleets of battery powered devices sensing environmental parameters and reporting their measurements to the cloud. Many such devices are mounted in hard to reach places, such as on streetlamps or in basements; therefore, manufacturers aim for long battery life times up to 10 years. Due to the large number of nodes typically present in a WSN network, installing large and costly batteries is generally not an option; rather, the energy consumption must be reduced to the absolute minimum necessary to fulfill the intended purpose. The resulting devices are constrained in terms of energy and processing power, which must be considered when designing protocols and applications. One of the main factors to think about when designing for energy efficiency is wireless network connectivity. Over the years, many approaches have been developed ranging from Personal Area Networks (PANs) to mesh networks, depending on the geographical area in which a WSN is deployed.

In recent years, a new group of wireless access technologies has emerged: LPWANs provide an energy efficient, long range and low cost communication interface for mMTC applications, which allows distributing WSNs over large geographical areas. Since there is a wide variety of mMTC use cases, an equally large variety of LPWANs has been developed to address individual network performance needs, which is called the QoS. As a result, operators need to select an LPWAN that provides an appropriate QoS level for the use case in question. LPWANs can be categorized into two main groups based on the spectrum they occupy. Licensed technologies are deployed by MNOs along existing networks such as LTE. In contrast, unlicensed technologies work in the ISM and SRD bands and can be deployed freely within the limits of local regulation. In Europe, the bands most commonly used for LPWANs are the 433 MHz ISM Band (EU-433) and the 868 MHz SRD Band (EU-868); within each band, there are sub-bands which are regulated in terms of the maximum permissible transmission power and duty cycle [20].

One of the most popular licensed LPWAN technologies is NB-IoT, which provides a long-range, energy efficient mobile network service with up to 10 years of battery life. It is a clean-slate technology that reuses mechanisms found in LTE, but is

not compatible with traditional mobile networks[1]. NB-IoT is attractive for low- to mid-range mMTC use cases by providing a large Maximum Coupling Loss (MCL) of 164 dB and up to 10 years of battery life, while simultaneously achieving comparatively high data rates of tens of kilobits per second. Its support for massive device densities and integration into the 5G architecture ensures long-term relevance [44].

On the side of unlicensed LPWANs, LoRaWAN is one of the most widespread candidates. This LPWAN technology employs a Chirp Spread Spectrum (CSS) based modulation called Long Range (LoRa) to mitigate the interference in the ISM and SRD bands [50]. The higher layers are handled by an open protocol called LoRaWAN [1], which is commonly used as a synonym for the entire stack. LoRaWAN targets low-end IoT applications exclusively; similar to NB-IoT, it provides a high MCL at 151 dB and long battery life times of multiple years, but it is limited severely in terms of the data rate by the duty cycle regulations.

While the wide range of QoS provided by different LPWANs allows selecting an appropriate technology for most use cases, it still does not address devices with volatile QoS requirements. For example, a device may send regular measurement data and occasional bursts of alarm messages; in this case, the LPWAN selection depends on the maximum QoS that may be required. As a result, all data must be sent using a high performance link, which increases energy consumption and cost. In traditional networks, such challenges can be addressed by a bundling protocol, which combines multiple network links into one logical connection; only when the preferred link can no longer provide the desired QoS a second link will be used. One of the most popular bundling protocols is Multipath Transmission Control Protocol (MPTCP), which is an extension of the classic Transmission Control Protocol (TCP). Similar to its base protocol, MPTCP provides a reliable, congestion controlled and in-order transport service; additionally, it allows applications making use of multiple network paths[2] and can fall back to regular TCP for hosts that are not multipath capable. There are many other protocols that focus on different use cases, e.g., Multipath Datagram Congestion Control Protocol (MP-DCCP) provides non-reliable delivery. What is common to the existing bundling protocols is that they are not optimal for the use over narrowband and unreliable networks, either due to their overhead or their reliable and in-order characteristics. However, a purpose-built LPWAN bundling protocol could provide considerable QoS benefits.

1.1 HYPOTHESES AND OBJECTIVES

In this thesis, the applicability of bundling protocols to LPWAN technologies is explored. There are two central questions that will be analyzed:

Hypothesis 1: The QoS provided by different LPWANs depends on their internal mechanisms (e.g., coverage extension techniques) and the radio conditions.
LPWANs make use of different Coverage Extension (CE) mechanisms, such as signal

1 The NB-IoT standard provides multiple strategies for co-existence with previously deployed mobile networks, which are explained in Section 2.2.2.

2 A path usually corresponds to a physical network link, however this is not strictly necessary; it is possible to establish multiple subflows (i.e., flows that are part of a multipath connection) over one link.

repetitions and variable symbol times, to a much larger degree than traditional mobile networks. Therefore, it can be assumed that these algorithms are responsible for at least part of the observed QoS differences between individual LPWAN technologies. Furthermore, the radio conditions such as signal level and external interference could influence the QoS observed in different scenarios.

Hypothesis 2: Using a novel LPWAN bundling protocol, multiple LPWANs can be combined to improve the QoS and alleviate drawbacks.
While traditional bundling protocols are not applicable to LPWANs due to their overhead or other undesirable characteristics, a special purpose LPWAN bundling protocol could improve the QoS and address use cases such as mMTC applications with varying QoS requirements. In the LPWAN context, a number of special considerations are necessary, e.g., addressing endpoints over LPWAN links that do not support IP.

From the Hypotheses, two objectives are derived that define specific steps to address the research questions.

Objective 1: Analysis and comparison of the most relevant LPWAN technologies
The first objective includes a systematic analysis of NB-IoT and LoRaWAN, which have been selected as representatives of a licensed and an unlicensed LPWAN due to their popularity.

NB-IoT employs up to 2048 signal repetitions for CE, which are coherently added in the receiver. This improves the sensitivity and thus the coverage range of the network. It is expected that the time needed for these transmissions affects the QoS considerably; therefore, the radio conditions that trigger the mechanism are of special interest. In Chapter 4, NB-IoT is evaluated in terms of physical and application layer QoS parameters in a commercial mobile network. Furthermore, the results are compared against theoretical specifications and existing simulations and the QoS is discussed in different smart metering use cases. The key characteristics observed in the NB-IoT QoS analysis are considered during the NBP design phase as an example of a licensed, IP-capable LPWAN technology. In the future, the raw data obtained from the measurements could be used to derive scheduling algorithms.

LoRaWAN employs Spreading Factors (SFs) for CE; each SF step doubles the symbol duration, but only encodes one additional bit. Therefore, higher SFs decrease the expected data rate. In Chapter 5, LoRaWAN is analyzed using the methodology developed during the NB-IoT measurements. Since LoRaWAN networks provide extensive configuration options, the QoS can be explored for a wide signal level range at individual SFs. Special consideration is given to how the regulations in the unlicensed bands limit the QoS. The results are compared to the theoretical specifications and LoRaWAN is discussed for smart metering use cases. Afterwards, LoRaWAN is compared to NB-IoT for central QoS parameters. The results of this evaluation are considered in the NBP design phase as an example for an unlicensed, non-IP LPWAN technology. Similar to the NB-IoT measurements, the raw data could be used to construct scheduling algorithms in the future.

Objective 2: Development of a novel bundling protocol that enables the efficient aggregation of different LPWAN technologies
The second objective consists of the design, implementation, and evaluation of an LPWAN-specific bundling protocol. Furthermore, a control mechanism is designed that allows applications communicating their QoS needs to NBP. Special consideration is given to backwards compatibility with the existing IoT ecosystem.

The evaluations of NB-IoT and LoRaWAN have revealed considerable QoS differences between the individual LPWAN technologies. Especially the duty cycle regulations drastically limit the QoS of LPWANs operating in the unlicensed bands. At the same time, these technologies provide a cost effective transport service for small amounts of data. Therefore, an approach that combines both licensed and unlicensed technologies could improve the QoS and reduce costs. In Chapter 6, a novel LPWAN-specific bundling protocol is presented: the *Narrowband Bundling Protocol* provides an efficient multipath transport service to IoT devices. It works over both IP and non-IP paths that are constrained in terms of data rate, payload size and energy by reducing overhead wherever possible. NBP aims to improve QoS characteristics such as latency, reliability and coverage by bundling multiple LPWAN links. Furthermore, a proxy architecture is proposed that enables legacy applications to continue working by intercepting UDP packets. NBP is evaluated in simulations and a field test and future protocol extensions are proposed. Finally, a mechanism is presented that allows applications communicating per-packet QoS requirements (e.g to mark a packet as emergency data and ensure priority treatment).

1.2 SCIENTIFIC CONTRIBUTIONS

Part of this thesis has previously appeared in journal papers. The related publications are listed hereafter:

[1] Andreas Philipp Matz, Jose-Angel Fernandez-Prieto, Joaquin Canada-Bago, and Ulrich Birkel. "The Narrowband Bundling Protocol." In: *IEEE Wireless Communications Letters* 11.9 (2022), pp. 1900–1904. DOI: 10.1109/LWC.2022.3186392.

[2] Andreas Philipp Matz, Jose-Angel Fernandez-Prieto, Joaquin Cañada-Bago, and Ulrich Birkel. "A Systematic Analysis of Narrowband IoT Quality of Service." en. In: *Sensors* 20.6 (Mar. 2020), p. 1636. ISSN: 1424-8220. DOI: 10.3390/s20061636.

1.3 ORGANIZATION OF THE THESIS

This thesis consists of three parts. In the first part, the state of the art and the research approach is presented. The second part details the research development, describing all the major findings of this thesis. Finally, in the third part an overall conclusion is drawn and future work opportunities are explored.

Chapter 1 provides a brief overview of the major topics of this work. The thesis hypotheses are stated, and the scope and scientific contributions of this works are explained.

Chapter 2 describes the physical properties and the protocol stack of NB-IoT and LoRaWAN. These fundamentals are essential for the analysis of the QoS measurement results.

Chapter 3 introduces the history and principles of bundling protocols. The basic components of a typical bundling architecture are explained and considerations for creating a new bundling protocol are discussed.

Chapter 4 contains a systematic analysis of the QoS of NB-IoT. Physical and application layer measurements are conducted and the results are compared to the theoretical specifications. Afterwards, the applicability of NB-IoT is discussed in typical smart metering use cases.

Chapter 5 presents a similar analysis for LoRaWAN. Furthermore, LoRaWAN is compared to NB-IoT in different QoS aspects.

Chapter 6 introduces the Narrowband Bundling Protocol, which is an LPWAN-specific multipath protocol that provides a low-overhead transport service for IoT applications.

Chapter 7 retrospects on the research performed in this thesis and highlights future research opportunities.

2

LOW POWER WIDE AREA NETWORKS

Before an in-depth analysis of NB-IoT and LoRaWAN can be performed, it is necessary to recall the theoretical fundamentals of operation, physical properties and the protocol stack of the two technologies. This information is helpful in designing a measurement procedure and discuss the observed results. Only the relevant concepts are explained; sources for further reading on specific topics are listed in the respective sections.

2.1 PROPERTIES OF WIRELESS RECEIVERS

When transmitting information over a wireless channel, a certain Bit Error Rate (BER) must be maintained to correctly decode the signal. The BER depends on the Signal to Noise Ratio (SNR); with decreasing SNR, the BER increases. For purposes of signal analysis, it is useful to consider the ratio of the bit energy and the noise power spectral density E_b/N_0, which is a normalized form of SNR [92]. E_b can be written as the signal power P_S multiplied with the bit time T_b; N_0 describes the noise power P_N per bandwidth B.

$$\frac{E_b}{N_0} = \frac{P_S \cdot T_b}{P_N/B} = \frac{P_S/R_b}{P_N/B} = \frac{P_S}{P_N} \cdot \frac{B}{R_b} = SNR \cdot \frac{B}{R_b} \tag{1}$$

with:
$R_b = 1/T_b$: Bit Rate
$SNR = P_S/P_N$: Signal to Noise Ratio

The significance of this parameter can be seen in the *waterfall curves*, which plot the Bit Error Probability P_B versus E_b/N_0. These diagrams are commonly used to quickly assess the performance of a digital communication system; a system that requires a lower E_b/N_0 to detect the signal at a given error probability demodulates the data more efficiently [92].

The receiver (RX) sensitivity ($P_{RX,min}$) describes the minimum input power level at the receiver antenna port related to a QoS threshold. This threshold is defined based on the communication system in question; it could relate to the BER or the Block Error Rate (BLER). To calculate $P_{RX,min}$, the thermal noise needs to be considered. For the analysis in this thesis, it is useful to choose a logarithmic representation:

$$P_N\,(dBm) = 10 \cdot \log\left(\frac{k \cdot T \cdot B}{1\,mW}\right) = 10 \cdot \log\left(\frac{k \cdot T}{1\,mW}\right) + 10 \cdot \log(B) \tag{2}$$

where k is the Boltzmann constant, T is the temperature, and B is the bandwidth.

At room temperature (T = 20 °C = 293.15 K), this can be written as:

$$P_N\,(dBm) = N_0 + 10 \cdot \log(B) \tag{3}$$

with:
$N_0 = -174\ \frac{dBm}{Hz}$: Logarithmic thermal noise spectral density

The effective noise floor ($P_{N,eff}$) takes into account the Noise Figure (NF) of the receiver front end, which adds noise to the thermal Noise Power (P_N):

$$P_{N,eff} = P_N + NF \tag{4}$$

Furthermore, a minimum SNR is needed to achieve the aforementioned QoS threshold.

$$SNR_{min} = \frac{E_b}{N_0} \cdot \frac{R_b}{B} \tag{5}$$

When thinking in terms of receiver sensitivity, it is useful to consider the logarithmic representation of the SNR_{min}.

$$\begin{aligned} SNR_{min}\,(dB) &= \left(\frac{E_b}{N_0}\right)\,(dB) + 10 \cdot \log\left(\frac{R_b}{B}\right) \\ &= \left(\frac{E_b}{N_0}\right)\,(dB) - 10 \cdot \log(B \cdot T_B) \end{aligned} \tag{6}$$

This relationship reveals the two major strategies how communication systems can improve the SNR_{min}. The first possibility is increasing the system bandwidth B, as it is common in spread spectrum modulation. The second option is increasing the bit time T_B, which can be done either by extending the bits in time as employed by LoRa or using coherent addition of signal repetitions as in NB-IoT.

Using the previous definitions, the receiver sensitivity ($P_{RX,min}$) is defined as:

$$P_{RX,min} = P_N + NF + SNR_{min} = P_{N,eff} + SNR_{min} \tag{7}$$

As a result, there are different ways to optimize $P_{RX,min}$:

- Decrease the P_N by reducing the bandwidth or the temperature.
- Decrease the SNR_{min} by increasing the bandwidth or bit time.
- Decrease the noise figure through optimized receiver design.

While some of these points appear to contradict each other, the optimization strategy depends on the use case in question. For example, LPWANs must reduce the power consumption to support battery powered applications. Therefore, they generally try to minimize P_N by choosing a narrow system bandwidth, which conserves power at the expense of a low data rate. Since most LPWAN modems are built to a minimum cost, it is difficult to achieve significant NF improvements. Therefore, the only remaining option to improve $P_{RX,min}$ and thus extend the coverage is increasing the T_B by extending the bits in time or using coherently added repetitions.

Finally, another important measurement is the MCL. It is a common metric to evaluate the coverage of a radio access technology and is calculated as the difference between the transmitted power level at the antenna connector (P_{TX}) and the receiver sensitivity.

$$MCL = P_{TX} - P_{RX,min} \tag{8}$$

2.2 FUNDAMENTALS OF NB-IOT

Narrowband IoT is a clean-slate LPWAN standard specified by the Third Generation Partnership Project (3GPP) standardization organization.[1] It was introduced in LTE Release 13 [74], with significant improvements made in Release 14 and 15 [80]. NB-IoT is designated as a 5G mMTC enabler in Release 15 [45] and will thus stay relevant in future mobile networks. In this section, only the fundamental aspects of the NB-IoT Release 13 standard[2] are introduced, which are of relevance for the analysis in this thesis. Additional details can be found in [44, 64].

2.2.1 *Introduction*

In recent years, the rise of the IoT and its associated platforms has directed the interest of researchers and companies alike to the exchange of information between IoT endpoints such as sensors and servers. These communication patterns are referred to as Machine Type Communication [44]. The communication requirements within the MTC category are diverse; therefore, it is common to distinguish between cMTC and mMTC technologies. A widespread mMTC use case is WSNs, which consist of many battery powered devices distributed across heterogeneous environments.

The diverse applications of WSNs have sparked the development of a new group of IoT-centric access technologies. LPWANs provide long-range communication for many devices simultaneously, while enabling battery life for many years. In general, LPWANs can be categorized based on the license of the occupied spectrum. Unlicensed technologies, such as LoRaWAN [51], work in the license-free and region-specific ISM and SRD bands. Licensed technologies are standardized by an industry consortium such as the 3GPP [78] and are commonly deployed by MNOs along with existing mobile networks. Some LPWANs are developed based on existing standards, such as LTE Cat-M1 [44]; others, like NB-IoT, [44] are referred to as clean-slate technologies and do not retain full compatibility with existing technologies. As a result, these standards interfere with neighboring networks.

NB-IoT has emerged from a 3GPP study on cellular IoT technologies [72] and was first standardized as part of LTE Release 13 [74]. It is designed to integrate with existing Global System for Mobile Communications (GSM) and LTE networks and provide a long-range, energy-efficient mobile service, with a MCL of 164 dB and up to 10 years of battery life. Additionally, NB-IoT supports massive device densities and guarantees delivery of high-priority exception reports within 10 seconds. In

1 Section 2.2 is an excerpt from the previously published journal paper "A Systematic Analysis of NB-IoT Quality of Service" by A. Matz et al. [57].

2 The theory and evaluations presented in this thesis are based on 3GPP LTE Release 13, since this version was deployed in Germany during the measurements.

the future, NB-IoT will be part of the 5G architecture, providing support for mMTC use cases [45].

Whenever new mobile network standards are introduced, their performance is approximated in system-level simulations, which include the air interface, as well as higher protocol layers. In the case of NB-IoT, extensive simulations have already been performed for different scenarios [44, 45].

NB-IoT Release 13 has been implemented by MNOs and gained considerable interest; however, little attention has been paid to systematic real-life QoS evaluations. While some measurements have been done, not all of them were performed in real-world environments, which is essential to verify the end-user QoS. Mozny et al. [62] analyzed attachment delay, latency, and message overhead under laboratory conditions. The analysis was based on five samples per measurement, which is a low sample size that may limit statistical relevance. Khan et al. [38] performed long-term analysis of NB-IoT SNR and Received Signal Strength Indicator (RSSI) on different floors of a university building. The evaluation showed that SNR and RSSI levels changed during the day. The authors proposed that this may be caused by LTE activity generated by mobile network users. Malik et al. [53] analyzed NB-IoT coverage in indoor, outdoor, and underground scenarios. The results indicated that NB-IoT worked well for indoor and outdoor scenarios, while underground coverage was limited to locations close to the base station, which is referred to as Evolved Node B (eNodeB) in LTE systems. Martinez et al. [54] performed analysis on energy consumption, reliability, and latency in NB-IoT networks. While these measurements were performed in an operative network, the authors focused on the power consumption of NB-IoT, and they found no correlation between the repetitions caused by the Enhanced Coverage Level (ECL) and latency. A thorough search of the relevant literature revealed that only Basu et al. [14] have performed detailed application layer QoS measurements in a commercial NB-IoT network. The results showed that decreasing signal levels resulted in increased latency and a reduction in throughput, but overall indicated a reliable performance of NB-IoT. However, the work considered a single tone downlink scenario, which did not align with the NB-IoT specification. Furthermore, only limited analysis has been performed on the physical NB-IoT network parameters and their influence on the application layer QoS.

This work expands on the previous publications by systematically examining both physical and application layer NB-IoT QoS parameters in a commercial mobile network. The influence of the physical QoS parameters and the underlying mechanisms on the end-user QoS is analyzed and compared against theory, as well as simulations. A long-term measurement allows concluding on how the radio conditions change between day and night. Moreover, this thesis presents the suitability of NB-IoT in four use cases of smart metering applications in the uplink (UL) and downlink (DL) direction.

To the best of the authors' knowledge, this thesis contains one of the first systematic analyses of the relevant physical and application layer QoS parameters and contributing factors in a real NB-IoT network and is the first study of NB-IoT QoS in Germany.

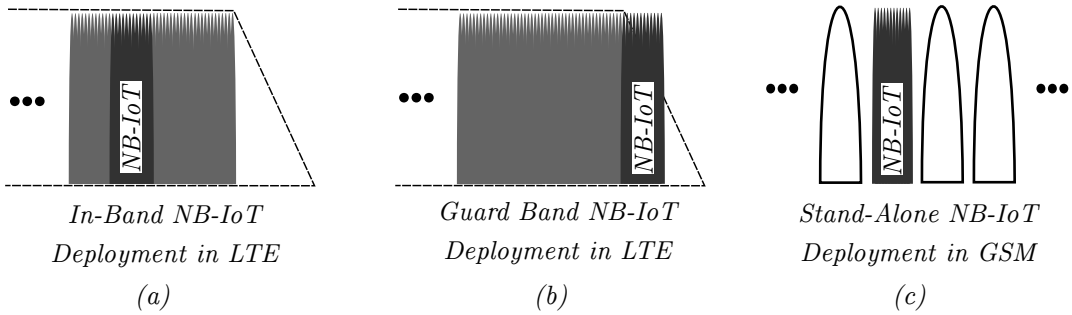


Figure 1: NB-IoT allows flexible deployment along existing mobile networks: **(a)** LTE in-band deployment, **(b)** LTE guard band deployment, **(c)** GSM stand-alone deployment (adapted from [85]).

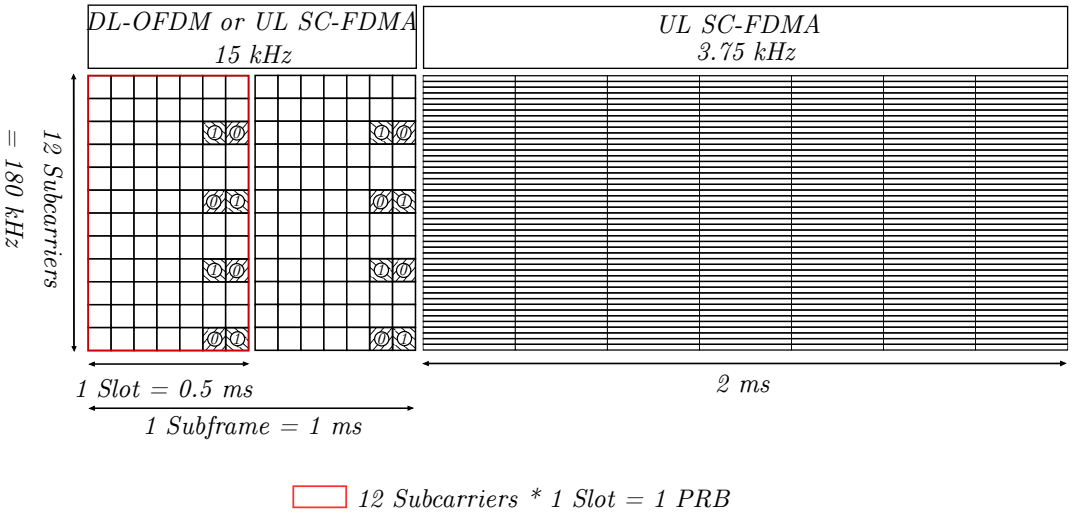


Figure 2: LTE resource grid. The NB-IoT downlink is based on 15 kHz carrier spacing and OFDM, carrying reference symbols. The uplink is based on 15 kHz or 3.75 kHz carrier spacing and SC-FDMA.

2.2.2 *Deployment Modes and Resource Grid*

NB-IoT occupies a total system bandwidth of 180 kHz and can be deployed in-band within a single LTE Physical Resource Block (PRB) or operate in the LTE guard band on each side of the LTE carrier. Alternatively, it can occupy a single GSM carrier of 200 kHz in stand-alone mode as shown in Figure 1.

The integration of NB-IoT into existing Fourth Generation (4G) networks is facilitated by reusing the LTE resource grid as shown in Figure 2. The NB-IoT downlink is based on Orthogonal Frequency Division Multiplexing (OFDM) with 15 kHz carrier spacing and always uses all 12 subcarriers. Downlink resources are allocated as subframes, which span 1 ms and contain two slots with seven OFDM symbols each. As such, each allocated subframe contains two PRBs. The downlink radio channel is estimated using Narrowband Reference Symbols (NRSs) that are present in the last two OFDM symbols of each slot. In the uplink direction, Single Carrier Frequency Division Multiple Access (SC-FDMA) is used, and the resource grid can be configured to use either 15 kHz or 3.75 kHz carrier spacing.

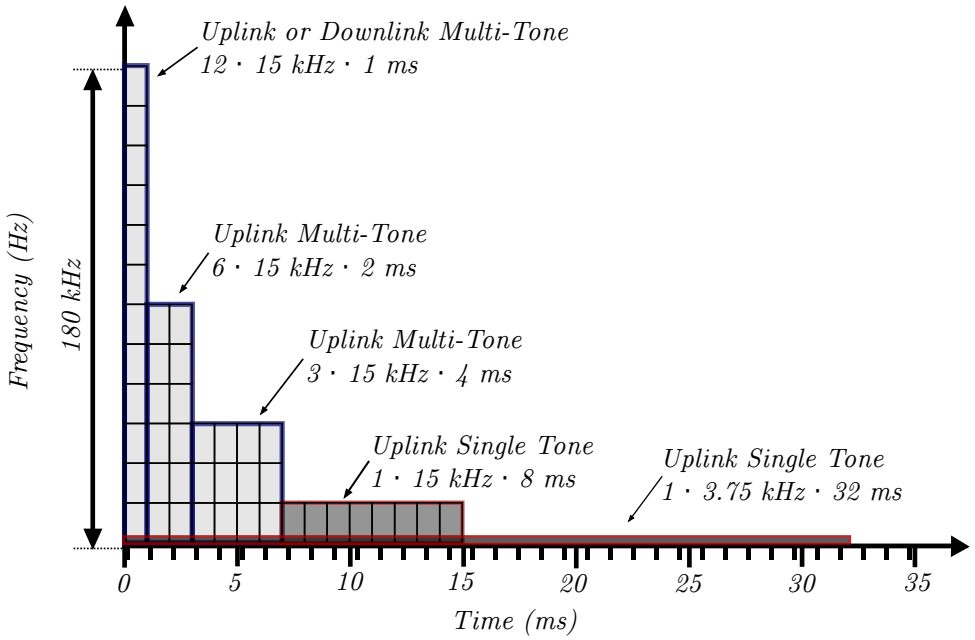


Figure 3: Resource allocation in NB-IoT uplink and downlink directions. Multi tone allocations are highlighted in light gray with blue borders, and single tone allocations are dark gray with red borders.

Additionally, NB-IoT introduces the concept of Resource Units (RUs) in the uplink, which allows allocating one or more subcarriers of a single PRB to different User Equipments (UEs). These subcarriers are referred to as tones, and the resulting configurations are called single and multi-tone operation. The possible allocation schemes [75] are illustrated in Figure 3.

2.2.3 *NB-IoT Physical Layer Radio Measurements*

The LTE specification defines a set of radio QoS parameters, which are used to estimate the channel between the eNodeB and the UE [73]. This section introduces the parameters that are most relevant to evaluate end-user QoS, as well as their relationship to each other.

- The *Reference Signal Received Power (RSRP)* is a linear average of the power of the resource elements carrying NRS in a given frequency bandwidth, expressed in Watts. Since NB-IoT downlink is based on a 15 kHz carrier spacing, the RSRP is the power of a single 15 kHz NRS.
- The *Received Signal Strength Indicator (RSSI)* is a linear average of the total power received in the measurement bandwidth from all sources, including external interference, noise, and others, expressed in Watts. In NB-IoT, the measurement bandwidth is exactly one PRB or 180 kHz. The RSSI depends on the cell load; it increases with the number of allocated subcarriers.

- The *Reference Signal Received Quality (RSRQ)* describes the ratio between RSRP and RSSI, where both measurements shall be made over the same set of resource blocks:

$$\mathrm{RSRQ} = \frac{\mathrm{RSRP}\,(W)}{\mathrm{RSSI}\,(W)} \tag{9}$$

- The *Signal to Interference and Noise Power (SINR)* is the ratio between the received signal level and the interference power (P_I) from external sources, as well as the effective noise power ($P_{N,eff}$). If all 12 Resource Elements (REs) of the broadband channel are occupied with the same signal power as the NRS, the narrowband and broadband SINR are identical.

$$\begin{aligned}\mathrm{SINR} &= \frac{\mathrm{RSRP}\,(W)}{P_{I,15\,\mathrm{kHz}}\,(W) + P_{N,eff,15\,\mathrm{kHz}}\,(W)} \\ &\approx \frac{\mathrm{RSSI}\,(W)}{P_{I,180\,\mathrm{kHz}}\,(W) + P_{N,eff,180\,\mathrm{kHz}}\,(W)}\end{aligned} \tag{10}$$

The SINR and RSRQ parameters are related to each other via the subcarrier activity factor x, which defines the ratio of occupied REs in a Resource Block (RB): $x = \mathrm{RE}/\mathrm{RB}$. In an unloaded NB-IoT cell, only the NRS are active, so $\mathrm{RE} = 2$ and $x = 2/12$, while a fully loaded cell uses all subcarriers ($x = 1$).

The relation between the parameters SINR, RSRQ, and the number of occupied REs or the subcarrier activity factor was derived in [65]:

$$\mathrm{SINR} = \frac{12}{\frac{1}{\mathrm{RSRQ}} - \mathrm{RE}} = \frac{12}{\frac{1}{\mathrm{RSRQ}} - 12 \cdot x} \tag{11}$$

This relationship is validated in Section 4.2.4 by analyzing the RSRQ and SINR measurements of the NB-IoT UE, which allows concluding on the subcarrier activity factor x.

2.2.4 *Data Rate and Latency Boundaries*

NB-IoT dynamically adjusts to the radio conditions by configuring the Modulation and Coding Scheme (MCS), which is a combination of a modulation type and a coding rate applied to a given PRB. NB-IoT supports MCS 0 to 12, using Quadrature Phase Shift Keying (QPSK) or Binary Phase Shift Keying (BPSK) modulation [75] and a variable Transport Block Size (TBS). Higher MCS implies reduced coding redundancy and thus provides increased TBS at the same number of RUs. Additionally, the coverage range can be improved by applying signal repetitions N_{Rep}, which increase the receiver sensitivity. User data are transmitted via two physical channels, which are called Narrowband Physical Uplink Shared Channel (NPUSCH) in format F1 (NPUSCH F1) and Narrowband Physical Downlink Shared Channel (NPDSCH).

When calculating an upper bound for the data rate in the uplink (NPUSCH F1) and downlink (NPDSCH) direction, the maximum possible MCS and TBS must be considered, which correlates to the minimum number of resource units (N_{RU}) and subframes (N_{SF}) used for the transmission [44]:

$$\text{NPUSCH single tone}: \ TBS_{max} = 1000\,\text{bits},\ MCS_{max} = 10 \rightarrow N_{RU,min} = 6 \tag{12}$$

$$\text{NPUSCH multi tone}: \ TBS_{max} = 1000\,\text{bits},\ MCS_{max} = 12 \rightarrow N_{RU,min} = 4 \tag{13}$$

$$\text{NPDSCH}: \ TBS_{max} = 680\,\text{bits},\ MCS_{max} = 12 \rightarrow N_{SF,min} = 3 \tag{14}$$

Afterwards, the minimum time needed to transmit the maximum TBS without repetitions ($N_{Rep} = 1$) can be calculated. In the uplink, the time needed to transmit one RU (T_{RU}) is 1 ms for a 180 kHz multi-tone transmission and 8 ms in the case of a 15 kHz single tone transmission. In the downlink direction, the time needed to transmit one subframe (T_{SF}) is 1 ms.

The minimum time needed to transmit 1000 bits in uplink (NPUSCH F1) is:

$$T_{min,15\,kHz} = N_{rep} \cdot N_{RU,min} \cdot T_{RU} = 1 \cdot 6 \cdot 8\,ms = 48\,ms \quad \text{for 15 kHz single tone} \tag{15}$$

$$T_{min,180\,kHz} = N_{rep} \cdot N_{RU,min} \cdot T_{RU} = 1 \cdot 4 \cdot 1\,ms = 4\,ms \quad \text{for 180 kHz multi-tone} \tag{16}$$

The minimum time needed to transmit 680 bits in downlink (NPDSCH) is:

$$T_{min,180\,kHz} = N_{rep} \cdot N_{SF,min} \cdot T_{SF} = 1 \cdot 3 \cdot 1\,ms = 3\,ms \quad \text{for a 180 kHz PRB} \tag{17}$$

NB-IoT employs a scheduling cycle to allocate resources to the individual UEs, which reduces the Physical Layer (PHY) effective data rates and results in the Medium Access Control (MAC) Layer data rates. This scheme is illustrated in Figure 4a for an uplink and in Figure 4b for a downlink transmission. Periodic scheduling information is transmitted on the Narrowband Physical Downlink Control Channel (NPDCCH) and user data are sent via NPUSCH F1 and NPDSCH. NB-IoT uses Hybrid Automatic Repeat Request (HARQ) feedback for forward error correction and automatic retransmissions, which is transferred in the uplink on the NPUSCH in Format F2. The green dashed line marks the earliest possible start of a new scheduling cycle and defines the minimum time needed for the transmission of the TBS.

Accordingly, the upper boundaries for the MAC layer data rates can be calculated. These values are used as a benchmark for the throughput evaluations in Section 4.2.6 of this thesis.

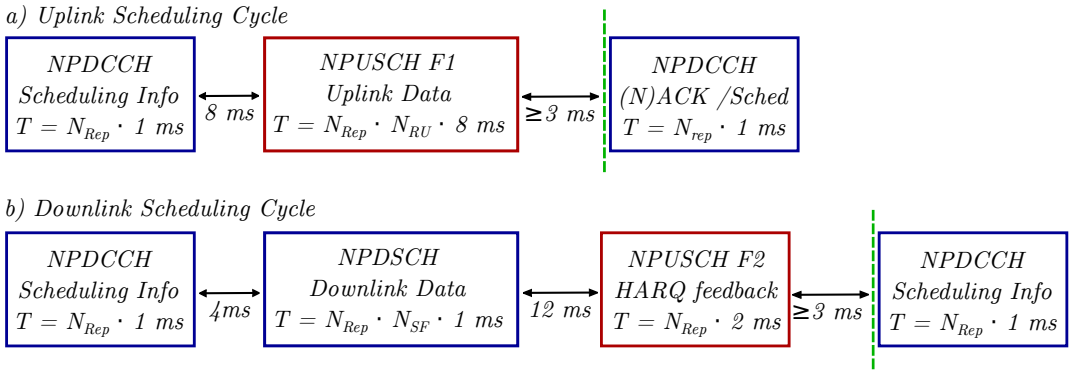


Figure 4: NB-IoT scheduling cycle for a single **(a)** uplink and **(b)** downlink transmission. The green dashed line marks the earliest possible start of a new scheduling cycle (adapted from [45]).

The maximum physical data rate (R_{max}) in NPUSCH is calculated as follows:

$$\begin{aligned} T_{min,15\,kHz} = 60\,ms, TBS = 1000\,bit \rightarrow R_{max} = TBS/T_{min} \\ = 16.7\,kbps \text{ for } 15\,kHz \text{ single tone} \end{aligned} \tag{18}$$

$$\begin{aligned} T_{min,180\,kHz} = 16\,ms, TBS = 1000\,bit \rightarrow R_{max} = TBS/T_{min} \\ = 62.5\,kbps \text{ for } 180\,kHz \text{ multi-tone} \end{aligned} \tag{19}$$

On the other hand, R_{max} in NPDSCH is:

$$\begin{aligned} T_{min} = 25\,ms, TBS = 680\,bit \rightarrow R_{max} = TBS/T_{min} \\ = 27.2\,kbps \text{ for a } 180\,kHz \text{ PRB} \end{aligned} \tag{20}$$

In NB-IoT, there are multiple factors that contribute to the total system latency. Liberg et al. [44] have performed system-level simulations to evaluate the latency under different signal conditions and deployment modes, as shown in Table 1. The simulation results indicate that at 144 dB coupling loss, NB-IoT provided a low latency of about 300 ms, which was mostly determined by the time needed to access the network and acquire the necessary configuration information. At the MCL of 164 dB, the latency was dominated by the signal repetitions used for CE, but still met the limit of 10 s for delivering exception reports, as defined by 3GPP [72]. The differences between the individual deployment modes could be attributed to the output power restrictions that apply for in-band and guard band installations. Overall, the latency estimation can serve as a useful benchmark for real-life evaluations.

2.2.5 *Receiver Sensitivity, ECL, MCL, and Transmission Power Control in NB-IoT*

The receiver sensitivity ($P_{RX,min}$) defines the minimum input power level at the receiver antenna port related to a QoS threshold as defined in Equation (7). In NB-IoT, this threshold is commonly specified at a Block Error Rate (BLER) of 10 % for the NPUSCH and NPDSCH NB-IoT channels [44]. The BLER depends on the MCS and the amount of symbol repetitions (N_{Rep}) applied for a given SNR. Accordingly,

Table 1: NB-IoT exception report latency (adapted from [44]).

Coupling Loss (dB)	Stand-Alone Mode Latency (s)	Guard Band Mode Latency (s)	In-Band Mode Latency (s)
144	0.3	0.3	0.3
154	0.7	0.9	1.1
164	5.1	8.0	8.3

SNR_{min} defines the minimum power ratio needed in order to achieve this QoS target.

The analysis of the measurement results requires the calculation of the effective noise floor of the RSRP, the RSSI, and the 15 kHz single-tone NPUSCH. These can be calculated at room temperature, considering the corresponding bandwidths B and the assumed NFs of UE and eNodeB:

$$P_{N,eff,NPUSCH} = P_{N,15\,kHz} + NF_{eNodeB} = -132.2\,dBm + 5\,dB = -127.2\,dBm \tag{21}$$

$$P_{N,eff,RSSI,DL} = P_{N,180\,kHz} + NF_{UE} = -121.4\,dBm + 7\,dB = -114.4\,dBm \tag{22}$$

$$P_{N,eff,RSRP,DL} = P_{N,15\,kHz} + NF_{UE} = -132.2\,dBm + 7\,dB = -125.2\,dBm \tag{23}$$

In NB-IoT, the receiver sensitivity is improved by applying signal repetitions [75]. The sensitivity is expected to improve by 3 dB for each doubling of the repetitions due to coherent addition of the symbols and incoherent addition of thermal noise. NB-IoT employs up to 128 repetitions in the uplink and 2048 repetitions in the downlink direction. The large number of repetitions in downlink compensates the higher effective downlink noise floor in order to balance the link budget. However, accumulating the signal repetitions takes time and is thus a trade off between latency and sensitivity.

The appropriate number of repetitions is adjusted dynamically by the eNodeB, which assigns an ECL to each device based on the received uplink and reported downlink signal level, where a higher ECL corresponds to more problematic radio conditions and a higher number of repetitions. The exact number of repetitions per ECL is defined by the MNO. This thesis refers to the three available ECL levels as ECL 0, ECL 1, and ECL 2.

The MCL is a common metric to evaluate the coverage of a radio access technology and is calculated as the difference between the transmitted power level at the antenna connector (P_{TX}) and the receiver sensitivity. Since the downlink NRS are used for channel estimation, their constant transmit power is indicated to the UE by the eNodeB and used for all physical channels. In the uplink direction, the transmission power depends on the coverage situation. For up to two repetitions, the

Table 2: Reference link budget configuration achieving the minimum requirement of MCL = 164 dB for 5G mMTC systems as defined by IMT-2020 [34] with 32 and 128 repetitions in uplink and downlink (adapted from [45]).

	NPUSCH 164 dB MCL Ref. Performance	**NPDSCH 164 dB MCL Ref. Performance**
Transmit Power P_{TX} (dBm)	23	35
TBS (bits)	1000	680
Repetitions N_{Rep}	32 (max. 128)	128 (max. 2048)
Resource Units N_{RU}	10	8
BLER (%)	10	10
SNR_{min} (dB)	−13.8	−14.7
Noise Figure NF (dB)	5	7
Interference Margin (dB)[1]	0	0
Sensitivity $P_{Rx,min}$ (dBm)	−141	−129
MCL (dB)	164	164

[1] In mobile networks, additional interference degrades the MCL especially in dense urban areas.

transmission power is a function of multiple cell parameters including coupling loss. For more than two repetitions, the cell-specific maximum transmission power on the i^{th} slot ($P_{CMAX,c(i)}$) is used. This thesis refers to the transmit power of the UE as $P_{TX,UE}$ [76]. The MCL is thus calculated as:

$$MCL = P_{TX} - P_{RX,min} \tag{24}$$

Table 2 shows a reference link budget configuration with the uplink $N_{Rep,UL}$ = 32 and the downlink $N_{Rep,DL}$ = 128 achieving the minimum requirement of MCL = 164 dB as defined by the International Mobile Telecommunications 2020 Standard (IMT-2020) [34], which specifies the requirements for upcoming 5G systems.

2.3 FUNDAMENTALS OF LORAWAN

The LoRa modulation and the LoRaWAN protocol are components of an LPWAN technology operating in the unlicensed spectrum. In Europe, LoRaWAN networks are commonly installed in the license-free EU-433 and EU-868 bands. Therefore, networks can be operated by groups or individuals without obtaining a license for the spectrum, which reduces cost and makes LoRaWAN a popular choice for WSN applications. While often used interchangeably, the terms LoRa and LoRaWAN define two different aspects of the wireless network. LoRa refers to a proprietary CSS modulation and forms the basis of most LoRaWAN networks. It is a proprietary technology owned by the Semtech corporation, which produces the modems necessary to make use of LoRa. LoRaWAN is an open standard that defines a MAC layer

protocol needed to create a network of many devices. LoRaWAN is maintained by the LoRa Alliance, which publishes new versions of the protocol in the LoRaWAN specification [1].

On the physical layer, a LoRaWAN network can either use the LoRa modulation or a Frequency Shift Keying (FSK) modulation [50]. While FSK provides higher data rates (up to 50 kbps before duty cycle regulations), many applications prefer the LoRa modulation due to its wide coverage range and robustness against interference. The measurements in this thesis are conducted using the LoRa modulation; therefore, the term *LoRaWAN network* is used to describe a network that employs both LoRa and LoRaWAN. This section covers the fundamentals of LoRa and LoRaWAN operation needed for the analysis in this work. Additional details can be found in [1, 86].

2.3.1 *Introduction*

The popularity of low-end IoT use cases such as environmental monitoring has created a demand for low-cost, wide-coverage LPWAN technologies. The LoRa modulation and the corresponding LoRaWAN protocol address such use cases; they provide an unlicensed and cost-effective LPWAN service that can be installed by anyone. These properties have attracted interest from industry and researchers alike, who continue to create novel use cases, many of which have found their way into everyday situations. Over the past years, the performance of LoRaWAN has been evaluated in a number of publications that each focus on a particular QoS aspect. There is a number of simulations that explore the scalability and delivery success rate of LoRaWAN in various configurations and environments [27, 102]. Additionally, LoRaWAN has been evaluated in physical deployments installed in different environments. Most publications do not cover a wide range of QoS parameters and there is little previous exploration on how the elevated interference levels in the ISM and SRD bands influence the LoRa QoS, which is essential to understand how and why the deployed performance differs from the specifications. Pötsch et al. [71] analyzed the elements that contributed to the end-to-end latency of LoRaWAN and compared it to the latency of a Universal Mobile Telecommunications System (UMTS) connection. They found that the end-to-end latency was composed of the Time on Air (ToA) and the LoRaWAN backhaul latency. While the ToA was increasingly dominant at higher SFs, the backhaul latency declined, which could not be explained. Petric et al. [69] conducted a drive test in a suburban environment to analyze the LoRaWAN packet loss for different SFs and signal levels. They found that LoRaWAN provided good outdoor coverage that improved with antenna height. The packet loss correlated with falling SNR, but not with falling RSSI. Under optimum conditions, packet loss ratios below 3 % were observed. Sanchez-Iborra et al. [83] evaluated the LoRaWAN QoS in urban, suburban and rural conditions using simulations and drive tests. The authors confirmed that the coverage predicted by the Okumura–Hata propagation model closely matched the real network. Furthermore, both higher SFs and higher Code Rates (CRs) improved the packet loss, especially in urban areas. Augustin et al. [10] performed a field test with different packet sizes and spreading factors. The authors found that coverage increased with higher SFs, but the RSSI was not a good indicator of MCL.

The measured throughput was significantly below the specifications and was further reduced for small packets. The authors attributed the reduction to the receive windows preventing the device from sending the next packet.

This thesis expands on the previous works by performing a systematic analysis of both physical and application layer QoS parameters using the methodology developed during the NB-IoT measurements [57]. This analysis is made possible by exploring options for a linear signal level representation and therefore circumventing the limitations of existing analyses based on the RSSI or SNR only. Special consideration is given to the influence of the CE mechanisms on the end-user QoS. The analysis is conducted both in a private LoRaWAN network and in a shielded laboratory setup to explore the influence of the interference on the performance of the system. Afterwards, the results are compared to the specifications and the applicability of LoRaWAN is discussed in four different smart metering use cases. Finally, LoRaWAN is compared to NB-IoT in terms of its physical and application layer QoS parameters.

To the best of the author's knowledge, this thesis presents one of the first systematic analyses of LoRaWAN that explore a wide range of physical and application layer QoS parameters, consider both a deployed and a shielded setup and compare LoRaWAN to other technologies. Unlike the existing analyses, all measurements are conducted using the Packet Strength as a linear signal level representation.

2.3.2 *Physical Properties of the LoRa Modulation*

The LoRa modulation is derived from existing principles of CSS modulation and encodes information in linear frequency chirps. Each chirp corresponds to one symbol, and the chirp duration equals the symbol time T_S. The chirps cover the full system bandwidth B (f_{min} to f_{max}). LoRa employs different bandwidths, usually 125 kHz, 250 kHz and 500 kHz; not all configurations can be used in every region [50]. Given the linear property of the LoRa chirps, they can be described using their steepness $\mu = B/T_S$. A chirp with $\mu > 0$ is called an upchirp, while a chirp with $\mu < 0$ is a downchirp. Figure 5 illustrates the time and frequency domain representation of an upchirp (a, b) and a downchirp (c, d).

The theory in the following sections is limited to what is necessary to analyze the QoS properties; a full analysis of the LoRa baseband processing, modulation and demodulation process can be found at [68].

2.3.3 *Receiver Sensitivity, Spreading Factor, MCL and Transmission Power Control in LoRa*

Like other LPWANs, LoRaWAN networks need to cover a wide variety of installation locations. Therefore, it is necessary to provide a CE mechanism that allows deploying devices in difficult environments such as basements. LoRa addresses this challenge by offering SFs, which double the T_S for each SF step and improve the SNR_{min} (and therefore also $P_{RX,min}$) by 2.5 dB [89]. Since LoRa is limited to a fixed bandwidth B, the steepness μ halves with each SF step. The doubling in symbol energy E_S improves the sensitivity at the cost of a decreased data rate.

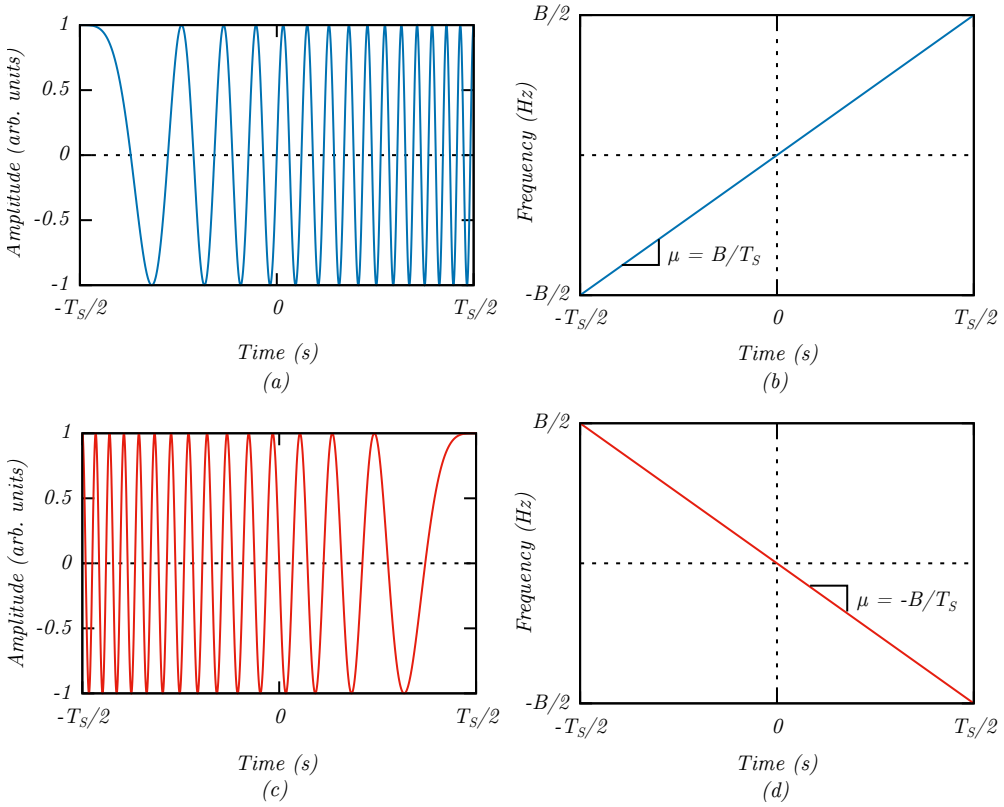


Figure 5: Chirp signals change their frequency over time: A sinusoidal linear upchirp **(a)** increases its frequency over the symbol duration T_S **(b)**. A sinusoidal linear downchirp **(c)** decreases its frequency over one symbol duration T_S **(d)**.

Overall, increasing the SF leads to:

- Improved sensitivity
- Longer transmission times for a given payload
- Lower data rates
- Higher power consumption

Similar to NB-IoT, LoRa employs a narrow system bandwidth to provide a long range and deep indoor coverage. As described in Section 2.1, the receiver sensitivity $P_{RX,min}$ and power consumption of LPWANs is improved by reducing the system bandwidth B, which also decreases thermal noise. At the same time, optimizing the modems for low cost makes improving the NF difficult, so an NF of 6 dB can be assumed [87]. In the case of LoRa, different system bandwidths are supported. In the lower ISM and SRD bands below 1GHz, the most commonly used system bandwidths are $B = 125\,kHz$, $B = 250\,kHz$ and $B = 500\,kHz$; the upper limit is defined by local regulations.

The theoretical LoRaWAN performance is illustrated in Table 3, which lists the SNR_{min}, $P_{RX,min}$ and MCL of a LoRaWAN network with a system bandwidth of $B = 125\,kHz$. The SNR_{min} is specified by the manufacturer [89] and $P_{RX,min}$ is calculated using Equation (7). The MCL definitions are identical for both LoRa and NB-IoT, so the definition in Equation (8) is reused and the MCL is calculated for the

Table 3: SNR_{min}, $P_{RX,min}$ and MCL of LoRa at different SFs [89]. The MCL is calculated for $P_{TX} = 14\,dBm$.

SF	SNR_{min} (dB)	$P_{RX,min}$ (dBm)	MCL (dB)
7	−7.5	−124.5	138.5
8	−10.0	−127.0	141.0
9	−12.5	−129.5	143.5
10	−15.0	−132.0	146.0
11	−17.5	−134.5	148.5
12	−20.0	−137.0	151.0

Table 4: An exemplary calculation of data encoded in a LoRa transmission using $SF = 7$. The chip k at which the frequency jump occurs directly correlates to the binary value of the encoded data bits. In this table, the least significant bit is on the right.

Frequency Jump Location k	Encoded Symbol
0	$(0000000)_2$
2	$(0000010)_2$
65	$(1000001)_2$
127	$(1111111)_2$

maximum permissible $P_{TX,ISM} = 14\,dBm$ according to the European SRD-band regulations [20].

2.3.3.1 *Encoding Data in LoRa Chirps*

The LoRa modulation employs a fixed scheme to synchronize the communication endpoints and encode data. Every transmission starts with the Preamble, which is a series of upchirps used for frequency and time synchronization between sender and receiver. Afterwards, downchirps form the Start of Frame Delimiter (SFD), which indicates the start of the data portion of a LoRa frame. It is the only time in a LoRa transission when downchirps are employed; all data are exclusively transmitted in upchirps, which simplifies decoding at the receiver.

Payload data are encoded by a jump in frequency at specific points in the chirp, as shown in Figure 6. For this purpose, each CSS symbol is split into 2^{SF} chips, which mark possible locations for a frequency jump to occur. The chip number $k \in [0; 2^{SF} - 1]$ at which the jump occurs directly represents the binary value of the data bits, as shown in Table 4. As a result, there are 2^{SF} possible symbols encoding SF bits each.

LoRa transceivers encode the payload for improved resiliency against interference. Gray Indexing is used to mitigate off-by-one errors during transmission, Whitening to improve clock recovery, Interleaving to improve interference resis-

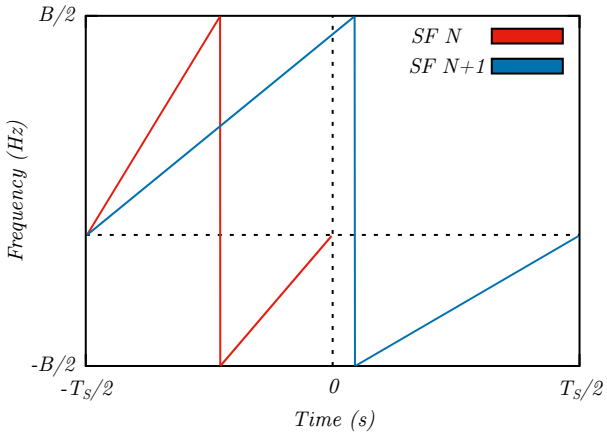


Figure 6: The LoRa modulation encodes data using instantaneous frequency jumps. The location of the jump directly represents the encoded bits. Increasing the SF by one doubles the ToA and the number of chips per symbol, but only encodes one additional bit.

tance and Forward Error Correction (FEC) to detect and/or correct bit errors [39]. In the receiver, these encodings need to be reversed to obtain the original data bits.

2.3.4 *Calculating the LoRa Data Rate*

Using the properties of the LoRa modulation discussed in Section 2.3.2, the theoretical maximum bandwidth can be calculated [87]. For base band encoding, a chip rate of $1\frac{\text{chip}}{\text{Hz}\cdot\text{s}}$ is used. The LoRa chip rate R_C can thus be calculated.

$$R_C = 1\frac{\text{chip}}{\text{Hz}\cdot\text{s}} \cdot B, \quad T_C = \frac{1}{R_C} \tag{25}$$

with:
B: Modulation Bandwidth

The symbol time T_S can be calculated by multiplication of the number of chips per symbol 2^{SF} and the chip duration T_C. This confirms that increasing the SF increases the time on air, but not the bandwidth like in other technologies such as Direct Sequence Spread Spectrum (DSSS) or Frequency Hopping Spread Spectrum (FHSS).

$$T_S = 2^{SF} \cdot T_C = \frac{2^{SF}}{R_C} \quad \rightarrow \quad T_S = \frac{2^{SF}}{B} \tag{26}$$

with:
SF: Spreading Factor

Each symbol encodes SF bits, and blocks the channel for the symbol time T_S, which allows calculating the modulation bit rate R_b. Furthermore, while each additional SF step encodes one bit more per symbol, the time on air doubles; therefore, the data rate decreases.

$$R_b = \frac{SF}{T_S} = SF \cdot \frac{1}{\frac{2^{SF}}{B}} \tag{27}$$

Table 5: Number of bit errors that can be detected and corrected by the Hamming Forward Error Correction used in LoRa. [39]

CR	R_{Code}	Error Detection (bits)	Error Correction (bits)
1	4/5	1	0
2	4/6	1	0
3	4/7	1	1
4	4/8	2	1

Since LoRa works in the unlicensed bands, interference needs to be expected and handled appropriately. LoRa implements error detection and correction using Hamming FEC with a variable code word size of 5–8 bits and a fixed data word size of 4 bits [39]. The additional parity bits reduce the effective bit rate $R_{b,eff}$. The CR is defined as the number of parity bits inserted per 4 bits of useful data [87]. The network administrator can configure a CR of 1–4 bits.

$$R_{b,eff} = SF \cdot \frac{\frac{4}{4+CR}}{\frac{2^{SF}}{B}} \tag{28}$$

with:
CR: Code Rate

In signal processing theory, the CR is more commonly expressed as the ratio of useful bits k over total bits n. In the context of LoRa, this ratio is called *Rate Code* (R_{Code}) [87]. The different levels of error detection and correction are shown in Table 5.

$$R_{Code} = \frac{4}{4+CR} \tag{29}$$

This allows simplifying Equation (28) to calculate the effective bit rate for a specific spreading factor and code rate.

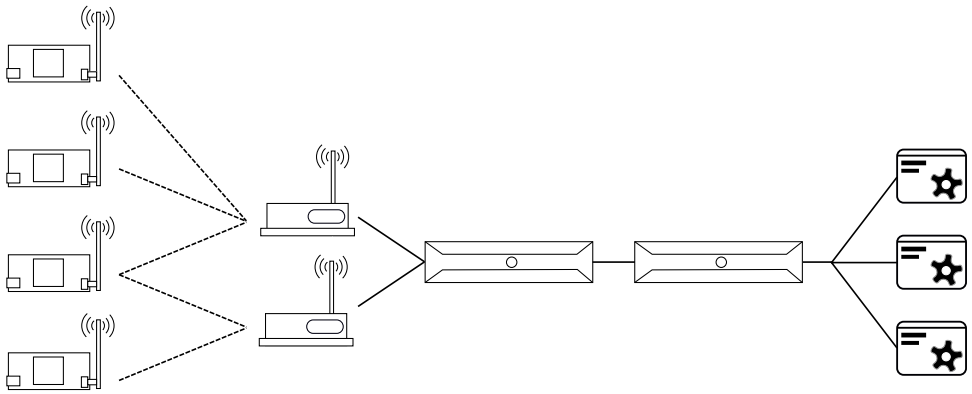
$$R_{b,eff} = SF \cdot \frac{R_{Code}}{\frac{2^{SF}}{B}} = SF \cdot \frac{B}{2^{SF}} \cdot R_{Code} \tag{30}$$

with:
R_{Code}: Substituted code rate

Using Equation (30), the theoretical LoRa data rate can be calculated. Table 6 lists the peak data rates of LoRa using different SFs at a bandwidth $B = 125\,kHz$ and a $CR = 4/5$.

Table 6: Coded bit rate of LoRa assuming B = 125 kHz and CR = 4/5

SF	Coded Bitrate (bps)
7	5469
8	3125
9	1758
10	977
11	537
12	293



(a) Nodes / Motes *(b) Gateways* *(c) Network Server* *(d) Application Server* *(e) Applications*

Figure 7: A generic LoRaWAN architecture (adapted from [47]).

2.3.5 *The LoRaWAN Architecture*

LoRaWAN networks are scalable from small local installations to a global network infrastructure. For that purpose, LoRaWAN networks are controlled by a central LoRaWAN stack that manages locally installed gateways, which provide the wireless networks that LoRaWAN end devices connect to. Figure 7 illustrates the architecture of a LoRaWAN network[3]. It is composed of the following elements [47]:

(a) LoRaWAN end devices are called *nodes* or *motes*. Usually, these devices are based on a microcontroller and include sensors or actors. Multiple nodes may form a WSN.

(b) LoRaWAN networks include one or more gateways, which provide the LoRaWAN air interface. Gateways act as packet forwarders between the nodes and the central LoRaWAN stack. For very small networks, gateways are available that provide an integrated stack.

(c) The network server manages the gateways, provides encryption and acts as a bridge between the gateways and the application server. When a node transmits an

3 The architecture described in this section refers to LoRaWAN v1.0.x, since this version was used in the LoRaWAN evaluations. The later standard LoRaWAN v1.1 introduces an additional Join Server component that verifies end devices and distributes keys.

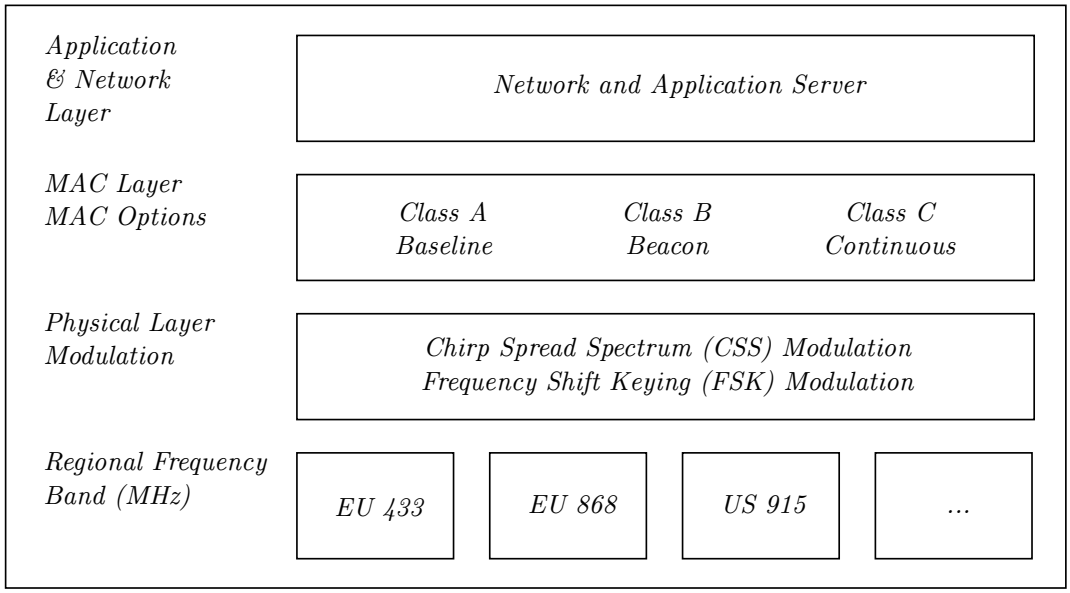


Figure 8: A LoRaWAN network is built from layers that stack on top of each other.

uplink message, it can be received by multiple gateways to improve the chance of a successful transmission. The network server will then deduplicate the message and forward only one copy to the application server. In the downlink direction, the message is transmitted only by the gateway that previously received the strongest signal from the target node, which avoids unnecessary transmissions. The network server implements the Automatic Data Rate (ADR) algorithm, which adjusts the wireless parameters of the nodes based on the channel conditions.

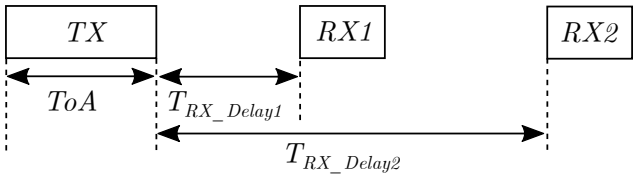
(d) The application server manages the nodes and provides an external interface to the LoRaWAN stack. From there, applications can collect the data using protocols such as Message Queuing Telemetry Transport (MQTT). Additionally, the application server handles encryption of the payload.

(e) Applications are external components that are operated by a third party. They exchange data with the nodes via the LoRaWAN application server.

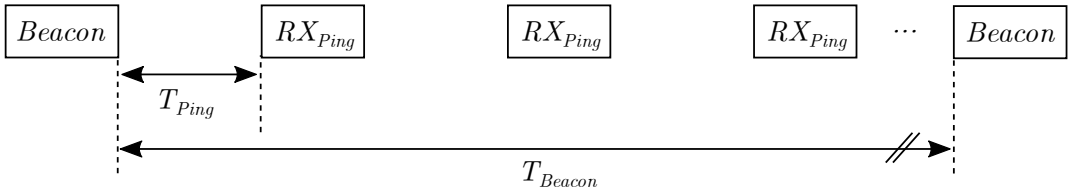
2.3.6 *The LoRaWAN Protocol Stack*

LoRaWAN networks are built from layers that stack on top of each other as shown in Figure 8. Compared to the well-known Open Systems Interconnection (OSI) model, the modulation (usually LoRa) forms the physical layer, while LoRaWAN forms the upper layers. LoRa is deployed in different region-specific frequency bands. At the PHY, a LoRaWAN network can either employ the LoRa modulation (sec. 2.3.2) or an FSK modulation. While the long range and interference resistance make LoRa an attractive choice for many use cases, there may be applications where the increased data rate of FSK is beneficial. For peer-to-peer connections, using only the LoRa PHY is sufficient to transmit information; if more devices need to be connected, the LoRaWAN protocol is required.

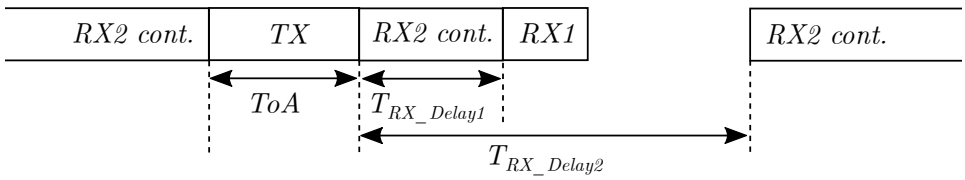
At the MAC layer, the LoRaWAN standard defines the behavior of the devices in a network. Due to the wide variety of use cases covered by LoRaWAN, the stan-



(a) LoRaWAN class A slot timing.



(b) LoRaWAN class B slot timing.
Regular uplink transmissions are possible as shown in (a).



(c) LoRaWAN class C slot timing.

Figure 9: LoRaWAN offers three device classes. **(a)** Class A is designed to conserve power for battery operated devices. **(b)** Class B is optimized for actuators, which do not have to wait for an uplink transmission to receive downlink data. The gateway sends periodic beacons, to which the node aligns its ping slots (RX_{Ping}). **(c)** Class C is meant for mains powered devices that can afford or need to be reachable at all times (adapted from [48]).

dard defines three device classes as shown in Figure 9. Not all devices support all classes, as higher classes increase the power consumption. However, higher classes incorporate the functionality of the lower classes.

The baseline *class A* is supported by all LoRaWAN devices and is primarily targeted at battery powered sensors that only send data occasionally. Devices using this class can send uplink data at any time (TX), but downlink data can only be received in two receive windows RX1, RX2 following an uplink transmission. The windows are opened after the configurable delays T_{RX_Delay1} and T_{RX_Delay2} and the gateway can use either one to transmit pending downlink data.

Some devices need to be reachable in regular intervals, such as battery powered actuators. For this use case *class B* was defined. To avoid having to wait for the next upload transmission, the device periodically opens receive windows called ping slots (RX_{Ping}). The ping slots are synchronized to beacons that are periodically sent by the LoRaWAN network and transmitted by the gateways. Uplinks are

possible at any time according to the slot timing defined in LoRaWAN class A; they are omitted from Figure 9 for simplicity .

Finally, some devices must be reachable at any time or can afford it due to being mains powered. A common example is a mains powered actuator. For these devices, the *class C* was defined. In this mode, the devices can transmit at any time; whenever a device is not transmitting, it keeps the receive window open and waits for downlink transmissions.

The LoRaWAN MAC layer design introduces a strong uplink bias to LoRaWAN networks. While nodes can send uplink data at any time, the gateway has to wait for the next receive window; depending on the selected class, this can take minutes or hours. Furthermore, the uncoordinated ALOHA medium access approach used in LoRaWAN increases the collision probability beyond a certain network load [2, 31].

On the gateway side, the duty cycle regulations hinder the scalability and throughput of LoRaWAN networks. While most LoRaWAN traffic usually is uplink oriented, the gateway still has to forward acknowledgments for confirmed uplink transmissions, the replies from the application, as well as any MAC commands from the Network Server. The network attempts to compensate for the increased load by using multiple channels at the gateway, yet the limited downlink capacity limits scalability beyond a certain point [2, 70].

2.3.7 *The LoRaWAN Packet Layout*

The LoRaWAN protocol defines the packet layout used in messages. The LoRaWAN protocol stack is composed of independent layers that use encapsulation to add the necessary control information. This generates additional overhead and must be considered when making QoS measurements. Since LPWAN applications try to minimize the payload to conserve battery power, the headers make up a significant portion of the packet. Therefore, the brutto data rate is usually considerably larger than the netto data rate. Figure 10 illustrates the layout of a LoRaWAN message.

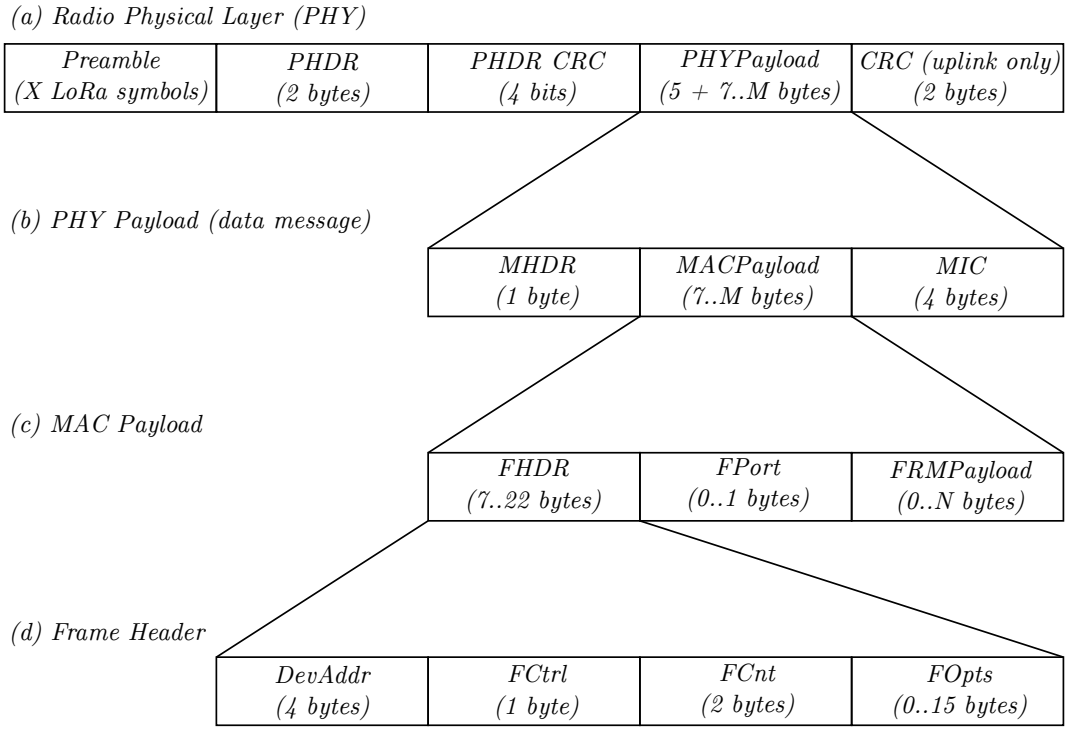
(a) On the PHY, the message consists of the following elements:

- The Preamble provides synchronization in time and frequency[4].
- The Physical Header (PHDR) contains information about the payload.
- The Physical Header Cyclic Redundancy Check (PHDR-CRC) is used to detect transmission errors.
- The Physical Layer Payload (PHYPayload) consists of a fixed 5-byte part and variable-length higher layer data.
- The Cyclic Redundancy Check (CRC) tail provides error detection.

(b) The *PHY Payload* consists of the following parts:

- The Medium Access Control Header (MHDR) defines the protocol version and message type.

4 The preamble length can be configured, ranging from 6+4 to 65535+4 Symbols. By default, LoRa uses 12 Symbols and LoRaWAN uses 8 Symbols. [89]



Legend:
X: The number of Symbols in the Preamble, configuration dependent.
M: The maximum MAC Payload length, subject to regional regulation.
N: The maximum FRM Payload length, calculated as N ≤ M - 1 - length(FHDR)

Figure 10: The LoRaWAN PHY and MAC layer message layout. In this example, the explicit header and the uplink-only CRC is included.

- The Medium Access Control Payload (MACPayload) contains either the LoRa frame, a join request, or a join accept message.
- The Message Integrity Code (MIC) enables node authentication using the Network Session Key (NwkSKey).

(c) A closer look at the *MAC Payload* reveals additional information:

- The Frame Header (FHDR) is detailed in paragraph (d).
- The Frame Port (FPort) indicates whether the frame contains MAC commands.
- The Frame Payload (FRMPayload) contains the user data.

(d) Finally, the *Frame Header* consists of the following elements:

- The Device Address (DevAddr) identifies the target device.
- The Frame Control (FCtrl) byte is used to transmit information about the network:
  - The data rate (i.e., the spreading factor) used for uplink transmissions.
  - An optional acknowledgment for previous transmissions.

  - A "last packet" identifier that indicates whether more packets will be sent from the gateway.

- The Frame Counter (FCnt) indicates whether a message has been received more than once, e.g., in case of multiple gateways receiving the same packet.
- The Frame Options (FOpts) field transports MAC commands.

Overall, the MAC overhead of a message that carries data ranges from 13–28 bytes; the exact size depends on the length of the FOpts field [49].

2.3.8 *The LoRaWAN MAC Commands*

LoRaWAN employs MAC commands to exchange configuration information between the node and the network server. These commands are invisible to the user and are sent as an encrypted message either within the FOpts field or the FRMPayload. MAC commands may reduce the Maximum Transmission Unit (MTU) of a LoRaWAN transmission.

The following is a selection of MAC commands commonly used in LoRaWAN networks. A full list of commands is available at [49].

- Acknowledgment of previous messages are piggybacked onto the next transmission
- Connectivity checks
- Automatic Data Rate (ADR) configuration of nodes
- Setting node parameters, such as data rate, transmit power, transmit frequency channel, duty cycle...
- Update the channel definitions

2.3.9 *LoRa Physical Layer Radio Measurements*

In contrast to NB-IoT, LoRa does not employ reference symbols to measure the signal level. Instead, it relies on the RSSI as defined in Section 2.2.3, which measures the total received power in the system bandwidth B including noise and interference. Therefore, the RSSI cannot fall below the effective noise floor. For a LoRa transmission with $B = 125\,\text{kHz}$ and $NF = 6\,\text{dB}$, the $P_{N,eff}$ can be calculated using Equation (4).

$$P_{N,eff,125\,\text{kHz}} = N_0 + 10 \cdot \log(125\,\text{kHz}) + 6\,\text{dB} = -117\,\text{dBm} \tag{31}$$

Furthermore, LoRaWAN modems generally report the SNR rather than the SINR, which is the ratio of the signal power P_S and the effective noise power $P_{N,eff}$, excluding the interference from other sources.

$$\text{SNR} = \frac{P_S}{P_{N,eff}} \tag{32}$$

Table 7: Duty Cycle regulations in the EU-868 band [20].

Sub-band	Frequency (MHz)	Duty Cycle (%)
g	863.00 – 868.00	1.0
g1	868.00 – 868.60	1.0
g2	868.70 – 869.20	0.1
g3	869.40 – 869.65	10.0
g4	869.70 – 870.00	1.0

Table 8: Example calculation of a repeater compatible LoRaWAN transmission at 1 % duty cycle and various spreading factors. The data rate considers the presence of MAC commands inside the FOpts field, which is a common scenario in deployed networks.

SF	Max. MAC Payload without FOpts (bytes)	Max. MAC Payload with FOpts (bytes)	Max. Goodput with FOpts (bps)
7	230	222	50.74
8	230	222	29.11
9	230	222	16.00
10	123	115	8.00
11	59	51	3.12
12	59	51	1.65

2.3.10 *Data Rate Boundaries*

While the throughput values calculated according to Equation (30) are commonly listed in literature, these are brutto data rates that can only be achieved under laboratory conditions. A deployed LoRaWAN network must adhere to the duty cycle regulations of the license-free spectrum used for LoRa transmissions. These regulations vary by region; in Europe, the EU-433 and EU-868 bands are used. Since anyone can operate a network in these bands, coexistence of devices is ensured by mandating a duty cycle, which describes the percentage of time a device is allowed to send. While a higher duty cycle improves the data rate, the collision risk increases considerably; even a low number of devices sending at 10 % duty cycle could cause enough collisions to render the channel useless. The duty cycle is defined for each individual sub-band of an ISM or SRD band. For example, the EU-868 band is divided into five sub-bands as shown in Table 7. The LoRaWAN network operator can further divide the regulated sub-bands into LoRaWAN channels, which are frequency bands that a LoRaWAN device can send in. In the LoRaWAN EU-868 regional parameters [50], the channels 868.2 MHz, 868.3 MHz and 868.5 MHz are predefined.

If a node wants to take advantage of a larger duty cycle, it must use channels that span multiple sub-bands, since all LoRaWAN channels within a sub-band share

the same duty cycle. While there are LoRaWAN devices that do enforce duty cycle regulations, it is common especially for cheaper devices not to provide such a mechanism. It is the responsibility of the operator to make sure that the duty cycle is maintained. In Europe, frequency regulations are assigned by the European Telecommunications Standards Institute (ETSI), while local governments can define additional requirements [20].

In a deployed network, the goodput is therefore calculated by applying the duty cycle regulations and subtracting the overhead caused by the LoRaWAN packet headers and FOpts field. Table 8 shows an example calculation of the maximum goodput in a deployed network in the EU-868 band. It assumes a duty cycle of 1 % and maintains compatibility with an optional repeater encapsulation layer [50]. It is evident from the data rate calculations that the data rates commonly advertised for LoRaWAN devices can only be achieved under laboratory conditions; the dominating factor is the regulation preventing a device from sending 100 % of the time. Therefore, LoRaWAN links are only suitable for transmitting individual packets rather than continuous data streams.

3 BUNDLING PROTOCOLS

Through the history of computer networking there has been a wide variety of mechanisms that combine multiple network links to improve the QoS while saving cost by leveraging existing infrastructure. These techniques are commonly referred to as link aggregation. In recent years, special consideration was given to protocols that enable applications to take advantage of multiple network links; these protocols are commonly called multipath, bonding, or bundling protocols. The great number of use cases has led to the creation of many protocols that each address a particular use case or traffic category. While most of them follow a similar architecture, there are many ways to design a bundling protocol for a specific application. In this section, the basic components of a bundling architecture and their functionality are explained. Afterwards, strategies for creating new protocols and options for backwards compatibility are discussed.

3.1 INTRODUCTION

Over the last decade the explosive growth of the IoT has generated a demand for IoT-centric network access technologies[1]. Numerous business cases have created vast WSNs, which create special traffic patterns that can be categorized as cMTC or mMTC [44]. For the latter LPWANs provide a long-range and energy efficient service for the massive number of devices common to WSNs. LPWANs can be categorized based on the spectrum they use. Licensed technologies such as NB-IoT [44] are deployed along existing mobile networks, while unlicensed technologies such as LoRaWAN [51] are deployed by third parties and work in the ISM bands. In general, licensed LPWANs provide improved QoS due to the exclusivity of the spectrum. At the same time they tend to be more costly, so companies have to balance cost and performance when designing new products.

In traditional networks such challenges are addressed by using multipath protocols that aggregate multiple network paths and select the optimum path using bandwidth, latency or cost metrics. However, existing multipath protocols do not satisfy the needs of a constrained network access. Many multipath protocols transmit payload reliably and in-order, which can lead to problems such as TCP Meltdown and Head-of-Line (HoL) blocking, ultimately degrading the goodput considerably in volatile networks such as LPWANs [105]. One of the most prominent reliable and in-order multipath protocols is MPTCP [29]. Its advanced functionality and complex header introduce significant overhead. Multipath QUIC (MP-QUIC) [24] provides low-latency connection establishment, but replicates much of MPTCP's functionality and thus inherits its challenges in constrained scenarios. Recently, a new group of multipath protocols emerged that addresses non-reliable transmissions. Stream Control Transmission Protocol (SCTP) [93] combines aspects of reli-

1 Parts of Section 3.1 have previously been published in the journal paper "The Narrowband Bundling Protocol" by A. Matz et al. [56]. © 2022 IEEE.

able and non-reliable protocols, but its complex header introduces considerable overhead and it is an uncommon protocol which may be filtered by middleboxes. Another option is MP-DCCP [7]. While it offers IP compatibility using a proxy framework [4] and avoids middlebox problems with U-DCCP [6], the necessary encapsulation and complex header introduce significant overhead. Huawei's Generic Routing Encapsulation (GRE) Tunnel Bonding Protocol [43] is limited to two paths, one of which must have a fixed bandwidth, which is not possible in volatile LPWANs. Finally, there is a number of protocols based on UDP [36] using additional infrastructure or encapsulation [22, 30, 42, 46]. Overall, the existing multipath protocols are unsuitable for LPWAN applications; reliable protocols suffer from congestion and packet loss, while others are incompatible with small MTUs that are common in LPWANs.

There is a number of works that present IoT endpoints equipped with multiple LPWAN interfaces, which is referred to as Multi-Radio Access Technology (RAT). These works focus on the design of an algorithm that assigns packets to LPWAN interfaces according to a specific QoS goal, e.g., the reduction of energy consumption. Since they do not implement a multipath protocol, they lack typical bundling features and provide only limited flexibility. There are two general categories of Multi-RAT devices. The first category consists of an application that implements a traditional scheduling algorithm, such as redundant transmission. Leenders et al. [41] perform analysis of LoRaWAN and NB-IoT using theoretical specifications and measurements of the energy consumption and latency. The authors discuss the potential benefits and drawbacks of Multi-RAT devices in terms of QoS and energy consumption. Mikhaylov et al. [61] construct a Multi-RAT device equipped with LoRaWAN and NB-IoT and analyze the energy consumption of a redundant transmission. Mozny et al. [63] present a field trial using a Multi-RAT device for tracking public transport vehicles. The authors report up to 92 % coverage using only LoRaWAN, with the rest of the locations being served by NB-IoT after switching the RAT manually. Ballal et al. [13] propose the use of a Multi-RAT device for critical tracking applications, using one main and one backup LPWAN link. Mikhaylov et al. [60] propose the concept of Multi-Radio mMTC, which allows redundant, sequential or individual transmission of a packet on LPWAN links. The second group of devices employ Machine Learning (ML) approaches to build scheduling algorithms for energy reduction. Stusek et al. [95] employ Multi-Armed Bandit (MAB) based reinforced learning algorithms to dynamically select the LPWAN interface with the lowest energy consumption. Sanchez-Iborra et al. [82] compare different ML approaches for the special case of mobile IoT devices. Sandoval et al. [84] propose a 5G mMTC focused approach and compare it to other intuitive approaches, such as random link selection and 5G first. Overall, the presented Multi-RAT approaches are optimized for specific WSN applications, while the bundling protocol and its supporting architecture presented in this work is universal to a wide range of use cases, e.g., by providing connectivity across arbitrary LPWAN technologies to any server on the Internet. Furthermore, following a traditional multipath protocol approach allows reusing existing innovations from other protocols. However, the two research directions compliment each other: thanks to the proposed modular scheduling architecture in this work, the ML based al-

gorithms developed in previous works could be integrated to optimize energy consumption.

This work expands on the existing solutions by proposing the novel Narrowband Bundling Protocol, which is tailored to the needs of a narrowband and energy-constrained network access. NBP provides an efficient multipath transport service to IoT devices that works over both IP and non-IP paths. It aims at improving the latency, reliability, and coverage of the aggregated connection. Similar to UDP it provides non-reliable and unordered transmission of packets. Along with NBP itself, a transparent proxy architecture is proposed, which enables legacy applications to continue working by intercepting UDP packets. The QoS benefits of NBP are verified in a simulation and the benefits and challenges are demonstrated in a real-world prototype aggregating multiple LPWANs. Finally, a protocol extension is proposed that allows applications to communicate their QoS requests to NBP.

To the best of the author's knowledge, this work contains the first general-purpose methodology to bundle multiple LPWANs and the first proposal of an LPWAN specific bundling protocol.

3.2 A SHORT HISTORY OF BUNDLING PROTOCOLS

Bundling protocols and their predecessors have been applied to redundant network paths for a long time. An early example is the Link Aggregation Control Protocol (LACP) [32], which is used to combine multiple Ethernet interfaces into a Link Aggregation Group (LAG). The LAG provides increased throughput and reliability; if a link fails, a failover mechanism redirects traffic to other links. LACP employs load balancing to assign complete data flows to physical interfaces based on a hash of their metadata; this avoids out-of-order packets due to latency differences, but limits the maximum per-flow data rate to that of a single interface [96].

The rise of high-resolution video streaming and live video conferences in the last decade has generated the need for ubiquitous high-speed connectivity. Many residential Internet connections, e.g., Digital Subscriber Lines (DSL) in rural areas, were no longer able to provide the required data rates, which resulted in poor QoS for video based applications. While the existing LTE infrastructure could deliver high peak data rates, the cost and limited capacity made replacing DSL impossible. Since upgrading the DSL infrastructure took time, short-term alternatives were required. From this situation, the idea was born to supplement the existing DSL connections with LTE when the data rate demand exceeded the DSL capacity. Since the two technologies are heterogeneous and LTE is volatile in terms of QoS properties (e.g., latency and data rate), traditional link aggregation mechanisms did not apply. Therefore, it was necessary to develop new protocols that could adjust to dynamic environments and perform scheduling on a per-packet basis. While this came at the cost of increased processing requirements, it was the only way to make use of the capacities of multiple volatile links in a single logical connection. The resulting network protocols were named multipath protocols or bundling protocols, and using them to aggregate Wide Area Network (WAN) connections was named access bundling. Such a device (e.g., a router) with multiple Internet access links is called multi-homed. One of the earliest examples of access bundling

was Huawei's GRE Tunnel Bonding Protocol [43], which bundles two links, one of which must provide nearly-static QoS properties such as DSL.

In the following years, a wide range of multipath protocols was developed for different applications. One of the earliest and most prominent examples is MPTCP [29], which extends the classic TCP [99] to support multi-homing via TCP options. MPTCP achieves high performance [67] and is used in smart phones [16], where it aggregates WiFi and LTE links. Other multipath protocols include MP-DCCP [7], which aggregates unreliable traffic by bonding multiple Datagram Congestion Control Protocol (DCCP) tunnel connections; since DCCP is an uncommon protocol and few applications use it natively, guest traffic is usually encapsulated before transport. A recent example of a high-end multipath protocol is the UDP-based MP-QUIC protocol [35], which is based on QUIC[2]. It provides a reliable and in-order multipath transport service and was originally designed to replace TCP for Hypertext Transfer Protocol (HTTP) applications, where it reduces the handshake latency by including Zero Round-trip Time (0-RTT) data and encryption keys in the first message.

Starting from 3GPP Release 16, multipath protocols will also be a part of state-of-the-art 5G networks. Now, the 5G Core network includes an Access Traffic Steering, Switching, and Splitting (ATSSS) service [18]. It allows combining non-3GPP access networks (e.g., WiFi) with 3GPP access networks to exchange data with a Packet Data Network (PDN). ATSSS allows sending traffic on a specific path (steer), provide session continuity between network links (switch) or aggregate multiple non-3GPP and 3GPP links (split); for this purpose, it makes use of existing multipath protocols such as MPTCP.

3.3 COMPONENTS OF A GENERIC BUNDLING ARCHITECTURE

Many bundling mechanisms follow a generic network architecture, which is shown in Figure 11. For simplicity, a traffic flow from left to right is assumed; a physical bundling architecture usually includes all components on both sides of the connection to enable bidirectional communication. The architecture consists of the following elements:

(a) A *traffic generator* (e.g., the user application) produces the Protocol Data Unit (PDU) to be transported.

(b) The PDU is forwarded to the *local aggregation point*. Generator and aggregation point can be part of the same application or separate entities, in which case one aggregation point can serve multiple applications.

(c) The *sequencing mechanism* is an optional component that assigns a sequence number to each PDU, which allows performing reordering and path estimation.

(d) The *scheduler* selects one or more paths to transmit the PDU on.

2 QUIC originally stood for "Quick UDP Internet Connections", but is no longer an acronym in the current Internet Engineering Task Force (IETF) RFC [35].

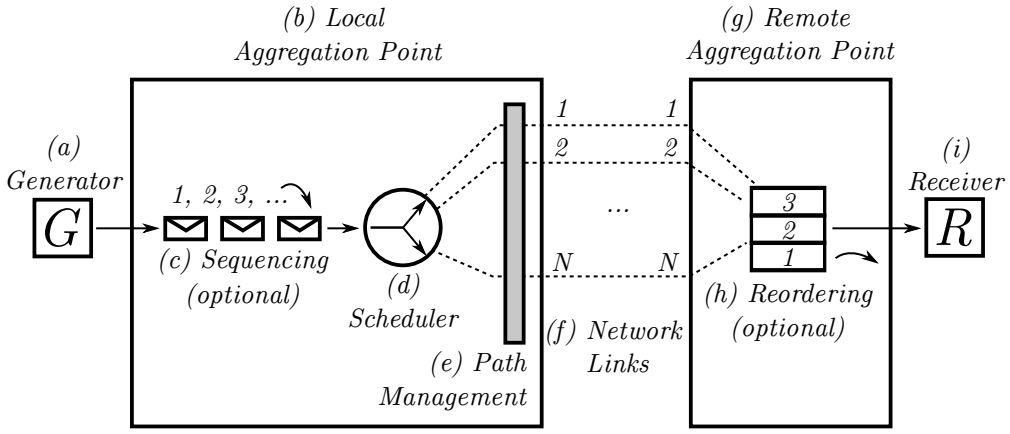


Figure 11: A generic bundling architecture. Image adapted from [55].

(e) The *path management* selects the network links that are part of the multipath connection. On each link, one or more subflows can be established. It may contain an *address management* component, which communicates changes in the available network interfaces to the remote endpoint.

(f) Multiple *network paths* interconnect the two aggregation endpoints. It is not necessary that each network path maps to a physical network link; rather, it is possible that multiple network paths terminate on the same network interface, especially if one peer has less interfaces than the other. A bundling protocol establishes one or more *subflows* (i.e., traffic flows that together form a multipath connection) over each path.

(g) The *remote aggregation point* receives the PDU and forwards it to the destination endpoint. One remote aggregation point can serve many local aggregation points.

(h) The *reordering mechanism* is another optional component that depends on sequence numbers to restore the original PDU order. This mechanism is particularly useful in the case of heterogeneous links, where it compensates for packet scrambling caused by latency differences.

(i) The packet arrives at the *receiver* and is processed by the remote application.

3.3.1 *Path Management*

The first step in creating a new multipath connection is selecting the network links that should be bundled. On each selected link, one or more subflows can be established to the remote peer. Bundling protocols use a path management mechanism that selects this mapping between local and remote interfaces. For example, MPTCP offers different algorithms to choose from, such as a full-mesh mechanism that creates subflows between all possible combinations of interfaces among other options [66]. Since each peer only knows about its own interfaces, a full-mesh can only be established if either both endpoints create subflows originating from every local interface or information about remote interfaces is exchanged, which is called address management.

3.3.2 *Scheduling Algorithms*

The subflow over which a PDU is sent is determined by the scheduling algorithm. It is the core component of a multipath architecture and is responsible for most of the QoS improvements commonly associated with bundling protocols. There is a wide range of schedulers to choose from that each optimize for a different QoS goal. In the following, some of the most common scheduler types are explained; additional algorithms and details can be found in [17].

1. A *round robin* scheduler distributes traffic equally by alternating between the available links. The round robin scheduler introduces considerable packet scrambling if the paths are heterogeneous, e.g., in terms of latency.

2. The *weighted round robin* scheduler is a special case of the round robin scheduler, which assigns PDUs based on a ratio (e.g., 10:1 for a 10 Gbps and a 1 Gbps Ethernet link).

3. A *redundant* scheduler duplicates packets on all or a subset of the available links. Sending a packet on multiple links guarantees the highest QoS level: it arrives as quickly as possible, with the highest reliability and greatest coverage area if multiple wireless links are available. Drawbacks include the high energy consumption and cost, as well as packet duplicates at the receiver.

4. A *lowest round trip time first* scheduler employs path estimation to prefer the link that currently provides the lowest latency. This algorithm works well in most networks and is the default in the reference MPTCP implementation [66], but the requirement of path estimation might be a challenge in constrained environments.

5. A *strict priority* or *cheapest pipe first* scheduler selects links according to a predefined priority value. It is a popular choice in access bundling solutions of network operators, since it allows prioritizing links with lower monetary cost.

6. Finally, there are application-specific schedulers based on reinforced learning, where a link model is derived from a large data set (see sec. 3.1). The scheduler can then optimize for a certain QoS goal such as reducing the energy consumption.

3.3.3 *Path Estimation*

The instantaneous QoS of a network path depends on many unpredictable factors, such as bottlenecks and other user activity. Therefore, many protocols employ path estimation mechanisms to derive how many packets can be sent without overwhelming the path. Over the years, a considerable number of algorithms have been developed. For example, TCP [37] and DCCP [40] reuse their built-in acknowledgments as a path estimation mechanism. The number and timing of received acknowledgments provide insight on the path capacity and allow reacting to adverse events such as packet loss. Many algorithms such as the lowest round trip

time first scheduler depend on an up-to-date path estimation. The constraints in terms of energy, data rate and MTU commonly found in LPWANs complicates finding an appropriate algorithm.

3.3.4 *Reordering*

Finally, many high-end multipath protocols such as MPTCP and MP-DCCP include sequencing and reordering mechanisms that detect packet loss and decrease packet scrambling at the receiver. For reliable and in-order protocols such as MPTCP, implementing a reordering mechanism is easy; the aggregation point simply waits until the PDU with the expected sequence number arrives. For unreliable protocols such as MP-DCCP, this task is more challenging; after all, there is no guarantee that a particular PDU arrives at all. Therefore, reordering mechanisms in unreliable protocols need to balance between waiting long enough for missing packets to arrive or moving on with the transmission and omitting the packets [5].

3.4 STRATEGIES FOR DESIGNING A NEW BUNDLING PROTOCOL

When creating a new bundling protocol, there are three basic approaches to choose from. The first approach is extending an existing protocol such as TCP [99] or UDP [36], e.g., by adding options or new header fields. This strategy is employed by MPTCP [29], which extends TCP using TCP Options. Extending an existing protocol allows reusing functionality such as congestion control and path estimation; furthermore, if designed appropriately all applications that support the base protocol can enjoy a seamless and low overhead multipath service. However, the new bundling protocol inherits both wanted and unwanted attributes from the base protocol, which may require workarounds to maintain compatibility with the protocol itself and existing applications.

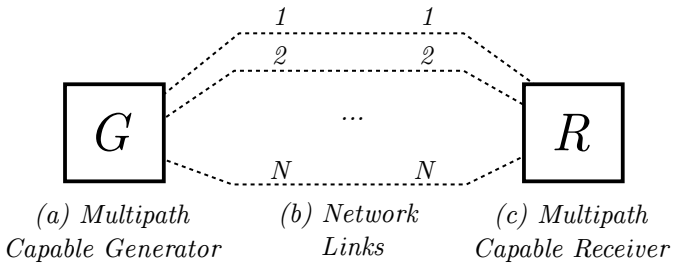
The second option consists of an application that employs tunneling, Software Defined Networking (SDN) or routing mechanisms to aggregate multiple independent connections of an existing network protocol. Similar to a Virtual Private Network (VPN), applications can then route their traffic to a virtual interface, where it is encapsulated in the host protocol and transported to the remote endpoint. Due to the encapsulation process, applications can choose whichever protocol fits their needs best. The main disadvantage is the overhead introduced by the necessary encapsulation of user traffic and the need for supporting network infrastructure.

The third option is creating a completely new protocol that is tailored to the individual needs of a particular use case. This approach provides the greatest degree of freedom to adjust the protocol to the QoS needs of the application and the characteristics of the expected network links, but initially there are no devices that support the protocol. Therefore, it is necessary to consider mechanisms for backwards compatibility. Furthermore, uncommon protocols can be subject to middlebox filtering, which may result in packets being dropped or header fields manipulated.

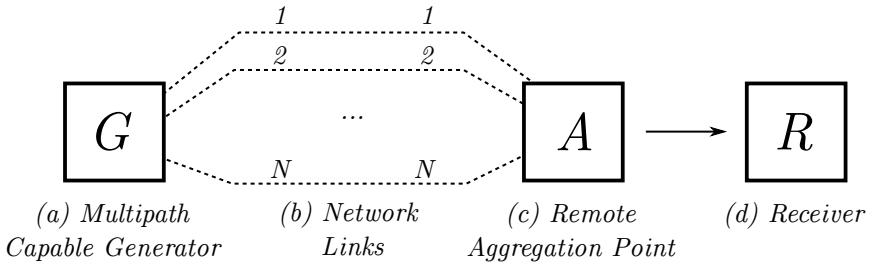
3.5 OPTIONS FOR BACKWARDS COMPATIBILITY

When new protocols are introduced to existing networks, there is usually a long period where most endpoints do not support the protocol. Therefore, it makes sense to design mechanisms that provide backwards compatibility with legacy endpoints. For multipath protocols, one solution to this problem is including a fallback mechanism that reverts to a well-known protocol if no support for the multipath protocol is detected. This strategy is implemented by MPTCP, which silently reverts to TCP if the remote endpoint does not indicate MPTCP support during the handshake [29]. The drawback of this approach is that it limits multipath operation to connections where both endpoints support the protocol. If multipath connectivity to all hosts is desired regardless of their protocol support, compatibility mechanisms such as proxy architectures are required [4]. Since the proxy imprints its QoS characteristics onto the guest traffic (e.g., in terms of reliability and automatic retransmissions), the multipath protocol must be chosen carefully to match the application.

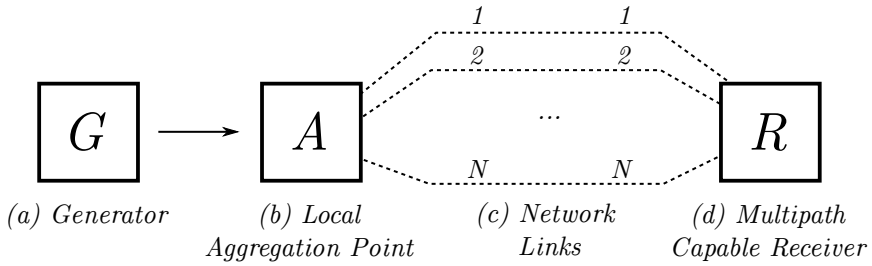
Figure 12 illustrates different compatibility options ranging from end-to-end multipath connectivity to a transparent proxy solution. In scenario 1, both endpoints support a multipath protocol and at least one of them is multi-homed. In this ideal case a connection can be directly established across multiple network links and no further supporting architecture is required. Especially for new protocols this is a rare scenario, so backwards compatibility options should be explored. Scenarios 2 and 3 consist of one endpoint that does support a multipath protocol (e.g., a smart phone running MPTCP) that communicates to another endpoint that does not support the protocol (e.g., a web server). In this case, an aggregation point can terminate the multipath connection and forward the traffic using a well-known protocol, e.g., TCP. This architecture is called one-end transparent, since the multipath connection is hidden from one of the endpoints. Scenario 4 is based on two endpoints that both do not support a multipath protocol, but should profit from bundling functionality. In this case, both a local and a remote aggregation point must be installed. This architecture is called a transparent proxy, since the endpoints are both unaware that a multipath connection exists between them.



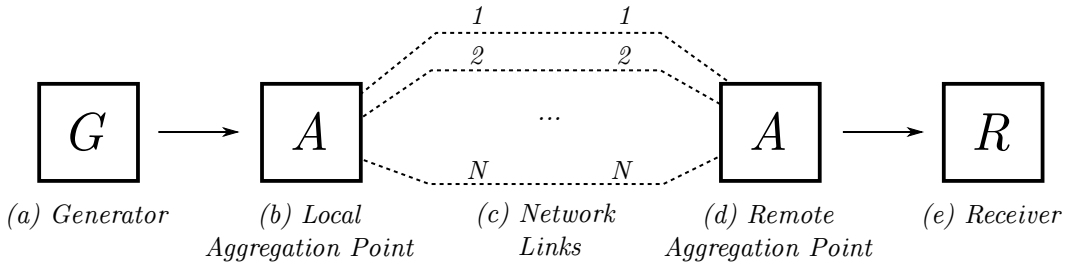
Scenario 1: End-to-end multipath connectivity



Scenario 2: A multipath capable generator connects to a single path receiver



Scenario 3: A single path generator connects to a multipath capable receiver



Scenario 4: A transparent proxy architecture hides the multipath operation from both endpoints.

Figure 12: Different options for backwards compatibility using proxy architectures.

Part II

RESEARCH DEVELOPMENT

4 A SYSTEMATIC QUALITY OF SERVICE ANALYSIS OF NB-IOT

Narrowband Internet of Things is part of a novel group of access technologies referred to as LPWANs, which provide energy-efficient and long-range network access to IoT devices.[1] Although NB-IoT Release 13 has been deployed by MNOs, detailed Quality of Service evaluations in public networks are still rare. In this chapter, systematic physical layer measurements are conducted, and the application layer performance is verified. Special consideration is given to the influence of the radio parameters on the application layer QoS. Additionally, NB-IoT is discussed in the context of typical smart metering use cases. The results indicate that NB-IoT meets most theoretical 3GPP design goals in a commercial deployment. NB-IoT provides a wide coverage by using signal repetitions, which improve the receiver sensitivity, but simultaneously increase the system latency. The maximum data rates are consistent over a wide range of coverage situations. Overall, NB-IoT is a reliable and flexible LPWAN technology for sensor applications even under challenging radio conditions. Four smart metering transmission categories are analyzed, and NB-IoT is verified to be appropriate for applications that are not latency sensitive.

4.1 EXPERIMENTAL SETUP AND METHODS

The complexity of mobile networks makes a systematic QoS evaluation challenging for end-users. In most cases, the measurements performed by the base station cannot be obtained from the network. As a result, QoS parameters such as the received signal level and quality, SNR, eNodeB transmission power, as well as various radio configuration parameters like the number of ECL repetitions are not available to estimate the network performance. However, modems perform extensive lower layer measurements for channel estimation, which can be employed to understand the system behavior under different coverage situations. Furthermore, the application layer measurements can be used to analyze the data rates and latency. In this section, the setup used for the experiments is presented.

All measurements in this research were performed from an end-user perspective, without access to the base station or control of the serving network. As such, the experiments could be reproduced by application developers, and could be employed to estimate the QoS that could be expected for a given use case in different coverage scenarios. Most physical radio parameters, such as ECL and the number of repetitions, were controlled by the serving network or the eNodeB. One notable exception was the UE transmission power ($P_{TX,UE}$), which was set by the UE algorithm, but was subject to cell-specific and user limits. In this evaluation, the maximum user limit of 23 dBm was used. Figure 13 shows the architecture used for measuring the physical and application layer NB-IoT QoS parameters.

1 Chapter 4 is an excerpt from the previously published journal paper "A Systematic Analysis of NB-IoT Quality of Service" by A. Matz et al. [57].

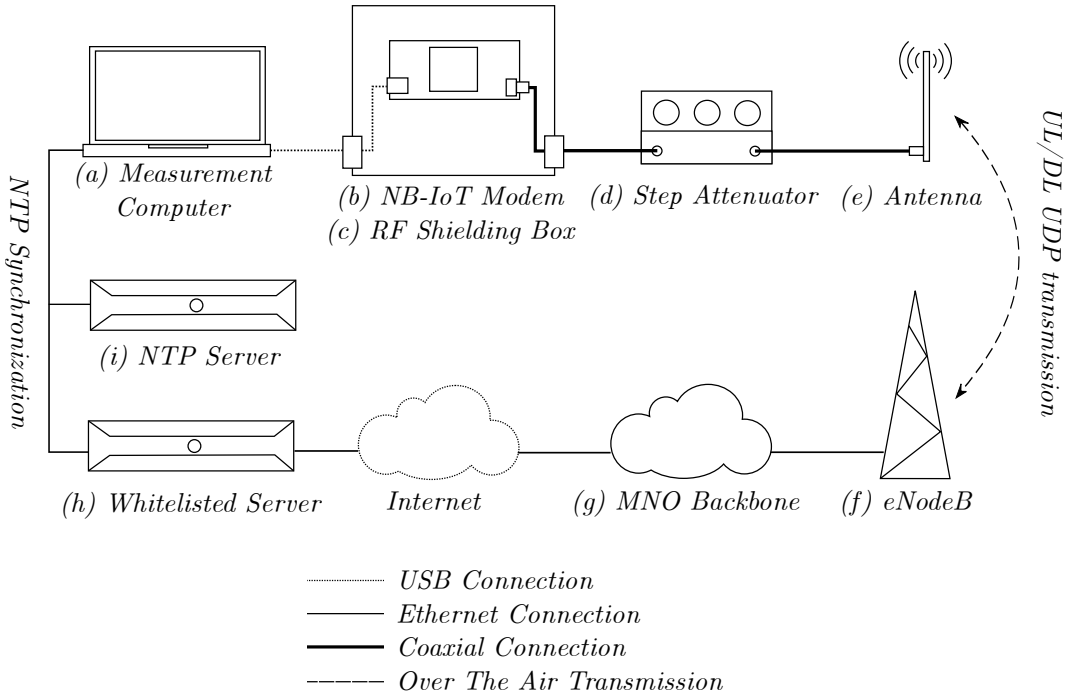


Figure 13: The measurement architecture used for the physical and application layer evaluations.

The architecture consisted of the following elements:

(a) A *measurement computer* controlled the NB-IoT modem via a Universal Serial Bus (USB) connection. A UDP client was running on the computer to send and receive test packets.

(b) The *NB-IoT modem* connected to the public NB-IoT network and performed lower layer measurements. The Exelonix NB|DEV kit [25], which is based on the uBlox SARA-N211 modem [100], was selected for the physical layer measurements due to its wide range of reported QoS values.

(c) A *Radio Frequency (RF) shielding box* prevented the mobile carrier signal from coupling into the modem Printed Circuit Board (PCB) when high attenuation was applied.

(d) A *Wavetek 5080.1 step attenuator* (0–81 dB attenuation, $f_{max} = 1000$ MHz) allowed simulating different signal levels.

(e) A *Delock 88571 omnidirectional antenna* (A = -0.8 – 0.2 dBi) was connected via a coaxial cable to the step attenuator.

(f) The *eNodeB* was the MNO base station providing the NB-IoT carrier signal.

(g) The *MNO backbone* contained provider internal functionality, such as subscriber management and packet routing.

(h) A *whitelisted server* was required by the MNO as a packet destination for security reasons. The server ran a UDP server to send and receive test traffic.

(i) A *Network Time Protocol (NTP) server* was part of the same local network as the measurement computer and the whitelisted server. It was used to synchronize the real-time clocks and enabled high-precision latency measurements.

The measurements were performed in a public NB-IoT network in Germany, which was installed as a guard band deployment in Band 20 (800 MHz). All measurements were conducted indoors with a Non Line of Sight (NLOS) connection to the eNodeB. The modem was kept stationary during the measurements. Since the modem implemented a full network stack, all IP connections were created between the whitelisted server and the modem itself. As such, the modem served as the source for all uplink packets, while the whitelisted server was the source for all downlink packets. Two groups of experiments were performed: physical and application layer QoS measurements.

(a) In the first group of measurements, the physical layer QoS was analyzed using the downlink parameters provided by the NB-IoT modem. For each 5 dB attenuation step, 30 UDP packets were transmitted via the user plane both in the uplink and downlink direction. The packets were identical in the uplink and downlink and included UDP, IP, Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), and MAC headers. Each packet carried a 27-byte payload that was comprised of the following elements:

- An indicator for uplink or downlink transmissions
- A sequence number, which was incremented by one for each packet
- The current attenuation step in dB
- A timestamp in milliseconds

For each transmission, all available QoS parameters were collected along with the timestamps of both endpoints. Between each measurement point, a delay was introduced to prevent two measurements from interfering with each other.

(b) The second group of measurements was conducted to examine the application layer QoS. There were two transmission categories that covered most IoT use cases:

- Uplink and downlink transfer of small, individual packets (e.g., sensor data) with a focus on latency
- Uplink and downlink transfer of a continuous data stream (e.g., a software update) with a focus on throughput

The first measurement employed the previously acquired packet timestamps to analyze the total system latency for individual packets at varying signal levels, which were simulated using the step attenuator. This allowed identifying conditions that significantly affected the QoS.

In the second measurement, the maximum data rate of NB-IoT was evaluated in different coverage situations, which were also enforced using the attenuator. The measurement employed UDP, which is a common protocol in IoT applications and lacks mechanisms that could have influenced the

measurement like congestion and flow control. The following parameters were adjusted to evaluate a wide range of potential use cases:

- Signal level (10 dB steps)
- Packet size (8 bytes, 64 bytes, 512 bytes, 1024 bytes)
- Uplink and downlink direction

For each combination of attenuation step, packet size, and direction, 100 UDP packets were sent to simulate a continuous data transmission. Since the Exelonix NB|DEV module imposed a packet size limit of 237 bytes and restricted the packet rate, the measurement was conducted using the PyCom GPy 4 NB-IoT modem [79]. The NB-IoT standard offers optimizations for non-IP communication, which reduce the protocol overhead by omitting IP and UDP headers. At the time of writing, these optimizations were not publicly available, so only IP-based transmissions were evaluated.

4.2 RESULTS

In this section, the measurement results are presented and analyzed in terms of physical and application layer QoS. For each evaluation, the measurement procedure is described. Afterwards, the observations are discussed, and an initial conclusion is drawn.

4.2.1 *PHY Measurements: Coupling Loss*

The first measurement explored the RSRP and RSSI signal level parameters. This evaluation was essential to find a suitable representation of path loss for later measurements. A wide range of signal levels was simulated by manually introducing attenuation in 5 dB steps until the signal was lost. For each attenuation level, 30 packets were sent, and the downlink RSRP and RSSI (DL-RSRP and DL-RSSI) were recorded for each packet.

Figure 14a shows that the RSRP was proportional to the coupling loss under all conditions, including signal levels below the effective noise floor (see Section 2.2.5). The ability to measure signal levels below the noise floor was enabled by coherently adding signal repetitions, which increased the receiver sensitivity. This measurement confirmed the sensitivity of the SARA-N211 modem of -135 dBm [101].

Figure 14b compares the RSRP and RSSI for different signal levels. While the RSRP was proportional to coupling loss, the RSSI converged to the 180 kHz effective noise floor $P_{N,eff,180\,kHz} = -114.4$ dBm. Since the RSSI included signal, noise, and interference power, it could not fall below the effective noise floor. As such, it was unsuitable for coupling loss evaluations.

Table 9 provides additional statistical detail on the RSRP and RSSI measurements. The RSRP mean values (μ_{RSRP}) followed the manual attenuation closely for all signal levels, with a maximum error of 3 % at 50 dB attenuation. On the other hand, the RSSI mean values (μ_{RSSI}) converged to the effective noise floor, with a minimum average value of -112.2 dBm at 50 dB attenuation. As expected, the

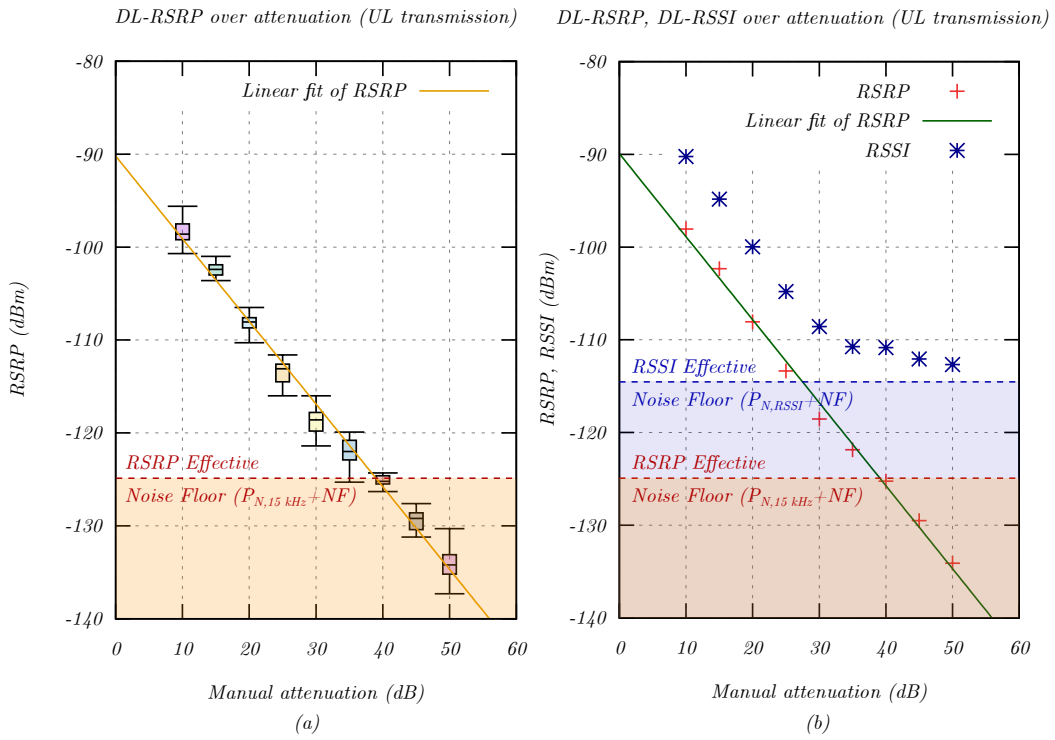


Figure 14: Analysis of different signal level measurements as an expression for coupling loss. **(a)** Distribution of RSRP signal levels for artificial attenuation in 5 dB steps. **(b)** Comparison between RSRP and RSSI for various levels of artificial attenuation.

standard deviation of the 15 kHz RSRP signal level (σ_{RSRP}) was higher than for the RSSI (σ_{RSSI}), which had a larger bandwidth of 180 kHz.

In the 3GPP specification, the Coupling Loss (CL) is estimated using the RSRP measurement [76]:

$$CL = P_{TX,NRS} - RSRP_{filtered} \tag{33}$$

where $RSRP_{filtered}$ is the higher layer filtered RSRP.

The power $P_{TX,NRS}$ allocated to an NRS depends on the system configuration [76] and was approximated hereafter proportional to the bandwidth:

$$P_{TX,NRS} \approx P_{TX,NRS} + 10 \cdot \log(1/12) = 35\,\text{dBm} - 10.8\,\text{dB} = 24.2\,\text{dBm} \tag{34}$$

The minimum observed DL-RSRP value during the measurement was -137.3 dBm at 50 dB attenuation. The maximum observed MCL could therefore be approximated:

$$MCL_{max,observed} \approx 24.2\,\text{dBm} - (-137.3\,\text{dBm}) = 161.5\,\text{dB} \tag{35}$$

In a following evaluation related to the exception report latency, a minimum DL-RSRP level of -139.3 dBm was measured, which resulted in a $MCL_{max,observed}$ of 163.5 dBm. As such, the MCL closely matched the 3GPP specification [72].

Table 9: Mean values (μ_{RSRP}, μ_{RSSI}) and standard deviation (σ_{RSRP}, σ_{RSSI}) of the RSRP and RSSI measurements. For each attenuation step, 30 measurements were performed.

Attenuation (dB)	μ_{RSRP} (dBm)	σ_{RSRP} (dB)	μ_{RSSI} (dBm)	σ_{RSSI} (dB)
10	−98.052	1.406	−90.237	0.909
15	−102.325	0.720	−94.827	0.592
20	−108.063	1.183	−99.966	0.631
25	−113.352	1.163	−104.787	0.680
30	−118.548	1.647	−108.580	0.647
35	−121.862	1.513	−110.743	0.429
40	−125.244	1.513	−110.834	0.293
45	−129.202	1.009	−111.768	0.613
50	−133.635	1.716	−112.210	0.932

Overall, the RSRP and RSSI measurements behaved as expected. There was a slight reduction from the 3GPP specified MCL of 164 dB, which could be attributed to external interference. The MNO could compensate for this by configuring more ECL repetitions, which in turn would increase the latency.

4.2.2 *PHY Measurements: Coverage Extension*

In this evaluation, the selection of different ECLs by the NB-IoT network was examined. This allowed understanding how NB-IoT used signal repetitions to adjust to changing signal levels. Various coverage situations were simulated using the attenuator. For each 5 dB attenuation step, 30 packets were transmitted in the uplink and downlink direction, and the selected ECL was recorded for each packet.

Figure 15 illustrates the ECL assigned to the UE by the eNodeB depending on the DL-RSRP signal level for (a) uplink and (b) downlink transmissions. In general, higher ECLs were selected for lower RSRP levels. There was an overlap between the individual ECL regions, which indicated a hysteresis in the ECL selection algorithm of the MNO. This technique avoided constant switching between different ECL levels for minor variations in signal level. The network used ECL 0 for RSRP signal levels above approximately -115 dBm, ECL 1 between -126 dBm and -97 dBm, and ECL 2 below -125 dBm RSRP. This allocation indicated that ECL 2 was selected for detecting signals at, or below the effective noise floor $P_{N,eff} = P_N + NF$.

4.2.3 *PHY Measurements: Transmission Power*

This measurement analyzed the transmission power $P_{TX,UE}$ selected by the modem in different coverage situations. Accordingly, conclusions could be drawn on the signal level needed for output power reduction. Various signal levels were simulated in 5 dB steps using the attenuator, and 30 packets were sent for each step

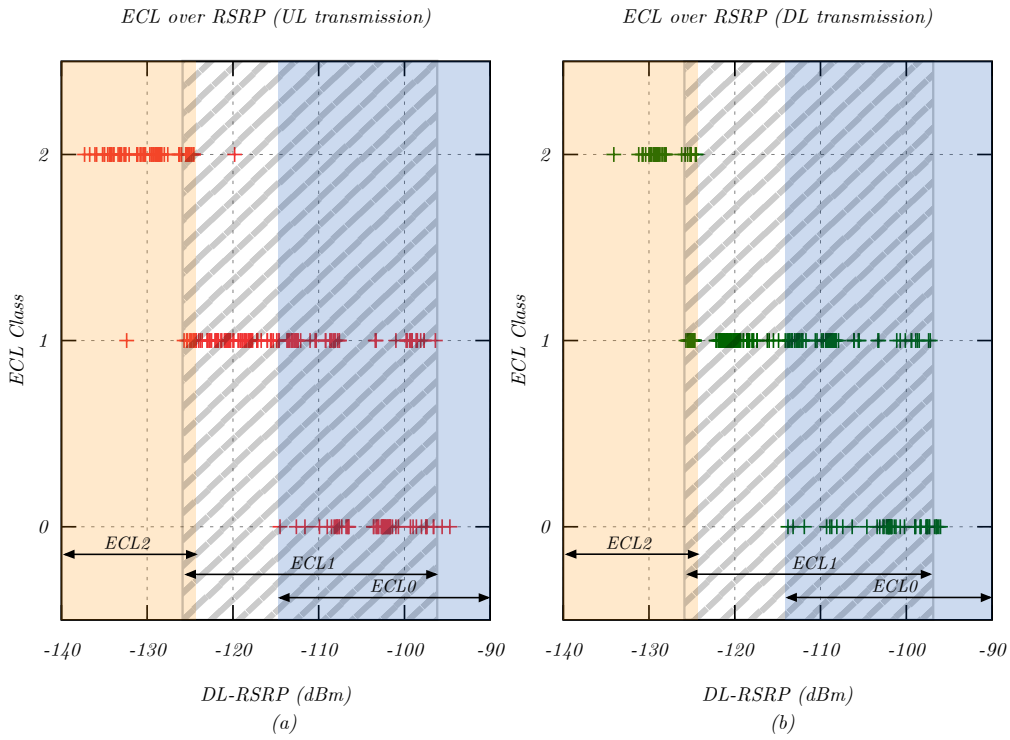


Figure 15: Relationship between ECL selected by the modem and downlink RSRP signal level during **(a)** uplink and **(b)** downlink transmissions.

in the uplink and downlink direction. For each packet, $P_{TX,UE}$ and the RSRP signal level were recorded.

Figure 16 shows the transmission power $P_{TX,UE}$ of the NB-IoT UE at different signal levels for (a) uplink and (b) downlink transmissions. For RSRP levels below -102 dBm, the maximum cell-specific output power of $P_{CMAX,c}$ = 23 dBm was used. Above this level, the UE reduced its output power in 1 dB steps. This technique conserved power for battery operated devices, while maintaining a reasonable performance at high received signal levels. The power reduction depended on the estimated coupling loss, which was calculated using the current RSRP value [76]. Overall, this measurement indicated that an improvement in received signal level (e.g., by installing an external antenna) could significantly improve battery life.

4.2.4 *PHY Measurements: Signal Quality Analysis*

In this section, the signal quality parameters SNR, SINR, and RSRQ were used to analyze the Interference Margin (IM), the noise figure of the modem, as well as the cell load during the measurements. Various signal levels between excellent coverage and total signal loss were simulated in 5 dB steps using the attenuator. For each attenuation step, 30 packets were sent in the uplink and downlink, and the RSRP, SNR, and RSRQ were recorded. Additionally, a long-term measurement was conducted over 24 hours. Every minute, one uplink and one downlink message

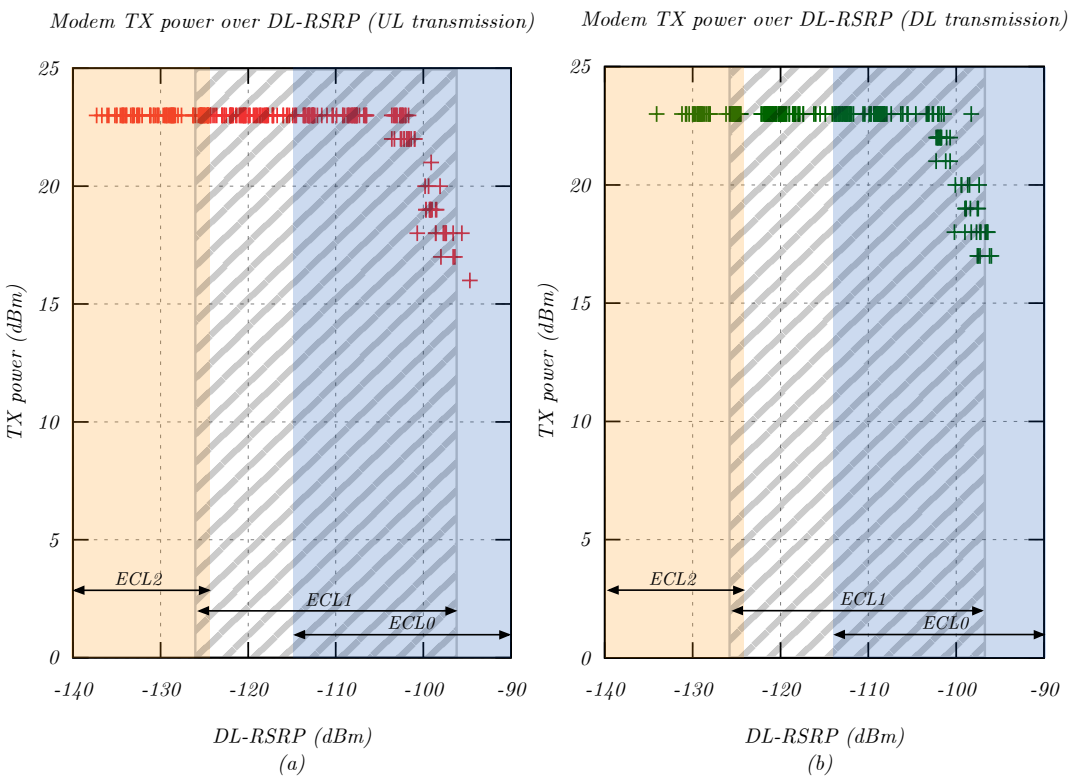


Figure 16: UE transmit power $P_{TX,UE}$ as a function of downlink RSRP signal level for **(a)** uplink and **(b)** downlink transmissions.

were sent for a total of 1440 measurements per direction. For each packet, the SNR, RSRP, RSSI, and latency measurements were recorded.

Figure 17a compares the theoretical SNR to the measured SINR as a function of RSRP. The SNR could be derived for an effective noise floor $P_{N,eff,15\,kHz} = -125.2\,dBm$ and a noise figure $NF = 7\,dB$:

$$SNR = RSRP - P_{N,eff,15\,kHz} = RSRP - P_{N,15\,kHz} - NF = RSRP - 125.2\,dBm \quad (36)$$

The SINR additionally included interference, so the SNR was an upper bound for the measured SINR:

$$SINR\,(dB) = SNR\,(dB) - 10 \cdot \log\left(1 + \frac{P_{I,15\,kHz}\,(W)}{P_{N,eff,15\,kHz}\,(W)}\right) \quad (37)$$

The IM, which was considered in the link budget in Section 2.2.5, could then be derived as:

$$IM = 10 \cdot \log\left(1 + \frac{P_{I,15\,kHz}\,(W)}{P_{N,eff,15\,kHz}\,(W)}\right) \quad (38)$$

Figure 17a shows an offset between the theoretical SNR and the measured SINR samples, which was an indication of the presence of network interference during

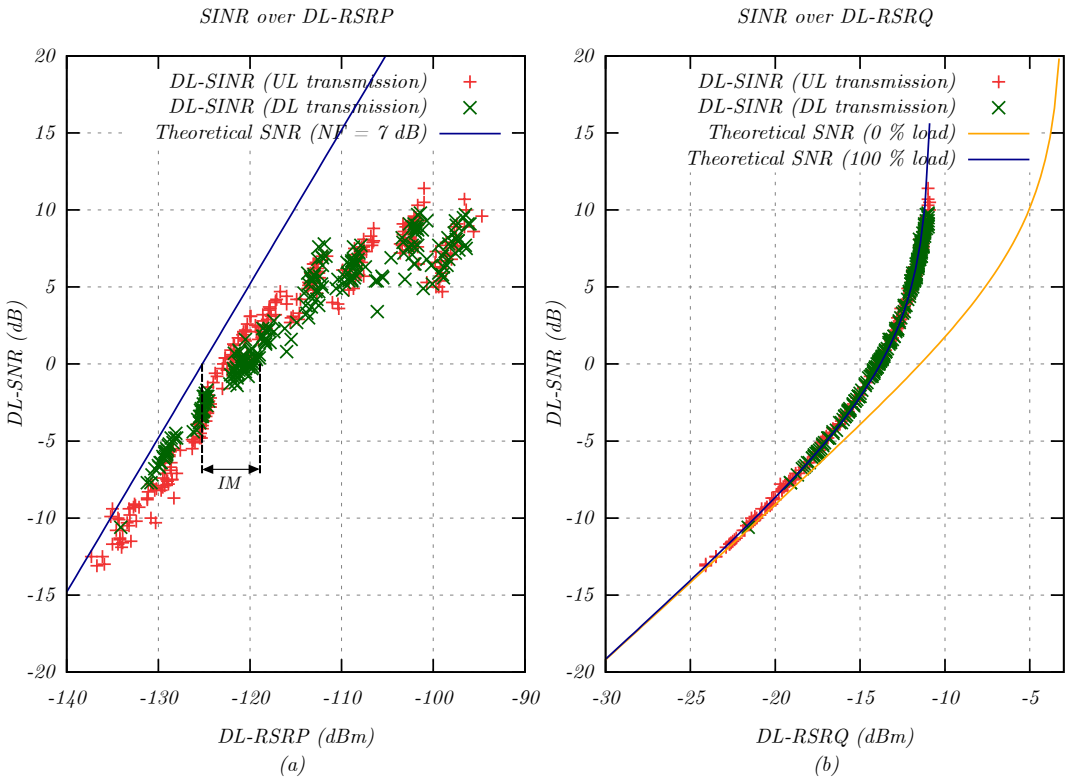


Figure 17: Analysis of the DL-SINR due to the presence of network interference for varying attenuation. **(a)** Measured DL-SINR over reported RSRP. **(b)** Measured DL-SINR over RSRQ. The solid lines show the theoretical curves for $x = 2/12$ (0 % load) and $x = 1$ (100 % load) according to Equation (11).

the measurement. The resulting IM of 2–6 dB as defined in Equation (38) increased the effective noise floor and reduced the MCL in a real NB-IoT network. Furthermore, Figure 17a provides an upper bound of 7 dB for the NF, since a higher NF would shift the slope of the theoretical SNR to the right and violate the condition that SNR was an upper bound for SINR.

In Figure 17b, the theoretical SNR for 0 % cell load ($x = 2/12$) and 100 % cell load ($x = 1$) is compared to the measured SINR values for varying RSRP signal levels. The measurement matched the theoretical curve for 100 % load as defined in Equation (11), which confirmed that all 12 subcarriers were allocated in every modem QoS measurement. This was expected, since all modem measurements were performed in the downlink, and the 3GPP NB-IoT specification [75] requires that downlink resources are allocated in subframes over the full system bandwidth.

Figure 18a illustrates the signal conditions over a period of 24 hours. The measurement points were filtered to only include one cell tower. There were significant fluctuations in the RSRP level, which reflected changes in the signal power received by the modem. These changes could be the result of human activity and created an effect similar to small-scale fading. At night, the signal conditions stabilized dramatically. As such, UEs in extreme coverage might benefit from collecting the sensor values during the day and transmitting at night.

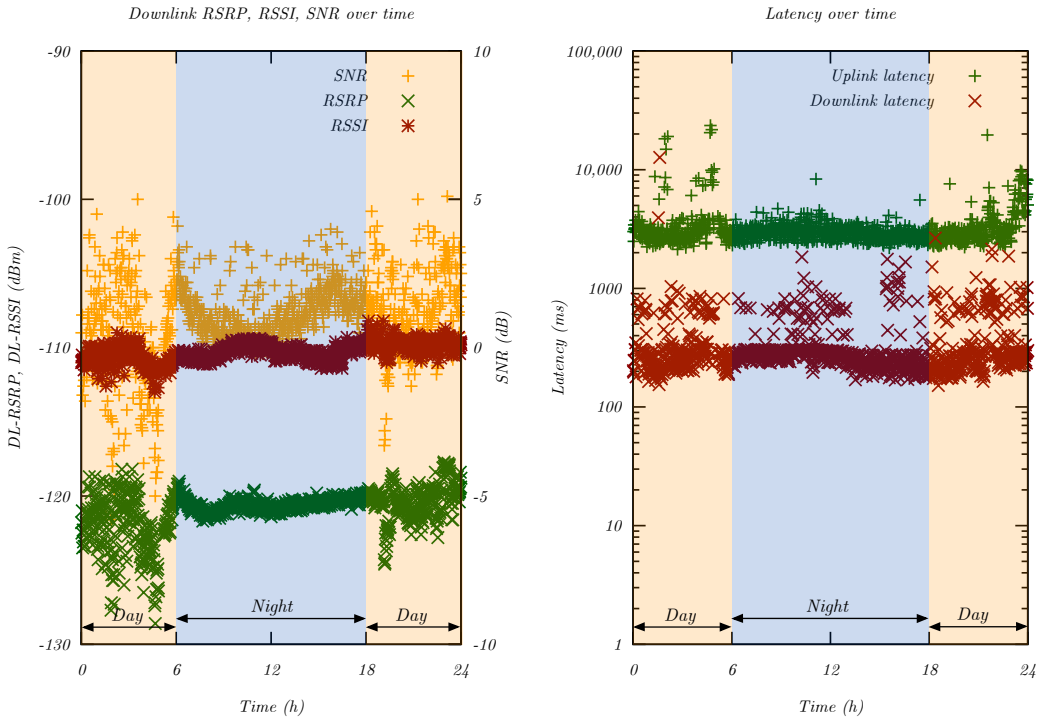


Figure 18: Long-term measurement of NB-IoT QoS parameters over 24 hours. **(a)** Signal quality parameters. **(b)** Uplink and downlink latency.

Figure 18b compares uplink and downlink latencies over time. Once again, the measurement points were filtered to only include one cell tower. Despite the significant changes in the RSRP and SNR levels shown in Figure 18a, the latency remained relatively constant. While there were latency spikes in the downlink, these values were well within the 10 s latency requirement of 3GPP [72].

4.2.5 *Application Layer Performance: Latency*

In the first application layer measurement, the factors impacting system latency were analyzed under varying coverage situations. The total system latency is the most important QoS criterion for small and infrequently transmitted packets, such as sensor data. Similar to previous evaluations, the signal level was varied in 5 dB steps using the attenuator, and 30 packets were sent per attenuation step and per direction. In a first analysis, the effect of the ECL mechanism on the uplink and downlink latency was considered. For this purpose, packets were grouped by their ECL class and analyzed in terms of latency.

Figure 19 illustrates the effect of different ECL levels on the total system latency. The measurement confirmed that higher ECLs, and thus more signal repetitions, had an impact on the total system latency. The latency rise between ECL 0 and ECL 1 was less significant than between ECL 1 and ECL 2, which indicated a more drastic increase in the repetition number in the latter case. Furthermore, there was a significant difference between the latency in the uplink and downlink direction.

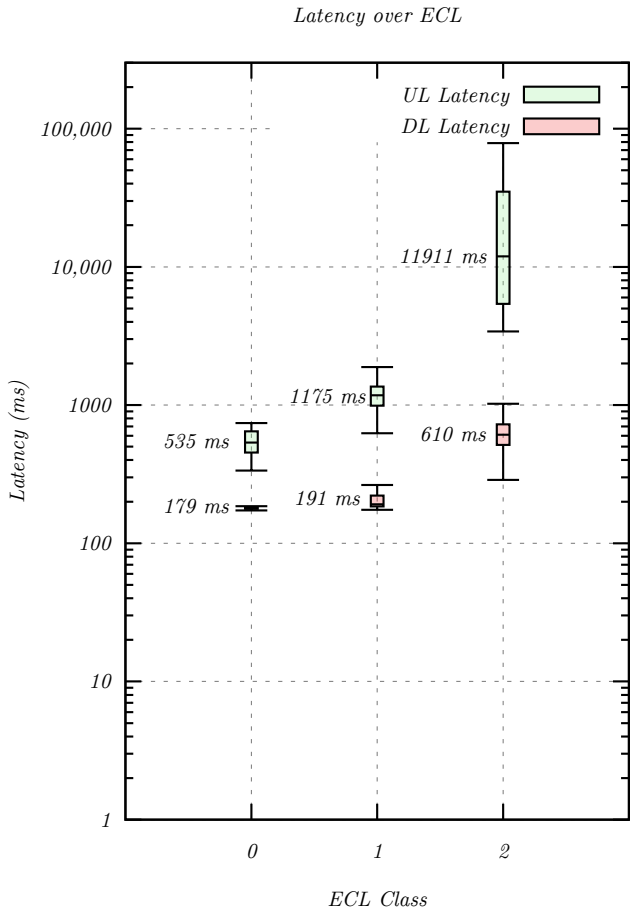


Figure 19: The influence of ECL classes on the total system latency. The median latency value is given for each ECL and transmission direction.

This could be explained by the allocation of 1 ms subframes over 12 subcarriers in the downlink, and 8 ms RUs on a single tone in the uplink, which overcompensated the higher number of repetitions in the downlink direction. In order to obtain a better understanding of the latency contribution of the individual systems, the round-trip-time between the whitelisted server and the MNO backbone (see Figure 13h and Figure 13g) was measured to be 17.776 ms (N = 30, σ = 0.163 ms), which indicated a latency of about 9 ms. As such, the latency contribution of the Internet link remained below 5 % of the total system latency for all use cases.

The second analysis was concerned with the influence of the RSRP signal level on the total system latency. In addition to transferring regular user data, NB-IoT implements the ability to send exception reports, which are high-priority messages that can be used for reporting an alert condition. The 3GPP has specified a maximum latency of 10 s for waking up the modem and delivering an exception report [44, 72]. In order to evaluate if exception reports could reduce the system latency, the modem was configured to send 30 exception reports per attenuation step in the uplink direction. For each step, packets were grouped together and assigned to their average signal level, and the latency was analyzed.

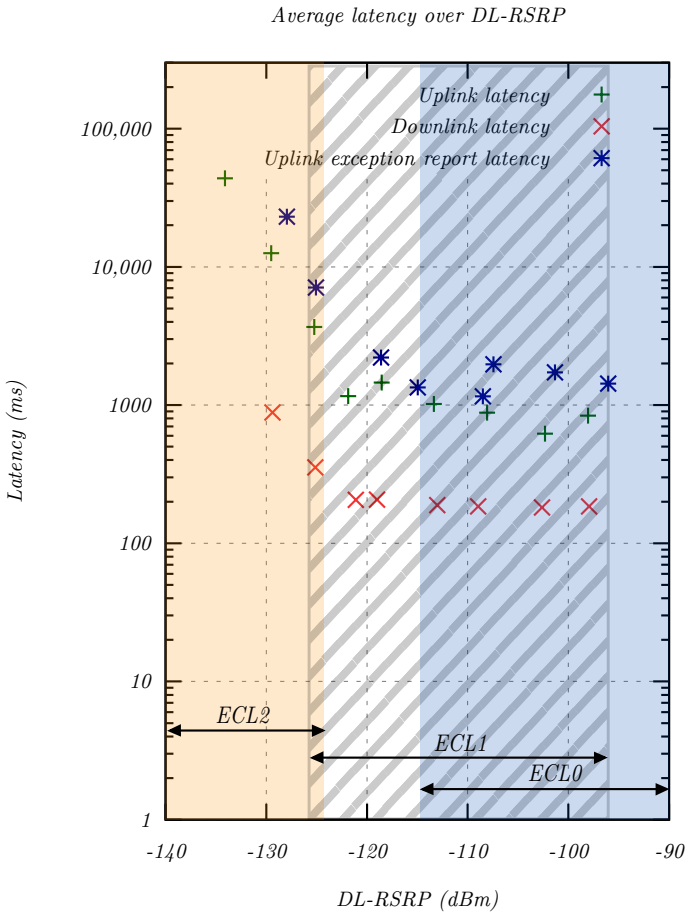


Figure 20: The total system latency as a function of RSRP signal level in the uplink and downlink direction for regular user data and exception reports.

Figure 20 confirms the relationship between the RSRP signal level and latency. In both the uplink and downlink direction, the latency rose significantly below the noise floor. There were two main reasons for this growth: (a) similar to LTE, NB-IoT used MCSs to adapt to varying channel conditions; with lower MCS, the coding redundancy was increased, and the TBS decreased, which required more RUs per data volume; (b) additionally, ECL 2 was employed for signals below the noise floor as discussed in Section 4.2.2.

The comparison of the latency of regular user traffic and exception reports showed that there was no latency benefit, and the 3GPP target of 10 s was still exceeded for signals significantly below the noise floor. Further research is required to identify what was causing this behavior. A possible reason might be, that exception reports had not yet been implemented in this NB-IoT network. Overall, the evaluation revealed a significant number of uplink samples with tens of seconds of latency in the uplink direction, which must be considered during application development.

4.2.6 *Application Layer Performance: Data Rate*

The second application layer evaluation was focused on the data rates provided by NB-IoT under different signal situations. This parameter was especially relevant for continuous data streams. The measurement considered the application layer goodput, which was the amount of user data that could be transmitted per second, as well as throughput, which also included protocol overhead of lower layers. Different signal levels were simulated in 10 dB steps using the attenuator, and 100 packets were transmitted for each attenuation step, in both uplink and downlink, as well as for different packet sizes of 8 bytes, 64 bytes, 512 bytes and 1024 bytes to simulate multiple use cases.

Figure 21a illustrates the goodput for different packet sizes and signal levels in the uplink and downlink direction. The highest data rates were possible with large packet sizes due to the comparatively small overhead. In uplink direction, 22.4 kbps could be achieved for 1024-byte packets, while the downlink provided up to 22.8 kbps. The data rates remained nearly constant until about -115 dBm RSRP, which allowed installation of NB-IoT UEs in a wide variety of locations.

For smaller packets, the maximum goodput was significantly lower in both the uplink and downlink direction. This was a result of the protocol overhead consuming larger portions of the total packet size. Furthermore, there was a large asymmetry between uplink and downlink data rates. This could be confirmed by analyzing the throughput, which consisted of the goodput plus the bandwidth consumed by the overhead:

$$\text{Throughput} = \text{Goodput} \cdot \left(1 + \frac{\text{Overhead}}{\text{Payload}}\right) \tag{39}$$

For NB-IoT, the following lower layer protocol headers must be considered: MAC 2 bytes, RLC 2 bytes, and PDCP 5 bytes [44]. For IP and UDP, which would normally consume 20 bytes and eight bytes, respectively, NB-IoT PDCP implements Robust Header Compression (ROHC) [77]. During signaling, a compression context was set up between UE and eNodeB, which reduced the size of the IP and UDP headers to a minimum of two bytes. As such, the total protocol overhead was decreased to 11 bytes. Using Equation (39), the throughput was calculated and is shown in Figure 21b.

The results showed a maximum throughput of 22.6 kbps in uplink direction, which exceeded the limits of a single tone transmission. At the same time, this value was much lower than the multi-tone maximum of 62.5 kbps. In the downlink direction, the throughput of 23.0 kbps was well aligned with the theoretical limit of 27.2 kbps (see Section 2.2.4). The comparison with the goodput highlighted that for very small packets, the overhead was the most significant part of the physical data rate. For example, an 8-byte UDP packet experienced a total overhead of 11 bytes, or 137.5 % of the payload. As such, transmissions of small packets (e.g, sensor data) could benefit from the non-IP optimizations of the NB-IoT standard, which further reduce the overhead.

Table 10: Comparison of different NB-IoT QoS parameters in a guard band deployment.

Metric	3GPP Specification [72]	Simulated Performance [44]	Measured Performance
MCL (dB)	164.0	164.0	163.5
Max. uplink throughput (kbps)	62.5 [1]	62.6 [1]	22.6 [2]
Max. downlink throughput (kbps)	27.2 [1]	26.2 [1]	23.0
Min. data rate (kbps) @ 164 dB MCL	0.160	UL: 0.320 DL: 0.370	>0.707 @ 160 dB MCL
Exception report latency (s) @ 164 dB MCL	10.0	8.0	UL: 11.2 DL: 0.610 for user traffic

[1] The specification and simulation assume 180 kHz multi-tone transmissions for maximum throughput.

[2] The uplink data rate depends on the allocation scheme and the number of subcarriers used.

4.3 DISCUSSION

In this section, the real-life performance of NB-IoT is discussed and correlated with the 3GPP requirements [72] and existing system-level simulations [44]. Table 10 provides an overview of the most relevant QoS parameters, which were used to discuss the suitability of NB-IoT in different scenarios. For example, in smart metering applications, four transmission categories were identified depending on the functionality of the end device:

- Uplink transfer of individual sensor data from a single sensor
- Downlink transfer of individual commands
- Uplink transfer of small packets from many sensors
- Transfer of software updates in the downlink and bulk data to the cloud in the uplink direction

The first transmission category corresponded to a typical IoT use case, which was a single sensor transmitting data to the cloud for further processing and storage. Sensors are commonly deployed in hard to reach locations, such as basements and metal enclosures. NB-IoT was designed to provide long-range communication with 164 dB MCL. While this coverage was confirmed in the simulations as shown in Table 10, the measurements showed that the MCL could be reduced by interference, with an observed best case MCL of 163.5 dB.

The second important QoS criterion was latency. For NB-IoT, an exception report latency of 10 s was defined, which was fulfilled in a simulated environment at 8.0 s. In most real-world applications however, the latency of regular user traffic is of interest. NB-IoT provided a low user traffic latency at high signal levels with an uplink median of 535 ms for ECL 0 transmissions. Below the noise floor, ECL 2 was employed, and the latency increased significantly to a median of 11.2 s. Since there were many samples with a latency of tens of seconds, the application must be designed accordingly. For applications that employ exception reports, no latency benefit was found, which suggested that this functionality may not have been implemented in the network under investigation. Furthermore, the high repetition count (up to 128) and corresponding high output power (up to 23 dBm) decreased the battery life of sensors without a mains connection. Since the long-term measurement revealed more stable radio conditions during night time, it might be beneficial for sensors in extreme coverage to collect sensor readings and transmit them at night. No packet loss was observed for transmissions of individual packets, which relaxed the need for application layer retransmissions.

In the second category, an actuator received commands from a control server on the Internet. The traffic patterns and QoS requirements matched those of the sensors closely, except that commands were transferred in the downlink direction. As such, most of the QoS conclusions made for the transmission of sensor data also applied to actuators. For good signal levels with ECL 0, NB-IoT provided very low median downlink latencies of 179 ms. In extreme coverage scenarios with ECL 2, the target latency was still satisfied at a median of 610 ms. As such, NB-IoT fulfilled the needs for transferring commands to actuators.

For the third category, a deployment of many sensors in a limited area was considered. One NB-IoT UE could aggregate and transmit all sensor readings. In this use case, many small packets would be sent within a short period of time, so the system bandwidth was an important QoS characteristic. For eight-byte packets, the maximum uplink goodput was limited to 712 bps, while the corresponding throughput was at 1.7 kbps. This could be explained by the significant protocol overhead of 137.5 %. As such, reducing the overhead and aggregating sensor data prior to transmission could significantly increase the efficiency. Overall, the sensor aggregation scenario was a prime use case for the NB-IoT non-IP optimizations, which were not yet supported by all networks.

The final category of sending software updates to end devices and bulk data to the cloud is challenging for many LPWAN technologies, since it requires the transmission of a large data volume in a certain time frame to prevent timeouts. As such, the maximum goodput rate is of interest. NB-IoT provided up to 22.8 kbps in the downlink direction. Table 10 shows that the corresponding throughput of 23.0 kbps was in line with theoretical calculations and simulations. In the uplink direction, the evaluation found a maximum goodput of 22.4 kbps and a throughput of 22.6 kbps. This value was significantly below the theoretical maximum of 62.5 kbps, which could be due to device limitations. However, both uplink and downlink throughput found in this evaluation significantly exceeded the values of previous measurements [14]. In both directions, the data rates were stable over a wide range of signal levels, before decreasing at an RSRP signal level of -120 dBm. Since the installation location of an IoT device can not be freely chosen in most

cases, this increases installation flexibility. In extreme coverage situations, the 3GPP specifies a minimum throughput of 160 bps at 164 dB MCL, which was confirmed in the simulations [44], but could only be approximated in the evaluation at 160 dB MCL. This aspect will be subject to further investigations in the future.

4.4 CONCLUSIONS

In this chapter, a detailed evaluation of the end-user QoS of the NB-IoT LPWAN standard was presented. One of the first systematic analyses of the relevant physical and application layer QoS parameters was conducted in a real network. Moreover, the influence of these parameters on the end-user QoS was studied in the uplink and downlink direction.

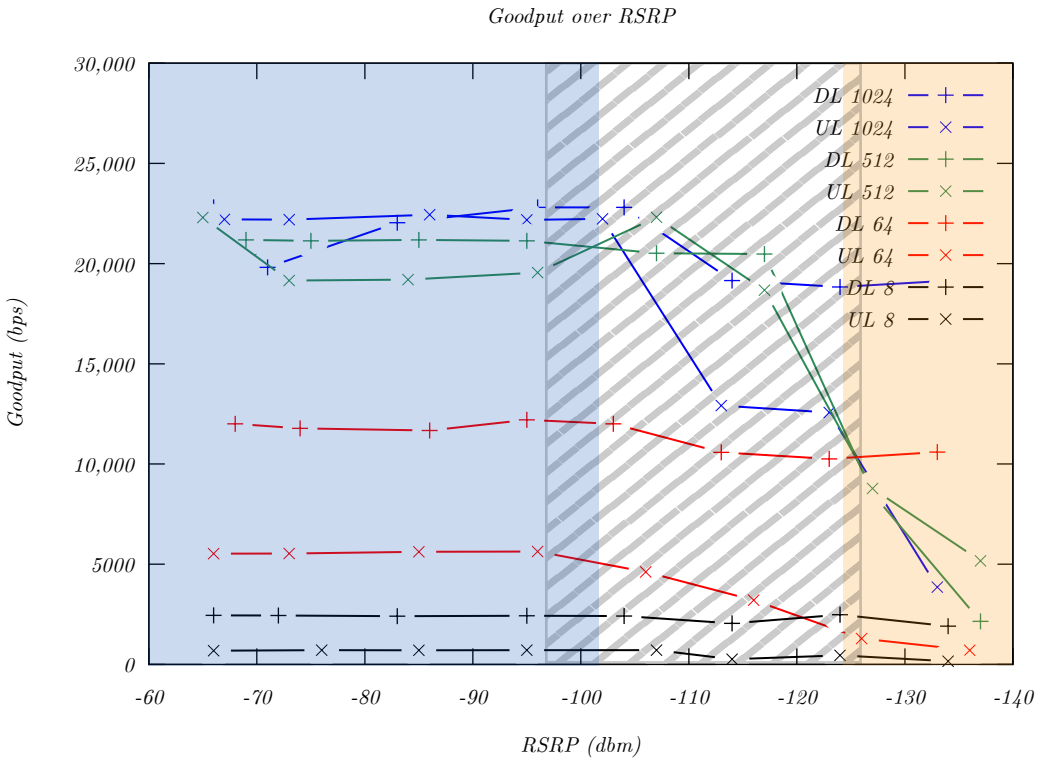
Different experiments were conducted to verify the performance of NB-IoT in a public network using commercially available hardware. Analyzing the NB-IoT coverage showed that the 3GPP target MCL of 164 dB [72] was closely approximated, but could be reduced by external interference. The necessary high receiver sensitivity was achieved by using discrete levels of signal repetitions, which were referred to as ECL. Using the highest repetition level ECL 2, signals below the noise floor could be detected, which allowed sensors installations under difficult conditions.

On the other hand, the number of ECL repetitions was a trade-off between system latency and coverage. In the downlink direction, NB-IoT maintained a latency below one second in all coverage scenarios. In the uplink, NB-IoT provided low latency for ECL 0 and ECL 1 transmissions. However, when ECL 2 was used below the noise floor, the latency increased to tens of seconds. This behavior could be attributed to the high number of signal repetitions and the allocation of individual subcarriers over a longer duration. Configuring the modem to send exception messages as specified by the 3GPP requirements did not improve the latency, which might be because the mobile network did not yet implement this functionality.

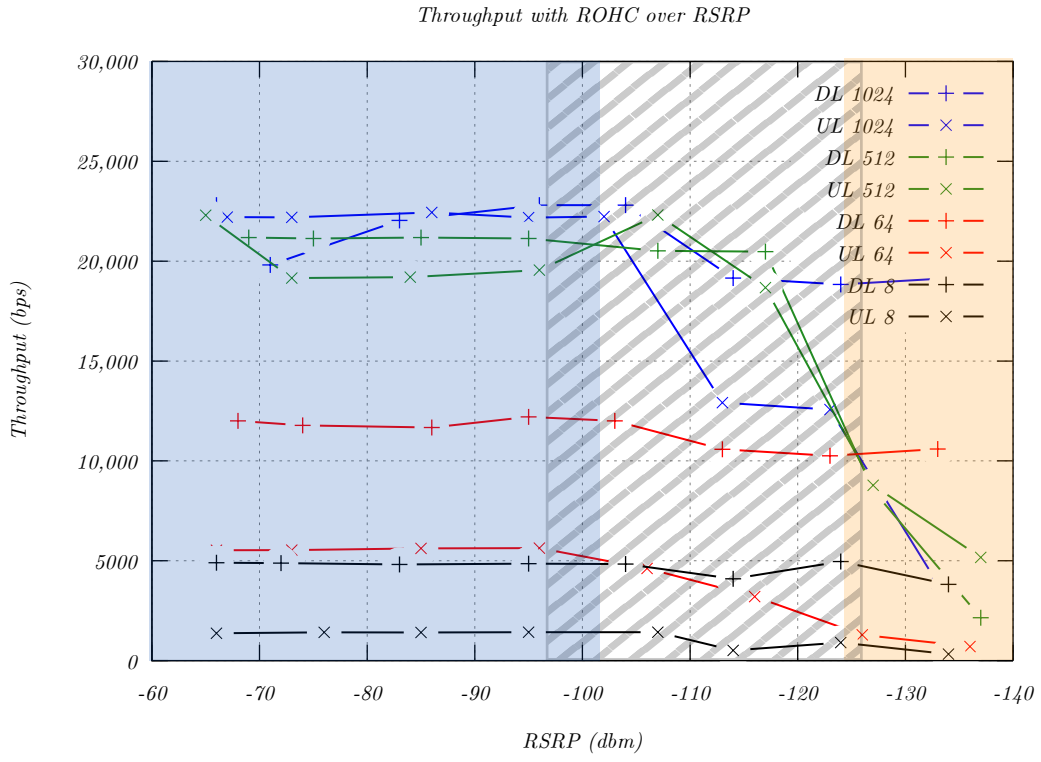
NB-IoT provided a consistent data rate over a wide range of signal levels. In the downlink direction, the throughput of 23.0 kbps was within the expected range. However, in the uplink, the data rate depended on the allocation scheme. The obtained throughput of 22.6 kbps was significantly lower than the theoretical maximum for a transmission using 12 subcarriers, which requires further investigation in a future work.

Overall, NB-IoT satisfied most of the theoretical 3GPP QoS requirements [72] in a commercial deployment, and was a suitable technology for a wide range of sensor applications. The QoS might be reduced for signal levels below the noise floor, which must be taken into consideration during application development. In the case of smart metering, four transmission categories were analyzed, and NB-IoT was verified to be appropriate if latency values above 10 s were acceptable for signal levels below the noise floor.

In a future research, NB-IoT will be compared to other LPWAN technologies in the context of smart metering use cases. This analysis will include the currently unavailable NB-IoT non-IP optimizations, as well as the 3GPP Release 14 improvements.



(a)



(b)

Figure 21: Maximum NB-IoT data rates for various packet sizes and signal levels. **(a)** Goodput excluding protocol headers for different packet sizes. **(b)** Throughput including protocol overhead for different packet sizes. The ECL regions are marked in blue (ECL 0), striped (ECL 1), and yellow (ECL 2).

5

A SYSTEMATIC QUALITY OF SERVICE ANALYSIS OF LORAWAN

LoRaWAN is an LPWAN technology that operates in the unlicensed ISM and SRD bands. It provides a low-cost, energy-efficient and long-range network service to low-end IoT endpoints. In this chapter, a detailed analysis of physical and application layer LoRaWAN QoS parameters is presented. Measurements are performed both in a deployed network, as well as in a shielded, interference-free laboratory environment. This approach permits to assess the influence of external interference on the performance of LoRaWAN, which is especially important for LPWANs that operate in the unlicensed bands. The measurements indicate that LoRaWAN meets most theoretical specifications in a setup built from commercial hardware. At the same time, both the coverage range and the packet delivery ratio is negatively influenced by the elevated levels of interference in the EU-868 band used in the measurements. LoRaWAN employs SFs to improve the sensitivity; since the ToA only accounts for part of the total latency, messages arrive in a reasonable time frame even at high SFs. LoRaWAN provides very low data rates, which are limited by the duty cycle regulations. After the measurements, the LoRaWAN applicability is discussed in typical WSN use cases, and it is confirmed to be appropriate for low-end IoT applications without hard requirements on data rate and packet loss. Finally, LoRaWAN is compared to NB-IoT for different QoS parameters.

5.1 METHODS

Since LoRaWAN operates in the unlicensed bands, a private LoRaWAN network can be deployed to provide control over the network parameters and obtain PHY measurements from both the nodes and the gateways. Installing a private network is essential for QoS measurements, since public networks such as The Things Network [97] commonly implement fair use policies, which limit the number of messages that a LoRaWAN node can send per day. This reduces the QoS even beyond the limitations already imposed by the duty cycle regulations. Therefore, a private LoRaWAN network was deployed for the measurements presented in this chapter. This setup represents a best case scenario; in large-scale LoRaWAN networks, collisions with uplink messages from other nodes and the downlink capacity of the gateway limit the performance (see Section 2.3.6).

The measurements in this chapter were performed in a private LoRaWAN network operating in the EU-868 band in Germany[1]. Two measurement architectures were used: first, the network was deployed at the university campus for realistic radio conditions. Afterwards, the measurements were repeated in a shielded environment to explore the system performance boundaries without external interference.

1 The plots and analyses in this chapter are created from the raw measurement data obtained in a Master's Thesis [58] that was supervised by Ulrich Birkel and Andreas Philipp Matz as part of this PhD thesis.

All measurements in this chapter were performed using the LoRaWAN 1.0.2 specification; while the ChirpStack LoRaWAN network stack supported both LoRaWAN 1.1 and 1.0, the Microchip RN2483 LoRaWAN node was limited to LoRaWAN 1.0.2.

Both the LoRaWAN node and the gateway performed PHY measurements whenever a transmission occurred. The measurements performed by the gateway were automatically attached to the message as metadata and could be obtained from the LoRaWAN backend; the node measurements needed to be manually obtained from the modem. In contrast to LTE, only a few basic measurements were available, which limited the amount of data from which conclusions could be drawn. Application layer measurements could still be made as usual within regulatory limits in terms of power and duty cycle.

During the measurement process, some architecture limits were discovered that limited the QoS analysis. A private LoRaWAN network should provide full control over typical LoRaWAN parameters such as the SF. In practice, the parameters that can be configured depend on the LoRaWAN stack in use. The LoRaWAN network used in the measurements was based on the Chirpstack [21] due to its advanced features. The Chirpstack did not allow deactivating the ADR adjustment in the downlink direction; rather, the gateways would continuously adjust the SF and the transmit power according to the channel conditions, which prevented collecting 30 measurements at the same SF in a reasonable time frame and therefore limited the analysis to uplink transmissions.

While the physical channel can be considered to be symmetric, the selected LoRaWAN class creates asymmetric MAC layer conditions. For example, a class A device can only receive downlink data in two RX windows after an uplink transmission, while a class C device can receive downlink messages at all times, except when uplink traffic is sent (half-duplex operation). A future analysis of the LoRaWAN downlink must therefore include measurements of all classes at different spreading factors.

5.2 EXPERIMENTAL SETUP AND RAW DATA

In the first experiment, the gateway was installed on top of the university building. The LoRaWAN node was placed in a neighbor building in Line-of-Sight (LOS) conditions in a distance of ca. 36 m to ensure a high signal level. The architecture in Figure 22 consisted of the following components:

(a) A *measurement computer* ran the QoS analysis application. It controlled the LoRaWAN modem (b) via USB and collected the measurement data from the Application Server (h) via MQTT.

(b) A *Microchip RN2483 LoRaWAN node* [59] in LoRaWAN class A mode was connected to a private LoRaWAN network and sent test messages to the LoRaWAN gateway (f). It implemented the LoRaWAN v1.0.2 specification [48].

(c) An *RF shielding box* prevented the LoRa signal from bypassing the attenuator and coupling directly into the PCB of the modem.

(d) A *Wavetek 5080.1 step attenuator* (0–81 dB attenuation, $f_{max} = 1000$ MHz) simulated a wide range of signal levels.

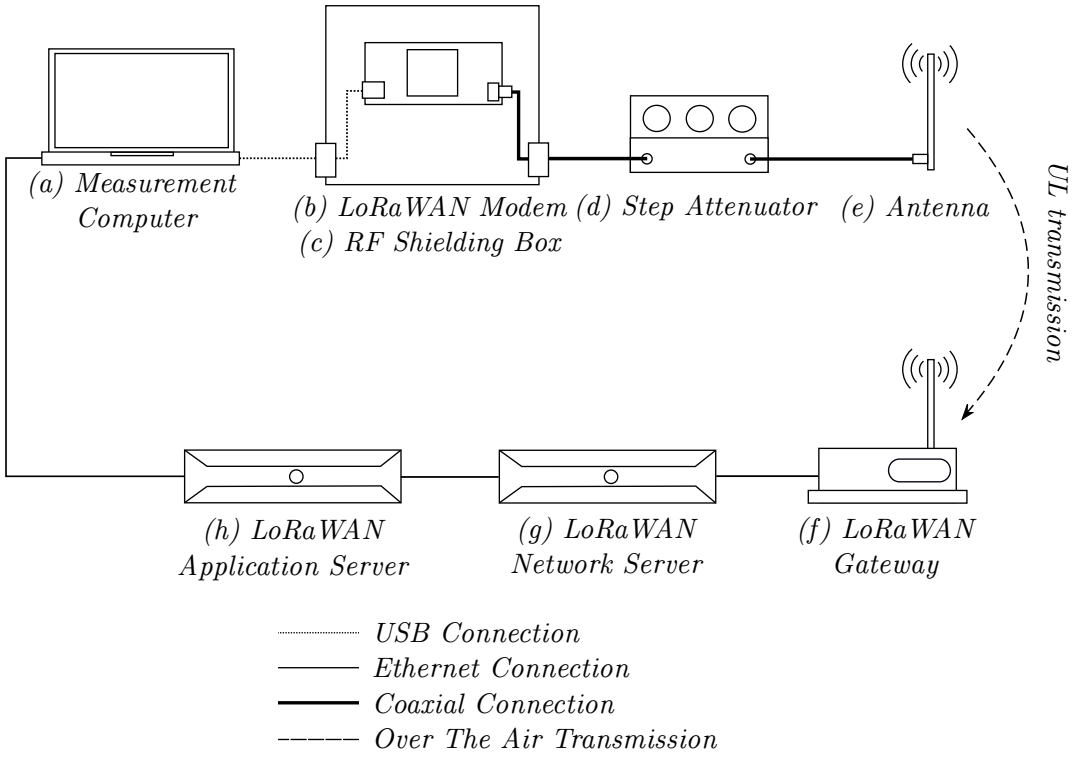


Figure 22: LoRaWAN measurement setup in a deployed, private LoRaWAN network.

(e) A *Delock 88571 omnidirectional antenna* (A = –0.8 – 0.2 dBi) was connected to the step attenuator via coaxial cable.

(f) A *Multitech Conduit LoRaWAN gateway* with an MTAC-LoRa-H 1.5 module was placed on top of the university building. It provided local coverage and performed QoS measurements in the uplink direction. The QoS parameters were attached to the message as metadata and forwarded to the network server (g).

(g) A *LoRaWAN network server* managed all connected devices and forwarded messages between the gateway (f) and the application server (h). It was part of the Chirpstack [21], which integrated all necessary backend components to operate a private LoRaWAN network.

(h) An *application server* was also part of the Chirpstack platform. It forwarded messages and provided external interfaces to retrieve the measurement data via MQTT.

For each measurement, 30 messages were sent from the LoRaWAN node (b) to the application server (h) under consideration of the duty cycle regulation in the EU-868 band. The messages were collected via MQTT by the application installed on the measurement computer (a). After each measurement, the attenuation (d) was increased by 10 dB; this process was repeated until the connection was lost.

In order to cover a wide range of potential use cases, the measurements were conducted for all combinations of the following payload sizes and SFs:

- Payload sizes
  - 8 bytes (representative of a sensor value)
  - 51 bytes (MTU of a LoRaWAN SF 12 connection)
  - 222 bytes (MTU of a LoRaWAN SF 7 connection, not available at SF 12)
- Spreading factor
  - SF 7 (provides the highest QoS but the lowest range)
  - SF 12 (provides the lowest QoS but the highest range)

The one-way application layer latency was measured by inserting a timestamp into each packet, which was compared to the system time when the packet arrived back at the laptop. A sequence number was used to track the packet loss. This information was captured along with the individual packets and their metadata.

In the second experiment, the LoRa wireless link was shielded to exclude external interference as shown in Figure 23. Most components remained unchanged; however, the antennas and the wireless link were replaced by coaxial cables and attenuators and the shielding was improved. The modified architecture consisted of the following elements:

(a) A *measurement computer* ran the QoS analysis application. It controlled the LoRaWAN modem (b) via USB and collected the measurement data from the Application Server (i) via MQTT.

(b) A *Microchip RN2483 LoRaWAN node* [59] in LoRaWAN class A mode was connected to a private LoRaWAN network and sent test messages to the LoRaWAN gateway (f). It implemented the LoRaWAN v1.0.2 specification [48].

(c) An *RF shielding box* prevented external interference from coupling directly into the LoRaWAN node PCB.

(d) A *Wavetek 5080.1 step attenuator* (0–81 dB attenuation, $f_{max} = 1000$ MHz) simulated a wide range of signal levels.

(e) Multiple *static attenuators* between the step attenuator and the gateway simulated the free space loss that would be present in a deployed network. Due to the high MCL of LoRa, the attenuation was spread across multiple attenuators to avoid the signal coupling over the attenuator and back into the coaxial cable.

(f) A *Multitech Conduit LoRaWAN gateway* with an MTAC-LoRa-H 1.5 module provided the LoRa signal and performed QoS measurements in the uplink direction. The QoS parameters were attached to the message as metadata and forwarded to the network server (h).

(g) An *RF shielding box* improved the attenuation of the LoRaWAN gateway housing against external interference, which was especially important at high coupling loss.

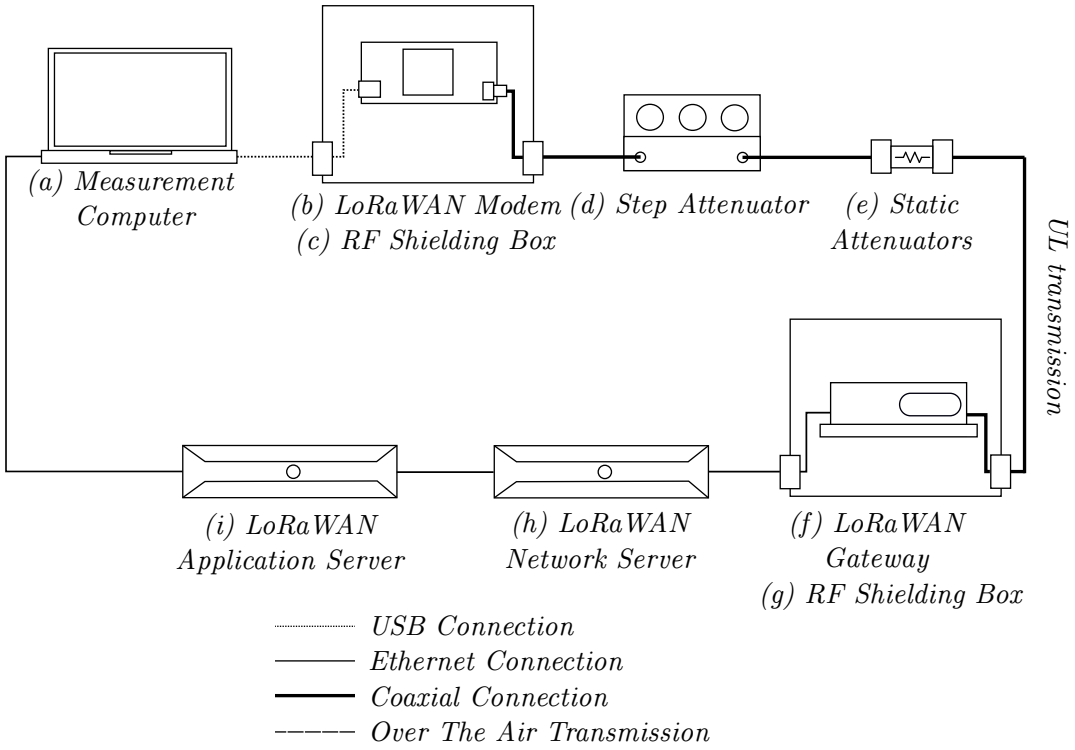


Figure 23: LoRaWAN measurement setup in a shielded, interference-free laboratory setup.

(h) A *LoRaWAN network server* managed all connected devices and forwarded messages between the gateway (f) and the application server (h). It was part of the Chirpstack [21], which integrated all necessary backend components to operate a private LoRaWAN network.

(i) An *application server* was also part of the Chirpstack platform. It forwarded messages and provided external interfaces to retrieve the measurement data via MQTT.

The measurement parameters remained unchanged from the first evaluation. In the second experiment, special attention was given to how the interference influenced packet loss and signal quality in terms of RSSI and SNR.

5.3 RESULTS

In this section, the evaluation results are illustrated and analyzed. The physical and application layer QoS is discussed using the approach developed during the NB-IoT measurements (see Section 4.2).

5.3.1 *PHY Measurements: Coupling Loss*

Since the LoRa modulation does not employ reference symbols, the RSRP is not available to express the signal level. Rather, the RSSI is used to represent the total power level in the channel bandwidth B, including the signal power P_S, the effective noise floor $P_{N,eff}$ and the interference power P_I. In the first measurement, the

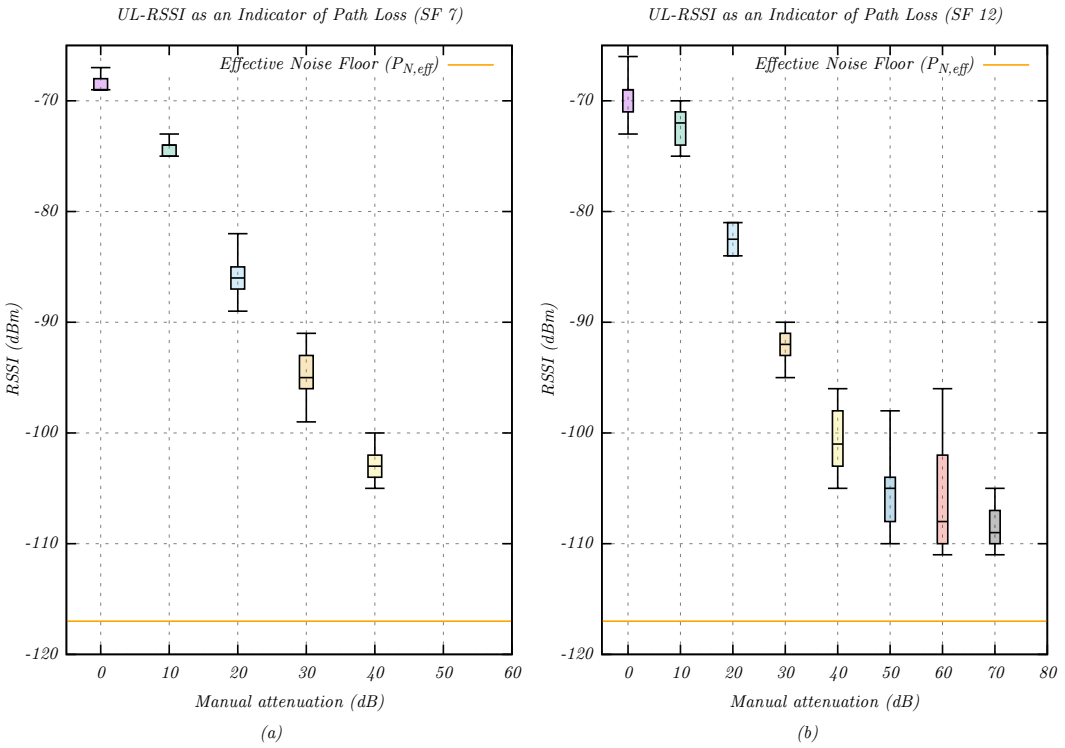


Figure 24: Distribution of RSSI as an expression of artificial path loss, using 8 bytes of payload **(a)** at SF 7 and **(b)** at SF 12.

RSSI was explored as a representation of path loss for later measurements. Using a step attenuator, a wide range of signal levels were simulated in 10 dB steps until the connection was lost. For each step, 30 packets were sent. For each packet, the UL-RSSI was measured by the LoRaWAN gateway and recorded.

Figure 24a illustrates the UL-RSSI for SF 7 at different signal levels. The RSSI followed the manual attenuation linearly, but at a reduced slope, i.e. 10 dB of manual attenuation reduced the RSSI by less than 10 dB. This is not unexpected, since the RSSI parameter measured asymptotically along the packet [104] is known to exhibit some linearity limitations and to diverge from the ideal 1 dB/dB curve [89]. The signal was lost before approaching $P_{N,eff} = -117$ dBm (as calculated in Equation (31)).

Figure 24b shows the same measurement for SF 12. The RSSI converged to the sum of the $P_{N,eff}$ and the P_I, both of which are included in the RSSI measurement. It was thus not a linear expression of the path loss. To improve the path loss approximation, the measurement needed to be repeated in a shielded, interference-free environment that excluded the interference P_I. Even then, it would be limited by noise, since the RSSI is not able to take advantage of the CE mechanism to measure signal levels below the $P_{N,eff}$.

In the second step, the measurement was repeated in a shielded, interference-free environment. This removed the unpredictable influence of the interference P_I. Once again, different signal levels were simulated in 10 dB steps; additionally, one measurement at 75 dB attenuation was conducted to further approach the MCL. 30 packets were sent per step, and the UL-RSSI and UL-SNR were recorded. Since

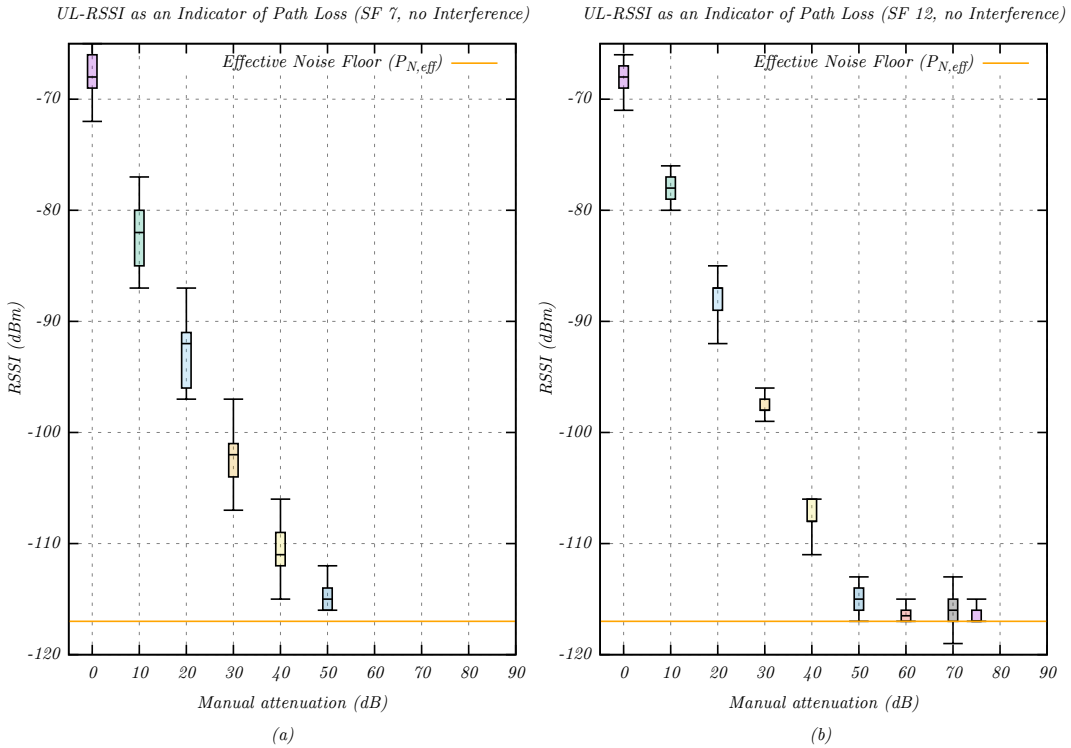


Figure 25: Distribution of RSSI in a shielded environment as an expression of artificial path loss, using 8 bytes of payload **(a)** at SF 7 and **(b)** at SF 12.

the RSSI included signal, noise and interference components, it was expected to converge to the effective noise floor $P_{N,eff}$.

Figure 25a illustrates a measurement performed at SF 7 in a shielded, interference-free environment. Similar to the previous evaluation in the first step, the RSSI followed the manual attenuation but was lost shortly before reaching the noise floor.

Figure 25b illustrates the same measurement at SF 12. As expected, the RSSI approached the $P_{N,eff} = -117\,\mathrm{dBm}$. While this was an improvement over the first measurement, the nonlinearity still makes the RSSI unsuitable to represent path loss by itself. Notably, there were a number of samples at 70 dB attenuation ($RSSI = -119\,\mathrm{dBm}$) below the $P_{N,eff}$. This could either have been a result of the limited measurement precision, or the modem exceeded its $NF = 6\,\mathrm{dB}$ specification, which in turn would also have reduced the $P_{N,eff}$.

The previous evaluations revealed that the RSSI is not an appropriate representation of path loss. Semtech proposes another measurement that addresses this challenge: the *Packet Strength* is equal to the RSSI until the signal falls below the noise floor ($SNR < 0\,\mathrm{dB}$), at which point it includes the SNR measurement to approximate the signal level [89].

$$\mathrm{PacketStrength} = \begin{cases} \mathrm{RSSI} & \mathrm{SNR} \geqslant 0 \\ \mathrm{RSSI} + \mathrm{SNR} & \mathrm{SNR} < 0 \end{cases} \tag{40}$$

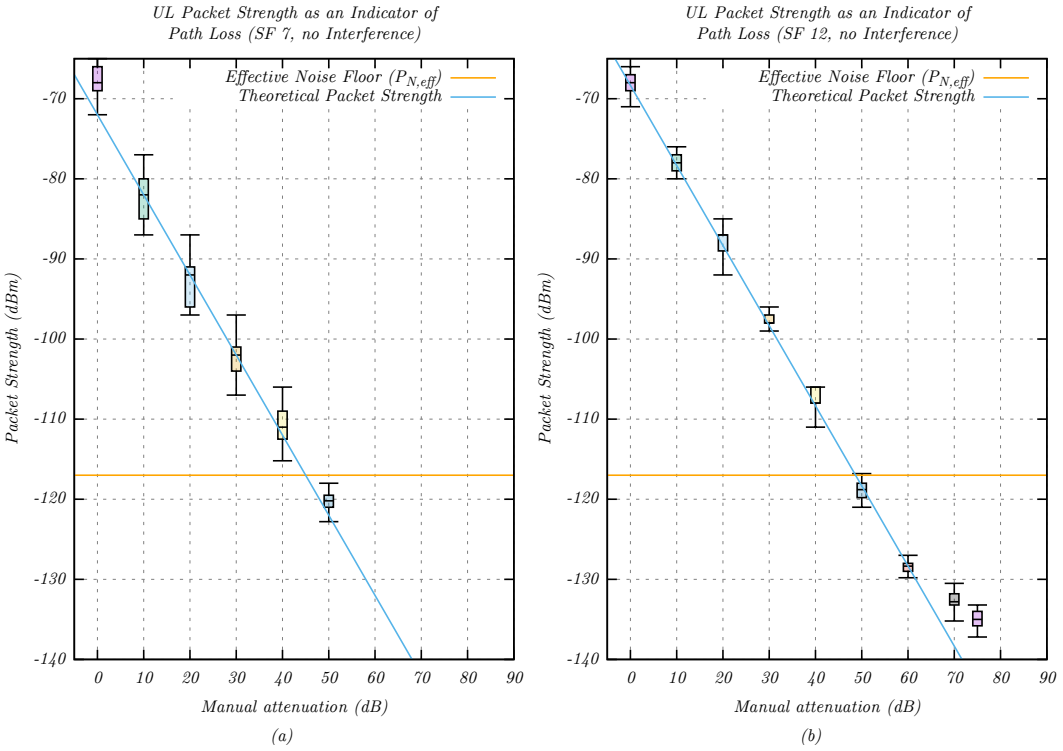


Figure 26: Distribution of Packet Strength in a shielded environment as an expression of artificial path loss, using 8 bytes of payload **(a)** at SF 7 and **(b)** at SF 12.

Figure 26a illustrates the Packet Strength at SF 7 as an indicator of path loss in an interference-free environment. Unlike the RSSI, the Packet Strength is able to approximate the signal level below the noise floor.

Figure 26b visualizes the same measurement for an SF 12 link. There are deviations from linearity at very low signal levels below the noise floor, which can be attributed to a combination of shielding limitations and limited measurement precision.

Table 11 provides statistical detail on the RSSI and Packet Strength measurements for different levels of attenuation. For each attenuation step, 30 packets of 8 bytes payload were transmitted. For both SF 7 and SF 12, the dominating factor that limits the RSSI linearity is the $P_{N,eff}$. Especially the SF 12 measurement converged to the $P_{N,eff}$, with a $\mu_{RSSI,min} = -119$ dBm at 75 dB attenuation. Despite the shielded setup excluding most interference, there was a considerable variation in the RSSI levels ($\sigma_{RSSI,max,SF7} = 3$ dB, $\sigma_{RSSI,max,SF12} = 2.6$ dB). At the same time, the Packet Strength measurement provided a much better approximation of the path loss. SF 12 connections in particular benefit from the ability to represent signals below the noise floor; however, there are linearity limitations below a Packet Strength of -130 dBm. Starting from 70 dB attenuation, shielding imperfections and measurement limitations can influence the measured signal strength levels. For the maximum attenuation of 75 dB, there are additional censorship effects, i.e. packets with lower SNR are more likely to be lost, which causes an overestimation of the average SNR [23].

Table 11: Statistical comparison of the mean values (μ_{RSSI}, μ_{PS}) and standard deviations (σ_{RSSI}, σ_{PS}) of the RSSI and Packet Strength measurements in a shielded environment. For each attenuation step, 30 packets of 8 bytes payload were transmitted.

SF	Attenuation (dB)	μ_{RSSI} (dBm)	σ_{RSSI} (dB)	μ_{PS} (dBm)	σ_{PS} (dB)
7	0	−67.690	2.601	−67.690	2.601
7	10	−82.100	2.879	−82.100	2.879
7	20	−92.533	3.008	−92.533	3.001
7	30	−102.310	2.627	−102.310	2.627
7	40	−110.933	2.337	−111.010	2.422
7	50	−114.810	1.096	−120.195	1.723
12	0	−68.333	1.972	−68.333	1.972
12	10	−78.200	2.135	−78.200	2.135
12	20	−88.267	2.581	−88.267	2.581
12	30	−98.133	1.607	−98.133	1.607
12	40	−107.833	1.951	−107.833	1.951
12	50	−114.933	1.123	−118.703	1.119
12	60	−116.500	0.764	−128.530	0.797
12	70	−116.100	1.274	−132.627	1.304
12	75	−116.684	1.079	−134.916	1.107

Overall, the Packet Strength seems to be an appropriate representation of path loss; when evaluating very low signal levels, the nonlinearity must be considered when approaching the modem sensitivity.

5.3.2 *PHY Measurements: Coverage Extension*

The LoRa modulation employs spreading factors to extend the coverage at the expense of the latency and data rate. In this section, the coverage is analyzed in terms of the MCL. To evaluate whether the large interference present in unlicensed bands had a significant influence on the MCL, both the measurements from the deployed network and the shielded setup were used. The Multitech MTAC-LORA-H 1.5 card is based on the Semtech SX1301 digital baseband chip [104] and SX1257 transceiver [88] for which the typical sensitivity values are shown in Table 12. Furthermore, the MCL was calculated using the maximum permitted transmit power $P_{TX,modem} = 14\,dBm$ [20]. In the measurements, the samples with the minimum observed Packet Strength are used to find the maximum MCL.

In a first step, the measurements from the deployed network are used to determine the MCL under realistic conditions. For SF 7, the lowest observed RSSI sample $RSSI_{min,obs,SF7} = -105\,dBm$ and its corresponding SNR value $SNR_{min,obs,SF7} = -7.8\,dB$ were considered. The resulting MCL value revealed a considerable reduction from the specified LoRa SF 7 MCL of 140.5 dB [104].

Table 12: Properties of the SX1301 reference implementation: $P_{RX,min}$ [104] and MCL calculated for $P_{TX} = 14$ dBm.

SF	$P_{RX,min}$ (dBm)	MCL (dB)
7	−126.5	140.5
8	−129.0	143.0
9	−131.5	145.5
10	−134.0	148.0
11	−136.5	150.5
12	−139.5	153.5

$$\begin{aligned} MCL_{deployed,SF7} &= P_{TX,modem} - PacketStrength_{min,SF7} \\ &= P_{TX,modem} - (RSSI_{min,SF7} + SNR_{min,SF7}) \\ &= 14\,\text{dBm} - (-105\,\text{dBm} - 7.8\,\text{dB}) = 126.8\,\text{dB} \end{aligned} \tag{41}$$

Performing the same calculation for SF 12 provides similar results. In this case, the minimum observed RSSI is $RSSI_{min,obs,SF12} = -111$ dBm and its corresponding SNR is $SNR_{min,obs,SF12} = -17.2$ dB. Once again, the observed MCL did not reach the specified LoRa SF 12 MCL of 153.5 dB [104].

$$\begin{aligned} MCL_{deployed,SF12} &= P_{TX,modem} - PacketStrength_{min,SF12} \\ &= P_{TX,modem} - (RSSI_{min,SF12} + SNR_{min,SF12}) \\ &= 14\,\text{dBm} - (-111\,\text{dBm} - 17.2\,\text{dB}) = 142.2\,\text{dB} \end{aligned} \tag{42}$$

Afterwards, the same analysis was performed using the measurements from the shielded environment. At SF 7, the RSSI sample $RSSI_{min,obs,SF7} = -116$ dBm and its corresponding SNR value $SNR_{min,obs,SF7} = -7.8$ dB were selected to calculate the MCL. The observed MCL approached the specified LoRa SF 7 MCL of 140.5 dB [104].

$$\begin{aligned} MCL_{shielded,SF7} &= P_{TX,modem} - PacketStrength_{min,SF7} \\ &= P_{TX,modem} - (RSSI_{min,SF7} + SNR_{min,SF7}) \\ &= 14\,\text{dBm} - (-116\,\text{dBm} - 7.8\,\text{dB}) = 137.8\,\text{dB} \end{aligned} \tag{43}$$

For SF 12, the same calculation was performed. The minimum observed RSSI is $RSSI_{min,obs,SF12} = -119$ dBm and its corresponding SNR is $SNR_{min,obs,SF12} = -18.2$ dB. Again, the observed MCL approached the specified LoRa SF 12 MCL of 153.5 dB [104].

$$\begin{aligned} MCL_{shielded,SF12} &= P_{TX,modem} - PacketStrength_{min,SF12} \\ &= P_{TX,modem} - (RSSI_{min,SF12} + SNR_{min,SF12}) \\ &= 14\,\text{dBm} - (-119\,\text{dBm} - 18.2\,\text{dB}) = 151.2\,\text{dB} \end{aligned} \tag{44}$$

Overall, the analysis in this section revealed a considerable influence of the interference on the MCL of LoRaWAN. This is not unexpected, since the shared nature

of the unlicensed bands creates a harsh environment that needs special techniques to mitigate interference from other users of the band.

5.3.3 *PHY Measurements: Signal Quality Analysis*

In this section, the IM and the NF are analyzed using the SNR and Packet Strength. The measurements were performed in two environments; first, a deployed LoRaWAN network was used to provide realistic signal conditions. Second, a shielded environment was used to exclude interference. In both cases, a step attenuator was used to simulate different signal levels in 10 dB steps until the connection was lost. For each attenuation level, 30 packets were sent in uplink direction and the SNR and RSSI were measured by the LoRaWAN gateway and recorded. Similar to previous experiments, measurements were performed for SF 7 using 8-byte, 51-byte and 222-byte packets, as well as SF 12 with 8-byte and 51-byte packets.

Figure 27 compares the measured UL-SNR to the theoretical SNR as a function of the Packet Strength. The theoretical SNR was calculated for a system bandwidth $B = 125\,\text{kHz}$ using the parameters specified by the manufacturer ($NF = 6\,\text{dB}$). A $SNR = 0\,\text{dB}$ meant that the signal power P_S was equal to the effective noise power $P_{N,eff,125\,\text{kHz}} = -117\,\text{dBm}$. Using Equation (40), the Packet Strength for $SNR = 0\,\text{dB}$ can be calculated:

$$\begin{aligned} PacketStrength_{SNR=0\,\text{dB}} &= RSSI = P_S + P_{N,eff} + P_I \\ &= -117\,\text{dBm} + (-117\,\text{dBm}) + P_I = -114\,\text{dBm} + P_I \end{aligned} \tag{45}$$

The measurement revealed a significant influence of the interference on the Packet Strength, which was expected due to its RSSI component. The SNR curve was thus shifted towards higher Packet Strength values. The interference could be quantified using the IM, a measurement defined as the P_I at $SNR = 0\,\text{dB}$. It varied from $IM_{min} = 7\,\text{dB}$ to $IM_{max} = 20\,\text{dB}$ between measurements, which was expected in the ISM bands. With increasing signal levels, the SNR approached a limit of ca. 12 dB, which was expected according to the documentation provided by the manufacturer; SNR levels above 5 dB were described as "meaningless, and should be considered as there being 'plenty of signal'" [91]. The SNR parameter was able to represent signal levels below the noise floor, which indicated that it employed the CE mechanism to determine the signal power P_S.

Figure 28 illustrates the same measurement in a shielded and interference-free environment. The IM was reduced to a range of $IM_{min} = 0\,\text{dB}$ to $IM_{max} = 4\,\text{dB}$. The large number of samples beyond the theoretical SNR line could be caused by the modem exceeding its specified $NF = 6\,\text{dB}$, which would lower the $P_{N,eff}$ and result in a $PacketStrength_{SNR=0\,\text{dB}} < -114\,\text{dBm}$, thus shifting the theoretical SNR to lower Packet Strength values. Similar to what was observed in the path loss measurements, there were linearity limitations that reduced the slope of the SNR measurements. As previously observed in the deployed LoRaWAN network, the UL-SNR did not exceed 12 dB.

Figure 29 compares the UL-SNR between the shielded environment and the deployed network. The measurement verified the significant amounts of interference that LPWANs generally face in unlicensed bands, which commonly limits the per-

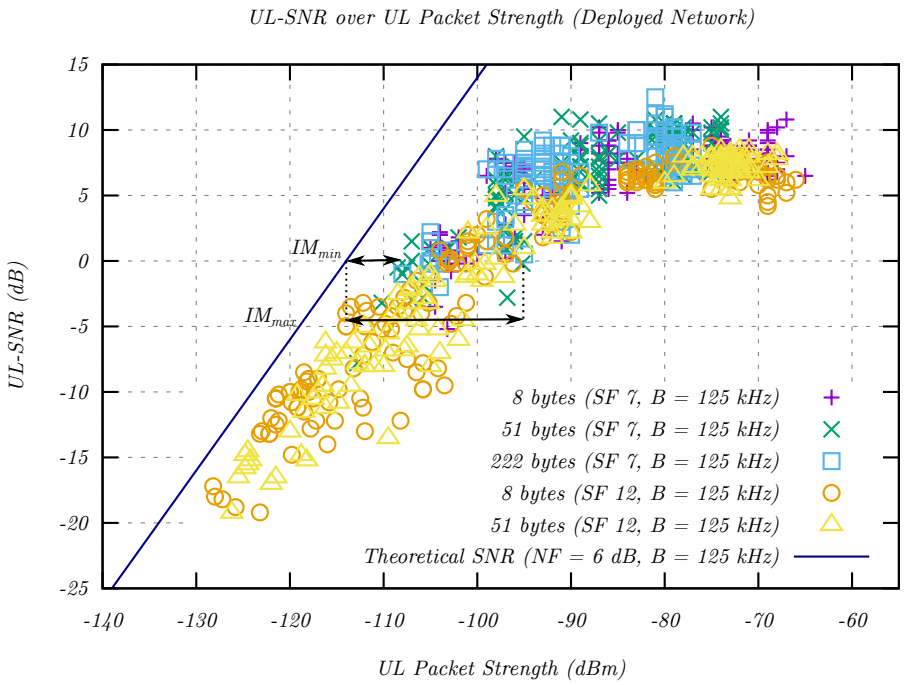


Figure 27: Analysis of the UL-SNR as a function of the Packet Strength under the influence of external interference for different SFs and packet sizes. The solid line represents the calculated SNR for NF = 6 dB.

formance especially in terms of the MCL. External interference must thus be considered when analyzing whether an LPWAN technology is fit for a particular application.

5.3.4 *Application Layer Performance: Latency*

LoRaWAN is subject to duty cycle regulations which limit applications to infrequent transmissions. As a result, LoRaWAN is commonly used for WSNs, where individual messages are transmitted in regular intervals. The one-way latency is one of the most relevant QoS metrics, since it determines the reaction time of the system.

In this section, an analysis is presented to identify the key factors that influence the uplink LoRaWAN latency. The measurements were performed in a deployed network to provide realistic results. For SF 7, three packet sizes of 8 bytes, 51 bytes and 222 bytes were analyzed. For SF 12, only 8 bytes and 51 bytes were supported. For each combination of packet size and SF, 30 packets were sent in uplink direction and the latency, the RSSI, and the SNR were recorded.

Figure 30 compares the latency distribution for SF 7 and SF 12 at different packet sizes. As expected, choosing a higher SF increased the end-to-end latency, since the LoRa chirp signal was spread over a longer duration to improve the sensitivity. For SF 7, an 8-byte message could be delivered in 362 ms, which enabled timely reaction to changing sensor values and alarm messages. For SF 12, the improved sensitivity came at the cost of an increased latency of 2029 ms. This measurement highlighted an important characteristic of the SF mechanism: Even for applications

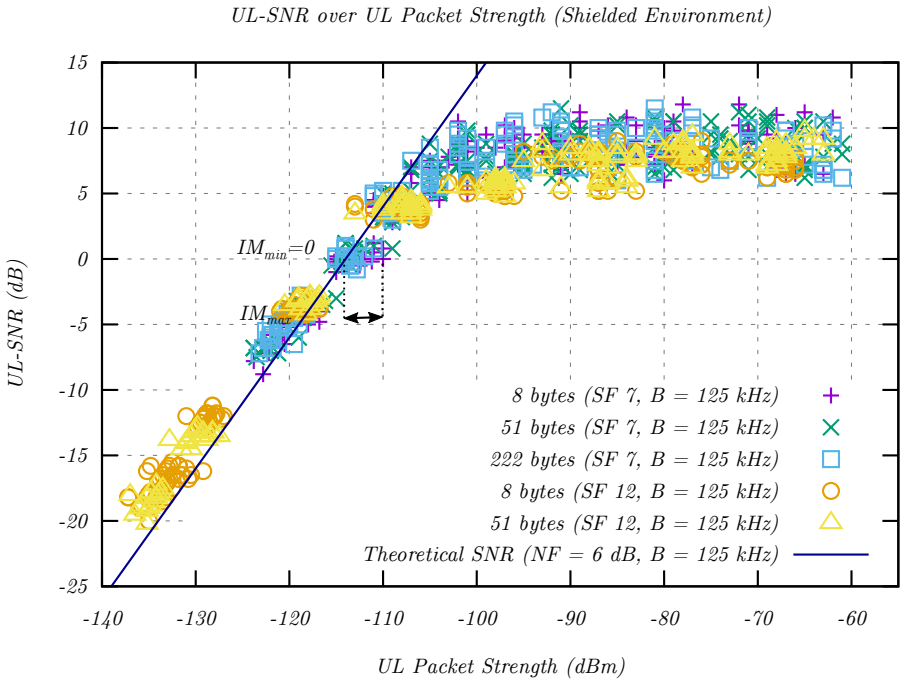


Figure 28: Analysis of the UL-SNR as a function of the Packet Strength in an interference-free environment for different SFs and packet sizes. The solid line represents the calculated SNR for NF = 6 dB.

that only transmit very little data, a low SF was desirable since it decreased the end-to-end latency.

For larger packets, more information had to be encoded which increased the time on air. To avoid excessive ToA and inter-packet wait times, LoRaWAN limits the MTU at high SFs. A packet filling the SF 7 MTU of 222 bytes took 1806 ms on average, while the SF 12 MTU of 51 bytes was transmitted in 3435 ms.

Figure 31 compares the latency of different SFs and packet sizes as a function of the Packet Strength. The SFs are the primary CE mechanism of LoRaWAN. Therefore, decreasing the signal level at a fixed SF did not increase the latency, since the time on air remained the same. In most networks, the ADR mechanism manages the SF, bandwidth and transmission power; in these situations, decreasing the signal power may trigger a switch to a higher SF, which in turn increases the latency.

Table 13 provides a closer look at the individual components of the end-to-end latency. The ToA is defined as the time that the message is transmitted between the LoRaWAN node and the gateway. It is deterministic and can be calculated using the configuration-dependent LoRaWAN header size and the payload length. There is a number of programs to automatically calculate the ToA; a popular tool is the *LoRa Air Time Calculator* [52], which was used to calculate the ToA in this Table.

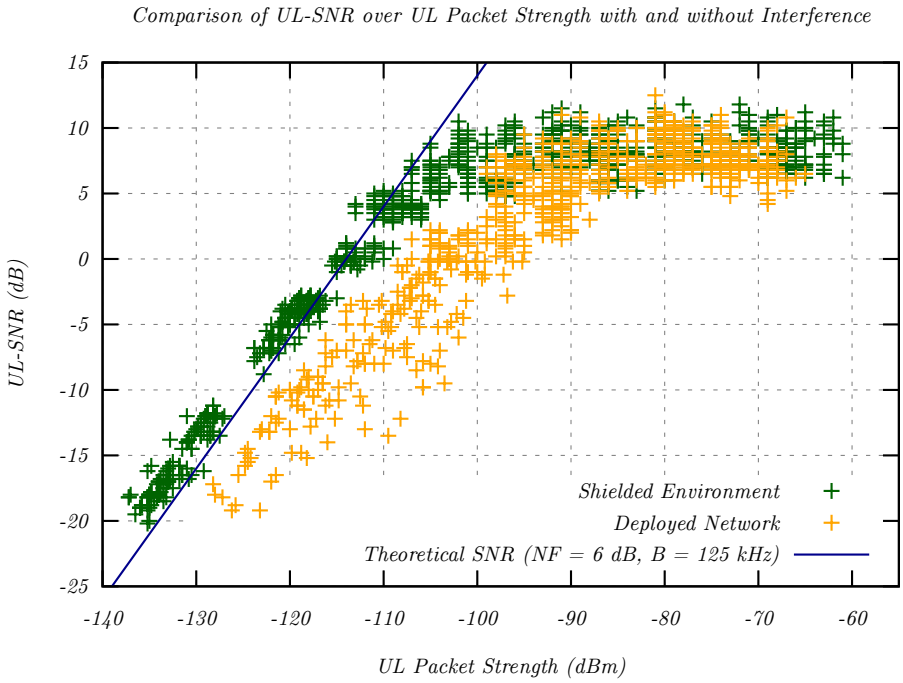


Figure 29: Comparison of the UL-SNR as a function of the Packet Strength with and without the influence of external interference. The solid line represents the calculated SNR for NF = 6 dB.

The remaining latency was named "network latency" and could be further divided into the following elements:

- The processing delay until the LoRaWAN node started transmitting the message
- The processing delay of the LoRaWAN gateway and the LoRaWAN stack
- The network delay caused by forwarding the packet to the measurement computer.

The network delay between the measurement computer and the LoRaWAN stack could be neglected, since they resided on the same physical Ethernet network. The remaining processing delays correlated with the packet size and the spreading factor; for identical packet sizes, a higher SF introduced a higher network delay, which indicated a higher processing delay. For SF 7, a higher packet size also introduced a higher network delay; this effect could not be observed for SF 12. Further research is required to determine the magnitude of the remaining latency components.

5.3.5 *Application Layer Performance: Data Rate*

Besides limiting the frequency with which messages can be sent, the duty cycle regulations also influence the maximum data rates that LoRaWAN can provide. While WSNs and their typical traffic patterns of regular, small messages are a common

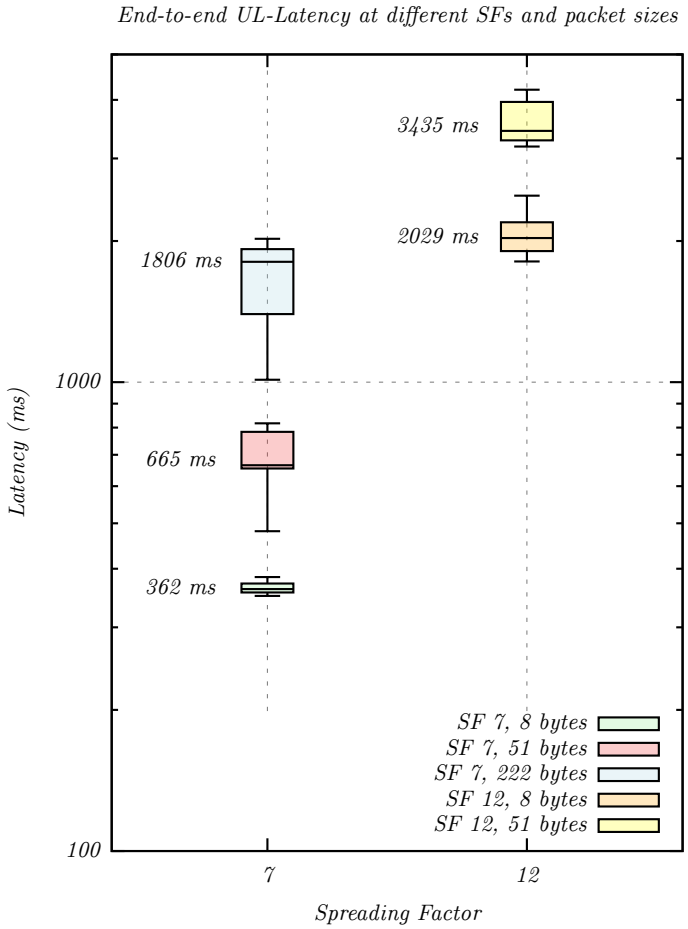


Figure 30: The influence of the spreading factor on the total system latency at different packet sizes. The median latency is given for each packet size and spreading factor.

use case for LoRaWAN, a systematic data rate measurement is still important to determine whether a particular application can be supported.

In this section, the theoretical data rates calculated in Section 2.3.10 are verified in a deployed network. For this purpose, both the goodput and the throughput were analyzed. The previous definition in Equation (39) is reused:

$$\text{Throughput} = \text{Goodput} \cdot \left(1 + \frac{\text{Overhead}}{\text{Payload}}\right) \tag{46}$$

The expected goodput and throughput under EU-868 regulations are deterministic and can be calculated. For this purpose, the ToA from Table 13 was used to calculate the required packet interval ($T_{interval}$) for a 1 % duty cycle regulation.

$$T_{interval} = \left(\frac{\text{ToA}}{\text{DutyCycle}}\right) - \text{ToA} \tag{47}$$

Table 14 lists the resulting theoretical goodput and throughput data rates.

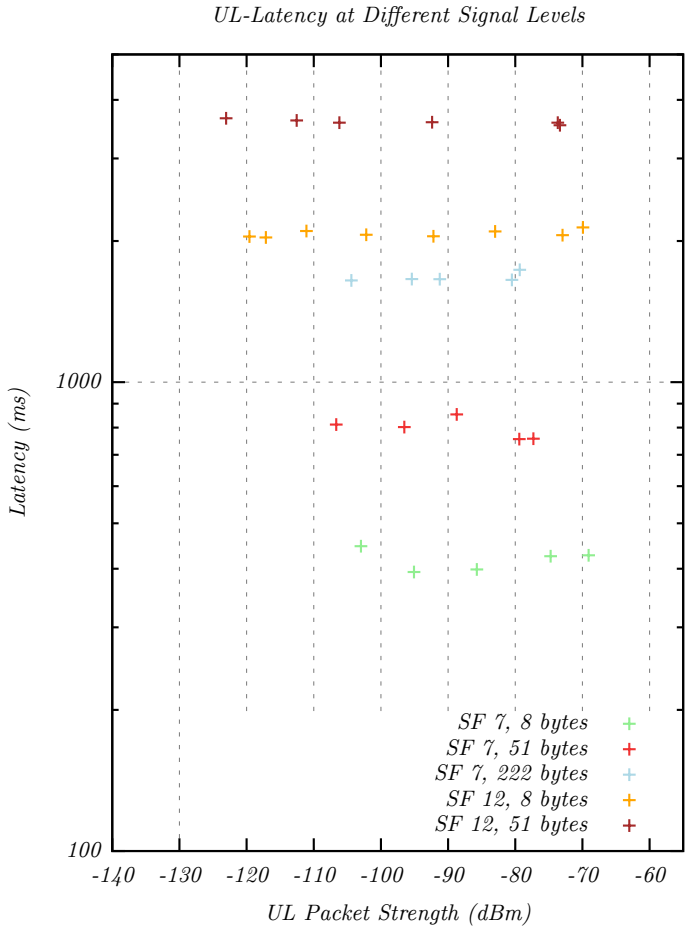


Figure 31: The total system latency as a function of the Packet Strength for SF 7 and SF 12 at various packet sizes.

Afterwards, the goodput and throughput were measured for a wide range of signal levels that were simulated using a step attenuator in 10 dB steps. For each attenuation step 30 packets of 8 bytes, 51 bytes and 222 bytes were sent at SF 7 and 8 bytes and 51 bytes at SF 12. The goodput was calculated from the packet size, the number of successful transmissions and the total transmission time.

Figure 32a illustrates the goodput for the aforementioned combinations of SF and packet size as a function of the Packet Strength. The obtained results matched the theoretical goodput limits closely. As expected, a higher goodput could be achieved with larger packet sizes, since the overhead made up a smaller percentage of the total packet size. The goodput exhibited little variation over the range of signal levels, since it was determined only by static factors (duty cycle, SF, packet size and header size). Shortly before the connection was lost, the goodput decreased, which was an artifact of the increasing packet loss. While the lost packets could be retransmitted, these retransmissions would be subject to the same duty cycle and packet loss conditions and block the channel for future transmissions, which would further reduce the goodput.

Figure 32b illustrates the throughput of the same measurement. The throughput was limited by the channel capacity, which was in turn restricted by the duty

Table 13: Components of the end-to-end latency for SF 7 and SF 12 at various packet sizes.

SF	Packet Size (bytes)	End-to-end Latency (ms)	ToA (ms)	Network Latency (ms)
7	8	362	36	326
7	51	665	103	562
7	222	1806	348	1458
12	8	2029	991	1038
12	51	3435	2466	969

Table 14: LoRaWAN goodput and throughput for SF 7 and SF 12 at various packet sizes.

SF	Packet Size (bytes)	1 % Duty Cycle Packet Interval (s)	Expected Goodput (bps)	Expected Throughput (bps)
7	8	4	16.00	72.0
7	51	10	40.80	63.2
7	222	35	50.70	57.1
12	8	99	0.65	2.9
12	51	247	1.65	2.6

cycle. Therefore, the throughput was independent from the packet size at each spreading factor; once the channel was full, no more data could be encoded for a given modulation scheme.

5.3.6 *Application Layer Performance: Packet Loss*

While many WSN use cases can tolerate some packet loss, it is still necessary to analyze the connection reliability in different scenarios to determine whether LoRaWAN is fit for a particular application. Especially if an application requires a gapless recording of sensor values, the lost messages need to be retransmitted. Depending on the message rate, this might conflict with the limits imposed by the duty cycle. In this section, the packet loss is analyzed as a function of the Packet Strength. The measurement was first conducted in a deployed network and then repeated in a shielded, interference-free setup. Using a step attenuator, different signal levels were simulated in 10 dB steps until the connection was lost. The measurement was repeated for different packet sizes and SFs (8 bytes, 51 bytes, 222 bytes at SF 7 and 8 bytes, 51 bytes at SF 12). The modem was configured to send unconfirmed messages, which were not acknowledged by the network server. Therefore, each message is sent only once without any retransmissions.

Figure 33 illustrates the packet loss for different packet sizes (a) at SF 7 and (b) at SF 12. The variation in packet loss at Packet Strength levels above -110 dBm could be attributed to external interference; there was a sharp increase in loss shortly

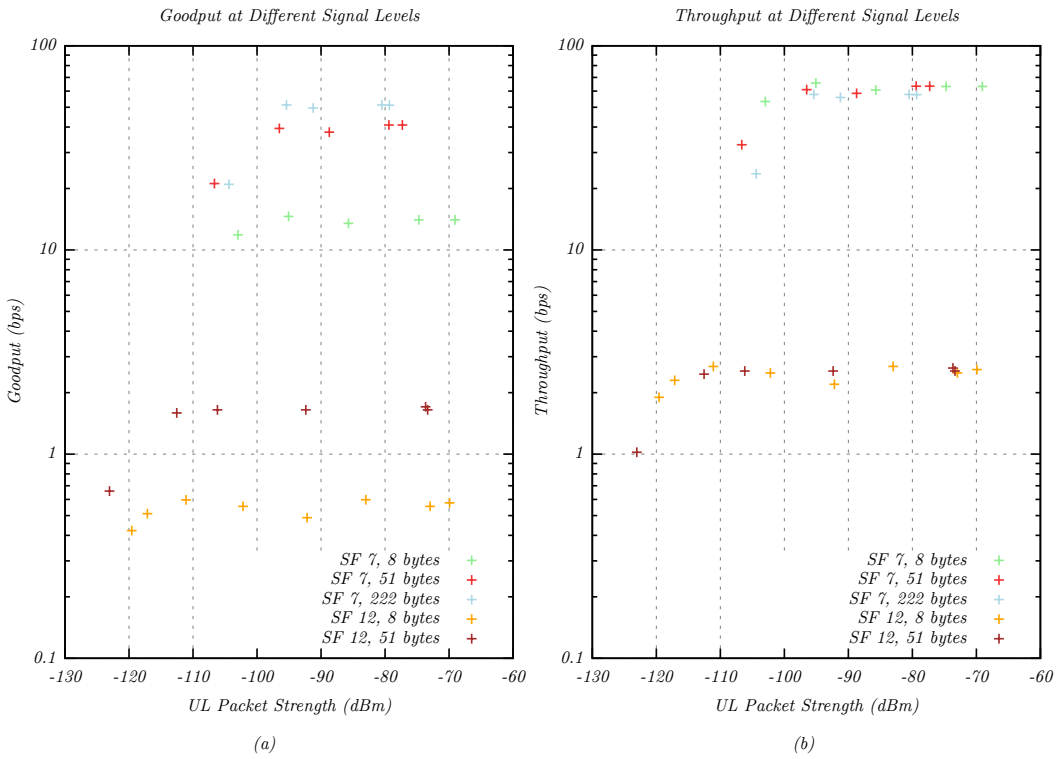


Figure 32: Maximum LoRaWAN data rates as a function of the Packet Strength at SF 7 and SF 12 for various packet sizes. **(a)** Goodput excluding LoRaWAN protocol headers. **(b)** Throughput including LoRaWAN protocol overhead.

before the connection was lost. The transmission at SF 12 extended to much lower signal levels due to its improved sensitivity. For both SF 7 and SF 12 connections, small 8-byte packets were subject to higher packet loss over 10 % even at excellent signal levels. At the same time, in most cases less than 5 % of all 51-byte and 222-byte packets were lost. Finding the source of the increased loss of 8-byte packets requires further investigation.

Figure 34 compares the packet loss in a deployed network and in a shielded environment. The comparison is made for 8-byte packets, which previously suffered the highest packet loss. The graph clearly shows that in the interference-free setup packet loss was virtually nonexistent; it could thus be deducted that external interference was the major cause of packet loss in the network, which was expected due to the unlicensed nature of the SRD band.

5.4 DISCUSSION

In this section, the LoRaWAN measurements are discussed and compared to the theoretical specifications [104]. Table 15 compares the LoRa and LoRaWAN specifications with the measurements in the deployed network and the shielded setup. Using the previously acquired data, the applicability of LoRaWAN for realistic smart metering use cases is discussed.

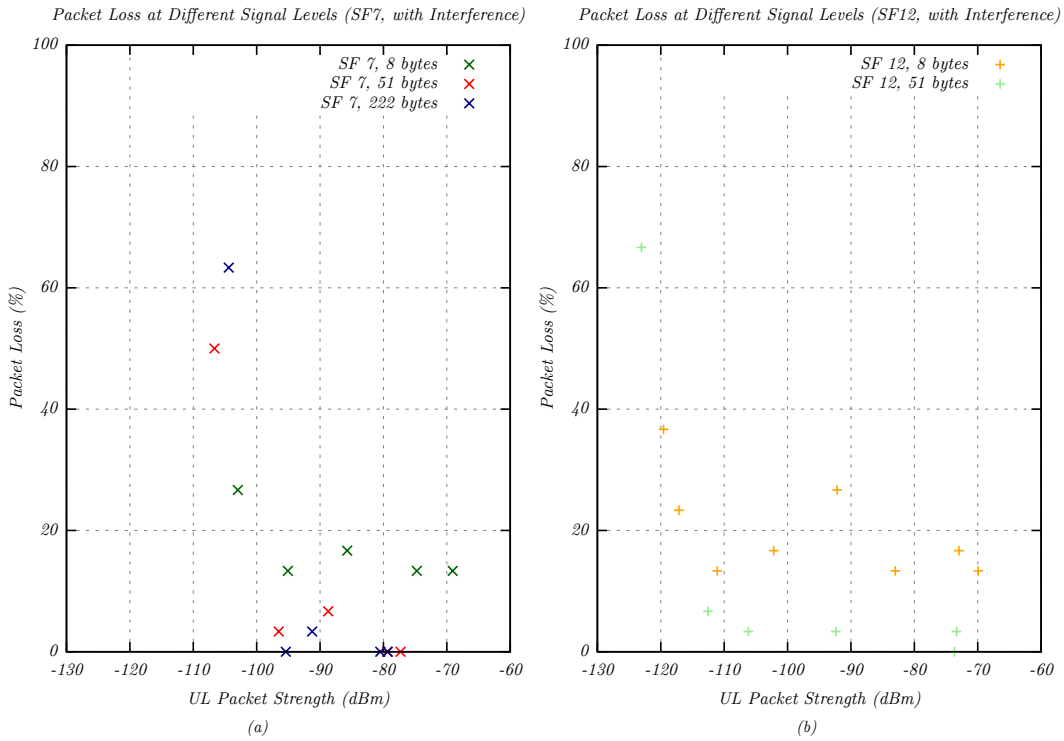


Figure 33: Packet loss as a function of the Packet Strength at different packet sizes **(a)** at SF 7 and **(b)** at SF 12 .

The following traffic patterns are considered:

- Uplink transfer of individual sensor data from a single sensor
- Downlink transfer of individual commands
- Uplink transfer of small packets from many sensors
- Transfer of software updates in the downlink and bulk data to the cloud in the uplink direction

The first use case represented a classic WSN application: a sensor was connected via LoRaWAN. It transmitted individual messages to the cloud for analysis and permanent recording. WSNs are commonly installed in a wide range of environments. In the evaluations, the deployed LoRaWAN network was able to provide a high MCL up to 142.2 dB. However, its coverage was considerably reduced compared to the specified value due to interference. LoRaWAN provided acceptable latency for individual packets, delivering an 8-byte message in 362 ms at SF 7 and in 2029 ms at SF 12. The measurements in the deployed network revealed about 15 % loss of small packets, which may require retransmissions. This reliability ceiling was confirmed by other researchers, however they found no improvements at larger packet sizes [9]. The biggest limitation was caused by the duty cycle regulations. There were long periods where no transmission was possible, which severely limited the sensor sampling rate. For example, an 8-byte message could be sent every 4 s at SF 7 and every 99 s at SF 12. If timely arrival of measurement values was not necessary, multiple samples could be aggregated into one message.

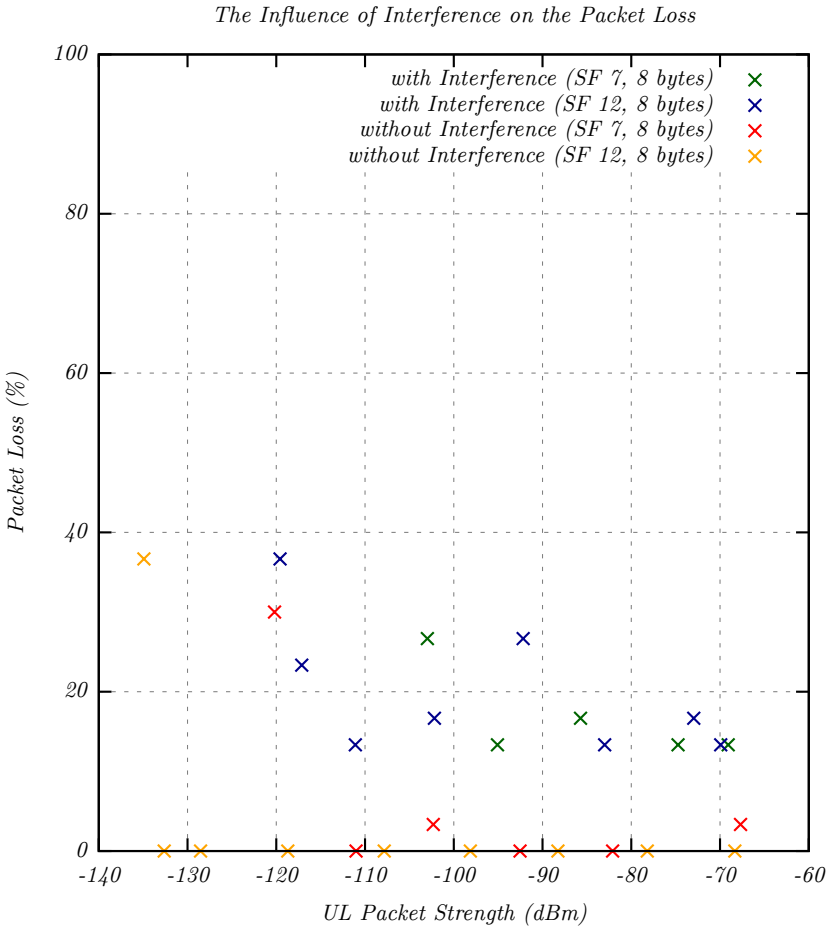


Figure 34: Comparison of the packet loss as a function of the Packet Strength in a deployed network to a interference-free laboratory environment.

The second traffic category was concerned with downlink commands, e.g., to control an actuator. While the measurements in this chapter did not consider downlink scenarios, it was still possible to provide an estimation of the QoS that could be expected. There were two mechanisms that dominated the downlink QoS: the selected LoRaWAN class and the gateway downlink capacity. The first factor determined the distribution of the receive windows at the receiver, which governed when downlink data could be transmitted. A device that listened for downlink commands needed to be reachable at least periodically, which excluded class A operation and increased power consumption. Depending on the required latency, the device either needed to operate in class B mode and open periodic receive windows or even work in class C mode, which kept the receiver active permanently. The second mechanism resulted from the fact that gateways also had to adhere to the duty cycle; if all channels were blocked by acknowledgments of uplink messages or other transmissions, the QoS of downlink commands could be severely impacted [9].

The third use case consisted of a fleet of sensors that were connected via a single LoRaWAN node, which aggregated and forwarded the messages. There were two basic strategies for such an installation: forwarding the sensor messages individually

Table 15: Comparison of the LoRaWAN specification to the performance in a deployed and a shielded scenario.

Metric	Specifications[104]	Deployed performance	Shielded performance
SF 7			
MCL (dB)	140.5	126.8	137.8
SNR_{min} (dB)	-7.5	-7.8	-8.8
Latency [2] (ms)	36 ms ToA + system latency	362	
$Throughput_{max}$ [1] (bps)	57.1	57.7	
SF 12			
MCL (dB)	153.5	142.2	151.2
SNR_{min} (dB)	-20.0	-19.2	-20.2
Latency [2] (ms)	991 ms ToA + system latency	2029	
$Throughput_{max}$ [1] (bps)	2.6	2.6	

[1] Throughput is given for 1 % duty cycle, B = 125 kHz, payload SF 7: 222 bytes, SF 12: 51 bytes.

[2] Latency is given for an unconfirmed transmission of 8 bytes of payload.

or aggregating multiple readings into a single packet. The goodput and packet interval limitations caused by the EU-868 duty cycle limited the first approach to very low sampling rates. The second approach provided much better scalability and could improve reliability, since larger packets were found to have lower packet loss. Due to the severe throughput limitations at high SFs, this use case was only feasible when the WSN was installed close to the gateway.

Finally, the transfer of large amounts of data was problematic at all configurations and channel conditions. The low throughput of LoRaWAN combined with the duty cycle limitations at the gateway would result in an extremely long transmission time for both software updates in the downlink and bulk data transmission in the uplink direction, in addition to blocking the channel for other communication. Overall, this use case would only be feasible if a prioritization mechanism was implemented in the network server that ensured other communication was still possible (i.e., software updates were served on a best effort basis) and long transmission times were acceptable.

5.5 COMPARISON OF LORAWAN AND NB-IOT QOS

Both LoRaWAN and NB-IoT (which was analyzed in Chapter 4) address low-end, energy efficient IoT use cases. It is therefore appropriate to compare them directly and analyze the applications that can benefit from one or the other technology. In

this section, the two LPWANs are compared based on their physical and application layer QoS properties.

One central advantage of LPWANs over their traditional counterparts is the considerably improved MCL. Where other networks such as LTE may cover the inside of a building but struggle in the basement, LPWANs can provide reliable coverage of hard to reach places. In the evaluations, both LoRaWAN and NB-IoT closely matched their specified MCL, with $MCL_{LoRaWAN} = 151.2$ dB and $MCL_{NB-IoT} = 163.8$ dB. NB-IoT exceeded LoRaWAN in terms of coverage range and was at an advantage due to the exclusivity of the spectrum. While the LoRa modulation was designed to be interference resistant, the deployed LoRaWAN network in the evaluation still experienced an MCL reduction to $MCL_{LoRaWAN,deployed} = 142.2$ dB, which was the result of the elevated interference in the unlicensed bands ($IM_{NB-IoT} = 2-6$ dB, $IM_{LoRaWAN} = 7-20$ dB). Depending on the application, this could be compensated for by installing additional LoRaWAN gateways to improve the signal level.

LoRaWAN and NB-IoT use different mechanisms for CE. While LoRaWAN uses SFs that double the symbol duration with each level, NB-IoT uses ECLs that define increasing numbers of signal repetitions. In the evaluations both techniques were able to extend the coverage range considerably and enable signal detection below the noise floor. At the same time, the mechanisms came at a cost in terms of QoS and energy consumption. Each LoRaWAN SF doubled the ToA, leading to an exponential increase in both the time the channel was blocked and the waiting time between packets. For example, a 51-byte message could be sent every 10 s using SF 7, while a transmission at SF 12 must wait 247 s between messages. As a result, LoRaWAN restricted the MTU at high spreading factors to limit the ToA and power consumption. Since the different SFs were orthogonal to each other, the LoRaWAN ADR algorithm avoided collisions and improved capacity by assigning different SFs to individual devices based on their link quality. For NB-IoT, the exclusivity of the spectrum and the eNodeB resource allocation mitigated collision problems, but the high repetition count at ECL 2 (up to 128 in uplink direction) could cause excessive energy consumption. For this reason, both the LoRaWAN ADR algorithm and NB-IoT exhausted their P_{TX} budget before increasing the CE level (see sec. 4.2.3 and [23]).

During the evaluations, both technologies were able to transmit small packets in an acceptable time frame. At strong signal levels, a typical sensor packet was delivered below one second of latency. LoRaWAN delivered a 51-byte packet in $T_{LoRa,SF7} = 665$ ms, while NB-IoT took $T_{NB-IoT,ECL0} = 535$ ms for a 27-byte payload[2]. At the same time, both technologies struggled to provide low latency at low signal levels. Due to the exponential ToA increase caused by the SFs, transmitting the same 51-byte packet took 3.4 s at SF 12. The end-to-end latency was consistent within the individual SFs and did not increase exponentially when switching to a higher SF, since a large part consisted of the network latency outside of the wireless LoRa link. For NB-IoT, the high number of repetitions caused a median latency of 11.9 s at ECL 2; there was a large standard deviation, with some packets taking over 100 s to arrive. Overall, NB-IoT was an attractive choice if the signal level is

2 At the time of the measurement, the NB-IoT network did not yet support non-IP transmissions. Therefore, an additional 28 bytes must be considered for IP and UDP headers, bringing the total MAC layer payload to 55 bytes.

sufficient to avoid ECL 2. For small messages and low signal levels, LoRaWAN could provide a lower latency.

One central aspect that hinders the QoS of LoRaWAN is the low data rate, which is the result of the duty cycle regulations in the ISM and SRD bands. The maximum LoRaWAN data rates can be achieved at SF 7 using the MTU of 222 bytes per packet. Assuming a 1 % duty cycle, a theoretical uplink throughput of 57.1 bps is reached, which translates to 50.7 bps of goodput. While these values can be reached in practice, it limits the technology to use cases that have no special data rate requirements. Furthermore, the device must wait 35 s before transmitting the next packet. In contrast, NB-IoT allows transmitting larger packets (a MTU of 1500 bytes is usually possible) and provides a much higher uplink throughput. The NB-IoT measurements revealed a throughput of 22.6 kbps using 1024-byte packets, which results in a 22.4 kbps goodput. This enables most sensor applications, especially since there is no forced wait time between packets. There might be limitations for large transmissions, e.g., of firmware updates, especially if the cell is heavily loaded with many devices; this scalability analysis is beyond the scope of the evaluation in this work.

Especially in hard to reach locations, packet loss is a common problem in LPWAN environments. Previous evaluations of LoRaWAN packet loss found a constant reliability ceiling at 15 % Packet Loss Ratio (PLR) for all spreading factors and packet sizes [9]. In this work, a correlation was found between packet size and PLR; while small 8-byte packets exhibited the same PLR as in previous works, there was a considerable improvement at larger packet sizes of 51 bytes and 222 bytes, which exhibited a PLR < 10 %. Additional measures were necessary to enable applications that relied on a reliable transmission, such as confirmed messages that were retransmitted if no acknowledgment was received. These retransmissions would come at the cost of increased energy usage and block the channel for future transmissions. Additionally, retransmissions would not improve the PLR if the gateway was out of reach [23]. NB-IoT provided a very high reliability, which was achieved due to the provider-configured HARQ mechanism. Furthermore, the provider could adjust the repetition count at each of the ECL levels, providing tailored levels of CE. The measurements in this work confirmed that NB-IoT exceeded LoRaWAN in terms of reliability.

Overall, NB-IoT QoS exceeded LoRaWAN QoS in most scenarios. NB-IoT provided wider coverage, higher data rate, better reliability and was not bound by duty cycle regulations. One critical limitation of NB-IoT was the exponential latency increase at low signal levels, which could exceed 100 s. NB-IoT provided an exception report mechanism which should provide a latency of 10 s using prioritization, but the corresponding evaluation found no such improvement; a possible reason was an incomplete implementation in the provider backend. At the same time, LoRaWAN provided a consistent QoS across a wide range of signal levels. Its biggest limitations were the low data rate and long wait time between packets, in addition to the considerable influence of interference. The latter could partially be mitigated by installing additional gateways and thus increasing the signal level. LoRaWAN was suitable for low-end applications that only transmitted packets occasionally, while NB-IoT addressed used cases with higher QoS demands.

5.6 CONCLUSIONS

In this chapter, a detailed evaluation of the uplink LoRaWAN physical and application layer QoS was performed. The measurements were conducted both in a deployed private network and in a shielded laboratory environment. In both cases, the complete measurement setup was constructed from commercially available hardware.

An in-depth analysis of the LoRaWAN coverage revealed that the specified MCL of 138.5 dB for SF 7 and 151 dB for SF 12 was closely matched in an interference-free environment, but was drastically reduced in the presence of interference. The high sensitivity was achieved using a chirp spread spectrum modulation with variable symbol duration T_S. Using the highest SF 12, a $SNR_{min} = -20dB$ could be achieved, which enabled sensor deployments in hard to reach locations.

While choosing a high SF appeared attractive to ensure a wide coverage, it was necessary to take latency and energy consumption into consideration. Every SF step doubled the ToA and therefore increased the energy consumption of the modem. Furthermore, the latency increased considerably; sending an 8-byte packet took 362 ms at SF 7, but 2029 ms at SF 12. However, LoRaWAN did not suffer from the exponential latency increase to tens of seconds that was observed with NB-IoT, which made it an attractive choice even at high SFs.

One of the central shortfalls of LoRaWAN was the limited data rate. In the ISM and SRD bands, the duty cycle is commonly limited to 1 %. In the evaluations, LoRaWAN provided consistent, but low goodput of 51.3 bps at SF 7, and only 1.7 bps at SF 12. This excluded all applications that require frequent transmission or need to transmit large amounts of data such as firmware updates.

Overall, LoRaWAN mostly satisfied the specifications [104] in a private LoRaWAN network built from commercial components. It was a suitable technology for WSN applications, yet was held back by the severe duty cycle regulations which limit its packet and data rates. Since LoRaWAN operates in the ISM band, interference played a significant role and reduced the MCL considerably.

A future evaluation of the LoRaWAN downlink QoS should analyze the QoS impact of the three device classes A, B and C. While nodes can send uplink messages at all times in every class, the distribution of downlink windows differs significantly. Since the gateway has to wait until the device opens a receive window, there are implications for the latency and data rate. Furthermore, the ADR mechanism should be analyzed to provide insight into the performance of nodes under difficult conditions.

6 THE NARROWBAND BUNDLING PROTOCOL

LPWANs provide an energy efficient, wide-range network service to IoT devices.[1] The QoS varies significantly between LPWANs in the licensed and unlicensed band, requiring to choose between cost and performance. Combining LPWANs with a multipath approach could improve the QoS, but existing protocols are unsuitable for LPWAN applications. In this chapter a novel Narrowband Bundling Protocol is proposed, which is tailored to the characteristics of LPWANs. It improves latency, reliability and coverage of IoT networks by aggregating multiple links. First experiments with a real prototype demonstrate promising QoS improvements.

6.1 DESIGN APPROACH

Unlike traditional networks such as LTE, LPWANs prioritize battery life and a long range over data rate and latency. For many IoT use cases, this is not a problem, as these endpoints spend most of the time in deep sleep to conserve energy and only wake up occasionally to transmit small amounts of data. Bearing in mind these characteristics, NBP was designed to retain the properties that make LPWANs an attractive choice for developers.

NBP is a light-weight and IoT-focused protocol that establishes a multipath connection between an end device and a server over multiple LPWAN links. Similar to UDP, it provides a low-overhead, non-reliable network service, which reduces energy consumption and prevents HoL blocking problems. Unlike UDP however, NBP introduces multipath functionality and is a connection-oriented protocol, which is necessary to exchange state information. NBP only includes features absolutely necessary for multipath operation in bandwidth and energy constrained use cases, such as endpoint addressing and subflow management; traffic is then distributed onto the links using a scheduling mechanism. This enables a wide range of use cases where constrained applications save overhead, while advanced functionality such as packet reordering or de-duplication can be implemented at the application layer. The time needed for connection establishment is partially compensated by providing 0-RTT data transmission in the first packet.

As the IoT ecosystem consists of a diverse set of preexisting applications, a proxy architecture is proposed along with NBP itself that provides backwards compatibility by intercepting UDP/IP packets. This proxy architecture has been implemented as a prototype to illustrate the compatibility features of the NBP architecture and promote adoption of the protocol.

1 Parts of Chapter 6 have previously been published in the journal paper "The Narrowband Bundling Protocol" by A. Matz et al. [56]. © 2022 IEEE.

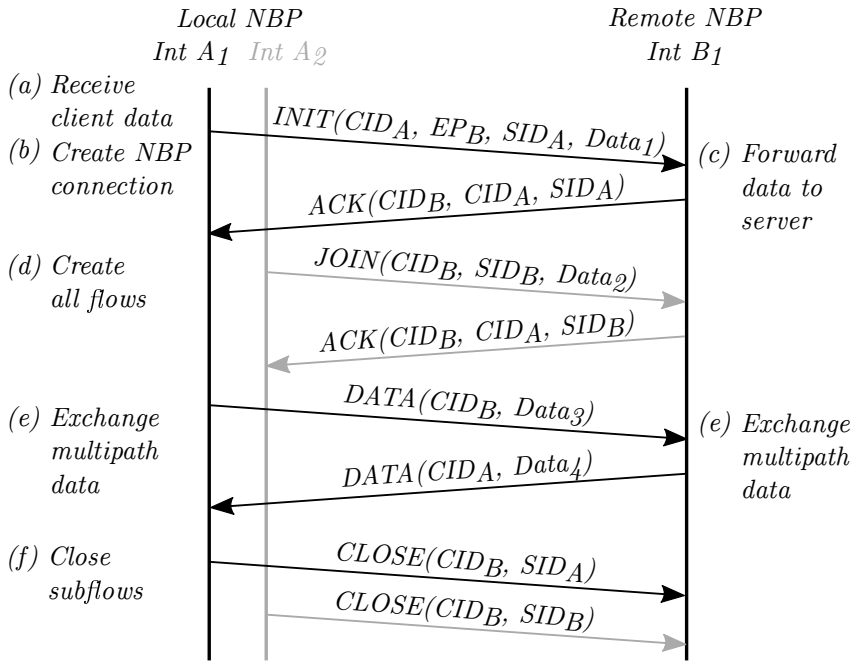


Figure 35: An exemplary NBP connection flow between two NBP instances. For simplicity reasons the client generating data and the server receiving the data have been omitted. © 2022 IEEE.

The packet flow of an exemplary NBP connection is illustrated in Figure 35. For simplicity reasons the diagram omits the client and server endpoints communicating through NBP. NBP itself consists of two instances with their respective network interfaces $\mathtt{IntA}_1$, $\mathtt{IntA}_2$ and $\mathtt{IntB}_1$.

(a,b) When the local NBP receives payload for a new connection, an *Initialize (INIT)*, *Acknowledgment (ACK)* handshake is initiated. The instances exchange the endpoint address and port ($\mathtt{EP}_\mathtt{B}$) and associate it with Connection Identifiers (CIDs), which are asymmetric between local ($\mathtt{CID}_\mathtt{A}$) and remote NBP ($\mathtt{CID}_\mathtt{B}$). This approach ensures uniqueness of the CIDs at the remote NBP; otherwise, if CIDs were symmetric and chosen by the local NBP, collisions could occur if two hosts selected the same value.

(c) The o-RTT payload is forwarded to the server endpoint.

(d) Once an initial flow is established, additional subflows can be created using JOIN packets. They are identified by an Subflow Identifier (SID), which enables NBP to create and destroy individual subflows as needed.

(e) Multipath data are exchanged using DATA packets and distributed via a scheduling algorithm. Data transmission is possible as long as at least one subflow is available. Using a CID removes the need to include the destination address in each packet, which improves efficiency.

(f) Individual subflows are terminated using CLOSE packets; closing the final subflow also terminates the connection.

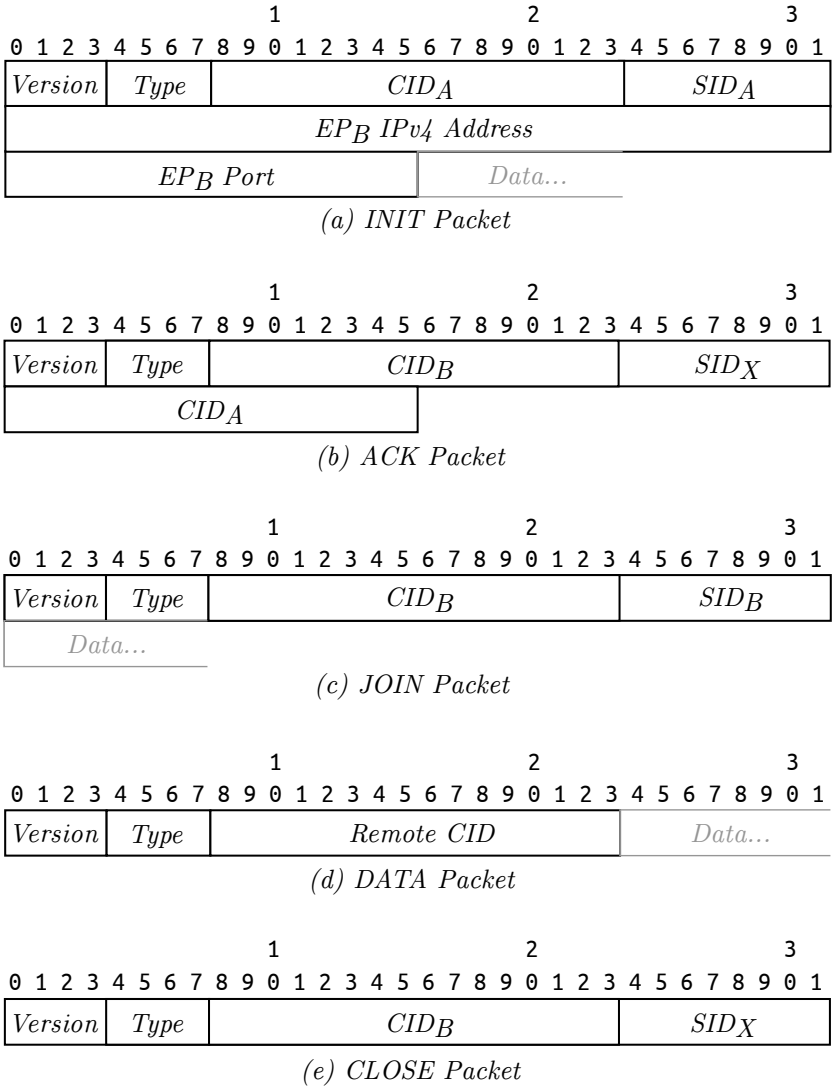


Figure 36: An overview of NBP packet types for IPv4 traffic. The numbers above each packet represent the bit offset of the individual header fields. © 2022 IEEE.

Figure 36 provides further detail on the packet types employed by NBP. Each packet has a *Version* field to allow for future protocol revisions, as well as a *Type* field that identifies the packet type. Due to significant MTU limits in some LPWANs (e.g., 51 bytes for LoRaWAN at high SFs [90]), NBP minimizes overhead:

- NBP assigns a 2-byte CID as part of the *INIT*, *ACK* (a,b) handshake to represent the endpoint addresses and ports in regular transmissions.
- Individual subflows are assigned a per-connection SID to efficiently add and delete flows using *JOIN* and *CLOSE* packets (c,e).
- *ACK* packets (b) are only sent during handshakes.
- As *DATA* packets (d) are the most common type, their header has been minimized to a size of 3 bytes.

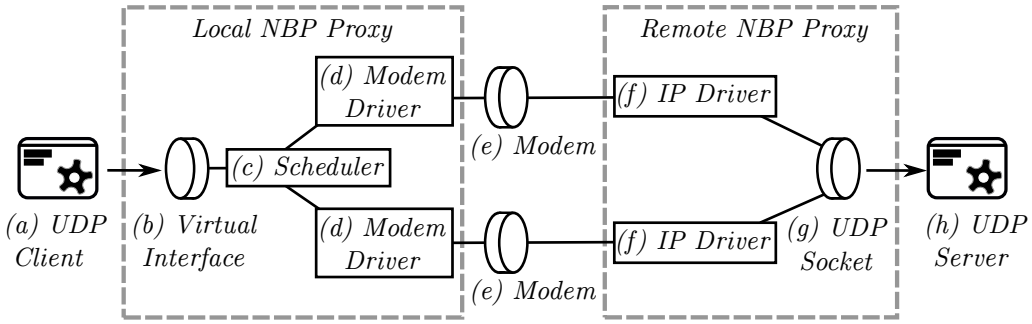


Figure 37: Example of an NBP architecture using dedicated aggregation hosts as implemented in the NBP prototype. Intermediate networks such as the provider backbone have been omitted for simplicity. © 2022 IEEE.

- NBP does not employ padding to fit the traditional 32-bit format used e.g., with TCP options.

The data overhead of an NBP connection depends on the average packet size, the total packet count, the scheduler in use and the number of subflows. For example, a 242-byte *DATA* packet (corresponding to the LoRaWAN MTU at SF 7) contains 3 bytes or about 1.2 % of overhead caused by NBP. For a LoRaWAN connection with SF 7 and 125 kHz bandwidth, this reduces the theoretical data rate from 5.47 kbps to 5.4 kbps[2]. Since NBP does not perform any cryptographic operations, the energy overhead of an NBP device is dominated by the power consumption of the modems necessary for multipath operation, especially if redundant transmission is employed. Finally, the individual data overhead of an NBP connection keeps the modems active for longer periods to transmit the user payload plus the NBP overhead.

6.2 INTRODUCING NBP INTO EXISTING NETWORKS

Whenever new protocols are introduced into existing networks, backwards compatibility is a major concern. A protocol that requires rewriting existing applications may have difficulties getting adopted, even if it was objectively a better choice for a specific use case. Therefore, a transparent proxy architecture is proposed along with NBP itself, which allows developers choosing a protocol that matches their needs. This provides multipath functionality to the diverse landscape of existing IoT applications and protocols without the need for fallback mechanisms. For efficiency reasons NBP proxies should extract the UDP payload and forward it as NBP traffic. Overall, this architecture enables deployment flexibility – NBP can be installed on the communication endpoints themselves or as a dedicated aggregation host, which collects the data from many IoT endpoints.

Figure 37 illustrates the architecture of the Linux-based NBP prototype. For this implementation, a transparent proxy concept was chosen; this allows to illustrate the applicability of NBP to a wide range of existing and novel use cases. For simplic-

2 In addition to the NBP overhead, the LoRaWAN lower layer overhead of 13–28 bytes (depending on the size of the Frame Options field) must be considered. Furthermore, the final goodput depends on the duty cycle regulations in the relevant frequency band.

ity a communication from UDP client to server is assumed. The prototype consists of the following components:

(a) *UDP Client*: Generates a UDP PDU, which is routed to the local NBP proxy.

(b) *Virtual Network Interface (VINF)*: An IP-based interface that accepts UDP PDUs. The traffic that is transmitted by NBP can be controlled via simple routing rules.

(c) *Scheduler*: A multipath scheduling algorithm that selects the subflow the next NBP PDU will be sent on. This decision can be made based on static criteria, based on the QoS parameters reported by the modems or historical data.

(d) *Modem driver*: There is a great variety of LPWAN modems, each with their own proprietary interface. The NBP prototype offers a standardized interface to implement modem drivers to easily integrate new hardware into the architecture.

(e) *Modem*: One or more LPWAN modems provide connectivity to different networks.

(f) *IP driver*: A special case of the *Modem driver* that encapsulates NBP into UDP/IP.

(g) *UDP Socket*: Created by the remote NBP proxy to communicate with servers on behalf of the UDP client.

(h) *UDP Server*: Receives and processes the UDP PDU.

Another challenge that novel protocols face is filtering by middleboxes, which are employed by network operators for performance optimization. Common mitigation strategies of other protocols are adapting an existing protocol or using encapsulation. Since NBP must also work in MTU-constrained non-IP networks, encapsulation is not an option. Instead, an experimental UDP extension called *UDP Options* [98] was considered as a base protocol for NBP; however, many LPWAN modems integrate a fixed UDP stack that does not allow modifications. The final design of NBP is a completely new application layer protocol that makes use of existing transport mechanisms of non-IP LPWANs and employs UDP encapsulation for transport across IP networks, which reduces middlebox filtering considerably [81].

6.3 EXPERIMENTAL SETUP AND METHODS

NBP was implemented as a prototype to evaluate its QoS benefits in various IoT scenarios. Using a simple scheduler that duplicates packets on multiple LPWAN links, NBP could improve the latency and delivery reliability of individual packets, as well as expand the total coverage area beyond what a single LPWAN can provide. These properties could be useful e.g., for alarm messages, which usually need to arrive quickly and reliably; duplicates are less of a problem.

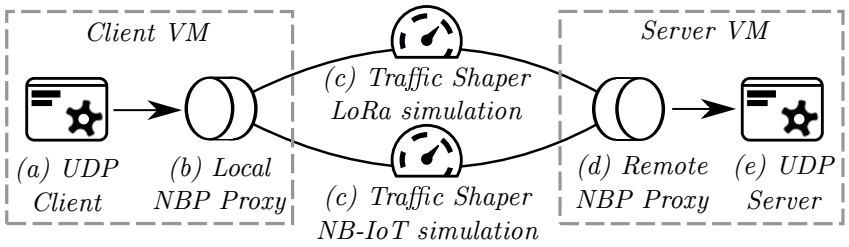


Figure 38: Experiment 1 - The simulation environment is based on virtual machines. © 2022 IEEE.

The overall *latency* $T_{L,NBP}$ of the end-to-end connection consists of the latency of the quickest path plus the latency generated by NBP packet buffers $T_{L,buf}$:

$$T_{L,NBP} = \min(T_{L,P1}, T_{L,P2}, \ldots) + T_{L,buf} \tag{48}$$

The *packet loss ratio* PLR_{NBP} of an NBP system depends on the correlation of the individual paths. In a worst case scenario, where every packet loss on the most reliable path is accompanied by a loss on all other paths (e.g., due to broadband Electromagnetic Interference (EMI)) and thus cannot be compensated for, the resulting PLR_{NBP} is that of the most reliable path:

$$PLR_{NBP} \leqslant \min(PLR_{P1}, PLR_{P2}, \ldots) \tag{49}$$

Finally, NBP extends the *coverage area* A_{NBP} by compensating a temporary coverage loss on one path by sending duplicates on other paths. Once again the amount of coverage extension depends on the coverage overlap of the individual paths; in the worst case (100 % overlap of all links) it is equal to the largest individual coverage area:

$$A_{NBP} \geqslant \max(A_{P1}, A_{P2}, \ldots) \tag{50}$$

The QoS benefits of NBP were analyzed in two experiments: a simulation (experiment 1) and a field trial (experiment 2).

Experiment 1 employed the Virtual Machine (VM) based simulation setup shown in Figure 38. This approach ruled out unpredictable external influences such as other network users. The UDP client (a) generated traffic, which was sent via the local NBP proxy (b) on both links (c) redundantly. The link properties were controlled using the Linux Traffic Control (`tc`) utility. Once the traffic arrived at the remote NBP proxy (d), it was converted back to UDP and forwarded to the UDP server (e). The experiment considered a LoRaWAN and an NB-IoT link, whose QoS is generally dominated by their CE mechanisms: NB-IoT employs signal repetition levels known as ECLs [44], while LoRaWAN networks use SFs that double the time on air for each SF step. In both cases a processing gain is achieved at the cost of increased latency and reduced data rate, resulting in distinct QoS levels. The simulation considered two scenarios: the highest and the lowest QoS level that LoRaWAN and NB-IoT can provide. For these scenarios, a simplified traffic model was derived from field measurements of NB-IoT [57] and LoRaWAN [58]. Table 16 details the

Table 16: Experiment 1 - The Link Parameters Simulated by the Traffic Shaper. © 2022 IEEE.

Parameter	**Highest QoS Level**		**Lowest QoS Level**	
LPWAN	LoRaWAN	NB-IoT	LoRaWAN	NB-IoT
CE Level	SF 7	ECL 0	SF 12	ECL 2
RSSI (dBm)	> -98	> -106	< -96	< -108
Packet Loss (%)	2	0	16	7
Latency (ms)	660	521	3513	6500
Standard Deviation (ms)	30	100	345	10000
Distribution	Normal	Normal	Normal	Paretonormal

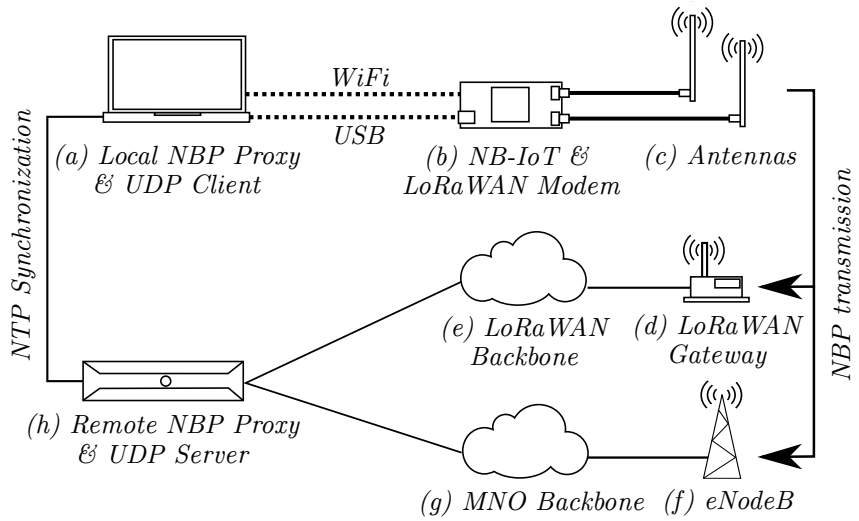


Figure 39: Experiment 2 - The setup of the NBP field trial in a university building. © 2022 IEEE.

statistical parameters of the model including latency, latency standard deviation, latency distribution and packet loss for both scenarios. Additionally, the network parameters RSSI and CE level are given. The two scenarios were then simulated using tc. For each measurement 1000 packets were transferred and latency, as well as packet loss were recorded. For NBP the latency was defined as the time after which the first copy of an alarm message arrived.

Experiment 2 consisted of a field trial, which was conducted in a deep indoor environment in the basement of our university building to explore the coverage enhancements provided by NBP. The network architecture is illustrated in Figure 39. The UDP client and local NBP proxy were installed on a laptop (a) that was connected to a PyCom FiPy NB-IoT/LoRaWAN modem (b,c). The modem was configured as a LoRaWAN class A device; the SF was controlled by the gateway using ADR. For LoRaWAN, a MultiTech Conduit LoRaWAN gateway (d) was installed on the roof of the building; it was connected to a private The Things Network

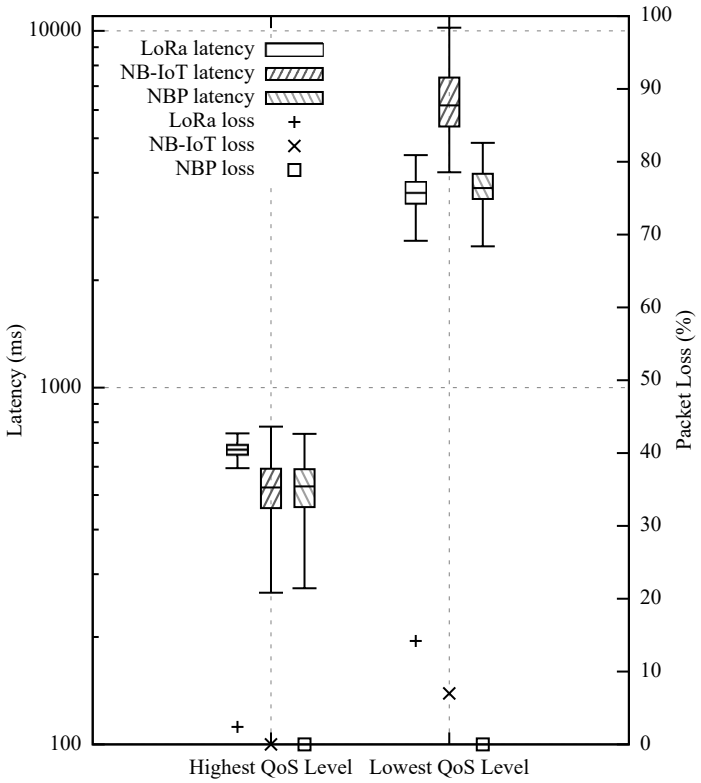


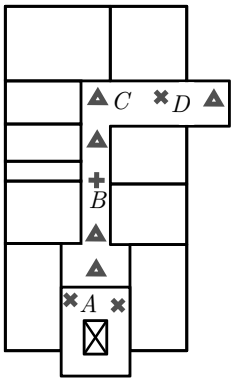
Figure 40: Experiment 1 - Latency and packet loss of individual LoRaWAN and NB-IoT links compared to an NBP connection that combines both. © 2022 IEEE.

LoRaWAN stack (e). This setup avoided duty cycle issues and ensured local coverage. The NB-IoT network link was provided by a public NB-IoT network (f,g). The UDP server and remote NBP proxy were installed on a private cloud server (h). For each measurement location 30 packets of 51 bytes each were transmitted. On the cloud server the packet loss was recorded along with the network link that was used.

6.4 RESULTS

Figure 40 illustrates the simulation results of experiment 1. In general, there was a latency increase for both LoRaWAN and NB-IoT links at the lowest QoS level, which could be attributed to the CE mechanisms. NB-IoT provided comparatively lower median latency at the highest QoS level, while LoRaWAN was at an advantage in extreme coverage scenarios. Overall, LoRaWAN latency was less variable compared to NB-IoT. When both paths were combined using NBP, the resulting median latency and latency spread were dominated by the path that provided the lowest latency. The observed results aligned with Equation (48).

In this experiment the packet loss introduced by the traffic shaper was random and uncorrelated across the individual links. NBP therefore was expected to improve the reliability significantly. As shown in Figure 40, NBP was able to improve reliability by compensating for all lost packets on both links. This was expected whenever there was no simultaneous loss on all paths and matched Equation (49).



Coverage with packet loss < 30 %:
- ✚ *LoRaWAN only*
- ✖ *NB-IoT only*
- ▲ *NB-IoT & LoRaWAN*

Packet Loss (%):

Point	*LoRaWAN*	*NB-IoT*	*NBP*
A	*69*	*17*	*14*
B	*0*	*31*	*0*
C	*14*	*7*	*0*
D	*45*	*3*	*3*
all	*32*	*15*	*4*

Figure 41: Experiment 2 - Coverage of LoRaWAN and NB-IoT in the basement of a university building. The table compares the packet loss of the individual LPWANs to NBP for the example points A through D. © 2022 IEEE.

Figure 41 shows the results of the field test in experiment 2. Both LoRaWAN and NB-IoT experienced high packet loss in parts of the basement. A closer analysis of the packet loss at the sample points A through D revealed that NBP reduced the loss according to Equation (49). The coverage was extended to the superset of all LPWANs according to Equation (50): if at least one LPWAN provided coverage, NBP provided a functional link. Overall, NBP was able to significantly improve the QoS parameters latency, reliability and coverage, which confirmed its benefit in many IoT scenarios.

6.5 DISCUSSION

After NBP was evaluated in simulations and in a physical test bed, its performance improvement over the individual LPWAN links can be discussed in practical scenarios. For this purpose, the four smart metering traffic categories introduced in the Sections 4.3 and 5.4 are recalled:

- Uplink transfer of individual sensor data from a single sensor
- Downlink transfer of individual commands
- Uplink transfer of small packets from many sensors
- Transfer of software updates in the downlink and bulk data to the cloud in the uplink direction

The first use case of individual messages from a single sensor was a typical WSN scenario. Especially sensors installed in hard to reach locations benefited from the coverage extension provided by NBP, since it increased deployment flexibility. Furthermore, some LPWANs like NB-IoT introduced high latency of tens of seconds under such conditions, which could be compensated for by a second path. Finally, the transmission redundancy improved the packet loss, which required fewer retransmissions that would otherwise block the channel and further increase the latency.

The second traffic category was concerned with individual messages in the downlink direction, usually commands to an actuator device. This category especially profited from the coverage extension provided by NBP, since actuators usually need to be installed in a fixed location. If a redundant scheduler was used, the command latency could be reduced considerably; in this case, the application should include a packet deduplication mechanism (e.g., a timestamp or packet number) to avoid performing an action multiple times.

The third category was concerned with one NBP device collecting the data from many sensors. This approach could significantly reduce the number of modems required to connect a local WSN. Since such a gateway device would be a single point of failure that could take a large number of devices offline, NBP could improve the reliability of the network access by providing failover in case one LPWAN link becomes unavailable. Furthermore, NBP could enhance the capacity by distributing packets on multiple LPWAN links.

Finally, the fourth category of traffic was bulk data transmission in the uplink and downlink directions such as software updates and transmitting large log files to the cloud. This traffic category was similar to the use cases addressed by traditional multipath protocols such as MPTCP; as a result, similar QoS requirements existed. The transmission of bulk data is generally not time critical and served on a best effort basis; the primary goal is that the transfer is completed as soon as possible, without interfering with regular operation. Therefore, it did not make sense to duplicate packets on all paths. Using the application QoS requests mechanism described in Section 6.6, the application could instruct NBP to activate the fastest available LPWAN link for the transfer of bulk data; if the capacity of multiple links should be aggregated, a path estimation mechanism may be added as described in Section 6.8.2.

6.6 A FIRST NBP EXTENSION - APPLICATION QOS REQUESTS

Traditional bundling protocols such as MPTCP usually employ a static scheduler that distributes packets onto the network paths according to predefined criteria, optimizing cost, latency or other parameters. This strategy works well for applications such as access bundling, where links are prioritized according to their monetary cost and expensive LTE resources are conserved. While it is possible to adjust schedulers at run time, e.g., to steer traffic away from overloaded links, this adjustment is applied to all traffic that passes a device and therefore does not cover use cases with dynamic QoS requirements.

The heterogeneous QoS of different LPWAN technologies and the dynamic QoS requirements of IoT applications that use edge computing techniques require a bundling mechanism that can react on application demands. For example, an end device might send regular measurement values that are not time critical, and a single alarm message which must arrive quickly and reliably. This scenario requires special treatment of the alarm message, yet the messages look the same from the outside; only the application knows which one is important. In order to inform NBP about the QoS requirements of individual packets, a suitable interface is necessary.

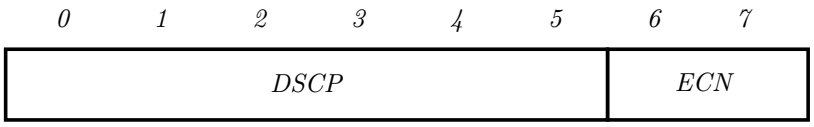


Figure 42: The DS header field consisting of the DSCP field [12] and the ECN field [28].

1. If NBP is implemented as a library, it must provide an Application Programming Interface (API) to set a per-packet QoS optimization strategy.

2. If NBP is implemented as a proxy, it must provide a mechanism to classify individual UDP/IP packets.

Since the proxy architecture needs to be compatible with both NBP aware and unaware applications, introducing a packet mark must not interfere with the regular operation of IP and UDP. Therefore, the Payload cannot be changed without compromising backwards compatibility. Unlike TCP, UDP also does not include an option field that could be used to communicate a QoS request. While there was a proposal for a UDP Options extension [98], it was never implemented by modern operating systems and would thus require a custom UDP stack. The final option to consider is the IP header, which already provides a field to communicate QoS requests[3]: The eight-bit DS field [12] provides a mechanism to classify and manage traffic across large networks. It replaces the historic Type of Service (ToS) field, which saw various uses and redefinitions over the years [3, 19, 33]. In large deployments, the DS architecture [15] allows service providers examining and marking individual packets at the edge of their network. As the traffic crosses the network, each router applies the appropriate Per-Hop Behavior (PHB) according to the traffic class using scheduling and queue management techniques. As a result, traffic with higher QoS requirements (such as video conferencing) can be given preferential treatment without the need for computationally expensive traffic classification at every node. This approach is subject to the same middlebox challenges as MPTCP and other protocols; however, it is unlikely to cause problems in local deployments, where the NBP proxy is part of the same device or network as the IoT node. In the cloud, it may be necessary to implement additional measures, e.g., installing both NBP and the application in the provider network.

The layout of the DS field is illustrated in Figure 42. It consists of a six-bit Differentiated Services Code Point (DSCP) [12], which can represent 64 different traffic classes and a two-bit Explicit Congestion Notification (ECN) [28] field.

The Internet Assigned Numbers Authority (IANA) maintains a registry of mandatory and recommended DSCP values [26]. The available DSCP space is divided into three pools as listed in Table 17. All DS implementations must implement a *Default PHB*, which covers traffic not assigned to any other class and usually implements a best effort PHB. The DSCP value *0* must be assigned to the Default PHB. Furthermore, seven Class Selector (CS) PHBs are reserved to provide backwards compati-

3 Another approach considered for communicating QoS requirements was introducing a new IP Option. Similar to TCP Options, the IP header can be extended using type-length-value style options that communicate additional information. The Differentiated Services (DS) field was preferred as it already defines many suitable code points and does not increase the packet size.

Table 17: The three DSCP pools assigned by IANA [26].

Pool	Pool Size	Codepoint Space	Assignment Policy
1	32	xxxxx0	Standards Action
2	16	xxxx11	Experimental / Local Use
3	16	xxxx01	Standards Action

bility with the Packet Precedence field, which consisted of the first three bits of the historic ToS field.

There are two other predefined PHB classes, which address common use cases. First, the Expedited Forwarding (EF) [94] PHB class addresses low-loss, low-jitter and low-delay applications by receiving priority queuing over other types of traffic. Second, the Assured Forwarding (AF) [103] PHB class enables operators to offer different levels of forwarding assurance to a previously subscribed maximum data rate. The AF PHB defines four *AF classes* for which resources are reserved at each intermediate node. Within each class, traffic can be further marked with a drop precedence according to its importance; packets with higher drop precedence values are dropped first in case of congestion.

Table 18 lists the DSCP values used by NBP along with their meanings. Most codepoints follow the recommended values of the Default, EF and AF PHBs [11]. Since NBP faces unique challenges in a multipath LPWAN environment, new class names and a description have been added to communicate the QoS optimization strategy. As NBP-unaware applications cannot be expected to always set the DSCP field to all zero, experimental code points have been used wherever the original meaning of the DSCP values conflicted with the new NBP meaning. How a specific QoS goal is achieved is left to the implementation. Some examples include:

1. The "Low-cost" class could direct traffic to unlicensed LPWANs to avoid consuming expensive licensed data budget.
2. The "High-throughput" class could enable a high-capacity link that was asleep to conserve energy.
3. The "Emergency Data" class could cause a packet to be queued immediately on all interfaces, which is expensive in terms of energy and overhead but ensures that the packet arrives as soon and as reliable as possible.

6.7 USE CASE EXAMPLES

As energy providers transition to monitoring grid components using modern wireless technologies, LPWANs can be an interesting option due to their enhanced coverage and ability to penetrate buildings. However, selecting an appropriate LPWAN for a specific use case is not trivial due to the large QoS variations between individual technologies and depending on the signal level at the installation location. Energy providers that want to build their own LPWAN applications thus require

Table 18: NBP employs most of the predefined DSCP values; the meaning of some traffic classes has been redefined to better communicate the QoS optimization goal.

NBP Name	DSCP Name	DSCP value	Optimization
Standard	DF	0	none
Low-cost	CS1	8, 11[1], 15[1]	Lowest monetary cost
High-throughput	AF11, AF12, AF13	10, 12, 14	Highest link capacity for bulk data transfer
Low-latency	AF21, AF22, AF23	18, 20, 22	Lowest end-to-end latency
High-reliability	-	35[1], 39[1]	Highest link reliability
Priority Signaling	CS5	40	Queue with higher priority
Emergency data	EF	46	Maximum priority and reliability

[1] The marked codepoints have been assigned from the experimental DSCP pool 2. If more than one codepoint is specified, they provide additional prioritization similar to the AF class.

considerable expertise to match their application with an appropriate LPWAN or need to rely on proprietary solutions from external vendors.

A ready-made, NBP-enabled device that automatically adapts to the local radio conditions and individual QoS needs supports the creation of novel use cases such as gathering live data from transformer stations in the low voltage grid or large-scale deployment of smart meters in heterogeneous conditions. Simultaneously, the use of multiple different LPWAN links can improve coverage and latency, while reducing packet loss, which enables the use of LPWANs in use cases that require reliability beyond what a single technology can provide. If an installation location is covered by two or more networks, the NBP device can select the best network and improve the QoS.

For legacy systems that are still controlled using low-bandwidth solutions such as tone control, combining one cost-effective and one powerful LPWAN via NBP can add the ability to combine cost-efficient transmission of regular sensor data and timely arrival of software updates and full control schedules. For example, electrical vehicles could be instructed to follow a specific charging schedule in duration and power draw, depending on when the vehicle needs to be fully charged and the current load profile of the grid.

6.8 FUTURE WORK

Beyond the first NBP prototype there is a range of features that would improve the usefulness of NBP in real-world scenarios. Some of these challenges are universal to bundling protocols, such as the challenge of securely establishing multiple subflows between endpoints or providing robustness against link failures. Other aspects, such as the question of how to perform path estimation in an energy and MTU-constrained environment, are specific to narrowband wireless environments.

6.8.1 *IPv6 Support*

The current version of NBP is limited to transport Internet Protocol Version 4 (IPv4) addresses in the INIT packet. In the future, a new INIT6 packet should be introduced to enable Internet Protocol Version 6 (IPv6) connectivity. At the same time, the IPv4-only INIT packet should be renamed as INIT4 to avoid ambiguity. The overhead introduced by IPv6 addresses will reduce the available room for o-RTT data, so its use should be carefully evaluated if low-MTU LPWANs are used. Since the IPv6 header contains a Traffic Class field that is functionally identical to the DS field, the application can communicate QoS requests in the same way as when using IPv4.

6.8.2 *Considerations for Path Estimation and Congestion Control*

mMTC traffic patters often consist of a small number of one-way messages that are sent in between long idle periods to conserve battery power. These small packet groups are unlikely to exhaust the link capacity, so path estimation is commonly omitted. However, in a multipath context there may be multiple different LPWAN technologies to choose from, at which point it is beneficial to have an estimation of the QoS of each link before making a scheduling decision. Over the years, many different path estimation mechanisms have been designed; for example, TCP employs acknowledgments to assess the channel properties. There is a number of aspects that make designing an effective path estimation mechanism for LPWANs challenging:

1. WSN nodes are commonly battery powered and thus limited in terms of energy and processing budget.
2. LPWANs are strongly uplink oriented, and many messages do not receive a server reply.
3. Many LPWAN technologies provide asymmetric QoS between uplink and downlink direction (e.g., LoRaWAN class A), which complicates one-way latency measurements.
4. Especially unlicensed LPWANs are limited in their MTU and thus sensitive to large protocol header sizes.

Since a full path estimation can overwhelm a link that is limited in terms of energy and MTU, the channel probing should be reduced to the absolute minimum

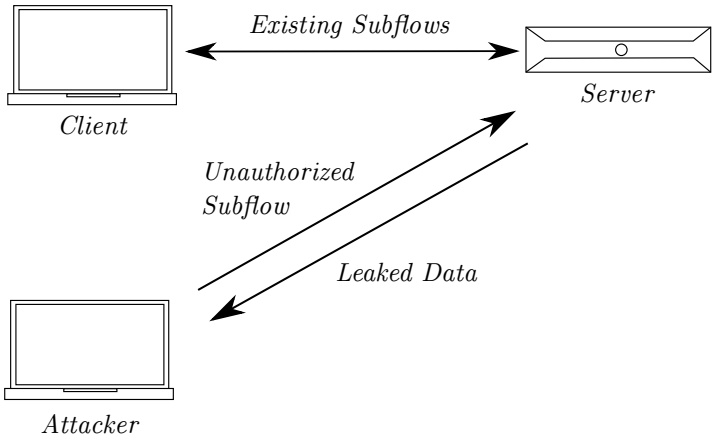


Figure 43: If an attacker is able to add unauthorized subflows to an existing connection, data may be leaked by the server.

necessary. In most LPWAN networks, the air interface is the bottleneck of the connection[4]; therefore, the current QoS properties of the LPWAN link provide a reasonable estimation of the end-to-end link QoS.

In the case of NBP, these properties could be leveraged: First, the initial connectivity check is performed automatically during connection establishment and could be repeated periodically, being omitted whenever two-way communication is detected. Second, a predictive approach could be used that employs the current modem QoS parameters, as well as QoS measurements of individual LPWAN technologies to derive a link quality metric for each LPWAN interface. This metric could then be used to perform scheduling decisions.

Once a suitable path estimation mechanism has been developed, the historical data could be collected and used for QoS optimizations; for example, if the algorithm reports unsatisfactory QoS in a certain region, the affected nodes could be expanded with another LPWAN modem or the signal level could be improved by installing additional base stations.

6.8.3 *Authentication of New Subflows*

Traditional protocols such as TCP bind themselves to a 5-tuple, which describes the combination of source and destination IP addresses and ports, as well as the protocol in use. If a client is equipped with multiple network interfaces and a multipath protocol is used to aggregate the links, the server needs to verify that these interfaces actually belong to the same host. If such verification is not in place, an attacker could add subflows and obtain part of the user data meant for the original client. Figure 43 illustrates this principle; in fact, the attacker could create a large number of new subflows or even close existing flows to starve the legitimate client.

This is a common challenge to all multipath protocols and has been discussed and solved in a number of different ways. MPTCP exchanges keys in plain text [29] at the beginning of a connection. Other protocols such as MP-QUIC employ

4 A notable exception to this rule is when multiple wireless networks are cascaded, e.g., when backhauling a LoRaWAN gateway using a 3G wireless connection [71].

cryptographic handshakes to exchange key material [24]. The latter is a very secure approach but introduces additional computation and energy overhead.

The energy and MTU constraints in LPWAN applications complicate the implementation of a secure handshake, as excessive cryptographic operations could increase the power consumption and reduce the opportunity to send o-RTT data. As such, NBP does not yet implement a subflow authentication mechanism. Further research is required into the best option of creating a secure and energy efficient authentication process.

6.8.4 *Robustness Against Link Failures*

The long range and high building penetration of LPWANs has generated a wealth of new use cases especially in the area of WSNs. At the same time sensors that are installed in deep indoor environments such as a basement may experience poor QoS with considerable packet loss. For multipath protocols this is a challenging situation; in most cases, a minimum amount of state information must be exchanged between the hosts before additional subflows can be established. For example, MPTCP creates an initial subflow to exchange keys and sequence numbers, after which additional subflows may be created. This ensures backwards compatibility with existing TCP hosts, which simply ignore the additional MPTCP header and answer with a plain TCP SYN/ACK, at which point MPTCP falls back to regular TCP [29]. There is a proposed MPTCP extension called Robust Establishment (RobE) that discusses ways to remove the dependency on the initial path [8]. Essentially, four proposals are discussed:

1. *Retransmission Timer*: Creates a single initial flow and starts a timer; when it expires, an initial flow is created on another link.
2. *Simultaneous Initial Paths (simple)*: Creates initial subflows on all paths and terminates all but the quickest one.
3. *Simultaneous Initial Paths (extended)*: Creates initial subflows on all paths and once the quickest flow is established, all other flows are fast joined using a new option type.
4. *Heuristic Initial Path Selection*: Heuristically selects the path with the highest success probability based on path estimation information.

The challenge that NBP faces regarding connection establishment robustness is twofold: First, WSN end devices are usually operated from battery, which places restrictions on the energy consumption that is acceptable to establish a connection. For example, the third proposal of RobE requires the creation of N initial flows on all paths, then $N-1$ flows are terminated and finally $N-1$ subsequent flows are added to the successful connection. This keeps the modems active for an unnecessary long time, especially since LPWANs can reach tens or even hundreds of seconds of latency in extreme coverage situations. This approach is currently employed in MP-QUIC [24], where the benefits of using the quickest path early outweighs the drawback of creating and terminating multiple subflows.

The second constraint is that it is difficult to predict the instantaneous latency of an LPWAN connection. The third RobE proposal requires a timer to be defined before attempting a connection; however, the wide latency range of LPWANs complicates finding a value that does not break the connection prematurely while not waiting excessively long either.

Finally, downgrading a connection could be an option; however, once authentication is built into NBP, an additional verification process would be required, which might negate the overhead reduction achieved by downgrading the subflows.

6.9 CONCLUSIONS

In this chapter a novel Narrowband Bundling Protocol was presented. It was specifically designed to address the bandwidth and energy constraints that are common among LPWAN technologies. NBP provides a network service similar to UDP and a proxy architecture is proposed along with the protocol itself that provides backwards compatibility to legacy applications. Given the diverse nature of LPWANs, NBP supports both IP and non-IP links and introduces little overhead to work in constrained networks with small MTUs. For regular *DATA* packets, only 3 bytes of additional application layer overhead is caused by NBP. Both simulations and a field trial were performed to analyze the QoS improvement that IoT devices can expect from using NBP. Using a simple redundant scheduler, NBP was capable of approaching the latency of the quickest path, compensate for packet loss and expand the coverage to the superset of the individual LPWAN technologies. In the future NBP will receive additional features to improve its usefulness in IoT scenarios. Some improvements are based on other protocols, such as a robust connection establishment [8] and authorization of subflows [24, 29]. Other challenges are IoT-specific: narrow channels permit only limited channel probing, which could be addressed using a predictive approach based on historic data. Finally, applications will be able to communicate their QoS needs, which NBP will act on dynamically.

Part III

CONCLUSIONS AND FUTURE WORK

7 CONCLUSION

In this thesis, the potential of using a bundling protocol approach in LPWAN use cases was explored. To establish a baseline of typical LPWAN characteristics, two popular technologies were selected for a systematic QoS analysis at a wide range of signal levels. For the group of licensed LPWANs, NB-IoT was selected due to its popularity and wide availability. Its high building penetration and reliable transmission characteristics make it an attractive choice for developers looking to install end nodes in hard to reach places. As an example of unlicensed LPWAN technologies, LoRaWAN was chosen for an in-depth analysis. The massive ecosystem of end nodes including all kinds of sensor and actor devices, as well as the availability of off-the-shelf gateways ensures that LoRaWAN is a consideration for many WSN use cases. A comparison of the two technologies revealed that both have strengths and weaknesses, and it is not possible to address all use cases with a single LPWAN.

From the results of the evaluations, the idea was developed to combine multiple LPWAN links using a bundling protocol. This approach could enable a wide range of use cases and compensate for individual drawbacks of the links. Furthermore, existing innovations from previous multipath protocols could be reused. Since the available solutions did not fit the characteristics of typical LPWAN connections, a novel bundling protocol was developed that was appropriately named the *Narrowband Bundling Protocol*.

7.1 FINDING THE BEST LPWAN TECHNOLOGY FOR A USE CASE

The evaluations of NB-IoT and LoRaWAN have revealed considerable QoS differences between the two technologies. In most cases, NB-IoT was at an advantage due to its superior specifications, however LoRaWAN revealed some properties that make it an attractive choice for certain applications.

The evaluation results confirmed the hypothesis that the CE mechanisms were central to the QoS provided at different signal levels. NB-IoT employed two CE mechanisms. First, it adjusted the MCS to match the channel conditions, which required more RUs and therefore a longer time to transmit the same amount of data; second, it employed ECL levels that determined the number of coherently added signal repetitions, which took time to transmit and process. The effect could be observed in the uplink latency measurements, which increased exponentially from below 1 s at ECL 0 to tens of seconds when switching to ECL 2 below the noise floor. Depending on the application, such a link could be functionally broken, even if the packets did arrive eventually. This did not mean that the CE mechanisms were ineffective; in the evaluations, NB-IoT achieved its specified MCL of 164 dB with very low packet loss. In terms of data rate, NB-IoT was able to provide up to 22.6 kbps of uplink throughput across a wide range of signal levels; this value was highly sensitive to MNO resource allocation and UE selection.

In contrast, LoRaWAN relied on SFs, where each step doubled the ToA and improved the sensitivity. Since the LoRaWAN stack introduced considerable latency itself, the ToA only accounted for part of the total latency and the increase was less drastic than in NB-IoT networks. Transmitting 51 bytes took less than 1 s at SF 7 and remained below 5 s at SF 12. In terms of coverage, LoRaWAN had to mitigate the higher interference levels in the unlicensed bands, which caused an MCL reduction from 151.2 dB in a shielded setup to 142.2 dB in a deployed network. Furthermore, 5 % to 15 % of packet loss could be expected. Since LoRaWAN operates in the unlicensed spectrum, the aforementioned QoS parameters could be improved by installing additional gateways. There was another side effect of the CE: the exponential ToA increase required a matching inter-packet interval to maintain the duty cycle regulations. Even for small payloads, this waiting time could reach hundreds of seconds, which was why LoRaWAN limited the MTU to just 51 bytes at the highest SFs. This reduced the data rate considerably; while LoRaWAN was commonly advertised to provide up to 5.4 kbps at SF 7, the measurements revealed a throughput of only 57.1 bps at 1 % duty cycle.

Overall, the measurements confirmed the higher performance of NB-IoT in most use cases. It generally provided higher MCL, better data rates and lower packet loss. Therefore, NB-IoT should be considered for applications that transmit large messages or require continuous transmissions (e.g., firmware updates). However, LoRaWAN has a number of desirable properties that make it worth considering for individual messages in long intervals. First, the latency increase below the noise floor is less drastic than for NB-IoT. Second, the unlicensed nature of LoRaWAN allows placing gateways as desired, reducing the need for CE. Finally, the availability of low-cost hardware and lack of monthly fees make it an attractive choice for cost constrained applications.

7.2 COMBINING LPWANS TO IMPROVE THE QOS

The evaluations have shown that even a low-end LPWAN such as LoRaWAN could be a worthwhile addition for use cases with volatile QoS requirements. For example, a multi-homed device using LoRaWAN and NB-IoT could offload regular transmissions to the cost effective unlicensed access, while software updates and high-priority messages could be sent via the licensed link. Implementing such functionality as a bundling protocol provides the necessary flexibility to adjust to different use cases and allows benefiting from the extensive research already performed for high-end protocols such as MPTCP and MP-QUIC.

In this thesis NBP was presented, which is a lightweight and LPWAN-specific bundling protocol. The narrowband and energy-constrained nature of LPWANs introduces a number of challenges when designing novel protocols. NBP addresses these characteristics by providing a low-overhead, unordered and non-reliable network service similar to UDP. This approach conserves energy and prevents HoL blocking that can occur in channels with high packet loss. Since LPWANs vary widely in terms of QoS, a universal bundling protocol needs to support both powerful IP and constrained non-IP paths. NBP employs an explicit handshake to enable efficient end-to-end addressing over any path by using CIDs to represent IP addresses. While this introduces additional latency before a connection is fully

formed, a o-RTT mechanism allows including data in the *INIT* and *JOIN* packets, making new subflows available immediately.

During the design phase of NBP, a primary goal was to minimize overhead in order to enable its operation in MTU constrained channels. For example, NBP introduces just 3 bytes of overhead for *DATA* packets, which results in about 1.2 % of overhead for a 242-byte packet corresponding to the LoRaWAN SF 7 MTU. The current version of NBP does not perform cryptographic operations and thus the energy overhead is dominated by the presence of multiple LPWAN modems. In general, NBP only provides the minimum functionality necessary to enable a functional multipath connection such as scheduling, path management and addressing; applications that require additional functionality such as packet deduplication and reordering can implement it at the application layer, while low-end applications save on overhead.

Along with NBP, a proxy architecture is proposed to support the diverse existing IoT ecosystem and ease the adoption of the new protocol. In this architecture, NBP is implemented in multi-homed aggregation points, which accept UDP packets from legacy hosts and forward their payload as native NBP traffic. The setup enables one-end or fully transparent operation, which allows hiding the multipath connection from one or both endpoints. Using this setup, it is possible to serve many IoT nodes from one aggregation point and to contact arbitrary servers on the Internet.

Since multipath protocols lend themselves to applications with volatile QoS requirements, a mechanism was designed to allow the application communicating per-packet QoS requests to NBP. This mechanism works both when NBP is installed as part of the application and when using the proxy architecture. In the latter case, backwards compatibility to applications that are not aware of NBP is maintained by reusing the DS field of the IP header.

NBP was analyzed in a simulation and a field trial to verify the hypothesis that a bundling protocol approach can improve the QoS of an LPWAN connection. Using redundant transmission over NB-IoT and LoRaWAN, NBP was able to improve the latency to that of the quickest flow, eliminate most packet loss and expand coverage to the superset of all LPWAN links. The experiment results highlight the potential of an LPWAN bundling approach in general and can be used as a baseline for future measurements.

7.3 FUTURE RESEARCH OPPORTUNITIES

The journey for multipath protocols in LPWAN contexts has only begun. During the development of the first NBP prototype, many promising research directions were identified that can improve the usefulness of an LPWAN bundling protocol such as NBP. Some of these challenges are common to all bundling protocols, but most are complicated by the limitations imposed by low-end LPWAN technologies such as LoRaWAN. On the other hand, finding efficient solutions to these problems can benefit existing protocols as well.

First, the question of securing the NBP handshake in extremely MTU constrained environments is of special interest. While traditional protocols exchange data only with a single peer address, multipath protocols have to verify that multiple discontinuous addresses belong to the same peer device to avoid leaking traffic to an

attacker. In MTU constrained scenarios, exchanging long encryption keys quickly fills the payload, reducing the opportunity to send o-RTT data.

Second, general-purpose bundling protocols such as MPTCP employ path estimation and congestion control to make use of the full link capacity of all subflows. While WSN applications usually do not exhaust the channel, some schedulers rely on path measurements to make scheduling decisions. The unique properties of LPWANs such as drastic variations in latency, asymmetric links and energy constraints require a suitable approach to estimate the channel.

Third, the current NBP setup is designed around one initial subflow that is joined by multiple additional subflows. While this ensures consistency of state at both connection endpoints, it creates a dependency on the functionality of the initial path that cannot be guaranteed in a volatile LPWAN environment. The discussion on the best solution for this problem is still ongoing for other protocols [8].

Finally, once NBP is mostly feature complete, a systematic evaluation of its usefulness and energy overhead should be conducted for a wide range of real use cases.

BIBLIOGRAPHY

[1] A. Yegin and O. Seller. *LoRaWAN Technical Specifications*. Specification. accessed: 2022-07-07. 2022. URL: https://resources.lora-alliance.org/technical-specifications.

[2] Ferran Adelantado, Xavier Vilajosana, Pere Tuset-Peiro, Borja Martinez, Joan Melia-Segui, and Thomas Watteyne. "Understanding the Limits of LoRaWAN." In: *IEEE Communications Magazine* 55.9 (2017), pp. 34–40. DOI: 10.1109/MCOM.2017.1600613.

[3] Philip Almquist. *Type of Service in the Internet Protocol Suite*. RFC 1349. July 1992. DOI: 10.17487/RFC1349. URL: https://www.rfc-editor.org/info/rfc1349.

[4] Markus Amend, Eckard Bogenfeld, Anna Brunstrom, Andreas Kassler, and Veselin Rakocevic. *A multipath framework for UDP traffic over heterogeneous access networks*. Internet-Draft. accessed: 2022-04-28. July 2019. URL: https://datatracker.ietf.org/doc/html/draft-amend-tsvwg-multipath-framework-mpdccp-01.

[5] Markus Amend, Eckard Bogenfeld, Milan Cvjetkovic, Veselin Rakocevic, Marcus Pieska, Andreas Kassler, and Anna Brunstrom. "A Framework for Multiaccess Support for Unreliable Internet Traffic using Multipath DCCP." In: *2019 IEEE 44th Conference on Local Computer Networks (LCN)*. 2019, pp. 316–323. DOI: 10.1109/LCN44214.2019.8990746.

[6] Markus Amend, Anna Brunstrom, Andreas Kassler, and Veselin Rakocevic. *Lossless and overhead free DCCP - UDP header conversion (U-DCCP)*. Internet-Draft. accessed: 2022-04-28. July 2019. URL: https://datatracker.ietf.org/doc/html/draft-amend-tsvwg-dccp-udp-header-conversion-01.

[7] Markus Amend, Anna Brunstrom, Andreas Kassler, Veselin Rakocevic, and Stephen Johnson. *DCCP Extensions for Multipath Operation with Multiple Addresses*. Internet-Draft draft-ietf-tsvwg-multipath-dccp-05. accessed: 2022-09-14. Internet Engineering Task Force, July 2022. 40 pp. URL: https://datatracker.ietf.org/doc/draft-ietf-tsvwg-multipath-dccp/05/.

[8] Markus Amend and Jiao Kang. *Multipath TCP Extension for Robust Session Establishment*. Internet-Draft. accessed: 2022-04-28. Mar. 2022. URL: https://datatracker.ietf.org/doc/html/draft-amend-tcpm-mptcp-robe-02.

[9] Takwa Attia, Martin Heusse, Bernard Tourancheau, and Andrzej Duda. "Experimental Characterization of LoRaWAN Link Quality." In: *2019 IEEE Global Communications Conference (GLOBECOM)*. 2019, pp. 1–6. DOI: 10.1109/GLOBECOM38437.2019.9013371.

[10] Aloÿs Augustin, Jiazi Yi, Thomas Clausen, and William Mark Townsley. "A Study of LoRa: Long Range & Low Power Networks for the Internet of Things." In: *Sensors* 16.9 (2016). ISSN: 1424-8220. DOI: 10.3390/s16091466. URL: https://www.mdpi.com/1424-8220/16/9/1466.

[11] Fred Baker, Jozef Babiarz, and Kwok Ho Chan. *Configuration Guidelines for DiffServ Service Classes*. RFC 4594. Aug. 2006. DOI: 10.17487/RFC4594. URL: https://www.rfc-editor.org/info/rfc4594.

[12] Fred Baker, David L. Black, Kathleen Nichols, and Steven L. Blake. *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*. RFC 2474. Dec. 1998. DOI: 10.17487/RFC2474. URL: https://www.rfc-editor.org/info/rfc2474.

[13] Kalpit D Ballal, Lars Dittmann, Sarah Ruepp, and Martin Nordal Petersen. "IoT Devices Reliability Study: Multi-RAT Communication." In: *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*. 2020, pp. 1–2. DOI: 10.1109/WF-IoT48130.2020.9221163.

[14] Subho Shankar Basu, Ashish Kumar Sultania, Jeroen Famaey, and Jeroen Hoebeke. "Experimental Performance Evaluation of NB-IoT." In: *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). Barcelona, Spain: IEEE, Oct. 2019, pp. 1–6. ISBN: 978-1-72813-316-4. DOI: 10.1109/WiMOB.2019.8923221. (Visited on 01/13/2020).

[15] David L. Black, Zheng Wang, Mark A. Carlson, Walter Weiss, Elwyn B. Davies, and Steven L. Blake. *An Architecture for Differentiated Services*. RFC 2475. Dec. 1998. DOI: 10.17487/RFC2475. URL: https://www.rfc-editor.org/info/rfc2475.

[16] Olivier Bonaventure. *Apple uses Multipath TCP*. Dec. 2018. URL: http://blog.multipath-tcp.org/blog/html/2018/12/15/apple_and_multipath_tcp.html.

[17] Olivier Bonaventure, Maxime Piraux, Quentin De Coninck, Matthieu Baerts, Christoph Paasch, and Markus Amend. *Multipath schedulers*. Internet-Draft draft-bonaventure-iccrg-schedulers-02. Work in Progress. Internet Engineering Task Force, Oct. 2021. 14 pp. URL: https://datatracker.ietf.org/doc/draft-bonaventure-iccrg-schedulers/02/.

[18] Mohamed Boucadair et al. *3GPP Access Traffic Steering Switching and Splitting (ATSSS) - Overview for IETF Participants*. Internet-Draft draft-bonaventure-quic-atsss-overview-00. Work in Progress. Internet Engineering Task Force, May 2020. 29 pp. URL: https://datatracker.ietf.org/doc/draft-bonaventure-quic-atsss-overview/00/.

[19] Robert T. Braden. *Requirements for Internet Hosts - Communication Layers*. RFC 1122. Oct. 1989. DOI: 10.17487/RFC1122. URL: https://www.rfc-editor.org/info/rfc1122.

[20] Bundesnetzagentur. *Allgemeinzuteilung von Frequenzen zur Nutzung durch Funkanwendungen geringer Reichweite (SRD)*. URL: https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Frequenzen/Allgemeinzuteilungen/FunkanlagenGeringerReichweite/2018_05_SRD_pdf (visited on 07/06/2022).

[21] Chirpstack. *Chirpstack Homepage*. accessed: 2022-07-06. 2022. URL: https://www.chirpstack.io/ (visited on 07/06/2022).

[22] Laurent Coustet. *MLVPN - MultiLink Virtual Public Network*. accessed: 2022-04-28. Oct. 2021. URL: https://zehome.github.io/MLVPN/ (visited on 04/28/2022).

[23] Ulysse Coutaud, Martin Heusse, and Bernard Tourancheau. "High Reliability in LoRaWAN." In: *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*. 2020, pp. 1–7. DOI: 10.1109/PIMRC48278.2020.9217220.

[24] Quentin De Coninck. "Flexible Multipath Transport Protocols." English. PhD thesis. Louvain-la-Neuve, Belgium: Université catholique de Louvain, Mar. 2020. URL: https://inl.info.ucl.ac.be/system/files/thesis%5C_deconinck.pdf (visited on 04/28/2022).

[25] Exelonix GmbH. *Exelonix NB | DEV Narrow-Band IoT Development Kit*. Available online: https://www.exelonix.com/products/nb-iot-dev-kit/. (Visited on 03/28/2022).

[26] Gorry Fairhurst. *Update to IANA Registration Procedures for Pool 3 Values in the Differentiated Services Field Codepoints (DSCP) Registry*. RFC 8436. Aug. 2018. DOI: 10.17487/RFC8436. URL: https://www.rfc-editor.org/info/rfc8436.

[27] Arshad Farhad, Dae-Ho Kim, and Jae-Young Pyun. "Scalability of LoRaWAN in an Urban Environment: A Simulation Study." In: *2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN)*. 2019, pp. 677–681. DOI: 10.1109/ICUFN.2019.8806140.

[28] Sally Floyd, Dr. K. K. Ramakrishnan, and David L. Black. *The Addition of Explicit Congestion Notification (ECN) to IP*. RFC 3168. Sept. 2001. DOI: 10.17487/RFC3168. URL: https://www.rfc-editor.org/info/rfc3168.

[29] Alan Ford, Costin Raiciu, Mark J. Handley, Olivier Bonaventure, and Christoph Paasch. *TCP Extensions for Multipath Operation with Multiple Addresses*. RFC 8684. accessed: 2022-04-28. Mar. 2020. DOI: 10.17487/RFC8684. URL: https://www.rfc-editor.org/info/rfc8684.

[30] Adrien Gallouët. *Glorytun*. accessed: 2022-04-28. Oct. 2020. URL: https://github.com/angt/glorytun (visited on 04/28/2022).

[31] Orestis Georgiou and Usman Raza. "Low Power Wide Area Network Analysis: Can LoRa Scale?" In: *IEEE Wireless Communications Letters* 6.2 (2017), pp. 162–165. DOI: 10.1109/LWC.2016.2647247.

[32] Stephen Haddock and Jessy Rouyer. *802.1AX-2020 – Standard for Local and Metropolitan Area Networks – Link Aggregation*. Jan. 2020. URL: https://1.ieee802.org/tsn/802-1ax-rev/.

[33] *Internet Protocol*. RFC 791. Sept. 1981. DOI: 10.17487/RFC0791. URL: https://www.rfc-editor.org/info/rfc791.

[34] ITU-R. *Report ITU-R M.2410-0, Minimum requirements related to technical performance for IMT-2020 radiointerface(s)*. Available online: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2410-2017-PDF-E.pdf. Nov. 2017. (Visited on 02/01/2020).

[35] Jana Iyengar and Martin Thomson. *QUIC: A UDP-Based Multiplexed and Secure Transport*. RFC 9000. accessed: 2022-04-28. May 2021. DOI: 10.17487/RFC9000. URL: https://www.rfc-editor.org/info/rfc9000.

[36] J. Postel. *User Datagram Protocol*. RFC 768. accessed: 2022-04-28. Aug. 1980. DOI: 10.17487/RFC0768. URL: https://www.rfc-editor.org/info/rfc768.

[37] J. Postel. *Transmission Control Protocol*. RFC 793. accessed: 2022-08-30. 1981. DOI: 10.17487/RFC0793. URL: https://www.rfc-editor.org/info/rfc793.

[38] Sikandar Zulqarnain Khan, Hassan Malik, Jeffrey Leonel Redondo Sarmiento, Muhammad Mahtab Alam, and Yannick Le Moullec. "DORM: Narrowband IoT Development Platform and Indoor Deployment Coverage Analysis." In: *Procedia Computer Science* 151 (2019), pp. 1084–1091. ISSN: 18770509. DOI: 10.1016/j.procs.2019.04.154. (Visited on 01/13/2020).

[39] Matthew Knight and Balint Seeber. "Decoding LoRa: Realizing a Modern LPWAN with SDR." In: *6th GNU Radio Conference*. Vol. 1. 1. 2016. URL: https://pubs.gnuradio.org/index.php/grcon/article/view/8.

[40] E. Kohler, M. Handley, and S. Floyd. *Datagram Congestion Control Protocol (DCCP)*. RFC 4340. accessed: 2022-08-30. 2006. DOI: 10.17487/RFC4340. URL: https://www.rfc-editor.org/info/rfc4340.

[41] Guus Leenders, Gilles Callebaut, Geoffrey Ottoy, Liesbet Van der Perre, and Lieven De Strycker. "Multi-RAT for IoT: The Potential in Combining LoRaWAN and NB-IoT." In: *IEEE Communications Magazine* 59.12 (2021), pp. 98–104. DOI: 10.1109/MCOM.008.2100382.

[42] Gábor Lencse, Szabolcs Szilagyi, Ferenc Fejes, and Marius Georgescu. *MPT Network Layer Multipath Library*. Internet-Draft. accessed: 2022-04-28. Dec. 2021. URL: https://datatracker.ietf.org/doc/html/draft-lencse-tsvwg-mpt-09.

[43] Nicolai Leymann, Cornelius Heidemann, Mingui Zhang, Behcet Sarikaya, and Margaret Cullen. *Huawei's GRE Tunnel Bonding Protocol*. RFC 8157. accessed: 2022-04-28. May 2017. DOI: 10.17487/RFC8157. URL: https://www.rfc-editor.org/info/rfc8157.

[44] Olof Liberg, Mårten Sundberg, Y.-P. Eric Wang, Johan Bergman, Joachim Sachs, and Gustav Wikström. *Cellular internet of things: from massive deployments to critical 5g applications*. 1st ed. Cambridge: Elsevier Inc, 2019. ISBN: 978-0-08-102902-2.

[45] Olof Liberg et al. "Narrowband Internet of Things 5G Performance." In: *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*. 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall). Honolulu, HI, USA: IEEE, Sept. 2019, pp. 1–5. ISBN: 978-1-72811-220-6. DOI: 10.1109/VTCFall.2019.8891588. (Visited on 01/13/2020).

[46] Shaowei Liu, Weimin Lei, Wei Zhang, and Hao Li. "MPUDP: Multipath Multimedia Transport Protocol over Overlay Network." In: *5th Int. Conference on Machinery, Materials and Computing Technology (ICMMCT 2017)*. Mar. 2017. URL: https://www.atlantis-press.com/article/25873703.pdf (visited on 10/28/2021).

[47] LoRa Alliance. *A technical overview of LoRa and LoRaWAN*. accessed: 2022-07-06. 2015. URL: https://lora-alliance.org/wp-content/uploads/2020/11/what-is-lorawan.pdf (visited on 07/06/2022).

[48] LoRa Alliance. *LoRaWAN 1.0.2 Specification*. English. accessed: 2022-07-19. 2016. URL: https://lora-alliance.org/wp-content/uploads/2020/11/lorawan1_0_2-20161012_1398_1.pdf (visited on 07/19/2022).

[49] LoRa Alliance. *LoRaWAN 1.1 Specification*. English. accessed: 2022-07-04. 2017. URL: https://lora-alliance.org/wp-content/uploads/2020/11/lorawantm_specification_-v1.1.pdf (visited on 07/04/2022).

[50] LoRa Alliance. *RP002-1.0.2 LoRaWAN Regional Parameters*. English. accessed: 2022-07-04. 2020. URL: https://lora-alliance.org/wp-content/uploads/2020/11/RP_2-1.0.2.pdf (visited on 07/04/2022).

[51] LoRa Alliance. *LoRa Alliance Home Page*. accessed: 2022-04-28. Apr. 2022. URL: https://lora-alliance.org/.

[52] LoRaTools. *LoRa Air Time Calculator*. accessed: 2022-07-06. 2022. URL: https://loratools.nl/#/airtime (visited on 07/06/2022).

[53] Hassan Malik, Sikandar Zulqarnain Khan, Jeffrey Leonel Redondo Sarmiento, Alar Kuusik, Muhammad Mahtab Alam, Yannick Le Moullec, and Sven Parand. "NB-IoT Network Field Trial: Indoor, Outdoor and Underground Coverage Campaign." In: *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. 2019 15th International Wireless Communications and Mobile Computing Conference (IWCMC). Tangier, Morocco: IEEE, June 2019, pp. 537–542. ISBN: 978-1-5386-7747-6. DOI: 10.1109/IWCMC.2019.8766568. (Visited on 01/13/2020).

[54] Borja Martinez, Ferran Adelantado, Andrea Bartoli, and Xavier Vilajosana. "Exploring the Performance Boundaries of NB-IoT." In: *IEEE Internet of Things Journal* 6.3 (June 2019), pp. 5702–5712. ISSN: 2327-4662, 2372-2541. DOI: 10.1109/JIOT.2019.2904799. (Visited on 01/13/2020).

[55] Andreas Matz. "MPDCCP - Moving Towards a Multipath Capable DCCP." MA thesis. Department of Information Technology, Electrical Engineering and Mechatronics, THM University of Applied Sciences, Friedberg, Germany, Apr. 2018.

[56] Andreas Philipp Matz, Jose-Angel Fernandez-Prieto, Joaquin Canada-Bago, and Ulrich Birkel. "The Narrowband Bundling Protocol." In: *IEEE Wireless Communications Letters* 11.9 (2022), pp. 1900–1904. DOI: 10.1109/LWC.2022.3186392.

[57] Andreas Philipp Matz, Jose-Angel Fernandez-Prieto, Joaquin Cañada-Bago, and Ulrich Birkel. "A Systematic Analysis of Narrowband IoT Quality of Service." en. In: *Sensors* 20.6 (Mar. 2020), p. 1636. ISSN: 1424-8220. DOI: 10.3390/s20061636.

[58] Lukas Metzger. "A Systematic Evaluation of Quality of Service Parameters for LoRaWAN and NB-IoT." MA thesis. Department of Electrical and Information Engineering, THM University of Applied Sciences, Gießen, Germany, Apr. 2020.

[59] Microchip. *Microchip RN2483 LoRaWAN Modem Datasheet*. accessed: 2022-07-19. 2020. URL: https://ww1.microchip.com/downloads/en/DeviceDoc/RN2483-Data-Sheet-DS50002346E.pdf (visited on 07/19/2022).

[60] Konstantin Mikhaylov, Vitaly Petrov, Rohit Gupta, Maria A. Lema, Olga Galinina, Sergey Andreev, Yevgeni Koucheryavy, Mikko Valkama, Ari Pouttu, and Mischa Dohler. "Energy Efficiency of Multi-Radio Massive Machine-Type Communication (MR-MMTC): Applications, Challenges, and Solutions." In: *IEEE Communications Magazine* 57.6 (2019), pp. 100–106. DOI: 10.1109/MCOM.2019.1800394.

[61] Konstantin Mikhaylov, Martin Stusek, Pavel Masek, Vitaly Petrov, Juha Petajajarvi, Sergey Andreev, Jiri Pokorny, Jiri Hosek, Ari Pout-tu, and Yevgeni Koucheryavy. "Multi-RAT LPWAN in Smart Cities: Trial of LoRaWAN and NB-IoT Integration." In: *2018 IEEE International Conference on Communications (ICC)*. 2018, pp. 1–6. DOI: 10.1109/ICC.2018.8422979.

[62] Radek Mozny, Pavel Masek, Martin Stusek, Krystof Zeman, Aleksandr Ometov, and Jiri Hosek. "On the Performance of Narrow-band Internet of Things (NB-IoT) for Delay-tolerant Services." In: *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*. 2019 42nd International Conference on Telecommunications and Signal Processing (TSP). Budapest, Hungary: IEEE, July 2019, pp. 637–642. ISBN: 978-1-72811-864-2. DOI: 10.1109/TSP.2019.8768871. (Visited on 01/13/2020).

[63] Radek Mozny, Martin Stusek, Pavel Masek, Konstantin Mikhaylov, and Jiri Hosek. "Unifying Multi-Radio Communication Technologies to Enable mMTC Applications in B5G Networks." In: *2020 2nd 6G Wireless Summit (6G SUMMIT)*. 2020, pp. 1–5. DOI: 10.1109/6GSUMMIT49458.2020.9083791.

[64] Collins Burton Mwakwata, Hassan Malik, Muhammad Mahtab Alam, Yannick Le Moullec, Sven Parand, and Shahid Mumtaz. "Narrowband Internet of Things (NB-IoT): From Physical (PHY) and Media Access Control (MAC) Layers Perspectives." In: *Sensors* 19.11 (June 8, 2019), p. 2613. ISSN: 1424-8220. DOI: 10.3390/s19112613. (Visited on 01/13/2020).

[65] Nokia Simens Network. *RF Measurements Quantities and Optimization*. Available online: https://www.slideshare.net/MuhammadNauman22/01-rf-measurementandoptimization-50334309. (Visited on 01/02/2020).

[66] C. Paasch. *MultiPath TCP Linux Kernel Implementation - Configure MPTCP*. June 2021. URL: https://www.multipath-tcp.org/pmwiki.php/Users/ConfigureMPTCP.

[67] C. Paasch, G. Detal, S. Barré, F. Duchêne, and O. Bonaventure. *The fastest TCP connection with Multipath TCP*. Mar. 2013. URL: https://multipath-tcp.org/pmwiki.php?n=Main.50Gbps.

[68] Gianni Pasolini. "On the LoRa Chirp Spread Spectrum Modulation: Signal Properties and Their Impact on Transmitter and Receiver Architectures." In: *IEEE Transactions on Wireless Communications* 21.1 (2022), pp. 357–369. DOI: 10.1109/TWC.2021.3095667.

[69] Tara Petrić, Mathieu Goessens, Loutfi Nuaymi, Laurent Toutain, and Alexander Pelov. "Measurements, performance and analysis of LoRa FABIAN, a real-world implementation of LPWAN." In: *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. 2016, pp. 1–7. DOI: 10.1109/PIMRC.2016.7794569.

[70] Alexandru-Ioan Pop, Usman Raza, Parag Kulkarni, and Mahesh Sooriyabandara. "Does Bidirectional Traffic Do More Harm Than Good in LoRaWAN Based LPWA Networks?" In: *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*. 2017, pp. 1–6. DOI: 10.1109/GLOCOM.2017.8254509.

[71] Albert Pötsch and Florian Hammer. "Towards End-to-End Latency of LoRaWAN: Experimental Analysis and IIoT Applicability." In: *2019 15th IEEE International Workshop on Factory Communication Systems (WFCS)*. 2019, pp. 1–4. DOI: 10.1109/WFCS.2019.8758033.

[72] 3rd Generation Partnership Project. *Technical Report 45.820 v13.1.0, Cellular System Support for Ultra-low Complexity and Low Throughput Internet of Things*. Available online: https://www.3gpp.org/DynaReport/45820.htm. Dec. 21, 2015. (Visited on 02/01/2020).

[73] 3rd Generation Partnership Project. *Technical Specification 36.214 v13.5.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical Layer Measurements*. Available online: https://www.3gpp.org/DynaReport/36214.htm. Sept. 26, 2017. (Visited on 02/01/2020).

[74] 3rd Generation Partnership Project. *3GPP - Release 13*. 3GPP - Release 13. Available online: https://www.3gpp.org/release-13. Feb. 1, 2020. (Visited on 02/01/2020).

[75] 3rd Generation Partnership Project. *Technical Specification 36.211 v13.13.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical channels and modulation*. Available online: https://www.3gpp.org/DynaReport/36211.htm. Jan. 6, 2020. (Visited on 02/01/2020).

[76] 3rd Generation Partnership Project. *Technical Specification 36.213 v13.15.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Physical Layer procedures*. Available online: https://www.3gpp.org/DynaReport/36213.htm. Jan. 6, 2020. (Visited on 02/01/2020).

[77] 3rd Generation Partnership Project. *Technical Specification 36.323 v13.6.0, Evolved Universal Terrestrial Radio Access (E-UTRA), Packet Data Convergence Protocol (PDCP) specification*. Available online: https://www.3gpp.org/DynaReport/36323.htm. Feb. 25, 2020. (Visited on 02/25/2020).

[78] 3rd Generation Partnership Project. *About 3GPP*. Available online: https://www.3gpp.org/about-3gpp. (Visited on 02/01/2020).

[79] PyCom. *GPy - Triple-network WiFi, Bluetooth and LTE-M dev board*. Pycom. Available online: https://pycom.io/product/gpy/. (Visited on 02/01/2020).

[80] Rapeepat Ratasuk, Nitin Mangalvedhe, Zhilan Xiong, Michel Robert, and David Bhatoolaul. "Enhancements of narrowband IoT in 3GPP Rel-14 and Rel-15." In: *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*. 2017 IEEE Conference on Standards for Communications and Networking (CSCN). Helsinki, Finland: IEEE, Sept. 2017, pp. 60–65. ISBN: 978-1-5386-3070-9. DOI: 10.1109/CSCN.2017.8088599. (Visited on 01/13/2020).

[81] Jim Roskind. *QUIC: Design Document and Specification Rationale*. English. accessed: 2022-10-26. 2013. URL: https://docs.google.com/document/d/1RNHkx_VvKWyWg6Lr8SZ-saqsQx7rFV-ev2jRFUoVD34/edit (visited on 10/26/2022).

[82] Ramon Sanchez-Iborra, Luis Bernal-Escobedo, and Jose Santa. "Machine learning-based radio access technology selection in the Internet of moving things." In: *China Communications* 18.7 (2021), pp. 13–24. DOI: 10.23919/JCC.2021.07.002.

[83] Ramon Sanchez-Iborra, Jesus Sanchez-Gomez, Juan Ballesta-Viñas, Maria-Dolores Cano, and Antonio F. Skarmeta. "Performance Evaluation of LoRa Considering Scenario Conditions." In: *Sensors* 18.3 (2018). ISSN: 1424-8220. DOI: 10.3390/s18030772. URL: https://www.mdpi.com/1424-8220/18/3/772.

[84] Ruben M. Sandoval, Sebastian Canovas-Carrasco, Antonio-Javier Garcia-Sanchez, and Joan Garcia-Haro. "SMART USAGE OF MULTIPLE RAT IN IOT-ORIENTED 5G NETWORKS: A REINFORCEMENT LEARNING APPROACH." In: *2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K)*. 2018, pp. 1–8. DOI: 10.23919/ITU-WT.2018.8597940.

[85] J. Schlienz and D. Raddino. *Narrowband Internet of Things Whitepaper*. Available online: https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_application/application_notes/1ma266/1MA266_0e_NB_IoT.pdf. Aug. 8, 2016. (Visited on 01/13/2019).

[86] Semtech. *LoRa Technical Documents*. Specification. accessed: 2022-07-07. 2022. URL: https://lora-developers.semtech.com/documentation/technical-documents.

[87] Semtech Corporation. *AN1200.22 LoRa™ Modulation Basics*. accessed: 2022-07-06. 2015. URL: https://semtech.my.salesforce.com/sfc/p/E0000000JelG/a/2R0000001OJu/xvKUc5w9yjG1q5Pb2IIkpolW54YYqGb.frOZ7HQBcRc (visited on 07/06/2022).

[88] Semtech Corporation. *SX1257 Transceiver Datasheet*. accessed: 2022-10-24. 2018. URL: https://www.semtech.com/products/wireless-rf/lora-core/sx1257#documentation (visited on 10/24/2022).

[89] Semtech Corporation. *SX1276/77/78/79 Transceiver Datasheet*. accessed: 2022-07-11. 2020. URL: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R0000001Rbr/6EfVZUorrpoKFfvaF_Fkpgp5kzjiNyiAbqcpqh9qSjE (visited on 07/11/2022).

[90] Semtech Corporation. *Packet Size Considerations*. accessed: 2022-04-28. 2021. URL: https://lora-developers.semtech.com/documentation/tech-papers-and-guides/the-book/packet-size-considerations (visited on 04/28/2022).

[91] Semtech Corporation. *What are the typical ranges of good and poor values for the RSSI and SNR one should expect for an uplink?* English. accessed: 2022-07-06. July 2022. URL: https://lora-developers.semtech.com/community/faq/faq-lora (visited on 07/06/2022).

[92] Bernard Sklar and Fredric J. Harris. *Digital Communications: Fundamentals and Applications, 3rd Edition*. Pearson, 2020. 1136 pp. ISBN: 9780137569076.

[93] Randall R. Stewart. *Stream Control Transmission Protocol*. RFC 4960. accessed: 2022-04-28. Sept. 2007. DOI: 10.17487/RFC4960. URL: https://www.rfc-editor.org/info/rfc4960.

[94] Dimitrios Stiliadis, Kent Benson, William Courtney, Jon Bennett, Shahram Davari, Jean-Yves Le Boudec, Victor Firoiu, Dr. Bruce S. Davie, and Anna Charny. *An Expedited Forwarding PHB (Per-Hop Behavior)*. RFC 3246. Mar. 2002. DOI: 10.17487/RFC3246. URL: https://www.rfc-editor.org/info/rfc3246.

[95] Martin Stusek, Dmitri Moltchanov, Pavel Masek, Jiri Hosek, Sergey Andreev, and Yevgeni Koucheryavy. "Learning-Aided Multi-RAT Operation for Battery Lifetime Extension in LPWAN Systems." In: *2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. 2020, pp. 26–32. DOI: 10.1109/ICUMT51630.2020.9222440.

[96] The Huawei Documentation Team. *Technical Documentation: What Is LACP? How Does LACP Work?* June 2019. URL: https://support.huawei.com/enterprise/en/doc/EDOC1100086560#EN-US_TOPIC_0169439602.

[97] The Things Industries. *The Things Network Homepage*. accessed: 2022-07-06. 2022. URL: https://www.thethingsnetwork.org/ (visited on 07/06/2022).

[98] Joseph D. Touch. *Transport Options for UDP*. Internet-Draft. accessed: 2022-04-28. Mar. 2022. URL: https://datatracker.ietf.org/doc/html/draft-ietf-tsvwg-udp-options-18.

[99] *Transmission Control Protocol*. RFC 793. accessed: 2022-04-28. Sept. 1981. DOI: 10.17487/RFC0793. URL: https://www.rfc-editor.org/info/rfc793.

[100] u-Blox. *SARA-N2 Series Power-optimized NB-IoT (LTE Cat NB1) Modules*. u-blox. Available online: https://www.u-blox.com/en/product/sara-n2-series. June 23, 2016. (Visited on 02/01/2020).

[101] u-Blox. *SARA-N2 Series Data Sheet*. Available online: https://www.u-blox.com/sites/default/files/SARA-N2_DataSheet_(UBX-15025564).pdf. Nov. 4, 2019. (Visited on 02/01/2020).

[102] Floris Van den Abeele, Jetmir Haxhibeqiri, Ingrid Moerman, and Jeroen Hoebeke. "Scalability Analysis of Large-Scale LoRaWAN Networks in ns-3." In: *IEEE Internet of Things Journal* 4.6 (2017), pp. 2186–2198. DOI: 10.1109/JIOT.2017.2768498.

[103] Walter Weiss, Dr. Juha Heinanen, Fred Baker, and John T. Wroclawski. *Assured Forwarding PHB Group*. RFC 2597. June 1999. DOI: 10.17487/RFC2597. URL: https://www.rfc-editor.org/info/rfc2597.

[104] Semtech Wireless and Sensing Products Division. *Semtech SX1301 Digital Baseband Chip Datasheet*. English. accessed: 2022-10-17. June 2017. URL: https://www.semtech.com/products/wireless-rf/lora-core/sx1301#documentation (visited on 10/17/2022).

[105] Joschka Wirges and Uwe Dettmar. "Performance of TCP and UDP over Narrowband Internet of Things (NB-IoT)." In: *2019 IEEE International Conference on Internet of Things and Intelligence System (IoTaIS)*. BALI, Indonesia: IEEE, Nov. 2019, pp. 5–11. ISBN: 978-1-72812-516-9. DOI: 10.1109/IoTaIS47347.2019.8980378. URL: https://ieeexplore.ieee.org/document/8980378/ (visited on 04/28/2022).

DECLARATION

I, Andreas Philipp Matz, hereby declare that this doctoral thesis titled: "Low Power Wide Area Networks Bundling" and the work presented herein are my own. I declare that the research described is original except where otherwise indicated. Any work of others is denoted and acknowledged as such and referenced in the bibliography section.

Jaén, November 2022

Andreas Philipp Matz