Université Saint-Joseph de Beyrouth
Faculté d'ingénierie

Low Power Wide Area Networks for the Internet of Things

Framework, Performance Evaluation, and Challenges of LoRaWAN and NB-IoT

Samer Lahoud   Melhem El Helou

ESIB, Saint Joseph University of Beirut, Lebanon

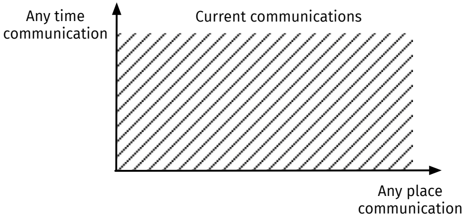
WPMC 2019, Lisboa, Portugal

Tutorial Outcomes

- What are the main drivers for Low Power Wide Area Networks?
- What are the fundamental mechanisms that enable to meet the LPWAN requirements?
- What are the main design principles of LoRaWAN and NB-IoT?
- How do we evaluate the performance of LoRaWAN and NB-IoT deployments in terms of coverage and capacity?
- What are the challenges for optimizing the performance of LoRaWAN and NB-IoT?

Outline

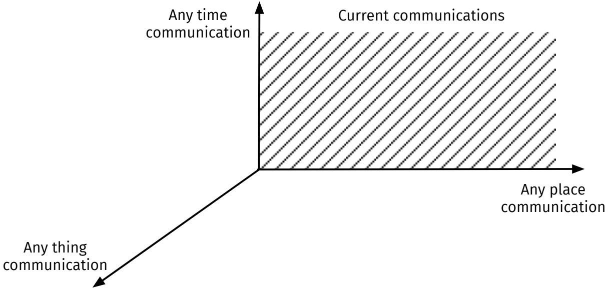
A New Dimension in Communications



Source: The Internet of Things, ITU Internet Reports, 2005

- Current communications brought the ABC (Always Best Connected) paradigm
- The Internet of Things (IoT) explores a new dimension in communications

A New Dimension in Communications



Source: The Internet of Things, ITU Internet Reports, 2005

- Current communications brought the ABC (Always Best Connected) paradigm
- The Internet of Things (IoT) explores a new dimension in communications

IoT Scenarios

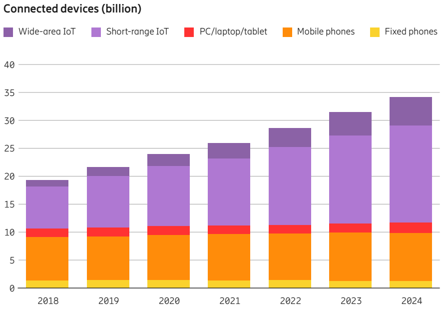
Internet of Things

The Internet of Things (IoT) generally refers to scenarios where network connectivity and computing capability extends to devices, sensors, and everyday items (ISOC IoT Overview, 2015).

Scenario	Example
Human	Wearables for health monitoring
Home	Heating, security automation
Retail	Self-checkout, inventory optimization
Vehicles	Condition-based maintenance
Cities	Traffic control, environmental monitoring

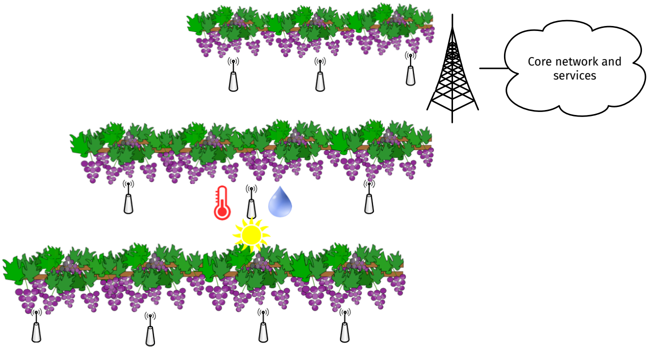
Evolution of IoT Devices

- The number of IoT devices will exceed the number of PCs and mobile phones



Source: Ericsson mobility report, 2019

The Case of IoT for Smart Agriculture



- Periodic sensing of microclimates in vineyards

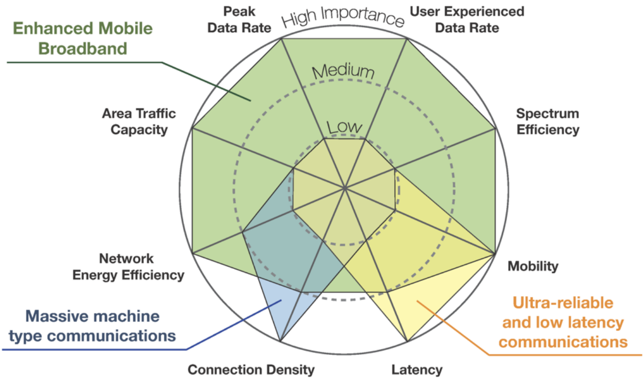
Constraints on the Device and Network Layers

- Difficult physical accessibility and limited access to power sources
  - Wireless communications
  - Autonomy and long battery life operation
- Wide area coverage with a large number of communicating devices
  - Scalable deployment
  - Cost efficient devices
- Very loose bandwidth and latency constraints
  - Adaptive radio and access mechanisms

Challenge

Do existing wireless networking technologies satisfy these constraints?

Capabilities of IMT-2020



Source: IMT 2020 Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond, ITU-R M.2083

Emergence of Low Power Wide Area Technologies

Low Power Wide Area Networks

Low power refers to the ability of an IoT device to function for many years on a single battery charge, while at the same time it is able to communicate from locations where shadowing and path loss would limit the usefulness of more traditional cellular technologies (3GPP Low Power Wide Area Technologies, GSMA White Paper, 2016)

- Low Power Wide Area (LPWA) technologies are designed to support mMTC
- Typical scenarios[1] for LPWAN:
  - Smart grid
  - Industrial asset monitoring
  - Critical infrastructure monitoring
  - Agriculture

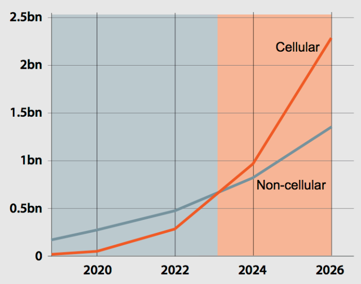
[1] Usman Raza *et al.*, Low Power Wide Area Networks: An Overview, IEEE Communications Surveys & Tutorials, 2017

LPWAN Requirements

Indicator	Requirement
Power consumption	Devices operate for 10 years on a single charge
Device unit cost	Below $5 per module
Dependability	Completely unattended and resilient operation
Coverage	Improved outdoor and indoor penetration coverage
Security	Secure connectivity and strong authentication
Data transfer	Supports small, intermittent blocks of data
Design complexity	Simplified network topology and deployment
Network scalability	Support of high density of devices

LPWAN Candidate Technologies

- A large number of candidate technologies:
  - LoRaWAN, Sigfox, Ingenu, NB-IoT, etc.
- Two main design approaches:
  - Adapt and leverage existing cellular technology (Cellular IoT)
  - Develop a clean-slate technology



LPWAN connections forecast. Source: ABI Research, 2019

Outline

Low Device Complexity and Cost

Device Complexity and Cost

- Devices are mainly composed of:
  - a processing unit: usually a microcontroller with a limited amount of memory
  - a sensing unit: sensors and analog to digital converters
  - a radio unit: usually a transceiver capable of bidirectional communications
- The radio unit complexity and cost are primarily related to the complexity of:
  - digital baseband processing
  - radio-frequency (RF) analog processing

Digital Baseband Processing

- Reduce baseband processing complexity through:
  - limiting message size
  - using simple channel codes
  - not using complex modulations or multiple-input multiple-output (MIMO) transmissions
  - supporting only half-duplex operation: no simultaneous transmission and reception

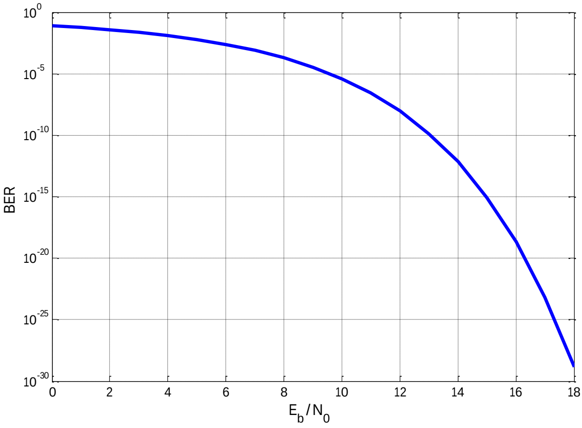
RF Processing

- Reduce RF processing complexity and cost through:
  - using one transmit-and-receive antenna
  - not using a duplexer (since only half-duplex operation is supported)
  - on-chip integrating power amplifier $\Rightarrow$ single-chip modem implementation (since transmit power is limited)

Reliability Under Extreme Coverage Conditions

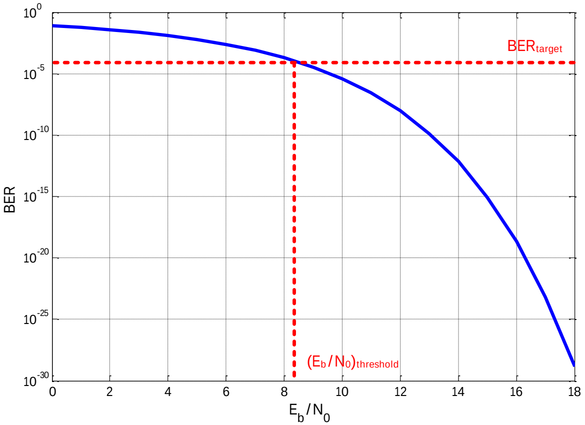
Reliability and Radio Quality

- Reliability $\Rightarrow$ bit error rate (BER) $\leq$ target BER
- The energy per bit to noise power spectral density ratio (E_b/N_0) is defined as the ratio of the energy per bit (E_b) to the noise power spectral density (N_0)



Reliability and Radio Quality

- Reliability $\Rightarrow$ bit error rate (BER) $\leq$ target BER
- The energy per bit to noise power spectral density ratio (E_b/N_0) is defined as the ratio of the energy per bit (E_b) to the noise power spectral density (N_0)



Reliability and Radio Quality

$$BER \leq BER_{target} \Leftrightarrow \frac{E_b}{N_0} \geq \left(\frac{E_b}{N_0}\right)_{threshold}$$

- $(E_b/N_0)_{threshold}$ does not depend on the signal bandwidth and data rate
- The SNR, or equivalently the carrier-to-noise ratio (CNR or C/N), is defined as the ratio of the received signal power C to the power of the noise N within the bandwidth of the transmitted signal

$$SNR = \frac{C}{N} = \frac{E_b/T_b}{N_0 B} = \frac{E_b}{N_0}\frac{R_b}{B}$$

  - B is the signal bandwidth in Hz
  - R_b is the data rate (or bit rate) in b/s.

Receiver Sensitivity

$$BER \leq BER_{target} \quad \Leftrightarrow \quad SNR \geq \underbrace{\left(\frac{E_b}{N_0}\right)_{threshold} \frac{R_b}{B}}_{SNR_{threshold}}$$

$$\Leftrightarrow \quad S\,(\mathrm{dBm}) \geq \underbrace{SNR_{threshold}\,(\mathrm{dB}) \; + N\,(\mathrm{dBm})}_{\text{Receiver sensitivity}}$$

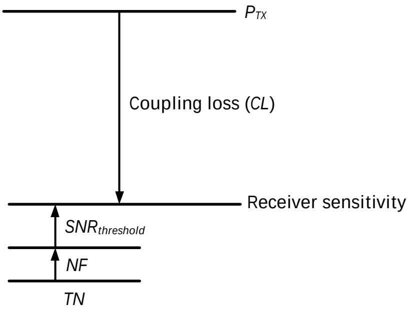
- $N\,(\mathrm{dBm})$ is the background noise power at the receiver $= TN\,(\mathrm{dBm}) + NF\,(\mathrm{dB})$
  - TN is the thermal noise caused by thermal agitation of charge carriers: $-174 + 10\log_{10}(B)$
  - NF is the noise figure caused by RF components

Maximum Coupling Loss

- The Maximum Coupling Loss (MCL) defines the maximum loss the system can cope with between a transmitter and a receiver:

$$MCL\ (\text{dB}) = P_{Tx} - \underbrace{\left(SNR_{threshold} - 174 + 10\log_{10}(B) + NF\right)}_{\text{Receiver sensitivity}}$$

  - P_{Tx} is the transmit power in dBm.



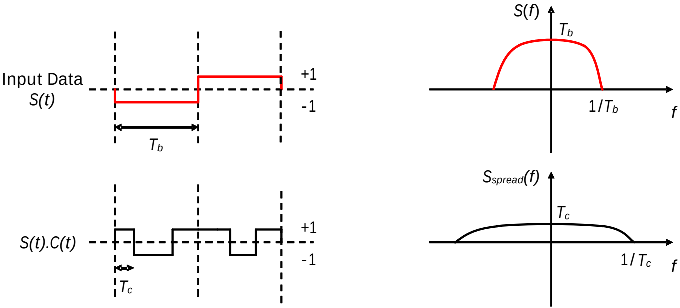
How to Improve Coverage?

- Coverage targets are usually specified in terms of *MCL*
- Increasing P_{TX}, or lowering *NF*, leads to higher device complexity and cost
- Reducing *B* leads to lower network capacity
- Improve coverage by reducing $SNR_{threshold}$ through:
  - Spread spectrum techniques (LoRaWAN)
  - Repetitions and efficient HARQ retransmissions (NB-IoT)

Chirp Spread Spectrum in LoRaWAN

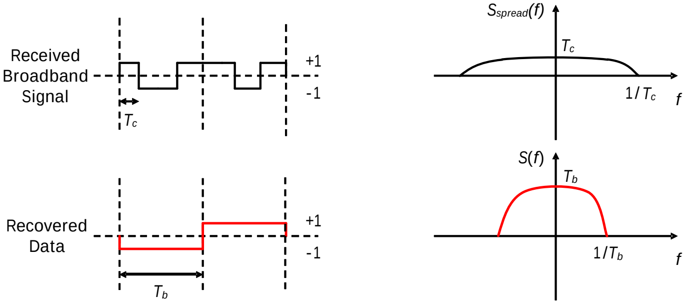
What is Spread Spectrum?

- Spread spectrum techniques deliberately spread a signal in the frequency domain, resulting in a signal with a wider bandwidth
- Direct-sequence SS (DSSS), frequency-hopping SS (FHSS), time-hopping SS (THSS), and chirp SS (CSS) are forms of spread spectrum
- Spreading process in DSSS systems: at the transmitter, the input data $S(t)$ is multiplied with a spreading code $C(t)$



What is Spread Spectrum?

- De-spreading process in DSSS systems: at the receiver, $S(t)$ is re-covered by re-multiplying with the same spreading code $C(t)$



Why Spread Spectrum?

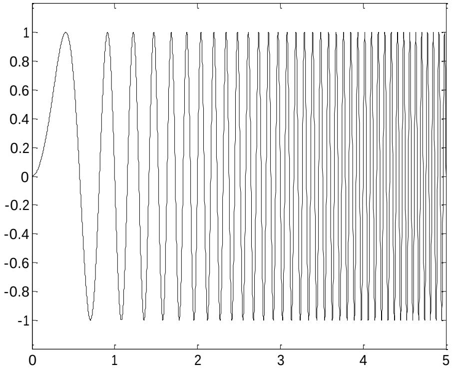
- Spread spectrum compensates for the *SNR* degradation

$$SNR = \frac{E_b}{N_0}\frac{R_b}{B} \Rightarrow SNR_{threshold} = \left(\frac{E_b}{N_0}\right)_{threshold} - \underbrace{10\log_{10}\left(\frac{B}{R_b}\right)}_{G_p}$$

  - $G_p = 10\log_{10}\left(\frac{B}{R_b}\right) = 10\log_{10}(T_b B)$ is the processing gain
- The higher G_p is
  - the lower $SNR_{threshold}$ is $\Rightarrow$ larger radio coverage
  - the lower R_b is (for a constant B)

Linear Chirp

- A linear chirp is a sinusoidal signal whose frequency linearly increases (*up-chirp*) or decreases (*down-chirp*) over time



A sinusoidal linear up-chirp in the time domain

Linear Chirp Theory

- A linear chirp waveform can be written as:

$$x(t) = a(t)sin(2\pi f_0 t + \pi\mu t^2 + \phi_0)$$

  - $a(t)$ is the chirp envelope which is zero outside a time interval of length T
  - f_0 is the initial frequency
  - μ is the chirp rate, or chirpyness
  - ϕ_0 is the initial phase.

- The instantaneous frequency $f(t)$ is defined as:

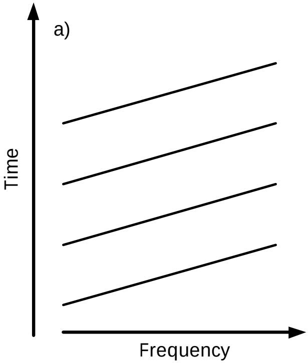
$$f(t) = \frac{1}{2\pi}\frac{d(2\pi f_0 t + \pi\mu t^2 + \phi_0)}{dt} = f_0 + \mu t$$

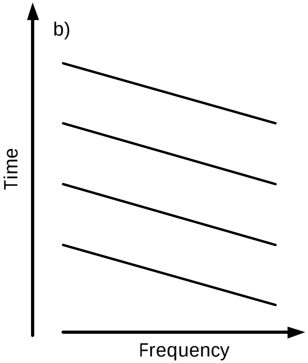
- The chirp rate μ represents the rate of change of the instantanous frequency:

$$\mu = \frac{df(t)}{dt}$$

Spectrograms of Linear Chirps

- $\mu > 0 \Rightarrow$ *up-chirps*, $\mu < 0 \Rightarrow$ *down-chirps*

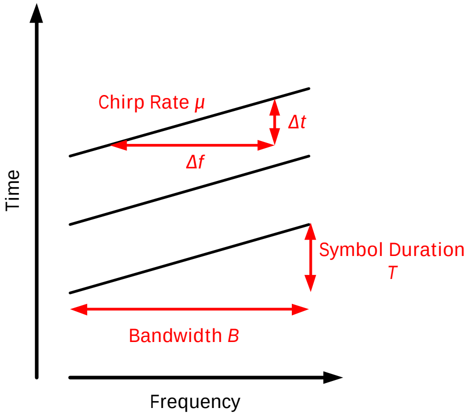




Spectrograms of linear *up-chirp* (a) and *down-chirp* (b)

Bandwidth Spreading

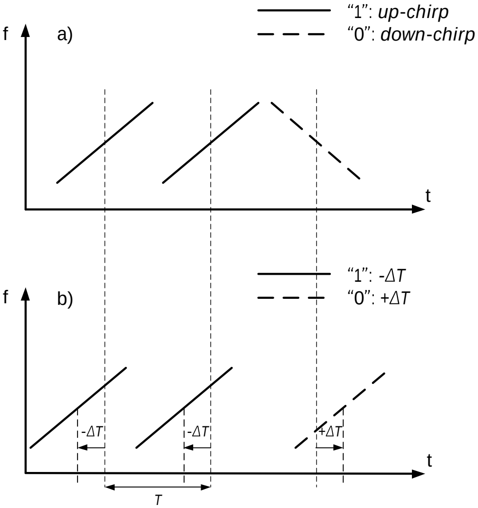
- The bandwidth B is defined as the range of the instantaneous frequency: $B = |\mu| T$
- The processing gain is given by the time-bandwidth product TB



What is Chirp Spread Spectrum?

- Chirp Spread Spectrum (CSS) is a spread spectrum technique that uses wideband linear frequency modulated chirps to encode information
- Encoding information using *up-chirp* and *down-chirp* signals:
  - Example: "1" $\Rightarrow$ transmit an *up-chirp*, "0" $\Rightarrow$ transmit a *down-chirp*
  - Chirps are transmitted in equidistant time steps
- Encoding information using only one chirp waveform with Pulse-Position Modulation (PPM):
  - M bits are encoded by transmitting a single *chirp* in one of 2^M possible time shifts $\Rightarrow$ data rate = M/T in b/s
  - Chirps are not transmitted in equidistant time steps
- At the receiver, the wanted information is re-covered through de-chirping

Example: Binary Orthogonal Keying (BOK) Schemes



a) BOK using *up*- and *down*- chirps b) BOK using PPM

Advantages of CSS

- CSS is robust to interference, multipath fading, and Doppler effect
- Time and frequency offsets between transmitter and receiver are equivalent, greatly reducing the complexity of the receiver design

Why CSS?

CSS provides a low-complexity, low-cost, low-power, yet robust alternative to the traditional SS techniques

Repetitions in NB-IoT

Signal Combination

- Devices in extreme coverage conditions blindly repeat information (without any feedback from the receiver)
- The receiver accumulates the blindly transmitted signals and combines all the repetitions
- Repetitions compensate for the SNR degradation

$$(SNR)_R \text{ (dB)} = \underbrace{10\log_{10}(R)}_{G_p} + (SNR)_1$$

- $(SNR)_R$ is the ideal SNR after combining R transmissions
- $(SNR)_1$ is the SNR of a single transmission
- $G_p = 10\log_{10}(R)$ is the processing gain

$$(SNR)_R \geq SNR_{threshold} \Rightarrow (SNR)_1 \geq \underbrace{SNR_{threshold} - 10\log_{10}(R)}_{\text{Reduced } SNR_{threshold}}$$

Processing Gain

- The higher G_p is
  - the lower the SNR threshold of a single transmission is $\Rightarrow$ larger radio coverage
  - the lower the useful data rate is
- In practice, channel estimation (CE) is rarely perfect
- CE errors result in lower processing gain: realistic $G_p < 10\log_{10}(R)$
- Improve CE and consequently G_p through averaging the channel estimates over multiple consecutive subframes (very slowly time-variant channel): *Cross-subframe CE*

Low Power Consumption

Deep Sleep Mode

- Most of the IoT applications require infrequent transmission of small data volumes
- Idle devices may enter a deep sleep mode. They:
  - shut down their transceiver
  - keep track of time and scheduled events via a low-power oscillator (that is kept running)
- Devices wake up from deep sleep mode to:
  - transmit data
  - open receive windows, or monitor paging channels

Battery Lifetime

- Battery lifetime is further increased through:
  - optimizing data transmission:
    - No signaling is required/exchanged when a device needs to transmit data (LoRaWAN)
    - Signaling messages are reduced, in comparison with existing cellular technologies, when a device needs to transmit data (NB-IoT)
  - optimizing device reachability: devices are reachable either periodically, or only after a device-originated data transfer (for a short period of time)

High Capacity

Support for Massive Number of Low-Rate Devices

- Trading off data rate for coverage
- How to increase network capacity?
  - LoRaWAN uses multiple orthogonal spreading factors simultaneously on the same frequency channel
  - NB-IoT uses single-tone transmissions in the UL, when coupling loss is high

Why Single-Tone Transmissions?

- The channel capacity C is given by:

$$C = B\log_2(1 + \frac{S}{N}) = B\log_2(1 + \frac{S}{N_0 B})$$

- When coupling loss is high, $\frac{S}{N_0 B} \ll 1 \Rightarrow \ln(1 + \frac{S}{N_0 B}) \approx \frac{S}{N_0 B}$.

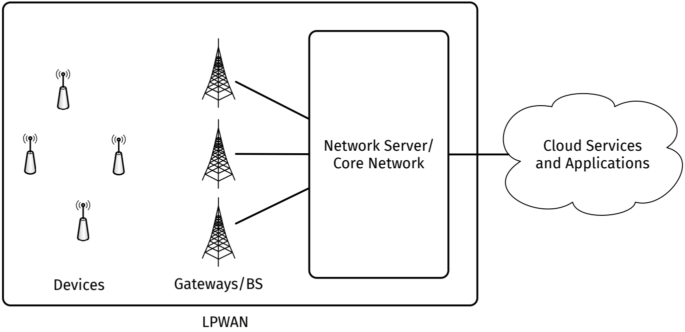
$$\Rightarrow C = \frac{S}{N_0}\log_2(e)$$

- As C no longer depends on B, allocating a single tone (subcarrier) for devices in bad coverage increases network capacity without loss of performance
- The uplink transmit power is concentrated on a narrower bandwidth, thus boosting the received SNR

Simplified Network Topology and Deployment

Network Topology and Deployment

- LoRaWAN has a simple network architecture and operates in license-free bands $\Rightarrow$ low-cost deployment
- NB-IoT reuses LTE frequency bands and infrastructure (through software upgrade) $\Rightarrow$ fast time-to-market



Outline

LoRaWAN

From LoRa to LoRaWAN

- LoRa
  - Modulation technique for LPWAN
- LoRaWAN
  - Uses LoRa modulation on physical layer
  - Proposes a MAC layer for access control
  - Specified by LoRa Alliance

LoRaWAN Timeline

- Cycleo first introduced LoRa in 2009
  - M2M communications
  - Large coverage
- Semtech acquired Cycleo in 2012
  - Patents filed in 2014
- LoRa Alliance initiated in 2014
  - Actility, Cisco, Bouygues, IBM, Orange, SK Telecom, KPN, ZTE, Semtech, La Poste, SoftBank, Swisscom, etc.
  - LoRaWAN 1.0.3 specification in 2018

LoRaWAN Radio Interface

What is LoRa?

Definition of LoRa

LoRa is a wireless modulation technique that uses Chirp Spread Spectrum (CSS) in combination with Pulse-Position Modulation (PPM)

- Processing gain given by $g_p = BT$
- Variable number of bits encoded in a symbol

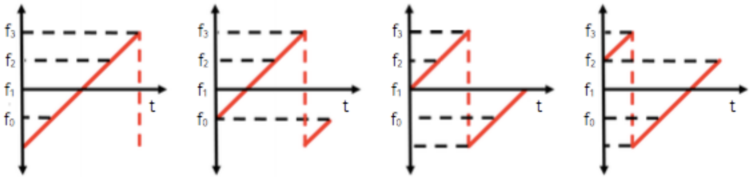
$$R_b = \frac{\log_2(g_p)}{T} = \log_2(g_p) \cdot \frac{B}{g_p}$$

- Spreading factor SF given by $\log_2(g_p)$

$$R_b = SF \cdot \frac{B}{2^{SF}}$$

LoRa Symbols

- $\log_2(g_p)$ bits are encoded by transmitting a single *chirp* in g_p possible cyclic time shifts
- Example: $g_p = 4 \Rightarrow 2$ bits/symbol



LoRa Data Rate

- LoRa includes a variable error correction scheme based on Hamming code
  - Improve the robustness of the transmitted signal at the expense of redundancy
- Given a coding rate CR, the data rate is given by:

$$R_b = SF \cdot \frac{B}{2^{SF}} \cdot CR$$

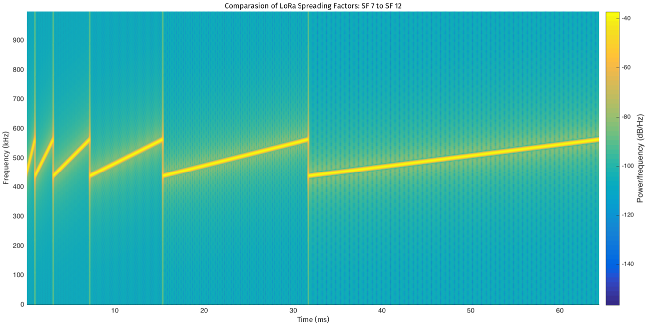
- R_b can also be written as:

$$R_b = SF \cdot \frac{B}{2^{SF}} \cdot \frac{4}{4 + CR}$$

with $1 \leq CR \leq 4$, and $6 \leq SF \leq 12$

LoRa Spreading Factors

- LoRa uses spreading factors from 6 to 12 (6 is not used in LoRaWAN)



LoRa Radio Performance

Spreading Factor	Data Rate[2] (kb/s)	Sensitivity (dBm)
6[3]	9.375	-118
7	5.468	-123
8	3.125	-126
9	1.757	-129
10	0.976	-132
11	0.537	-134.5
12	0.293	-137

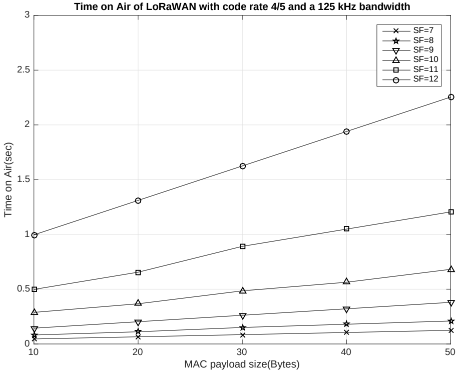
- Higher spreading factors lead to lower sensitivity and larger coverage
- Lower spreading factors lead to higher data rates

[2] CR = 1 and B = 125 kHz

[3] Spreading factor 6 is not used in LoRaWAN

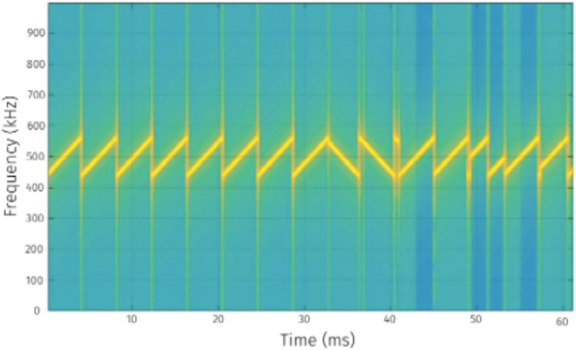
Spreading Factor and Time on Air

- The Time on Air is defined as the time required to transmit a packet in a sub-band
- The selection of the spreading factor impacts the Time on Air



LoRa Physical Layer

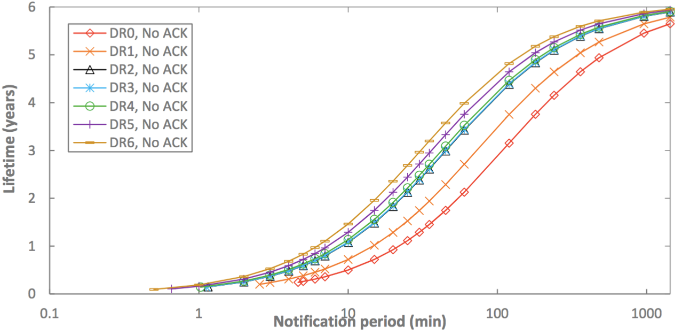
- LoRa transmission consists of:
  - 8 preamble (*up-chirp*) symbols
  - 2 synchronization (*down-chirp*) symbols
  - 5 modulated symbols (payload)
  - Up to 255 bytes of payload



LoRaWAN Data Rates

Data rate	Configuration	Indicative Physical Data Rate
0	LoRa SF12 / 125 kHz	250
1	LoRa SF11 / 125 kHz	440
2	LoRa SF10 / 125 kHz	980
3	LoRa SF9 / 125 kHz	1760
4	LoRa SF8 / 125 kHz	3125
5	LoRa SF7 / 125 kHz	5470
6	LoRa SF7 / 250 kHz	11000
7	FSK: 50 kbps	50000
8...14	RFU	
15	Defined in LoRaWAN	

Data Rates and Energy Consumption



Source: Lluís Casals *et al.*, Modeling the Energy Performance of LoRaWAN, Sensors, 2017

- For an end-device sending packets every 100 minutes, changing the spreading factor from 12 (DR0) to 7 (DR6) increases its lifetime by almost 1.5 years

LoRaWAN Channels

- Operates in license-free bands all around the world
  - 433, 868 (EU), 915 MHz
- EU 863-870MHz ISM Band
  - Default radiated transmit output power by devices: 14 dBm
  - Minimum set of three channels, maximum of 16 channels

Modulation	Bw [kHz]	Freq [MHz]	Data Rate	Nb Channels	Duty cycle
LoRa	125	868.10 868.30 868.50	DR0 to DR5 0.3-5 kbps	3	<1%

ETSI Limitations

- Restrictions on the maximum time the transmitter can be on or the maximum time a transmitter can transmit per hour
- Choice between
  - Duty-cycle limitation
  - Listen Before Talk Adaptive Frequency Agility (LBT AFA) transmissions management
- The current LoRaWAN specification exclusively uses duty-cycled limited transmissions to comply with the ETSI regulations

Duty Cycle Limitation

- The LoRaWAN enforces a per sub-band duty-cycle limitation
  - Each time a frame is transmitted in a given sub-band, the time of emission and the on-air duration of the frame are recorded for this sub-band
  - The same sub-band cannot be used again during the next T_{off} seconds where:

$$T_{off} = \frac{TimeOnAir}{DutyCyleSubband} - TimeOnAir$$

  - During the unavailable time of a given sub-band, the device may still be able to transmit on another sub-band
  - The device adapts its channel hopping sequence according to the sub-band availability

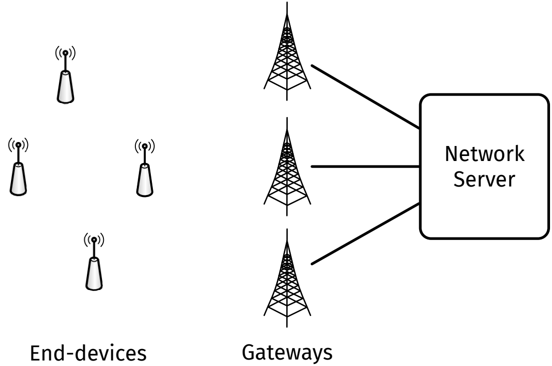
Example

A device just transmitted a 0.5 s long frame on one default channel. This channel is in a sub-band allowing 1% duty-cycle. Therefore this whole sub-band (868 – 868.6) will be unavailable for 49.5 s

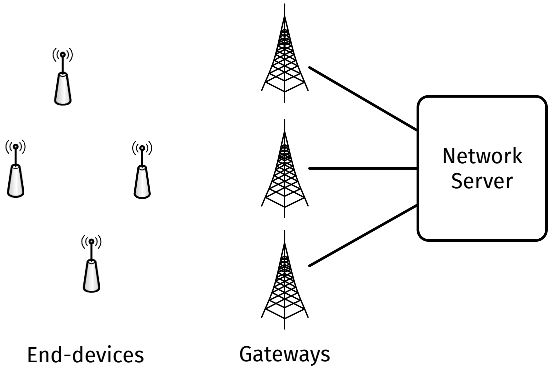
LoRaWAN Physical Architecture

LoRaWAN General Architecture

- LoRaWAN network architecture is typically laid out in a star-of-stars topology
  - End-devices
  - Gateways
  - Network server

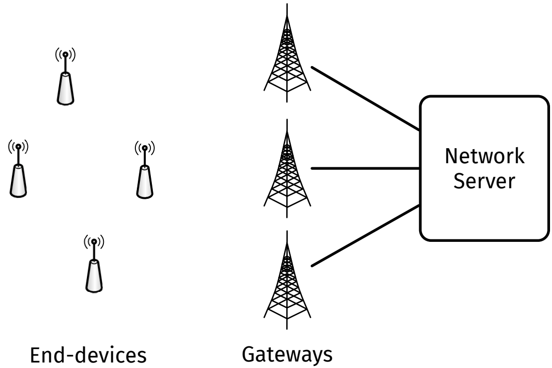


End-Devices



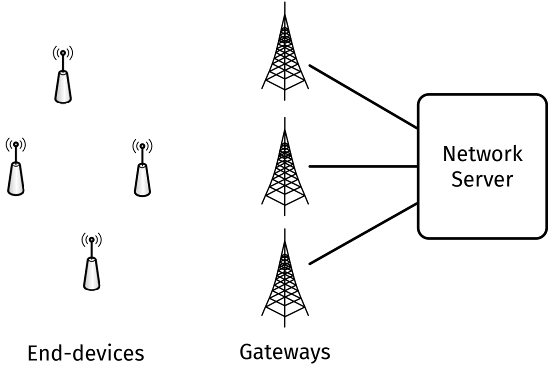
- End-devices are also called motes or devices
  - Communicate to one or more gateways via a wireless interface using single hop LoRa or FSK

Gateways



- Gateways are also called concentrators or base stations
  - Forward Frames between devices and network server
  - Connected to the network server via IP interfaces

Network Server



- Network server is a central server located at the backend
  - Provides mobility, frame control, and security functions
  - Adapts data transmission rates

LoRaWAN Protocol Architecture

Uplink transmission

- Uncoordinated data transmission
  - Devices transmit without any coordination on a randomly chosen channel
  - Regulated maximum transmit duty cycle
  - Regulated maximum transmit duration (or dwell time)
- Collisions occur in LoRaWAN
  - Simultaneous transmissions on the same channel and spreading factor collide

LoRaWAN Access Method

LoRaWAN is an ALOHA-type protocol: transmission by the device is based on its own communication needs with a small variation based on a random time basis

Device Classes

- Class A
  - Each uplink transmission is followed by two short downlink receive windows
  - Adapted for applications that only require downlink communication from the server shortly after the end-device has sent an uplink transmission
- Class B
  - In addition to class A, receive windows are opened at scheduled times
  - A time synchronized Beacon is sent by the gateway
- Class C
  - Nearly always open receive windows (unless transmitting)

Receive Windows for Class A Devices

- First receive window
  - Same channel (and data rate) as the uplink
- Second receive window
  - Predefined channel and data rate, and possibility to modify it by MAC commands

Messages

- Uplink messages
  - Sent by devices to the NS
  - Relayed by one or multiple gateways
  - [Preamble, PHDR, PHDR_CRC, Payload, CRC]
- Downlink messages
  - Sent by the NS to only one device and is relayed by a single gateway
  - [Preamble, PHDR, PHDR_CRC, Payload]
- Each device has two frame counters
  - Uplink frames, incremented by the device
  - Downlink frames, incremented by the NS

MAC Header

- Format
  - [MAC type, ..., Device Address, Frame Control, Frame Counter, Frame Options, Frame Port, Payload]
- Message Types
  - Join Request
  - Join Accept
  - Unconfirmed Data Up
  - Unconfirmed Data Down
  - Confirmed Data Up
  - Confirmed Data Down
  - RFU
  - Proprietary

ACK in Frame Control for Confirmed Mode

- If the ACK (demanding acknowledge) sender is an end-device
  - The network will send the acknowledgement using one of the receive windows opened by the end-device after the send operation
- If the sender is a NS
  - The end-device transmits an acknowledgment at its own discretion, possibly piggybacked with the next Data message
- A message is retransmitted (predefined number of times) if an ACK is not received

MAC Commands

- Commands are exchanged between devices and NS, not visible to the application layer
- Examples
  - Indicate the quality of reception of the device
  - Indicate the battery level of a device
  - Request the device to change data rate, transmit power, repetition rate or channel
  - Sets the maximum aggregated transmit duty-cycle of a device
  - Change to the frequency and the data rate set for the second receive window (RX2) following each uplink

Data Stored in Each device

- Device address
  - 7 bit network identifier
  - 25 bit network address arbitrarily assigned by the admin
- Application Identifier
  - 64 bits that uniquely identify the owner of the device (EUI-64)
- Session key
  - Used for integrity check and encryption/decryption of MAC only messages
- Application Session key
  - Used for integrity check and encryption/decryption of application data messages

Two Ways of Activation

- Over the air activation
  - Necessitates a globally unique end-device identifier (DevEUI), the application identifier (AppEUI), and an AES-128 key (AppKey)
  - Two MAC messages between NS and devices: Join and Accept
- Activation by Personalization
  - No MAC messages
  - The DevAddr and the two session keys NwkSKey and AppSKey are directly stored into the end-device

Adaptive Data Rate

- Objectives
  - Increase battery life
  - Maximize network capacity
- Data rate validation
  - A device periodically sets the ADR acknowledgment bit and waits for an acknowledgment from the network
  - If an ACK is not received, the device switches to the next lower data rate that provides a longer radio range

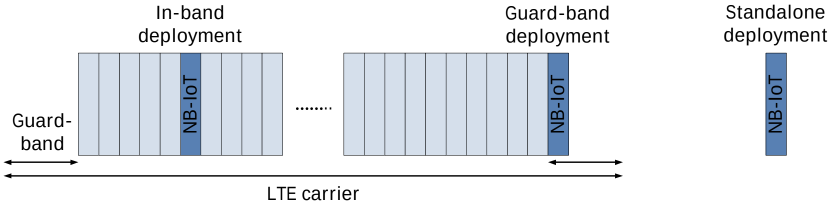
NB-IoT

NB-IoT and LTE

- NB-IoT is part of the 3GPP specifications: Releases 13, 14, and 15
  - NB-IoT fulfills the 5G mMTC requirements
- NB-IoT adapts and leverages the LTE ecosystem:
  - it reuses many LTE design principles:
    - Transmission schemes
    - Protocol architecture
    - Bearer management
    - Security management
    - Mobility management
  - it reuses LTE infrastructure through software upgrade

Deployment Flexibility

- NB-IoT supports three operation modes: in-band (LTE), guard-band (LTE), and standalone (*e.g.*, refarm the GSM carrier)



- NB-IoT supports frequency division duplex (FDD) and time division duplex (TDD) modes

Radio Interface

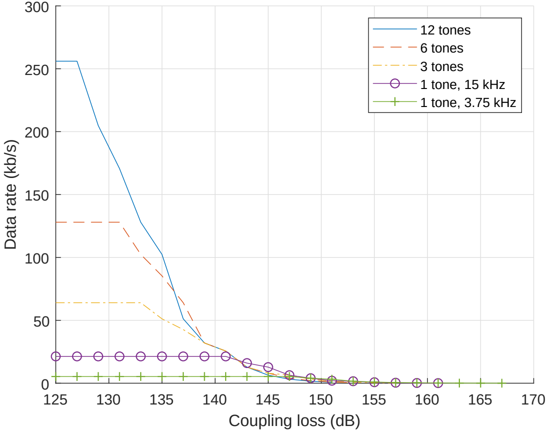
- Channel bandwidth: 180 kHz $\equiv$ 1 LTE Physical Resource Block (PRB) in the frequency domain
- Transmission schemes:
  - OFDMA (subcarrier spacing $\Delta f =$ 15 kHz) in the DL
  - SC-FDMA ($\Delta f =$ 15 kHz or 3.75 kHz) in the UL
- Smallest schedulable unit:
  - 1 PRB = 180 kHz (12 subcarriers) over 1 ms (1 subframe) in the DL
  - 1 Resource Unit (RU) in the UL
    - 180 kHz (12 subcarriers) over 1 ms
    - 90 kHz (6 subcarriers) over 2 ms
    - 45 kHz (3 subcarriers) over 4 ms
    - 15 kHz (1 subcarrier) over 8 ms
    - 3.75 kHz (1 subcarrier) over 32 ms

Radio Interface

- Maximum Transport Block Size (TBS):
  - 680 bits (R13), or 2536 bits (R14), mapped over up to 10 subframes (10 ms) in the DL
  - 1000 bits (R13), or 2536 bits (R14), mapped over up to 10 RUs in the UL
- Modulations:
  - QPSK in the DL
  - QPSK (for multi-tone transmission), $\pi/4$-QPSK, or $\pi/2$-BPSK (for single-tone transmission) in the UL
- Channel codes:
  - LTE tail-biting convolution code (TBCC) in the DL
  - LTE turbo code (for data transfer) and repetition code (for HARQ feedback) in the UL
- Repetitions for coverage enhancement:
  - up to 2048 repetitions in the DL and up to 128 repetitions in the UL
  - 20 dB coverage enhancement over GPRS $\Rightarrow$ sevenfold increase in coverage area (in an open environment), or (deep) indoor penetration
  - Transmission gaps can be configured to avoid long transmissions

Uplink Transmission Formats

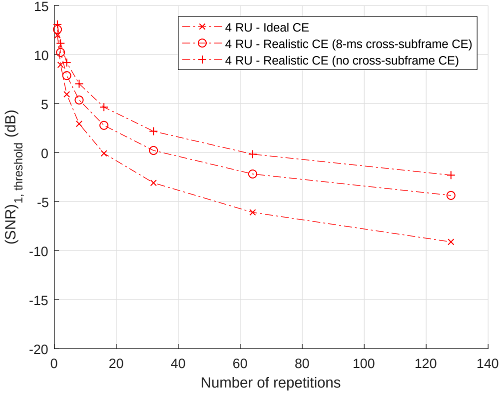
- When coupling loss is low, multi-tone transmissions provide higher data rates
- When coupling loss is high, all transmission formats provide similar data rates
- When coupling loss is very high, only single-tone transmissions are possible



TBS = 1000 bits, *CRC* = 24 bits, 8-ms cross-subframe CE

Repetitions: Ideal vs. Realistic Processing Gain

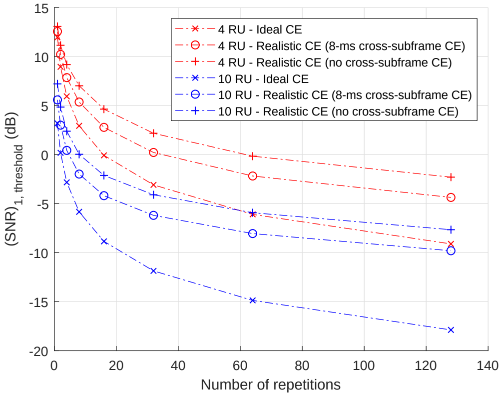
- Cross-subframe CE improves realistic G_p and decreases $(SNR)_{1,\ threshold}$
- As R is higher, $(SNR)_{1,\ threshold}$ decreases leading to an increase in the CE error
- As N_{RU} is higher, the channel coding rate decreases leading to a decrease in $(SNR)_{1,\ threshold}$



TBS = 1000 bits, CRC = 24 bits, T_{RU} = 1 ms, and B = 180 kHz

Repetitions: Ideal vs. Realistic Processing Gain

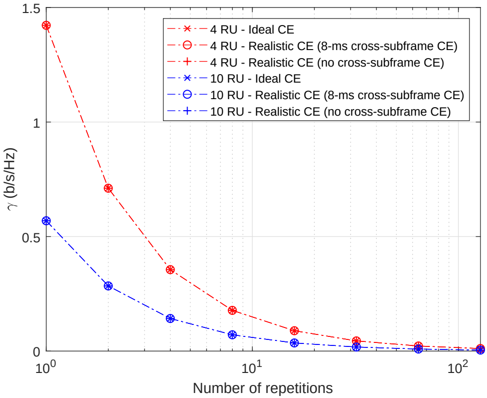
- Cross-subframe CE improves realistic G_p and decreases $(SNR)_{1,\ threshold}$
- As R is higher, $(SNR)_{1,\ threshold}$ decreases leading to an increase in the CE error
- As N_{RU} is higher, the channel coding rate decreases leading to a decrease in $(SNR)_{1,\ threshold}$



TBS = 1000 bits, CRC = 24 bits, T_{RU} = 1 ms, and B = 180 kHz

Spectral Efficiency γ

- Ideal CE achieves similar spectral efficiency with lower $(SNR)_{1,\ threshold}$
- As N_{RU} is higher, the channel coding rate decreases leading to a decrease in γ

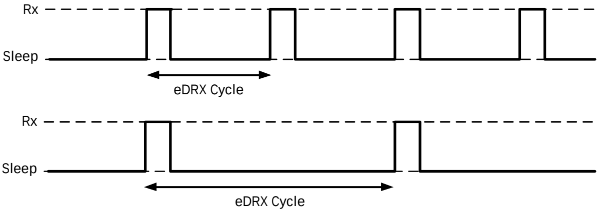


TBS = 1000 bits, CRC = 24 bits, T_{RU} = 1 ms, and B = 180 kHz

Device Reachability

extended Discontinuous Reception (eDRX)

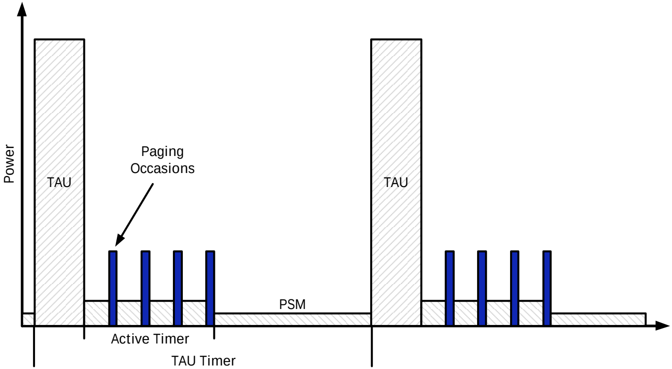
- An eDRX cycle is the time period between two paging occasions a device needs to monitor (up to 2 h, 54 min, and 46 s)
- In between these two occasions, the device is assumed to be in deep sleep mode
- The eDRX cycle is negotiated on a per-device basis



Two possible eDRX cycle configurations

Power-Saving Mode (PSM)

- In PSM, idle devices do not monitor paging channels $\Rightarrow$ unreachability
- A device leaves PSM to send application data or a periodic tracking area update message



Operation in PSM including periodic TAU

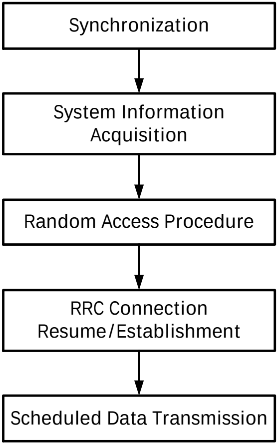
Power-Saving Mode (PSM)

- After data transfer, the device monitors paging occasions until an active timer expires
- When the active timer expires, the device re-enters PSM and is unreachable until the next device-originated event
- The tracking area update period is configurable on a per-device basis (up to a year)

Data Transmission

Cell Access

- From idle to connected mode:

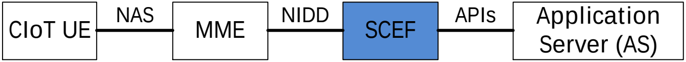


Data Transmission Optimization

- Signaling messages, that are required before a device transmits data, are reduced:
  - User Plane Cellular IoT (CIoT) Evolved Packet System (EPS) optimization procedure
    - Suspend/resume RRC connection (rather than release/re-establish RRC connection)
    - The device context is maintained at the UE, eNB, and MME during idle mode
  - Control Plane CIoT EPS optimization procedure
    - Transfer data over non-radio signaling (DoNAS, Data over Non-Access Stratum)
    - The IP packets are encapsulated in non-radio signaling messages and are sent to the MME
    - The MME extracts the IP packets and forwards them to the S-GW

Non-IP Data Delivery (NIDD)

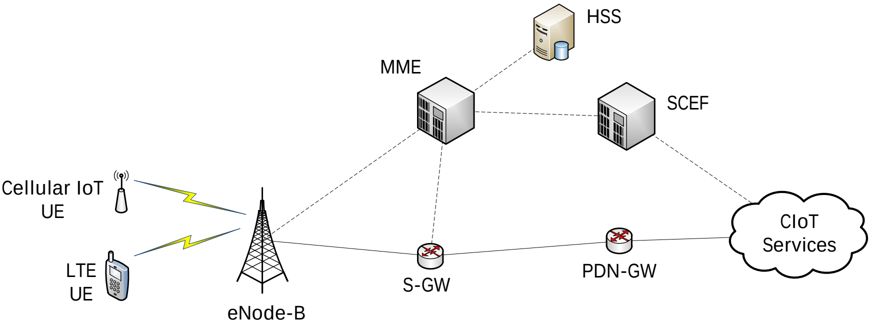
- To further reduce device power consumption, non-IP data transfer is also supported
- Non-IP data is transferred over non-radio signaling:
  - Non-IP data is encapsulated in non-radio signaling messages and is sent to the MME
  - The MME extracts the data and forwards it to the SCEF (Service Capability Exposure Function)



Service Capability Exposure Function (SCEF)

- SCEF is defined in Release 13
- SCEF provides APIs for small data transfers and control messaging
- The APIs securely expose network capabilities and services. They enable many use cases:
  - Device trigger delivery: wake up and notify a UE to connect to the AS
  - UE reachability and monitoring: check if a UE is currently reachable. If not, send back a notification when it becomes reachable.
  - Network configuration and parameters: set the PSM and eDRX parameters

Physical Architecture



SCEF: Service Capability Exposure Function

Mobility Management

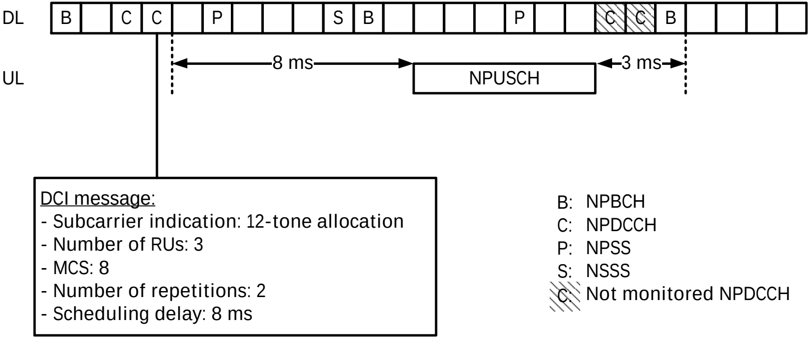
- Mobility management is limited to idle mode and is performed through cell reselection (with tracking area updates)
- Handover management is not supported, as NB-IoT is designed for infrequent and short messages
  - Connected devices do not perform mobility measurements
  - In case of connection loss (persistent link-layer failures), they switch to idle mode
  - In idle mode, they initiate cell reselection and then switch back to connected mode

Uplink Scheduling

- Scheduling information is transmitted in the Downlink Control Information (DCI) message
- UL scheduling information includes:
  - resource allocation (in time and frequency domains): subcarrier indication and number of RUs
  - Modulation and Coding Scheme (MCS)
  - number of repetitions
  - scheduling delay: time gap between the last DCI and the first scheduled UL subframe ($\geq$ 8 ms)
- A TB can be mapped over multiple RUs, allowing more redundancy bits for channel coding
- The scheduling delay allows the device to decode the DCI message, switch to transmission mode, and prepare for the UL transmission
- After data transmission, the device has at least 3 ms to switch to reception mode and monitor the next DCI message

Uplink Scheduling

- An uplink scheduling example[4]:



[4] O. Liberg et al., *Cellular Internet of Things - Technologies, Standards, and Performance*. Cambridge, MA, USA: American Press, 2017.

Downlink Scheduling

- The general aspects of DL scheduling are similar to those of UL scheduling
- DL scheduling information includes:
  - resource allocation: number of subframes per repetition
  - MCS
  - number of repetitions
  - scheduling delay ($\geq$ 4 ms)
  - HARQ-Ack resource: subcarrier index and time offset
- Resources for HARQ feedback are also scheduled
- After HARQ feedback transmission, the device has at least 3 ms to switch to reception mode and monitor the next DCI message

Power Control

- Closed-loop power control requires constant feedback and measurements, and is consequently power consuming
  - Open-loop power control is supported
- Power control for UL data channels:
  - If the number of repetitions is greater than 2, the transmit power P is the maximum device power: $P = P_{max}$
    - R13 defined two device power classes: $P_{max} = 20$ and 23 dBm
    - R14 introduced one additional device power class: $P_{max} = 14$ dBm
  - If the number of repetitions is 1 or 2, the transmit power is determined by:

$$P\,(\text{dBm}) = \max\left\{P_{max}, P_{target} + \alpha L + 10\log_{10}(M)\right\}$$

    - P_{target} is the target received power
    - L is the estimated path loss
    - α is a path loss adjustment factor
    - M is a bandwidth adjustment factor

Power Control

- M relates P_{target} to target SNR

Bandwidth (kHz)	M
3.75	1/4
15	1
45	3
90	6
180	12

- P_{max}, P_{target}, and α are provided by higher-layer configuration signaling

Outline

Cell Range Estimation

Link Budget

- The link budget is a measure of all the gains and losses from the transmitter, through the propagation channel, to the target receiver
- The link budget can be expressed as:

$$P_{Rx} = P_{Tx} + G_{System} - L_{System} - L_{Channel} - M$$

  - P_{Rx} = median received power
  - P_{Tx} = transmit power
  - G_{System} = system gains (*e.g.,* antenna, diversity, and amplifier gains)
  - L_{System} = system losses (*e.g.,* cable and connector losses)
  - $L_{Channel}$ = median path loss
  - M = additional margins

Additional Margins

- Fading margin
- Interference margin
- Penetration margin:
  - indoor penetration loss (first wall): $\sim$ 18 dB (in dense urban environment), $\sim$ 15 dB (in urban environment), and $\sim$ 10 — 12 dB (in rural environment)
  - deep indoor penetration loss (second wall): +3 dB
- Body loss

Maximum Allowable Path Loss and Cell Range

- The maximum allowable path loss (*MAPL*) is expressed as:

$$MAPL = \max L_{Channel} \mid P_{Rx} \geq \text{receiver sensitivity}$$

$$\Rightarrow MAPL = P_{Tx} + G_{System} - L_{System} - M - \text{receiver sensitivity}$$

- The maximum allowable distance between the transmitter and the target receiver (cell range) depends on the *MAPL* and the channel model[5]
- Transmission and reception parameters are determined so as to achieve balanced links (*i.e.*, uplink *MAPL* = downlink *MAPL*)

[5] R. El Chall, S. Lahoud, and M. El Helou, "LoRaWAN Network: Radio Propagation Models and Performance Evaluation in Various Environments in Lebanon," in *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2366-2378, April 2019.

Illustration: NB-IoT Uplink Link Budget Analysis

(a)	P_{Tx} (dBm)	23
(b)	G_{System} = base station antenna gain (dBi)	12
(c)	L_{System} = base station cable loss (dB)	3
(d)	M (dB)	16
(e)	B (Hz)	3750
(f)	Receiver NF (dB)	3
(g)	Receiver N (dBm) = $-174 + 10\log_{10}(e) + (f)$	-135.3
(h)	Required SINR[6] at the receiver (dB)	-12.74
(i)	Receiver sensitivity (dBm) = $(g) + (h)$	-148
	Uplink *MAPL* (dB) = $(a) + (b) - (c) - (d) - (i)$	**164**
	Uplink cell range[7] - urban environment (km)	**11**

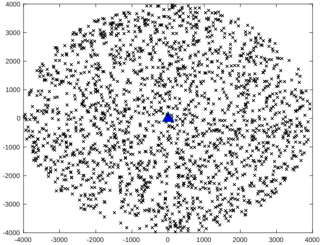
[6] Target transmission data rate = 160 b/s

[7] Okumura-Hata model: base station antenna height = 30 m, device antenna height = 1 m, carrier frequency = 862 MHz

Coverage of LoRaWAN

Evaluation Scenario

- Study area
  - Surface: circular cell of radius 4 km
  - Distribution of devices: uniform
  - Single gateway
  - Environment type: urban
- Uplink link budget
  - Parameters as in link level study
  - Shadow fading: $\mathcal{N}(0, 8)$
  - Interference: $IF = 3$ dB
  - Penetration loss: $L_{penetration} = 15$ dB
  - 50% of indoor devices



Signal to Noise Ratio Computation

- We consider the following specific parameters for LoRaWAN:
  - Transmit power: $P_{Tx} = 14$ dBm
  - Noise floor: $N = -153$ dBm
- We can now compute the uplink $SNR(i)$ at the gateway for end-device i:

$$SNR(i) = P_{TX} + G_{system} - L_{system} - L_{channel}(i) - \beta(i) L_{penetration} - N - IF$$

  - $L_{channel}(i)$ channel loss with shadow fading
  - $\beta(i) = 1$ for an indoor device, and 0 otherwise

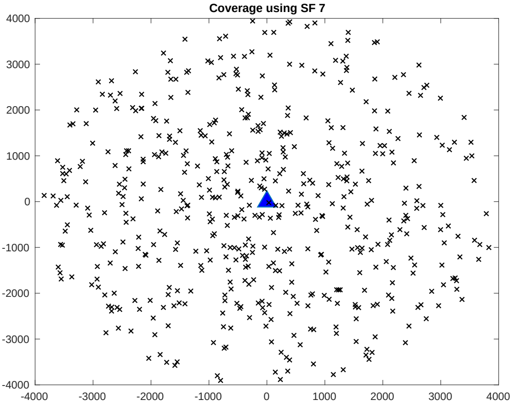
Spreading Factor Selection

- The spreading factor for each end-device is selected using the following matching table (Source: SX1276/77/78/79 Semtech datasheet):

SNR Interval (dB)	Spreading Factor
[-7.5, $+\infty$[	7
[-10, -7.5[	8
[-12.5, -10[	9
[-15, -12.5[	10
[-17.5, -15[	11
[-20, -17.5[	12

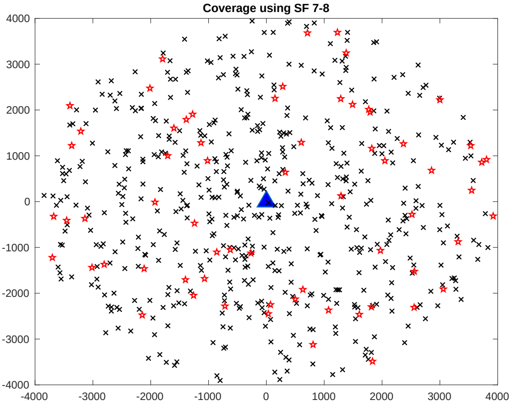
- Note that for SNR values lower than -20 dB, the end-device is considered out of coverage of the gateway

Coverage Study of LoRaWAN



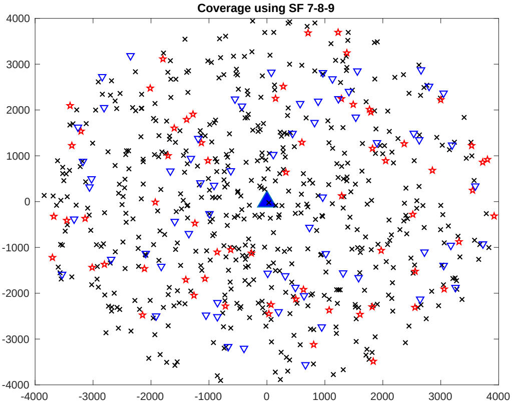
Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage Study of LoRaWAN



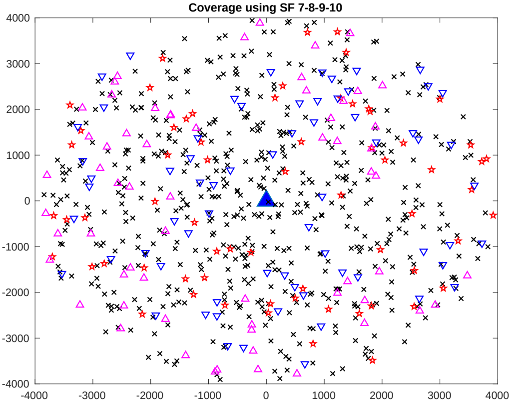
Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage Study of LoRaWAN



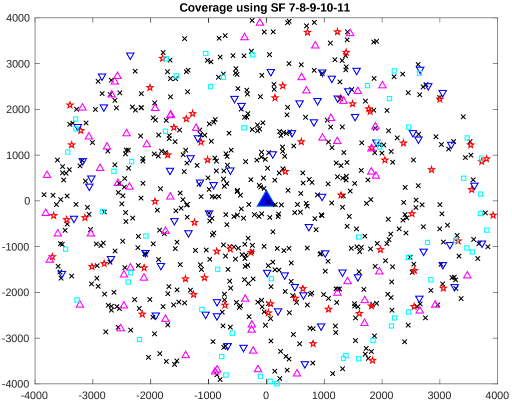
Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage Study of LoRaWAN



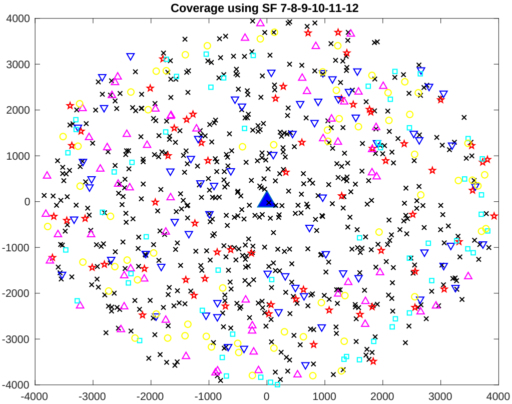
Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage Study of LoRaWAN



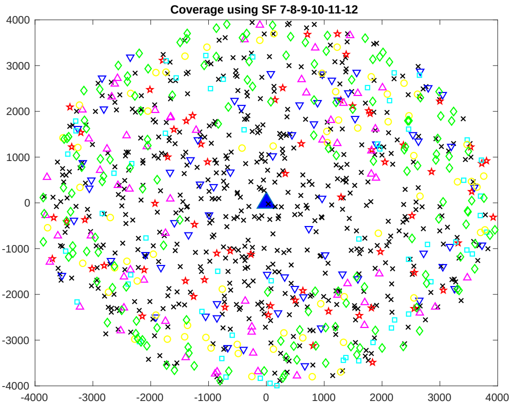
Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage Study of LoRaWAN



Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage Study of LoRaWAN



Spreading Factor	7	8	9	10	11	12
Cumulative coverage (%)	52.60	59.10	65.80	72.00	77.60	84.10

Coverage of NB-IoT

Device Rate Computation (1/2)

- We consider a similar evaluation scenario as for LoRaWAN
- The uplink SINR for each device i and each transmission format t is given by:

$$SINR(i,t) = P_{TX} + G_{system} - L_{system} - L_{channel}(i) - \beta(i)L_{penetration} - N(t) - IF$$

  - $P_{TX} = 23$ dBm
  - $N(t) = -174 + 10\log_{10}(B(t)) + NF$

Device Rate Computation (2/2)

- We compute the corresponding maximum data rate after link adaptation

$$D(i,t) = \max_{m}(\max_{r}(D(m,r,t)))$$

$$\text{with } SINR(i,t) \geq SINR_{threshold}(m,r,t)$$

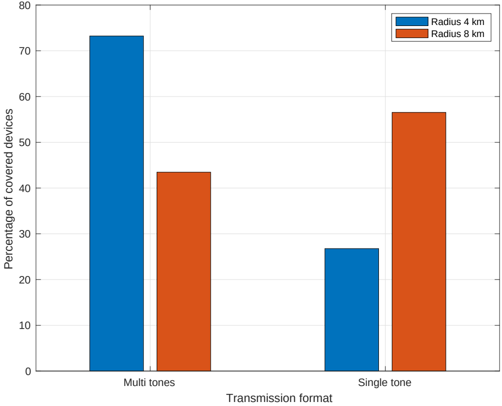
- The maximum data rate[8] for each device i is given by:

$$D(i) = \max_{t} D(i,t)$$

[8] We assume that multi-tone and single-tone transmissions provide similar spectral efficiencies respectively

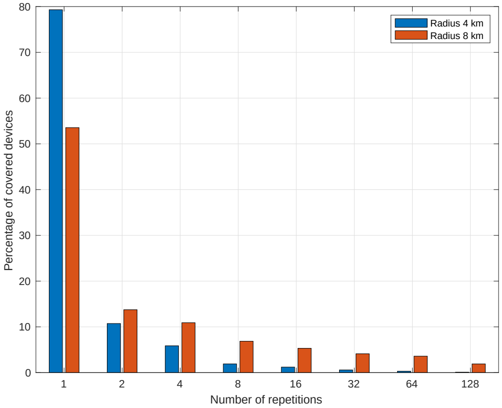
Transmission Formats

- For a cell radius of 4 km, good radio conditions enable to exploit the spectral efficiency of multi-tone transmissions
- For larger cells, single tone transmissions achieve better signal quality and become more attractive



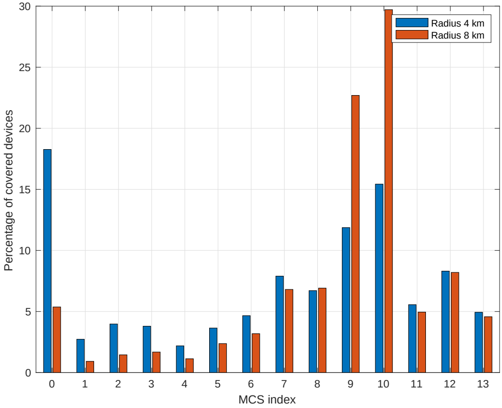
Repetitions

- Only 20% of devices use more than one repetition for a cell radius of 4 km
- 45% of devices use two or more repetitions in harsh radio conditions



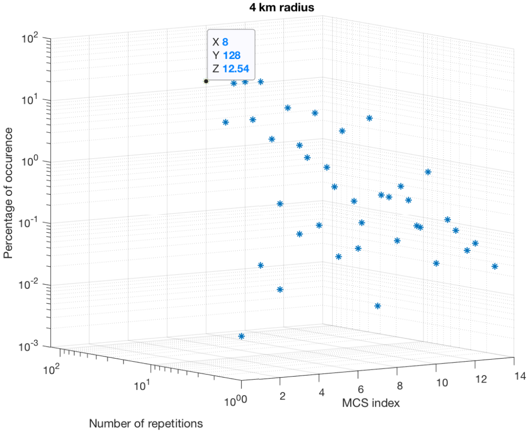
Modulation and Coding Schemes

- For a cell radius of 4 km, 50% of devices use high MCS index (greater or equal than 9) in order to increase data rates
- For larger cells, 70% of devices use high MCS index!



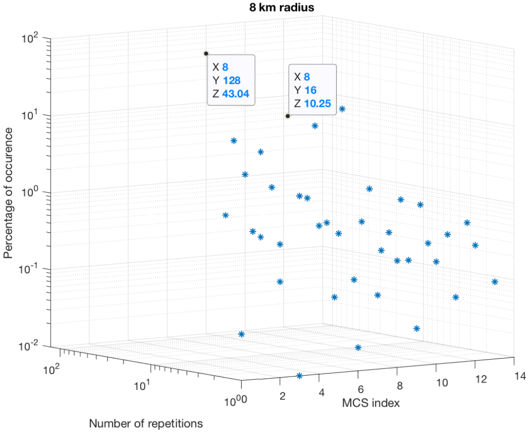
Link Adaptation: MCS index and Repetitions

- In order to combat harsh radio conditions and maximize rates, high MCS index is used jointly with a large number of repetitions



Link Adaptation: MCS index and Repetitions

- In order to combat harsh radio conditions and maximize rates, high MCS index is used jointly with a large number of repetitions



Coverage Comparison of NB-IoT and LoRaWAN

Outage Probability and Receiver Sensitivity

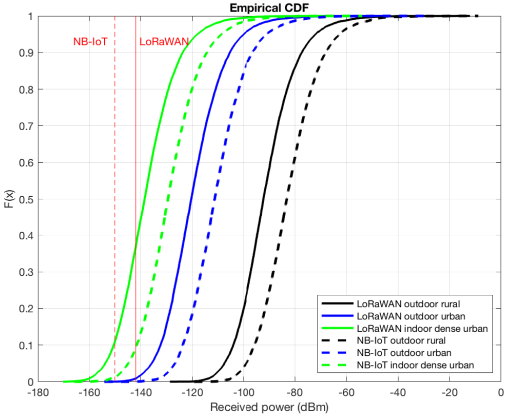
- Coverage outage is observed when the received power is less than the sensitivity
- Receiver sensitivity is given at different target rates:

	100 b/s	200 b/s
NB-IoT	-150.1 dBm	-147 dBm
LoRaWAN	-142 dBm	-139 dBm

- Coverage is computed for different environments (path loss formulas)
  - Outdoor rural
  - Outdoor urban
  - Indoor dense urban (penetration margin of 18 dB)
- Best coverage is computed for single-tone transmission and $\Delta f = 3.75$ kHz

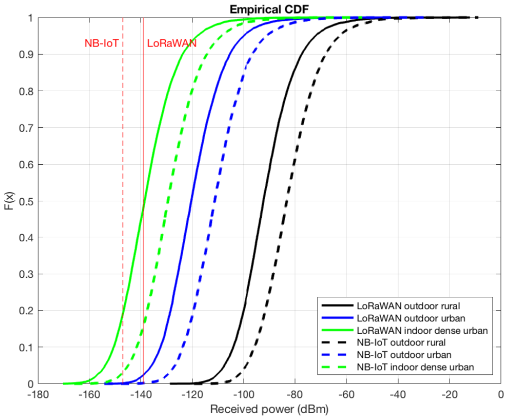
Coverage Comparison (1/3)

- The outage probability of NB-IoT at 100 b/s is almost null for all cases whereas it reaches 36% for LoRaWAN in indoor dense urban environments



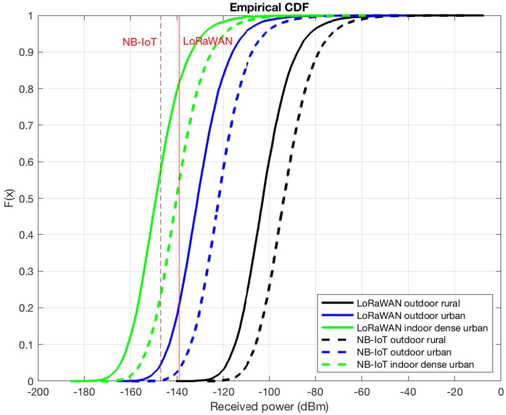
Coverage Comparison (2/3)

- The outage probability of NB-IoT at 200 b/s is slightly degraded in indoor dense urban environments



Coverage Comparison (3/3)

- For a larger cell radius (8 km), LoRaWAN exhibits severe coverage limitation at 200 b/s in comparison with NB-IoT



Capacity of LoRaWAN

Capacity Model for LoRaWAN

- The start times of the packets in an ALOHA channel are modeled as a Poisson point process with parameter λ packets/second
- Given a duty cycle limitation of $d = 1\%$, we must verify for spreading factor s:

$$\lambda(s) \leq \frac{d}{T_a(l, s)}$$

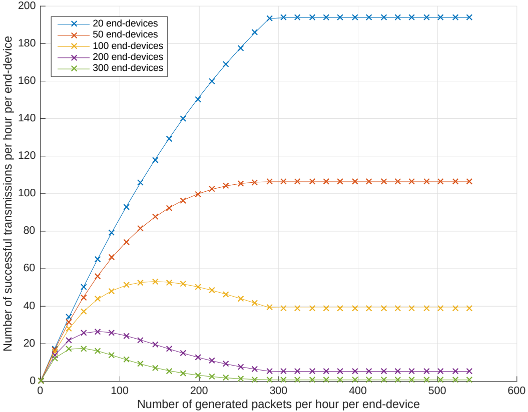
- The time to transmit a packet of l bytes (size of MAC payload) on spreading factor s is denoted $T_a(l, s)$
- The normalized total throughput of the LoRaWAN network with $N(s)$ end-devices on spreading factor s is given by:

$$S = \sum_s G(s)\exp(-2G(s))$$

with $G = N(s)\lambda(s)T_a(l, s)$

Packet Generation Rate

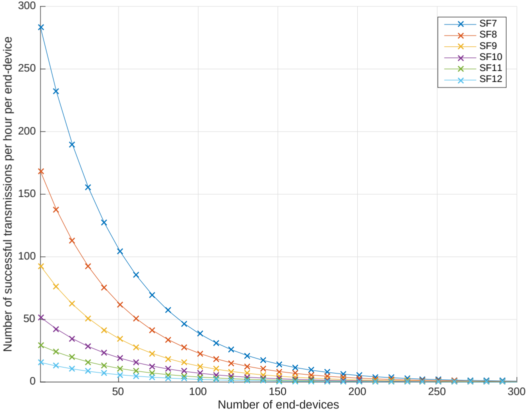
- For small number of end-devices, the throughput is limited by the duty cycle
- For large number of end-devices, the throughput is limited by collisions



l=50 bytes, SF=7

Spreading Factors and Packet Delivery

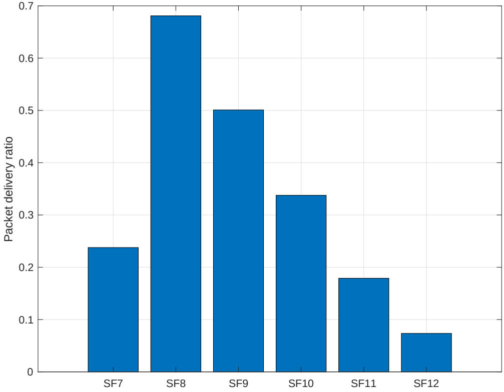
- For 50 end-devices, the average number of delivered packets per end-device per hour increases from 6 to 106 when SF decreases from 12 to 7



l=50 bytes, $\lambda(s) = \frac{d}{T_a(l,s)}$

Traffic Load and Packet Delivery

- The traffic load for high spreading factors induces a low delivery ratio



l=50 bytes, $\lambda = \frac{d}{T_a(l,s=12)}$

Capacity of NB-IoT

Average Data Rate Computation

- We compute the maximum data rate $D(i)$ for each device i
  - Link adapation
  - Choice of transmission format
- The device data rate for average radio conditions is given by:

$$\overline{D} = \frac{\sum_i D(i)}{I}$$

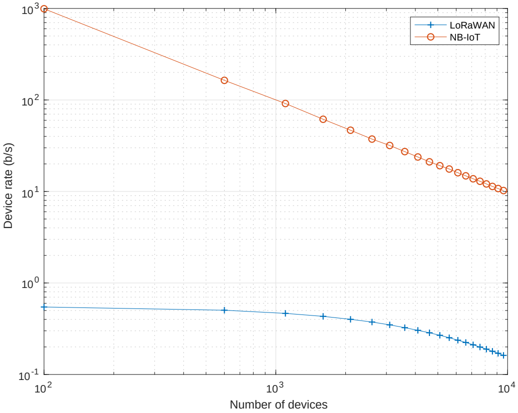
- After scheduling, the average device rate is computed as follows:

$$d = \frac{\overline{D}}{I}$$

Capacity Comparison of NB-IoT and LoRaWAN

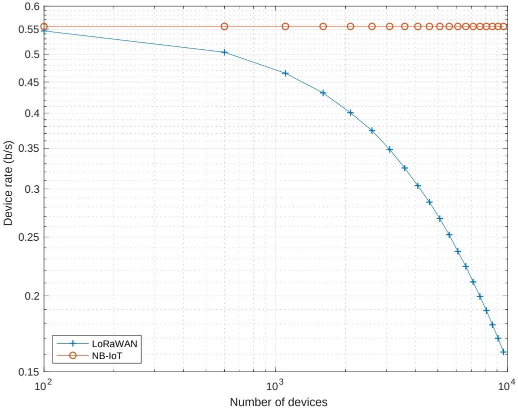
Network Capacity for NB-IoT and LoRaWAN

- The average rate decreases linearly with the number of devices for NB-IoT, while it drastically drops beyond 1000 devices for LoRaWAN



Network Capacity with Constant Arrival Rate

- If we consider an arrival rate of 5 packets per hour on each device, LoRaWAN does not succeed in delivering such service for more than 1000 devices



Outline

Revisiting LPWAN Requirements

- How do LoRaWAN and NB-IoT meet the LPWAN requirements?
  - Complexity: short messages, simple channel codes, and low-complexity modulations
  - Coverage: low receiver sensitivity
  - Energy consumption: optimized device reachability and data transmission
  - Capacity: multiple transmissions on the same channel
  - Simplified network topology and deployment

Typical Performance Results

- LoRaWAN shows coverage limitations, in comparison with NB-IoT, especially in indoor dense urban environments
- NB-IoT shows no coverage outage for a 4 km cell radius in indoor dense urban environments
- In order to maximize data rates, multi-tone transmissions are preferred for good radio conditions, while single-tone transmissions enable to combat harsh conditions
- Collisions hinder the performance of LoRaWAN in heavy loaded networks
- The capacity of an NB-IoT deployment can scale up to 200 000 devices sending 5 packets per hour
- NB-IoT requires more signaling messages than LoRaWAN, leading to higher energy consumption

Challenges

- What are the challenges for optimizing the performance of LoRaWAN and NB-IoT?
  - Adaptation of the radio propagation models
  - Interference mitigation
  - Link adaptation and resource allocation
  - Support for quality of service such as delay-bounded transmissions
  - Energy efficiency maximization and power management

Feedback and Tutorial Material

- We appreciate if you could take five minutes to complete the following evaluation form:

  http://tiny.cc/lpwan-iot

- Your feedback is important to us in order to continually improve our tutorial
- We made the tutorial sources available under Creative Common license CC BY-NC-SA 4.0

  https://github.com/samerlahoud/tutorial-lpwan-iot

- This is a human-readable summary of the license:

  https://creativecommons.org/licenses/by-nc-sa/4.0/