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

ICT 2018, Saint-Malo, France

Tutorial Outcomes

- Questions we are going to answer
- Feedback form
- Presentation slides are available

Outline

1. General Framework
2. Design Rationale
3. Technical Specification
4. Performance Evaluation
5. Research Challenges

Defining the Internet of Things

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, October 2015).

- Characteristics of IoT devices
  - Physical world interface
  - Computing capability
  - Communication interface

A New Dimension in Communications

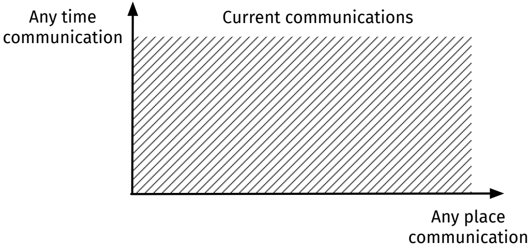


Figure 1: ITU Internet Reports (2005), The Internet of Things

- 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

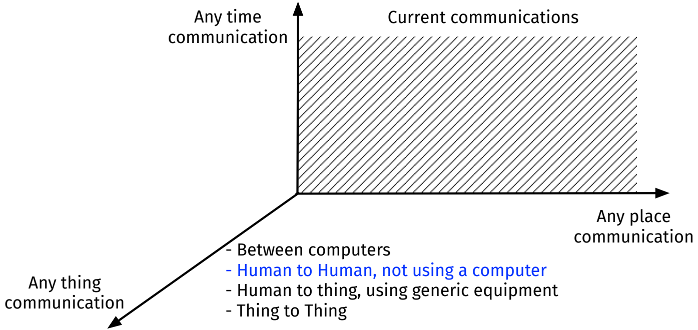
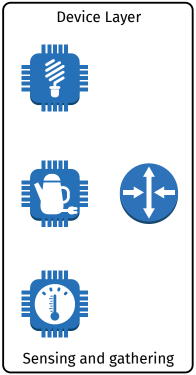


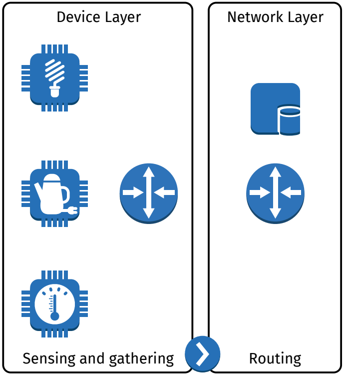
Figure 1: ITU Internet Reports (2005), The Internet of Things

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

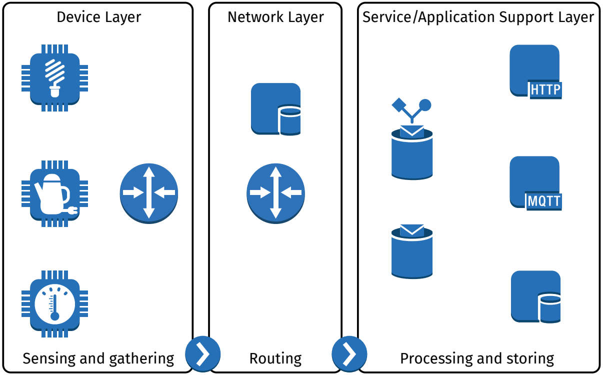
End-to-End IoT Chain



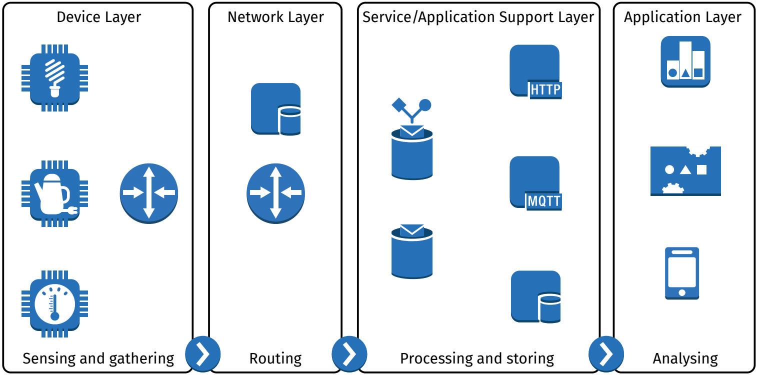
End-to-End IoT Chain



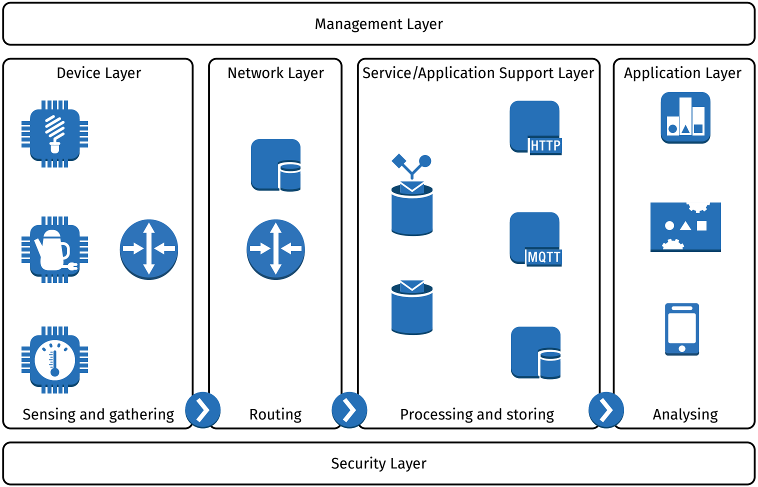
End-to-End IoT Chain



End-to-End IoT Chain



End-to-End IoT Chain



IoT for Smart Agriculture

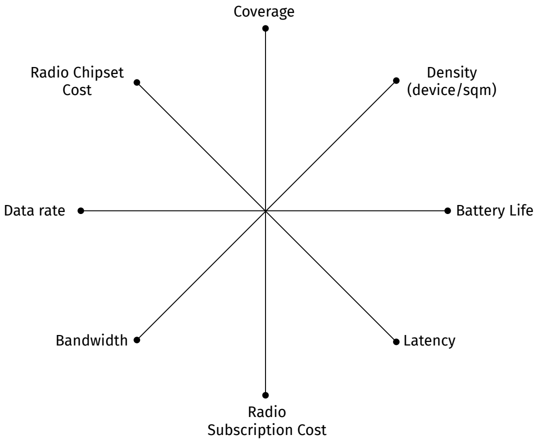
Constraints on the Device and Network Layers

- 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

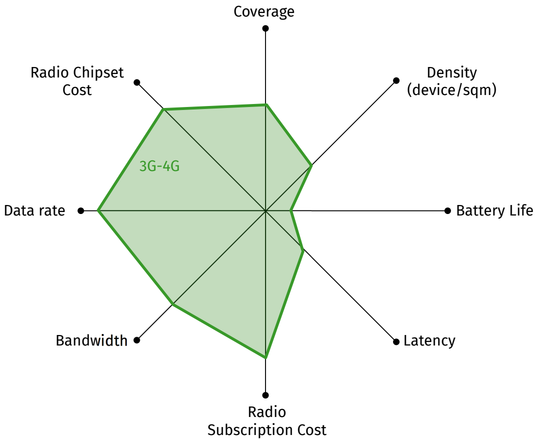
Challenge

Do existing wireless networking technologies satisfy these constraints?

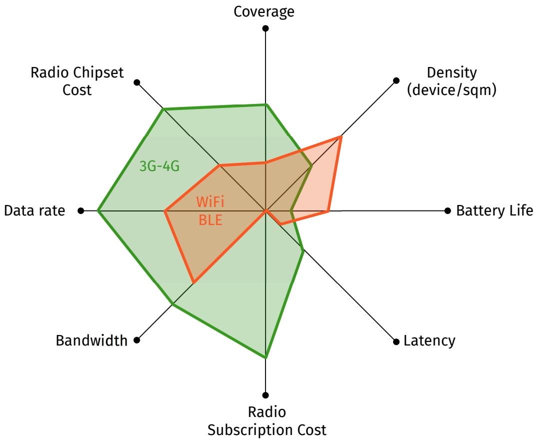
LPWAN Sweet Spot



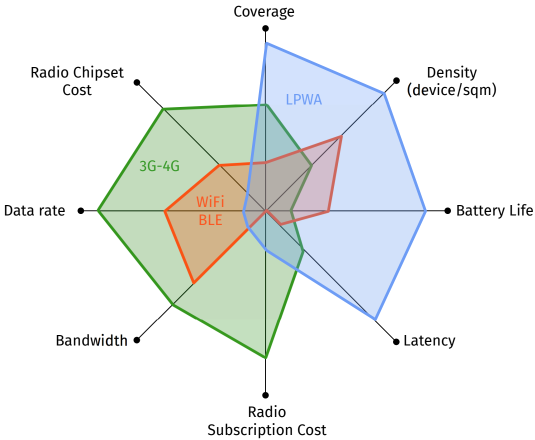
LPWAN Sweet Spot



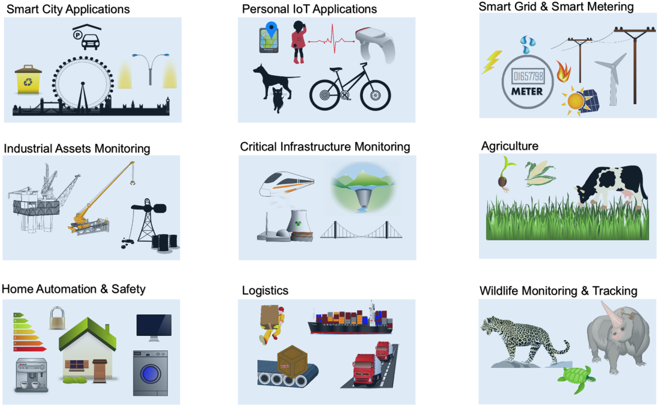
LPWAN Sweet Spot



LPWAN Sweet Spot

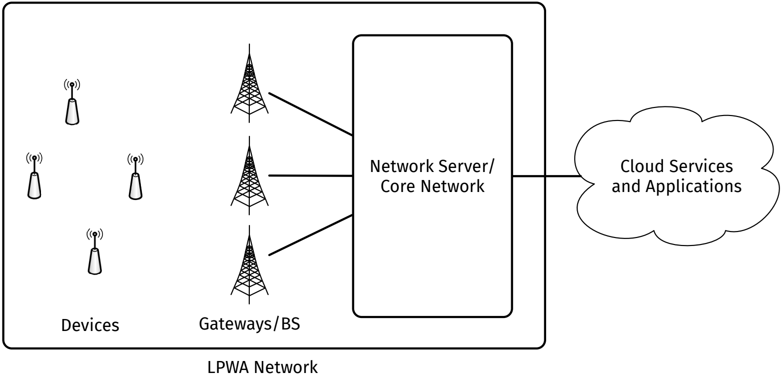


LPWAN Typical Use Cases



Usman Raza *et al.*, Low Power Wide Area Networks: An Overview, IEEE Communications Surveys & Tutorials, Issue 99, 2017

LPWAN Architecture



Common Characteristics of LPWAN Technologies

- Optimised radio modulation
- Star topology
- Frame sizes in the order of tens of bytes
- Frames transmitted a few times per day at ultra-low speeds
- Mostly upstream transmission pattern
- Devices spend most of their time in low-energy deep-sleep mode

LPWAN Technologies

Various technologies are currently candidating for LPWA: LoRaWAN, NB-IoT, Sigfox, Wi-SUN, Ingenu, etc.

Outline

Need for Optimised Radio Modulation

- End-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 power, that includes the transceiver circuit power and transmit signal power, is a main contributor to device power consumption
- Reducing the transmit signal power leads to lower signal-to-noise ratio (*SNR*) particularly at long distances

Modulation Requirement

Low *SNR* threshold to achieve acceptable radio transmission quality even with reduced transmit signal power

What is Spread Spectrum?

Spread Spectrum (SS)

Spread-spectrum techniques deliberately spread a signal in the frequency domain, resulting in a signal with a wider bandwidth

- Frequency-hopping SS (FHSS), direct-sequence SS (DSSS), time-hopping SS (THSS), and chirp SS (CSS) are forms of spread spectrum

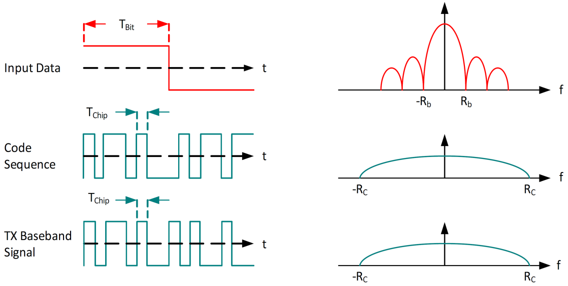


Figure 2: Spreading process in DSSS systems

What is Spread Spectrum?

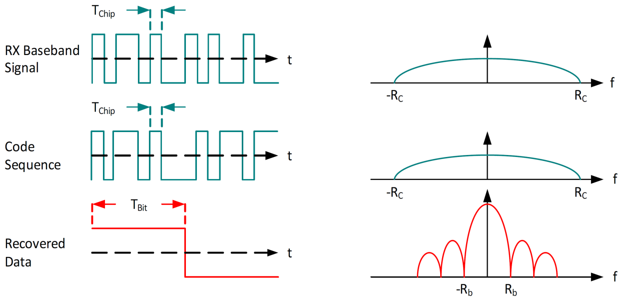


Figure 3: De-spreading process in DSSS systems

Direct Sequence Spread Spectrum

- At the transmitter, the wanted signal is multiplied with a spreading code
- At the receiver, the wanted signal is re-covered by re-multiplying with the same spreading code

Radio Quality Indicators

- 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
- 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)
- E_b/N_0 is a normalized *SNR* measure, also known as the "*SNR* per bit"
- The E_b/N_0 (or *SNR*) needs to be greater than the E_b/N_0 (or *SNR*) threshold for acceptable transmission quality (bit error rate (BER) $\leq$ target BER)

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

SNR threshold vs. E_b/N_0 threshold

Unlike the *SNR* threshold, the E_b/N_0 threshold does not depend on the signal bandwidth and bit-rate (including any use of spread spectrum)

Why Spread Spectrum?

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

where B is the signal bandwidth in Hz, R_b the bit-rate in b/s, and R_c the chip rate in chip/second

$$\Rightarrow \left(\frac{E_b}{N_0}\right)_{dB} = (SNR)_{dB} + G_p$$

where G_p is the processing gain given by: $G_p = 10log_{10}(BT_b) = 10log_{10}(\frac{R_c}{R_b})$

- Spread spectrum compensates for the SNR degradation

The higher the processing gain is...

- the lower the SNR threshold is $\Rightarrow$ lower receiver sensitivity $\Rightarrow$ larger radio coverage (assuming a fixed transmit signal power)
- the lower R_b is (assuming a fixed W)

G_p and Receiver Sensitivity

- The receiver sensitivity is given by:

$$S = SNR_{threshold} + \nu$$

where ν is the background noise power at the receiver: sum of the thermal noise (generated by the thermal agitation of charge carriers) and the noise figure (caused by RF components)

R_c/R_b	G_p	Receiver sensitivity (dBm)
128	21.0721	-121
256	24.0824	
512	27.0927	
1024	30.1030	
2048	33.1133	
4096	36.1236	

G_p and Receiver Sensitivity

- The receiver sensitivity is given by:

$$S = SNR_{threshold} + \nu$$

where ν is the background noise power at the receiver: sum of the thermal noise (generated by the thermal agitation of charge carriers) and the noise figure (caused by RF components)

R_c/R_b	G_p	Receiver sensitivity (dBm)
128	21.0721	-121
256	24.0824	-124
512	27.0927	
1024	30.1030	
2048	33.1133	
4096	36.1236	

G_p and Receiver Sensitivity

- The receiver sensitivity is given by:

$$S = SNR_{threshold} + \nu$$

where ν is the background noise power at the receiver: sum of the thermal noise (generated by the thermal agitation of charge carriers) and the noise figure (caused by RF components)

R_c/R_b	G_p	Receiver sensitivity (dBm)
128	21.0721	-121
256	24.0824	-124
512	27.0927	-127
1024	30.1030	
2048	33.1133	
4096	36.1236	

G_p and Receiver Sensitivity

- The receiver sensitivity is given by:

$$S = SNR_{threshold} + \nu$$

where ν is the background noise power at the receiver: sum of the thermal noise (generated by the thermal agitation of charge carriers) and the noise figure (caused by RF components)

R_c/R_b	G_p	Receiver sensitivity (dBm)
128	21.0721	-121
256	24.0824	-124
512	27.0927	-127
1024	30.1030	-130
2048	33.1133	
4096	36.1236	

G_p and Receiver Sensitivity

- The receiver sensitivity is given by:

$$S = SNR_{threshold} + \nu$$

where ν is the background noise power at the receiver: sum of the thermal noise (generated by the thermal agitation of charge carriers) and the noise figure (caused by RF components)

R_c/R_b	G_p	Receiver sensitivity (dBm)
128	21.0721	-121
256	24.0824	-124
512	27.0927	-127
1024	30.1030	-130
2048	33.1133	-133
4096	36.1236	

G_p and Receiver Sensitivity

- The receiver sensitivity is given by:

$$S = SNR_{threshold} + \nu$$

where ν is the background noise power at the receiver: sum of the thermal noise (generated by the thermal agitation of charge carriers) and the noise figure (caused by RF components)

R_c/R_b	G_p	Receiver sensitivity (dBm)
128	21.0721	-121
256	24.0824	-124
512	27.0927	-127
1024	30.1030	-130
2048	33.1133	-133
4096	36.1236	-136

Chirp Spread Spectrum

Linear Chirp

Linear Chirp

Sinusoidal signal whose frequency linearly increases (*up-chirp*) or decreases (*down-chirp*) over time

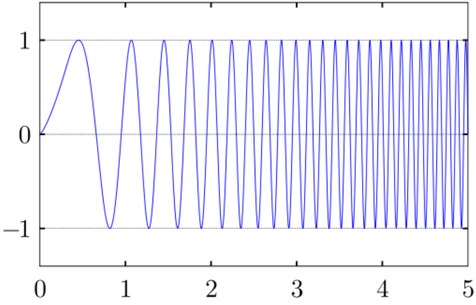


Figure 4: A sinusoidal linear up-chirp in the time domain

Linear Chirp Theory

- A linear chirp waveform can be written as:

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

where $a(t)$ is the envelope of the chirp signal which is zero outside a time interval of length T, f_0 the initial frequency, μ the chirp rate, or chirpyness, and ϕ_0 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*

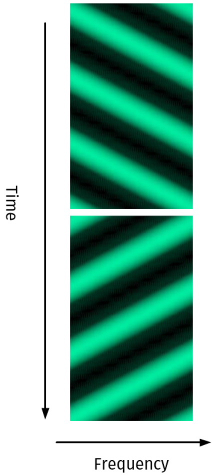
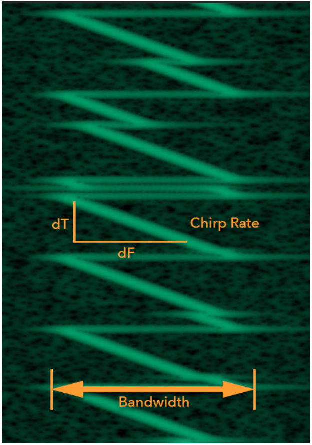


Figure 5: Spectrograms of linear *up-chirp* (top) and *down-chirp* (bottom)

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)

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 $\Rightarrow$ bit-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

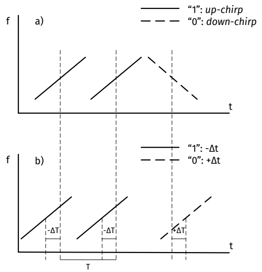


Figure 6: a) BOK using *up*- and *down*- chirps b) BOK using PPM

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

Outline

LoRa 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

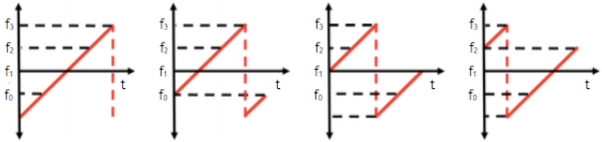
$$R_b = \frac{log_2(g_p)}{T} = log_2(g_p) \cdot \frac{B}{g_p}$$

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

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

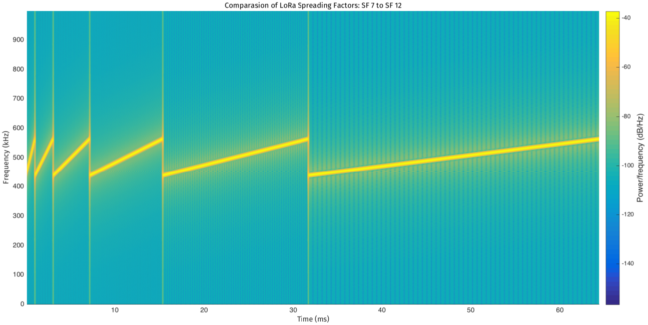
LoRa Symbols

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



LoRa Spreading Factors

- LoRa uses spreading factors from 7 to 12



LoRa Bit-Rate

- LoRa includes a variable error correction scheme that improves the robustness of the transmitted signal at the expense of redundancy
- Given a coding rate CR, the bit-rate is given by:

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

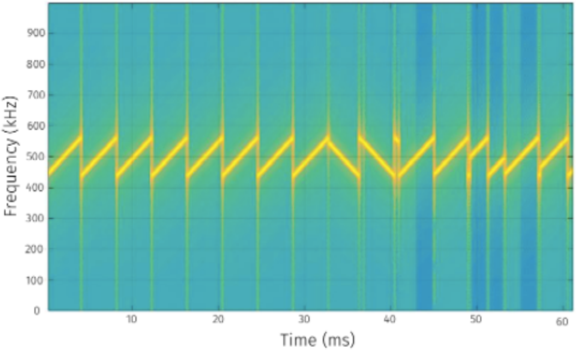
- R_b can also be written as:

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

with $1 \leq CR \leq 4$, and $7 \leq SF \leq 12$

LoRa Physical Layer

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



LoRa Characteristics

- Operates in license-free bands all around the world
  - 433, 868 (EU), 915 MHz
- Spectrum regulation for EU
  - Transmit power (EIRP) is limited to 14 dBm (25 mW)
  - 1% per sub-band duty-cycle limitation
- Receiver sensitivity: -142 dBm
- Link budget: 156 dB
- Uses spreading factors and channel coding rates to set the modulation rate

LoRa Radio Optimization

Spreading Factor	Bit Rate (kb/s)	Sensitivity (dBm)
6	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

(CR = 1 and B = 125 kHz)

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

Channels

- 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:

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

  - 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

From LoRa to LoRaWAN

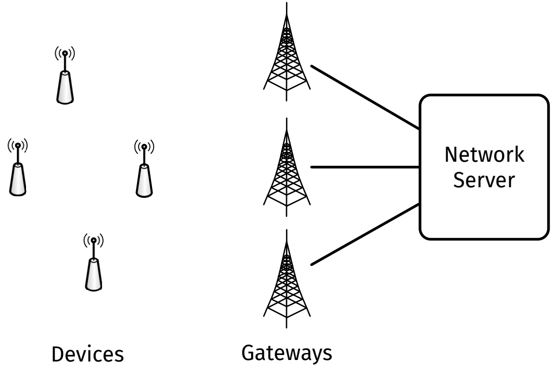
- LoRa
  - Modulation technique for LPWAN
- LoRaWAN
  - Uses LoRa modulation on physical layer
  - Proposes a MAC layer for access control
  - Specified by LoRa Alliance

LoRaWAN Timeline

- Cycleo first introduced LoRa in 2009
  - M2M communications
  - Large coverage
- Semtech acquired Cycleo in 2012 for 5 M$!
  - 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 specification in 2015

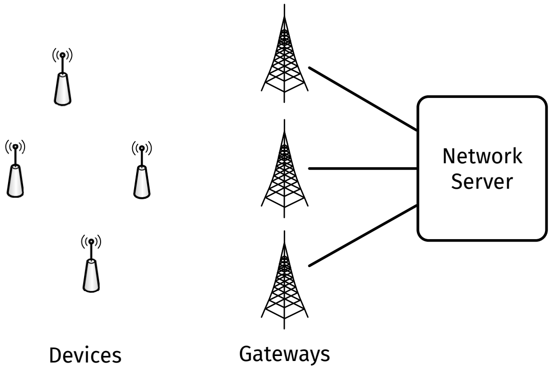
LoRaWAN Physical Architecture

End-Devices



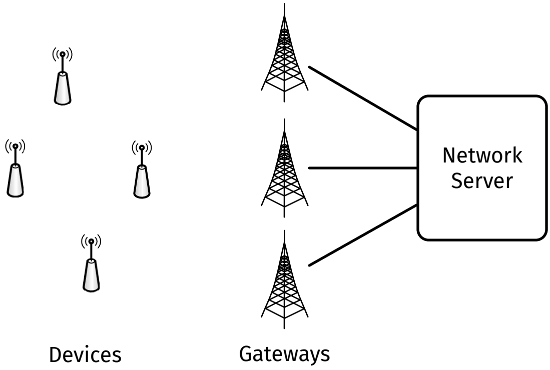
- End-devices are also called motes or devices
- Communicate to one or more gateways via a wireless interface using single hop LoRa or FSK

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 General Characteristics

- LoRaWAN network architecture is typically laid out in a star-of-stars topology
- All end-point communication is generally bi-directional
  - Uplink communications are predominant
- Data rates ranging from 300 bps to 5.5 kbps
  - Two high-speed channels at 11 kbps and 50 kbps (FSK modulation)
  - Eight channels: bandwidth 125 kHz or 250 kHz
  - Support for adaptive data rate (power and spreading factor control)
- Secure bi-directional communication, mobility, and localization
  - Device authentication, message encryption, and frame counter

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)

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

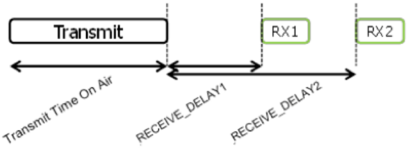
- Class A
  - Each uplink transmission is followed by two short downlink receive windows
  - Adapted for applications that only require downlink communication from the server shortly after the end-device has sent an uplink transmission
- Class B
  - In addition to class A, receive windows are opened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

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]

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



MAC Layer

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

- 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

Frame Counter

- Each device has two frame counters
  - Uplink frames, incremented by the device
  - Downlink frames, incremented by the NS

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

LoRa Radio Optimization

Spreading Factor	Bit Rate (kb/s)	Sensitivity (dBm)
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

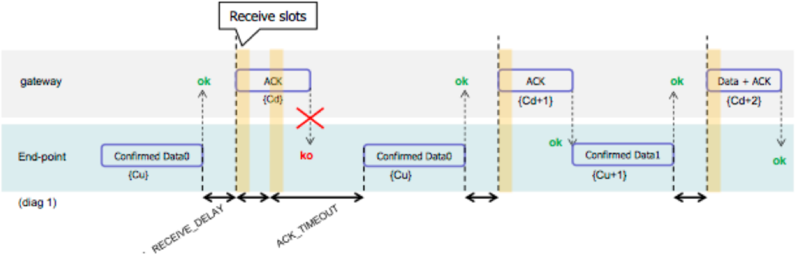
(RC = 1 and B = 125 kHz)

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

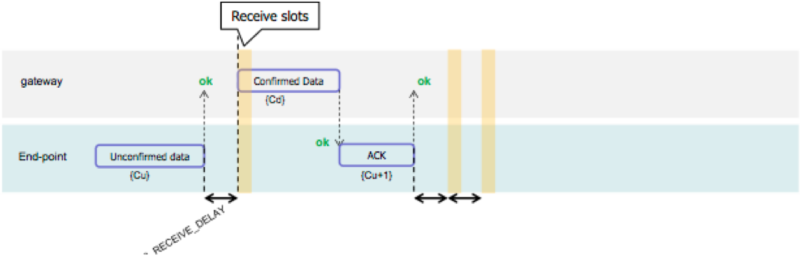
Adaptive Data Rate

- 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

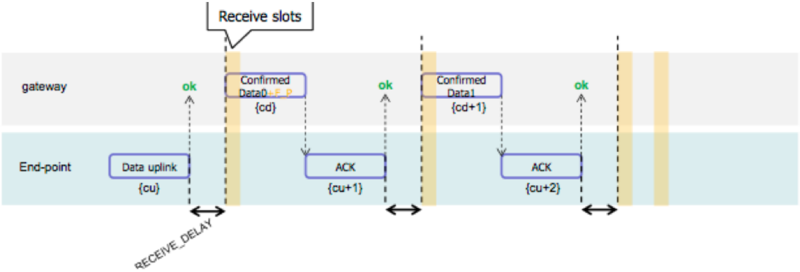
Wrap-up Example (1/3)



Wrap-up Example (2/3)



Wrap-up Example (3/3)



Security

Online Video Tutorial on Security

- https://www.youtube.com/watch?v=Nu_yZelDMZI

Outline

Link Budget

Enhanced Network Capacity

- LoRa employs orthogonal spreading factors which enables multiple spread signals to be transmitted at the same time and on the same channel
- Modulated signals at different spreading factors appear as noise to the target receiver
- The equivalent capacity of a single 125 kHz LoRa channel is:

$$SF12 + SF11 + SF10 + SF9 + SF8 + SF7 + SF6$$

$$= 293 + 537 + 976 + 1757 + 3125 + 5468 + 9375$$

$$= 21531 \text{ b/s} = 21.321 \text{ kb/s}$$

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 of a network wireless link can be expressed as:

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

where:

P_{Rx} = the expected received power

P_{Tx} = the transmitted power

G_{System} = system gains such as antenna gains

L_{System} = system losses such as feed-line losses

$L_{Channel}$ = losses due to the propagation channel

M = fading margin and protection margin

Outline

Test

Test