sigfox

Machine Learning in IoT

Olivier Isson

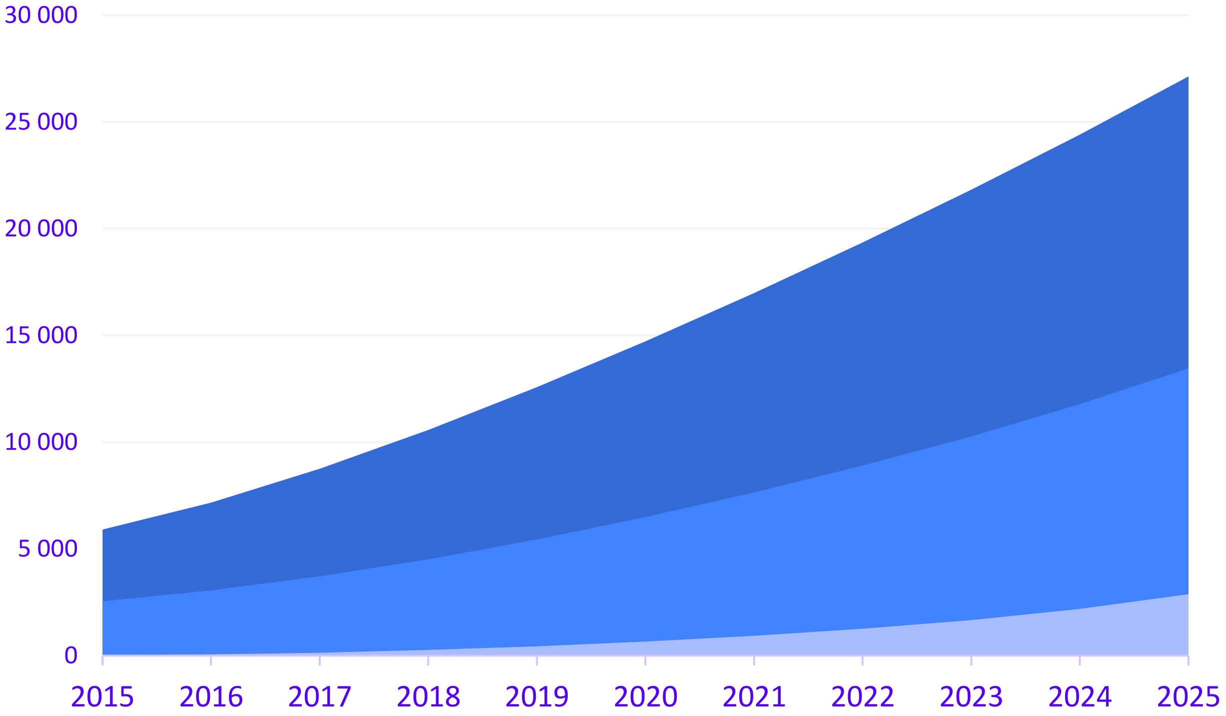
Data Science

Contents

1

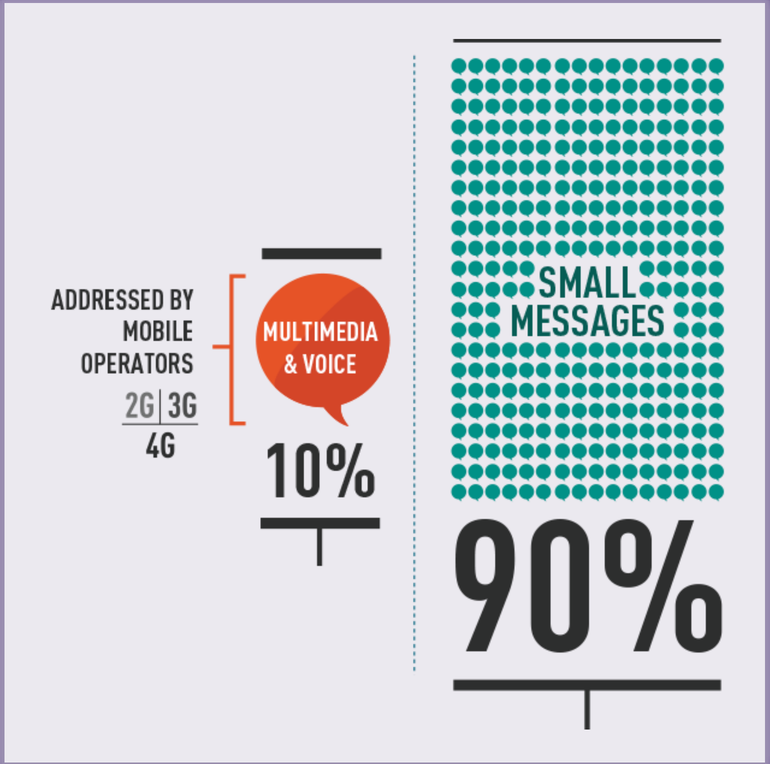
IoT - Sigfox

IoT market size worldwide



Source: Machina Research

The Internet of Things needs its Twitter!



IoT applications need

- ✓ small messages
- ✓ low cost
- ✓ low battery

2G, 3G and 4G are not optimized for IoT

sigfox

Small messages…

- ✓ **6 bytes**: GPS coordinates

  Location report with below 3m precision (GPS technical accuracy is above 3m)

- ✓ **2 bytes**: temperature reporting

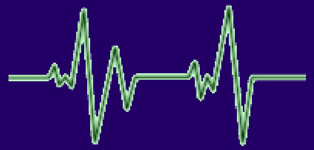
  Lab thermometer with -100°/+200° range, 0.004° precision

- ✓ **1 byte**: speed reporting

  Speed Radar up to 255km/h

- ✓ **1/8 byte**: object state reporting

  Switch report like set in day/night, hot/cold, on/off



- ✓ **0 byte**: heartbeat

  Object is in working state, battery is OK….

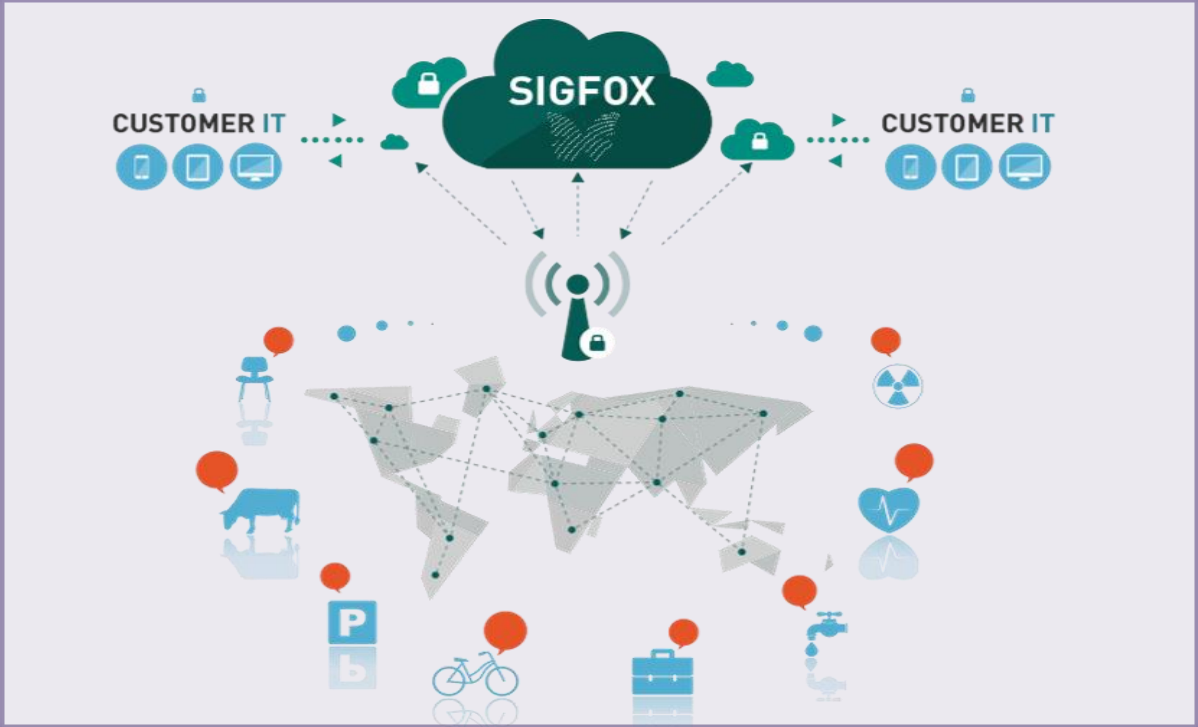
- ✓ **0 byte**: Request for duplex operation

  Do you have some information for me?

sigfox

0G - Low Power Wide Area Network

- Ultra low power
- Ultra long range
- Energy efficient
- Cost effective



sigfox

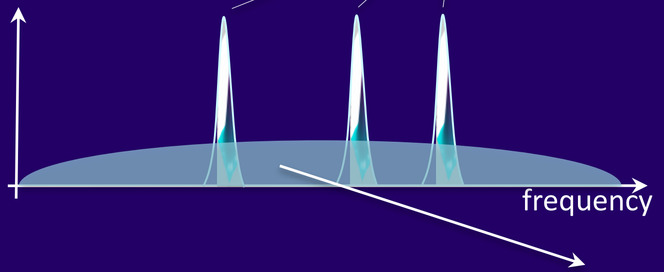
0G, how does it work?

Ultra **N**arrow **B**and signal

100-600 Hz per message

Space & **F**requency & **T**ime diversity

- ✓ High noise resilience
- ✓ Very low energy
- ✓ Long range capabilities



Use of free ISM band (868 MHz in Europe)

0G, how does it work?

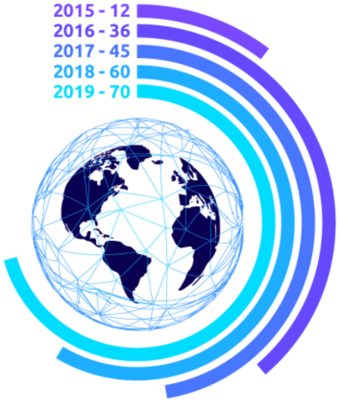
Very **Efficient** protocol

- ✓ Low redundancy to transfer a message
- ✓ No negotiation
- ✓ No Handover
- ✓ Bi-directional

12 bytes / message
Up to **140 messages** / day

A device can work up to **20 years** off two AA batteries

sigfox



Present in
70
Countries & Regions

2019 **5.3M km²**

2018 **3.6M km²**

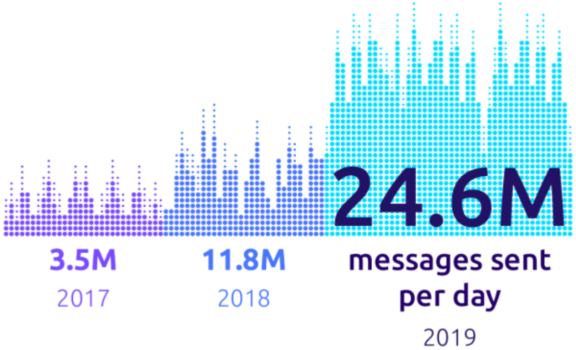
2017 **1.2M km²**

1.1 billion people & 5+ million km² covered

sigfox

January 2020

15.4M
Connected Devices

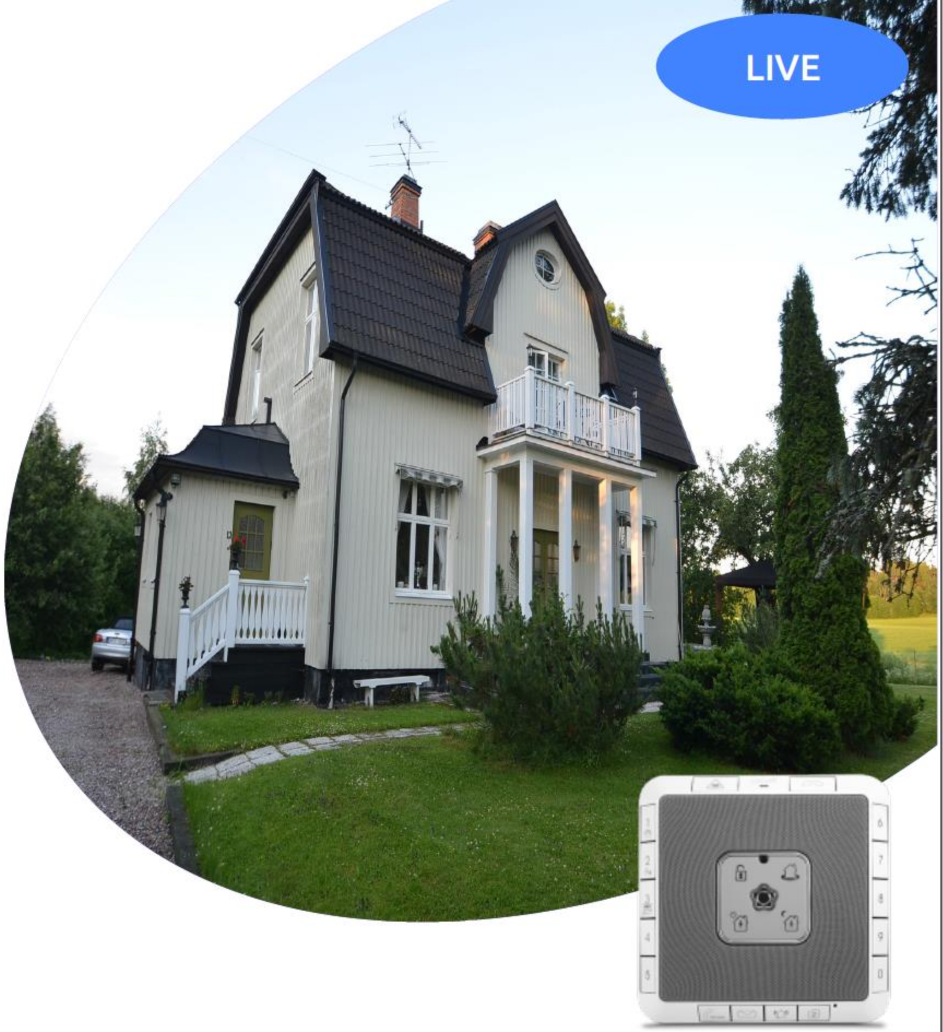



Visible stuff
Water metering
silver economy
Loiret
fastprk
Smart parking
Less visible stuff

Home Alarm System

Challenge

Secure better households



sigfox

Stolen Car Recovery

Challenge

Locate and recover stolen assets through a small and discrete GPS tracking device.

sigfox

free

2

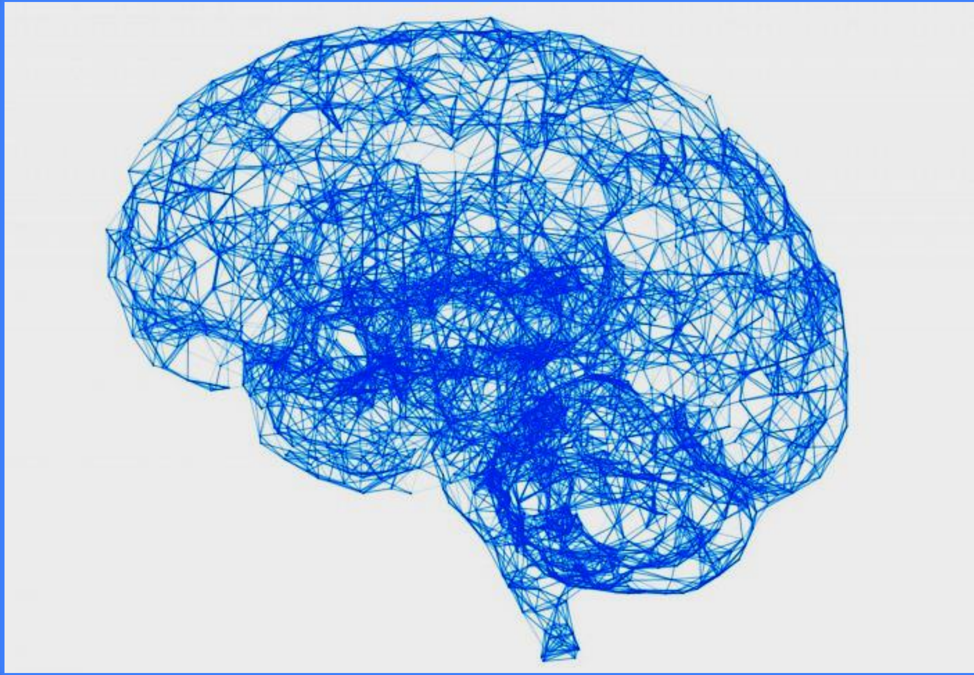
Machine Learning exemples

Data activities

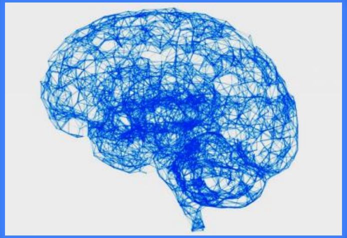
- Data Collection / storage
- Data knowledge / Analytics
- Data Intelligence / Machine learning

Data activities

- Data Intelligence / Machine learning



Data type

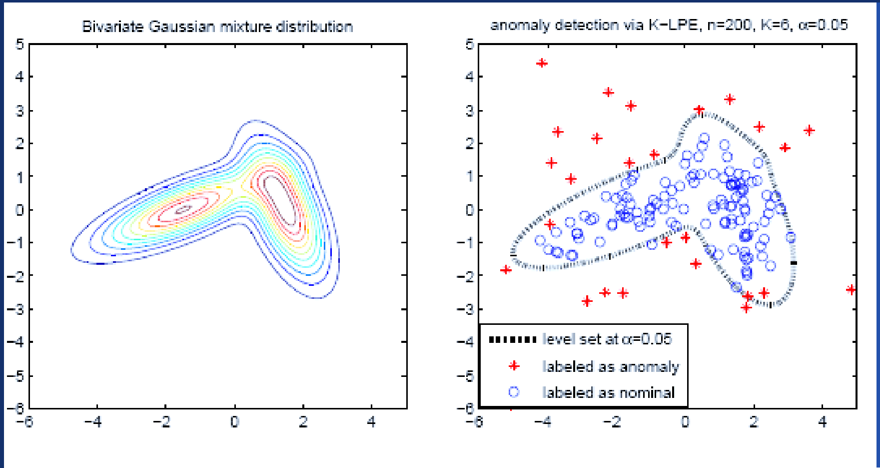


- Machine learning using data in the network
  - Payload data: the value of the transported data
    - Temperature, pressure, position,…
  - Metadata: all about data except its value
    - Reception date
    - Reception level
    - …

Data Intelligence / Machine learning

Anomaly detection /

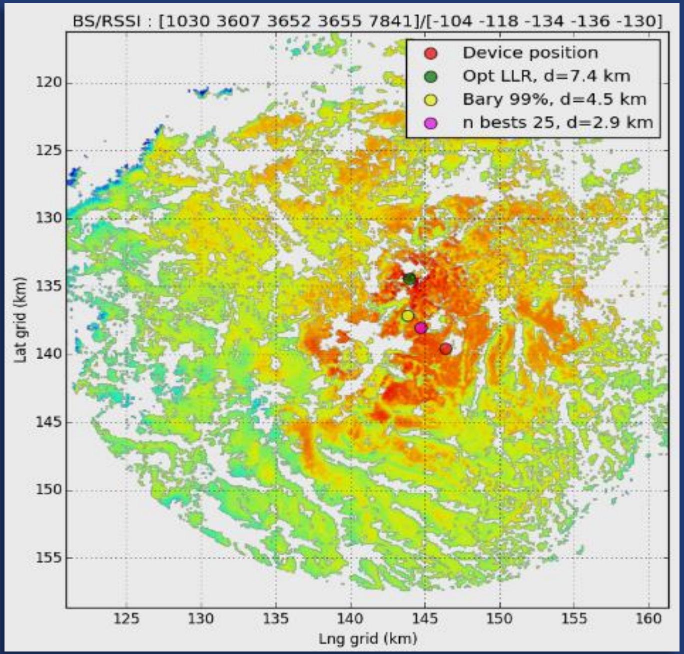
Predictive Maintenance



sigfox

Data Intelligence / Machine learning

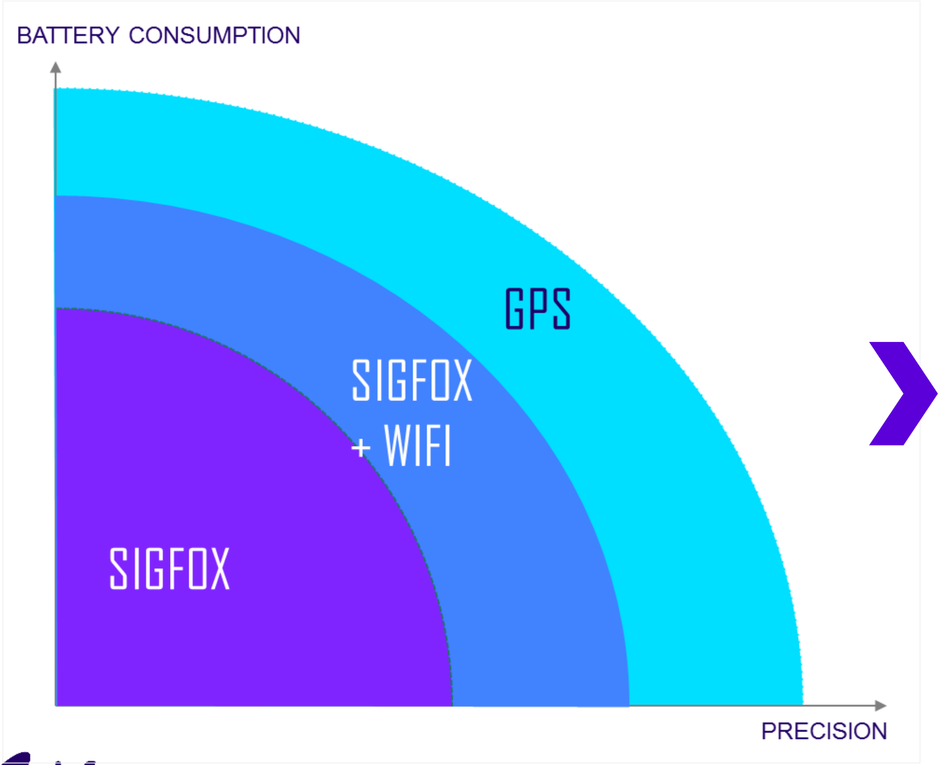
Geolocation



Geolocation business case: Louis Vuitton

DHL
DHL
DHL
NH

The Geolocation Challenge

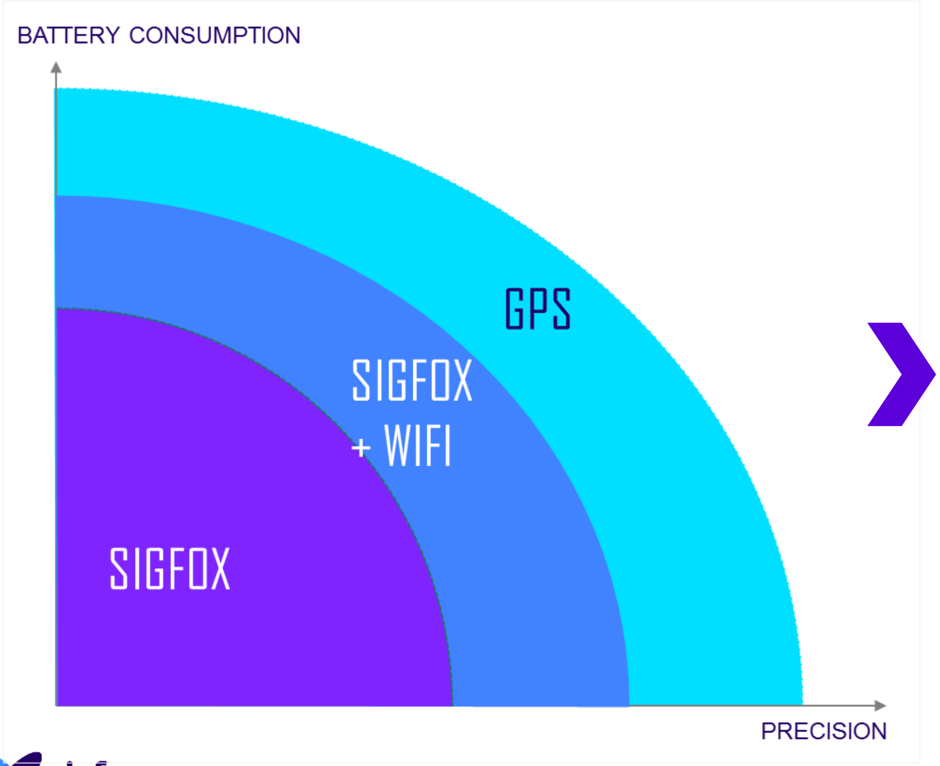


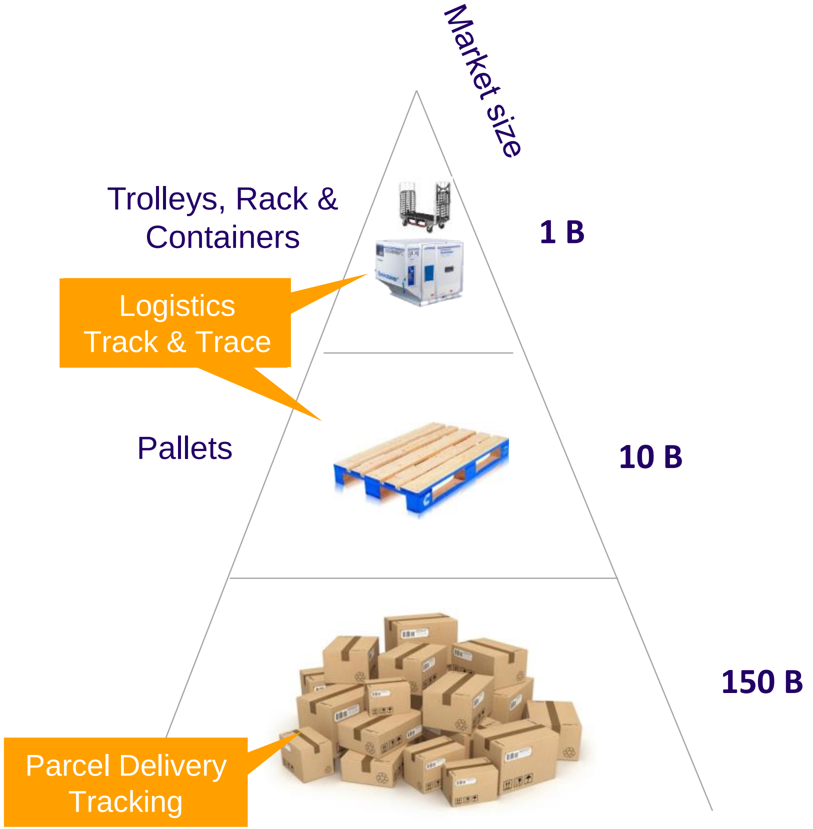
Need for a location solution without GPS:

- *Low cost module*
- *High battery life*
- *Precision target ~ 1km*

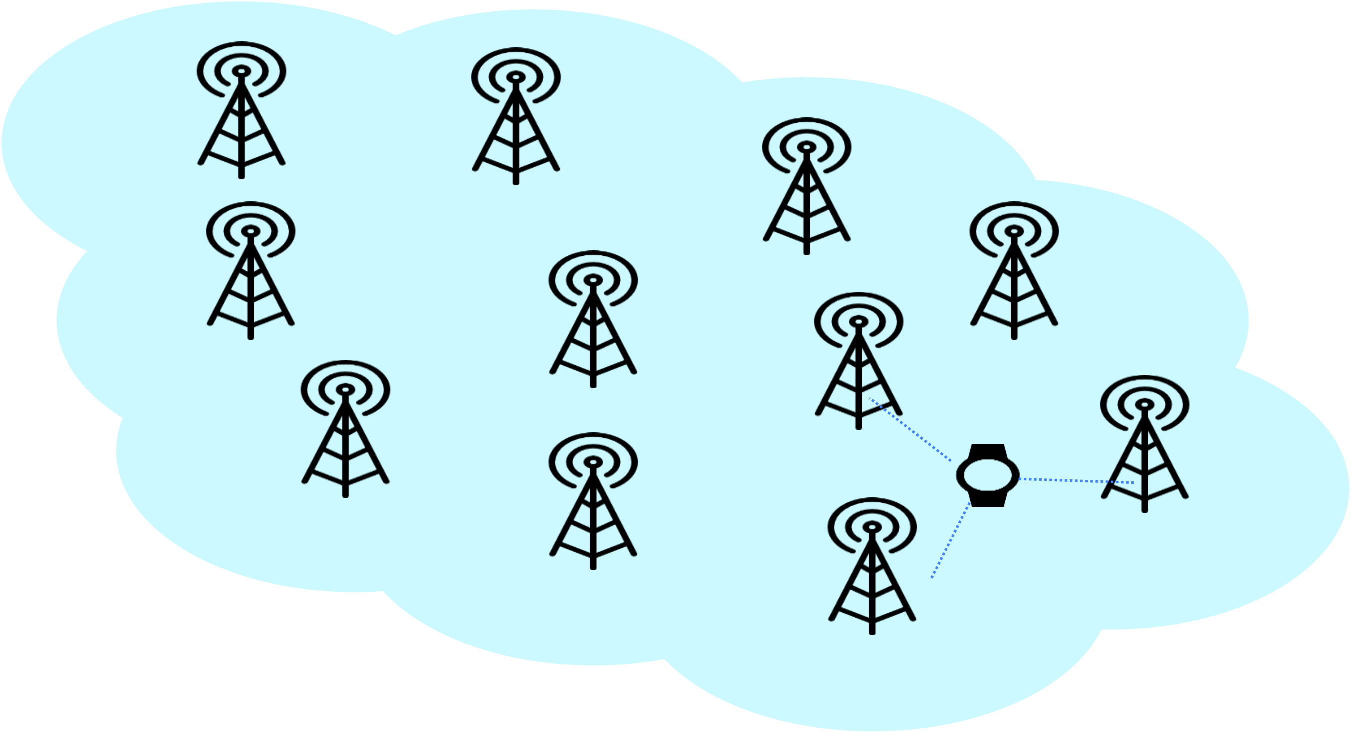
sigfox

The Geolocation Challenge

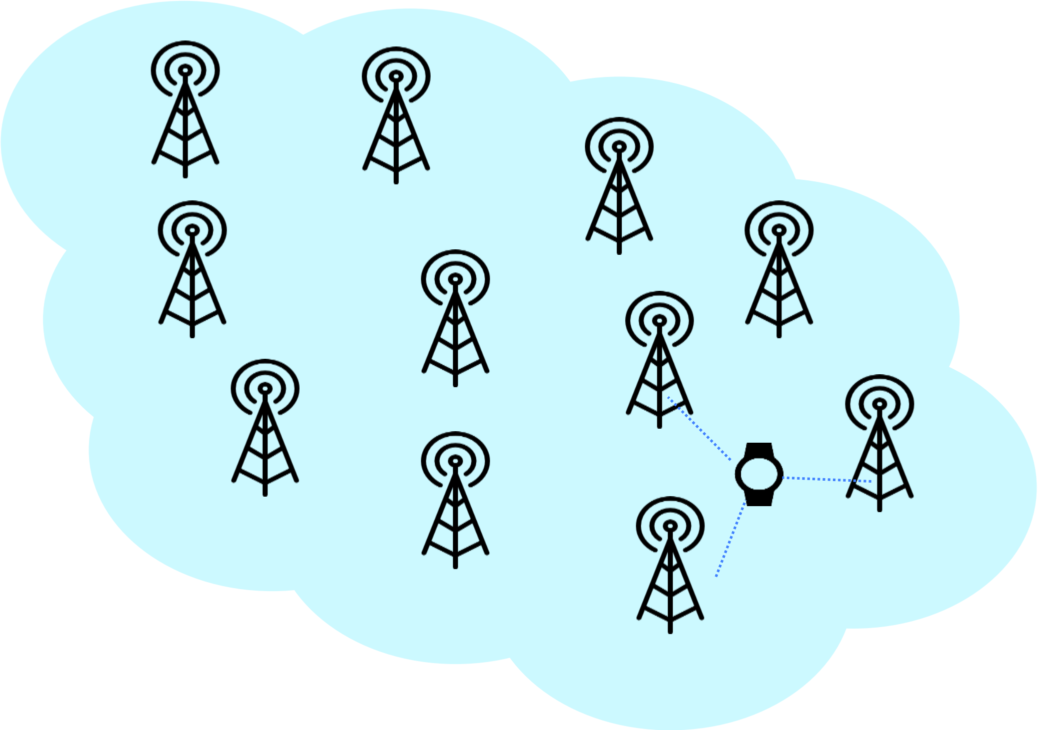




Network based geolocation



Geolocation state of the art – Time of flight

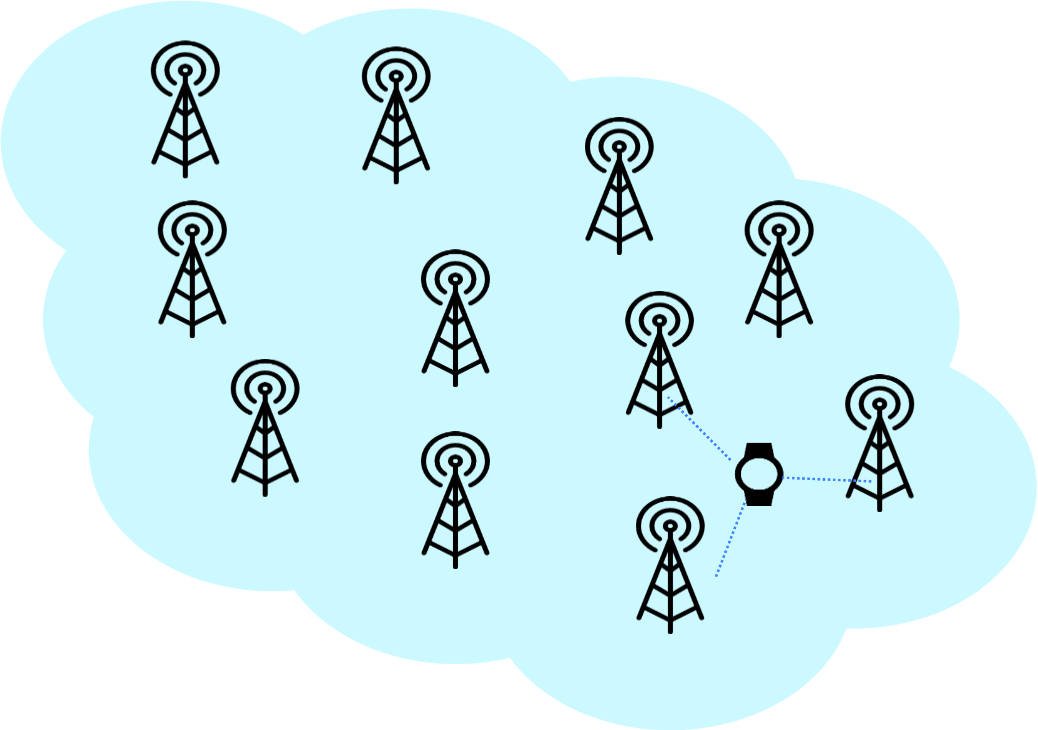


Calculate signal Time of Flight

- *Use of TDOA*
- *Estimate distances BS - Device*
- *Solve equation system: device area*

sigfox

Geolocation state of the art – Time of flight



Calculate signal Time of Flight

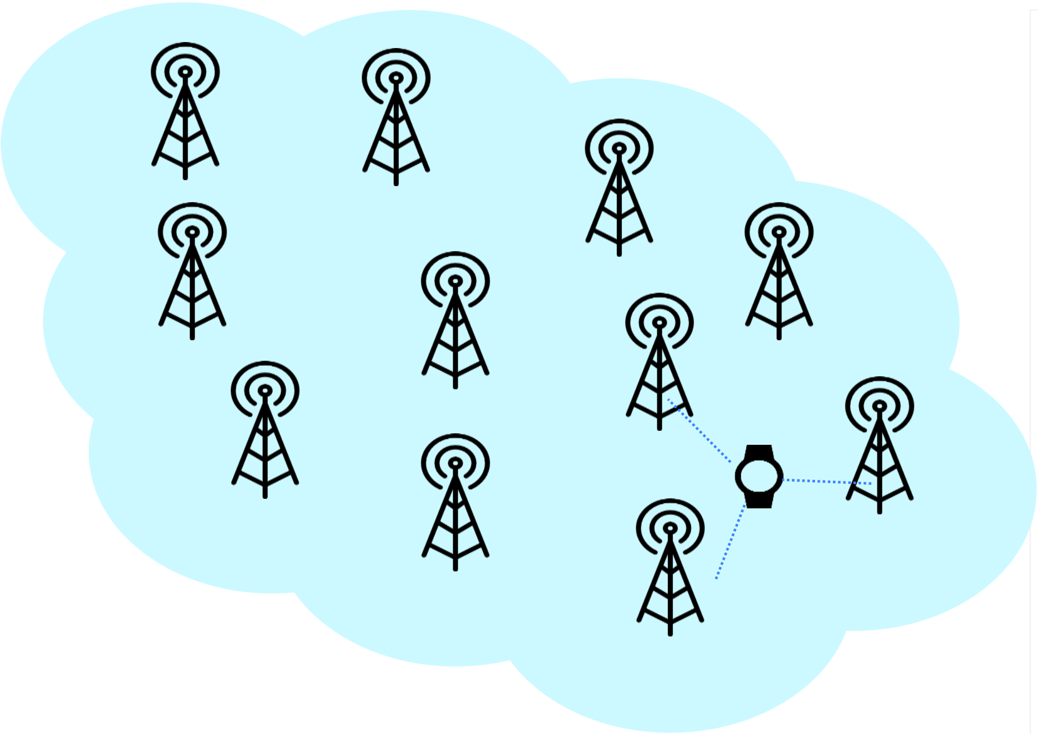
- *Use of TDOA*
- *Estimate distances BS - Device*
- *Solve equation system: device area*

Drawbacks for LPWAN

- *UNB not well suited for precise TDOA*
- *Need network synchronization time domain (~µs).*
- *Multipath channel destroy perfs*

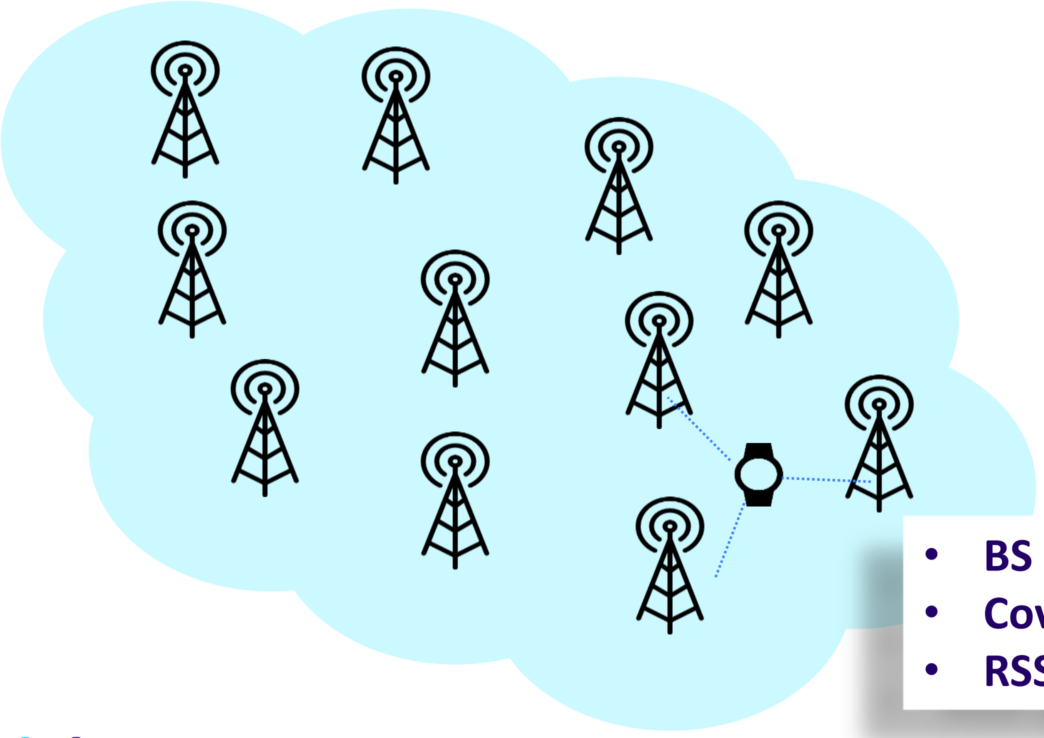
sigfox

Network based geolocation



- **Use of metadata to estimate a device position**
  - *Message reception → feature vector :* X

Network based geolocation– RSSI



Received Signal Strength Indicator

- *Use the received signal power*
- *Try to make a link between RSSI and device position*
- *No need for synchronized network*
- *Hard to link distance – RSSI in multipath environment*

- **BS Locations**
- **Coverage Map of each BS**
- **RSSI @ each BS**

Data oriented geolocation

- Sigfox received message → X a feature vector $X = (x_0,\dots,x_n)$
- x_i : Network information concerning the message
  - Base station position, received power, SNR,…

Data oriented geolocation

- Sigfox received message → X a feature vector $X = (x_0,\ldots,x_n)$
- x_i : Network information concerning the message
  - Base station position, received power, SNR,…
- Let Y be the position of the object (Lat, Lng)
- Then we want to make the link $X \rightarrow Y$

3

Lab: Geolocation

Geolocation Train Set

Input:

Message Id

Base station Id

Device Id

Base station position (Lat, Lng)

	messid	bsid	did	nseq	Rssi (dBm)	time_ux (ms)	bs_lat	bs_lng
0	573bf1d9864fce1a9af8c5c9	2841	473335	0.5	-121.5	1.463546e+12	39.617794	-104.954917
1	573bf1d9864fce1a9af8c5c9	3526	473335	2.0	-125.0	1.463546e+12	39.677251	-104.952721
2	573bf3533e952e19126b256a	2605	473335	1.0	-134.0	1.463547e+12	39.612745	-105.008827
3	573c0cd0f0fe6e735a699b93	2610	473953	2.0	-132.0	1.463553e+12	39.797969	-105.073460
4	573c0cd0f0fe6e735a699b93	3574	473953	1.0	-120.0	1.463553e+12	39.723151	-104.956216

Output: device position

	lat	lng
0	39.606690	-104.958490
1	39.606690	-104.958490
2	39.637741	-104.958554
3	39.730417	-104.968940
4	39.730417	-104.968940

sigfox

How to apply ML techniques to Geolocation

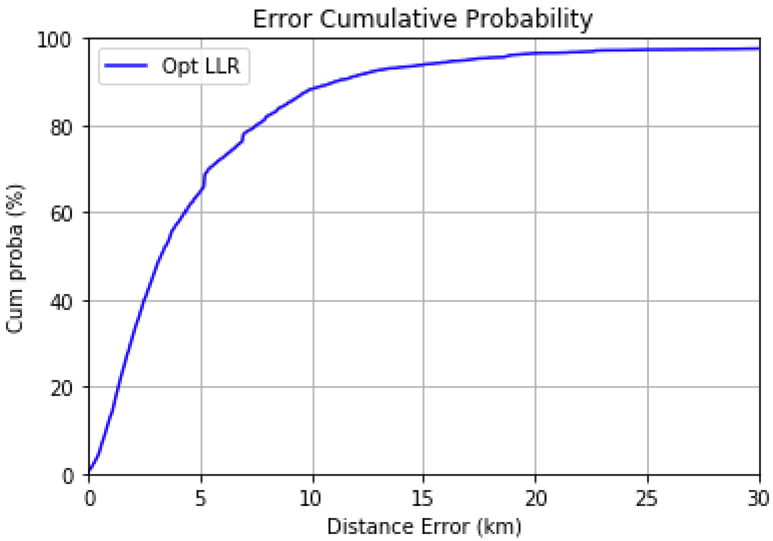
What kind of ML problems do we have?

What is the feature matrix / Ground truth ?

What kind of algorithm will we use ?

Goals

- Build feature matrix
- Build ground truth
- Plot error cumulative probability
- Compute prediction criterion: error @ 80%
- Extract prediction for the test set
  - Save result in csv file
- Cross validation: Build a « leave 1 device out » strategy



Error @ 80% = 7521 m

sigfox

Send me your results before 28/02/2020:
Olivier.Isson@gmail.com
Groups of 3-5 people

1. Python code used to generate previous goals
2. Predicted position for test set in csv format: **pred_pos_test_list.csv**
3. Short explanation of your approach and your choices: ~ 1-2 pages, can be included into the notebook or a separate document

Leave 1 device out predictor

- In practice we will not use the same devices to learn and to predict:
  - GPS devices to learn
  - Non GPS devices to predict

- To limit overfitting a « leave 1 device out » strategy must be used
  - To find the best model use cross-validation
    - Do not use the same device for the train and the validation sets

sigfox

Thank You