# Development of machine learning algorithms to improve performance in LoRaWAN and creation of a model for the analysis of experimental result This project was partially funded by the region Paris Ile-de-France Region via a grant from DIM RFSI (EPINE project). 

July 2021

## Contents

1 Introduction ..... 1
2 Specifications of LoRa technology ..... 2
3 Different ways to choose SF ..... 4
3.1 Calculation of minimal SF ..... 4
3.2 Estimate the rewards ..... 4
$3.3 \varepsilon$-Greddy ..... 5
3.4 Boltzmann exploration ..... 5
3.5 BaseSTEPS ..... 5
4 Description du processus de décision Markovien ..... 8
4.1 Finite state of states ..... 9
4.1.1 Transmit subset ..... 9
4.1.2 Waiting subset ..... 9
4.1.3 Initial state, success and failure ..... 10
4.2 Finite action state ..... 10
4.3 Probabilities of transition ..... 10
4.3.1 Choose of SF ..... 10
4.3.2 Probability of success at transmission ..... 11
4.3.3 Probabilities of not losing the packet ..... 11
4.3.4 Probability of no collisions ..... 13
4.4 Success probability ..... 16
4.5 Cost and Rewards ..... 16
5 Mise en place et résolution du MDP ..... 18
5.1 Nodes distribution ..... 18
5.2 MDP Resolve ..... 19
5.3 Data recovery and usage ..... 19
6 Simulation results ..... 21
6.1 Differents ways to initialize the SF table ..... 21
6.1.1 BaseSTEPS ..... 21
6.1.2 Proportionnal ..... 21
6.1.3 Ordder of appearance ..... 22
6.1.4 $50 \%$ premium ..... 22
6.1.5 $25 \%$ premium ..... 22
6.2 First Result ..... 22
7 Model Checking ..... 27
8 Conclusion ..... 29

## List of Figures

3.1 Collisions in greedy, Boltzmann and STEPS ..... 6
3.2 Number of ACK for greedy,Boltzmann and STEPS ..... 7
3.3 Ratio of ACK for greedy, Boltzmann and STEPS ..... 7
4.1 Loss of packet i if transmit of k in $T_{C}$ ..... 15
4.2 Model of MDP ..... 17
6.1 Number of collisions using the MDP with $\alpha=0.10$ ..... 23
6.2 Number of ACK using the MDP with $\alpha=0.10$ ..... 24
6.3 Ratio of ACK using the MDP with $\alpha=0.10$ ..... 25
6.4 Number of collisions using the MDP with $\alpha=0.20$ ..... 25
6.5 Number of ACK using the MDP with $\alpha=0.20$ ..... 26
6.6 Ratio of ACK using the MDP with $\alpha=0.20$ ..... 26
7.1 Probabilities to reach a success State ..... 28
7.2 Probabilities to reach success at the try k ..... 28

## List of Tables

3.1 Sensibility of recevier depending of SF and BW ..... 5
4.1 SNR by SF ..... 11
4.2 Loss path for different environment ..... 12
6.1 Average energy consumption in a day ..... 22
6.2 value given to a success using SF i ..... 23
7.1 Probabilities ..... 27

## Chapter 1

## Introduction

In view of the new importance of technology using connected smart device and the new policies to build smart cities, with the objective to create more ecological citites, we are facing a new technological challenge. Indeed it will be necessary to build a netword of devices whom constantly share data in a vast area and on a long term.

To solve this issue a solution already exist in the form of the internet of thing, indeed internet of thing technology such as LoRa, SIgfox and Weigthless have been created in the objective of wireless data exchange on long range. Another objective of this technology is to reduce the consumption of energy to save the battery of the device because one of the main point of using internet of thing network is to minimize the contribution of human input once the network is set.

During this internship we will be focusing on improving the performance of a LoRa network which we will first implement in a simulator before applying them to our experimental model.

## Chapter 2

## Specifications of LoRa technology

We therefore need to understand the specifications of LoRa technology, LoRa is an internet of thing technology using a star of star topology where each end devices is equiped with a LoRa chip to communicates with gateways whom then communicates with server. Although what interests us in particularity are the parameters used for the transmissions which are highly configurables directly at the level of the device.

The first of the modifiable parameter that impact performance in the device is the transmission power which can be adjusted from -4 dBm to 20 dBm in steps of 1 dBm , this modification is often limited by the equipment of the legislation of the region where the network is located.

The main modifiable parameter of LoRa technology is the spreading factor SF which is the ratio of the symbol rate (the original message to transmit) to the chip rate (the modulated message to transmitted).

The spreading factor can take 6 values between SF7 and SF12, their use presents a compromise, indeed the use of a small value increases the throughput which permit to reduce the transmission time and therefore the energy cost but when using this small value it leads to a reduction of the possible range, we are thus faced with a compromise between the energy efficiency of the transmission against the transmission range, by opposition high value of SF give a longer transmission range but have a high energy cost.

We have to tke into account a third parameters : the bandwidth which, in the case of LoRa technology, can take three different values $125 \mathrm{kHz}, 250 \mathrm{kHz}$ or 500 kHz . As for the use of spreading factor there is a link between the bandwidth, the speed data and the receiver sensibility, indeed when using a high value of banwidth the data speed will be better, which will reduce the airtime, but the sensibility will be reduced which once again lead to a reduction of the transmission range.

The fourth parameter is the carier frequency which can be included between 137 MHz and 1020 MHz , but must be include between 863 MHz and 870 MHz with at least the three channel which are $868.1,868.3$ and 868.5 MHz to be in accord of european legislation.

Once we know the parameters that can be chaqnged, we have to understand how changing them can upgrade the performance of our network. For that we have to look into the differents types of events that will make us lose our packet because two or more collide, in particular we need to look into SF collisions and the frequency collisions.

First we need to define the SF collision, during a transmission if several packets arrive at the base station at the same time, the use of different SF allows the reception of the two messages, but if the messages use the even SF there are then collisions which induces the loss of the two messages, it is thus in theory possible if one distributes the SFs on the
devices to reduce the risks of collisions due to an abundant use of one and the same SF.
The other collision that we have to take care of is the frequency collision, this collision can take place despite the use of different SF if several messages arrive at the same instant on the same channel (therefore using the same carrier frequency), it is It is therefore conceivable to standardize the use of the channels to reduce the risks of having several messages arriving on the same channel.

For this we first have to look for a method for the choice of the SF at each attempt by setting up at each node a table which for each of the SFs gives the probability, which changes over the transmissions, of choosing this last. In a second steps to ensure good performance, we must give good values whean creating the table of initial probabilities that would reduce the time necessary for learning to obtain optimal performance for this we decided to use a Markovian Decision Process.

## Chapter 3

## Different ways to choose SF

### 3.1 Calculation of minimal SF

The minimum SF represents the $S F \in\{7,8,9,10,11,12\}$ for which the receiver will be able to receive the transmission this SF depends on the distance between the node and the base station it is for We are interested in the reception power, which depends on the transmission power, and in the sensitivity of the receiver, which depends on the SF and the bandwidth because indeed the transmission can only be successfully received if the receiving power is greater than the sensitivity.

First, we must observe the reception power Prx which is given to us with the equation:

$$
P_{r x}=P_{t x}+G L-L p l
$$

where $P_{t x}$ is the transmission power which is configurable at the level of each transmitter and is theoretically adjustable from -4 dBm to 20 dBm in steps of 1 dB but is limited by hardware implementations and frequency regulation. $G L$ is the sum of the transmission and reception gains and losses and $L p l$ is the following path loss calculation obtained:

$$
L p l=L p l_{d 0}+10 n \log _{10}\left(\frac{d}{d_{0}}\right)
$$

where Lpld0 and $n$ are respectively the loss at the reference distance $d 0$ and the path loss coefficient and are obtained experimentally, $d$ is the distance between the node and the base station.

Now that we have seen how the reception power is obtained, we must check that it is much higher than the sensitivity of the receiver, the sensitivity depends on the chosen SF as well as on the bandwidth.

Using the LoRa SX1272 calculator [1] we obtain the following sensitivity table 3.1 (in dBm):

The minimal SF is therefore the smallest SF for which the reception power is greater than the sensitivity of the receiver, so we have for each node a set of available SFs S such that $S \in\left[S F_{\min }, 12\right]$

### 3.2 Estimate the rewards

Each node uses a table containing estimated rewards, Eri for each of the 6 available SFs, the table is initialized either from the database experiments or according to the distance from the node to the gateway if the database is empty.

|  | Bandwidth |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | 125 | 250 | 500 |
|  | 7 | -123 | -120 | -117 |
|  | 8 | -126 | -123 | -120 |
|  | 9 | -129 | -126 | -123 |
|  | 10 | -132 | -129 | -126 |
|  | 11 | -134.5 | -131.5 | -128.5 |
|  | 12 | -137 | -134 | -131 |

Table 3.1: Sensibility of recevier depending of SF and BW

On each transmission the node will receive a reward $R$ equal to 1 in the event of acknowledgment and to 0 in all other cases. After the transmission the Eri reward estimate for i chosen in the transmission will be updated according to the formula:

$$
E_{r_{i}}=E_{r_{i}}+\alpha\left(R-E_{r_{i}}\right)
$$

where $\alpha$ is the learning factor and $\alpha \in(0,1]$ can be choose by the user.

## $3.3 \quad \varepsilon$-Greddy

When selecting an SF for the transmission attempt, a $S F \in S$ will be chosen at random with a probability $\varepsilon$, otherwise the SF with the best $E_{r}$ will be chosen. We therefore have $P_{i}$ the probability of choosing the SF i such that:

$$
\mathbb{P}_{i}=\left\{\begin{array}{cc}
1-\varepsilon+\frac{\varepsilon}{K} & \text { if } i=\underset{\operatorname{argmax}^{\operatorname{argmax}} E_{r_{i}}}{\frac{\varepsilon}{K}}
\end{array}\right.
$$

where $\varepsilon \in(0,1]$ is choose by the user and K is the cardinal of S .

### 3.4 Boltzmann exploration

When selecting an SF for the transmission attempt, the SF will be chosen at random according to a probability Pi depending on the estimate $E_{r_{i}}$ of each SF as well as on a parameter $\tau \in \Re^{+}$such that:

$$
\mathbb{P}_{i}=\frac{e^{\frac{E_{r_{i}}}{\tau}}}{\sum_{k=S F_{m i n}-7}^{6} e^{\frac{E_{r_{k}}}{\tau}}}
$$

where $\tau$ is choose by the user, plus $\tau$ approaches infinity the more the algorithm behaves randomly and the closer $\tau$ is to 0 the more the algorithm behaves like a greedy algorithm. for all $S F<S F_{\text {min }}$ we have $P_{i}=0$.

### 3.5 BaseSTEPS

In the BaseSTEPS method each node has an array that represents the probability of choosing each of the SFs during a transmission attempt, for any $S F_{i}<S F_{\min }$ we have $P_{i}=0$.

At each transmission attempt, the probability of the SF used k is modified depending on the result of the transmission, but also on a random draw $\varepsilon$ which aims to increase the punishment of SFs having a low probability, after each attempt we have:

$$
\mathbb{P}_{k}=\left\{\begin{array}{cc}
\mathbb{P}_{k} *\left(1+3 e^{|k-\operatorname{minSF}|}\right) & \text { si } A C K \\
\mathbb{P}_{k} * 0.8 & \text { si } \varepsilon>P_{k} \\
\mathbb{P}_{k} *\left(0.9 * e^{\mid k-\operatorname{minSF|}}\right. & \text { sinon }
\end{array}\right.
$$

We then need to normalize the table so that the sum of the probabilities is equal to 1 .
We then set up simulations of 100 nodes to compare performance using the following set of parameters:

- Bandwidth $B W=125 \mathrm{kHz}$
- Transmit power $P i=14 \mathrm{dBm}$
- Loss Coefficient $\eta=2.32$
- Deviation $\sigma=7.8$
- Reference distance $d 0=1000 \mathrm{~m}$
- Traject loss at distance d0, $P L 0=128.95$
- Pauload $p l=10$ bytes
- Average transmission time $\tau=5 \mathrm{~s}$


Figure 3.1: Collisions in greedy, Boltzmann and STEPS
Although having more collisions and a slightly lower acknowledgment rate, we notice that the number of total packets acknowledged using the BaseSTEPs method is greatly greater than the other two methods over short distances, so we will favor this method when setting up the initial probability table using the predictions of the Markovian decision process.


Figure 3.2: Number of ACK for greedy,Boltzmann and STEPS


Figure 3.3: Ratio of ACK for greedy, Boltzmann and STEPS

## Chapter 4

## Description du processus de décision Markovien

In our objective to improve the transmission performance of nodes in a LoRaWAN network, we propose a method which allows to choose different SFs according to the SFs used in the previous attempts for the different transmission attempts in case of failure. The goal of changing SF during the various attempts is a priori to reduce collisions to allow a better acknowledgment ratio to be obtained and therefore better performance in the system in the long term.

In order to obtain a prediction for the choice of SF at each retransmission attempt, we propose to represent the transmission process by a Markovian decision process, which will then allow us to obtain by the iteration by value algorithm, which we will explain later, the optimal strategy, that is to say a prediction of the SFs to be chosen to minimize the cost of transmissions while maximizing the success of the transmission procedure.

A Markovian decision process is defined by a set of states, a set of actions, as well as for each combination of state and action the state transitions for the action as well as the cost associated with this transition. .

It is therefore possible to formally describe the Markovian decision process as a tuple $M=<S, A, P, R>$, such as:

- $S$ is a finite set of states ;
- A is a finite set of actions ;
- $\mathrm{P}: \mathrm{S} \times \mathrm{S} \times \mathrm{A}$ is the matrix of transitions between states S , which for each state $s$ inS and given an action $a \in A$ gives us the probability $P\left(s, a, s^{\prime}\right)$ to be at state $s^{\prime} \in S$, it is noted that for all and for all $\operatorname{ain} A$ available in state s we have $\operatorname{sum}_{s^{\prime}=0}^{S} \mathbb{P}\left(s, a, s^{\prime}\right)=$ 1 such that the sum of the probabilities of transitions of action a in state $s$ is equal to 1 ;
- $\mathrm{R}: \mathrm{S} \times \mathrm{S} \times \mathrm{A}$ is the reward matrice which given each state $s \in S$ and and action $a \in A$ give the reward $R\left(s, a, s^{\prime}\right)$ which is the expected reward when we reach the state $s^{\prime} \in S$ from the state s using action a.

In the following sub-parts we will focus on the definition as well as on the construction of the different elements of M .

### 4.1 Finite state of states

The goal of the Markovian decision process being to represent the mechanism of choice of SF and transmission by devices using LoRaWAN technology, we must therefore separate the states into two distinct groups.

The first state group will represent the states during which the devices transmit while the second group will represent the standby states, that is to say the states following a transmission failure and during which the device will choose the SF to use on the next attempt.

### 4.1.1 Transmit subset

We will denote by T the finite state subset representing the states representing transmission attempts such as $T \subset S$, the technology on which we are working being the LoRaWAN technology each transmission has the right to a maximum of 8 attempts before d 'be abandoned, so we can denote by k the current transmission attempt such that $1 \leq k \leq 8$.

The long-term goal being to change our choice of SF according to previous transmission failures we will therefore have for each transmission state $s \in T$ at the transmission attempt k the information of the SF used for the current attempt but also the number of times that each SF will have been used during the attempts preceding the attempt k .

Each state of the subset T will therefore allow the system to know the spreading factor for the current transmission attempt as well as to allow the analysis of past failures to know a priori the best SF to use.

The state for the transmission attempt $\mathrm{k}=3$ using SF10 and having used 7 and 12 during the second attempt during the first attempt will be represented for example by "T $/ 7,12,10$ " or even by "T / 10,11,9,7"the state for the transmission attempt $k=4$ using SF7 and having used SFs 11, 9 and 7 respectively during attempts 1, 2 and 3.

### 4.1.2 Waiting subset

We will denote by W the finite state subset representing the waiting times between two transmission attempts such as $W$ subset $S$, waiting time between 2 transmission attempts during which the device will decide which SF to use for transmission $\mathrm{k}+1$ with k the last transmission attempt made and having failed where 1 leqk leq 7.

The goal is always to choose the best theoretical SF from the point of view of the apparatus, each waiting state $S \in W$ will have the information of the SF used during the attempt k but also of the number of times each SF will have been used during the attempts preceding the attempt k .

We will therefore have for each state $s \in W$ the knowledge of the SF of the last attempt made k as well as of all the previous attempts at k which will make it possible to know the previous failures in order to make a better choice of SF during transmission $\mathrm{k}+1$. In the same way as for the states of the transmission sub-space, the states of the waiting sub-space will be represented by the SF of the attempt k as well as all the spreading factors used during the attempts previous ones.

We will therefore denote for example by "W / 7,12,10" the waiting state after a failure during the attempt 3 using the SF 10 following the first two attempts using respectively the spreading factors 7 and 12 . .

Likewise, the notation "W9,8,11,12" represents the state of waiting after a failure during the attempt number 4 using the SF 12 and having used the factor 9 during the first attempt, the factor 8 during the the second attempt and the factor 11 on the third try.

### 4.1.3 Initial state, success and failure

After having defined the two subsets of states T and W , we must now look at the last three states belonging to the finite set of state S :

- L'état $S_{0}$ which represents the state before the first transmission attempt in which the device will choose the SF to use for the first attempt;
- The "Success" state which is one of the two terminal states that will be reached if the transmission is successful during one of the attempts;
- The "Failure" state which is the second terminal state and which will be reached if all 8 transmission attempts are unsuccessful.


### 4.2 Finite action state

To best represent the transmission mechanism in a LoRaWAN technology device, we have decided to represent the set of actions by
$A=\{T r, a 7, a 8, a 9, a 10, a 11, a 12\}$, these actions can be subdivided into two categories and are only available depending on the current state.

First, we have the first category made up of the action Tr , which is the action representing a transmission attempt, for this the action Tr will be the only action available in the states of the subspaces T which represents the states during which we carry out transmission attempts.

The second category is made up of the remaining actions (a7, a8, a9, a10, a11, a12) each of these actions represents the choice of a certain SF by the device (ie a7 represents the selection of $\mathrm{SF}^{\circ}{ }^{\circ}$ ) these actions will therefore be the actions available in the waiting states of the sub-spaces W as well as during the initial state $S_{0}$.

### 4.3 Probabilities of transition

Just like the action space, we must subdivide the probabilities of transitions into two groups, these two groups being the same as for the action space with on one side the probabilities of transitions during transmissions and 1 the other is the probabilities of transitions when choosing the spreading factors.

We recall that the probability of reaching state $s$ from state $s$ as a function of the action a is denoted by $\mathbb{P}\left(s, a, s^{\prime}\right)$.

### 4.3.1 Choose of SF

When choosing SF, which is an action $a \in\{a 7, a 8, a 9, a 10, a 11, a 12\}$, in a wait state $s \in W$ or in the initial state $S_{0}$ the probability of transitions between state s and associated state $s$ ' is equal to 1 .

We must therefore define what is the associated state s ', we recall that each state carries the information of the SF used for the current attempt but also of all the SFs used for the previous attempts, a state s' associated with a state s by the action a is therefore the transmission state having the same SF for the previous attempts and using for the new attempt the SF of the action a ie the transmission state $\mathrm{s}^{\prime}=\mathrm{T} / 7,10$ is a state associated with the waiting state $\mathrm{s}=\mathrm{W} / 7$ by the action $\mathrm{a}=\mathrm{a} 10$ and we have:

$$
\mathbb{P}\left(s, a, s^{\prime}\right)=1
$$

### 4.3.2 Probability of success at transmission

We have therefore just seen that the probability of transition in the case of a choice of SF is easy to define, it is not the same for the probability of successful transmission.

In fact, during a transmission in the absence of acknowledgment, the node can only assume that the attempt to transmit the message was successful, so we will theoretically obtain the probability of success of a transmission in order to know the probabilities of transitions between a state of transmission and the state of success (and therefore the probability of transitions to the associated wait state or to the state of failure).

When attempting to transmit there are two important conditions to be met [2] :

- The message send by the node is not lost
- The message send by the node did not enter in collisions

We will therefore see the models and the prediction functions which make it possible to verify these two conditions.

### 4.3.3 Probabilities of not losing the packet

So that the transmitted message is not lost, it is sufficient that the signal to noise ratio (SNR) received is greater than a threshold specific to the SF used this factor is noted qSF and is available in the table 4.1 [3].

| SF | $q_{s} f(\mathrm{~dB})$ |
| :---: | :---: |
| 7 | -7.5 |
| 8 | -10 |
| 9 | -12.5 |
| 10 | -15 |
| 11 | -17.5 |
| 12 | -20 |

Table 4.1: SNR by SF

We thus have a probability $H_{i}(s f)$ that the message sent by our node i at the distance $d_{i}$ using the SF is not lost which can be expressed by the equation (4.1) [4]:

$$
\begin{equation*}
H_{i}(S F)=\mathbb{P}\left[S N R \geq q_{S} F \mid d_{i}\right] \tag{4.1}
\end{equation*}
$$

[2], [4] et [5] give the theoretical calculation of the formula (4.1),thus assuming that we have $h_{i} \sim \exp (1)$, the probability of $H_{i}(S F)$ that the message sent by our node i is not lost can thus be expressed as:

$$
\begin{equation*}
H_{i}\left(S F, d_{i}\right)=\exp \left(-\frac{N q_{s f}}{\mathbb{P}_{i} g\left(d_{i}\right)}\right) \tag{4.2}
\end{equation*}
$$

Where we have $P_{i}$ the transmit power of device i in milliwatts, where $\left|h_{i}\right|^{2}$ is the gain of the channel which is modeled as an exponential random variable with an average of 1 , $N=-174+N F+10 \log B W \mathrm{dBm}$ with NF the receiver noise figure which is equal to 6 dB for our device, $-174 \mathrm{dBm} / \mathrm{Hz}$ is the noise spectral density and where $g\left(d_{i}\right)$ is the function of path loss attenuation.

The path loss attenuation function $g\left(d_{i}\right)$ is derived from the Friis transmission equation which can be written as:

$$
P_{r}=\frac{P_{T} G_{T} G_{R} c^{2}}{(4 \pi f)^{2}} *\left(\frac{1}{d}\right)^{\eta}
$$

where $P_{r}$ is the received signal power, $P_{T}$ is the transmitted signal power, $G_{T}$ and $G_{R}$ are respectively the gains of the transmit and receive antennas, c is the speed of light, $f$ is the transmission frequency, $d$ is the distance between the transmitting antenna and the receiving antenna and finally $\eta$ is the path loss exponent which depends on the environment and can be given by the table4.2 [6]


Table 4.2: Loss path for different environment

After modifying the Friis equation, by considering the transmission and reception gains to be equal to 1 , the authors of the articles [2] and [4] obtain $g\left(d_{i}\right)=\lambda /\left(4 \pi d_{i}\right)^{\eta}$ for their part the authors of [6] obtain as a result, in addition to having two different equations these present a dimension inconsistency.

To be able to perform our calculation of the probability of not having a loss $H_{i}$ we therefore had to find another model, this is how we found the logarithmic path loss model which is used in [7], [8], [4] and [9] and which calculates the SNR with the formula (4.3).

$$
\begin{equation*}
S N R=P_{r x}-N=P_{t x}-\left(P L_{0}+10 \eta \log _{10}\left(\frac{d_{i}}{d_{0}}\right)+X_{\sigma}\right)-N \tag{4.3}
\end{equation*}
$$

Where $P_{t x}$ and $P_{r x}$ are the transmit power and receive power in dBm respectively, $P L_{0}$ is the path loss at the reference distance $d_{0}, \eta$ is the loss exponent of path and $X_{\text {sigma }}$ is the decibel attenuation which is modeled by a Gaussian random variable $X_{\sigma} \sim N\left(\mu, \sigma^{2}\right)$ with $\mu=0$. Thanks to the formula (4.3) it is therefore possible to write the probability
$H_{i}$ as such:

$$
\begin{aligned}
H_{i}\left(S F, d_{i}\right) & =\mathbb{P}\left[S N R \geq q_{S F} \mid d_{i}\right] \\
& =\mathbb{P}\left[\left.P_{t x}-\left(P L_{0}+10 \eta \log _{10}\left(\frac{d_{i}}{d_{0}}\right)+X_{\sigma}\right)-N \geq q_{S F} \right\rvert\, d_{i}\right] \\
& =\mathbb{P}\left[\left.X_{\sigma} \geq q_{S F}+N-P_{t x}+P L_{0}+10 \eta \log _{10}\left(\frac{d_{i}}{d_{0}}\right) \right\rvert\, d_{i}\right]
\end{aligned}
$$

With the function of repartion: $F_{x_{\sigma}}(x)=\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{x}{\sigma \sqrt{2}}\right)\right)$, the probability $H_{i}$ can be noted as (4.4).

$$
\begin{equation*}
H_{i}\left(S F, d_{i}\right)=1-\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{q_{S F}+N-P_{t x}+P L_{0}+10 \sigma \log _{10}\left(\frac{d_{i}}{d_{0}}\right)}{\sigma \sqrt{2}}\right)\right) \tag{4.4}
\end{equation*}
$$

The value of $P L_{0}, \eta$ et $\sigma$ can be calculated empirically by performing experiments as well in [5] and [7] the results of the measurement experiments give the following values such as $P L_{0}=129.95 d B, \eta=2.32$ and $\sigma=7.8 d B$ with distance $d_{0}=1000 \mathrm{~m}$ and $P_{t} x=14 d B m$. It is important to note that the calculation of the probability $H_{i}$ of nonloss of the message does not take into account the distribution of nodes in space and that is why in the following section we will take it into account.

### 4.3.4 Probability of no collisions

During a communication in LoRaWAN technology it is possible for a collision to occur between several terminal devices, so that the collision between the messages of the node i studied and of the node $k$ induces the loss of the message $i$ there. some conditions must be satisfied [4]:

1. Frequency $_{i}=$ Frequency $_{k}$
2. $S F_{i}=S F_{k}$
3. $\frac{P_{r x_{i}}}{P_{r x_{k}}}<4 \mathrm{~mW} \Leftrightarrow P_{r x_{i}}-P_{r x_{k}}<6 \mathrm{~dB}$ where $P_{r x}$ is the receive power, this condition is called the capture effect
4. $t_{\text {pream }_{i}}-t_{\text {fin }_{k}}<t_{c}$ où $t_{\text {pream }_{i}}$ is the time that the receiver takes to finish processing the preamble of packet i, $t_{\text {end }}{ }_{k}$ is the time that the receiver needs to finish processing packet k and $t_{c}$ is the minimum time that is often defined as the duration of the receiver to process 5 preamble symbols [4]

To be able to check these conditions we must therefore have a distribution of the nodes, when creating the MDP the only information we will know about the nodes is their distance from the base.

During the creation of the model, knowing only the distance compared to the base, it will be necessary to attribute to each node values of frequency and SF , for the value of frequency this one is given in a random way and is included in 868.1,868.3,868.5 which are the 3 frequencies in MHz implemented by all LoRa devices for the European region, the frequency used by the node that we are studying is on its side chosen and not random.

The allocation of SFs for each device is done from the calculation of Hi, for each node the probability of no loss $H_{i}$ will be calculated for each of the SFs, the SF that will be assigned to the node will be the smallest SF for which $H_{i}(S F) \geq$ threshold where the threshold is a value that we will set during resolutions.

Not being able to know the transmission time of the other nodes, our device can only assume that it will compete on each transmission attempt, knowing this and knowing the conditions if we consider them independent we can denote by $Q_{i}(k)$ the probability that the message from node $i$ is not lost as a result of the collision with the message from node k as in (4.5)

$$
\begin{equation*}
Q_{i}(k)=1-\left(X_{S F_{i}}\left(S F_{k}\right) X_{f r e q_{i}}\left(\text { freq }_{k}\right) P(\text { condition } 3) P(\text { condition } 4)\right) \tag{4.5}
\end{equation*}
$$

where $X_{a}(x)=1$ if $\mathrm{a}=\mathrm{x}, 0$ otherwise.
The probability $\mathbb{P}($ condition 3$)$, which is the probability of gaining the capture effect can be obtained as proposed in [4] :

$$
\begin{equation*}
\mathbb{P}(\text { Condition } 3)=\mathbb{P}\left[\left.\frac{\left|h_{i}\right|^{2} g\left(d_{i}\right)}{\left|h_{k^{*}}\right|^{2} g\left(d_{k^{*}}\right.} \geq 4 \right\rvert\, d_{i}\right] \tag{4.6}
\end{equation*}
$$

Where $k^{*}=\operatorname{argmax}_{k \neq i}\left(P_{k} X k^{S F}\left|h_{k}\right|^{2} g\left(d_{k}\right)\right)$ is the node with the strongest received signal except node i which we are studying, for the SF currently in use. This probability therefore depends on the distribution of the nodes that we have seen.

We can obtain the analytic value $\mathbb{P}($ condition 3$)$ which can be found in [4] :

$$
\begin{equation*}
\mathbb{P}(\text { Condition } 3)=\mathbb{E}_{\left|h_{i}\right|^{2}}\left[F_{X_{k^{*}}}\left(\frac{\left|h_{i}\right|^{2} g\left(d_{i}\right)}{4}\right)\right]=\int_{0}^{\infty} e^{-z} F_{X_{k^{*}}}\left(\frac{z g\left(d_{i}\right)}{4}\right) d z \tag{4.7}
\end{equation*}
$$

This probability is an analytical formula for a random number of nodes distributed at random locations. It is therefore possible for us, knowing the number of nodes as well as their distances to the base of transformed the formula 4.7 by the following formula:

$$
\begin{equation*}
\mathbb{P}(\text { Condition } 3)=\int_{0}^{\infty} e^{-z} F_{X_{k^{*}}}\left(\frac{z g\left(d_{i}\right)}{4}\right) d z=\prod_{k \neq i, S F_{k}=S F_{i}}\left\{\int_{0}^{\infty} e^{-z} F_{X_{k}}\left(\frac{z g\left(d_{i}\right)}{4}\right) d z\right\} \tag{4.8}
\end{equation*}
$$

Where $F_{X_{k}}(z)$ is the cumulative distribution function(CDF) of $X_{k}$ and $X_{k}=\left|h_{k}\right|^{2} g\left(d_{k}\right)$.
The equation (4.8) allow us to obtain $\mathbb{P}$ (condition 3) using the formulas in mW but it is also possible to obtain it by using the formulas of dBm :

$$
\begin{align*}
\mathbb{P}(\text { Condition } 3) & =\mathbb{P}\left(P_{r x_{i}}-P_{r x_{k}}<6\right) \\
& =\mathbb{P}\left(-10 \eta \log _{10}\left(\frac{d_{i}}{d_{0}}\right)-X_{\sigma_{i}}+10 \eta \log _{10}\left(\frac{d_{k}}{d_{0}}\right)+X_{\sigma_{k}}<6\right) \\
& =\mathbb{P}\left(10 \eta \log _{10}\left(\frac{d_{k}}{d_{i}}\right)-X_{\sigma_{i}}+X_{\sigma_{k}}<6\right) \tag{4.9}
\end{align*}
$$

To obtain the analytical value of (4.9) we must use the following formula:

$$
\begin{align*}
\mathbb{P}(\text { condition } 3) & =F_{X_{\sigma_{k}}-X_{\sigma_{i}}}\left(6+10 \eta \log _{10}\left(\frac{d_{i}}{d_{k}}\right)\right) \\
=F_{X_{\sqrt{2 \sigma}}}\left(6+10 \eta \log _{10}\left(\frac{d_{i}}{d_{k}}\right)\right) & =\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{6+10 \eta \log _{10}\left(\frac{d_{i}}{d_{k}}\right)}{2 \sigma}\right)\right) \tag{4.10}
\end{align*}
$$

Where $F_{X_{\sqrt{2 \sigma}}}(z)$ if the cumulative distribution function of $X_{\sqrt{2 \sigma}} \sim N(0, \sqrt{2 \sigma})$.
The equation (4.10) therefore makes it possible to know the probability of gaining the capture effect for node i facing another node $k$, we must now proceed to the analysis of the condition 4, which allows to know if another message is already processed and if our message coming from node i can be processed by the database or if it is lost due to a temporal collision. To calculate the probability $\mathbb{P}($ condition 4$)$, we must assume that the nodes transmit according to an exponential distribution with the time $\tau$ as the average.

If we suppose that the packet i is weaker than the packet k (compared to the capture effect of condition 3), then we can thanks to the figure 4.1 know the duration during which the packet i would be lost if a new k packet arrives.


Figure 4.1: Loss of packet i if transmit of k in $T_{C}$
When calculating this probability we will have to assume that the size of the packets is the same as well as the parameters used to send it (which is the case since we are doing the calculation SF by SF), so we get a constant value denoted by $T_{r} e c$ which represents the time necessary to process a packet, the collision duration is therefore the duration during which the two packets will be processed even if we lose the information concerning one of the preambles, so we have $T_{c}=2 T_{\text {rec }}-T_{\text {pream }}$.
The calculations for the processing time of packets $T_{\text {rec }}$ and the processing time of a preamble $T_{\text {pream }}$ are given in cite LoRaScale and are as follows:

$$
\begin{align*}
T_{\text {pream }} & =\left(N_{\text {pream }}+4.25\right) * T_{\text {sym }}  \tag{4.11}\\
T_{\text {payload }} & =\left[8+\max \left(\text { ceil }\left(\frac{8 p l-4 S F+28+16}{4(S F-2 D E)}\right) *(c r+4), 0\right)\right] \\
T_{\text {rec }} & =T_{\text {pream }}+T_{\text {payload }} \tag{4.12}
\end{align*}
$$

Where $T_{\text {sym }}=\frac{2^{S F}}{B W}, N_{p}$ ream is the number of preamble symbols which is often $8, p l$ is the packet size in bytes, $c r$ is the CodingRate which in our model is 1 and $\mathrm{DE}=1$ if BW $=125 \mathrm{kHz}$ and $\mathrm{SF} \in 11,12,0$ otherwise.

Once the processing times have been calculated, it is possible to obtain $\mathbb{P}$ (condition 4 ) by the following formula:

$$
\begin{equation*}
\mathbb{P}(\text { Condition } 4)=\mathbb{P}\left(t_{k} \in T_{c}\right)=F_{X_{k}}\left(T_{c}\right)=1-e^{-\frac{T_{c}}{\tau}} \tag{4.13}
\end{equation*}
$$

Now that we have defined the probabilities of not colliding for the conditions 3 (4.10) and 4 (4.13) we can denote the probability $Q_{i}$ of not colliding as:

$$
Q_{i}=\prod_{\substack{k \neq i \\ S F_{i}=S F_{i} \\ \text { reqeqk }_{k}=\text { freq }_{i}}}(1-\mathbb{P}(\text { condition } 3) \mathbb{P}(\text { condition } 4))
$$

$$
\begin{equation*}
\prod_{\substack{k \neq i \\ S F_{k}=S F_{i} \\ f r e q_{k}=f r e q_{i}}}\left\{1-\left(\frac{1}{2}\left(1+\operatorname{erf}\left(\frac{6+10 \eta \log _{10}\left(\frac{d_{i}}{d_{k}}\right)}{2 \sigma}\right)\right)\right)\left(1-e^{-\frac{T_{c}}{\tau}}\right)\right\} \tag{4.14}
\end{equation*}
$$

However, it should be noted that the duration $\tau$ must respect a regional duty-cycle limit, so we will take as value $\tau$ :

$$
\tau=\max \left(\tau, \frac{T_{\text {rec }}}{p_{0}}\right)
$$

Where $p_{0}$ is the regional duty cycle limit and is 0.01 in Europe.

### 4.4 Success probability

Now that we have calculated the probability of not losing the packet $H_{i}$ as well as the probability of not colliding $Q_{i}$, we can calculate the probability of success $\mathbb{P}_{i}$ as the product of the two probabilities.

As we need for the studied node the probability of success for each SF, we will have to calculate $H_{i}$ and $Q_{i}$ by assigning to the studied node in turn the SFs from 7 to 12 , without changing the distribution of the other node, we will thus obtain the probabilities of success individually for each of the SFs.

So for any state $s \in T$ we have the probability of transition to the successful state using the transmission action Tr which is:

$$
\mathbb{P}(s, T r, S u c c e s)=H_{i}(S F) * Q_{i}(S F)
$$

Likewise the probability of transition to the state of failure or associated waiting noted $s^{\prime}$ is:

$$
\mathbb{P}\left(s, T r, s^{\prime}\right)=1-H_{i}(S F) * Q_{i}(S F)
$$

For example, when attempting to transmit the " $\mathrm{T} / 7$ " state, we have the probability of success :

$$
\mathbb{P}(T / 7, T r, \text { Success })=H_{i}(7) * Q_{i}(7)
$$

and the probability of success:

$$
\mathbb{P}(T / 7, T r, W / 7)=1-H_{i}(7) * Q_{i}(7)
$$

### 4.5 Cost and Rewards

In order to be able to evaluate the best policy in our CDM, therefore the best spreading factors to choose for each k transmission attempt, we need to find a reward function on successes and at the same time a cost function on failures.

To fix our reward function we base ourselves on the energy cost of each transmission attempt, knowing that the energy cost depends on the transmission time and we have seen that this transmission time depends on the symbol time such as:

$$
T_{s y m}=\frac{2^{S F}}{B W}
$$

An attempt to transmit on SF12 therefore has an energy cost over transmission airtime approximately $2^{5}$ times greater than when using SF7.

To represent this we then have to during the resolutions assign values V to each SF such as $V(7)>V(8)>V(9)>V(10)>V(11)>V(12)$

On success during a transmission attempt the system will therefore receive a reward V (i) with i the SF used in the transmission. On a successful transmission we therefore have the reward $R$, for a state $s \in T$ and using the SF i:

$$
R(s, T r, \text { Success })=V(i)
$$

But the objective being to succeed in our transmission, we must sanction the failures with a cost function, the sanction will be proportional to the gain V (i) due to a success and dependent on the number of past uses of the SF i.

We therefore have for any state $s \in T$, for any SF i, for any associated state s' $\in \mathrm{W}$, for any punishment rate $\alpha \in[0 ; 1]$, and n (i) the number of uses of the SF during previous attempts, the reward function in case of failure (which is therefore a sanction):

$$
R\left(s, \operatorname{Tr}, s^{\prime}\right)=-\alpha * n(i) * V(i)
$$

This cost is intended to penalize continued use of the same SF if it fails and therefore reduce the total value of using a small SF and gradually increasing the punishment in order to tip the system on the use of a SF having less value but having a theoretically higher chance of success.


Figure 4.2: Model of MDP

## Chapter 5

## Mise en place et résolution du MDP

All MDP construction and resolution is done on Scilab using the cite MDPScilab MDP toolbox.

Before coming to the complete construction of the MDP, we must come back to the form it takes, in fact we initially planned to have for each of the states the SF used for the current attempt as well as the set of SFs used by previous attempts, but during the initial creation of this model we reached a technical limit for the resolution, in fact this initial model had a number of states which exceeded two million which prevented the resolution due to a lack of memory. We therefore decided to change the model, although we no longer have the information on the order of use of the SFs during the previous attempts, each state still knows the SF of the current attempt and this time knows the number of using each of the SFs during the previous attempts, this change of model made it possible to divide by approximately 100 the number of states present without changing the reward and probability matrices.

It is then possible for us to create and resolve an MDP for each of the nodes of the distribution and to retrieve the best possible policy as we will see below.

As we have seen a number of parameters are considered for the formation of the CDM in the rest of this report we will use the following values:

- Bandwidth : $B W=125 \mathrm{kHz}$
- transmit power : $P i=14 \mathrm{dBm}$
- Loss path coefficient $\eta=2.32$
- Deviation $\sigma=7.8$
- Reference distance $d 0=1000 \mathrm{~m}$
- Loss at d $0, P L 0=128.95$
- Payload $p l=10$ Bytes
- Average transmission time $\tau=5 \mathrm{~s}$


### 5.1 Nodes distribution

Before building the MDP, we need a set N of nodes which is the distribution, we decided in our results first to create a distribution of 100 nodes.

The distribution of the nodes is done uniformly in a circle having for radius a maximum distance maxDist $=8921.35 \mathrm{~m}$, this value of the maximum distance is calculated with the following equation:

$$
\max \text { Dist }=d_{0} * 10^{\frac{P L-P L_{0}}{10 \eta}}
$$

Where PL is the maximum pathloss for which a sent message can arrive using SF12 with antenna sensitivity using $\mathrm{BW}=125 \mathrm{kHz}$ which is sensi $=-137 \mathrm{dBm}$ and is obtained by the calculation given by the equation:

$$
P L=P_{i}-\text { sensi }
$$

Once the 100 nodes have been placed, each one is assigned a frequency which is chosen uniformly from among the 3 frequencies which are $868.1 \mathrm{MHz}, 868.3 \mathrm{MHz}$ and 868.5 MHz .

The last step necessary for the creation of the distribution of nodes concerns the attribution of the use of an SF to each of the nodes, indeed as we have seen it is necessary for us to know the SF used by the other node during the probability calculations for the creation of the MDP, we perform others for our 100 nodes a calculation of $H_{i}$ for all the SFs, the SF which will be allocated will be the smallest SF which exceeds a certain threshold that we have fixed at 0.7 , in the case where no $\mathrm{H}(\mathrm{SF})$ exceeds the fixed threshold then the SF assigned to the name will be as large as possible, i.e. the SF12 this SF will be used later in the simulations and will be named minSF.

### 5.2 MDP Resolve

To solve the MDP we use the iteration by value algorithm which allows to assign to each state the action having the best expected reward in the long term using Bellman's equations such that the algorithm tries to find the action a for which the value $\mathrm{V}(\mathrm{s})$ is the greatest:

$$
\left.V^{i}(s)=\max _{a}\left(\sum_{s^{\prime}} \mathbb{P}\left(s, a, s^{\prime}\right) * R\left(s, a, s^{\prime}\right)\right)+\gamma V^{i-1}\left(s^{\prime}\right)\right)
$$

Where $\mathrm{R}(\mathrm{s}, \mathrm{a}, \mathrm{s}$ ') is the reward value of the transition from state s to state s ' using action a, $\gamma$ is a discount value that we set at $0.95, \mathbb{P}\left(s, a, s^{\prime}\right)$ is the probability of reaching state s' by action a from state s and where $V^{i}\left(s^{\prime}\right)$ is the value of s' during iteration i. Solver will iterate through to a maximum number of iterations or until there is no significant change in any $\mathrm{V}(\mathrm{s})$ value.

Once the algorithm is finished, the action that will be assigned to each state as the best action will therefore be the action a for which we have obtained the best value V (s).

### 5.3 Data recovery and usage

To have the ability to use the data obtained by the MDP, it is above all necessary to know which data to recover, as we have seen a transmission in LoRaWAN technology can perform up to 8 attempts without acknowledgments before abandoning the sending, this is what we represented in our model.

Now that we have solved the CDM we know for each of the states that it is the best action to take in terms of reward, the actions that interest us are the actions of choice
of SF , to recover the plan on the 8 transmission attempts we will therefore follow the optimal path, i.e. recover the optimal action for state s and continue by following the transition linked to the action, in the case of transmission states the transition that we will follow will be that due to a failure, to be able to recover the optimal action during the previous attempt.

For example if the optimal action in the initial state " $S_{0}$ " is to take SF7, we will recover this value of 7 the go to the state linked to the transition by the action a7 which is the transmission state " $\mathrm{T} / 7$ ", in a transmission sate the only possible action being to transmit we do not recover any data and we the return to the state linked to the state " $\mathrm{T} / 7$ " by a failure,that is to say the state of waiting " $\mathrm{W} / 7$ ", in this state we will one again recover the optimal action, and as we are this time in a waiting state we will keep this value which is a choise of SF and so on until we get an array that contains the 8 values of the chosen SFs.

The objective being to fix an initial table for the nodes of the simulation, it is thus necessary for us to recover these values for the whole of the nodes which is why we must carry out an MDP and its resolution for the 100 nodes of our distribution in order to the exploit the results during the simulation, for this we recover all the information concerning the position, the minSF, the frequency, as well as the prédiction given by hte MDP in a file which will be used for the creation of the distribution and initial probability tables in the simulations.

## Chapter 6

## Simulation results

Once the MDP resolutions are done, we need to perform simulations using different methods of initializing the SF choice probability table, this table represents the probability for each of the 100 nodes to choose a SF to each attempt, we will denote by C (i) the probability of choosing SF i.

The different methods used will be as follows:

### 6.1 Differents ways to initialize the SF table

### 6.1.1 BaseSTEPS

In the baseSTEPS method, the table of probabilities of choosing SF depends on the minSF of the node with all the SFs less than the minSF having a probability of being chosen equal to 0 .

$$
C(i)=\left\{\begin{array}{cc}
0 & \text { if } \\
e^{-2 * \mid \operatorname{minSF-i|}} & \text { otherwise }
\end{array}\right.
$$

To know the probability, we then need to normalize the table C such that for each SF i:

$$
C(i)=\frac{C(i)}{\sum_{j} C(j)}
$$

This method ensures a reduction in the probability of choosing an SF the further it is from the minSF.

### 6.1.2 Proportionnal

In this method the initial probability table will be defined thanks to the table obtained with the MDP, each time the value of a certain SF appears in the choice of strategy of the MDP prediction, the value $\mathrm{C}(\mathrm{i})$ will be incremented by 1 , once the eight SFs of the prediction pass and the values assigned to $\mathrm{C}(\mathrm{i})$ we normalize table $\mathrm{C}(\mathrm{i})$ to have the probabilities.

For example if in the prediction the SF7 is used twice then the initial probability of choosing the $\mathrm{SF} 7 \mathrm{C}(7)=2 / 8=0.25$.

### 6.1.3 Ordder of appearance

Just like the previous method this method uses the results of the MDP for the initial probability table, the difference between this method and the previous method is that the score of each SF in the prediction depends on its placement, if we used the SF i at attempt k then we have $\mathrm{C}(\mathrm{i})=\mathrm{C}(\mathrm{i})+\mathrm{k}$, again to obtain probabilities we will have to normalize the table.

We therefore have whether for example the SF10 was predicted during attempts 5 and 7:

$$
C(10)=\frac{5+7}{36}=\frac{1}{3}
$$

### 6.1.4 $50 \%$ premium

This method uses the same process as the proportional method each of the appearances of SF i increments $\mathrm{C}(\mathrm{i})$, but in this method we want to make sure that $C(\min S F) \geq 0.5$, for this before normalizing the array we add at $\mathrm{C}(\mathrm{minSF})$ the sum of $\mathrm{C}(\mathrm{i})$, this ensures that during normalization we have $C(\min S F) \geq 8 / 16$.

### 6.1.5 $\quad 25 \%$ premium

This method uses the same process as the proportional method each of the appearances of SF i increments $\mathrm{C}(\mathrm{i})$, but in this method as in the previous method we want to ensure a minimum probability for the minSF which is in this method of $25 \%$ such that $C(\min S F)$ geq0.25, for that before normalizing the array we add to $\mathrm{C}(\mathrm{minSF})$ a third of the sum of $\mathrm{C}(\mathrm{i})$, this ensures that during the normalization we have $C(\min S F) \geq$ 2.33/10.33.

### 6.2 First Result

One of the last parameters to set to perform the MDP resolution and therefore the recovery of the values for the initialization of the probability tables are the values assigned to the successes for each of the SFs, for this we will again use the calculator [1] to obtain the average energy hit of a device over a day using each of the SFs, once these values have been obtained we will compare the energy ratio between the SF12 and all the other SFs. After using the calculator we obtain the results given in the table 6.1 with the bandwidth parameter 125 kHz , a duty cycle of 5 seconds and a battery voltage of 3 V .

| SF | Energy $(\mathrm{mJ})$ |
| :---: | :---: |
| 7 | 77.72 |
| 8 | 132.03 |
| 9 | 263.99 |
| 10 | 527.93 |
| 11 | 868.94 |
| 12 | 1737.82 |

Table 6.1: Average energy consumption in a day

| SF | $\mathrm{V}(\mathrm{i})$ |
| :---: | :---: |
| 7 | 22.36 |
| 8 | 13.16 |
| 9 | 6.58 |
| 10 | 3.29 |
| 11 | 1.99 |
| 12 | 1 |

Table 6.2: value given to a success using SF i

To obtain the values, we will therefore make for each of the SFs the ratio between the energy consumed by the use of SF12 with all of the other energy.

$$
V(i)=\frac{E(12)}{E(i)}
$$

where $\mathrm{E}(\mathrm{i})$ is the average consumption of energy by the device in a day using the SF i .
We obtain the value in the table 6.2


Figure 6.1: Number of collisions using the MDP with $\alpha=0.10$
We see in figures 6.1, 6.2 and 6.3 that the first two methods, the proportional and that of order of appearance, which ensure no probability of choosing the minSF have results much lower than the other methods and especially the baseSTEPS method, for their part the methods ensuring a minimum on the minSF, the premium of $50 \%$ and $25 \%$, have results that are close to baseSTEPS, for the following results we will therefore only keep these 2 methods and we will change the penalty value such that $\alpha=0.2$, we have also reduced the number of nodes of the simulations to 50 . .

We notice on these second results (figure 6.4, 6.5 and 6.6) that when we increase


Figure 6.2: Number of ACK using the MDP with $\alpha=0.10$
the penalty rate $\alpha$ the results of the different methods are come together, and are even sometimes almost confused which is very encouraging for our method of prediction.


Figure 6.3: Ratio of ACK using the MDP with $\alpha=0.10$


Figure 6.4: Number of collisions using the MDP with $\alpha=0.20$


Figure 6.5: Number of ACK using the MDP with $\alpha=0.20$


Figure 6.6: Ratio of ACK using the MDP with $\alpha=0.20$

## Chapter 7

## Model Checking

To analyze the performance of our MDP, we will use probabilistic model checking techniques which allows to obtain the probabilities of occurrence of certain events, here the events that interest us concern the arrival or not of the packet sent by the node to the base station level.

To do this Model Checking we will use PRISM [10] which is a tool developed at the University of Birmingham. PRISM allows the implementation and analysis of Markovian decision models by describing the model with guarded commands. The system is made up of one or more modules interacting with each other, the modules themselves being made up of a list of variables indicating their state and kept commands indicating their behavior in the form:

$$
[]<\text { garde }>\rightarrow<\text { commande }>
$$

The guard being a predicate on the variables of the system and the command being a transition which is applied in the case where the guard is valid. If the transition is probabilistic it will be specified in the form:

$$
<\text { prob }>:<\text { commande }>+\ldots+<\text { prob }>:<\text { commande }>
$$

Moreover PRISM accepts PCTL specifications which will allow us to obtain the probabilities that certain events occur.

To represent our MDP in PRISM in will use several variables, one of the variables allows to know if we are in a state of waiting, transmission, success or failure as well as variables which allows to know the number of use of each of the SFs.

The results that we will not give below were given using the probabilities of the table 7.1 that we calculated from the equations given in 4.3.2 and 4.3.4 for a node located 2600 m from the base station.

| SF | $\mathrm{P}(\mathrm{i})$ |
| :---: | :---: |
| 7 | 0.39 |
| 8 | 0.56 |
| 9 | 0.70 |
| 10 | 0.80 |
| 11 | 0.89 |
| 12 | 0.92 |

Table 7.1: Probabilities

First, we will see the minimum and maximum probabilities of reaching a state of success in at most k attempts, namely:

$$
\text { Pminmax }=?[\text { true } U(" \text { success" \& Numberoftry }<=k)]
$$

We obtain the results of the figure 7.1.


Figure 7.1: Probabilities to reach a success State
We can also be interested in the probabilities of reaching a state of failure:

$$
\text { Pminmax }=?[\text { true } U " F \text { Failure" }]
$$

We find the following probabilities : $\operatorname{Pmax}=0.019$ and $\operatorname{Pmin}=1.9 E-9$
Once the probabilities of success and failure have been obtained, it may be interesting to look at the probability of being in a state where the number of transmission attempts is equal to k either:

$$
\text { Pmaxmin }=?[\text { true } U \text { NombreTentatives }=k]
$$

with result being in figure 7.2


Figure 7.2: Probabilities to reach success at the try k

## Chapter 8

## Conclusion

In our objective of improving the performance of a LoRaWAN network we have implemented a method that changes the SFs during transmission attempts in order to reduce collisions based on probabilities that depend on previous attempts, but a One of the problems with this method was to initialize the probability table before the slightest attempt.

To solve this problem we have tried to propose a method of predicting the SF array using a Markovian decision process. The first results we are getting are encouraging but we need to keep experimenting by changing the number of nodes as well as potentially the alpha penalty ratio to confirm the first results we got.

Following the confirmation, we will therefore have to set up our method in real conditions by setting up a model using transmitters and an antenna using LoRa technology to confirm our experimental results.

## Bibliography

[1] Developers Semtech LoRa. SX1272 LoRa Calculator. https://www.semtech.com/products/wireless-rf/lora-core/sx1272.
[2] Muhammad Asad Ullah, Junnaid Iqbal, Arliones Hoeller, Richard Demo Souza, and Hirley Alves. K-means spreading factor allocation for large-scale lora networks. Sensors, 19(21), 2019.
[3] Semtech. SX1272/73 -860 MHz to 1020 MHz low power long range transceiver. Semtech, March 2015.
[4] Martin Bor, Utz Roedig, Thiemo Voigt, and Juan Alonso. Do lora low-power widearea networks scale? 112016.
[5] Madoune R. Seye, Bassirou Ngom, Bamba Gueye, and Moussa Diallo. A study of lora coverage: Range evaluation and channel attenuation model. In 2018 1st International Conference on Smart Cities and Communities (SCCIC), pages 1-4, 2018.
[6] Fernando Palacio, Ramon Agustí, Jordi Pérez-Romero, Miguel López-Benítez, Sebastien Grimoud, Berna Sayrac, Ioannis Dagres, Andreas Polydoros, Janne Riihijärvi, Jad Nasreddine, Petri Mähönen, Liljana Gavrilovska, Vladimir Atanasovski, and Jaap Beek. Radio environmental maps: information models and reference model. document number d4.1. 042011.
[7] Juha Petajajarvi, Marko Pettissalo, Konstantin Mikhaylov, Antti Roivainen, and Tuomo Hänninen. On the coverage of lpwans: Range evaluation and channel attenuation model for lora technology. 122015.
[8] Gilles Callebaut and Liesbet van der perre. Characterization of lora point-to-point path-loss: Measurement campaigns and modeling considering censored data. IEEE Internet of Things Journal, PP:1-1, 112019.
[9] Semtech. LoRaWAN Specification v1.1. Semtech, October 2017.
[10] M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of probabilistic real-time systems. In G. Gopalakrishnan and S. Qadeer, editors, Proc. 23rd International Conference on Computer Aided Verification (CAV'11), volume 6806 of LNCS, pages 585-591. Springer, 2011.

