Optimization of Gateway Deployment in Smart Metering LoRaWAN Networks

Madalena Maria Muller e Sousa Avillez Albuquerque
madalena.muller@ist.utl.pt

Instituto Superior Técnico, Lisboa, Portugal

March 2021

Abstract

Recently, Low Power Wide Area Networks (LPWANs) have attracted great interest due to the need of connecting more and more devices to the so-called Internet of Things (IoT). We have witnessed the development of Long Range (LoRa) technology as an emerging technology suitable for smart grids (SG). Therefore, this work uses theoretical considerations to develop a channel model of LoRa that considers propagation attenuation, shadowing, and fading effect.

Hence, a theoretical model developed in this study proposes to estimate the optimal gateways positions of LoRa. Each experiment considers smart meters with defined locations and plots the NSGA-II Pareto optimal curve with both objectives: the minimum number of gateways combined with the packet loss of the channel. The packet loss and distance between nodes are estimated theoretically.

Results show a significant decrease in the signal interference in the presence of fading or shadowing. This effect had a considerable impact on the network's optimization. Therefore, this led the study to find that the effect of fading and shadowing can reduce packet loss because of the spreading factor's orthogonality. The contribution of this work is the study of the impact of fading and shadowing on the optimization and deployment of LoRaWAN.

Keywords: LoRa, Gateway, Fading, Shadowing, NSGA-II, LoRaWAN

1. Introduction

Traditional networks weren't developed for a typical Internet of Things (IoT) scenario. The power consumption from the connected network devices is too high and so is the cost of its connectivity when the number of devices in the network scales. More recently Long Power Wide Area Network (LPWAN) technologies were developed to meet the requirements of the IoT, being able to cover huge numbers of low power devices, allowing device lifetimes in the order of years.

Long Range Wide Area Network (LoRaWAN) is such a LPWAN technology whose presence is increasing. This technology is popular in battery-powered systems that require transferring a small amount of data at short intervals over long range. A LoRaWAN network capacity depends on many factors such as the distance between the end-nodes. There must exist a trade-off between coverage and costs. Ideally, all devices would have a gateway (GW) close by. This increases the performance and decreases the consumption of the devices, increasing their lifetime. However, as the number of devices scale, better coverage requires more GWs, which increases the costs of the network's installation and maintenance. To optimize this problem, the distance from every end-device to the GW should be minimum without exceeding a large number of GWs. Optimizing this distance, the overall consumption decreases and the global network performance improves, reducing the economic cost. This study aims to analyze the impact of fading and shadowing phenomena in the trade-off between coverage and cost. Hence, the results will optimize the limited number of GWs locations and maximize the signal performance for a fixed number of smart meters. For this accomplishment, a LoRaWAN analytical model and implementation of an optimization algorithm to determine the best GW's positions were developed for different propagation models.

2. LoRa

Lora is a LPWAN modulation protocol designed for long distance communication. LoRa is the physical layer often used with the LoRaWAN MAC layer protocol. While LoRa is designed and patented by Semtech, LoRaWAN is open, non-profit and developed by the LoRa Alliance. This protocol supports bi-directional communication, mobility, localization and security required by IoT applications. Using a modulation technique known as chirp spread spec-

trum (CSS), the LoRa signal can vary depending on the message it is carrying. It also uses the entire channel bandwidth for broadcasting, allowing it to be more robust to noise and frequency offsets. LoRa CSS modulation is explained in section 2.1.

2.1. LoRa Modulation

LoRa modulation is based on a derivative of Chirp Spread Spetrum (CSS): a signal is spread by using wideband linear frequency modulated chirp pulses to encode information. It uses Frequency Shift Keying(FSK) in order to achieve lower consumption and it uses Chirp Spread Spectrum (CSS) for large area coverage. LoRa use of CSS improves resilience and robustness against interference, Doppler effect, and multipath. In CSS there are up-chirps when the frequency increases and down-chirps if the frequency decreases. These chirp signals (frequency varying sinusoidal pulses) are used as carrier signals where the message is encoded on. In addiction, the use of CSS modulation means the signals are orthogonal to each other, which allows multiple data rates simultaneously transmitting on the same channel. A representation of an up-chirp can be seen in Figure 1.

In LoRa, the starting frequency of a chirp, f_0, seems to be used to represent a symbol. In the case of an up-chirp, the frequency increases steadily up to f_{max}. Then it jumps back to f_{min}, growing steadily to $f0$ again. Afterwards, the next symbol is ready to be transmitted with a new f_0 frequency, and the process is repeated.

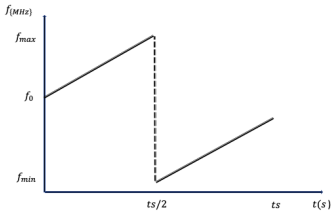

Figure 1: Up-chirp representation

By examining Figure 1 the bandwidth, which is the number of vibrations or wave cycles per second, is given by:

$$BW = f_{max} - f_{min}, \tag{1}$$

Lora uses three different bandwidths:125 MHz, 250 MHz, 500 MHz where symbols are modulated over a chirp of a chosen bandwidth and different spreading factors are used based on data rate requirement and channel conditions. The Spreading Factor (SF) represents the number of encoded bits in a symbol and be obtained by:

$$SF = log_2(\frac{Chirprate}{Symbolrate}), \tag{2}$$

This also means that every symbol is encoded in 2^{SF} chirps that cover the available bandwidth. In Lora SF can assume values from 7 to 12. The symbol duration, T_s, can therefore be expressed as:

$$T_s = \frac{2^{SF}}{BW}, \tag{3}$$

Considering that the symbol rate,R_s, is the inverse of T_s and that the same is related with the chip rate, R_c, by the expression:

$$R_c = R_s \times 2^{SF}, \tag{4}$$

Consequently. the bitrate is given by:

$$R_b = SF \times \frac{BW}{2^{SF}}, \tag{5}$$

It's important to highlight that by increasing the spreading factor, the time needed for the data to be received will be longer. Therefore, the robustness of the connection will be higher but the power consumption also increases. Another disadvantage of longer message transmission is the higher probability of collisions. This calls for a need to balance the wait time and power consumption according to the application used. The gateways have the ability of receiving data in different spreading factors which allows the end-nodes to choose the spreading factor that fits better. Furthermore, the SF affects the sensitivity S of the receiver that is defined as [6]:

$$S = \text{-}174 + 10log10(BW) + NF + SNR, \tag{6}$$

where -174 is due to the thermal noise at the receiver in 1 Hz bandwidth, NF is the Noise Figure at the receiver (which is fixed for a hardware configuration data) and SNR is the signal to noise ratio required for the modulation.

By spreading the signal in time domain it is possible to reduce the Bit Error Rate (BER) and achieve long-distance communication. LoRa can demodulate signals which are -7.5 dB to -20 dB below the noise floor.

To improve resilience against interference LoRa uses Forward Error Correction (FEC). FEC is the process where error correction bits are added to the transmitted data. The introduction of redundant data helps to restore the data when it gets corrupted. Making use of Hamming codes for FEC, LoRa offers code rates of 4/5, 4/6, 4/7, and 4/8. For example, a transmission with a coding rate of 4/5 means one bit of redundancy is added to each block of 4 bits of useful information. Code rate expression is given by:

$$CR = \frac{4}{4+n}, \tag{7}$$

Because of FEC coding, the number of used bits decreases, and consequently the bit rate is given by:

$$R_b = SF \times \frac{BW}{2^{SF}} \times CR, \tag{8}$$

Also, as refered in this section, a change of the SF has implications in the data rate transmission. In order to achieve high network capacity, LoRaWAN uses the Adaptive Data Rate (ADR) mechanism developed to optimize data rates, wait time and power consumption. This mechanism consists of selecting a SF value and bandwidth for each end-node based on the collected connection metrics. This means the SF can be changed to get better data rates for transmissions where the link is better. Furthermore, the transceivers can manage receiving different data rates in different channels. Lowering the SF means increasing the data rate meaning lowering the Time on Air. If a node needs less Time on Air, this time can be used by other nodes to transmit. Therefore, there will be an increase on battery life preservation.

Additionally, Transmission Power on a LoRaWAN device can usually be adjusted from 2 dBm to 16 dBm EIRP by step of 2 dB. The maximum allowed power is given by the local regulations. In Europe and most of the world, the value is 14 dBm.

2.2. LoRaWAN Network

LoRaWAN network architecture is usually a mesh architecture with star topology. This type of topology is the one that gives more advantages in terms of battery life of the end-nodes when long-range connectivity is achieved. Figure 2 shows a typical loRaWAN network composed by these type of elements:

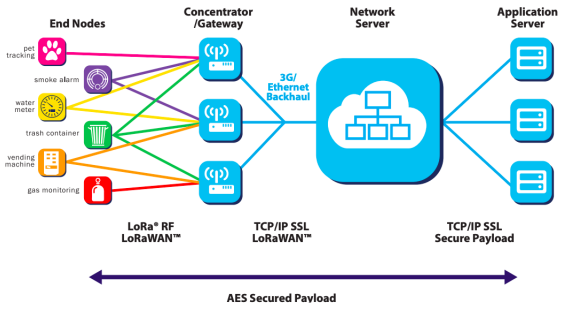

Figure 2: LoRaWAN Arquitecture. [1]

- **End node** - Consists of some sensor or other entity transmitting or collecting data. In the uplink scenario, the data is transmitted from the end node to the gateway. When the data is transmitted from the gateway to the end node it is called a downlink. In the LoRaWAN network, an end-node can send data to more than one gateway.
- **Gateway** - Receives the data coming from the end nodes and sends it to the network server. The connection to the server via some backhaul network(IP, Ethernet, WiFi, etc).
- **Network Server** - Collects the information from the gateway, where there is a filtering of redundancy data, the performance of security checks, and avoidance of collisions. The Network Server then forwards the information to the Application Server.
- **Join Server** - The Join Server (JS) handles the LoRaWAN join flow, including Network and Application Server authentication and session key generation.The JS manages the Over-the-Air (OTA) End-Device activation process. There may be several JSs connected to a NS, and a JS may connect to several NSs.
- **Aplication Server** - The final destination of the data, either in public or private clouds where the applications are running.

Looking over Figure 2 a particular device can be connected to more than one gateway by communicating over LoRa protocol. On the other hand, the communication between a gateway and the Network Server is over TCP/IP, meaning the gateway has to be connected to the Internet in some way.

3. Multiobjective Optimization Algorithms

Most real-life science problems require a Multiobjective Optimization (MOO), which involves several conflict objectives and aims to convert all objectives into a single objective (SO) function. However,this MOO has some limitations and a simple optimization process is no longer acceptable for systems with multiple conflicting objectives. In MO problems there is no single solution making the optimization more difficult to determine. Instead, there is a set of acceptable trade-off optimal solutions: Pareto front. The solution most desirable to the designer or decision maker (DM) is selected from the Pareto set. Generating a Pareto set allows the DM to make an informed decision with a wide range of options since it contains the solutions that are best for all objectives. The MOOA that was used in this study is NSGA-II.

3.1. NSGA-II

NSGA is a popular non-domination based genetic algorithm for multi-objective optimization. It is a very effective algorithm but has been generally criticized for its computational complexity, lack of elitism and for choosing the optimal parameter value for sharing parameter. A modified version, NSGA- II [2] was developed, which has a better sorting algorithm , incorporates elitism and no sharing parameter needs to be chosen a priori. To entirely explain NSGA-II, some essential operations

utilized throughout the optimization process need extra attention. These operations are a fast non-dominated sorting approach, a crowding distance assignment, and a crowded-comparison operator. Additionally, the standard genetic algorithm operators such as binary tournament selection, simulated binary crossover and mutation are crucial for the good performance of the NSGA-II algorithm.

3.2. General Description of NSGA-II

The population is initialized as usual. Once the population in initialized the population is sorted based on non-domination into each front. The first front being completely non-dominant set in the current population and the second front being dominated by the individuals in the first front only and the front goes so on. Each individual in the each front are assigned rank (fitness) values based on front in which they belong to. Individuals in first front are given a fitness value of 1 and individuals in second are assigned fitness value as 2 and so on. In addition to fitness value a new parameter called crowding distance is calculated for each individual. The crowding distance is a measure of how close an individual is to its neighbors. Large average crowding distance will result in better diversity in the population. Parents are selected from the population by using binary tournament selection based on the rank and crowding distance. An individual is selected in the rank is lesser than the other or if crowding distance is greater than the other 1. The selected population generates offsprings from crossover and mutation operators. The population with the current population and current offsprings is sorted again based on non-domination and only the best N individuals are selected, where N is the population size. The selection is based on rank and the on crowding distance on the last front. This procedure is explained in Figure 3.

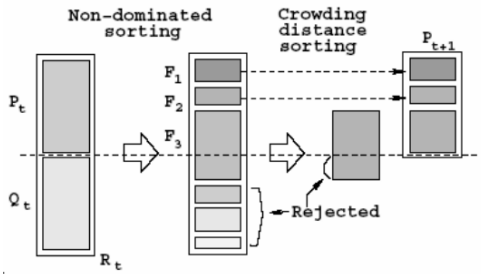

Figure 3: NSGA-II procedure. [2]

These are the three main innovations for the NSGA-II algorithm: a fast nondominated sorting procedure that ensures elitism, a fast crowded distance estimation procedure, and a simple crowded comparison operator that guarantees diversity preservation and estimation.

4. LoRaWAN Analytical Model

The LoRaWAN model developed in this study is based on the INESC-ID model developed in the WiMeCom project [3]. This model allows the computation of the Signal-to-Interference-plus-Noise-Ratio (SINR) and Packet Loss Ratio (PLR) for each device in a LoRaWAN network of Class A devices that may involve several gateways and support more than one frequency band. In the following sections, a detailed explanation of the propagation model and packet error model is provided.

4.1. Signal reception model and selection of spreading factors

This study assumes a log-distance path loss model, in which the received power P_r is given by the following expression:

$$P_{r[dBm]} = P_{t[dBm]} - PL_0 - 10.a.log_{10}\frac{d}{d_0} + 20.log_{10}(\alpha) + X(\sigma) \quad (9)$$

where P_t is the transmit power, PL_0 is the path loss at reference distance d_0, a is the path loss exponent, d is the distance between the device and the gateway, α represents the Rayleigh fading, is a random variable with exponential distribution and $E(\alpha^2) = 1$, and $X(\sigma^2)$ is a Gaussian random variable, modeled as log normal, with zero mean and variance parameter σ^2.

In the first part of this study, in order to simplify the model, the terms related with fading and shadowing were not included, so that P_r [dBm] becomes deterministic. Furthermore the effects of shadowing and fading will be included and an explanation of it's calculation for the propagation gain will be given in Section 5. Therefore, to calculate log-distance path loss we only consider the elements of the following equation:

$$P_{r[dBm]} = P_{t[dBm]} - PL_0 - 10.a.log_{10}\frac{d}{d_0} \quad (10)$$

The value of $PL0$ in Equation 10 used to be based on a free space calculation or experimental measurements performed at distance of reference, $d0$. Considering project WiMeCom [3] empirical parameters, the value used in this study is 8.1dB.

In this model the receiver sensitivity is taken into account when choosing the spreading factor, it being considered that the device will employ the lowest spreading factor possible:

$$SF_i = \min_g(j|P_{r^i g} \geq RS_j), \quad (11)$$

where SF_i is the spreading factor chosen by device i, Pr_i^g is the received power from device i at gateway g, and RS_j is the receiver sensitivity associated with spreading factor j. In this study, the allocation of SF depends on the distance to the gateways and the propagation model that is used.

4.1.1 Pathloss exponent

The following Table 1 contains the values for the path loss exponents in different environments. In this study the LoRaWAN model is analysed in a simulating shadowed urban cellular radio environment.

Environment	Path loss exponent
Free Space	2
Urban area cellular radio	2.7 to 3.5
Shadowed urban cellular radio	3 to 5 dB
In building line of sight	1.6 to 1.8
Obstructed in buildngs	4 to 6
Obstructed in facotries	2 to 3

Table 1: Path loss exponent [5]

4.2. Packet error model taking into account a single gateway

The packet loss model for the case of a single gateway was modeled in two ways. The first approach calculates the distribution of the total interfering power as the convolution between the probability distribution functions of the possible interfering devices. Then, it calculates the Signal-to-Interference-Ratio (SIR) distribution. Due to the computational complexity of this approach, another model was developed, which computes the packet loss probability based on the collision probability and average SIR during collision. Both models make the following assumptions:

1. The packet error rate depends only on the SIR, being independent of the specific fields of the packet that are affected by interference.

2. Interference occurs only between transmissions using the same channel and spreading factor. Different spreading factors are orthogonal.

3. All the packets have the same duration T_{tx}.

4.3. Packet error model based on a collision model

4.3.1 Sub-band occupancy and delay model

A LoRaWAN device may choose among different available sub-bands. Spectrum utilization regulations impose duty cycles that limit the load imposed by a device on each sub-band. Also, different sub-bands may have a different number of channels. If one assumes that a device chooses the transmission channel according to a uniform distribution over all the available channels just before transmission, utilization of the sub-bands will be asymmetric. These characteristics are taken into account in the work of René Søresen et al [7]. The model assumes a Poisson process of packet generation. According to this model, the total transmission latency is given by:

$$T_{total} = T_{tx} + T_{w'} \tag{12}$$

where T_{tx} is the time-on-air of the packet and T_w is the waiting time due to duty-cycling. T_tx can be calculated according to the LoRaWAN specification, based on the payload length, spreading factor, channel bandwidth, code rate and protocol overhead. Since different sub-bands may have different regulatory duty-cycles, an asymmetric M/D/c queuing model is considered, with c denoting the number of available sub-bands, which in practice is approximated by a jockeying M/M/c queue. An empirical assumption is made that the waiting line of the M/M/c queue is approximately twice that of an M/D/c queue. Knowing the packet rate λ produced by a device, T_w can thus be calculated based on Little's Law:

$$T_w = \frac{p_{busy,all}}{(\sum_{i=1}^{c} \mu_i - \lambda).2'} \tag{13}$$

where $p_{busy,all}$ is the Erlang-C probability that all sub-bands are busy and $_i$ is the service rate of sub-band i. The service ratio of a sub-band i, r_i is given by the following expression:

$$p_i = \frac{\mu_1}{\lambda}.(1 - p_{i,idle}) \tag{14}$$

The jockeying M/M/c queue is used to calculate $p_{busy,all}$ and $p_{i,idle}$.

4.3.2 Collision model

In order to determine the total traffic load and collision probability, the sub-band occupancy model [7] is used. Based on this model, for each sub-band s and SF j in a gateway, the total traffic load is calculated as follows:

$$L(s,j) = \frac{\lambda.p_s.T_{txj}.N.p_{SF_{s,j}}}{n_s} \tag{15}$$

where p_s is the service ratios of the sub-bands, T_{tx} is the time-on-air of the packet, p_{SF} is the percentage of devices N using spreading factor j in sub-band s and n_s is the number of channels in sub-band s.

Based on a simple ALOHA model, the probability of collision can be calculated as:

$$p_{col,s,j} = 1 - e^{-2L(s,j)} \tag{16}$$

The model defined in [7] assumes that a packet is lost every time there is a collision. However, this

doesn't take into account capture effects, with the possibility of successful packet reception in case the received powers of the colliding packets being too different. This would result in a SIR that is high enough for one of the packets to be received. As such, an extension of the collision model was developed, which takes SIR into account. Regarding manageability, the model assumes that the probability of collision between more than two transmissions is not significant compared to the probability of collision between two transmissions. Hence, only collisions between two transmissions are considered.

The average SIR resulting from a collision is calculated based on the average interference power during transmission of the reference packet.

Since, according to the ALOHA model, the starting times of interfering packets are uniformly random within the vulnerable period of $2.T_{tx}$, it is considered that the average interfering power is one half of the instantaneous power received from the interfering node:

$$SIR_{i,s} = \frac{P_{r_i}}{\frac{\sum_{l=1}^{N} 1_{l \neq i \Lambda SF(l)=SF(i)^{(l,i)}} . \frac{P_{rl}}{2}}{\sum_{l=1}^{N} 1_{l \neq i \Lambda SF(l)=SF(i)^{(l,i)}}}}, \quad (17)$$

Where $SIR_(i,s)$ is the average SIR of the reception of a packet from device i in sub-band s and $1_{l \neq i \Lambda SF(l)=SF(i)}$ is an indicator function to limit the interfering nodes to those who employ the same spreading factor. Again, it is considered that the packet is lost when $SIR_{(i,s)} < 6$ dB.

The average PER of packets transmitted from node i to a gateway g is then estimated as follows:

$$P_{e\,g}^{i} = \frac{\sum_{s=1}^{N_s} \sum_{i=1}^{N} 1_{<6dB}(SIR_{i,s}).p_s}{N.S} \quad (18)$$

4.4. Packet loss model with multiple gateways

When there are multiple gateways within range of a device, the packet loss probability tends to lower, since it is enough that at least one gateway receives the packet in order for the transmission to be successful. The Packet Loss Ratio (PLR) is thus the probability that all gateways have received the packet with errors. If it is assumed that the PERs of different gateways are independent (best case), PLR of device i is calculated as follows:

$$PLR_i^{min} = \prod_{g=1}^{N_G} P_{e\,g'}^{i} \quad (19)$$

where N_G is the number of gateways (it should be noted that for gateways out-of-range from the device, $P_e^{ig} = 1$) . However, the PERs of different gateways are not independent, since some interfering device that causes a collision in one gateway may also interfere at the same time with the reception of the same packet in other gateways. Consequently, at the other extreme, the PLR estimate corresponds to the minimum PER of all the gateways:

$$PLR_i^{max} = \min_g P_e^{i'_g} \quad (20)$$

5. Fading and Shadowing model

The propagation model described in Section 4 predicts the received power as a deterministic function of distance, where the communication range is represented as an ideal cycle.

Experimental results have shown that many well-designed protocols will fail simply because of fading and shadowing experienced in a realistic wireless environment. This section aims to explain how this study adopts a typical LoRaWAN operating scenario where the transmissions of LoRa Class A devices are affected by path-loss, shadowing and fading. With the purpose of evaluating LoRa's performance in Large Scale (path loss, Shadowing) and Small-Scale fading environments, we consider both Log-normal hadowing and Rayleigh fading.

5.0.1 Fading

Fading is caused by movement of transmitter, receiver or other object in the environment. Two common small-scale fading models are Rayleigh and Ricean. A Rayleigh distribution is normally used to describe the statistical time-correlation nature of the received signal envelope, or the envelope of an individual multipath component. When there is a dominant stationary (non-fading) signal component present, such as line-of-sight(LOS) propagation path, the small scale fading envelop distribution is Ricean [4]. In this study, only fast fading following Rayleigh distribution will be considered for the simulation results, considering that you rarely have a dominant LOS ray in a typical LoRaWAN network.

Fading channel can be characterized by a random variable α that describes random nature of envelope fading. Since the performance of wireless communication systems is mainly function of signal to noise ratio, the fading level has to be described in power. Hence, Rayleigh amplitude fading channels can be described as exponential fading distribution in power domain.

The amplitude of a signal subject to fast fading is assumed to be distributed according to a Rayleigh distribution. Therefore, the power of the fast fading effect distributed according to an Exponential distribution. The shorthand $X \sim$ exponential(α) is used to indicate that the random variable X has the exponential distribution with positive scale param-

eter α. The exponential distribution can be parameterized by its *mean* α with the probability density function

$$f(x) = \frac{1}{\alpha} e^{-x/\alpha} \qquad x > 0,$$

for $\alpha > 0$. An exponential random variable X can also be parameterized by its *rate* λ via the probability density function

$$f(x) = \lambda e^{-\lambda x} \qquad x > 0,$$

for $\lambda > 0$.

Regarding this study research the distribution of power in the receiver is an exponential distribution. This means, because of fading effect the power at the receiver due to path loss, will lead to the instantaneous power at the receiver having the probability distribution explained above.

Therefore, this propagation model can be represented by Equation 21.

$$P_{r[dBm]} = P_{t[dBm]} - PL_0 - 10.a.log_{10}\frac{d}{d_0} + 20.log_{10}(\alpha) \tag{21}$$

Where α, Rayleigh fading, is a random variable with exponential distribution and E(α^2)=1

5.0.2 Shadowing

Additionally to the fading effect, the log-distance path loss propagation model doesn't consider the fact that the surrounding environmental clutter may be vastly different at two different locations having the same T-R separation. This leads to measured signals which are vastly different than the average value predicted by Equation 10 in log-distance path loss model. Measurements have shown that at any value of d , the path loss $PL(d)$ at a particular location is random and log-normally distributed about the mean distance-dependent value. This is given by the following equation:

$$P_{r[dBm]} = P_{t[dBm]} - PL_0 - 10.a.log_{10}\frac{d}{d_0} + X(\sigma) \tag{22}$$

The log-normal distribution describes the random shadowing effects which occur over a large number of measurement locations which have the same T-R separation, but have different levels of clutter on the propagation path.

This phenomenon is referred to as log -normal shadowing. Simply put, log-normal shadowing implies that measured signal levels at a specific T-R separation have a Gaussian (normal) distribution about the distance-dependent mean of Equation 22 in Log-distance Path Loss model, where the measured signal levels have values in dB units.

The standard deviation of the Gaussian distribution that describes the shadowing also has units in dB . Thus, the random effects of shadowing are accounted for using the Gaussian distribution which lends itself readily to evaluation.

In this model, the values of n, and σ are based on empirical results.

5.0.3 Fading and shadowing validation

For the implementation of fading and shadowing effects some validations were made in Matlab to ensurethe models of propagation were giving the expected results. First, we generated random samples following Rayleigh fading and log-normal shadowing, using Matlab histogram function, the samples led to a exponential distributin in the case of fading and a normal distribution in the case of shadowing. This result,guarantees that the instances of signals generated with fading and shadowing propagation model will be correct.

6. LoRaWAN Gateway Placement Optimization for Smart Metering Infrastructures

Regarding the objective of this study to optimize the deployment of a Smart Metering (SM) network using a LoRaWAN solution, it is assumed that the locations of SM devices is an input to the problem. The objective is then to find the best trade-offs between the cost of the investment on network infrastructure and the quality of the connectivity. For this accomplishment, NSGA-II algorithm. explained in Section 3.1 was developed in Matlab also based on the model developed in the WiMeCom project [3].

6.1. Objective Functions

As already stated, the developed algorithm seeks to find the best trade-offs between the minimum cost for investment on the infrastructure and the quality of connectivity.

Analytically, this tasks aim for a balance between the number of gateways and number of loss packets. For this effect, the OFs should be the following:

OF_1: min Gw

OF_2: min $avg\ p_{loss}$

where G is the number of Gateways and p_{loss} the average of the number of loss packets. The number of lost packets can be calculated based on the computation of the collision probability for ALOHA systems. Since the OFs are defined, the NSGA-II algorithm can be executed. .

6.2. Initialization and Stopping Criteria

NSGA-II starts by generating an initial population set according to the propagation model that is used. In fading or shadowing model the population is initialized assuming an average of 5 channel instances

of the received power. With Log distance model the initial population stars without this channel instances calculation.

The first population is used to generate offspring chromossomes using geneting operations. These chromossomes are merged and sorted in order to select the non-dominated solution to be used in the next generation.

The outputs of the algorithm are monitored and analysed every 10 iterations. The algorithm stops if there isn't a change in the population of the candidate solution. This means that if the Pareto optimal Curve contains the same solution for 10 consecutive iterations it's assumed that the algorithm has converged, and therefore, the simulation is finished.

7. Simulations

The NSGA-II implementation was run in a scenario with 500 devices randomly deployed in a circle of radius 7362 m. This deployment was developed for all propagation models: log-distance path loss, fading, shadowing. An example of this devices deployment is depicted in Figure 4. The devices are represented in colors according to their SF assignment. The green color represents the devices with the lowest SF and the red color represents the ones with the highest. Devices that weren't assigned with a SF value are represented with a black color.

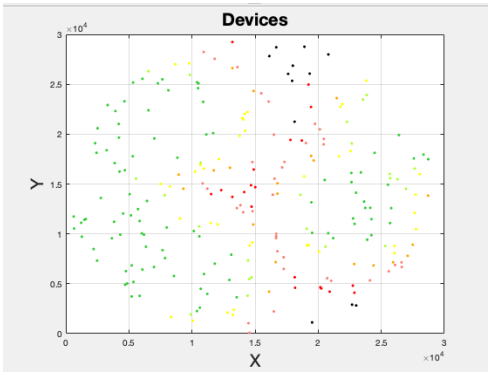

Figure 4: 500 deployment positions.

To simulate the LoRaWAN network with all the propagation models and explanations provided in previous sections, the following tables show the values chosen for the simulations. Table 2 shows which model distribution was used for each propagation effect.

The LoRaWAN parameters and NSGA-II parameters used for the simulations in section 7.1 are listed in Tables 3 and 4.

Path loss model	Log-distance
Fading model	Rayleigh
Shadowing model	Log-Normal
σ	3 dB
Interference model	ALOHA collision probability + SIR matrix

Table 2: Propagation model

Parameters	**Values**
Simulation scenario	circle of radius 7362 m
Transmission power	14dBm
Frequency	868 MHz
Bandwidth	125 KHz
d_0	1m
PL_0	8.1 dB
path loss exponent n	3.76
Code rate	4
Packey payload length	20 Bytes
overload	13 Bytes
nº bits preamble	8 bits
Packet arrival rate of each device	0.0017 packets/s
LoRaWAN sub-bands	G (3 channels)
Spreading factor	[7:12]

Table 3: LoRaWAN parameters

7.1. Simulation Results with different propagation model

Analysing the first model simulation result represented in Figure 5, the trade-off can be clearly distinguished, as an increase in capacity corresponds to an increase in the number of gateways of the solution, and vice-versa. Also, it is explicit that there is an performance improvement of the algorithm solutions from the initial population to the final population. The initial population hasn't as much solutions for the number of gateways as the final population. Figure

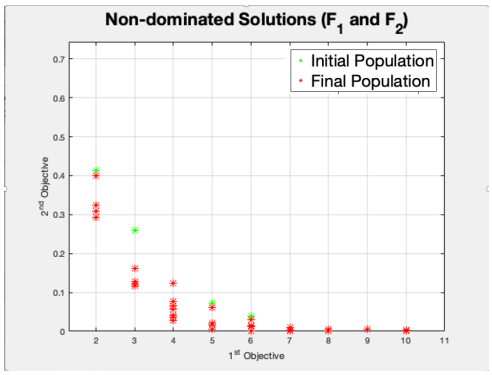

Figure 5: Initial and Final Population for Log-distance propagation model.

Parameters	Values
Minimum number of gateways	1
Maximum number of gateways	10
Population size N_{pop}	500
Crossover percentage p_{cross}	0.7
Mutation percentage p_{mut}	0.4
Mutation rate mu	0.06
Mutation Step Size fraction	0.5
Stopping criteria	$>$ 10 it. same pop
PM channel instances	5

Table 4: NSGA-II parameters

The results of the simulation in Figure 6, shows that for the same number of gateways, the packet loss decreases comparing to the log-distance propagation model.

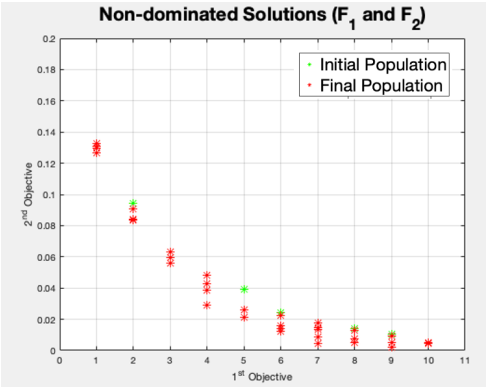

Figure 6: Initial and Final Population for Fading propagation model.

. Figure 7 illustrates shadowing propagation model simulation, where we can notice that the shadowing effect also reduces the interference of the signal propagation similar to the fading effect. However, in this case, the decrease of the packet loss is even more significant.

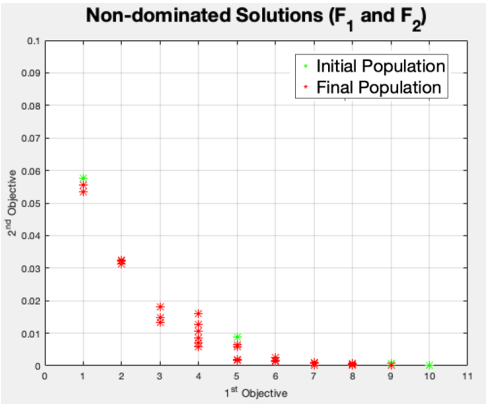

Figure 7: Initial and Final Population for Shadowing propagation model.

Taking into account the observations presented in the previous simulations illustrated in Figure 6 and Figure 7, the same conclusion is to be expected in this propagation model. Having both fading and shadowing effects, the packet loss has less impact on interference and packet loss. This result is represented in Figure 8.

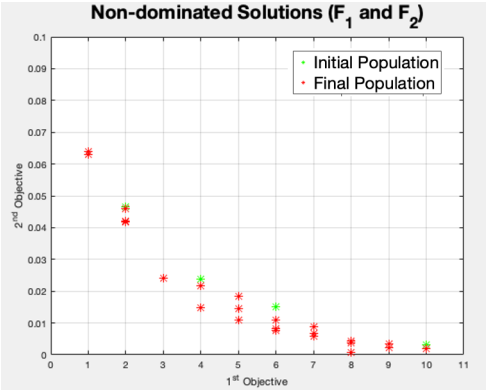

Figure 8: Initial and Final Population for Fading and Shadowing propagation model.

Therefore, it seems that the fading or shadowing effect tends to significantly reduce instantaneous interference and this effect has more impact than the effect on the main signal power, with the overall effect of reducing the packet loss ratio, which could be assumed to be counter-intuitive. In an effort of explaining this performance, numerical results were obtained for the average PLR as a function of the number of devices in a scenario in which the devices are randomly deployed within a circle of radius 7362 m around a single gateway. The propagation models of this study are represented in Figures 9 and 10.

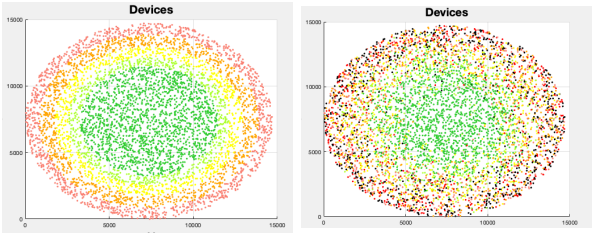

Figure 9: Graphics of SF distribution for log-distance PM (Right) and Fading PM (Left).

As mentioned previously, the fading or shadowing effect tends to reduce instantaneous interference as well as the packet loss ratio. This result is unpredictable at first since it would be presumed that these events would increase both interference and the packet loss ratio.

However, since the simulation area has a small radius and this study's SF allocation procedure is distance-dependent it may happen the decrease of

collision probability in the presence of fading or shadowing.

In fact, because of fading, spreading factors are randomly distributed over the simulated area, which reduces Co-SF interference in small cells.

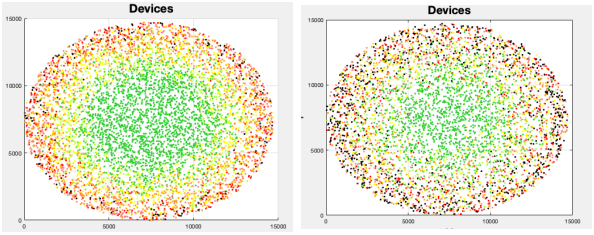

Figure 10: Graphics of SF distribution for Shadowing PM (Right) and Fading and Shadowing PM (Left).

This can be sustained by Figures 9 and 10, where the SF are uniformly distributed in the case of Log-distance propagation and the other cases, with fading or shadowing effect the distribution is not uniform. This means the SFs will have different values and consequently, its orthogonality leads to a decrease in collisions. This results is also related to the way SFs are assigned in the current model, which chooses the lowest SF possible for the expected received power. Different SFs enables simultaneous transmissions at different data rates on the same frequency channel.

8. Conclusions

The major purpose of this study was to implement a LoraWAN network analytical model and optimize the gateways positions using NSGA-II as well as study the impact of fading and shadowing on the network's deployment.

The challenges of the quality of the signal reception led to the development of different propagation models to study the network's deployment.

This study implemented a LoRaWAN analytical model developed by INESC in the WiMeCOM project [3]. The first contributions are proposed with the study of fading and shadowing in the quality performance of the network. Results show that the fading or shadowing effect leads to the decrease of instantaneous interference as well as the packet loss ratio. This has a significant impact on the gateway placement. This effect was validated and considered to be related to the orthogonality of the SFs assigned to the devices which leads to a decrease of collisions.

It must be highlighted that this performance may have different results in other simulation environments. In this study, we consider that transmissions that reach a gateway are independent from those that reach neighbor gateways. However, in real scenarios, the transmission from a node will likely reach more than one gateway, creating inter-dependency between packet vulnerable times. While the inability to consider this statistical dependency reduces the complexity of the WiMeCOM model, it may have a significant impact on the obtained results. such limitation should be corrected in the future. Furthermore, the current model assumes static SF assignment, which is adequate for a log-distance model, but not for a model that integrates fading effects. In a true scenario, the Adaptive Data Rate mechanism of LoRaWAN would respond to changes in the signal power, changing the SF used by each device. However, such dynamic mechanisms are difficult to model analytically, where worst-case or average conditions are usually assumed. Discrete event simulation should be used to find the correct parameters and correct the analytical model.

9. Future work

This research has some directions for future work. Firstly, this study should be developed for a 5G network and analyze if there are the same observations for the effect of fading and shadowing in this type of network.

Additionally, the interference between gateways should be taken into account as this study is considering each GW as an isolated device.

Adaptive optimization of the SFs should also be considered to ensure more realistic results. This development can consume time but the simulations would be more reliable.

Implement Machine Learning techniques to give more immediate answers. For example, use a neural network, that could be trained with the results of the NSGA-II and provide new configurations for the GWs.

References

[1] L. Alliance. *A technical overview of LoRa and LoRaWAN*. 2015.

[2] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

[3] A. M. R. C. Grilo. A moroccan wireless smart metering solution (wimecom). 2019.

[4] B. H. Liu, B. P. Otis, S. Challa, P. Axon, C. Chou, and S. Jha. The impact of fading and shadowing on the network performance of wireless sensor networks. *International Journal of Sensor Networks*, 3(4):211–223, 2008.

[5] J. Miranda, R. Abrishambaf, T. Gomes, P. Gonçalves, J. Cabral, A. Tavares, and J. Monteiro. Path loss exponent analysis in

wireless sensor networks: Experimental evaluation. In *2013 11th IEEE international conference on industrial informatics (INDIN)*, pages 54–58. IEEE, 2013.

[6] H. Mroue, A. Nasser, B. Parrein, S. Hamrioui, E. Mona-Cruz, and G. Rouyer. Analytical and simulation study for lora modulation. In *2018 25th International Conference on Telecommunications (ICT)*, pages 655–659. IEEE, 2018.

[7] R. B. Sørensen, D. M. Kim, J. J. Nielsen, and P. Popovski. Analysis of latency and mac-layer performance for class a lorawan. *IEEE Wireless Communications Letters*, 6(5):566–569, 2017.