Do LoRa Low-Power Wide-Area Networks Scale?

4 authors, including:

Martin Christiaan Bor
Lancaster University
9 PUBLICATIONS **192** CITATIONS

Utz Roedig
Lancaster University
143 PUBLICATIONS **2,013** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project Risk Management in Wireless Sensor Networks View project

Project IoT Infrastructure View project

Do LoRa Low-Power Wide-Area Networks Scale?

Martin Bor, Utz Roedig
Lancaster University, UK
m.bor@lancaster.ac.uk
u.roedig@lancaster.ac.uk

Thiemo Voigt
Uppsala University and SICS,
Sweden
thiemo@sics.se

Juan M. Alonso
Univ. Nac. de Cuyo and Univ.
Nac. de San Luis, Argentina
jmalonso@uncu.edu.ar

ABSTRACT

New Internet of Things (IoT) technologies such as Long Range (LoRa) are emerging which enable power efficient wireless communication over very long distances. Devices typically communicate directly to a sink node which removes the need of constructing and maintaining a complex multi-hop network. Given the fact that a wide area is covered and that all devices communicate directly to a few sink nodes a large number of nodes have to share the communication medium. LoRa provides for this reason a range of communication options (centre frequency, spreading factor, bandwidth, coding rates) from which a transmitter can choose. Many combination settings are orthogonal and provide simultaneous collision free communications. Nevertheless, there is a limit regarding the number of transmitters a LoRa system can support. In this paper we investigate the capacity limits of LoRa networks. Using experiments we develop models describing LoRa communication behaviour. We use these models to parameterise a LoRa simulation to study scalability. Our experiments show that a typical smart city deployment can support 64 nodes per 3.8 ha, which is not sufficient for future IoT deployments. LoRa networks can scale quite well, however, if they use dynamic communication parameter selection and/or multiple sinks. *Please note that this is an updated version of the corresponding MSWiM '16 paper.*[1]

Keywords

LoRa, Low-Power Wide-Area Network, Scalability Analysis

1. INTRODUCTION

Large scale Internet of Things (IoT) installations are becoming a reality and networks are being deployed to realise smart city, intelligent transportation system or environmental monitoring applications. Many of these IoT installations rely on Low-Power Wide-Area Network (LPWAN) technologies. New LPWAN technologies such as Long Range (LoRa) [5], Sigfox [13], RPMA [11] and Weightless [18] are emerging which enable power efficient wireless communication over very long distances. LPWANs generally form one-hop networks where every node can reach directly one (or more) Internet connected sink nodes. Network operators see this as beneficial as constructing and maintaining a multi-hop network can be avoided. However, given the fact that LPWANs cover a wide area and that all devices communicate directly to a few sink nodes a large number of nodes have to share the communication medium. The question arises how many nodes can be operated in the same area without dissatisfying application performance requirements.

In this paper we focus on LoRa as it is the currently most widely deployed emerging LPWAN technology and is considered by a large number of industries as a base for their IoT applications. To be scalable LoRa provides for a range of communication options (carrier frequency, spreading factor, bandwidth and coding rate) from which a transmitter can choose. Many settings are orthogonal and provide simultaneous collision free communications. Nevertheless, there is a limit regarding the number of transmitters a LoRa system can support. In this paper we investigate these capacity limits combining practical experimentation and simulation. The contributions are:

- *LoRa Link Behaviour:* Using practical experiments we develop models describing (i) communication range in dependence of communication settings Spreading Factor (SF) and Bandwidth (BW) and (ii) capture effect of LoRa transmissions depending on transmission timings and power.

- *LoRa Simulator:* We use the insight from our practical experimentation to build the simulator LoRaSim. This purpose built simulation tool captures specific LoRa link behaviour and enables evaluation of large-scale LoRa networks.

- *LoRa Scalability Evaluation:* Using LoRaSim we carry out an evaluation of the scalability of LoRa networks. We show that LoRa does not scale well when using it with static settings and a single sink as typically deployed in Long Range Wide Area Network (LoRaWAN). However, we show that the usage of multiple sinks and dynamic communication parameter settings can produce very scalable solutions.

The next section gives an overview of LoRa. Section 3 describes our experiments to understand LoRa link behaviour. Section 4 describes the simulator LoRaSim and our scalability evaluation of LoRa. Section 5 describes related work and Section 6 concludes the paper.

---

[1] Mariusz Slabicki found a bug in the simulator, whereby previously collided packets would be in some cases marked as 'not collided', thereby overestimating the goodput of the network. The bug was fixed and all figures and numbers have been updated accordingly. The overal conclusions, however, stays the same.

2. LORA

Long Range (LoRa) is a proprietary spread spectrum modulation technique by Semtech. It is a derivative of Chirp Spread Spectrum (CSS) with integrated Forward Error Correction (FEC). Transmissions use a wide band to counter interference and to handle frequency offsets caused by low cost crystals. A LoRa receiver can decode transmissions 19.5 dB below the noise floor, thus, enabling very long communication distances. LoRa key properties are: long range, high robustness, multipath resistance, Doppler resistance and low power. LoRa transceivers available today can operate between 137 MHz to 1020 MHz, and thus can also operate in licensed bands. However, they are often deployed in ISM bands (EU: 868 MHz and 433 MHz, USA: 915 MHz and 433 MHz). The LoRa physical layer may be used with any MAC layer; however, LoRaWAN is the currently proposed MAC which operates in a simple star topology.

2.1 Transmission Options

A typical LoRa radio provides five configuration parameters: Transmission Power (TP), Carrier Frequency (CF), Spreading Factor (SF), Bandwidth (BW) and Coding Rate (CR). Energy consumption, transmission range and resilience to noise is determined by the selection of these parameters.

Transmission Power (TP). TP on a LoRa radio can be adjusted from -4 dBm to 20 dBm, in 1 dB steps, but because of hardware implementation limits, the range is often limited to 2 dBm to 20 dBm. In addition, because of hardware limitations, power levels higher than 17 dBm can only be used on a 1% duty cycle.

Carrier Frequency (CF). CF is the centre frequency that can be programmed in steps of 61 Hz between 137 MHz to 1020 MHz. Depending on the particular LoRa chip, this range may be limited to 860 MHz to 1020 MHz.

Spreading Factor (SF). SF is the ratio between the symbol rate and chip rate. A higher spreading factor increases the Signal to Noise Ratio (SNR), and thus sensitivity and range, but also increases the airtime of the packet. The number of chips per symbol is calculated as 2^{SF}. For example, with an SF of 12 (SF12) 4096 chips/symbol are used. Each increase in SF halves the transmission rate and, hence, doubles transmission duration and ultimately energy consumption. Spreading factor can be selected from 6 to 12. As we have shown in previous work, radio communications with different SF are orthogonal to each other and network separation using different SF is possible [1].

Bandwidth (BW). BW is the width of frequencies in the transmission band. Higher BW gives a higher data rate (thus shorter time on air), but a lower sensitivity (because of integration of additional noise). A lower BW gives a higher sensitivity, but a lower data rate. Lower BW also requires more accurate crystals (less ppm). Data is send out at a chip rate equal to the bandwidth. So, a bandwidth of 125 kHz corresponds to a chip rate of 125 kcps. Although the bandwidth can be selected in a range of 7.8 kHz to 500 kHz, a typical LoRa network operates at either 500 kHz, 250 kHz or 125 kHz. Also, bandwidths lower than 62.5 kHz require a temperature compensated crystal oscillator (TCXO).
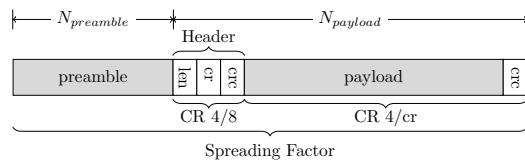
Figure 1: LoRa packet structure. Grey shaded areas are required, white shaded areas are optional.

Coding Rate (CR). CR is the FEC rate used by the LoRa modem that offers protection against bursts of interference, and can be set to either 4/5, 4/6, 4/7 or 4/8. A higher CR offers more protection, but increases time on air. Radios with different CR (and same CF, SF and BW), can still communicate with each other if they use an explicit header, as the CR of the payload is stored in the header of the packet.
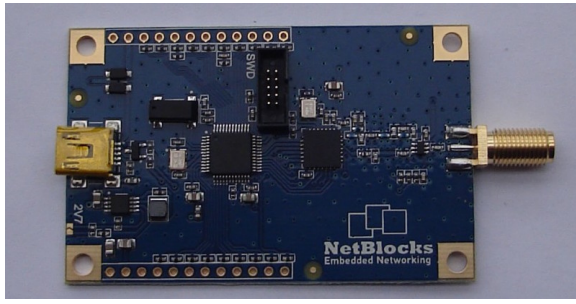
2.2 Transmissions

LoRa Packet Structure. The LoRa packet structure is shown in Figure 1. A packet starts with the preamble, programmable from 6 to 65535 symbols, to which the radio adds 4.25 symbols. Thereafter follows an optional header, which describes the length and FEC rate of the payload, and indicates the presence of an optional 16-bit CRC for the payload. The header is always transmitted with a 4/8 FEC rate, and has its own CRC. After the optional header, there is the payload, which can contain 1 to 255 bytes. At the end of the payload an optional 16-bit CRC may be included.

Airtime. The airtime of a LoRa transmission depends, besides the payload size, on the combination of SF, BW, and CR. The duration of a transmission can be calculated with the Semtech LoRa modem calculator [6]. It has to be noted that depending on the selected communication settings a data packet can have significant variations in airtime. For example, a 20 byte packet can vary between 9 ms and 2.2 s. Thus, the selection of communication parameters has a tremendous impact on scalability of a LoRa deployment.

2.3 Regulatory Constraints

LoRa is classified as a Short Range Device (SRD) and usually operates in license-exempt frequency bands. There are certain restrictions on access to the physical medium, imposed by the regulatory body for the particular region. These limitations have an impact on communication performance and hence have an impact on scalability of LoRa deployments. Scalability is therefore often limited due to regulatory constraints and not due to technical limitations. Next we describe in more detail EU and US regulations; other countries such as China have their own regulations with often are modelled on EU or US standards.

Europe. The constraints in Europe regarding frequency allocation and use for SRD are defined in CEPT/ERC/REC 70–03 [3]. The license-exempt band usable for LoRa (863 MHz to 870 MHz) is referred to as 'Annex 1 h', and is subdivided in 7 (overlapping) subbands. Each subband has specific requirements regarding maximum Effective Radiated Power (ERP), spectrum access and channel spacing. For the majority of the subbands, the ERP is 25 mW (14 dBm). For spectrum access there is the option of either using a

Figure 2: NetBlocks XRange RF module.

duty cycle (often $\leq 0.1\%$) or a Listen Before Talk (LBT) transmission scheme, combined with an Adaptive Frequency Agility (AFA) depending on the specific subband and/or ERP required (see [2, chap. 9] for details).

United States. The Federal Communications Commission (FCC) regulates the use of frequencies for wireless communications in United States. Rules and regulations are stated in Title 47 of the Code of Federal Regulations (CFR). Part 15 (often referred to as 'FCC Rule 15') of this code deals with devices operating in unlicensed frequency bands. The license-exempt band usable for LoRa is 902 MHz to 928 MHz. Compared to the European regulations, the FCC allows a higher peak output power of 1 W (30 dBm), but requires a bandwidth of at least 500 kHz. For lower bandwidths, the device operates in 'hybrid mode', that combines the regulations for digital modulation techniques (like LoRa) with those for Frequency Hopping Spread Spectrum (FHSS). An important limitation for FHSS systems, is the maximum dwell time of 400 ms. This makes the lowest LoRa datarates not usable, as transmitting the preamble already takes longer than 400 ms.

3. LORA LINK BEHAVIOUR

In this section we develop a model of LoRa communication behaviour which we use subsequently in our simulation environment LoRaSim. Specifically, we develop a model describing (i) achievable communication range in dependence of communication settings SF and BW and (ii) capture effect behaviour of LoRa transmissions depending on transmission timings and power. The development process of these models is supported by practical evaluation.

3.1 Experimental Platform

For our experiments we use the XRange SX1272 LoRa module from NetBlocks[1] as shown in Figure 2. The module consists of a low-power STM32L151CC ARM Cortex-M3 microcontroller (32 MHz CPU, 32 kB RAM, 256 kB flash) and a Semtech SX1272 LoRa transceiver. Our models describing link behaviour are validated on this particular platform. However, the used communication chip is the most commonly employed LoRa chip and we believe the used models can be easily tailored to other transceiver types.

3.2 Communication Range

A transmission is successfully received if the received signal power P_{rx} lies above the sensitivity threshold S_{rx} of

---

the receiver. The received signal power P_{rx} depends on the transmit power P_{tx} and all gains and losses along the communication path:

$$P_{rx} = P_{tx} + G_{tx} - L_{tx} - L_{pl} - L_m + G_{rx} - L_{rx} \qquad (1)$$

P_{rx} is the received power in dB, P_{tx} is transmitted power in dB, G_{tx} is the transmitter antenna gain in dBi, L_{tx} is the transmitter loss (RF switch, non-matching circuit, connectors) in dB, L_{pl} is the path loss in dB, L_m are miscellaneous losses (fading margin, other losses) in dB, G_{rx} is the receiver antenna gain in dBi and L_{rx} are receiver losses. For the purpose of this study we simplify this general equation to:

$$P_{rx} = P_{tx} + GL - L_{pl} \qquad (2)$$

Here, GL combines all general gains and losses while L_{pl} represents the path loss, determined by the nature of the communication environment.

On the transmitter side, range can only be changed by changing the transmit power. Other parameters like SF, BW and CR do not influence the radiated power, or any other gains and losses. On the receiver side, the range is limited by the sensitivity threshold S_{rx}, which is influenced by the LoRa parameters SF and BW.

Path Loss. Many models exist to describe path loss in dependence of different environments (built-up area, free space). We use the well known log-distance path loss model [9] which is commonly used to model deployments in built-up and densely populated areas. We choose this model as it matches environments in which we expect LoRa deployments are to be found. Using this model the path loss in dependence of the communication distance d can be described as:

$$L_{pl}(d) = \overline{L_{pl}}(d_0) + 10\gamma \log\left(\frac{d}{d_0}\right) + X_\sigma \qquad (3)$$

where $L_{pl}(d)$ is the path loss in dB, $\overline{L_{pl}}(d_0)$ is the mean path loss at the reference distance d_0, γ is the path loss exponent and $X_\sigma \sim N(0, \sigma^2)$, the normal distribution with zero mean and σ^2 variance to account for shadowing.

The advertised communication range of LoRa is more than 15 km for suburban environments. Petajajarvi et al. [8] have reported a range of 15 km to 30 km in a city, where the receiver was located in a 24 m tall tower and the transmitter was on the roof of a car, and in a boat on open water. Our own experiments with the aforementioned hardware show a range of 2.6 km in rural areas. From our studies [1] in built-up environments we deduce a range of 100 m. This is significantly less than other reported ranges, probably caused by less than ideal indoor deployment, hardware and antennas, and as such represents a worst-case deployment. We also performed all the simulations using parameters reported by Petajajarvi et al. [8], and obtained similar results in terms of scalability.

Obviously, the communication range and, hence, the exact path loss model is highly dependant on the environment and a generic figure cannot be given. However, using our own empirical measurements with d_0 at 40 m, we determined that in the built up environment $\overline{L_{pl}}(d_0)$ is 127.41 dB, γ is 2.08 and σ is 3.57. We use these values in our simulation, but set $\sigma = 0$ since otherwise some transmissions might not be able to reach the sink rendering our results inconclusive.

Table 1: Measured receiver sensitivity in dBm for different bandwidths and spreading factors.

	Bandwidth (kHz)		
SF	125	250	500
7	−126.50	−124.25	−120.75
8	−127.25	−126.75	−124.00
9	−131.25	−128.25	−127.50
10	−132.75	−130.25	−128.75
11	−134.50	−132.75	−128.75
12	−133.25	−132.25	−132.25

Sensitivity. The sensitivity of a radio receiver at room temperature, as found in [12], is given by:

$$S = -174 + 10\log_{10}(BW) + NF + SNR \qquad (4)$$

The first term describes thermal noise in $1\,\mathrm{Hz}$ of bandwidth and can only be influenced by changing the temperature of the receiver. BW is the receiver bandwidth. NF is the receiver noise figure, and fixed for a given hardware implementation. SNR is the signal-to-noise ratio required by the underlying modulation scheme, and is determined by the spreading factor SF. The higher the SF, the higher the SNR.

As BW is set in steps of powers of 2, we can derive from Equation 4 that increasing the bandwidth decreases the sensitivity by $3\,\mathrm{dB}$ and vice versa. Similar for SF, increasing the spreading factor doubles the chips per symbol, which increases the sensitivity by $3\,\mathrm{dB}$.

To determine the receiver sensitivity for our experimental platform, we carry out an experiment using two nodes. Both nodes are placed in different rooms on different floors of an office building. The distance between the nodes is approximately $40\,\mathrm{m}$. One node transmits a fixed number of packets, on all combinations of spreading factor (SF7 to SF12), bandwidth ($125\,\mathrm{kHz}$, $250\,\mathrm{kHz}$, $500\,\mathrm{kHz}$), coding rates (CR 4/5, CR 4/6, CR 4/7 and CR 4/8) and transmit powers ($2\,\mathrm{dBm}$ to $17\,\mathrm{dBm}$). We repeat the measurement over several days and of all the correctly received packets we record the minimal RSSI to determine the sensitivity. The results are shown in Table 1.

As expected, decreasing the bandwidth or increasing the spreading factor does improve sensitivity. The difference between each step, however, is not $3\,\mathrm{dB}$, but more in the range of $0\,\mathrm{dB}$ to $4\,\mathrm{dB}$, and $2\,\mathrm{dB}$ on average. Presumably this is caused by external interference, and hardware limitations other than the radio chip itself. We use these experimental determined values in our simulations.

Summary. Using Equation 2, Equation 3 and Equation 4 we can now estimate if a LoRa transmission will be received or not. The decision regarding transmission reception can be formally described as:

$$R = \begin{cases} 1, & P_{rx} > S_{rx} \\ 0, & else \end{cases} \qquad (5)$$

To determine P_{rx}, $\overline{L_{pl}}$, d_0, γ and σ must be set to parameterise the path loss model and the communication distance d must be known. In our simulations we set these parameters to the values previously described to reflect a built up environment. S_{rx} depends on the selected BW and SF. We use the measured sensitivity as shown in Table 1 in our simulations to determine sensitivity in dependence of BW and SF.

3.3 Collision Behaviour

When two LoRa transmissions overlap at the receiver, there are several conditions which determine whether the receiver can decode, one or two packets, or nothing at all. These conditions are Carrier Frequency (CF), Spreading Factor (SF), power and timing.

Reception Overlap. Packet reception starts at time a and ends at time b. We define reception interval (a_i, b_i) for packet $i \in \mathbb{N}$, that is reception i starts at a_i and ends at b_i. We define the midpoint $m_i = \frac{a_i + b_i}{2}$ and midpoint length $d_i = \frac{b_i - a_i}{2}$. Two packets, x and y, overlap when their reception intervals overlap, that is:

$$O(x, y) = |m_x - m_y| < d_x + d_y \qquad (6)$$

Carrier Frequency. When two transmissions overlap in time, but not in Carrier Frequency (CF), they do not interfere which each other and can both be decoded (assuming a receiver is listening at both carrier frequencies). The overlap in CF is defined as the absolute difference of these frequencies, and the tolerable frequency offset, which depends on the bandwidth. Therefore, we can define the condition when two transmissions collide on CF C_{freq} as:

$$C_{freq}(x, y) = \begin{cases} 1 & \text{if } |f_x - f_y| < f_{threshold} \\ 0 & \text{else} \end{cases} \qquad (7)$$

where f_x and f_y are the centre frequencies of transmission x and y, and $f_{threshold}$ is the minimum tolerable frequency offset. The minimum tolerable frequency offset for the Semtech SX1272 is $60\,\mathrm{kHz}$ for a bandwidth of $125\,\mathrm{kHz}$, $120\,\mathrm{kHz}$ for a bandwidth of $250\,\mathrm{kHz}$ and $240\,\mathrm{kHz}$ for a bandwidth of $500\,\mathrm{kHz}$.

Spreading Factor. The spreading factors (SF) used in LoRa are orthogonal. Transmissions with different SF (and same CF and BW) can thus be successfully decoded (assuming two available receive paths). Therefore, we define the condition on when two receptions collide on SF C_{sf} as:

$$C_{sf} = \begin{cases} 1 & \text{if } SF_x = SF_y \\ 0 & \text{else} \end{cases} \qquad (8)$$

where SF_x and SF_y are the SF of transmission x and y.

Power. As LoRa is a form of frequency modulation, it exhibits the *capture effect*. The capture effect occurs when two signals are present at the receiver and the weaker signal is suppressed by the stronger signal. The difference in received signal strength can therefore be relatively small. When the difference is too small, however, the receiver keeps switching between the two signals, effectively not able to decode either transmission. Therefore, we can define the condition on when packet x collides with packet y on received signal strength as:

$$C_{pwr}(x, y) = \begin{cases} 1 & \text{if } (P_x - P_y) < P_{threshold} \\ 0 & \text{else} \end{cases} \qquad (9)$$
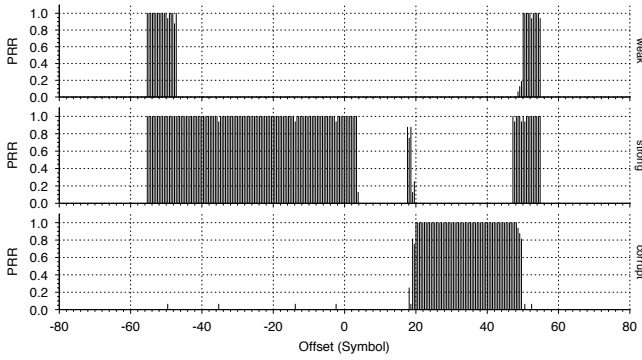
Figure 3: Capture effect. $SF = 12$, $BW = 250\,\text{kHz}$, 55.25 **symbols packet length. X-axis shows the transmission offset relative to the weak node in symbols, Y-axis shows the packet reception rate.**

where P_x is the received signal strength of transmission x and P_y is the received signal strength of transmission y and $P_{threshold}$ is the power threshold.

To verify the capture effect, we set up an experiment with one receiver and two transmitters (an extension of our previous experiments reported in [1]). One transmitter is set to transmit a 32-byte packet on a regular interval at a *weak* transmit power $(2\,\text{dBm})$, while another transmitter was set to transmit a 32-byte packet at a *strong* transmit power $(3\,\text{dBm})$. The strong transmitter sends a number of packets on a particular time offset relative to the weak transmitter, from one packet time early to one packet too late. The strong and weak transmitter are time synchronised by a packet send by the receiver that initiates the experiment. We repeat this experiment for all different combinations of SF, BW and offsets. The results of the experiment with SF12 and BW250 is shown in Figure 3. For all other combinations similar patterns are obtained.

From Figure 3 we can see that a strong transmission can be successfully decoded when it arrives one packet time early up to at most 3 symbols late, successfully suppressing the weak transmission. However, with an offset of more than +3 symbols up to the end of the packet, no transmission gets through. The receiver requires 5 symbols to detect the preamble and synchronise. The transmissions were sent with 8 preamble symbols. Therefore, after 3 symbols, the receiver has locked on to the weak transmission, but its signal is suppressed by the strong transmission and the packet is corrupted.

Timing. From the aforementioned experiments and Figure 3, we can also conclude that packets can overlap, as long there are at least 5 preamble symbols left intact (in case of a *weak* packet). In other words, the critical section of a packet reception starts at the last 5 preamble symbols, so we can redefine the interval for transmission x as $x_{cs} = (a_x + T_{sym} \cdot (N_{pp} - 5), b_x)$, where T_{sym} is the symbol time and N_{pp} is the number of programmed preamble symbols. Therefore, packet x collides with packet y when it overlaps in its critical section x_{cs}:

$$C_{cs}(x,y) = \begin{cases} 1 & \text{if } O(x_{cs}, y) \\ 0 & \text{else} \end{cases} \qquad (10)$$

Summary. When all conditions as defined in Equation 6, Equation 7, Equation 8, Equation 9 and Equation 10 are true, then packet x and y collide:

$$C(x,y) = O(x,y) \wedge C_{freq}(x,y) \wedge C_{sf}(x,y)$$
$$\wedge\, C_{pwr}(x,y) \wedge C_{cs}(x,y) \qquad (11)$$

We use this model of collision behaviour in our simulations.

4. LORA SCALABILITY

We use a simulator to examine and understand scalability of LoRa networks. It is not feasible to evaluate scalability of large-scale LoRa networks in practice as the deployment of such networks would be prohibitively expensive. Furthermore, a real deployment would not allow us to test a larger number of configurations and topologies as is needed for a general study on scalability. However, to ensure our results are of practical relevance we use the aforementioned practical experiments to calibrate our simulation.

4.1 Simulation Framework

For the purpose of this study we developed the simulation tool LoRaSim. The LoRaSim[2] is a custom-build discrete-event simulator implement with SimPy [14]. LoRaSim allows us to place N LoRa nodes in a 2-dimensional space (grid layout or random distribution). M LoRa sinks (the data collection points) can also be placed within the space.

Each LoRa node has a specific communication characteristic defined by the transmission parameters TP, CF, SF, BW and CR. For an experiment, each node's transmission behaviour is described by the average packet transmission rate λ and packet payload B. We assume a preamble length of 8 symbols, so packet airtime for a packet is given by B, SF, BW and CR. The behaviour of a node n during a simulation run is therefore described by the set $SN_n = \{TP, CF, SF, BW, CR, \lambda, B\}$.

Each LoRa sink is able to receive for a given CF multiple signals with different SF and BW combinations. This mimics the behaviour of LoRa sink chips such as the Semtech SX1301 which can receive 8 concurrent signals as long as these signals are orthogonal (i.e. as they are using different SF or BW settings). Two of such chips can be used in a sink node to ensure that concurrent signals on all orthogonal SF and BW settings can be received simultaneously.

The communication behaviour of LoRa nodes can be modelled using the equations for communication range (Equation 5) and collision behaviour (Equation 11). However, the simulator has the ability to replace both models with a simplified variant. The simple variant assumes infinite communication range and any two transmissions overlapping in time at the receiver with the same CF, SF and BW will collide and none of the two transmissions is received. The simple models allows us to establish a baseline which can be analytically described (See Experiment 1).

4.2 Evaluation Metrics

To evaluate scalability and performance of LoRa deployments we define two metrics: Data Extraction Rate (DER) and Network Energy Consumption (NEC).

DER: In an effective LoRa deployment all transmitted messages should be received by the backend system. This means that each transmitted message should be received

Table 2: Parameter setting for Experiment Set 1.

Parameter	Set		
	SN^1	SN^2	SN^3
TP (dBm)	14	14	14
CF (MHz)	868	868	868
SF	12	6	12
BW (kHz)	125	500	125
CR	4/8	4/5	4/5
λ (ms)	1×10^{-6}	1×10^{-6}	1×10^{-6}
B (byte)	20	20	20

Figure 4: Experiment Set 1 − Single Sink: Pure ALOHA and SN^1 (Simple Models) overlap. As the number of nodes increases, the DER decreases exponentially. With typical LoRaWAN settings (SN^3) and a typical $DER > 0.9$ requirement $N = 120$ nodes can be supported.

correctly by at least one LoRa sink. We define the Data Extraction Rate DER as the ratio of received messages to transmitted messages over a period of time. The achievable DER depends on the position, number and behaviour of LoRa nodes and sinks which is defined by N, M and SN. DER is a value between 0 and 1; the closer the value is to 1 the more effective the LoRa deployment is. In a perfect deployment one would expect $DER = 1$. The metric does not capture individual node performance and is a metric looking at the network deployment as a whole.

NEC: The energy consumption of a LoRa node will depend in most scenarios mainly on the energy consumption of the transceiver. As nodes will be deployed in many scenarios on batteries it is essential to keep energy consumption for transmissions to a minimum. Transmit energy consumption for each message depends on transmit power TP and transmission duration which is influenced by SF, BW and CR. We define Network Energy Consumption NEC as the energy spent by the network to successfully extract a message. The NEC depends on the number of nodes, frequency of transmissions and transmitter communication parameters. The lower the metric, the more efficient is the deployment as lifetime of nodes is longer. The energy required to extract a message should be independent of the number of nodes deployed in the network. Again, the metric does not capture individual node behaviour and is a metric looking at the network deployment as a whole.

4.3 Experimental Evaluation

Experiment Set 1 – Single Sink. In our first set of experiments we evaluate the principle capability of LoRa using a simple setup where N nodes transmit to one sink ($M = 1$). We use in these experiments homogeneous transmitter configurations; for an experiment run all nodes use the same configuration set $SN = \{TP, CF, SF, BW, CR, \lambda, B\}$. Nodes are placed randomly around the sink such that all nodes can reach the sink with the given setting SN.

We compare the three transmitter configurations SN^1 and SN^2 and SN^3 (see Table 2). In all settings we assume a 20 byte packet is sent by each node every 16.7 min representing a realistic application. With SN^1 we choose the most robust LoRa transmitter settings leading to transmissions with the longest possible airtime of 1712.13 ms. With SN^2 we choose the transmission setting leading to the shortest airtime of 7.07 ms. With SN^3 we choose the setting use by common LoRaWAN deployments as, for example, one
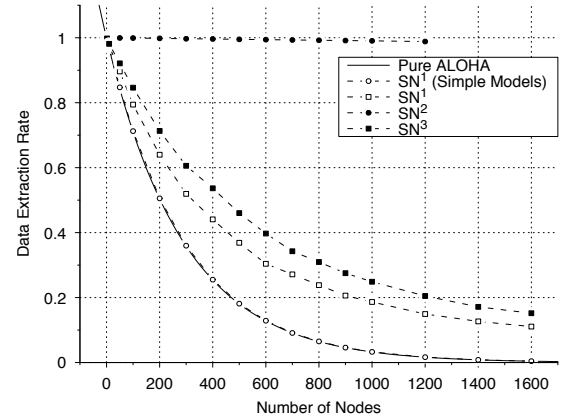
trialled in Amsterdam[3]. We use SN^1 with simple channel models and our LoRa channel models to analyse the impact of these more realistic channel representations. For all subsequent experiments we use the LoRa channel representation.

Figure 4 shows the result of our first set of experiments. Each data point represents a simulation run of approximately 58 days. With an increasing number of nodes the Data Extraction Rate (DER) drops exponentially in all cases. The difference in DER is significant when comparing the configuration with longest (SN^1) and shortest airtime (SN^2). The default LoRaWAN configuration (SN^3) is very close to the configuration with the longest airtime (SN^1). We also observe a significant difference between using simple channel models (SN^1 Simple Models) and the LoRa channel representation (SN^1).

If we would assume that an application requires a $DER > 0.9$ to provide useful functionality we would be able to support $N = 64$ nodes with the default LoRaWAN configuration (SN^3). The modelled communication range here is around 100 m (as observed in our experiments in a built up environment) and we can see that many applications (such as building automation) could not be supported by a LoRa system. It is likely that in such scenarios more nodes would have to be supported within the given range of a sink.

Obviously one could use less conservative transmission settings (the extreme represented by SN^2) to accommodate more nodes. However, in this case the transmission range is reduced and little protection against burst interference is provided. For example, the average transmission range for SN^1 is 98 m compared to 37 m for SN^2.

If we assume our deployment is located in Europe the regulator would require that each node can only use the channel for 0.1% of the time (duty-cycle limitation). For our experiment using the default LoRaWAN configuration (SN^3) we would obtain a channel duty-cycle of 0.13% that is above the regulator allowance. To comply we would need

---

[3]https://thethingsnetwork.org
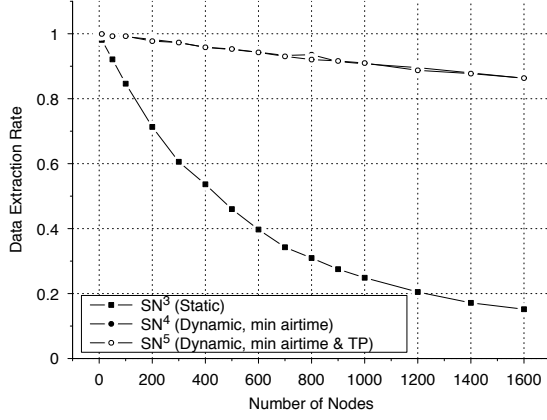
Figure 5: Experiment Set 2 − Dynamic Parameters: Lines for SN^4 and SN^5 overlap. When optimising transmission parameters for minimal airtime (or airtime and power) network capacity greatly improves. With minimised airtime (SN^4) and $DER > 0.9$, well over $N = 1100$ nodes can be supported (compared to $N = 64$ nodes with static settings).

to reduce the data transmission rate from one 20 byte packet every 16.7 min to every 22 min.

For SN^1 Figure 4 shows results using simple channel models and LoRa models. The more realistic channel representation leads to an increase in DER as colliding transmissions may still be received due to the capture effect. For example, for $N = 200$ the DER increases from 0.51 to 0.64. This effect is significant and cannot be neglected when analysing the capacity of a LoRa network.

The setup with simple channel models corresponds to Pure ALOHA [16]. The DER for such systems is:

$$DER = e^{(-2N \cdot T_{packet} \cdot \lambda)} \quad (12)$$

N is the number of transmitters, T_{packet} the packet airtime and λ is the transmission rate of all nodes. Figure 4 shows for SN^1 simulation results together with the analytic solution that match closely. This analytic solution can be used to describe the DER worst-case bound. More realistic channel models always result in a performance boost due to the capture effect. Equation 12 implies a lower DER for larger packets and higher transmission rates. We have verified that this is indeed the case, also in the more complex simulations such as those with multiple sinks.

Experiment Set 2 – Dynamic Parameters. In the second set of experiments we evaluate the impact of dynamic communication parameter selection on Data Extraction Rate (DER) and Network Energy Consumption (NEC). We compare three transmitter configurations SN^3, SN^4 and SN^5. SN^3 is the same as in Experiment Set 1 and is used as reference. For all settings we assume again a 20 byte packet is sent by each node every 16.7 min and CF is 868 MHz. N nodes transmit to a single sink ($M = 1$). Nodes are placed randomly around the sink within a radius that ensures that all nodes can reach the sink if they use the most robust settings. However, for each node the BW, SF, CR are set such that airtime is minimised (setting SN^4 with constant
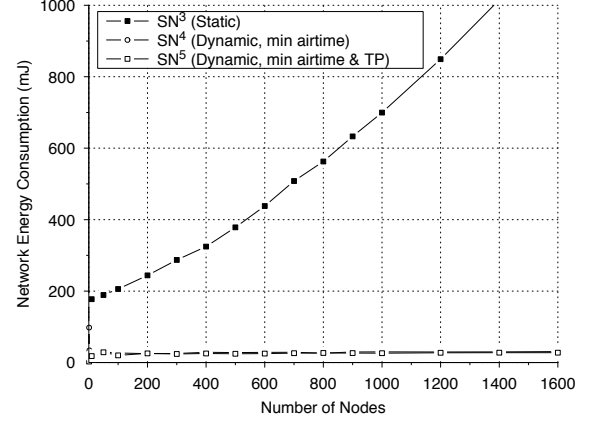
Figure 6: Experiment Set 2 − Dynamic Parameters: Lines for SN^4 and SN^5 overlap. Choosing communication parameters of nodes to minimise airtime (or airtime and power) has a significant impact on energy per extracted packet.

$TP = 14$ dBm) and then such that first airtime and then Transmission Power (TP) is minimised (setting SN^5). For all experiments we use the LoRa channel representation.

As shown in Figure 5 the optimal allocated settings in terms of airtime (and airtime plus TP) has a huge impact on achievable DER. With minimised airtime (SN^4) and a $DER > 0.9$ requirement well over $N = 1100$ nodes can be supported. This is a dramatic improvement compared to $N = 64$ nodes achieved with static conservative settings as used in LoRaWAN.

However, it has to be considered that this achievement is not practical and relies on quite optimistic assumptions. First, the simulation does not consider interference and a minimum airtime setting has a low CR setting which does not provide sufficient protection. Second, the minimum setting would need to be re-evaluated from time to time due to environmental changes. A protocol would need to be used in the LoRa network to determine and adjust the settings. Although LoRaWAN specifies a *Network Manager* component to specifically deal with this issue the implementation and its protocols are not yet defined. Thus, existing LoRaWAN deployments use static and conservative transmission settings represented by SN^3.

Figure 6 shows the impact of optimal allocated settings on Network Energy Consumption (NEC). Obviously, choosing settings with shorter airtime and less TP will not only help to improve DER but helps to achieve significant energy savings. For example, for $N = 200$ energy consumption in the network is reduced by 90%. This in turn translates to a proportional longer node lifetime if they operate on battery. Again, in practice these savings may only be achieved partially due to a lack of mechanisms for transmission setting adaptation and due to other constraints such as interference.

Experiment 3 – Multiple Sinks. We have seen in the previous experiments that LoRa communication settings have a huge impact on network performance. In this set of experiments we explore the impact of the number of sinks M.

We use the previously described setting SN^1 for each experimental run (a 20 byte packet is sent by each node every
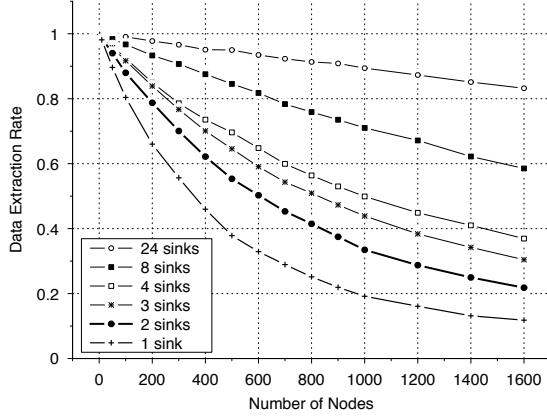
Figure 7: Experiment Set 3 – Multiple Sinks: multiple sink can significantly increase the DER.
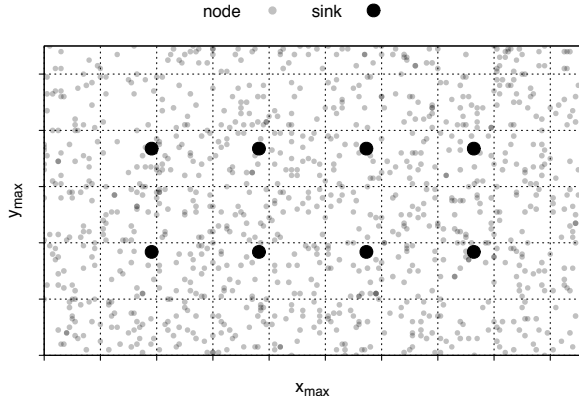
Figure 8: Example of a simulated deployment with 1000 nodes and 8 sinks.

16.7 min and CF is 868 MHz). For each run an increasing number of sinks M is used. The node placement strategy is changed as now multiple sinks are present. Nodes are placed in a rectangle with a diagonal twice the maximum transmission range d_{max}, and side lengths $x_{max} = \sqrt{3} \cdot d_{max}$ and $y_{max} = d_{max}$. This setup ensures that with communication settings SN^1 nodes within this rectangle can reach at least one sink. With four sinks or less, we space them equally over x_{max} on a straight line with $y = y_{max}/2$. Six or eight sinks are equally spaced over the two straight lines at $y = y_{max}/3$ and $y = 2 \cdot y_{max}/3$. 24 sinks are equally spaced over three straight lines at $y = y_{max}/4$ and $y = 2 \cdot y_{max}/4$ and $y = 3 \cdot y_{max}/4$. Figure 8 shows an example deployment with 1000 nodes and eight sinks. Here the bottom left point is placed at $(x_{max}/5, y_{max}/3)$ while the top right point is placed at $(x_{max} \cdot \frac{4}{5}, y_{max} \cdot \frac{2}{3})$. We intentionally chose this sink placement strategy for simplicity rather than optimality. Simulations with different node placements have led to similar results.

Our results in Figure 7 depict that increasing the number of sinks significantly increases the DER. For example, with 200 nodes, one sink is not able to support the typical $DER > 0.9$ requirement while eight sinks achieve this. With 24 sinks, more than 1000 nodes still obey this requirement while with one sink the DER would be as low as 0.19.

Our expectation was that with an increase in the number of sinks the network would get saturated, and the DER would actually decrease. The figure, however, shows that this is not the case. We believe that this is caused by the fact that there only needs to be one sink where the capture effect comes into play in order to ensure that a packet can eventually be received. With more sinks, the chances increase that a packet finds a sink where the capture effect plays for its advantage. With an infinite number of sinks, each node might find such a sink avoiding packet loss.

4.4 Findings

Our experiments lead to a number of findings regarding the scalability of LoRa networks:

- *Lower-Bound on Performance*: Pure ALOHA represents a good Data Extraction Rate (DER) lower-bound in single sink deployments. Equation 12. can be used to quickly estimate expected performance of a typical LoRaWAN deployment.

- *LoRaWAN Scalability*: With typical LoRaWAN settings (SF12, 125 kHz bandwidth, CR 4/5), the assumption of a 20 byte packet is sent by each node every 16.7 min and a $DER > 0.9$ requirement, $N = 64$ nodes can be supported. This is not a sufficient number for applications such as smart city deployments.

- *Dynamic LoRa Settings*: Dynamic allocation of LoRa communication settings has a tremendous impact on network scalability. However, to make use of this potential gain protocols and mechanisms for dynamic parameter selection are required. In LoRaWAN the *Network Manager* is envisioned to fulfil this role but a specification is yet to be given.

- *Capture Effect*: The capture effect has a significant impact on achievable Data Extraction Rate (DER). By far not all colliding transmissions are lost, in many situations at least one of the colliding transmissions can be received successfully. As this effect is significant it has to be taken into account when planning LoRa deployments. It also would have to be taken into account in simulation environments.

- *Multiple Sinks*: Adding additional sinks to a deployment improves Data Extraction Rate (DER). We have not observed that there is an upper bound below 1 in terms of DER when adding additional sinks.

4.5 Future Work

We have not yet investigated the full spectrum of parameters governing scalability of LoRa networks. For example, we have yet to investigate the impact of choosing different Carrier Frequency (CF) in a deployment on scalability. In a deployment area it can also be expected that we find multiple concurrent deployments that may interfere which each other; this aspect has yet to be investigated.

We have shown that dynamic transmission parameter selection and the introduction of more sinks has a dramatic impact on scalability. However, we have not investigated which of the two strategies yields a better return. Dynamic parameter selection requires implementation of complex (and potentially unstable or erroneous) protocols to facilitate this; deployment of multiple sinks is costly and one has to find

out where to deploy sinks best. For LoRa network operators it would be beneficial to determine which route to take.

5. RELATED WORK

There is limited published work discussing scalability of LoRa. Closest to this paper is the work by Petajajarvi et al. [8] and our own previous work reported in [1]. Petajajarvi et al. present an evaluation of LoRa link behaviour in open spaces. We evaluated LoRa link behaviour in built-up environments. We built upon the results reported in these papers when constructing our communication models for LoRaSim (see Section 3). This previous work, however, does not address general scalability questions of LoRa.

A vast number of generic wireless simulation tools such as ns-3 [10] or OMNet++ [17] exist. There are also simulators such as Cooja [7] or TOSSIM [4] designed for Wireless Sensor Networks (WSN) and IoT environments. These simulators can be extended by the components designed for our simulator LoRaSim to enable LoRa simulations.

The Semtech LoRa modem calculator [6] helps with analysis of LoRa transmission features (airtime of packets, receiver sensitivity) but does not enable network planning.

Siradel provides a simulation tool called *S_IOT* [15]. S_IOT relies on Volcano, a 3D-ray tracing propagation model and a portfolio of 2D and 3D geodata. The tool supports sink deployment decisions based on propagation models. This commercial tool considers the environment to a much greater detail than our simulator LoRaSim. However, it does not take into account actual traffic, collisions or details such as capture effect. Our models provided in Section 3 could be used to improve S_IOT.

6. CONCLUSION

Do Long Range (LoRa) Low-Power Wide-Area Network (LPWAN) scale? According to our study presented in the paper current installations based on LoRaWAN do not. The selected conservative transmission settings combined with the fact that in most scenarios only one sink is in range of a node scalability is limited. In a typical scenario, $N = 64$ nodes in an area of 3.8 ha can be handled which is not sufficient for future IoT deployments. However, our study also shows that LoRa networks can scale quite well if they use dynamic transmission parameter selection and/or multiple sinks. For both aspects more work is required as protocols for dynamic transmission parameter selection and strategies for useful sink deployment must be developed.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] M. Bor, J. Vidler, and U. Roedig. LoRa for the Internet of Things. In *Proceedings of the 2016 International Conference on Embedded Wireless Systems and Networks*, EWSN '16, pages 361–366, USA, 2016. Junction Publishing.

[2] Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices (SRD); Radio equipment to be used in the 25 MHz to 1000 MHz frequency range with power levels ranging up to 500 mW; Part 1: Technical characteristics and test methods, May 2012. EN 300 220-1 V2.4.1.

[3] ERC Recommendation 70-03: Relating to the use of Short Range Devices (SRD), Sept. 2015. CEPT/ERC/REC 70-03.

[4] P. Levis, N. Lee, M. Welsh, and D. Culler. Tossim: accurate and scalable simulation of entire tinyos applications. In *Proceedings of the International Conference on Embedded Networked Sensor Systems (ACM SenSys)*, Los Angeles, California, USA, 2003.

[5] LoRa. https://www.lora-alliance.org/What-Is-LoRa/Technology. Accessed: 2015-11-07.

[6] LoRa Calculator. http://www.semtech.com/apps/filedown/down.php?file=SX1272LoRaCalculatorSetup1\%271.zip. Accessed: 20-05-2016.

[7] F. Österlind, A. Dunkels, J. Eriksson, N. Finne, and T. Voigt. Cross-level sensor network simulation with cooja. In *First IEEE International Workshop on Practical Issues in Building Sensor Network Applications (SenseApp 2006)*, Tampa, Florida, USA, Nov. 2006.

[8] J. Petajajarvi, K. Mikhaylov, A. Roivainen, T. Hanninen, and M. Pettissalo. On the coverage of LPWANs: range evaluation and channel attenuation model for LoRa technology. In *ITS Telecommunications (ITST), 2015 14th International Conference on*, pages 55–59, Dec 2015.

[9] T. S. Rappaport et al. *Wireless communications: principles and practice*, volume 2. Prentice Hall PTR New Jersey, 1996.

[10] G. F. Riley and T. R. Henderson. Modeling and tools for network simulation. chapter The ns-3 Network Simulator, pages 15–34. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[11] Ingenu RPMA. http://www.ingenu.com/technology/rpma/. Accessed: 2016-05-09.

[12] Semtech. *AN1200.13 SX1272/3/6/7/8: LoRa Modem – Designer's Guide, Revision 1*, July 2013.

[13] Sigfox. http://www.sigfox.com. Accessed: 2015-11-07.

[14] SimPy – Event discrete simulation for Python. https://simpy.readthedocs.io. Accessed: 24-05-2016.

[15] Siradel S_IoT. http://www.siradel.com/portfolio-item/alliance-lora/. Accessed: 2016-05-29.

[16] A. S. Tanenbaum. *Computer Networks*. Upper Saddle River, NJ, USA, 3rd edition, 1996.

[17] A. Varga et al. The OMNeT++ discrete event simulation system. In *Proceedings of the European simulation multiconference (ESM'2001)*, volume 9, page 65, 2001.

[18] Weightless Open Standard. http://www.weightless.org. Accessed: 2015-11-07.