Abstract—
Currently used Internet of Things (IoT) devices mostly use WLAN, Bluetooth, and Zigbee, limiting the communication range and battery life. LoRaWAN is a new low-power, long-range wireless technology that enables battery-powered IoT devices with a lifetime of more than 10 years. This paper aims to give a good understanding of what LoRaWAN is and how it works. Particular attention is paid to performance in terms of flexibility and power efficiency. The modulation technology is also analyzed in more detail, which enables demodulation with a Signal to Noise Ratio (SNR) of less than -22 dB and a range of up to 15 km. Finally, the biggest problem of LoRaWAN, the scalability problem, is discussed, and current research is presented.

Index Terms—internet of things (iot), modulation, medium access control (mac), aloha, lora, lorawan

1. Introduction

Most of today’s IoT devices use WLAN, Bluetooth, and Zigbee, limiting the communication range to about 100m [1]. Another problem most devices share is their high power consumption, making battery-powered devices last only for a few years at best. Battery-powered long-range IoT devices would enable many new applications, especially in smart agriculture, smart cities, and industry 4.0.

Low Power Wide Area Networks (LPWAN) were introduced to address this issue. LoRaWAN is a relatively new LPWAN and stands for Long Range Wide Area Network [1], [4]. Compared to short-range transmission standards, LoRaWAN is designed for wide-area coverage, low energy consumption, and cost-effective deployment of End Devices (EDs) [2], [3]. There are also other LPWANs like SigFox and Weightless. Still, LoRaWAN is often of greater interest because of its open business model and ability to constantly optimize time on air (ToA), energy consumption, and data rate. This reduces the deployment cost even further and creates a network that can adapt to the environment, while optimizing the battery life of the EDs [2], [4].

This paper presents a profound overview of LoRaWAN and its capabilities. Additionally, we will analyze LoRaWAN’s performance regarding flexibility, energy efficiency, and scalability.

Section 2 gives an introduction to LoRaWAN and its network architecture. LoRaWAN’s modulation technique is explained in Section 3. Section 4 describes the different device classes of LoRaWAN. In Section 5, the performance of LoRaWAN is analyzed. Finally, section 6 presents the scalability problem of LoRaWAN and current research.

2. Introduction to LoRaWAN

Two of LoRaWAN’s key characteristics are high power efficiency and long transmission range. This enables battery-powered devices to last more than 10 years and a range of up to 15 km away from the next LoRa Gateway (GW) [1], [4]. Today, LoRaWAN is used worldwide in applications like reindeer tracking in Finland, smart fire alarms and fire detectors, smart bus schedule signs, and a city-wide network in Canada [5].

2.1. LoRaWAN vs. LoRa

LoRa and LoRaWAN are separate elements of the LoRa network, each associated with a different layer in the protocol stack [5].

LoRa, residing at the physical layer, is a wireless modulation technique that employs a variant of the Chirp Spread Spectrum modulation, providing the long-range communication link between the GWs and the EDs [1], [5].

LoRaWAN builds on top of LoRa. It defines a communication protocol and a system architecture for the LoRa network, specifying the Medium Access Control (MAC). As MAC, the ALOHA principle is used so EDs can initiate uplinks whenever they want, enabling the EDs to go into sleep mode the rest of the time to save power. The LoRa Alliance standardizes LoRaWAN, an open specification in contrast to the proprietary LoRa radio frequency modulation, which Semtech Corporation owns [2], [6], [7]. The protocol works in the unlicensed, worldwide, industrial, scientific, and medical (ISM) bands, so there is no licensing fee, but the devices have to follow the ISM rules [5]. The LoRaWAN protocol also manages routing, access control, and data encryption for EDs [7].

2.2. LoRaWAN Network

A LoRaWAN Network is a star-of-star network consisting of multiple network elements:

- The nodes of the star-of-stars topology are the LoRaWAN EDs, like temperature sensors or actuators such as valves. They are often battery-powered and use the LoRa radio frequency modulation to communicate with GWs [1], [5].
LoRa utilizes a patented modulation technique by Semtech known as Chirp Spread Spectrum. This technique encodes the data signal onto a chirp signal, a tone with a linearly increasing or decreasing frequency over time [1], [3]. The modulation technique spreads the modulated signal over a wide band beyond the original signal’s bandwidth, making the signal less sensitive to selective frequency fluctuations [1], [2]. The modulation technique allows the demodulation of signals for an SNR even below -22dB, a link budget of up to 157dB, and a receiver sensitivity of just -137 dBm [1], [9], [10].

3.1. Modulation

For digital modulation schemes, a finite set of symbols is required. A flexible parameter of LoRa is the spreading factor (SF), which can be seen as the chirp rate and will be analyzed in section 5. LoRa utilizes 2^SF different symbols, with each symbol carrying 2^SF bits of information. In LoRa, a symbol is encoded with a cyclically shifted chirp, so the modulation technique is also called frequency shift chirp modulation [11].

The transmitted waveform \( c(nT_s + kT) \) for a symbol \( s(nT_s) \) in LoRa is defined as:

\[
c(nT_s + kT) = \sqrt{\frac{2}{T_s}} \cdot \exp \left( j2\pi \left[ (s(nT_s) + k) \mod 2^\text{SF} \right] \right)
\]

Here, \( s(nT_s) \) represents the symbol number and \( k \) increments for each sample both take values in \( 0, 1, ..., 2^\text{SF} - 1 \). \( T_s \) is the sample duration, and \( T_s = 2^\text{SF} \cdot T \) is the symbol duration [11]. In the following, \( s(nT_s) = q \) is used for the symbol to be transmitted.

The formula shows that \( q \) can be seen as the starting frequency of the waveform. The modulo operation ensures that when the chirp reaches the end of the bandwidth, it starts again at the beginning, rising until frequency \( q \) is reached again. So, different starting frequencies represent different symbols, as seen in Figure 2.

3.2. Demodulation

Here, the challenge is in identifying the symbol that was transmitted. The idea is to compare the received signal \( r(nT_s + kT) \) with the ideal symbols and pick the one with the highest correlation. Correlation is a mathematical way of measuring the similarity between two signals and is defined as the dot product of the two signals, described with the formula:

\[
\sum_{k=0}^{2^\text{SF}-1} r(nT_s + kT) \sqrt{\frac{2}{T_s}} \exp \left( -j2\pi \left( (q + k) \mod 2^\text{SF} \right) \right) \frac{k}{2^\text{SF}}
\]

(2)

So, to find the correct symbol for a received LoRa symbol, \( 2^\text{SF} \) similarity checks have to be performed. Each similarity check would need \( 2^\text{SF} \) multiplications and \( 2^\text{SF} \) additions, making it very inefficient, especially for higher SFs. To overcome this problem, we use a mathematical trick and add \( +k-k \) to the exponential function, allowing us to separate the \( +k \) term:

\[
\exp \left( -j2\pi \left( (q + k) \mod 2^\text{SF} + k - k \right) \right) \frac{k}{2^\text{SF}}
\]

\[
= \exp \left( -j2\pi \frac{k^2}{2^\text{SF}} \right) \cdot \exp \left( -j2\pi ((q + k) \mod 2^\text{SF} - k) \right) \frac{k}{2^\text{SF}}
\]

\[
= \exp \left( -j2\pi k^2 \frac{1}{2^\text{SF}} \right) \cdot \exp \left( -j2\pi qk \frac{1}{2^\text{SF}} \right)
\]

(3)

In Equation 3 the following observations can be made:

- \( \exp \left( -j2\pi k^2 \frac{1}{2^\text{SF}} \right) \) does not depend on \( q \) anymore and is a simple down chirp.
- \( \exp \left( -j2\pi qk \frac{1}{2^\text{SF}} \right) \) is a pure sinusoidal waveform where the frequency depends on \( q \).

The correlation can now be rewritten as:

\[
\sum_{k=0}^{2^\text{SF}-1} r(nT_s + kT) \sqrt{\frac{2}{T_s}} \exp \left( -j2\pi qk \frac{1}{2^\text{SF}} \right)
\]

(4)

Demodulation of the signal becomes:

1) Multiplying the received signal \( r \) with the down chirp of the base signal, converting the chirp into a single-frequency tone, representing the transmitted symbol.

2) Performing correlation with the bank of frequencies depending on \( q \). This is functionally the same as performing the Fast Fourier Transform and picking the symbol with the highest peak.
This method of demodulating the signal is much more efficient than using the dot product. It’s a key reason why LoRa is known for its power efficiency. [11], [12]

4. Device Classes

A problem with LPWANs is the conflict between lower power consumption and network downlink latency [5]. LoRaWAN addresses this problem by allowing the EDs to act according to one of the three classes: A, B, and C [2].

4.1. Class A

This device class includes all LoRaWAN EDs, as the name implies. So, every ED acts like a Class A device but can implement additional behavior to become a Class B or C device. Class A devices prioritize power efficiency and are, thereby, normally battery-powered.

Following an uplink transmission to the GW, Class A devices open two short downlink windows (RX1 and RX2) before entering a low-power sleep mode [1]. While Class A devices offer good power efficiency, downlink communications from the server must wait until the next uplink from the ED, introducing latency [5]. So this Class is optimal for low-power sensors focused on uplink, which are by now most of all LoRaWAN EDs and are thereby well studied [2].

4.2. Class B

Class B devices, also typically battery-powered, begin as Class A devices and negotiate with the network server to switch to Class B mode [2]. In addition to the uplink transmission and downlink windows RX1 and RX2, Class B devices open extra receive windows called ping slots at scheduled times without significantly increasing power consumption [1], [13]. To do so, the EDs and the GW synchronize their time via beacons broadcasted every 128 seconds from the GW to all EDs in range. This gives the NS opportunities to initiate a downlink, drastically reducing the downlink access delay compared to Class A and providing a perfect solution for actuators or sensors requiring command interventions [2].

To avoid systematic collisions and problems of overhearing messages between Class B EDs, the NS calculates a ping offset for each ED and every beacon period. It is a pseudo-random offset added to the start of the first ping slot in the beacon period, so Class B EDs have their ping slots at different times [2].

Class B devices themselves provide a trade-off between low downlink access delay and low packet loss with lower power consumption. The number of ping slots opened by the end of each beacon period is called the ping number and can be chosen from the ED [2]. A low ping number reduces lost packets and power consumption due to fewer collisions and more sleep time, perfect for ED in large LoRa networks prioritizing power efficiency. A higher ping number enables low access delay while increasing the power consumption, well-suited for ED needing lower latency. However, there is a tipping point for higher ping numbers where the access delay and packet loss ratio rises with the number of EDs due to more collisions caused by the increased network load [2].

Other advantages of Class B devices over Class A devices are that they significantly reduce packet loss of downlink traffic and can receive firmware over the air way more efficiently [2], [13].

Despite potential benefits, real-world implementations of Class B devices remain limited, because there is only an unmaintained buggy version available, necessitating further study and development [2].

4.3. Class C

Unlike A and B, Class C devices listen continuously, sacrificing power efficiency for constant availability. The device’s receive windows remain open until the subsequent uplink transmission, ensuring uninterrupted communication availability. Therefore, the power consumption is relatively high, so Class C EDs are by default not battery-powered [1].

5. Performance

5.1. Time on Air

The ToA is the time it takes to transmit a message between the ED and the GW and is defined by the SF, the bandwidth, and the message size. A longer ToA results in higher power consumption because the ED has to transmit longer. Additionally, a higher ToA and the duty cycle limits of 1% in the ISM band result in a longer block duration, which is the time the ED has to wait until it can transmit again, which is especially critical for the GWs [7], [11], [14]. The MAC layer of LoRaWAN using ALOHA does not play well with a longer ToA because it increases the probability of collisions and retransmissions, resulting in a lower Packet Delivery Ratio (PDR) [2]. All in all, the ToA is a critical factor in the performance of LoRaWAN and should be as short as possible [7].

5.2. Spreading Factor

One parameter that appears all the time is the SF and, thereby, has a significant influence on the performance of LoRa. LoRa defines six orthogonal SFs, enabling simultaneous non-conflicting transmission on the same channel. The SF are defined as SF = \{7, 8, 9, 10, 11, 12\}, resulting in 2SF possible symbols [1], [2].

Increasing the SF by one also halves the rate at which the chirp changes its frequency, resulting in a halved data rate, as can be seen in Equation 1. A lower chirp rate additionally results in a higher receiver sensitivity and is less susceptible to noise and interference [15]. So, increasing the SF results in a higher receiver sensitivity, a higher ToA, and a lower data rate. Comparing SF7 and SF12, SF12 can typically still be demodulated with an SNR of -22 dB, while SF7 requires an SNR of -7 dB. However, SF12 has a ToA of 32 times the ToA of SF7 and a typical data rate of 0.3kbit/s compared to 5.5kbit/s of SF7 [9]. The choice of the SF enables a trade-off between good range/robustness and short ToA/high throughput [3].
5.3. Adaptive Data Rate (ADR)

One problem with LPWANs is the conflict between lower power consumption and wide range. A long transmission range also requires more energy in LoRaWAN, either caused by higher transmission power or a higher SF [7]. LoRaWAN’s solution for this conflict is the ADR algorithm, which the NS performs. ADR tries for every ED to determine the proper communication parameters to enable reliable communication while prioritizing low energy consumption. This is done by dynamically adjusting the SF to the lowest SF possible while still maintaining a stable connection between the GW and the ED. A lower SF reduces the receiver sensitivity and the ToA, so transmissions have a lower power consumption and are less likely to collide [1]. ADR decision is based on the estimated link margin, calculated by measuring the SNR over the last few uplinks [16].

Each ED can decide on its own if it wants to use ADR or can only activate it if it detects transmission problems or deactivate it if the connection to the GW is stable. This enables LoRaWAN networks to adapt to changes in network infrastructure and to varying path loss, which allows EDs like battery-powered GPS trackers [4].

Current research has shown that LoRaWAN’s implementation of ADR is not yet perfect. It was proposed to use a more sophisticated algorithm considering other objectives like scalability and throughput [17]. In [4], ADR was optimized to increase power efficiency by up to 25% and the packet success rate by nearly 7%. ADR can also cause problems even in small LoRaWAN networks where a few EDs have communication problems. Their SF and ToA will rise through the ADR algorithm, increasing communication problems and leading to network degradation, in which, ultimately, everyone uses SF12 [17]. The authors in [17] describe a different SF-management technique to avoid this problem, increasing the PDR up to 470%.

5.4. Energy Consumption

The energy consumption of LoRaWAN ED only depends on the ToA and transmission power, while the sleep current of the microcontroller of the ED can be neglected with $\mu$A at 3.3V in [7]. In [7], it was possible to only use 2.9mJ for a 23-byte transmission with SF7, bandwidth 125kHz, and a transmission power of 3dBm. For comparison, a standard 16850 Li-Ion battery with 3250 mAh has a capacity of about 43200000 mJ. To achieve the impact of the SF on the ToA [7].

- Avoid using Semtech’s PA Boost function, which increases the transmission power, for savings of up to 50% [7].
- Reduce the transmitter supply voltage. In [7], this was possible to lower the voltage from 3.3V to just 1.9V without reducing the transmission power, resulting in additional savings of 55%.
- Use newer generation transmitters [7].
- When higher transmission ranges are needed, the transmission power should be increased before raising the SF because of the significant negative impact of the SF on the ToA [7].
- Use smaller payload sizes and limit the amount of transmissions [1].

This shows that LoRaWAN is very power efficient and can be used for battery-powered devices with a lifetime of more than 10 years and is way more efficient than other standards like WiFi which uses about 90% more energy [1], [7], [18]. Another advantage of the low power consumption of LoRaWAN is that this allows to power the EDs using renewable energy sources, such as solar energy [18].

6. Scalability Problem

Scalability is a crucial aspect of IoT networks, and LoRa faces challenges due to its use of the ALOHA principle [1]. The ALOHA principle is a simple protocol where the EDs can transmit whenever they want, resulting in a high probability of collisions and retransmissions. This is especially a problem for LoRaWAN because of the long ToA and the duty cycle limits of the ISM band, resulting in a long block duration for the EDs and GWs [2].

Studies on the scalability of LoRa networks consistently suggest challenges in scaling, with notable sensitivity to increased network load [1]. This sensitivity manifests in a decrease in the packet delivery ratio (PDR) and an increase in network load due to retransmissions, primarily caused by collisions [2]. Simulations in [19] have shown that only 120 users per antenna can cause a PDR of only 90%. Classical collision avoidance mechanisms commonly used in wireless networks, such as Listen-Before-Talk using Channel Activity Detection (CAD) and closed-loop collision avoidance prove ineffective for LoRa due to specific characteristics, as discussed in [16].

6.1. Challenges in Existing Solutions

Attempts to address scalability challenges have encountered difficulties. Adaptive Data Rate (ADR) exhibits long convergence times, making handling increases in network density impractical. Enabling LoRaWAN’s acknowledgment mode heavily increases network load and decreases the PDR of most nodes [2], [20]. Efforts to adapt well-known carrier sensing approaches for LoRa networks face reliability issues, particularly with Semtech’s CAD, which becomes unreliable at distances less than 400 meters in dense urban environments. Slotted ALOHA or TDMA-like scheduling, while effective in low-density scenarios, struggle to scale due to high synchronization requirements and duty-cycle limitations [21].

6.2. Potential Solutions and Improvements

There are several strategies and improvements to mitigate the scalability challenges in LoRa networks:

- Directional antennas and multiple base stations can be advantageous in reducing communication interference [3].
- A new acknowledgment mode involving acknowledgment messages only for every N-received message and an instant message if a lost packet was
detected. This method was able to increase the PDR but not for all locations [20].

- Peer-to-peer mode, where nodes suffering from low PDR communicate with neighboring nodes for data forwarding, can overcome communication path issues. However, this may increase power consumption and only works if a neighbor in the range has a good connection to a GW [20].

- CANL LoRa, an open-loop collision avoidance mechanism employing a Listen-Before-Talk strategy, outperforms classical carrier sensing approaches in dense LoRa networks, as demonstrated in extensive simulations [16].

7. Conclusion

This paper provided insight into the functionality of LoRaWAN and its performance. LoRaWAN is a powerful LPWAN and stands out with its modulation technique and flexibility, which enables long-range communication with low power consumption, making it especially interesting for battery-powered EDs. LoRaWAN opens up many new possibilities for IoT devices, even if nearly only Class A devices are used today. Class B devices would make LoRaWAN even more attractive for actors and partly solve LoRaWAN’s scalability problem. ADR is a powerful tool to optimize the energy consumption of LoRaWAN EDs, but it is not perfect and can also not solve the scalability problem, which is the biggest problem of LoRaWAN and needs to be solved to enable the full potential of LoRaWAN.

References


