Improving LP-WAN performance in Dense Environments with Practical Directional Clients

Artur Balanuta*†, António Grilo†, Bob Iannucci*‡, Anthony Rowe*§,
*Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, USA
†INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
‡RAI Laboratory LLC, Palo Alto, CA, USA §Bosch Research, Pittsburgh,PA , USA

***Abstract*—Directional antennas are a promising solution for improving the range of client devices and capacity of wireless networks. Unfortunately, in LP-WAN systems directional antennas tend to be both large and expensive due to operating at sub-GHz frequencies. However, if a client device is willing to forgo improvements in antenna gain, it is possible to realize compact and low-cost antennas that provide spatial diversity control. In this paper, we show that by increasing spatial diversity in LP-WAN clients with limited (or no) client gain, we can dramatically increase overall network capacity and improve client battery life by avoiding re-transmissions. This type of directional control can also be used for hot-spot management by more effectively load balancing clients across gateways.**

We performed a sensitivity analysis using a combination of real hardware and simulation to explore the impact of various switchable antenna geometries on network capacity under a variety of deployment configurations. We then designed and evaluated three prototype multi-sector array clients: (1) a switchable patch antenna configuration, (2) a digital phase-shift nulling configuration, and (3) a low-cost switched PCB element phase-shift system. Each design explores a different hardware cost vs antenna beam performance operating point. We experimentally see that our real antenna beam patterns, captured in an anechoic chamber, perform in a similar manner to our simulated prediction models in terms of beam pattern and in simulation improve network capacity by up to 28% from interference isolation alone and up to 95% when offloading hot spots between four gateways. We also perform a small measurement study of how often our final design changes its configuration when deployed over multiple days on a campus testbed.

I. Introduction

Low-Power Wide-Area Networks (LP-WANs) are a promising solution for low data-rate Internet of Things (IoT) applications where base stations cover several kilometers supporting clients with multiple-year battery lives. As these systems are adopted in utility sensing, traffic monitoring, and other urban infrastructure applications, they will face significant challenges in terms of client density and the total number of nodes. One common approach for increasing network capacity is using directional antennas to improve spatial reuse. Directional antennas focus radio frequency (RF) signals, allowing clients to transmit farther, at lower power, and with more stable links. Unfortunately, at the sub-GHz frequencies used in most LP-WAN technologies, for client devices to adopt directionality, it would require large (tens of cm) and expensive antenna arrays.

One often overlooked attribute of directional antennas is that, by focusing the RF energy, the client is naturally reducing interference imparted on other nearby nodes and base stations. In this paper, we explore the impact of increased spatial diversity on client devices in LP-WANs in terms of overall network capacity as opposed to just client gains. Given the nature of LP-WAN networks, where there are thousands of clients per gateway and long transmission distances, we see that reducing interference among neighbors not only has a significant impact on overall system capacity, but also reduces message retries that are costly in terms of battery life. Our key insight is that it is possible to create low-cost and compact directional designs at the expense of gain. Normally, this is a counter intuitive operating point in antenna design for a client, since it requires adding complexity with no increase in range or transmit power. However, in an altruistic LP-WAN context, this capability has a significant impact on overall network performance. We demonstrate that it is possible to create a variety of directional antenna designs that have similar coverage to an omnidirectional antenna while remaining simple and compact. By reducing off-axis interference alone, we can improve overall LP-WAN capacity by as much as 28%. We see that at overloaded (i.e., hotspot) base stations, directional antennas can be used to more effectively shed load to alleviate congestion.

We explore the potential of client-side spatial diversity in LP-WAN systems by creating and profiling real antenna hardware and then simulating its performance at city-scale. As part of this effort, we design and evaluate three generations of antenna beam steering hardware with a software scheme that is protocol compatible with LoRaWAN networks: (1) a (rather expensive) switchable patch antenna array, (2) a digitally controllable phased array for dual antenna nulling, and (3) a compact and low-cost switchable PCB trace delay phase array. Each design trades beam steering accuracy for cost and simplicity.

For our switchable patch array (1), called **DoRa**, we used a miniature (undersized) ground plane to reduce the mechanical size. Normally, the ground plane would need to be larger than 10 cm to operate efficiently, so by using smaller ground planes we trade-off size for antenna gain. Second, we design and evaluate a digitally controllable phase offset front-end (2) as a proof-of-concept, called **NulLoRa v1**, that provides fine-grained control of phase but the hardware is expensive and power-hungry. Finally, we brought all of our design lessons together and built a discrete (as opposed to digitally controllable) phase delay system with RF switches and PCB

delay traces called **NulLoRa v2**. This final approach only adds $1.79 (priced at 1000 units) worth of additional cost over a standard LoRaWAN radio package which often totals around $15 with CPU, radio, antenna, and battery.

In summary, our main contributions are as follows:

1) We design and evaluate compact and low-cost LoRa antenna front-ends that offer the gain of an omnidirectional antenna, with the ability to dramatically reduce neighbor interference and boost network capacity;
2) We provide open-source extensions to an NS3 LoRaWAN simulator to support high-fidelity antenna simulations. We use the simulator to provide a sensitivity analysis of the impact of antenna geometry (i.e. number of switchable sectors) on overall network capacity;
3) Finally, we provide an initial proof-of-concept that directional antennas can aid in hotspot offloading even without increased gain; the proposed algorithm is evaluated in ns-3.

II. Directionality in LoRaWAN

Different LP-WAN implementations have unique architectural trade-offs that impact hardware and deployment costs. The cellular industry uses a licensed spectrum for LP-WAN, thus minimizing interference and ensuring QoS. Other standards such as LoRaWAN and SigFox, operate in an unlicensed ISM spectrum that is often easier to deploy, but its shared nature can lead to performance concerns, especially in densely populated areas. We have already seen problems in systems using other shared ISM bands such as WiFi, but the lower frequency and scale of LP-WAN exacerbate these challenges. Throughout this work, we use LoRaWAN because it is available as an open-source platform that works on the unlicensed spectrum. But the same techniques can be applied independently or in combination with other device-centric networks on the market. A more detailed overview of LoRa and LoRaWAN can be found here [1].

LoRaWAN networks consist of a set of well-provisioned gateways that are usually powered endpoints on a broadband network. These gateways often listen on multiple channels simultaneously and can benefit from high-performance (potentially directional) antennas. Each gateway can support thousands of low-cost battery-operated clients. Clients are low-cost, resource-constrained devices that transmit asynchronously to a selected gateway. A negotiation protocol at startup is used to establish the minimum transmit power and Spreading Factor (SF) for the target gateway, but each message can be overheard by other nearby gateways to improve reliability. The SF controls how much coding is applied to a packet allowing clients to trade off airtime for range and reliability (high SFs transmit farther but take significantly longer). Clients are assumed to use Isotropic antennas with the combination of power control and SF to reduce interference on neighboring gateways. Upstream and downstream traffic uses different channels, so client upstream traffic should not interfere with downstream traffic.

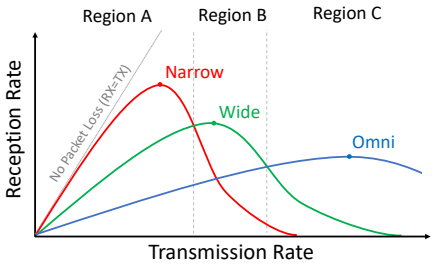


Fig. 1: Illustration of idealized network performance given load with different spatial diversity capabilities.

This architecture leads to the interesting trade-off where reception by multiple gateways increases the reliability of a single packet at the cost of increasing interference globally. In a local sense, transmitting a packet with higher transmit power should increase reliability. In a global sense, high-powered transmissions increase interference with other gateways, which decreases network capacity. This is similar to the principle behind CSMA network capacity [2], except that now it comes in the context of multiple receiver gateways. This scenario becomes more complex when one considers the energy cost of retransmitting lost packets in LoRaWAN. In practice, we see scenarios where the network capacity could be quite high, but nodes are transmitting multiple times which leads to poor energy efficiency. Control over spatial diversity allows clients to go beyond power control and SFs to direct RF energy toward a gateway of their choice. In theory, it is possible to leverage spatial diversity to allow for higher power transmissions at lower SFs without interfering with neighbors.

In this paper, we present two mechanisms for controlling spatial diversity. The first is switchable sector antennas that consist of an RF switch that can cycle through several patch antennas tiled in different directions. The patch antennas slightly overlap in terms of coverage with their adjacent antennas to avoid holes in coverage. In our designs, we assume that, as the number of sectors increases, the width of each sector decreases to increase pointing resolution. The second mechanism we explore is the use of phase-offsets in an array of antennas to programmatically create nulls in the antenna radiation pattern. In this case, the antenna can be configured to adjust where maximal power output should go with deep nulls in off-axis directions. Unlike the switchable patch antennas, nulling antenna configurations tend to be less regular. In Section III, we discuss how we design and evaluate these different configurations.

Figure 1 shows an ideal illustration of how we would expect spatial diversity to impact overall network performance. On the x-axis, we see transmission rate as a measure of load on the network. The y-axis shows the reception rate to indicate how many unique packets are received by the overall network across all gateways (i.e., multiple copies of the same packet only count as a single packet received). In an ideal environment, the transmission rate would match the reception rate shown by the ”No Packet Loss” line. In practice, contention and collisions result in packet loss. We show three different

antenna configurations with their respective curves that define network performance. The peak of each curve can be thought of as the maximum network capacity (peak goodput) for any configuration. The first line (shown in red) defines what we might expect from a network where all the client devices have fine-grained *Narrow* controllable antenna beams. This could be achieved in any number of ways including MIMO beamforming [3], switchable sectors, etc. It is no surprise that a network with highly directional ideal clients has the highest network capacity. If we look at the second *Wide* line (shown in green), we see the performance of the network if the clients had wider beam patterns. Keep in mind that the amplitude and position of these curves are highly dependent on several factors such as antenna gain and side-lobe interference. Generally, we might expect that all things equal the network capacity would be lower compared to the Narrow system. The last line (in blue) represents how the network would respond with omnidirectional antennas on each client.

One interesting aspect to note is that peak capacity might not tell the entire story about what is important for clients. As we can see, there are three regions labeled A through C that indicate different levels of increasing transmission rate. If the transmission rate is increasing, but the peak reception rate is lower, that indicates that packets are getting lost due to contention and require re-transmissions. This implies that each client is expending more energy per unique packet. In this example, the Narrow clients have both a higher peak reception rate and the peak occurring in Region A compared to the Wide clients that have a lower peak that occurs in Region B. It would be possible however for Wide to have a potentially equal or higher peak in Region B compared to the Narrow clients, but that would indicate this is coming at a penalty in terms of client energy (due to retries). Region C shows an example where the omnidirectional clients are still able to get data through compared to the Narrow and Wide clients that have already reached an overload state where most devices are constantly colliding and retrying. Normally this would be resolved with better MAC tuning like in p-persistent CSMA [2], but this is difficult in practice in LoRaWAN networks where each client is device-centric making mostly local decisions without a constant central coordinator.

Through the rest of this paper, we will explore how different mechanisms change the network capacity of various client and gateway configurations. We will highlight the trade-offs in terms of network capacity and average client energy to help inform the influence of various techniques under different scenarios.

III. System Design

In this section, we propose two distinct ways of reducing interference and increasing the overall network capacity of a LoRa network. We validate and optimize our solutions using the NS3 [4] discrete-event network simulator and build hardware prototypes of a directional antenna array (DoRa) and a nulling array front-end (NulLoRa)[1]. We also fine-tune our initial simulation models with real measurements captured in an anechoic chamber and compare their performance at scale under several different gateway and client configurations.

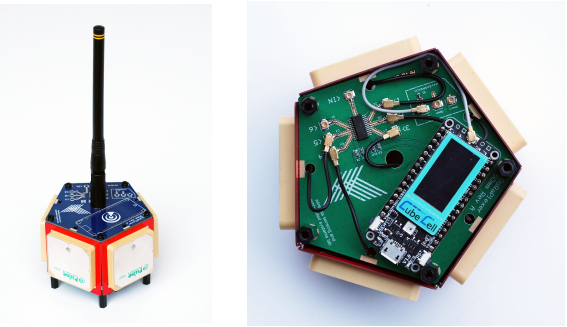


Fig. 2: DoRa client patch array prototype is composed of five PulseLARSEN W3215 Ceramic Patch antenna and a Linx ANT-916-PML whip antenna. Antenna selection is performed by the Analog Devices HMC252A SP6T RF switch using GPIO from a CubeCell HTCC-AB02 LoRaWAN end-device

A. DoRa: Directional Patch Array

Traditional directional antennas come with many benefits (increased gain and directivity), but also have drawbacks, such as the size. This is especially true when operating on sub-GHz bands, where antenna geometries are a function of wavelength. The use of Ceramic as a substrate in a ceramic patch antenna enables it to act as electrically larger when compared to its operating wavelength. Compared to an omnidirectional monopole antenna, the peak gain of a ceramic patch can be as high as 4.5 dBi [5] at sub-GHz bands, but to achieve this theoretical gain, large ground planes are required.

To keep our design compact, we opted to sacrifice the maximum achievable gain while maintaining beamwidth performance. Before designing our prototype, we evaluated the trade-off between the number of sectors and beamwidth, and its impact on network capacity (Section IV-A). Figures 5a and 5b show the estimated radiation pattern of our idealized 0 dB gain 5 patch array and the anechoic measurements from our prototype (shown in Figure 2). The prototype also includes an omnidirectional monopole that can be switched on as a fallback in cases when a deployed protocol doesn't support antenna selection.

We make use of an Analog Devices HMC252A SP6T RF switch to select the active patch antenna. When active, the switch adds 11 mW, which is negligible compared to the RF transmission but does add 0.8 dB of insertion loss. For this prototype, we make use of Pulse LARSEN W3215 ceramic patch antennas. The total Bill-of-Materials cost for the antennas and the switch is about $80 US (clearly cost-prohibitive for production systems). As seen in Figure 5b, the peak gain of each sector varies from 0.3 dB and -1.5 dB in the direction of interest.

[1]We will be releasing both open-source designs with Gerber files on GitHub

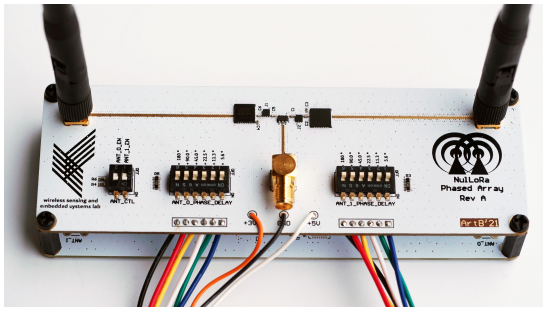


Fig. 3: NulLoRa programmable nulling front-end prototype is composed of a Mini-Circuits BP2C+ Power Splitter/Combiner, two Skyworks SKY13347-360LF RF Switches (diversity control), two Analog Devices HMC936ALP6E 6-bit digital phase shifters and two Linx ANT-916-CW-HWR-SMA antenna

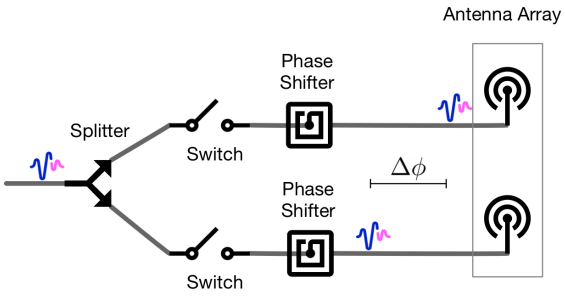


Fig. 4: Block diagram for adaptive steerable nulling.

B. Nulling Prototype

As an alternative to ceramic patch antennas, we explore the idea of using nulling based on a two isotropic antenna array with controllable phase offset between them. A relatively large number of patterns can be achieved by varying the phase and distance between the two omnidirectional antennas. In order to reduce the number of antenna elements and keep the design compact, there is a limitation in directivity that such a system can achieve. Three (or more) antenna elements could have been used to improve gain and directivity but this comes at a significant increase in cost and size. We opted for two Isotropic antennas since this would also enable more traditional diversity through non-phase-shifted antenna selection.

Figure 4 shows a block diagram and the functional parts of our nulling Prototype. It works by splitting/combining the signals between our transceiver and the two antennas while passing through programmable digital phase shifters. Figure 3 shows our Nulling prototype front-end that can accept the signal from any LoRa client via an SMA connector, which then transmits a nulling pattern controlled by some GPIO pins. The prototype board allows us to generate phase-offsets between the two antennas with up to 6° resolution and a range from -354° to 354°. We used the Skyworks SKY13347-360LF as an RF switch along with HMC936ALP6E chips to allow for fine-grained digital control of the phase going into each antenna. Though not ideal in a production system, this front-end allowed us to easily experiment with different nulling patterns. Each digital control pair costs $85 US, consumes 3mW and adds 5.0dB of insertion loss.

Even though we can generate a larger number of radiation patterns, in practice a limited number of patterns that can replicate the effective cumulative gain of Isotropic antenna equivalent is sufficient. We used the MATLAB Antenna Toolbox [6] to derive the separation distance between antennas that minimizes the number of patterns, as well as maximizes their nulling regions. Figure 5c shows the derived four nulling patterns by separating the two antennas by $0.37\lambda \simeq 12$ cm at -90°, 0°, 90° and 180° of phase-offsets. Our hardware prototype was also configured based on this configuration, and the radiation pattern of each phase offset is shown in Figure 5d.

Our NulLoRa nulling front-end also includes a pair of RF switches that can be used to disable one of the antennas. This allows to selectively choose the best antenna in a particular RF environment, which is also known as antenna Diversity. We compare the performance gains of antenna diversity with nulling in Section IV-B.

C. NulLoRa: Passive Nulling Antenna Array

We believe that cost plays a major factor in Internet of Things (IoT) hardware adoption. Thus, we have developed a passive version of the Nulling Antenna Array. As shown in Figure 6, we have replaced the expensive phase shifters used in the prototype with discrete trace delays and SP3T RF switches. NulLoRa adds 1.1 mW of energy consumption when active (LoRa Radio consumes 120-350 mW in TX mode), and adds 3.5 dB of insertion loss (which gets offset by the gain in the pattern beam forming). The passive Nulling Antenna Array design can be added to any commercial LoRaWAN design for as little as $2.96 in parts.

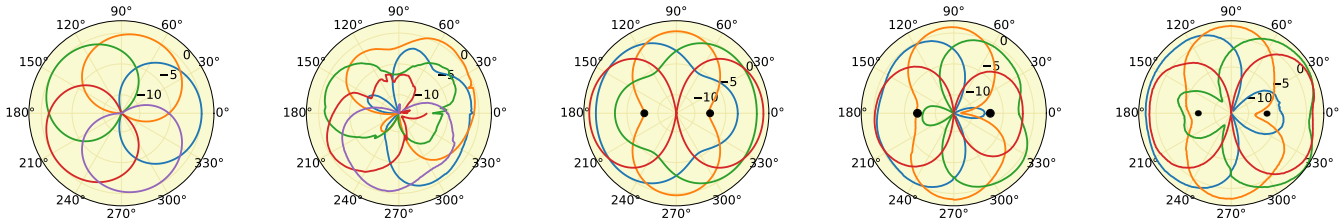
D. Anechoic Chamber Experiments

We performed back-validation of our prototype designs against our simulations and commercially available hardware in an anechoic chamber. As shown in Figure 7, the device-under-test (DUT) is placed on a computerized rotary table that measures the received signal strength under different incident angles (0° to 359° azimuth and -45° to 45° elevation at 1° increments) and across the US915 ISM band. For the simplicity of data interpretation in this paper, simulation antenna models and anechoic measurements use the values at the interception of the horizontal plane (0° elevation) and consider the average gain across the US915 ISM band.

When averaged across the US915 band, and due to hardware imperfections, we can observe the radiation patterns of our patch array (Figure 5b) and nulling array (Figure 5d) have the same effective gain as an omnidirectional antenna with 0dB gain (we use a Linx ANT-916-CW-HWR-SMA Whip antenna as our benchmark).

E. Client Antenna Selection Procedure

Client devices periodically scan the status of the surrounding gateways, by cycling through a ping message on each antenna. We extend the current LoRaWAN association protocol used for Adaptive Data Rate (ADR) to include cycling through multiple antennas (similar to mark and sweep algorithms).



(a) Ideal five sector patch antenna with a beamwidth of 110 deg and 0 dB gain.

(b) Measurements from our five sector directional patch antenna prototype.

(c) Select patterns used in our two antenna nulling-array simulations.

(d) Measurements from our two antenna nulling-array prototype.

(e) Measurements from our two antenna passive nulling-array.

Fig. 5: Antenna gain (in dBi) radiation patterns used in our simulations, derived from mathematical models (a, c) and compared to the measurements taken in an anechoic chamber (c, d, e). Black dots represent monopole locations.

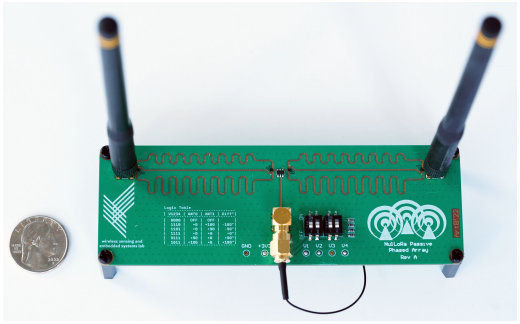


Fig. 6: Passive version of the NulLoRa programmable nulling front-end, composed of a Mini-Circuits BP2C+ Power Splitter/Combiner and four Infineon BGS13S4N9 RF switches

When a client is cycling through antennas, each gateway keeps track of the mean packet reception ratio (PRR) and RSSI of a sequence of messages for each client. The closest gateway coordinates with other nearby gateways and returns the average PRR and RSSI values to the client in an appended downstream message. This coordination assumes an interconnect between nearby gateways which is not uncommon given LoRaWAN's MAC in the cloud architecture. Gateways that have low PRR or RSSI values do not send their information to clients. The device can then use the set of gateways, PRR, and RSSI values to switch to the antenna where the highest PRR was achieved that was above a certain minimum RSSI threshold. The scan operation can also be triggered if the device stops receiving any response in the currently selected antenna, or if the RSSI reported by the gateway drops below the minimum RSSI threshold for a number of sequential packets. The final association decision can be made by the client or pushed down by the gateway depending on the architecture. Standard LoRaWAN systems would allow the client to decide while systems that employ hotspot offloading might leverage more information from the network. This protocol can easily be implemented as an application layer scheme in the context of LoRaWAN.

F. Hotspot Offloading Algorithm

Besides interference isolation, we can perform load balancing by directing transmissions away from saturated gateways. The challenge is to make sure that the underlying signaling is scalable and converges in a manner where nodes don't oscillate from one gateway to another unnecessarily.

When a new endpoint joins the network, it preforms a scanning routine where it sends one SF10 packet through each one of its antenna patterns that is then received by the surrounding gateways. We record the used antenna pattern with the received Received Signal Strength Indicators (RSSIs) and save them in a Cloud database. We can use this information along with our centralized algorithm to determine the endpoint allocation.

The algorithm starts with the endpoints in the default greedy allocation (where each is assigned to the closest gateway). We then compute the Airtime metric at each gateway. We define airtime as the total time required for each of the assigned endpoints to send one packet to their assigned gateway. This takes into account many parameters, such as the SF and transmission power required (which can be deduced from the scanning procedure).

We define the balance of the network as the difference between the gateway with the most and the least airtime. As a first step the algorithm migrates endpoints that require the most airtime (SF10) from the most loaded gateway each turn. We do this, until we reach our predefined unbalance limit or we ran out of candidates for migration. In some migrations there is a need to use a more robust modulation (i.e. SF7 to SF8), we allow this to happen only once per endpoint.

As a second step, the algorithm looks at optimizing the average network RSSI. Keeping the same airtime this implies exchanging neighboring endpoints with the same SF between two or more gateways. Since this can result in cycles and false true minimums we employ randomization in the migration order over multiple rounds.

A practical realization of hotspot offloading would have to be distributed, for sake of scalability. When nodes are performing their association operation, as described in Section III-E, gateways can pass an indicator of their current capacity (e.g., based on the ToA metric). Each node can then decide to switch to a gateway with additional slack capacity even if it is not the strongest in terms of signal quality. Such distributed schemes will be explored in future work.

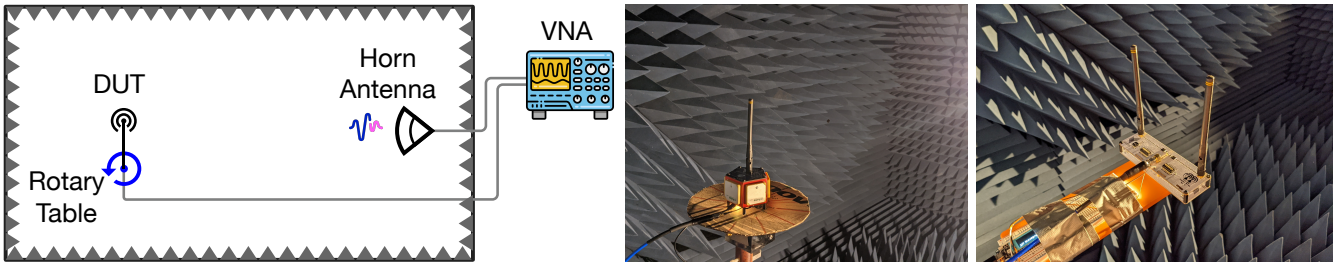
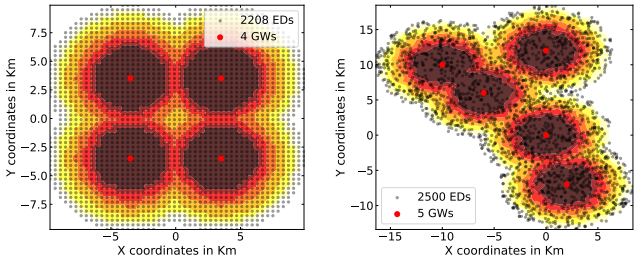


Fig. 7: Anechoic chamber experimental setup (left), DoRa five-sector board (middle) and Nulling front-end (right)



(a) Uniform layout (b) Random layout

Fig. 8: Gateway and Client placements used in NS3 simulations. The gradient shows range and data rate (SF7-SF12)

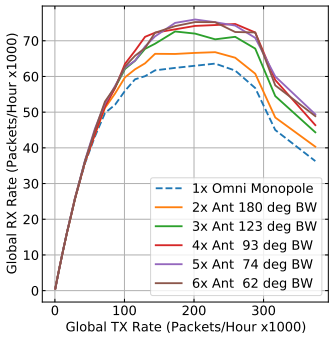


Fig. 9: Impact of switchable sectors on network capacity in uniform-gateway topology with random starting orientations.

IV. Evaluation

To evaluate network capacity, we built upon the NS3 [4], [7] LoRaWAN module [8]. We have added high-fidelity support for directional antennas, antenna arrays, and arbitrary antenna beam patterns that can be dynamically configured and adjusted at run-time[2]. These beam patterns are either based on ideal mathematical models of antennas or recorded radiation patterns from our antennas that we collected in the anechoic chamber. At the start of each simulation, each client performs an antenna gateway selection protocol searching for the antenna elements that yield the highest RSSI. In each test, we place clients at random locations and with random orientations. Clients slowly increase the rate of their traffic, which increases the Global Transmission Rate (TX) load found in many of our performance plots. The Global RX packets give an indicator for unique packets received by the network of gateways (duplicates are ignored). Figure 8 shows an example of two topologies used in our simulations. The first shows four uniformly spaced gateways where the color corresponds to different SF / RSSI levels, with 2208 clients placed on a grid. The second shows a more random gateway layout with 2500 clients also randomly placed within the gateway reach. Unless specified otherwise, we use these typologies with traffic rates scaled up globally to nearly 800,000 packets per hour (222 per second) across the 4 or 5 gateways. Each packet was fixed to 23 bytes long, which corresponds to between $77.1 - 493.6$ ms of on-air time, depending on the required SF. Each experiment measures between 2,500 and 1.6M packets per test point, depending on the rate and we average the results over 5 test runs. Standard deviation is shown in some plots as shaded regions around each line.

[2]We plan to release the simulator as open-source software on GitHub

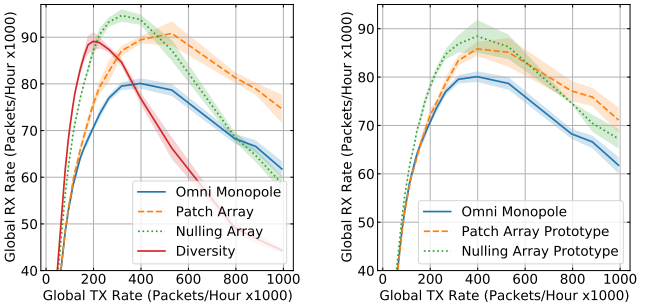
Since our focus is on evaluating how spatial diversity impacts a network as it scales, we were mostly confined to simulation experiments. However, it is worth mentioning that both of our DoRa and NulLoRa boards worked well as standalone LoRaWAN clients on our campus network. Due to the high complexity, and a large number of external factors we attempt to isolate each advantage of directional antennas and test their performance separately in the following sections.

A. Number of Antennas and their Beamwidth

First, we answer the question of how many sectors an end node needs before it gets a diminishing return in terms of overall network capacity. when using a directional patch antenna. Figure 9 shows the network capacity of the network as a function of the increasing number of sectors. As the number of sectors increases, we decrease the effective needed beamwidth of each sector ensuring a 10° overlap in coverage. The lowest capacity is the default omnidirectional antenna (at the bottom), with 3 or 4 sectors reaching a similar maximum network capacity of up to 19% more than the baseline. In practice, we chose five sectors because they appeared slightly more stable in terms of performance, and mapped nicely to commercially available patch antennas.

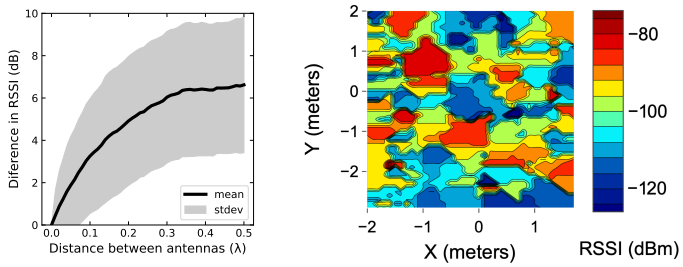
B. Antenna Diversity Gains

With our nulling prototype (NulLoRa v1 and v2), we also have the option of simply selecting one antenna or another to benefit from a small amount of spatial diversity even with an isotropic antenna. These gains result from the antenna moving from a peak or trough created by fading or shadowing in the environment. In prior LoRaWAN deployments, we had observed differences in signal strength as much as 19 dB from



(a) Simulation Results (b) Prototype Hardware Results

Fig. 10: Network capacity given a uniform gateway and client topology (4GWs 2208 EDs).



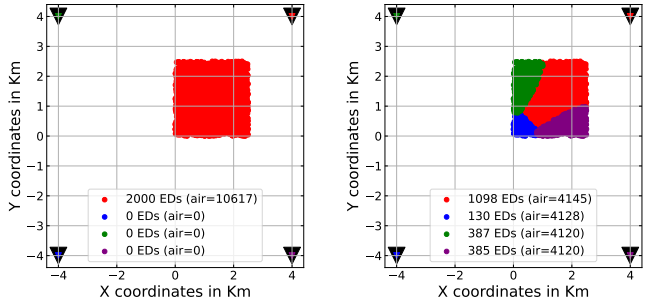
(a) Gains vs Distance (b) Example Location RSSI Heatmap

Fig. 11: Antenna Diversity Gains

relatively small location changes within campus buildings (Figure 11a). In order to better characterize the potential gain relative to antenna distance, we captured fine-grained maps of RSSI at 2 cm increments across 4×4 meters spaces in 6 locations (Figure 11b). For each map, we collected 40 thousand measurements uplinks across 4 GWs located on rooftops distributed across campus. We repeated this collection process at different times of the day to get a sense of the signal stability. As one might expect, during quiet periods like the middle of the night, the signals were reasonably stable, but they can change dramatically during the day.

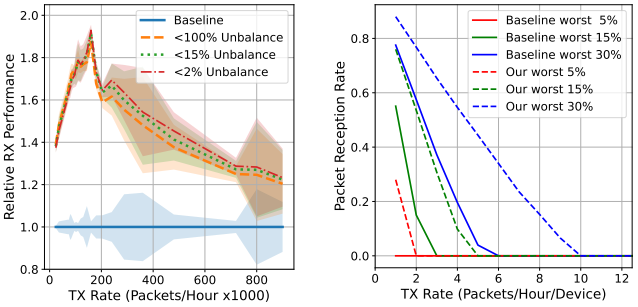
C. Isolation Capacity Gains

Next, we evaluate the end-to-end network performance given our baseline omnidirectional antenna, our five-sector patch antenna, our quad pattern nulling antenna, and an omnidirectional antenna with an average 6 dB of gain to represent what could be achieved through antenna diversity selection. Figure 10a shows the performance of ideal simulated antennas of each class in our uniform environment. We see that Diversity (selecting one of two antennas with a 0.35λ separation) provides a 13% boost over the omnidirectional baseline. This peak occurs early, which implies that, as load increases, there are likely many packets lost in retransmissions. This is to be expected, given that each node generates a large amount of interference with a strong signal arriving at the target gateway. We see that our patch design and the nulling design provide a 13% and 20% increase in capacity, respectively, over the omnidirectional baseline. Nulling has the highest capacity in this configuration by quite a large margin. When comparing plain Diversity to the Patch sector array, we see similar peak capacity, but the Patch array seems to be



(a) Default Greedy Allocation (b) Balanced Airtime Allocation

Fig. 12: Comparing End-device gateway allocation with and without Load Balancing Suggestions



(a) Performance Gains (b) Worst Performers

Fig. 13: Performance of Simple End-device gateway allocation with and without Load Balancing Suggestions

losing more packets due to collisions. We attribute this to the natural fact that even though Diversity is boosting power at one gateway, it is possible that the packet is not being received as strongly by other gateways. With the sectored antenna, the energy is focused.

Figure 10b shows the performance of our four measured hardware antenna patterns. We get slightly smaller gains of between 10% and 15% with less difference in practice between the sector and null hardware options. All things equal, the nulling antenna could be made more cheaply and provides the flexibility to also leverage standard antenna selection diversity.

We see that the performance gains seem consistent compared to the uniform topology. Note that the overall throughput is generally almost 10% higher, but there is an additional gateway, which indicates that per gateway performance has decreased. This can be attributed to imbalanced load and ideally, could be further improved with a hotspot load-balancing algorithm (which was not applied in this case).

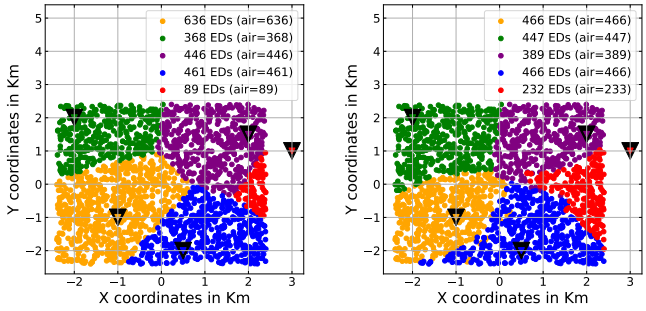
D. HotSpot Offloading Performance

With information about the traffic load at each gateway, it is possible to use our various hardware platforms to redirect traffic to alleviate hotspot gateways. As a simple proof-of-concept demonstration, we set up a simulation topology where we place 2000 client end-devices (EDs) close to one of the four gateways to simulate a dense concentration of nodes (HotSpot). Though not formally defined in the specification, LoRaWAN normally uses a greedy approach where nodes will associate with the gateway that has the highest signal strength.

Figure 12a shows this antagonistic topology where, by default, all of the clients would associate with the closer gateway (top right red triangle). This case might look contrived, but it captures a common deployment scenario where many devices are placed around a single gateway, for example, sensors in a large skyscraper. Other gateways nearby may have significantly less load and ideally should help in supporting additional traffic. We make use of the NulLoRa steering capabilities in order to better distribute the traffic across the available gateways. The distribution of clients between the Gateways is accomplished by our Load Balancing Algorithm formally described in Section III-F and the balanced allocation is shown in Figure 12b. We observe that the allocation of devices between gateways is proportional to the distance between the clients and their target gateway. This is because LoRa achieves a longer range by using higher Spreading Factors (SFs) (and consequently lengthens the transmission time of a message). Figure 13a demonstrates the performance gains associated with performing the allocation. As shown in the figure we compare the baseline greedy allocation with three degrees of *unbalance*. Unbalance is defined as the difference in airtime between the most and the least loaded gateway (i.e. 100% Unbalance represents double the traffic). We can observe a relative improvement in the capacity of up to 90%. We also want to demonstrate that is in the interest of the clients to participate in the load balancing of the network, even if this requires that some nodes use a slower transmission speed and make longer use of the spectrum (as shown by the cumulative air times in Figure 12). The directionality of the patterns isolates the interference and allows for better spatial reuse of the spectrum. Figure 13b shows the packet reception rate of the worst performing 5%, 15%, and 30% nodes in the network. By distributing the load across the gateways we minimize the number of collisions, which translates into a better packet reception rate. Figure 14 shows a more common device placement scenario. In this allocation, we have five gateways (triangles) and 2000 devices randomly distributed across a 5×5 Km space. In this allocation, the Gateways are more densely packed and the default allocation does a better job of distributing the nodes. As shown in Figure 15b we are still able to achieve up to 15% improvement over the baseline allocation. Figure 15b demonstrates that we can achieve this without deteriorating the performance of our worst-performing devices, (i.e. all devices benefit regardless of their distance to the gateway). Thus it is fair to state that is in the interest of all devices to cooperate for a balanced network.

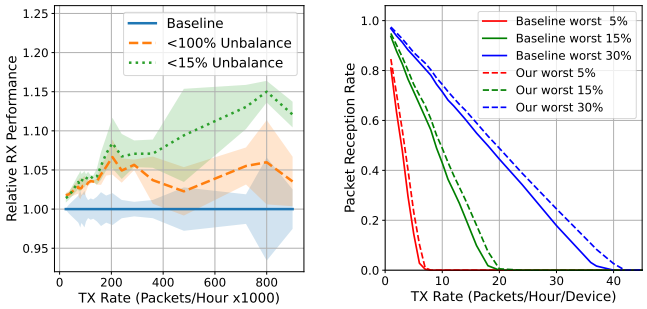
E. Load Balancing Performance in Practice

We evaluate the effectiveness of NulLoRa (v2) on a 10 sq. km. test-bed in a major U.S. city across our university campus network, and inside large office buildings over the course of 28 days in 12 different locations. In order to better characterize the performance of the NulLoRa Hardware over time and in a dynamic environment, we sampled the performance of all of its patterns every 5 minutes. Figure 16 shows the changes in RSSI of our two best performing patterns (from the four available)



(a) Default Greedy Allocation (b) Performance Results

Fig. 14: Common End-device gateway allocation with and without Load Balancing Suggestions



(a) Performance Gains (b) Worst Performers

Fig. 15: Performance of Common End-device gateway allocation with and without Load Balancing Suggestions

of one of such locations. We can see that during the day there are multiple changes in the best-performing pattern. We have observed that these changes in RSSI are highly related to the dynamics of the environment, such as large reflective objects near the source or sink of the signal. Furthermore, the variability in signal strength is slowed during the weekends. We observed up to 8.5 dB difference between the two patterns during this day. The dynamism of the environment can define the frequency we need to re-tune the best pattern and the search scope. For this particular location, Patterns 2 and 3 were always the two best performing patterns for the duration of the experiment (28 days). Most of the other 12 locations demonstrated the same proprieties, except for one location where a third pattern was also optimal for a small portion (3.5%) of the total time.

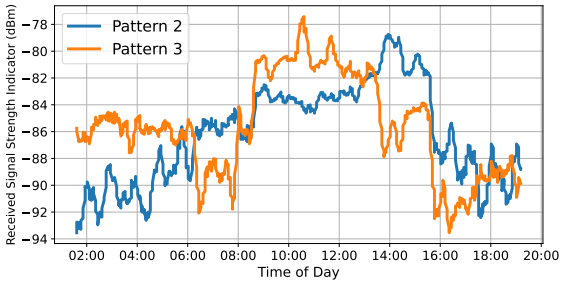


Fig. 16: Changes in the RSSI of the two best performing patterns over the course of a day. End-device is located inside a building office without direct Line-of-Sight to a Gateway.

V. Related Work

There has been a significant body of work around improving the performance of LP-WANs, in particular LoRaWAN at multiple layers of the network stack [9]–[13]. Some looked at combining multiple signal versions to improve the SNR [14]–[16]. Others exploit SF orthogonality to increase network capacity by performing adequate SF assignment [17], [18].

LPWANs typically rely on a simple hardware design – they rely on omnidirectional broadcasting to reach different gateways for sake of diversity – the use of directional antennas, being an apparent contradiction, deserves a careful evaluation.

[19] studied the effects of SPIDA antennas when mounted on the end devices. The study concludes that increasing the number of gateways leads to a more significant performance improvement than using directional antennas. However, no exhaustive evaluation is made of the impact of directional antennas, namely for interference mitigation.

By avoiding signal propagation towards null directions, directional antennas can also be used to mitigate interference and increase spatial reuse, which has been researched in the context of infrastructure WLANs. In [20], the authors propose a Multi-Channel Multi-Sector Directional Antenna WLAN. This scheme integrates multi-sector switched beam antennas in user terminals, a TDMA MAC, and a centralized scheduling algorithm to provide load balancing in infrastructure WLANs. The scheduling mixed-integer optimization problem targets the minimization of overall transmission time.

Regarding the use of directional antennas in LPWAN en devices, to the best of our knowledge, this is the first study specifically focused on this topic.

VI. Conclusions

This paper provides evidence that spatial diversity control, even without significant gain over omnidirectional antennas, can significantly increase LP-WAN network performance. This not only improves the number of clients, and volume of traffic, but also has implications on increasing battery life. The key insight in this work is that even at sub-GHz frequencies, it is possible to create simple and low-cost spatial diversity management hardware, which, when considered at scale, provides significant benefit to the overall system.

We first perform a design exploration in a simulation that looks at the impact of the number of sectors and sector beam width on overall network capacity. Using this simulation as a guideline, we design a five-sectored antenna with miniature patch antennas and an RF switch. Through testing in an anechoic chamber, we capture the hardware's real beam pattern that we then feedback into the simulator for comparison. We see that a five-sector design is between 13% and 20% better than an omnidirectional antenna in terms of overall capacity. Since patch antenna systems are still expensive and bulky, we prototype a phase-switched nulling front-end with two offset Isotropic antennas as a simpler alternative design. Based on testing in the anechoic chamber, we see that the nulling antenna does not as closely match our idealized simulated radiation pattern, but still performs quite well in network-wide simulations. This performance stems from the fact that, even though nulling patterns can be less regular, if the client can detect the strongest main direction with deep nulls, the interference is reduced on neighboring gateways. We see that the less expensive nulling design is between 16% and 28% better compared to our baseline omnidirectional design. Finally, we show that, with network coordination, it is possible to perform load shedding from gateways to offload hotspots to dramatically increase capacity around bottlenecks (95% and beyond in ideal situations).

As future work, we plan to extend this study considering the 3D radiation patterns. Moreover, we only show a teaser of how hotspot offloading is possible, but we do not actually dive into how to make a stable online algorithm. We believe that simple heuristics or more complex techniques from game theory could easily be applied. Finally, we plan to develop a gateway selection and configuration step that adjust to the network over time, which is a natural extension to LoRaWAN's Adaptive Data Rate (ADR) configuration mechanism.

Acknowledgment

This work was partially supported by Portuguese national funds through FCT, Fundação para a Ciência e a Tecnologia, under project UIDB/50021/2020.

References

[1] J. Haxhibeqiri, E. De Poorter, I. Moerman, and J. Hoebeke, "A survey of lorawan for iot: From technology to application," *Sensors*, 2018.
[2] H. Takagi, L. Kleinrock, and F. Ieee, "Throughput analysis for persistent csma systems," *IEEE Transactions on Communications*, vol. 33, 1985.
[3] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive mimo: Benefits and challenges," *IEEE journal of selected topics in signal processing*, vol. 8, no. 5, pp. 742–758, 2014.
[4] G. F. Riley and T. R. Henderson, "The ns-3 network simulator." in *Modeling and Tools for Network Simulation*, K. Wehrle, M. Günes, and J. Gross, Eds. Springer, 2010.
[5] (2021) Ceramic patch antenna w3215: www.pulseelectronics.com.
[6] (2021) Antenna toolbox: www.mathworks.com/products/antenna.html.
[7] J. M. Marais, A. M. Abu-Mahfouz, and G. P. Hancke, "A review of lorawan simulators: Design requirements and limitations," in *IMITEC'19*, 2019.
[8] D. Magrin, M. Capuzzo, and A. Zanella, "A thorough study of lorawan performance under different parameter settings," *IEEE Internet of Things Journal*, 2020.
[9] W. Gao, W. Du, Z. Zhao, G. Min, and M. Singhal, "Towards energy-fairness in lora networks," in *ICDCS'19*. IEEE, 2019, pp. 788–798.
[10] Z. Qin and J. A. McCann, "Resource Efficiency in Low-Power Wide-Area Networks for IoT Applications," in *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, Dec 2017, pp. 1–7.
[11] M. Rahman and A. Saifullah, "Integrating Low-Power Wide-Area Networks in White Spaces," in *IoTDI'18*, April 2018, pp. 255–260.
[12] A. Duda and M. Heusse, "Spatial issues in modeling lorawan capacity," in *Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, New York, NY, USA, 2019.
[13] L. Liu, Y. Yao, Z. Cao, and M. Zhang, "Deeplora: Learning accurate path loss model for long distance links in lpwan," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications*, 2021, pp. 1–10.
[14] C. Li, H. Guo, S. Tong, X. Zeng, Z. Cao, M. Zhang, Q. Yan, L. Xiao, J. Wang, and Y. Liu, "Nelora: Towards ultra-low snr lora communication with neural-enhanced demodulation," in *SenSys'21*, 2021. [Online]. Available: https://doi.org/10.1145/3485730.3485928
[15] A. Dongare, R. Narayanan, A. Gadre, A. Luong, A. Balanuta, S. Kumar, B. Iannucci, and A. Rowe, "Charm: Exploiting Geographical Diversity through Coherent Combining in Low-Power Wide-Area Networks," in *Proceedings - 17th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN 2018*, 2018.
[16] A. Balanuta, N. Pereira, S. Kumar, and A. Rowe, "A cloud-optimized link layer for low-power wide-area networks," in *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*. ACM, 2020.
[17] H. Fawaz, K. Khawam, S. Lahoud, S. Martin, and M. E. Helou, "Cooperation for spreading factor assignment in a multioperator lorawan deployment," *IEEE Internet of Things Journal*, vol. 8, no. 7, 2021.
[18] G. Premsankar, B. Ghaddar, M. Slabicki, and M. D. Francesco, "Optimal configuration of lora networks in smart cities," *IEEE Transactions on Industrial Informatics*, 2020.
[19] T. Voigt, M. Bor, U. Roedig, and J. Alonso, "Mitigating inter-network interference in lora networks," in *Proceedings of the 2017 International Conference on Embedded Wireless Systems and Networks*, ser. EWSN '17, USA, 2017.
[20] Y. Huang, W. Gong, and D. Gupta, "Mcmsda: a multi-channel multi-sector directional antenna wireless lan," in *2006 International Symposium on a World of Wireless, Mobile and Multimedia Networks(WoWMoM'06)*, 2006.