Bird Patrol

Rodrigo da Silva Freitas Rocha
rodrigo.freitas.rocha@ist.utl.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2019

Abstract

While global warming remains a pertaining problem, scientists are evermore interested in understanding the human impact on the ecosystem. Current technologies applied in ecology for understanding and sensing the marine species remain at very cost, providing opportunities for low cost microcontrollers and worldwide IoT communities. This study provides the experimental apparatus of a bio-tag, capable of collecting the environmental telemetry, intended to be used on marine fauna focusing on seabirds. Contribution is in providing the feasibility study, using the remote sensing through LoRa protocol and modeling the geolocation position while maintaining a competitive frequency of message transmissions when compared with technologies employed nowadays.

Keywords: IoT, Wildlife Monitoring, Ecology, LoRaWAN, LoPy, Biotelemetry, RSSI, Ranging, Multilateration, Battery optimization.

1. Introduction

Recently, the viability and widespread of the Internet of Things (IoT) became possible along the years with the ongoing extension of Internet connectivity beyond standard devices. Nowadays, these devices are used mostly for sensing and actuating in urban environments, allowing them to shift the traditional non-internet enabled objects into micro-electronics, communication and sensing units. Biggest contribution of these IoT devices is that they can be easily and remotely monitored and controlled, while maintaining low cost and low power policies. These devices, represent the edge or end-nodes of the IoT ecosystem, and fuel the idea of collaborative effort networks among nodes through production, transmission, and processing of the widest variety of environmental parameters. In the case of this study, focus will be on seabirds and specifically, in the region of Madeira, as there are numerous endemic species of seabirds in that region (e.g. Madeiran Zino's Petrel). This paper contributes in building a prototype and a pilot study, paving the way for the future use on local seabirds, with the purpose to collect more environmental data at this region.

1.1. Motivation

Up to the present time, technological improvements made during the last decades are ending what the industrial revolution began, creating economies based on information technology [2]. It is of common knowledge that the industrial era, along with the technology produced as a consequence of it, played a considerable part in changing the day-to-day lifestyle, while it is continuously increasing the human impact on nature on a global level (e.g, biodiversity loss, climate changes, deforestation, among others) [15]. Moreover our planet needs sustainable actions and more research and proof of the human impact in the ecosystem.

1.2. Research Questions and Contributions

While other studies focus on small distances, or in city environments, in this study, we explore the issues of using long range (LoRa) radio communications in ocean environments, focusing on a very low altitude for data collection using the sea vessels. We explore LoRa as a mean for oceanic environmental telemetry as well as to approximate location without the usage of high energy devices. Moreover we also maximize the efficiency of the developed prototype in terms of battery. To achieve these aims, we focus on following research questions:

1. Is it possible to produce a bio tag for birds with an off-the-shelf solution as-is and are the emerging LPWAN (Low Power Wide Area Networks) sufficient to face the constraints of the problem posed in the sections below?
2. How does distance estimation behaves in ocean settings?
3. What's the overall level of accuracy in position estimation that is possible to achieve with the

system developed?

2. Related Work

Bio-loggers are devices transported by animal subjects with the purpose of collecting data. Applications of such devices are denominated and known as bio-logging [7]. However, when doing the literature survey, bio-logging differs from the definition of bio-telemetry [3]. Bio-loggers are referred to devices with storage capacities to record measurements in order to create local stored logs. These devices have limited or no transmission capabilities while they are being affixed to the subject specie. Moreover, their recorded logs are collected after the end of the experiment and are mostly used for laboratory processing. Bio-telemetry devices, on the contrary, provide experts with access to real-time data. These devices may have storage capacity, while their primary focus is the real-time transmission of the acquired data to a remote server.

Archival Tags are of most rudimentary forms of bio-logging systems [3]. As typical geolocation tags do not transmit data remotely, they do however require retrieval when the experiment is performed. This allows the successful retrieval of sensing logs from the on-board memory. Location in archival tags is usually estimated with global location sensing (GLS), however as studied in [14], the use of positions obtained through light-based estimation presents a mean error that can range significantly from 40 to 380 km for free ranging animals (animals living outside of captivity). Such inaccuracies are due to passing storms or clouds and also due to large-scale movements that occur during their flights. Currently, the biggest advantage of data-loggers is the low-price that can go from US$ 12 to US$ 600, with devices weighing from 0.3 to 3.3 g. These devices are usually equipped with batteries providing a longevity as low as 5 days to as high as 2 years.

In bio-telemetry systems, there are three types of distinct systems used in radio-telemetry [6]: (1) Very High Frequency (VHF) radio tracking, (2) ARGOS satellite tracking and (3) Global Positioning System (GPS); VHF is a radio band located between 30 to 300 MHz and transmitters emit radio-frequency signals which are received by a receiving antenna with a ground-to-ground range of 5-10 km and a air-to-ground range of 15-25 km [11]. VHF enables animal tracking using two main methods, homing and triangulation. As explained in [10], former technique involves following a signal towards its strength, repeatedly until the subject of the experiment is in sight. Latter is presented below in the location estimation techniques. In ornithology, this is presented as a low accuracy method, due to large-scale movements of birds which will introduce significant error in position estimation. However, VHF reports a precision between 200-600 m for the location estimation through triangulation and homing, and has proven effective in the studies of species with low levels of movement with few up-front costs and transmitters going from US$ 180 to US$ 300 [16]. In general, VHF transmitter can weigh as little as 0.2 g [13] with a lifespan of 18-22 days or can weigh as much as 100 g with a lifespan up to 4 years. VHF is a very effective way of tracking animals, although when studying migratory birds or far ranging animals [17], the usage of these systems does not provide the necessary needed range, allowing the coverage of the area travelled by the specie.

ARGOS is a satellite-based system which began in the 1978 [1], and has global coverage through three subsystems: (1) the Platform Transmitter Terminal (PTT), (2) the space segment and (3) the ground segment; PTT attached to animals transmit radio signals in the Ultra High Frequency (UHF) band which are detected by ARGOS satellites located on a polar orbit at 850 km above the earth. Argos satellites cover 100% of the earth surface with a visibility diameter of 5000 km. When a transmitter begins to be in range of a passing satellite, it has approximately a window of 10 to 12 minutes to send the frequency data and the required timestamps (Doppler effect). These data are then downloaded and processed at the ARGOS processing centers, where the location is calculated. Also, these centers process the information received by the satellites and the location is assessed trough a least squares analysis and assigned to one of the several location classes which represent the range in which the position estimate is inserted. ARGOS satellites can locate any transmitter between 1 to 14 times per day with a high position accuracy of 1 km of standard deviation. Applicable to ornithology, an ARGOS PTT can be powered by solar cells which can weigh a total of 2 to 50 g having a base price of 2900 to 4450 US$ having an estimated lifespan of 2 - 3 years. When powered by batteries ARGOS PTT can weigh from 45 to 105 g costing between 2550 to 2950 US$ with a lifespan between 40 days up to 3 years.

Unlike the aforementioned technologies GPS telemetry systems have major disadvantages related to the costs. A single collar can range from around 2000 to 8000 US$ depending on the features of the unit itself [8]. Another problem related with GPS is the lifespan of the device, as GPS is a power hungry location method. To mitigate this problem, many devices have programmable duty-cycles to take only a few locations per day and proceed to sleep, achieving the conservation of power. Currently, solar combined GPS/ARGOS systems weigh between 17 to 50 g with lifetime up to 3 years.

2.1. Location Estimation Techniques

In general, several techniques can be used to estimate the position of ubiquitous devices. However, in some applications, this technology cannot be used due to hardware, power or location (e.g. indoor) constraints. In this section, we report the current techniques and methodologies used for location estimation using radio signal, while selecting the one for the study:

1. Triangulation is a technique that makes use of the Angle of Arrival (AoA) to determine target location by intersecting pairs of angle direction lines. Knowing the AoA on more than one antenna its then a problem of solving the trigonometric relationships from the intersection of the two bearings formed by a radial line to the receiving gateways, however this method is not used as it requires equipment capable of determining the AoA;

2. Multilateration is a technique that allows to estimate the distances d from the signal strength captured on receivers and draw a radius with that distance centered on the receivers;

3. Time of Flight (ToF) and Time Difference on Arrival (TDoA) are methodologies employed in order to discover distance between emitter and receivers, making use of time synchronization taking as principle the linear relation between the time of propagation, i.e ToF, and the speed of propagation c, in order to determine the distance traveled, however this method was not used as it requires high precision equipment;

4. Distance based on Received Signal Strength Indicator (RSSI) is another research perspective focused in discovering the distance separating emitters and receivers. Instead of using a time dependent method which will increase costs as well as the need of using more hardware and complexity, one can use RSSI measurements in order to estimate its distance by using a radio wave propagation model or path-loss model;

2.2. Radio Signal Propagation

Properties of radio signal propagation directly limit the performance of communication in Wireless Sensor Networks (WSNs). These same properties vary depending on the environment that the signal is subjected. Normally a way of quantify these losses is to apply a path-loss model. However these models can be applied inversely by knowing the transmitted and lost power in order to determine the distance d.

In telecommunications, when addressing radio signal propagation the free-space model is one reference in prediction models. This model works by assuming a clear Line of Sight (LoS) with no obstructions between emitter and receiver represented in equation 1.

$$L_{Pfree} = [\frac{4\pi \cdot d}{\lambda}]^2 \tag{1}$$

In the literature [9], the scenario that characterizes best, transmissions that take place over water, or at sea with LoS conditions are best described by the previously explained free space path-loss model and by the log path-loss model. The Log-distance path-loss model presents itself as a practical and simple way of estimating the signal strength as a function of distance. The equation that defines the model is,

$$PL = PL_{d_0} + 10 \cdot \gamma \cdot log\frac{d}{d_0} \tag{2}$$

where the γ parameter represents the path loss exponent, and d_0 is the reference distance in meters, usually at 1 m, for the path-loss PL_{d_0} reference and d the distance of interest. This model provides a practical way of adapting to each environment, although some references, of path-loss exponents are given [4], a good practice is to determine the exponent in order for the model to be best adapted to the surrounding environment.

2.3. Long Range Radio Communication

Most modern electronic communication systems use either electricity or electromagnetism as a way to carry information. One important factor of IoT is the microcontroller capable to be configured on the go. In this study, 3 microcontrollers were reviewed. Arduino UNO, since is since it is the most used board within the Arduino spectrum for in-situ IoT solutions, Particle Photon since Particle offers solutions into a fully integrated IoT platform that offers the hardware, paving the way for the constant connection to the internet, over WiFi, cellular, or mesh networks. And lastly the Pycom LoPy 4 given its recent launch in the market. In this study the LoPy 4 was used since its main advantage against the other two microcontrollers is the capability to allow 4 radio communication techonologies: LoRa, Sigfox, Bluetooth and Wi-Fi, and having a series of expansion boards for different purposes such as sensing, tagging and tracking.

In terms of LPWANS, LoRa and Sigfox were reviewed. Both work with similar frequencies (mostly around 800 and 900 MHz, depending on the territory) but use different encoding technologies, bandwidth and settings. The Sigfox is also for-profit with a royalty-based model while LoRa is an open platform. The available bandwidth is very limited (in the order of just a few hundred bits per second and with very limiting fair-usage policies) but the range can go up to a few tens of kilometers in open environments or a few kilometers in urban environments. This makes LPWAN technologies great for IoT solutions, but limited to non-real-time solutions that can wait very long to transmit small packets of information. In our study, LoRa will be used as the location of the study currently only supports it.

3. Methodology

In this study implementation, a LoPy 4 was used, moreover the system was tested for both geolocation in oceaninc settings and also tested for maximization of battery.

3.1. Research Apparatus

For the geolocation testing, 4 gateways on land were deployed and one sensor node was placed on-board of a vessel cruising in the south coast of the Madeira island. The LoPy microcontrollers were embedded into a plastic protective casing. Further, each microcontroller was equipped with PySense expansion board, including the sensors for ambient light, barometric pressure, humidity and temperature sensor, a LiPo battery charger, and a MicroSD card.

Maximizing the longevity of the sensor was done in a laboratory resorting to an oscilloscope as a way to measure and classify the different behaviours of the sensor explained below. Moreover maximizing the battery lifespan, was done using the final settings of the prototype, LoRaWAN was used to guarantee encrypted messages and also ack messages. In further a GPS module was attached to the sensor.

3.2. Software

In figure 1 the full the algorithm used in the development of the node is shown.

The concept behind all the terminating states depicted in 1 was to create a sensor with the capability of adapting its deep sleep time to any group of events occurred while executing the main cycle. Thus, each state represents a different collection of events, depending on the LoRa joining procedure, payload type, the events received by the callback on the first and second transmission and lastly by checking if there's any message stored in memory. Each state, depicted above served also as a basis for the battery optimization section, and illustrates that each behaviour will have different impacts on battery consumption.

Once the node's software was implemented, the development of the gateway began as well as establishing the conditions for the network server to receive the messages and forwarding them to the The Things Network (TTN) API.

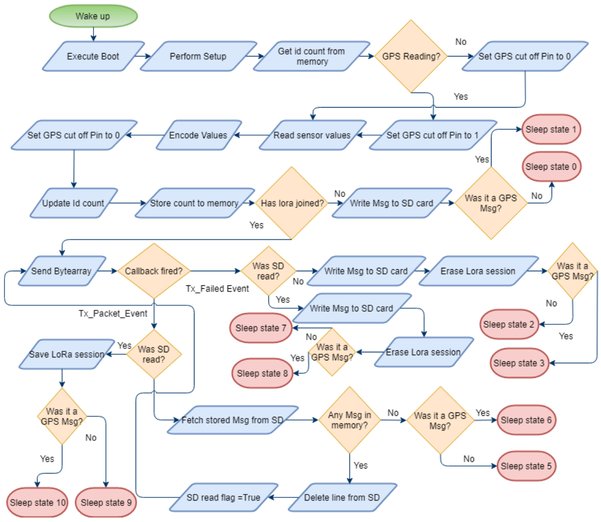

Figure 1: LoPy 4 node algorithm

3.3. Study Setup

For estimating positions as was mentioned there was no use for encryption, or for an API, as these tests were done with the single interest of obtaining the RSSI values per transmission. 4 gateways were placed on top of a PVC pole with an average height of 3 meters. The node was attached on the mast of a vessel transmitting the GPS ground truth for later processing. The first approach was used and the distance d and is calculated with the equation 3 presented below [12].

$$d = 10^{(RSSI/10)*n} \tag{3}$$

This equation uses the RSSI values, the distance d in meters and a tuning parameter n. The workflow in the location estimation tests was done as follows:

- Apply an average of the RSSI values per position as RSSI values can be influenced by the surrounding environment, and even the same ground truth position can suffer signal variation. This action smooths the data for modelling and increases the precision of the ranging method.
- As for ranging methods both the free-space and log path models were used, in addition linear regression was used to compute the distance vs linear RSSI.
- Use linear regression with the RANSAC method, in order to improve robustness against outliers in the data.

- Apply bilateration using the best model for the circles calculated in the ranging phase.
- From the two possible bilateration, choose the location calculated at sea, discarding the inland one.

Normally higher degree multilateration would, result in better location estimation. This is usually true, however the use of multilateration was not possible due to the gateways being aligned in an almost straight line with small curvature in-land. When ranging, and if the circles do not intersect this causes the multilateration equations to determine a positions equally apart from the circles although with a symmetrical orientation relative to the straight line, increasing the error of the model. Stated this, we chose to use a less accurate solution although more stable: bilateration.

4. Results

This test was performed for the smallest battery available at the moment, which was a 1800 mAh LiPo battery. Furthermore the tests, were made to fit the states presented in the algorithm (see fig. 1) with the corresponding behaviour. The first criteria was to leave a security margin for the available battery levels. As a consequence of this criteria, the system was designed to use 80% of the battery, in result the calculations made fit the equivalent to a 1440 mAh battery. As 3 months, represent 2160 hours of consumption the average current drained by the system must correspond to 0.66 mA. By measuring, the deep-sleep which corresponded to 60 μA and by establishing this balance, the values were related in equation 4 formed by the values of current consumption and to discover the percentage of time both in current consumption as in sleep consumption forming the whole duty cycle.

$$0.66 = \overline{I}_{State_i} \cdot t + \overline{I}_{Sleep} \cdot (1 - t) \qquad (4)$$

By obtaining the current consumption per measure and in function of the time of the whole cycle per state, we were able to derive a number of measures that the system will be able to send both in the best and worst case scenarios, by dividing it for the whole capacity of the battery. Table 1 depicts the best and worst case scenarios as well as the number of transmissions achievable with any one of the two cases.

State	Duty-Cycle (h)	TX p/ day
5 (BC Telemetry)	0.71	18
6 (BC with GPS)	2.66	4
7 (WC Telemetry)	0.95	12
8 (WC with GPS)	3.48	3

Table 1: Number of transmissions p/day based on best and worst scenarios

These tests, allowed the system to be tuned-up in terms of deep-sleep time and also in terms of conditioning the GPS reading to happen. The system was designed for 3 months, it might endure more time depending on which behaviour the sensor presents, but further conclusions in this matter will be pure assumptions with no basis of proof.

4.1. RSSI Distance Estimation

The measurements for distance estimation, were done in-situ and located in the south coast of Madeira island, although four gateways were placed inland, two of them failed and were unable to capture the greatest part of transmissions, which forced us to work with the data received by gateway 2 (green) and gateway 3 (yellow). However the full boat trip as well as the location of the gateways is depicted in the figure below (see fig. 2).

Figure 2: Full trip and gateway location

Once the trip was concluded, the data was processed. Although before applying the models, the data gathered by each gateway was analyzed individually.

By applying the path-loss models to the full data set we can observe, that the log-path model outperformed the free-space model. Furthermore, figures 3 and 4 depict the average error between the applied models as a function of the distance to the gateways. In these figures the free space model error can be seen as the blue line and the log-path error as the red dotted line.

One notable difference between the two graphs shown, is that gateway 2 shows a much more stable error throughout the time, this is due to the fact that the sensor was most of the time obstructed by the mast which provided noisy readings throughout the time and therefore with lesser oscillations. Gateway 3 was intermittently in LoS and non-LoS conditions which in average provided better values although with more oscillations.

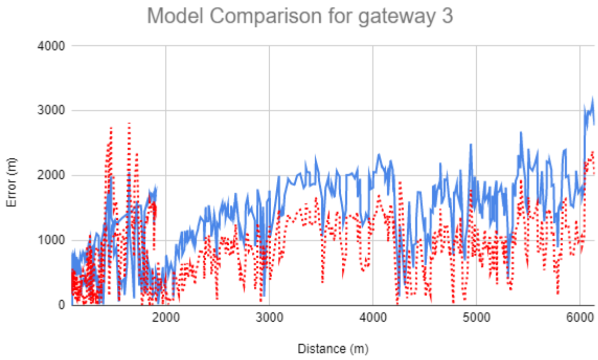

Figure 3: Gateway 3 average error per model

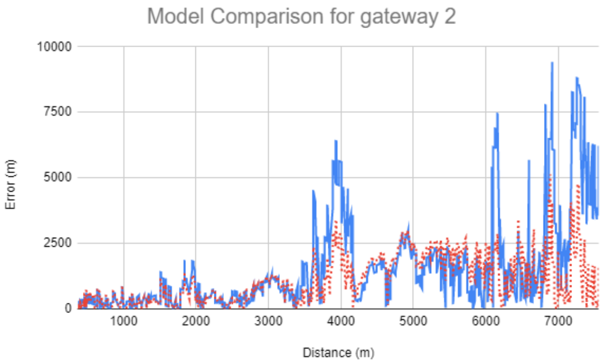

Figure 4: Gateway 2 average error per model

In the end the log-path model outperformed the free-space model by approximately by a factor of 1.15 in gateway 2 and of 1.5 in gateway 3.

Leaving aside the propagation loss models and in attempt to provide a better distance estimation, the RSSI and distance values were modeled by application of linear regression. The linear regression was modeled using the following as function of distance, (1) the RSSI in logarithmic scale (2) the linearized RSSI and (3) the averaged RSSI used in (1) in order to smooth oscillations of RSSI values that correspond to the same ground truth position. This way allowed to obtain a solution that was better in average for the whole set. The best model obtained from applying the linear regression to the grouped data as depicted above, corresponds to the linear regression that weighs the averaged RSSI with the distance to each gateway. However, at this point the mean error per measurement hasn't suffered any substantial decrease. By looking at figure 5 and figure 6 and apart from the fact that gateway 2 best fits its equation, one can still observe that each model still presents many outliers, that negatively influence each set.

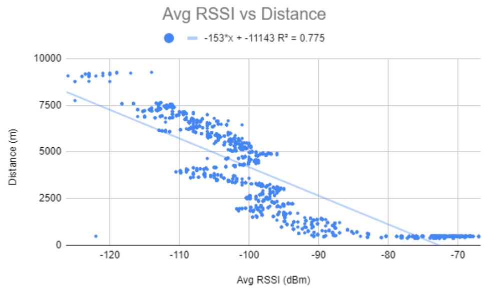

Figure 5: Avg RSSI in function of distance of gw 2

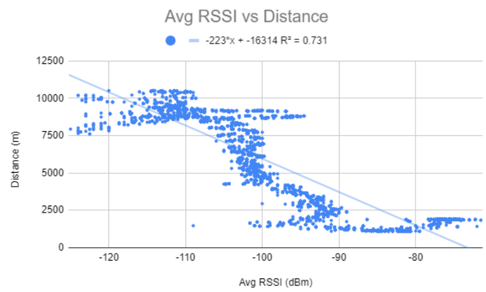

Figure 6: Avg RSSI in function of distance of gw 3

Due to the presence of outliers, the RANdom SAmple Consensus (RANSAC) was applied. This method is a general parameter estimation approach designed to cope with the presence of outliers in the input data. Furthermore this method generates candidate solutions by using a minimum number of observations [5]. The maximum residual/threshold for a data sample to be classified as an inlier was the MAD (Median Absolute Deviation). The figures 7 and 8, shows the respective improvement of the linearized model presented above, as well as the respective inlier and outlier points.

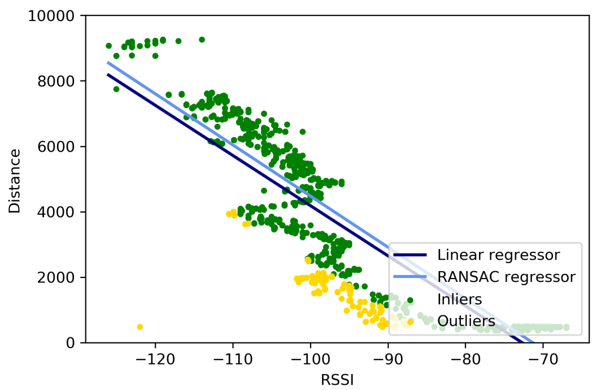

Figure 7: RANSAC regressor for gw 2

In this analysis with the linear regression plus the RANSAC method applied a mean error per measurement of 732 m for gateway 2, and a mean error per measurement of 885 m for gateway 3 was

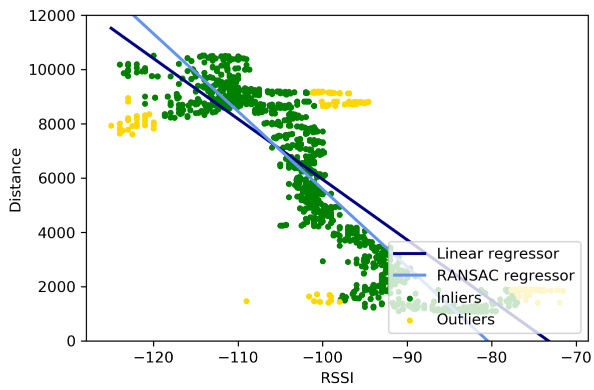

Figure 8: RANSAC regressor for gw 3

achieved. The RANSAC regressor outperformed the average RSSI linearized model which achieved a mean error per measurement 945 m for gateway 2 and a 1180 m mean error for gateway 3. Moreover by applying the RANSAC method the path-loss models were also outperformed. With the free-space model we've achieved a mean error per measurement of 1325 m for gateway 2 and a 1360 m mean error for gateway 3. Concerning the log-path model we've achieved a mean error per measurement of 1160 m for gateway 2 and a 960 m mean error per measurement of 960 m for gateway 3.

4.2. Position Estimation

By knowing the gateway positions in-land and once the estimated distances were calculated it became a problem of translating the cartesian coordinates, into Universal Transverse Mercator (UTM) coordinates. As it was expected, position estimation became, progressively better as distance to the gateways decreases. In figure 9 we can separate the locations estimated into three classes, the first class, represented in red color corresponds to the points were the LoS was obstructed by the mast of the vessel to both the gateways, which drastically influences the RSSI values obtained. The red class groups a set of points that correspond to a precision of approximately 4 km. The yellow class corresponds to the moment where LoS exists for the gateway 3 although remaining obstructed for gateway 2. The yellow class represents a precision of 3 km. Lastly, the green class corresponds to the situation were a clear LoS was achieved for both gateways a 650 m precision was achieved in the latter. While many estimated points in 10 are close to the GPS ground truth, we can see that the data is very much dispersed.

4.3. Discussion

The research findings presented in this discussion relate with the research questions posed in the introduction section.

Figure 9: Degree of precision in function of subset

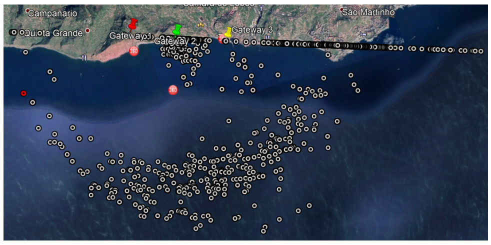

Figure 10: Estimated points

One of the features achieved that really stands out and elects these types of systems as a viable choice for a bio-telemetry system is the longevity that one can achieve with a small battery. For the tests made with the smallest battery of 1800 mAh, this system could achieve in the best-case scenario a total of 22 transmissions of telemetry per day, with 4 of those yielding a GPS ground-truth measure. In the worst-case scenario 15 transmissions are sent, being 3 of the total a GPS ground-truth. In terms of frequency of messages and exchange of telemetry, there is a competitive advantage in this system when facing ARGOS which only transmits a total of 14 messages per day. GPS tags provide by far the best accuracy, however the cost of the tag, and the price to pay in power to acquire and transmit is very high as it was depicted above when studying the power consumption.

In terms of deriving the distance d, from the gateway to the emitter, and since the RSSI values are inherently sensitive to the environment, as much as the signal is the post-processing of the data still left a lot of noise, as can be seen in figures 7 and 8. Although, despite the noise, the results presented an average error of 732m and 885m in the two gateways used. Which in the distances measured up to 10km represents a satisfactory error, although as expected the reduction of the distance also reduces the error per measurement. In addition results also show several points were the error per measurement is $\leq$ 1m. Although these are good points to obtain, we must ensure that these points remain common to the both gateways. As a point with minimum error in one gateway and with a greater error in other gateway will also induce error when multilateration takes place.

As it was stated before, higher degrees of multilateration will, usually, result in better location estimation. However in this case bilateration would be the only case considered. In the test presented 4 gateways were placed inland and 2 gateways failed at providing results. However there is other two reasons to support only the usage of bilateration which were verified in previous tests, and these are: 1) the coastal gateways are aligned in an almost straight line with a small curvature, which caused the multilateration equations to near a singularity where it is highly unstable and tends to give false results inland, and 2) the very unstable RSSI values don't provide coherent values throughout the time. This means that it is extremely difficult to fingerprint a determined distance with an RSSI value. By scaling the problem to more gateways in this situation translates in even more radius from each node resulting in more oscillating results. Problem 1) can be solved by better redistributing the gateways to form a better geometry, one that would contain the node. As for problem 2) better hardware and better testing conditions will result in better obtained values. Using the post-processed RSSI values, a range was determined and bilateration was executed using the well known location of the gateways. Afterwards the results were compared to the GPS derived values, in order to understand the accuracy and quality of the results obtained. By fixing problem 1) and 2) in the future, would enable a higher degree multilateration with better fingerprinting of RSSI per location. Another obstacle faced in this study was the limited access to boat trips for testing usually with no prior knowledge of trip planned, which depicts a more realistic situation. The best class obtained yielded results with a mean error of 650m up to the worst class which yielded results in the order of the 4 km. Being the results estimated using RSSI values all the improvements and drawbacks are shared. Although in absolute values these errors may seem large and prohibitive in a urban or semi-urban environment, an average error up to 5 km is still seen with the naked eye in open ocean. Thus this location estimation proves to be useful in oceanic scenario.

4.4. Research Contributions

The main focus in this study was to build a bio-telemetry tag that would be low-cost, and power efficient and would provide an efficient geolocation solution, to be used by marine seabirds in oceanic environments, namely to aid in the research and conservation of marine life. The LoRa technology and LoRaWAN protocol were used and its behaviour in oceanic settings was explored by resorting to RSSI values that comes at no cost in any kind of electromagnetic-based communication technology. The error related with the geolocation models elaborated are largely due to the fact that no control of the experiment was achieved, oscillating between LoS and non-LoS conditions which largely increased the error. Although one can't characterize a bird's flight as something controlled, this fact contributed to depict the experiment as realistic as possible, as a bird while flying would also be prone to encounter obstacles. Apart from these facts the results obtained range in average from 660 m to 4 km. This error is adequate for study migration, patterns, general location of animals among other situations where a pin-point location isn't needed. These studies not only benefit from the lifespan of the tags gained partially by employing this kind of passive geolocation, but also by removing the need for continued human interaction with the taxa.

4.5. Possible Applications

In the context of this study, for the position estimation system to function effectively in any given scenario, better geometry should be achieved as referred above, this would enable the sensor to reach more number of gateways and to apply higher degrees of multilateration, therefore reducing error. Otherwise the system fails at providing solid and consistent results. With this fact in mind, these type of systems could be applied, in the wildlife, but to species with lower degree of movement like turtles, or reptiles. This would allow, to sample more RSSI per location which would improve the error obtained therefore obtaining better results. Aside from the ecology field, the most promising applications would be in areas such as cities, where one could make use of the increasing population and their increasing usage of ubiquitous systems to take a role as gateways providing multiple endpoints for a system like these. In rural environments, this system would be feasible in managing crops, where spread sensors would make use of the LoRa technology to send status messages of the monitored crops to a central server. Suitable applications most likely include cases where limitations of power and cost overrule the need for a pin-point location accuracy. Many forms of asset tracking should also be explored and present great opportunities for these systems, like locating cargo containers or managing warehouses and inventory may not require GPS accuracy.

5. Conclusions

The challenge of this research was to use an off-shelf low-cost microcontroller and lightweight protocol as a way to study marine seabirds in oceanic environments as a way to produce telemetry of the surrounding environment. As marine seabirds access environments that are difficult to access the possibility of mounting a system of this type, would come as a great advantage. Our goal was to achieve and analyze the better lifespan of the system, as replacing batteries after deployment would be intrusive. Together with evaluating the lifespan of the system, a passive geolocation was developed so to have an understanding of the general system position. In addition, a GPS module was also used as a way to correct our perception of the system position and to receive a ground truth correction between telemetry readings. In addition, when in out of range for transmission conditions the GPS module will allow that the system doesn't perform blindly in terms of geographical position. In order to accomplish this, a sensor was developed resourcing to the LoRaWAN protocol, and a network server (The Things Network). These provided a practical and easy implementation of the system, and enabled us to rapidly test the system both in a laboratory for battery lifespan enhancement, by creating multiple states as a way of adapting the sensor sleep and maintaining an average low current supplying the sensor. Tests for the passive geolocation system were developed in-situ, on board of a touristic boat at the south coast of Funchal, Madeira island. From the data collected we've modeled our data and removed the noise of the model without decreasing the realistic side of the scenario that it represents. Position estimation was done using multilateration, as we've seen there's multiple methods to estimate distance separating emitters and receptors and multiple methods to determine the final calculated position. From the first contact with research in the area it becomes clear that, despite of the several choice range in the market for hardware, methods and methodologies, it is required to select the best technologies that fit the problem. From an initial point, in every choice one has to make compromises in order to best-fit the solution to the problem. It also became clear that the area of ecology would largely benefit with a system of this type, as technologies used, remain costly and don't provide flexibility to the experts working in the area. In addition, the systems used are costly partially due to the fact that high accuracy is employed when not needed, although when there is no viable alternative these end up conditioning experiments.

5.1. System Limitations

In terms of limitations, and being the RSSI an indication of the power of the received radio-frequency signal is susceptible, as mentioned, to obstacles, reflections, multipath and other types of disturbances as much as the signal itself is. This creates a problem when relying on these values in order to derive distances, as depicted in the plotted charts. Furthermore in open sea, it was extremely difficult to condition the antenna position to be in LoS conditions, this serves the purpose of attempting to depict a realistic scenario. Although these conditions result in a great distortion and oscillation of the values, therefore inducing error in the final results. Another thing to note is that RSSI values are integer values with an inherently very limited range of values (usually 0 to -128) in a logarithmic scale. This greatly limits the resolution of the data as small differences at low distances represent several oscillations between RSSI values and small differences at higher distances represent only small oscillations, thus becoming difficult to fingerprint RSSI values to higher distances. The latter problem was mitigated by averaging the RSSI per location. During this practical experiment 2 of the land gateways failed to deliver results. However with the very limited access to boats, and constantly being hindered of testing due to conditions at sea, the test proceeded with only 2 gateways. The 2 extra gateways might have increased overall quality of the study and reduced error in the final position estimation, by duplicating the amount of data to tune the developed model. The limited access to boats, largely affects the attempt to track species over the ocean, by forcing the study to be conducted with only land gateways and just one boat. By improving the overall geometry, the system will, for sure, increase the overall quality of the signal and decrease the final estimated point error.

5.2. Future Work

A study of this type, namely in ocean settings, with all the constraints that this environment offers to the application of IoT, opens possibilities for numerous future use cases. Moreover the reported techniques, using calculated models remain to be further verified using real trajectories of fishing boats, where geometry of the network could be further improved by implementing gateways on buoys and ships along the shore. In addition, implementation of an IoT dashboard providing real-time information, where it will be possible to provide additional back-end processing of the data obtained. In terms of battery autonomy, and accounting with the satisfactory result obtained, can easily be transformed into solar powered, since Pycom expansion boards all include a battery charger circuits, thus further extending the longevity of the deployment. Another component for future deployment is to invest in proper casing and miniaturization of the sensor, although these topics were not addressed throughout this study, it remains an important aspect as, ocean environments are harsh to deal with due to salt corrosion and, of course, the presence of water thus water and pressure proofing remains a concern. In addition, to cope with the immensity of the ocean a mesh network could be built based on the idea of data mules, that transmits the messages to other nodes, until they reach a gateway with internet connection. These middleman could have fixed positions, be carried by sea vessels or, in context, be a bird flying receiving messages and working as a monitor of tagged vessels, cetaceans or other marine species. Finally, the latter would create an opportunity to further scaling up the proposed systems at an affordable price enabling other sensors to operate far away from shore collecting levels of parameters such as: salinity, levels of dissolved oxygen, ocean depth among others, which are crucial to the conservation of the ecosystem, where a real-time monitoring system brings huge benefits to the marine research and conservation and the general community.

Acknowledgements

I would like to thank my family for their friendship and encouragement, without whom this project would not be possible at all, for education is the basis of an individual in society. To my girlfriend for her patience, friendship, love, affection and utmost dedication throughout this journey.

I would also like to acknowledge my dissertation supervisors Prof. Nuno Jardim Nunes and Prof. Marko Radeta for the possibility of being part of M-ITI and also for their valuables insights, support and sharing of knowledge that has made this Thesis possible.

Last but not least, to all my friends and colleagues that helped me grow as a person and are always there for me during the good and bad times in the journey of life. To each and every one of you, who without your contribution this work wouldn't be possible, for that I'm grateful and from the bottom of my heart, I thank you all.

References

[1] I. Argos and T. Argos. Give Vision. 1970.

[2] R. O. Charlie Karlsson, Gunther Maier, Michaela Trippl, Iulia Siedschlag and G. Murphy. *ICT and Regional Economic Dynamics: A Literature Review.* 2010.

[3] S. J. Cooke, S. Hich, M. Lucas, and M. Lutcavage. Chapter 18 - Biotelemetry and Biologging. *Fisheries Techniques*, (January):819–860, 2012.

[4] J. Danebjer. A Hybrid Approach to GPS-Free Geolocation over LoRa. pages 1–62, 2018.

[5] K. G. Derpanis. Overview of the RANSAC Algorithm. 4:1–2, 2005.

[6] V. Dyo, S. A. Ellwood, D. W. Macdonald, A. C. Markham, C. Mascolo, B. Pásztor, N. Trigoni, and R. Wohlers. Poster Abstract : Wildlife and Environmental Monitoring using RFID and WSN Technology. *Distributed Computing*, pages 7–8, 2009.

[7] G. Fehlmann and A. J. King. Bio-logging. *Current Biology*, 26(18):R830–R831, 2016.

[8] M. Hebblewhite and D. T. Haydon. Distinguishing technology from biology: A critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2303–2312, 2010.

[9] W. Khawaja, I. Guvenc, D. W. Matolak, U.-C. Fiebig, and N. Schneckenberger. A Survey of Air-to-Ground Propagation Channel Modeling for Unmanned Aerial Vehicles. *IEEE Communications Surveys & Tutorials*, pages 1–1, 2019.

[10] L. D. Mech. *Handbook of Animal Radio-Tracking.* University of Minnesota Press, ned - new edition edition, 1983.

[11] L. D. Mech and S. M. Barber. a Critique of Wildlife Radio-Tracking and Its Use in National Parks a Report To the U.S. National Park Service. (January 2002):1–83, 2002.

[12] P. D. I. Milano. Polo Regionale di Como Final Thesis for MSc . Computer Engineering RSSI Based Localization for Indoor Application by using 868 MHz Radio Signal. (737700), 2012.

[13] B. Naef-Daenzer. Miniaturization (0.2 g) and evaluation of attachment techniques of telemetry transmitters. *Journal of Experimental Biology*, 208(21):4063–4068, 2005.

[14] R. A. Phillips, J. R. D. Silk, J. P. Croxall, V. Afanasyev, and D. R. Briggs. Accuracy of geolocation estimates for flying seabirds. *Marine Ecology Progress Series*, 266:265–272, 2004.

[15] L. Rothacker, A. Dosseto, A. Francke, A. R. Chivas, N. Vigier, A. M. Kotarba-Morley, and

D. Menozzi. Impact of climate change and human activity on soil landscapes over the past 12,300 years. *Scientific Reports*, 8(1):1–7, 2018.

[16] G. M. Skupien, K. M. Andrews, and T. M. Norton. Benefits and biases of VHF and GPS telemetry: A case study of American alligator spatial ecology. *Wildlife Society Bulletin*, 40(4):772–780, 2016.

[17] B. K. Walkerl, G. Elliottl, D. Nichollsz, D. Murray, and P. Dilksl. ALBATROSS (Diomedea exulans) FROM THE AUCKLAND ISLANDS : PRELIMINARY RESULTS. (Gales), 1993.