Università degli Studi di Padova

Facoltà di Ingegneria

Corso di Laurea Magistrale in ICT for Internet and Multimedia

Master's Thesis

Environmental Monitoring of Coastal Waters with a Collaborative Underwater Acoustic and Above Water LoRaWAN Sensor Network

Advisor: Filippo Campagnaro

Author: Nicola Toffolo - 2019040

December 7, 2022

Alla mia famiglia,
a Valentina e Carlo

Abstract

Climate changes are transforming the world as we know it and have a devastating impact on frail areas, such as coasts, afflicted by catastrophic events (rise in seawater temperature, floods) deteriorating the local biodiversity. Between the strategies undertaken to mitigate these effects, the EU Biodiversity Strategy for 2030 is one of the most ambitious. In particular, a relevant point is the inclusion of new solutions to monitor the conditions of the water, measuring specific parameters and polluting agents. However, up to today, there is no common ground when dealing with low-cost and low-power devices to collect data related to the quality of the water in coastal areas: a dense deployment of sensors would be the best option, but the technology used for long-range underwater acoustic communication is indeed extremely expensive. Nonetheless, in the last few years researchers have been investigating the possibilities given by low-cost and low-power acoustic modems, in the attempt to provide a way to employ dense deployment of underwater nodes. Another major turn in long-range low-power communications is the introduction of Low-Power Wide-Area Networks (LPWAN), which can be regarded as one of the most crucial entries in Internet of Things (IoT) applications. With this dissertation, we propose a network infrastructure for the tracking and the study of water quality parameters, to understand the impact they have on biodiversity. Specifically, we envision a system where there are two types of sensor nodes; one underwater and another on the water surface, forwarding the data they aggregate to one or more gateways. The gateways are connected to the Internet so that the data can be saved in a database for further processing. Underwater nodes use a part of the surface nodes as relays being based on an acoustic communication protocol, while the remaining surface nodes generate sensor data themselves; LoRa (together with LoRaWAN) has been chosen as the core LPWAN, enabling the long-range communication between the surface nodes and the gateways. Finally, the gateways are connected to the Internet with LTE standard. Simulations have been run to estimate the traffic requirements of the network as well as the feasibility of the system and a functioning prototype of a surface node has been developed. We selected a section of the Venice lagoon as reference area where our network could eventually be put in place, thus the simulations have been set according to this scenario.

Sommario

I cambiamenti climatici stanno trasformando il mondo come lo conosciamo e hanno un impatto devastante sulle aree più fragili, come le coste, colpite da eventi catastrofici (aumento della temperatura dell'acqua marina, inondazioni) che impattano negativamente la biodiversità locale. Tra le strategie intraprese per mitigare questi effetti, la Strategia sulla biodiversità per il 2030 dell'UE è una delle più ambiziose. In particolare, un punto rilevante è l'inclusione di nuove soluzioni per monitorare le condizioni dell'acqua, misurando parametri specifici e agenti inquinanti. Tuttavia, ad oggi, non esiste un punto d'incontro quando si tratta di dispositivi a basso costo e a bassa potenza per la raccolta di dati relativi alla qualità dell'acqua nelle aree costiere: un fitto dispiegamento di sensori sarebbe l'opzione migliore, ma la tecnologia utilizzata per la comunicazione acustica subacquea a lungo raggio è estremamente costosa. Tuttavia, negli ultimi anni i ricercatori hanno studiato le possibilità offerte dai modem acustici a basso costo e a bassa potenza, nel tentativo di fornire un modo per favorire una distribuzione fitta di nodi subacquei. Un'altra svolta importante nelle comunicazioni a lungo raggio a bassa potenza è l'introduzione delle reti LPWAN (Low-Power Wide-Area Networks), che possono essere annoverate tra le più cruciali applicazioni dell'Internet of Things (IoT). Con questa tesi, proponiamo un'infrastruttura di rete per il tracciamento e lo studio dei parametri di qualità dell'acqua, per capire l'impatto che hanno sulla biodiversità. Nello specifico, immaginiamo un sistema in cui ci sono due tipi di nodi sensore, uno sott'acqua e uno sulla superficie dell'acqua, che inoltrano i dati che aggregano a uno o più gateway. I gateway sono collegati a Internet in modo che i dati possano essere salvati in un database per l'elaborazione successiva. I nodi subacquei utilizzano una parte dei nodi di superficie come relay, basandosi su un protocollo di comunicazione acustico, mentre i restanti nodi di superficie generano autonomamente i dati sensoristici; LoRa (insieme a LoRaWAN) è stato scelto come core LPWAN, e permette la comunicazione a lungo raggio tra i nodi di superficie e i gateway. Infine, i gateway sono collegati a Internet con lo standard LTE. Sono state eseguite moltpelici simulazioni per stimare i requisiti di traffico della rete e la fattibilità del sistema ed è stato sviluppato un prototipo di nodo di superficie. Abbiamo isolato un'area della laguna di Venezia e l'abbiamo scelta come area di riferimento in cui la nostra rete potrebbe essere installata. Le simulazioni sono state quindi studiate in relazione a questo scenario.

Contents

List of Tables

List of Figures

Introduction

Climate changes have a devastating impact on coastal and littoral areas, heavily affected by seaquakes and floods. Furthermore, global warming causes a dramatic change in the biodiversity of rivers, seas, and lakes, including fragile biodiversity hotspots and protected areas, such as the Venice Lagoon in Italy. A similar impact is due to pollutants: this called for a large-scale long-term action that aims to monitor aquatic environmental parameters in order to predict, manage and mitigate these effects. The new European Biodiversity Strategy for 2030, for example, is a comprehensive long-term plan whose aim is to protect nature and reverse the degradation of ecosystems, not only with immediate actions (e.g., the restoration of coastal ecosystems or the creation of consortia to remove waste from coastal areas), but also with the introduction of innovative solutions to monitor water quality parameters and pollutants.

Nevertheless, coastal systems are highly heterogeneous in space and variable over short (daily), medium, and long (seasonal, interannual) timescales, thus reliable but affordable monitoring can be considered quite a challenging task. Indeed, in such scenarios, a dense deployment of sensors is usually preferable to a coarse deployment of expensive sensors equipped with long-range high-power acoustic modems, due to unsustainable costs. Another crucial point is the evaluation of the performance of such a system by means of simulation. The underwater acoustic channel is strongly dependent on the scenario considered and the environmental conditions.

In fact, channel impairments differ significantly in shallow water with respect to deep water, and the presence of external factors such as bubbles, rain, or ships passing nearby, changes of temperature, and wind strength can change drastically the quality of the link in different seasons and even during the same day. Although legacy mathematical models already exist, they are usually not very accurate (Urick model) or very computationally demanding (Bellhop ray tracer). Deterministic models based on lookup tables (LUTs) of sea trial measurements are widely used by the research community to simulate the acoustic channel, but even though they could provide a good characterization of the acoustic channel, the corresponding simulation results are limited to a single channel realization, making it difficult to comprehensively evaluate the acoustic network under different conditions.

The goal of this Thesis is to propose a way to automate the water data collection process with the use of a low-power sustainable integrated underwater and above-water Internet of Things sensor network, capable of collecting water measurements in a cloud database and making them available to researchers to monitor the status of a certain area and develop their prediction models. Moreover, we discuss the development of a statistical channel model based on the analysis of real field experiment data, and how this model performs with respect to the other channel models available in the DESERT underwater network simulator. A functional prototype of a low-power, affordable sensor

node has been built as well, and therefore here presented. Extensive simulations using the aforementioned statistical channel model have been run, and the results highlight how Low-Power Wide-Area Networks can support the data collection from dense sensor deployments.

The Thesis is structured as follows.

- In Chapter 1 we introduce the concept of underwater acoustic networks and the reasons behind our choice of developing a hybrid system to monitor water parameters, as well as discussing currently used solutions. Also, we explain why we considered the implementation of a statistical channel model in the underwater network simulator we relied on necessary for the purposes of the application.
- In Chapter 2 we present the actual statistical model for the acoustic channel, examining both the formal aspects of its development as well as its expected accuracy with respect to preexisting models.
- In Chapter 3 our envisioned network and the employed technology is discussed in detail and we present a laboratory proof-of-concept (PoC) prototype of a possible sensor node.
- In Chapter 4 we show the meaningful results of numerous simulations that on the one hand prove the robustness of our system, on the other hand highlights the maximum number of nodes supported by our envisioned network and the need for retransmission to achieve a high Packet Reception Rate (PRR).
- In Chapter 5 the achievements of this work are summed up and, highlighting the obtained results and proposing future work ideas for an eventual improvement of the system are pointed out.

Chapter 1

Underwater Acoustic Networks Background and our Contribution

In this section we first offer a brief overview on underwater acoustic networks in Section 1.1, followed by a discussion on the choices that led us to design a hybrid network infrastructure to monitor coastal environments, with a focus on the currently employed solutions and on their deficiencies with respect to our proposal (Section 1.2). We also motivate the development of a statistical acoustic channel model for the underwater network simulator that we extensively used to assess the performance of our envisioned network architecture in Section 1.3.

1.1 Underwater Acoustic Networks

The main purpose of underwater communications is often long-term monitoring of chosen areas of the ocean. During coastal monitoring, deployed sensors gather data, and the measured information are typically obtained by physically retrieving the sensors. This may be inconvenient, since a significant amount of time may pass between the collection of the data and the moment this data becomes available to be processed by scientists. Therefore, our purpose is to reduce this enormous latency as much as possible and transmit the data to a sort of database structure on the final server with no unnecessary delays. This would also reduce the probability of losing relevant data due to network and node malfunctions.

The typical network consists of a number of sensors placed in the water and that may or may not share bidirectional acoustic links. A station on the surface could be the sink for the data collected by the sensors, and it might be connected to a backbone (generally via a radio frequency link) to further forward this data. The network parameters may be altered by means of control messages sent to the sensors [2].

There are several aspects that distinguish the underwater acoustic (UWA) channel from a classical radio channel. The first is certainly the bandwidth, quite limited in the UWA channel; thus the signals are affected by multipath time-varying effects [3], producing intersymbol interference (ISI) and Doppler shifts or spreads. The range and

bandwidth have therefore to be restricted even more to mitigate these impacts, in order to achieve a communication whose robustness is at least acceptable.

From a historical perspective, we can date the first uses of UWA communications to the development of submarines with human equipment and the subsequent need to communicate with them [4]. An underwater telephone using analog modulation with carrier frequencies in the between 2 and 15 kHz was initially employed and is still in use nowadays for both military and industrial purposes. The analog signal is filtered and modulated on the transmission carrier and then demodulated and filtered at the receiving side - with the final result heavily depending on the communication path characteristics. Digital UW communications, on the other hand, was first used in sonars working in the audible band, and in the 1960s researchers started to be interested in signaling and modulation for a challenging, non-perfect channel, such as the UWA channel. In particular, since throughput may be limited to $1/T_s$ (where T_s is the length of the delay spread in a non-compensated channel), a major interest was in how to increase the effective data rate. In 1971, Williams and Battestin [5] discussed an actual example of multipath compensation that allows for a throughput greater than $1/T_s$, exploiting the temporal coherence of the channel, which allows for estimation and compensation. Earlier systems (1980s) used incoherent modulation for the sake of simplicity, with one of the few exceptions being the vertical links in deep water, for which almost always coherent modulation was used due to the very little delay spread. However, in the 1990s, phase coherent communications drew attention due to the potentially increased efficiency, for this kind of communication would allow for an efficiency higher than 0.5 bits per Hz (that is, the limit for incoherent communication). Nevertheless, phase-incoherent systems (e.g., frequency-hopping based systems) still have their field of use, even if their counterpart has undergone relevant improvements, especially for their ease in terms of hardware and algorithms.

Thus, we can say that the underwater acoustic channel has three main properties that limit its capabilities, that are, the attenuation which increase proportionally to the signal frequency, time-varying multipath propagation of the signal and the relatively low speed of sound. We may now review them.

The path loss could be divided into an absorption loss, that is, when energy becomes heat underwater, and a spreading loss, proportional to the distance. If f is the frequency of the signal, d the distance of the transmission, d_r a reference value, $a(f)$ the absorption coefficient empirically computable [6] and k a coefficient whose purpose is to model the spreading loss, we have that the path loss is given by:

$$A(d, f) = \left(\frac{d}{d_r}\right)^k a(f)^{d-d_r}. \tag{1.1}$$

As far as the noise afflicting the acoustic underwater channel is concerned, it may be an ambient noise or a site-specific noise. While the former is always present and can be considered a background noise that can be modeled as a White Gaussian Noise, the latter depends heavily on the characteristics of the spot of interest. This implies that this is not easy to model properly. Therefore, since the attenuation is proportional to the frequency and the spectrum of the noise declines with the frequency, the Signal-to-Noise Ratio (SNR) in a narrow band of frequencies around f can be computed as:

$$SNR(d, f) = \frac{S_r(f)}{A(d, f)N(f)}, \tag{1.2}$$

where $S_r(f)$ is the Power Spectral Density (PSD) of the signal being transmitted,

and $N(f)$ the noise function. So, longer distances are truly an issue for underwater communications, and we can use only 1 kHz of bandwidth at a distance of 100 km. Also, the bandwidth is in the same order of magnitude of the central frequency f_c and the narrowband assumption $B << f_c$ does not hold true most of the times. This means that when working with the acoustic underwater channel, we can not forget that it is a wideband channel. Moreover, since the bandwidth depends on the distance, we can leverage multi node transmission (with short hops) between a source and a destination, gaining a higher bitrate and an inferior power consumption.

Another challenge that has to be tackled is the multipath effect. This is due to the sound reflection (at surface, bottom and with eventual obstacles in the water) and the sound refraction. This last is linked to the variability of the sound speed, which is related to temperature, pressure and salinity of the water. Temperature and pressure are almost constant near the surface, as it happens in our envisioned network, but usually temperature increases with depth, and it stops decreasing at 4°C. Lastly, the destination will receive multiple signals from the source, each of which might have followed a different path, and it is possible that non-direct rays reach the destination before the direct ones, for they may have had a higher speed.

Time variability is a not negligible characteristic of the acoustic channel. The main cause we can identify for it are the changes in the propagation medium (i.e., the water), which can both happen in a timescale that does not affect significantly the communication (such as very slow changes of temperature) or in a short timescale (e.g., waves, which modify the length of the paths for the rays since the reflection point moves). Time variability is particularly hard to model accurately and while for classic radio channels a number of statistical model exist and are well documented, the same does not hold for the acoustic underwater channel.

Finally, the Doppler effect, which may be triggered by the motion of the transmitter or the receiver, produces both a frequency shifting and spreading. It increases as the ratio v/c, being v the relative speed of the transmitter and receiver and c the sound speed in water. Equipment such as Autonomous Underwater vehicles (AUV) has a low speed (few meters per second) but also unintentional effects, such as tides and currents, have to be taken into account. Synchronization might be thus difficult to achieve.

1.1.1 Network Design Principles

We now discuss briefly general telecommunication principles that may be useful for the following analysis.

It is a known fact that an information network can be characterized by means of a layered architecture, with the first layers being the physical layer, the data link layer (DLL) and the network layer.

The physical layer converts the bits into signals and sends them over a channel on the transmitting side, while on the receiving side it decodes the signals into bits, and it should detect if the former have been corrupted by noise or interference.

The next layer is the DLL, which performs Framing, that is, the task of outlining the parts of a packet (e.g., payload, source and destination addresses), and Error Correction. The latter is generally performed through Cyclic Redundancy Checks (CRC), adding some bits - function of the other bits in a packet - at the end of the packet itself. The only purpose of them is to be checked on the receiving side to detect possible errors after reception. Automatic Repeat Request (ARQ) may follow to request retransmission of corrupted packets. Finally, if the medium is shared between more than two nodes,

Medium Access Control (MAC) protocols, such as ALOHA and Carrier Sense Media Access (CSMA) protocols, are responsible for avoiding collision events. Since in a network usually the nodes spend more time being idle than actually communicating, the frequency and time resources should be managed cleverly. Multiple Access Methods deal with this issue. The most important ones are the following.

- The *Frequency Division Multiple Access* (FDMA), which partitions the bandwidth into multiple segments, each of which is granted to a node. In UWA networks, the not negligible fading makes FDMA difficult to use, but this issue is mitigated by coding [7]. Still, when the traffic is bursty, FDMA may not be the best choice in our scenario of interest [8].

- The *Time Division Multiple Access* (TDMA), where time intervals are divided into slots, each of which is allocated to a single node. Guard times are employed to avoid collisions of packets from contiguous slots. The main advantages of this method are the possibility of saving energy by turning off a transmitting node during idle periods - since the communication with TDMA is generally bursty - and the fact that the hardware for all the nodes is the same, so data rates could be modified by adjusting the slots duration without any additional cost. However, TDMA requires precise synchronization, which is quite hard to achieve in a UWA network due to prominent delays, determining an overall low throughput.

Finally, on top of the DLL there is the network layer whose main task is routing the packets through a certain path linking the source and the destination nodes. The optimal path is found following a specified policy (i.e., we want to use the path that minimizes the average delay between the transmission of a packet and its reception or the total number of hops). The Dijkstra and Bellman-Ford algorithms are two of the most well-known static routing procedures.

As far as network topologies are concerned, there are three main architectures.

- A *centralized* network consists of end nodes which are individually linked to a primary node that is, in turn, connected to some kind of backbone. A critical aspect of this architecture is its unique failure point, as if the central node fails, the communication in the entire network is disrupted [9]. However, this topology can be employed quite straightforwardly to underwater acoustic networks: we may have a surface node (central station) that receives packets from sensors placed in the water via the acoustic channel, and the former could then forward the received data to a server or a similar entity, being connected to a traditional backbone.

- A *distributed, fully connected* network is made of nodes, each of which is directly linked to all the others. Routing would no longer be necessary, but the energy cost of such an infrastructure would be too high.

- A *multihop* network comprises nodes that are directly connected only to their immediate neighbors, and to reach the final destination, a packet should hop from node to node (following the routing policy). A disadvantage is that the delay increases with the number of hops, and we should pay attention to this depending on the specific end application.

1.2 Why an Affordable, Dense Deployment and not an Expensive, Sparse One?

The impact of climate changes on coastal areas and biodiversity is today more intense than ever, and extreme events such as floods must be fought with long-term large-scale strategies whose purpose would be to mitigate and predict these effects through monitoring. Yet, coastal systems are extraordinarily diverse in space and variable over short, medium, and long timescales, and a relevant thematic related to underwater communications is the realization of low-cost, low-power, and easy-to-handle equipment to sustain monitoring of coastal environments and maritime operations in general using low-cost sensors to retrieve data related to water quality in the areas of interest.

Due to the highly complex characteristics of water, the most suitable medium for achieving fruitful and meaningful underwater communication is the acoustic wave [10]. Other media suffer from extreme propagation loss and distortion of refraction; while electromagnetic waves could transmit messages very quickly in a (very) limited range, acoustic waves are subject to a number of detrimental effects (not negligible transmission loss, time-varying multipath propagation, Doppler spread, distance-dependent bandwidth, significant propagation delay), but the acoustic channel is the only one that allows communication at realistic distances. Unsurprisingly, this channel is one of the most complicated to work with if we aim for a robust and reliable communication.

However, the use of complex long-range, high-power acoustic modems in our scenario is prohibitive due to the high cost of this equipment. We would prefer a dense sensor deployment, which would enable us to gather a lot of data from a variety of areas - not necessarily far one from the others -, but we could not afford to build such an infrastructure with high-end commercial available devices. In fact, underwater networks are used mainly in large-scale applications, such as monitoring oil spills, remote control of underwater vehicles, and coastal protection. Their use in Internet of Things (IoT) applications, such as monitoring aquaculture and bathing sites, is not practical due to the high cost and power consumption of most commercial acoustic modems, which are usually designed to achieve long-range transmissions [11, 12, 13] rather than being used in small low-power nodes.

Unlike these legacy applications, in the last ten years, researchers [14, 15, 16, 17] as well as modem manufacturers [18, 19] have started to develop low-cost and low-power acoustic modems to bridge this technological gap that has made the use of acoustic modems so far tricky in dense sensor network deployments.

Furthermore, the use of long-range high-power radio antennas to send the data collected by surface nodes to shore should be avoided: we would like to minimize the power consumption in order to extend the battery duration of the nodes. In this context, the use of Low-Power Wide-Area Network (LPWAN) solutions seems to be a perfect fit.

1.3 The Need for a Statistical Channel Model in Underwater Simulations

Due to the various challenges resulting from the complex environment where an Underwater Acoustic Network (UAN) should operate, the most accurate way to evaluate its performance is to perform a sea experiment where the UAN we are interested in is actually deployed and its activity is analyzed thoroughly. Although sea trials are proven

to be the best way to evaluate UANs, their realization is not trivial; in fact, they are very demanding in terms of costs, time, personnel and equipment and are very prone to external factors that can cause failures of the trial, not only due to equipment issues caused by software failures and hardware damages, but also due to bad sea conditions. For these reasons, network simulators are often employed for a preliminary evaluation, in order to debug the protocol stack before the final sea trial, hence minimizing the probability of software faults and having an idea on how the new protocol works if compared to other benchmarks. However, in the underwater research community, simulations are still not considered a valuable tool to perform the final evaluation of UANs, as channel models are often unable to accurately describe the time-varying behavior of a real underwater acoustic channel [20]. In fact, the acoustic channel depends on several factors. First, changes in temperature, depth of the node, and salinity cause a variation in the sound speed along the water column and, therefore, propagation of the acoustic signal. Second, the presence of water currents, wind, and mobile nodes causes a strong Doppler effect that affects the received signal [21]. Last, the noise caused by wind waves, rain, snapping shrimps, bubbles caused by tidal inflow, and ship propellers [22] causes degradation of the signal-to-noise ratio (SNR). The use of realistic channel models, such as the Bellhop ray tracer [23] where a subset of these parameters can be included, is computationally demanding and, hence, restricted to networks with a limited number of nodes.

Given the large number of sea experiments performed by scientists in the last 15 years [1, 24, 25, 26, 27], a wide dataset of time-varying links has been collected and some measurements are publicly available. Data-driven models have gradually been used to predict the trend of channel performance; for example, in [28] the authors, considering as features the different environmental characteristics of the model, build a logistic regression network whose Packet Success Rate (PSR) estimates are quite accurate if restricted to the short-term variability of only one of the acoustic link features used to build the regression network. In several works [29, 30, 31] the authors mapped different modem performance figures of PSR versus range in the DESERT network simulator [32]. Although in some cases they have also included performance degradation due to interference, this model can only be used for a preliminary evaluation of the network, as the channel variability is not considered, and the modem performance is assumed constant in time.

The ASUNA dataset [1] is a collection of the evolution of acoustic link quality time evolution observed during many different sea trials carried out by Haifa University, Israel, the University of Padova, Italy and IMDEA Networks, Spain. These experiments have been carried out in different locations in Europe and Israel. The authors also show how the time-varying links stored in the dataset can be used in a Matlab network simulation to reproduce the link quality evolution experienced during those sea trials. Similarly, in [33] the authors included in the DESERT underwater network simulator the time evolution of the links of the multimodal acoustic mobile ad hoc network deployed in [26] and composed of low-frequency and high-frequency modems. They also included the impairments caused by interference and LUTs of the noise variability to test the adaptation of different modulation and coding schemes. Although, on the one hand, both solutions in [1] and [33] allow to reproduce the time evolution observed during sea trials, on the other hand they do not allow to test different channel realizations.

During the last decade, researchers [34, 35] demonstrated that the time evolution of underwater acoustic channels can be statistically well characterized with two- and four-state Markov models and with a two-state Hidden Markov Model (HMM) [36]. In fact,

the nature of the acoustic channel, whose error probability often changes during the day due, for instance, to the presence of rain, changes in wind speed and shipping activity, can be well described by HMM. Indeed, when analyzing real channel measurements [1, 24, 25, 26, 27] it is common to observe time intervals with a low PSR alternated by time intervals with a high PSR, rather than having an almost constant error probability during the whole experiment.

The evaluation study of which Markov and HMM models best fits the experimental data [35] showed that the HMM yields an accurate reproduction of channel metrics, following long-term channel behaviors, and making it a good choice for modeling the channel in UAN simulators.

Thus, we also present a statistical model based on the analysis of sea trial data, and to evaluate the effectiveness of this model with respect to already existing models. This statistical model is included in the DESERT Underwater simulator [32], that includes a wide set of protocols for best customizing the underwater network to the needs of a user. The model relies on measurements extracted from the ASUNA dataset [37], which presents a number of time series of link quality indicators (LQIs), measured during the aforementioned experiments. The model has been developed to provide the research community with an open-source framework for underwater network simulations where the acoustic channel is modeled with high reliability and low computation complexity.

Chapter 2

A Statistical Model for Acoustic Channel Variability

In this section we discuss our implementation of a statistical channel model based on Hidden Markovian Models as a physical layer of the DESERT simulator. We first discuss how the model has been designed and implemented in DESERT in Section 2.1 and we then show how we set up the simulations, together with relevant results, in Section 2.2.

2.1 Model Design and Implementation

2.1.1 Dataset Description

The statistical model used in this Thesis is trained using the measurement data of one of the sea trials collected in the ASUNA dataset: the Haifa harbor (Israel) test performed in May 2009 [38]. During this experiment, 4 meter rubber boats deployed the nodes in six distinct topologies for different periods of time. A spatial reuse TDMA protocol (each device had a 5 second slot dedicated for transmission) was tested, and the transmission rate of the modems was 600 bps without channel coding, using a B-PSK signal modulated by direct sequence spread spectrum (DSSS), which was created using a gold sequence-based pseudo random sequence of 128 chips, centered at 25 kHz, and bandwidth 5 kHz. The modem prototype was composed of ITC transceivers, a National Instrument data acquisition system, and a laptop for signal processing. The transceivers were deployed at a depth of 4 m.

The LQI observed during the trial is the Bit Error Rate (BER), defined as the ratio between the number of erroneous bits and the total number of transmitted bits. The dataset provides a set of time-varying BER per-link values collected into six Topology Matrix Information (TMI) (one for topology). A TMI consists of an `NxN` matrix, with N the number of nodes in a topology, where the entry (t, i, j) represents the BER value for the link from node i (transmitter) to node j (receiver) at time t: the time interval between two subsequent measurements is 5 s, at each measurement BER and GPS position (in UTM coordinates) of each node are recorded. During the sea trial, Topology 1 was tested for 30 minutes, Topology 2, 3, 4 and 5 were tested for 60 minutes while topology 6 was tested for 90 minutes. Table 2.1 provides the experiment details.

During the experiment the LQI of each link was varied in time. In some of the links

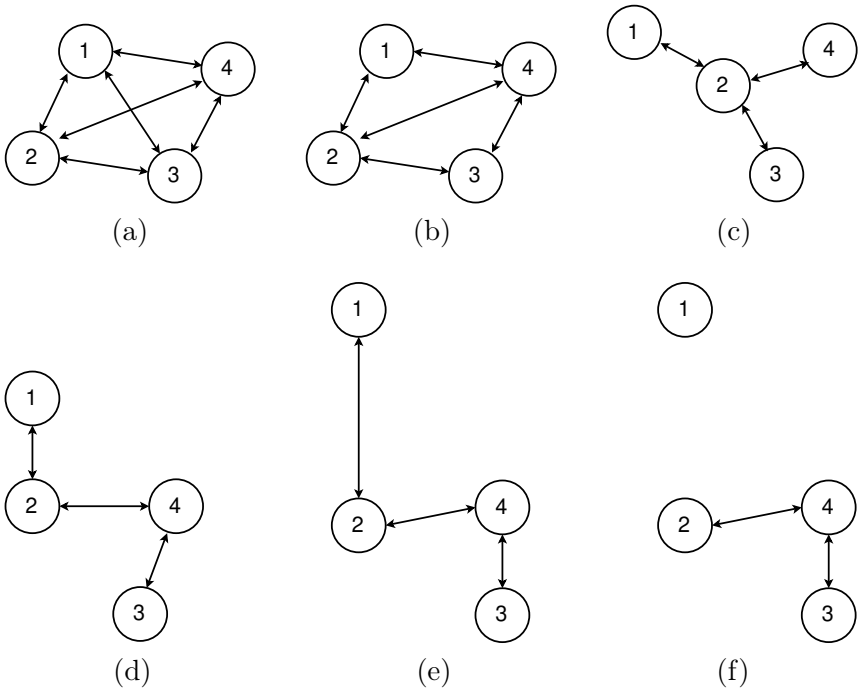


Figure 2.1. Topologies tested in the sea trial [1]: topology 1 (a), topology 2 (b), topology 3 (c), topology 4 (d), topology 5 (e), and topology 6 (f).

Table 2.1. Haifa Harbor sea trial details.

Location, Date	**Nodes**	**Topologies**	**Collection Time**
Haifa Harbor, 05/09	4	6	30-90 minutes
Rate	**LQI**	**Total Time**	**Interference**
Once every 5 s	BER	6 hours	No

the BER was very small for almost all the time, while other links had a higher error rate.

For instance, in Figure 2.2 we can observe the BER Cumulative distribution function (CDF), fitted with an exponential distribution, of three representative links observed during the trial.

Specifically, Figure 2.1.1 presents the CDF of the very stable link from node 4 to node 2 observed in topology 2, whose BER is lower than 0.02 for 90% of the time. Figure 2.1.1, instead, presents the CDF of the link from node 3 to node 2 observed in topology 2: in this case the BER is slightly higher than in the previous case but never exceeds 0.06. Finally, Figure 2.1.1 depicts the link from node 1 to node 3 observed in topology 3: this link has a BER that is definitely higher than the other two links.

2.1.2 Three-State Hidden-Markov Model

In this section we analyze the data measurements in order to obtain the statistic characterization of the acoustic channel experienced during the sea trial (Section 2.1.2) and

compute the transition probabilities of the three-state HMM used to model the channel variability (Section 2.1.2). We also present the two-state HMM used as benchmark (Section 2.1.2). We analyze only the time evolution of the acoustic links of the nodes in communication range of each other, as nodes that are not in range simply did not exchange any message and their analysis is therefore trivial. Although the analysis presented in this Thesis focuses on topologies 1 and 2, in the new release of DESERT we also included the link evolution statistics of topologies 3, 4, 5 and 6.

BER Thresholds

In order to analyze the link quality, we need to define when a link is assumed to be in "good", "medium" and "bad" state. For this reason, we set the following thresholds to the observed BER:

- **Good state**: $BER < 0.012$;
- **Medium state**: $0.012 < BER < 0.025$;
- **Bad state**: $BER > 0.025$.

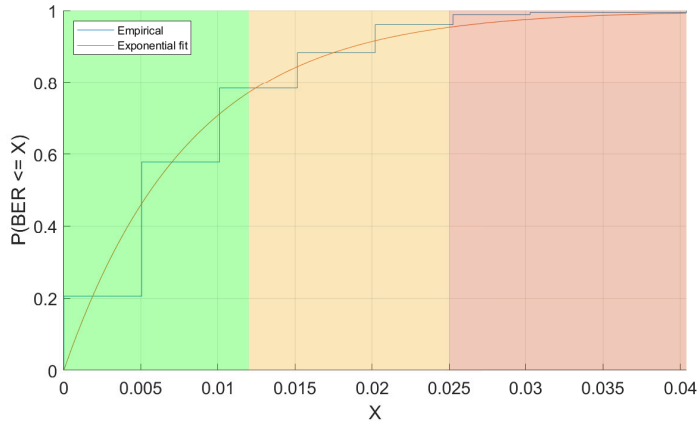
With these thresholds, considering a Hamming(7,4) Forward Error Correction (FEC) and a packet size of 16 bytes without FEC (i.e., 28 bytes with FEC), the resulting Packet Error Rate (PER) can be computed analytically as follows. If we define the probability of having no more than one error in 7 bits as

$$\mathbb{P}_{succ} = (1 - BER)^7 + 7 \cdot BER(1 - BER)^6, \tag{2.1}$$

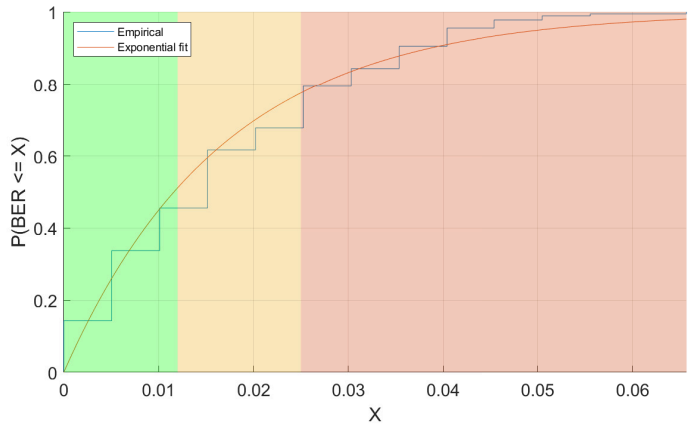
we can obtain

$$PER = 1 - (\mathbb{P}_{succ})^{224/7}. \tag{2.2}$$

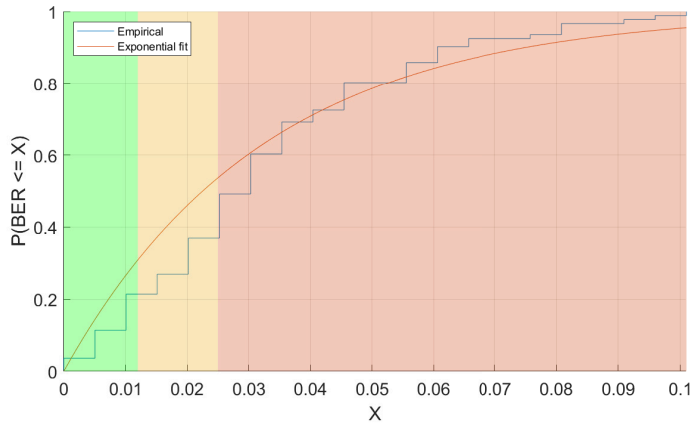
To check that these results are correct, we verified them via simulation. Given a topology, a link and its empirical BER observed at a fixed time t_x, a sequence of 224 uniform random values in $[0, 1]$ are extracted. Each of the values has been compared with the respective BER empirical value to generate a logical array with "0" in the cells where the number generated by the RNG was greater than the BER value, and "1" in the other positions.



(a)

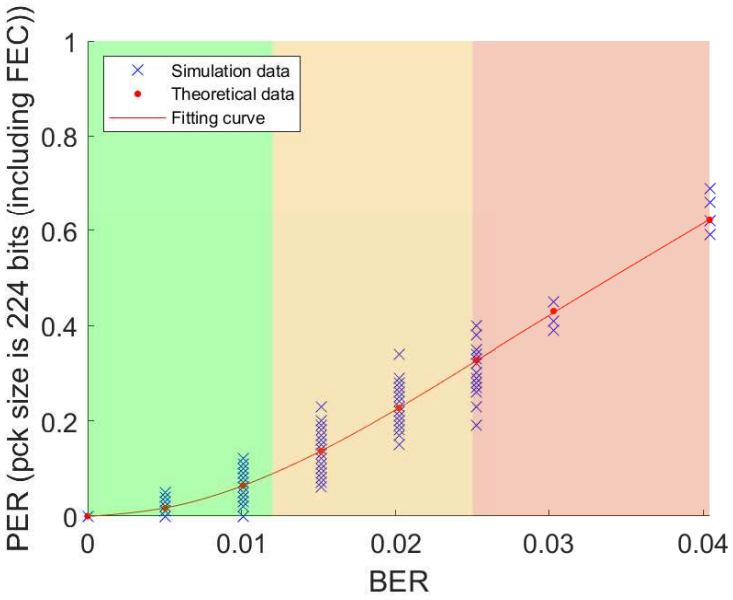


(b)

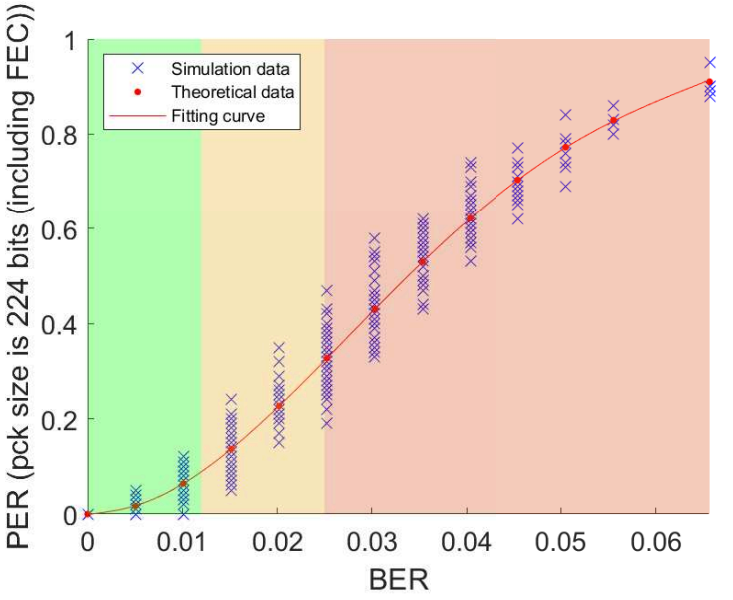


(c)

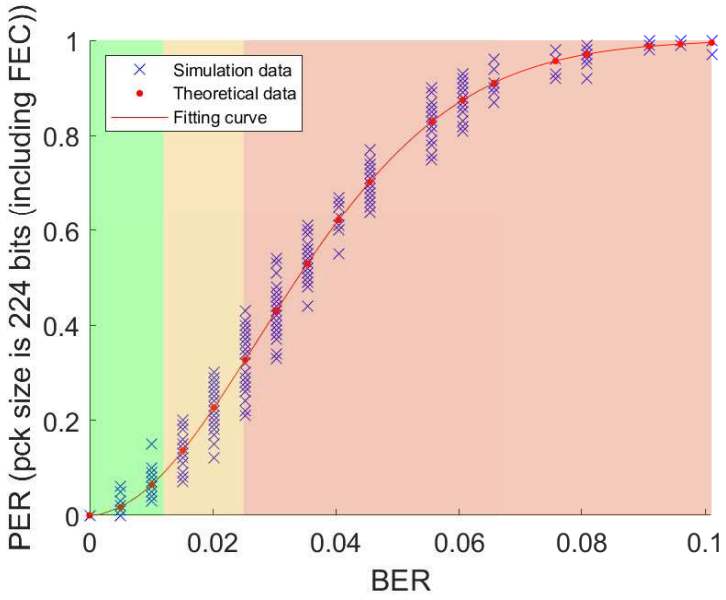
Figure 2.2. Examples of BER CDF fits for the stable link from node 4 to node 2 observed in topology 2 (a), the average link from node 3 to node 2 observed in topology 2 (b), and the challenging link from node 1 to node 3 observed in topology 1 (c).


Simulation data
Theoretical data
Fitting curve
PER (pck size is 224 bits (including FEC))
BER
0
0.2
0.4
0.6
0.8
1
0
0.01
0.02
0.03
0.04

(a)


Simulation data
Theoretical data
Fitting curve
PER (pck size is 224 bits (including FEC))
BER
0
0.2
0.4
0.6
0.8
1
0
0.01
0.02
0.03
0.04
0.05
0.06

(b)



(c)

Figure 2.3. PER vs BER values considering a packet with 16 bytes payload and Hamming(7,4) FEC for three different links: the "Good" link from node 4 to node 2 observed in topology 2 (a), the "Medium" link from node 2 to node 2 observed in topology 2 (b), and the "Bad" link from node 1 to node 3 observed in topology 1 (c).

This array can be interpreted as our 224 bit packet, where the bits set to "1" are wrong and the bits set to "0" are correct. Therefore, since we have adopted Hamming(7,4), the packet is scanned with a 7-bits step: since Hamming(7,4) cannot correct more than one error every 7 bits, whenever the sum of the bits in a block is greater than 1 we mark the block as compromised and the whole packet is considered corrupted. The process is iterated for $N = 1000$ times and the PER value is given by the number of corrupted packets divided by N.

We can observe in Figure 2.3 the PER-BER relationship obtained analytically (red line) and via simulation (blue crosses) for the three links presented in Figure 2.2.

With the BER thresholds presented above, the corresponding PER thresholds follow:

- **Good state**: $PER \leq 0.09$;
- **Medium state**: $0.09 < PER \leq 0.32$;
- **Bad state**: $PER > 0.32$.

We can finally observe that the stable link from node 4 to node 2 observed in topology 2 is 95% of the time in Good or Medium states, the average-performance link from node 3 to node 2 observed in topology 2 is only 80% of the time in Good or Medium states, and the challenging link from node 1 to node 3 observed in topology 1 is in Bad state 45% of the time.

With these fits, we can compute the generic probability that a link is in one of the three-states. Nevertheless, this is not enough to model the variability of the channels.

Transition Probabilities

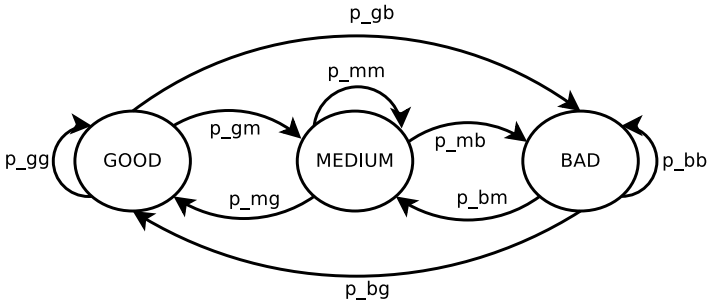


Figure 2.4. Three-state channel model.

From a visual inspection of the link BER time evolution we noted that, grouping the data on a per-state basis, a link in a state i is more likely to remain in that state in the successive time slot, rather than jump to another state. Once this was verified, we decided to model the PER time evolution of a generic link as a three-state Markov chain (Figure 2.4), with the three states $S = \{G, M, B\}$ that stand for "Good", "Medium", and "Bad", respectively. Specifically, if we denote as $X_0, \ldots, X_n, \ldots X_N$ a sequence of random variables where X_i takes values in the set S of the three states, $\mathbb{P}(X_{n+1} = j | X_n = i)$ is the *transition probability* from state i to state j at step n. Additionally, by the Markov property, we have that:

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \ldots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n), \tag{2.3}$$

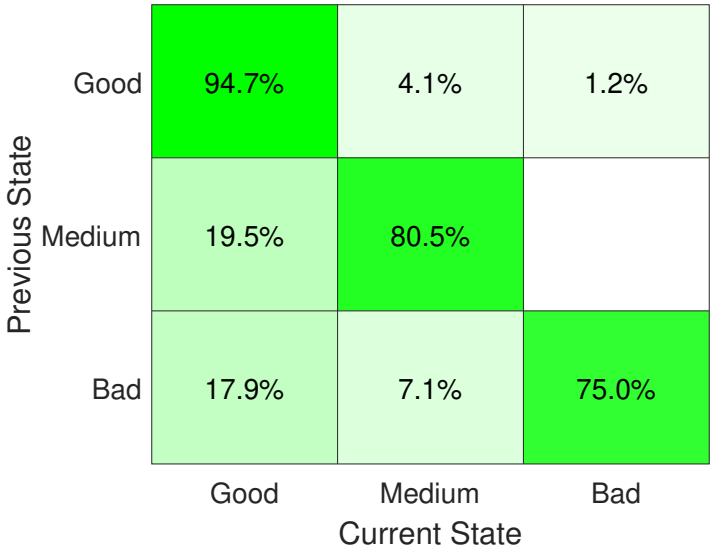
which can be interpreted as the fact that, if the current state $X_n = i_n$ is known, the probability of $\mathbb{P}(X_{n+1} = i_{n+1})$ does not depend on the previous states. If the transition probabilities do not depend on n but only on i and j, the Markov chain is *homogeneous* and we may compute every joint probability knowing only the initial distribution of the states $p_i^{(0)} = \mathbb{P}(X_0 = i)$ and the values of p_{ij}, where:

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i), \forall n. \tag{2.4}$$

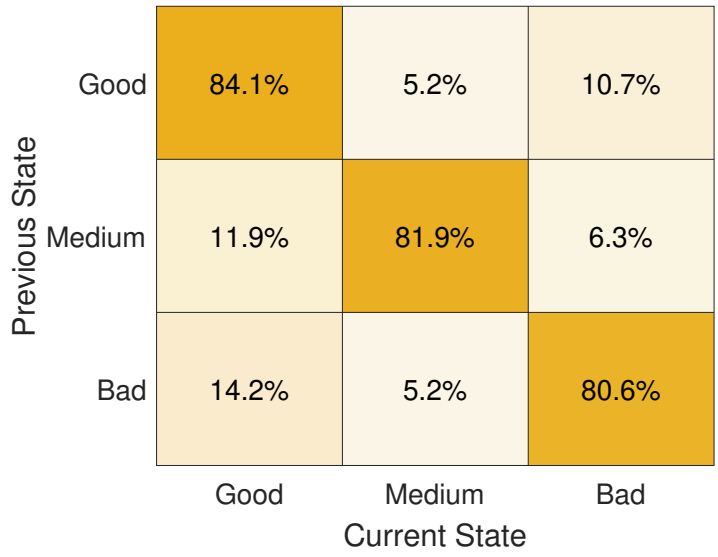
Exploiting matrix calculus, since we knew the frequencies of the BER values of each link, we found the *transition matrices* $P = (p_{ij})$, which have only non negative elements, are row-normalized to 1 and, in our case, have a size `3x3`. In Figure 2.5 we show the matrix charts presenting the transition matrices for the three links discussed so far.

A relevant result is that, given the transition matrix P^n at time n, it is possible to compute the t-step transition probabilities by means of matrix exponentiation:

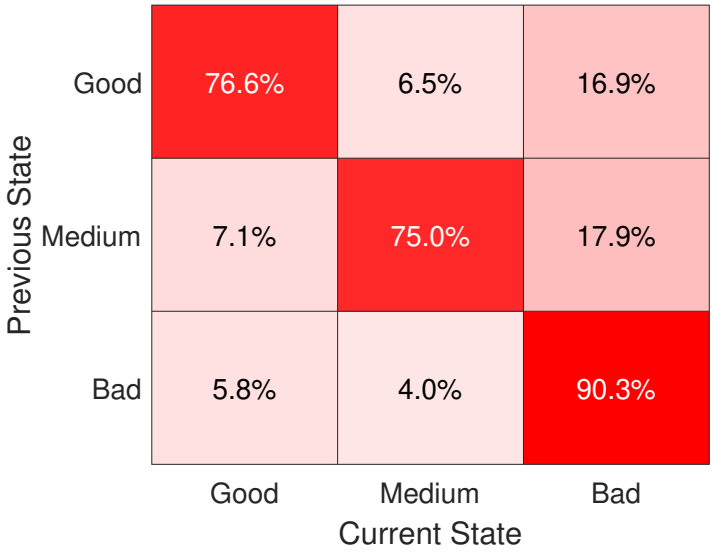
$$\mathbb{P}(X_{n+t} = j | X_n = i) = (P^t)_{ij}, \forall n \geq 0. \tag{2.5}$$


Good
Medium
Bad
Previous State
94.7%
4.1%
1.2%
19.5%
80.5%
17.9%
7.1%
75.0%
Good
Medium
Bad
Current State

(a)


Good
Medium
Bad
Previous State
84.1%
5.2%
10.7%
11.9%
81.9%
6.3%
14.2%
5.2%
80.6%
Good
Medium
Bad
Current State

(b)



(c)

Figure 2.5. Examples of transition matrices: transition matrix P of the links from node 4 to node 2 observed in topology 2 (a), from node 3 to node 2 observed in topology 2 (b), and from node 1 to node 3 observed in topology 1 (c).

The averaged values of the BER in the three states for the links we are considering are reported in Table 2.2.

Table 2.2. Average BER values, three-state HMM.

	avg good	**avg medium**	**avg bad**
Topology 2, link 4→2	0.0051	0.0174	0.0281
Topology 2, link 3→2	0.0048	0.0165	0.0338
Topology 1, link 1→3	0.0066	0.0184	0.0448

Two-State Hidden-Markov Model

As benchmark of the three-state HMM presented in Section 2.1.2, we now present the more used two-state HMM (Figure 2.6).

In the two-state model, we define a cumulative Bad state b' grouping together the Bad and the Medium states used in the three-state model. The probabilities of successful reception given a channel state are computed link-wise by taking the average PERs in each state. The transition probabilities, instead, are computed starting from the three-state model transition probabilities as:

- $p_{gb'} = 1 - p_{gg}$,
- $p_{b'g} = \frac{p_{mg} \cdot p_m + p_{bg} \cdot p_b}{1 - p_g}$,

- $p_{b'b'} = 1 - p_{b'g}$,

where p_{gg} is the probability of not having a transition at time $n+1$ when a link is in the Good state at step n for the three-state HMM and $p_s, s \in \{g, m, b\}$ is the generic probability a link finds itself in the Good, Medium or Bad state respectively.

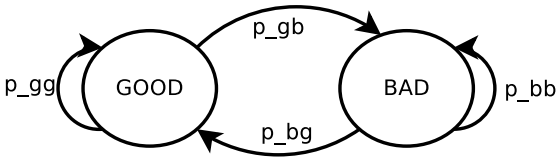
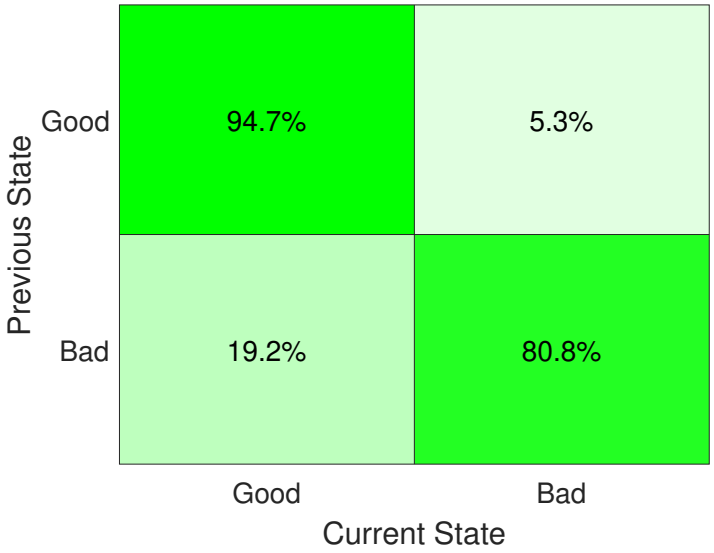


Figure 2.6. Two-state channel model.

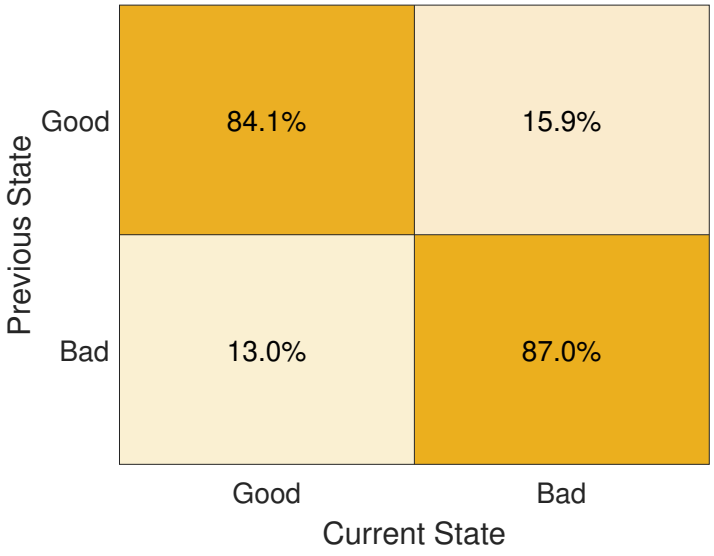
While with the three-state HMM the transition matrix P^n at step n needs to be computed with matrix exponentiation as presented in eq. (2.5), in the simple two-state model the transition probabilities at step n can be obtained via the closed formula [39]:

$$P^n = \frac{1}{p_{gb'} + p_{b'g}} \begin{pmatrix} p_{b'g} & p_{gb'} \\ p_{b'g} & p_{gb'} \end{pmatrix} + \frac{(1 - p_{gb'} - p_{b'g})^n}{p_{gb'} + p_{b'g}} \begin{pmatrix} p_{gb'} & -p_{gb'} \\ -p_{b'g} & p_{b'g} \end{pmatrix}. \tag{2.6}$$

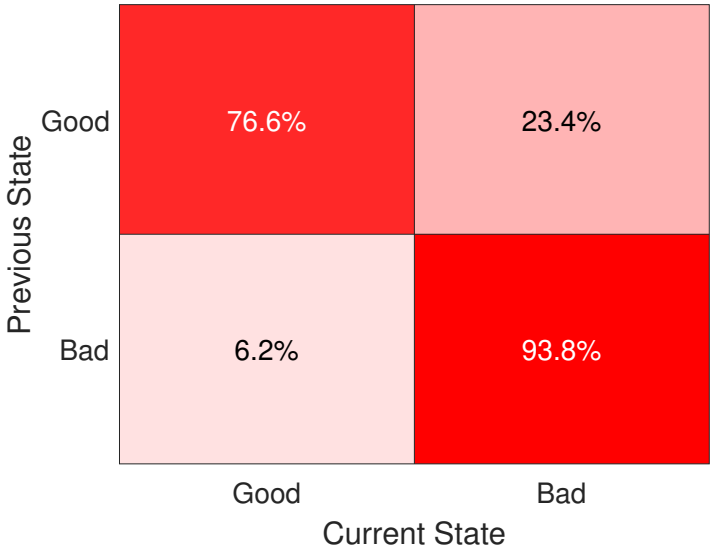
In Figure 2.7 we report the transition matrices in the two-state HMM for the links under analysis.



(a)



(b)



(c)

Figure 2.7. Examples of transition matrices for the two-state HMM: transition matrix P of the links from node 4 to node 2 observed in topology 2 (a), from node 3 to node 2 observed in topology 2 (b), and from node 1 to node 3 observed in topology 1 (c).

Table 2.3 shows the relevant averaged BER values for the two-state HMM.

Table 2.3. Average BER values, two-state HMM.

	avg good	**avg bad**
Topology 2, link 4→2	0.0051	0.0193
Topology 2, link 3→2	0.0048	0.0267
Topology 1, link 1→3	0.0066	0.0395

2.1.3 Model Implementation

In order to evaluate the models presented in the previous sections, we implemented the two- and three-state HMM in the DESERT Underwater framework [32], an open-source underwater network simulation and experimentation tool publicly available in [40]. Notably, the DESERT Underwater legacy physical module, called UWPhysical, models the path-loss with the Urick and Thorp formulas, and computes the signal to noise ratio using the model presented in [20]. Although this model is largely used by researchers, it does not address well the variability of the acoustic channel. Therefore, we implemented from scratch two new physical layers, one called UWHMMPhysical that uses the two-state HMM described in Section 2.1.2, and one named UWHMMPhysicalExtended that uses the three-state HMM presented in Section 2.1.2. In both physical layers we included the statistics of each link using the so called `link-stats` objects, and let the physical layer compute the probability that a packet is correctly received at a specific moment, hence providing a per-link channel variability. The `link-stats` objects are independent of each other: in the case of near nodes that share the same channel, the same `link-stats` object can be used to model the channel variability in the same way: in the case of the sea experiment considered in these simulations, the links between the nodes are considered independent, hence a different `link-stats` object is used to model the channel variability between every pair of nodes.

The most relevant difference between the two- and three-state HMM is the way the transition probabilities are computed. As explained in Section 2.1.2, the two-state HMM can be computed via a closed formula, while for the three-state HMM the transition probability can only be computed by means of matrix exponentiation. This implies that the exponentiation has to be performed efficiently, so that even with a big exponent n, the complexity is limited and not growing without bounds. Given that n monotonically increases during the simulation, it is not necessary to compute P^n starting from the initial transition matrix P^0, as this would cause a degradation in performance. Specifically, we save the aforementioned matrix each time we compute it, so that we can operate conveniently on the last available P^k and compute P^n with a number of exponentiations equal to $n - k$, that is strictly less than n. As a result, the computation time of a simulation using the three-state model is not much longer than the same simulation relying on the legacy physical model or on the two-state HMM.

2.2 Simulation and Results

2.2.1 Simulation Setting

In our simulations we analyze the system behavior with the nodes placed in the positions presented in topology 1 (Figure 2.1) and topology 2 (Figure 2.1). The simulation lasted 18000 s, and we switched from topology 2 to topology 1 in the middle of the simulation (i.e., at time 9000 s) by adding the link from node 1 to node 3 and changing the packet success probability per link and the transition probabilities of every link accordingly. The behavior of the three communication stacks depicted in Figure 2.8 is analyzed. All stacks use a constant bitrate application layer, static routing with all nodes transmitting to their 1-hop neighbors and a time division multiple access (TDMA) MAC layer. The first stack (Figure 2.2.1) uses the legacy DESERT physical layer, the second stack (Figure 2.2.1) uses the two-state HMM-based physical layer and, finally, the third stack (Figure 2.2.1) employs the three-state HMM-based physical layer.

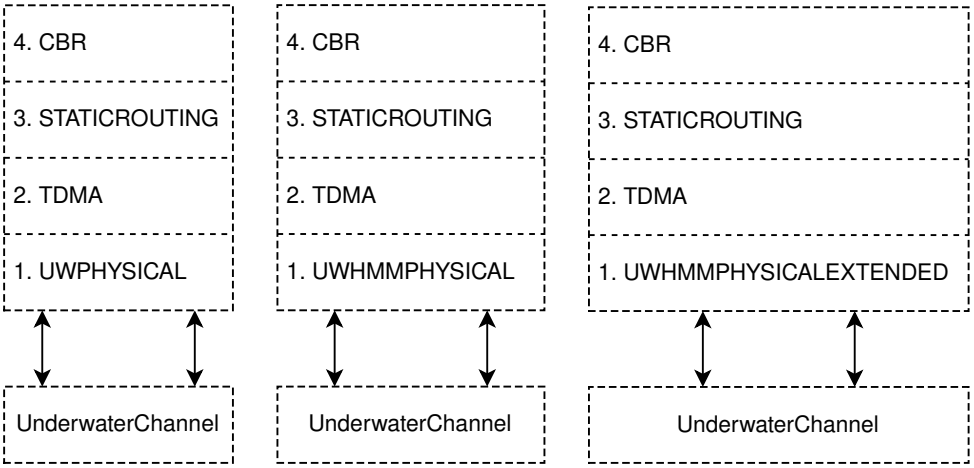


Figure 2.8. The three communication stacks compared in simulation, all composed of a constant bitrate application layer, static routing, and TDMA, and a different physical layer: `Uwphysical` (a), `UWHMMPhysical` (b) and `UWHMMPhysicalExt` (c).

The network is composed of 4 nodes and each node generates 28 bytes packets every 60 s. Bandwidth and carrier frequency are set to 5 kHz and 25 kHz, respectively, in order to best simulate the behavior of the modems used in the field experiment presented in Section 2.1.1. The simulation parameters are summarized in Table. 2.4.

Table 2.4. Simulation parameters.

Nodes	**Pkt Size**	**Tx Duration**	**Tx Power**
4	28 B	18000 s	165 dB
Frequency	**Bandwidth**	**Bitrate**	**Cbr Period**
25 kHz	5 kHz	600 bps	60 s

The TDMA MAC is configured with a frame duration of 8 s, equally divided between the four nodes that have a time slot of 2 s each to transmit their packets. A guard time

of 0.8 s is used to avoid interference caused by the propagation time and to consider possible synchronization errors between the nodes.

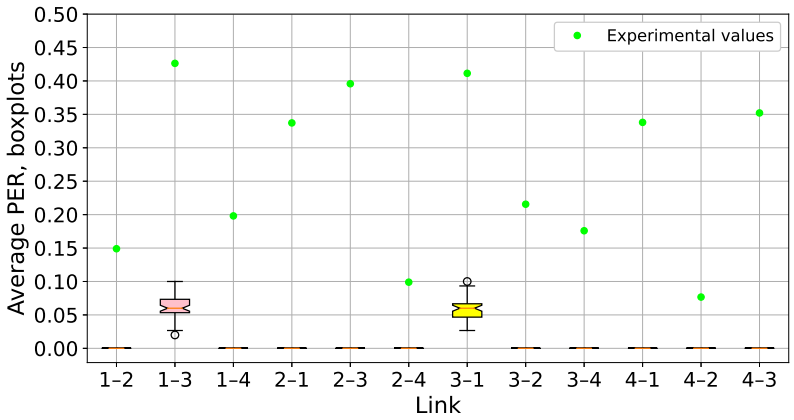
At the end of the simulations we observed the performance of each link of the network by computing PER and throughput averaging over 50 simulation runs and presenting the 95% confidence interval (CI).

2.2.2 Simulation Results

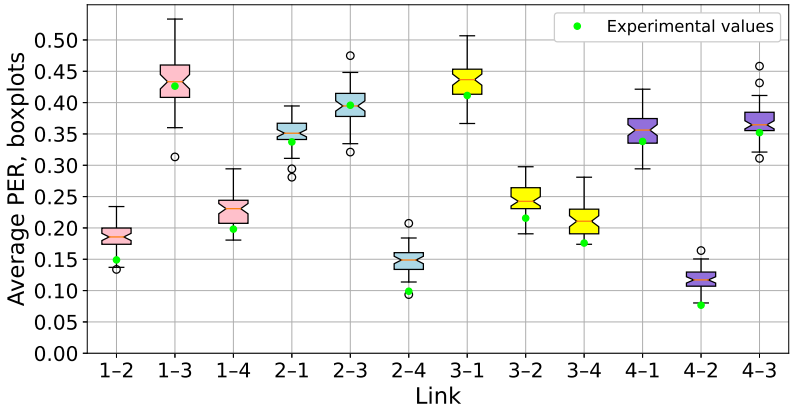
PER and throughput of each link are presented in Figures 2.9 and 2.10, respectively. Figure 2.9 compares the PER per link obtained with the three physical layers described in Section 2.2.1 with the PER measured during the sea trial (green diamond). Uwphysical (Figure 2.2.2) is extremely optimistic and provides a very low PER. In particular, with the considered settings the PER of the links is equal to zero up to a transmission range of 1.1 km, and increases to 1 when the distance between nodes is more than 1.6 km. This implies that the link connecting the two farthest nodes (node 1 and node 3, that are 1.2 km from each other) has a non-zero PER, but still the real values are underestimated. Conversely, the PER obtained both with the two-state (Figure 2.2.2) and with the three-state (Figure 2.2.2) models is very similar to the one observed in the sea trial, with the three-state model having a PER that matches almost perfectly (within the CI) the experimental one (depicted with green diamonds), definitely outperforming the other two models.

Similarly, the throughput observed with Uwphysical (Figure 2.2.2) is almost the same for all the links, and is equal to 3.7 bps: only in the link between node 1 and 3 the throughput is approximately 1.85 bps, as that link was removed at the simulation time 9000 s, when the network topology was changed from topology 1 to topology 2. With a higher PER per link, the throughput observed with the two- and three-state HMM is significantly different link by link, presenting results that are definitely closer to those that can be observed during a sea experiment.

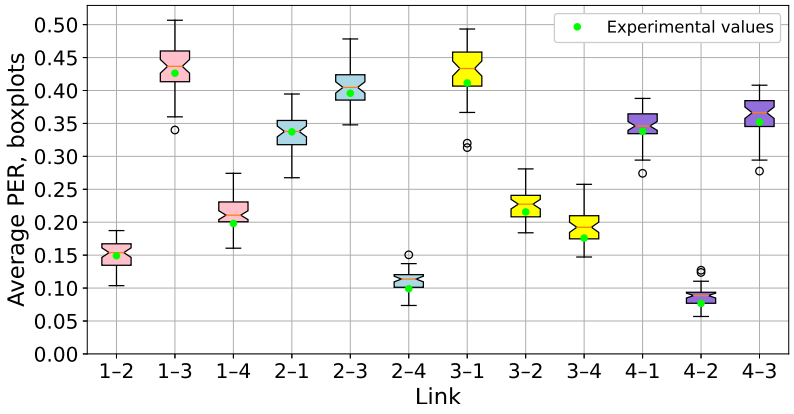
Finally, we report some plots showing the variability of throughput in time (i.e., computed every 300 seconds) for the links from node 4 to node 2 and from node 1 to node 3, and again we see how optimistic the results obtained using the simplest PHY module are. We can observe the jump at $9000s$ for the link $1 \rightarrow 3$, due to the switch from topology 2 (where the link was not in place) to topology 1. Besides, the values for the throughput are constant for the UWPhysical module, but for the links $1 \rightarrow 3$ and $3 \rightarrow 1$, which are the only ones having a PER greater than zero. Conversely, the throughput obtained with the two HMMs models is definitely lower, due to the higher PER, and has a higher variance, well characterizing the channel variability. While we could directly compare the PER obtained in simulation with that experienced during the experiment, we could not perform the same operation for the throughput, as the simulation used an application layer generating traffic with different rate than the one used during the sea trial. This tool can be used to test protocol stack configurations that are different from the one used in the experiment, exploiting the measures obtained during the sea trial to model the packet error rate time evolution and observing as a result other performance indicators, such as the throughput per link.



(a)

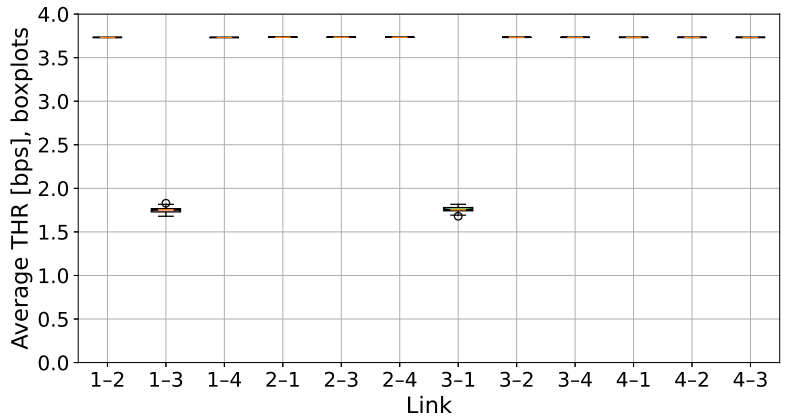


(b)

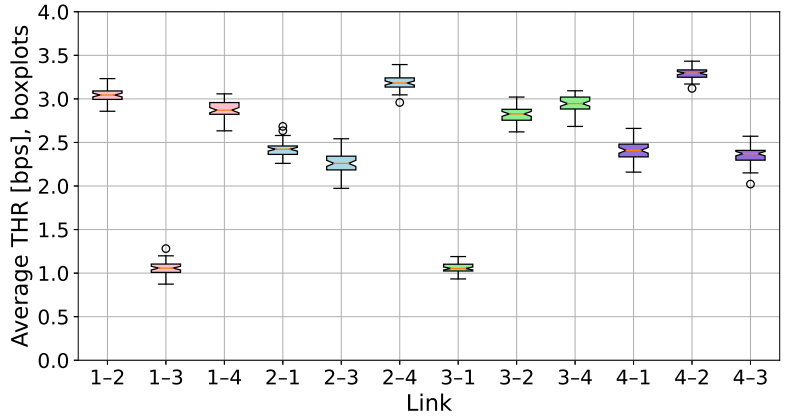


(c)

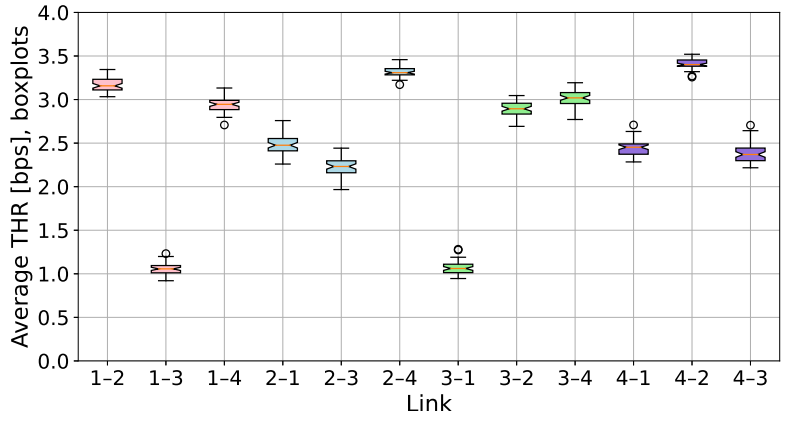
Figure 2.9. PER results yielded by the simulations (bars) with respect to Haifa Harbor measurements (green diamond) for `UWPhysical` (a), `UWHMMPhysical` (b) and `UWHMMPhysicalExt` (c).



(a)

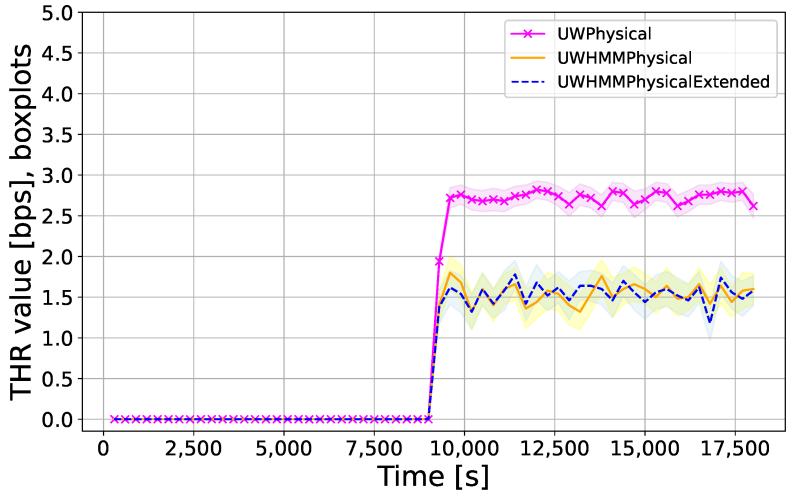


(b)

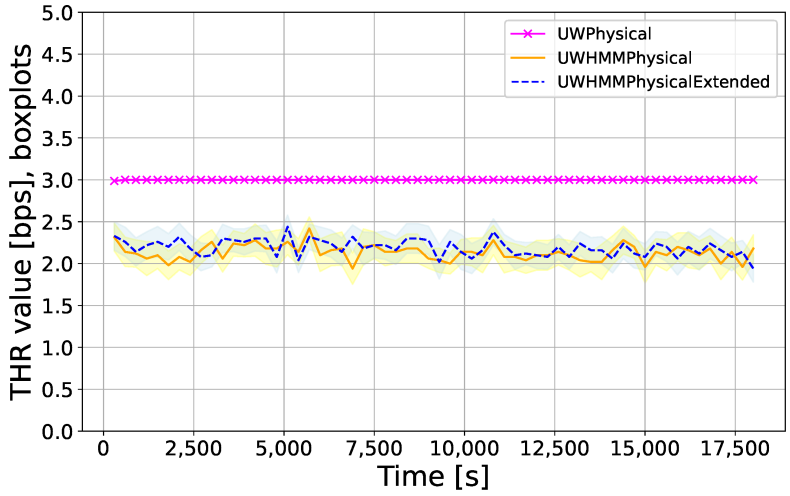


(c)

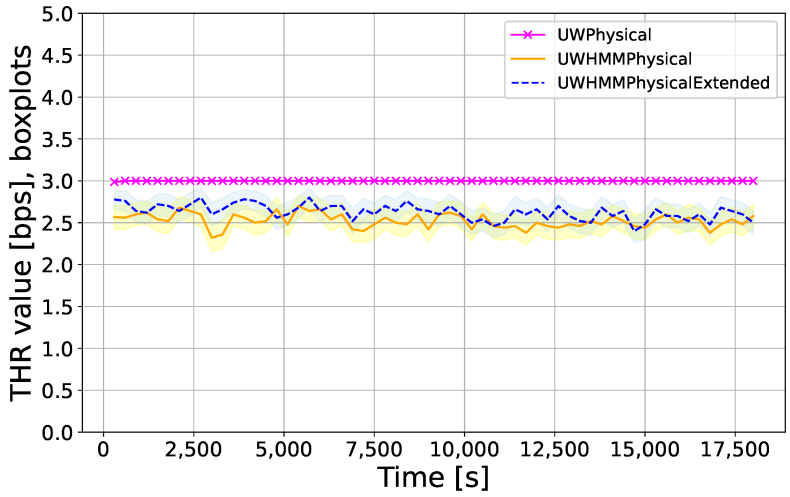
Figure 2.10. THR results yielded by the simulations with respect to Haifa Harbor measurements for `UWPhysical` (a), `UWHMMPhysical` (b) and `UWHMMPhysicalExt` (c).


UWPhysical
UWHMMPhysical
UWHMMPhysicalExtended
THR value [bps], boxplots
Time [s]
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
0
2,500
5,000
7,500
10,000
12,500
15,000
17,500

(a)


UWPhysical
UWHMMPhysical
UWHMMPhysicalExtended
THR value [bps], boxplots
Time [s]
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
0
2,500
5,000
7,500
10,000
12,500
15,000
17,500

(b)



(c)

Figure 2.11. Variability of throughput in time yielded by `UWPhysical`, `UWHMMPhysical` and `UWHMMPhysicalExt` modules for the links from node 4 to node 2 observed in topology 2 (a), from node 3 to node 2 observed in topology 2 (b), and from node 1 to node 3 observed in topology 1 (c).

Chapter 3

A Network Infrastructure for Monitoring Coastal Environments

In this chapter we will unfold the design of our envisioned network architecture, explaining in detail its functioning and we will present an operative, laboratory prototype of a low-cost, low-power sensor node. First, an overview of Low-Power Wide-Area Networks (LPWAN) is offered in Section 3.1. Traffic requirements for the architecture we propose are discussed in Section 3.2 and we present a prototype of a sensor node designed according to requirements of the network in Section 3.3.

In Figure 3.1 we depict the network infrastructure and the sensor nodes to collect the measurement data. Specifically, we envision the need for two types of surface nodes,

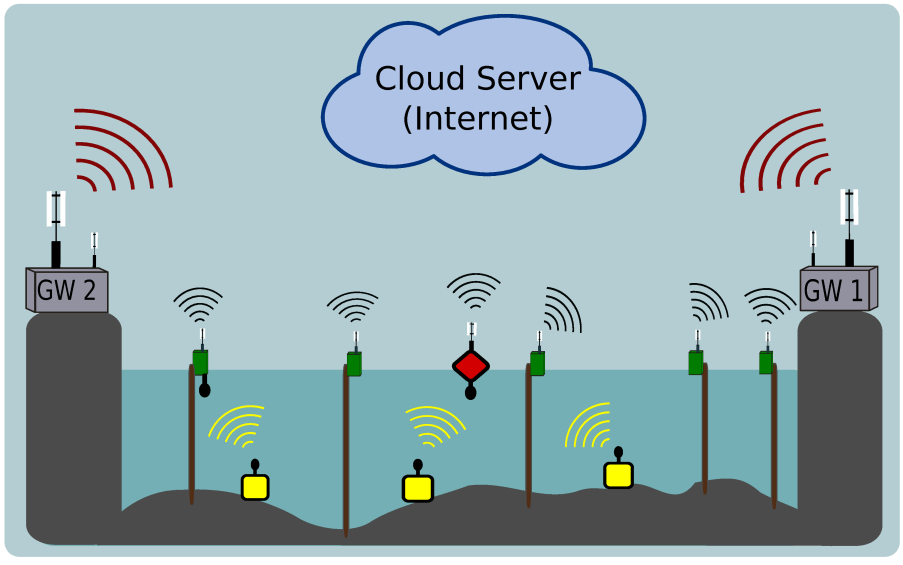


Figure 3.1. Network infrastructure.

a buoy node (depicted in red) equipped with sophisticated sensors to be deployed in a few strategic points of the area of interest, and a very low cost sensor node that can be installed either on the seafloor (the yellow nodes in Figure 3.1) or in Venetian "briccole" pillars and poles placed in the Venice lagoon (the small green nodes in Figure 3.1), and equipped with low power devices suitable for a long term dense deployment. While the former will focus on collecting data measurements of water properties that are usually constant in an area of 1 km^2, the latter will collect measurements of parameters that can strongly change within less than a hundred meters (due to morphological heterogeneity or habitat patchiness), for which a dense spatial granularity is required. The former can also be deployed in the lagoon salt marshes, that are a very interesting hot-spot to monitor floods and biodiversity. Both nodes are connected to a wireless gateway (gray node) that forwards the data to a cloud server accessible from the Internet by marine scientists and biologists. While buoys and nodes installed in pillars, from now onward called surface nodes, will transmit their data to the gateway through an LPWAN, the nodes deployed on the seafloor will first transmit their data to a surface node, using underwater communication devices (e.g., a low power acoustic link). The surface node will then act as a relay, forwarding the data to a gateway. Depending on the area to be covered, one or multiple gateways will be employed (for instance, two gateways are depicted in Figure 3.1).

The low cost sensor nodes that are part of our envisioned deployment should satisfy the system requirements in terms of types of measurements, power consumption and data transmission. The most common measurements they will collect are temperature, pressure (which can be translated to water depth), pH, turbidity, dissolved oxygen (DO) and electrical conductivity (EC). Some of these quantities are subject to rapid major variations in time (e.g., DO), while others are expected to remain more or less constant, or at least to experience very slow and smooth changes in time (e.g., temperature). This means that data measurements may be collected with a different time granularity depending on the type of sensor.

3.1 Low-Power Wide-Area Networks

Nodes located on the seafloor will transmit the data to surface nodes with underwater acoustic communication, which limits the maximum bitrate significantly, but allows long-range transmissions. Surface nodes, instead, could use standard protocols such as WiFi or GSM/LTE, but this would possibly cause the depletion of the scarce energy available to the nodes and constrain their deployment to locations covered by a WiFi gateway or the cellular network. In exchange, this solution would provide a very high bitrate, but this tradeoff is not convenient, as we are aiming at sending small amounts of data with a minimum waste of energy cost, as our devices run on batteries and we would like not to replace them too frequently, i.e., more than once a year. Thus, we decided to rely on robust protocols which have been developed specifically to handle networks akin to ours, making low-power communication in a wide area their key purpose. The three main competitors in this area are: NB-IoT, Sigfox and LoRaWAN. Their characteristics, as presented in [41], are summarized in Table 3.1. NB-IoT is likely an overkill standard for our purposes; albeit capable of guaranteeing higher data rates and bigger payload sizes, it constrains the position of the end nodes to the locations of LTE antennas. Moreover, its maximum range is the lowest among the technologies presented in the table, its resilience to interference is low, and it relies on licensed frequency bands. This

Table 3.1. Comparison of LPWAN technologies.

	NB-IoT	**Sigfox**	**LoRaWAN**
Modulation	QPSK	BPSK	CSS
Frequency	Licensed LTE	Unlicensed EU: 868 MHz NA: 915 MHz AS: 433 MHz	Unlicensed EU: 868 MHz NA: 915 MHz AS: 433 MHz
Bandwidth	200 kHz	200 kHz	125 kHz, 250 kHz 500 kHz
Max.data rate	200 kbps	100 : 600 bps	50 kbps
Message/day limit	Unlimited	140(up), 4(down)	Unlimited
Max.payload length	1600 bytes	12 bytes(up) 8 bytes(down)	243 bytes
Range urban/rural	1 km/10 km	10 km/40 km	5 km/20 km
Interference robust	Low	Very good	Very good
Security	LTE encryption	Not available	AES 128B
Adaptive data rate	N/A	N/A	Available
Private networks	No	No	Yes

means that the channel will be less affected by external interference than the ISM radio bands (used by Sigfox and LoRaWAN), but requires either cellular connectivity or a very expensive deployment due to the high cost of licensed spectrum. When comparing Sigfox and LoRaWAN, both technologies can be an appropriate fit for our network. Nevertheless, we feel that Sigfox imposes more rigorous and strict requirements on the data transmission, which may not be high enough according to the number of sensors and amount of data we have to manage. In addition, the hard limit on the maximum number of messages that can be sent per day (both in downlink and in uplink) is tight, and, although it provides the longest communication range among the analyzed technologies, it does not appear to provide relevant benefits over LoRaWAN. On the other hand, LoRaWAN seems to be a great fit for our system. The standard, supported by the LoRa industry alliance [42], is optimal when employed to connect (constrained) nodes in a network where an extended range and low battery consumption are both prime requisites [43]: for this reason we selected LoRaWAN as the best candidate for our LPWAN. It uses a Frequency Shift Chirp modulation [44] to achieve longer communication range than Frequency Shift Keying (FSK), without increasing the power consumption [45], and provides a wide set of customizable parameters which can be

different for each end device. Among these, we have the *transmission power*, the *carrier frequency*, seven *spreading factors* (SF, the higher the SF, the higher the range and the packet airtime), the *bandwidth* and the *coding rate* for the Forward Error Correction (FEC) [46].

Moreover, in [47] the authors established a LoRa communication link between a node deployed 1 meter below the water surface of a fresh water swimming pool and an in-land gateway, making it a promising technology for shallow water sensor deployments. It is quite remarkable that LoRa devices can be separated in three different "Classes" (A, B and C) depending on their energy and transmission requirements; this would be a relevant opportunity for the system we are designing, where there are extremely constrained nodes that run on small batteries (the low cost ones, typically "Class A") coexisting with more powerful devices that still should not waste energy (the buoy nodes, typically "Class B") and nodes which are plugged in and may receive as much energy as they need (the gateways, "Class C" nodes). Thus, from this analysis, we think that LoRa is the best option for the network we are developing, among those we have considered up to now.

3.2 Traffic Requirements

We envision our network to be deployed in a quite peculiar and challenging environment, the Venice lagoon. The water of the lagoon is brackish and mostly shallow, with numerous salt marshes, and an intense tidal cycle. The characteristics of this unique area make it hard for researchers to perform quantitative analysis of the water parameters with respect to a more stable environment, such as the open sea, where changes are more easily predictable. Currently, the technology used to perform measurements in the lagoon is very sparse, and the data are logged a few times per day in very few locations. Our aim is to provide the researchers with data that have a much finer granularity, both spatially and temporally. This would result in more and better data to work with. On average, a commercial LoRa gateway should be reachable by sensors in a 10-15 km range if in Line of Sight (LoS), and this is compliant with our scenario, where we do not expect to have big or tall objects (buildings, trees, etc.) to prevent the sensors from communicating with the gateway. However, we certainly have sources of interference (such as boats or even ships on major access points to the lagoon), so, in order not to make the Packet Error Rate (PER) too high, we assume that the maximum distance that would allow communication between a gateway and a sensor node is 5 km. In our simulations, we focused on a single area covered by one gateway. Also, the best distance between nodes we came up with is 500 m, which consists of a good trade-off between a high enough data granularity and the economic costs. Of course, in a real scenario, this distance should be tuned with respect to local changes: on the one hand, if we expect water parameters to change frequently and unexpectedly in a small area, we may have to decrease the distance between the sensors; on the other hand, the distance could be increased in areas where parameters vary slowly and more predictably. The same reasoning could be applied to temporal variability, that is, we may want to transmit sensor data more or less frequently according to the specific areas in which sensors are deployed. Also, some data are expected to vary more often than others, so we may want to transmit these more frequently, but not the others. The average time between data transmission from the sensors has been chosen in the range 600 s-1200 s.

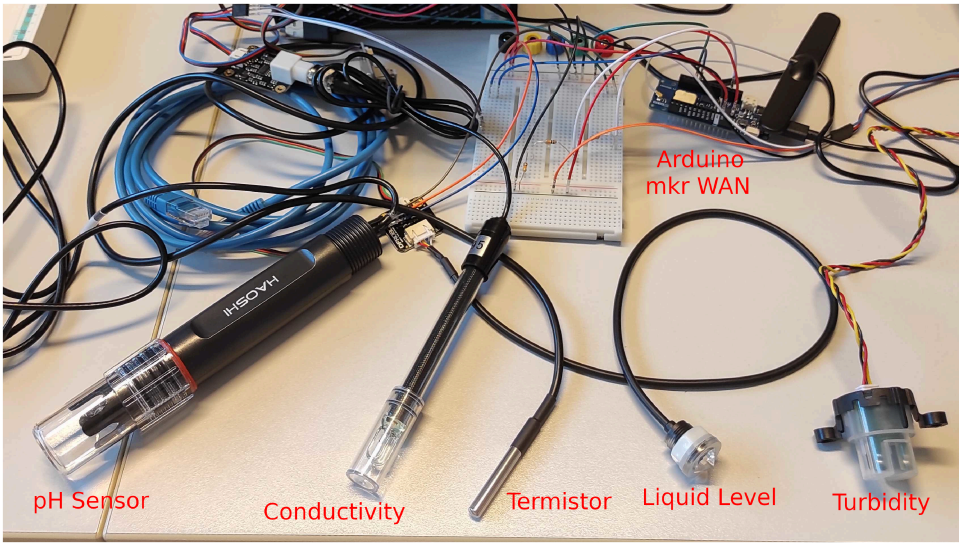


Figure 3.2. Sensor in-lab prototype.

3.3 Sensor Prototype

We decided to build a prototype of a surface sensor node according to the previous discussion; the hardware we used is shown in Figure 3.2. We can see how the prototype is only lab-grade up to now, since it is by no means weatherproof or waterproof, but for the sensors themselves, which are at least partially submersible. The core of the sensor is an Arduino MKR WAN 1310 [48], a microcontroller board with an integrated LoRa chip, chosen for its versatility, well-written documentation and extreme low-energy capabilities (the lowest power consumption is only 104 μA). The following sensors are attached to the Arduino:

- *pH sensor* [49]: used to measure pH values, needs calibration;
- *Electrical Conductivity sensor* [50]: used to measure EC values and to infer water salinity from these;
- *Turbidity sensor* [51]: used to measure the amount of total suspended solids (TSS) in water;
- *GPS sensor* [52]: used to determine the position of the node in GPS coordinates;
- *Termistor* [53]: used to measure temperature of the water with a ± 0.5°C accuracy;
- *Liquid Level sensor* [54]: used to determine if the sensor is in contact with water or not, possibly for tidal effects.

Chapter 4

Simulations

In this chapter we present the different scenarios we took into consideration, as well as the structure of the network and the parameters chosen for DESERT and ns3 simulations in Section 4.1. The relevant results are then highlighted and discussed in Section 4.2.

4.1 Simulation Scenario and Settings

The simulation of this hybrid underwater and above water network is divided in two parts, using the tool developed in [55]. Specifically, the underwater network part was simulated with the DESERT Underwater framework [32], while the LoRaWAN network with the ns3 [56] module for LoRaWAN: given that the data generated by the underwater nodes are sent to surface nodes that forward it to the cloud server using the LoRaWAN network, the output of the DESERT simulations are saved in a tracefile (containing the sequence of packets received by the surface nodes) that is used as input for the ns3 simulations. Both DESERT Underwater and the LoRaWAN ns3 module are opensource tools publicly available in [57] and [58], respectively, and are developed and maintained by the University of Padova. In particular, we analyzed the performance of the envisioned network by means of a very dense deployment that can be considered quite challenging with respect to the actual scenario, considering a fraction of the Venice lagoon. We distinguished among five different types of nodes, following the requirements seen in Section 3.2. Specifically, nodes can be either:

- *Underwater nodes*: nodes placed on the seabed collecting data from their sensors and sending them to a surface node by acoustic communication;
- *Forwarders*: surface nodes that relay the data received from a group of underwater nodes to the nearest gateway;
- *Generators*: surface nodes that generate data to send them to the nearest gateway and do not receive packet from other nodes;
- *Interferers*: surface nodes whose only role is to transmit dummy data in order to increase the load of the network, simulating the presence of LoRa nodes used for other services close to our deployment;

- *Gateways*: receiving data from all the other nodes. Note that in LoRa there is no handshake procedure between an end device and a gateway, so the former simply broadcasts relevant information, that will be received and processed by the nearest gateway.

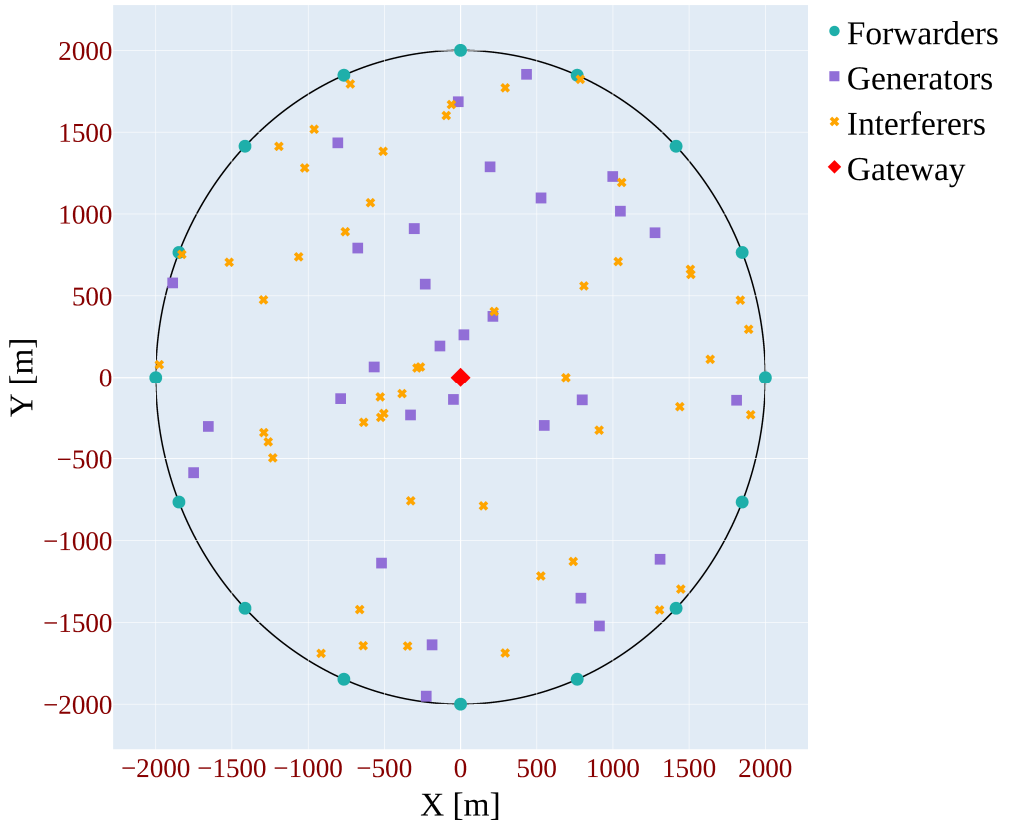


Figure 4.1. An example of a challenging deployment.

All of these nodes communicate by means of LoRa and their Spreading Factor is set individually by the ns3 LoRa module. In Figure 4.1 we can observe a possible deployment, where 16 forwarders are deployed uniformly in a circle of radius 2000 m, a single gateway is located in the center of this circle and a number of 30 generators and 50 interferers is randomly placed in the area delimited by the circle. Each forwarder is connected to 3 underwater sensor nodes (not shown in the figure for the sake of simplicity), transmitting their measurements exploiting an acoustic communication protocol and deployed uniformly in a small circle whose radius decreases as the number of forwarders increases - in this case, the radius is 200 m.

The underwater nodes transmit with a Constant Bit Rate (CBR) with a Poisson mean of 30 s, 60 s, 120 s, 240 s, 480 s or 960 s. Packet size is 30 bytes, while the transmission rate is equal to 4800 bps. Transmission frequency is 25 kHz and the bandwidth is 5 kHz. The protocol stack of these nodes is shown in Figure 4.2: they use

Table 4.1. DESERT simulations parameters.

Parameter	Value
UW nodes	3 for every forwarder
Packet Size	30 Bytes
Tx Duration	100000 s
Tx Power	135 dB re 1 μPa @1 m
Frequency	25 kHz
Bandwidth	5 kHz
Bitrate	4800 bps
CBR Period	variable in [30, 960] s range
TDMA Frame	8 s
TDMA Guard Time	0.8 s

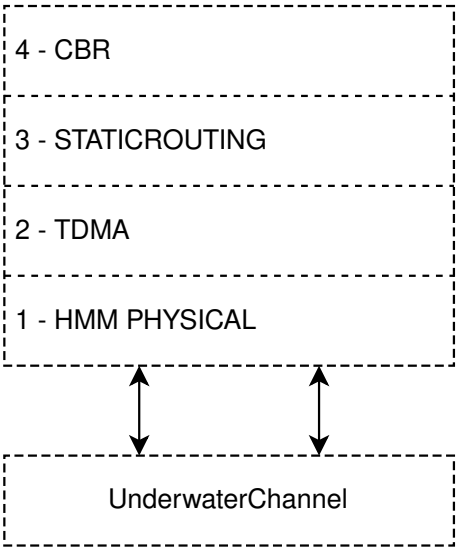


Figure 4.2. Stack of the underwater nodes.

a CBR application layer, static routing consisting in all the nodes transmitting directly to their 1-hop nearest receiver (i.e., a forwarder node), a Time Division Multiple Access (TDMA) MAC layer and, finally, a Hidden Markov Model (HMM) physical layer, as described in [59]. In particular, according to this physical model, each underwater node is linked to its destination by a channel whose behavior is statistically determined by means of its initial condition and the set of transition probabilities, which specifies the probability that, given the current state, the channel conditions in the next slot will be better, worse, or remain the same. We have set the initial state of all the links to a good (but not excellent) condition, and the transition probabilities are randomly extracted in ranges that would characterize well the evolution of a stable link.

As far as the surface nodes are concerned, we have that the forwarders relay the packets received from the underwater nodes as soon as they are correctly delivered. The

generators produce data with a CBR period equal to the one of the underwater nodes and with the same packet size, while the interferers transmit packets of 30 bytes with a CBR period of 30 s. Note that the LoRa overhead is 9 bytes for all of the packets.

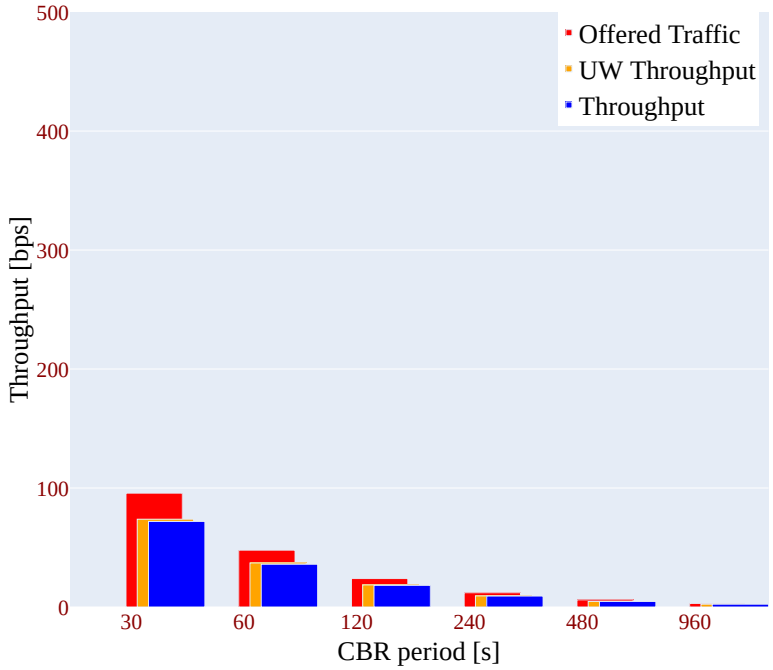
Table 4.2. ns3 simulations parameters.

Parameter	Value
Packet Size	30 Bytes
Tx Duration	100250 s
CBR Period (forwarders and generators)	variable in [60, 960] s range
CBR Period (interferers)	30 s

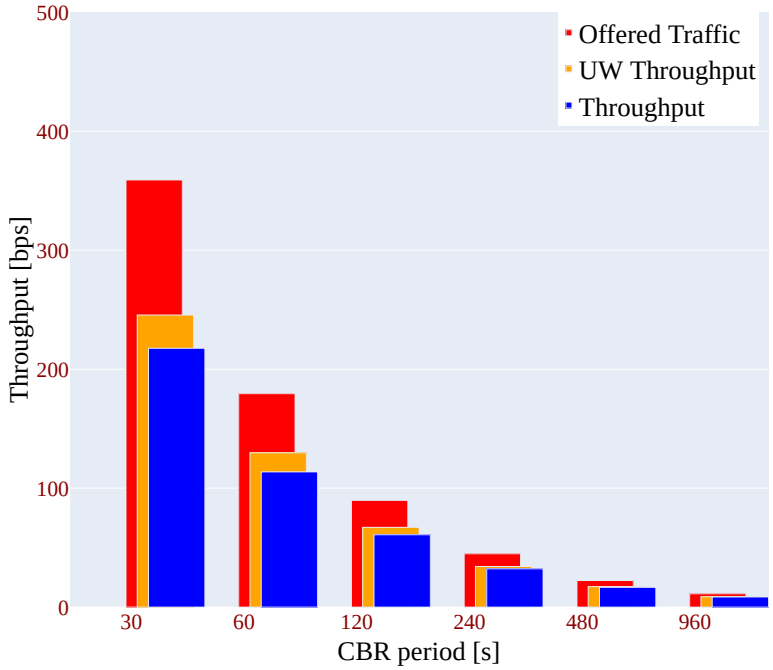
Simulations have been run extensively using SEM [60], the ns3 simulation execution manager, for different combinations of the CBR period, number of forwarders, number of generators and number of interferers. Results are averaged over 20 runs.

4.2 Results

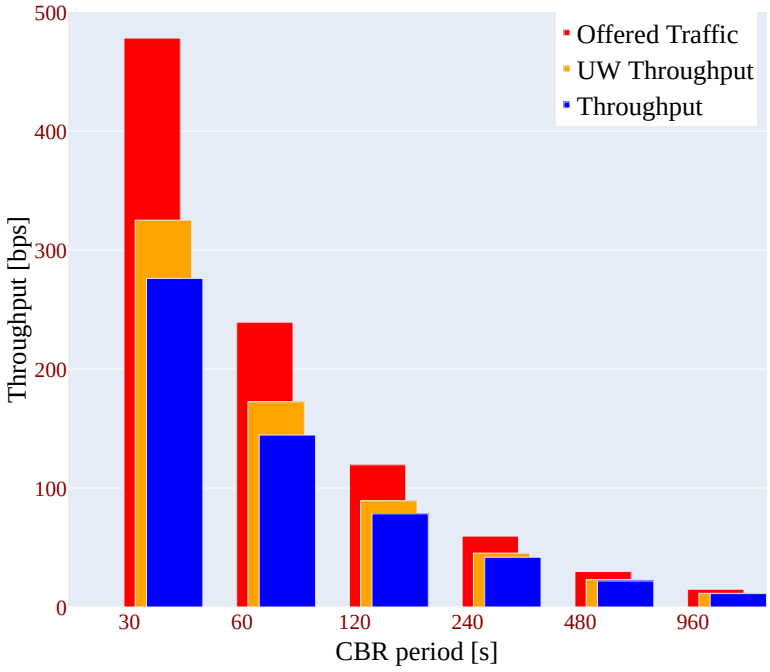
We now report some relevant results from the simulations. Figure 4.3 compares the throughput of different numbers of forwarder nodes, deployed as previously discussed, for multiple CBR periods for 4, 15 and 20 forwarder nodes. Neither generators nor interferers are present. The offered traffic (in red) is intended as the traffic generated by the 3 underwater source nodes linked to each forwarder, the underwater throughput ("UW", in yellow) is the throughput computed after the data from the underwater nodes have been received by the forwarder by means of acoustic transmission and the final throughput (in blue) is the actual throughput computed at the final destination - that is, the gateway - achieved by means of LoRa.


500
400
300
200
100
0
Throughput [bps]
30
60
120
240
480
960
CBR period [s]
Offered Traffic
UW Throughput
Throughput

(a)


500
400
300
200
100
0
Throughput [bps]
30
60
120
240
480
960
CBR period [s]
Offered Traffic
UW Throughput
Throughput

(b)



(c)

Figure 4.3. Forwarder nodes throughput analysis, no generators nor interferers, for 4 (a), 15 (b) and 20 (c) forwarders.

We can see how the major responsible for the throughput degradation is by far the underwater part, and not the above water one, and this happens in particular for low CBR periods (30 s, 60 s, 120 s) and with a higher number of forwarders. We have that, with 20 forwarders and a CBR period of 30 s, the UW throughput is just 67% of the offered traffic. Luckily, the degradation is mitigated for higher CBR periods, which is compliant with our scenario. An interesting analysis that could be carried on these data concerns the research of the best CBR period such that, in a fixed time, the probability of at least one packet reception by the gateway is maximized. Therefore, if we define p as the ratio between the throughput computed at the gateway and the offered traffic and if we set the fixed time to 960 s, for a given CBR period (CBR below) we can compute the aforementioned probability as:

$$P(succ) = 1 - (1 - p)^r,$$

where $r = \frac{CBR}{960s}$.

Table 4.3. Probabilities of successfully receiving a packet in 960 s, 4 forwarders.

CBR	30 s	60 s	120 s	240 s	480 s	960 s
P(succ)	~ 1	0.9999	0.9998	0.9971	0.9491	0.7767

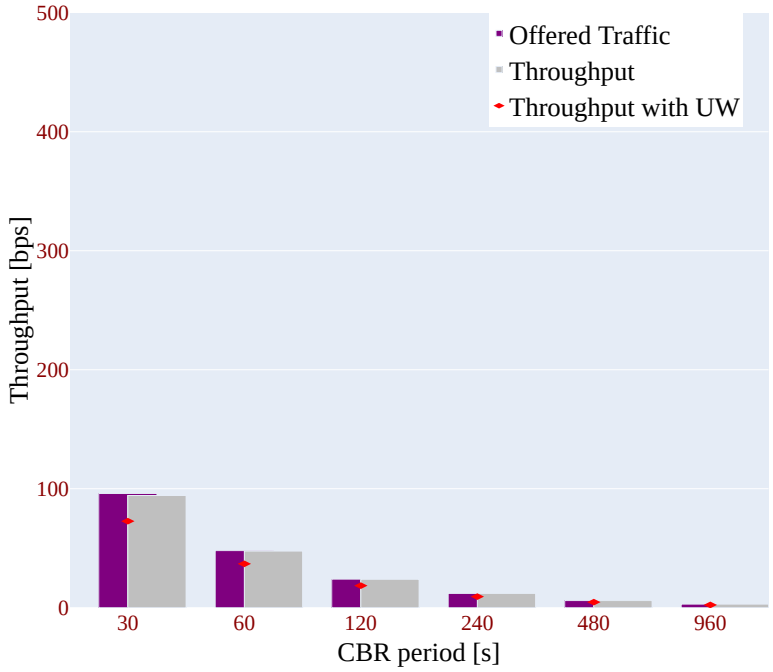
Table 4.4. Probabilities of successfully receiving a packet in 960 s, 15 forwarders.

CBR	30 s	60 s	120 s	240 s	480 s	960 s
P(succ)	0.9999	0.9999	0.9998	0.9937	0.9341	0.7613

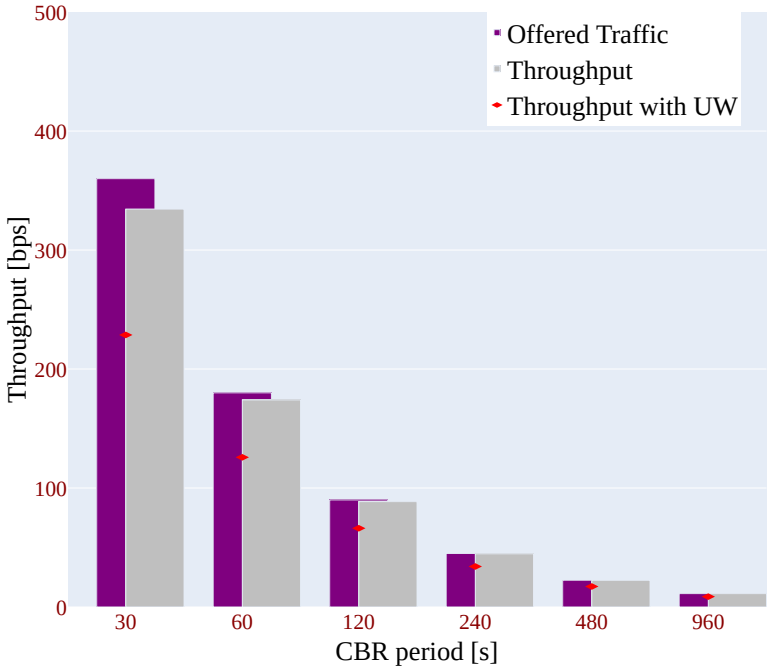
Table 4.5. Probabilities of successfully receiving a packet in 960 s, 20 forwarders.

CBR	30 s	60 s	120 s	240 s	480 s	960 s
P(succ)	0.9999	0.9999	0.9998	0.9921	0.9298	0.7546

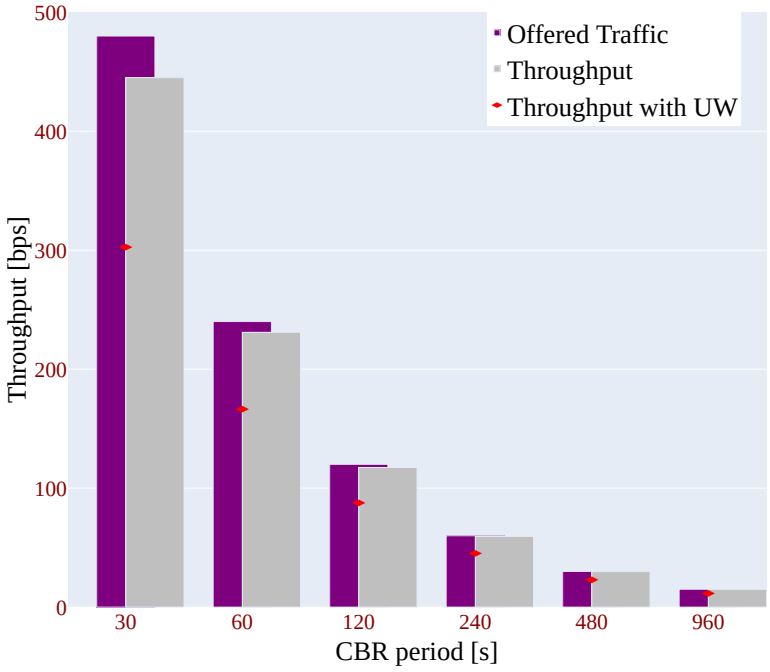
Tables 4.3, 4.4 and 4.5 report the probabilities computed with six different CBR periods for the deployments of Figure 4.3. We can see that it is not enough to send a single packet every 960 s in order to obtain an adequate packet reception probability, as it is never higher than 0.78. Conversely, a greater reception rate is achieved if we schedule four subsequent transmissions in this time lapse, that is, a transmission every 240 s. This would allow to receive at least a packet in 960 s with a probability that is never below 0.99, which is a remarkable result: a higher generation rate is then of utmost importance for the forwarders, since the throughput degradation caused by the underwater transmission is the critical aspect of the communications in the network. On the other hand, scheduling multiple transmissions will impact on the energy consumption of the devices, that is another key issue in the envisioned deployment. This means that the ideal time between transmissions will have to be chosen carefully in relation to the specific requirements.



(a)



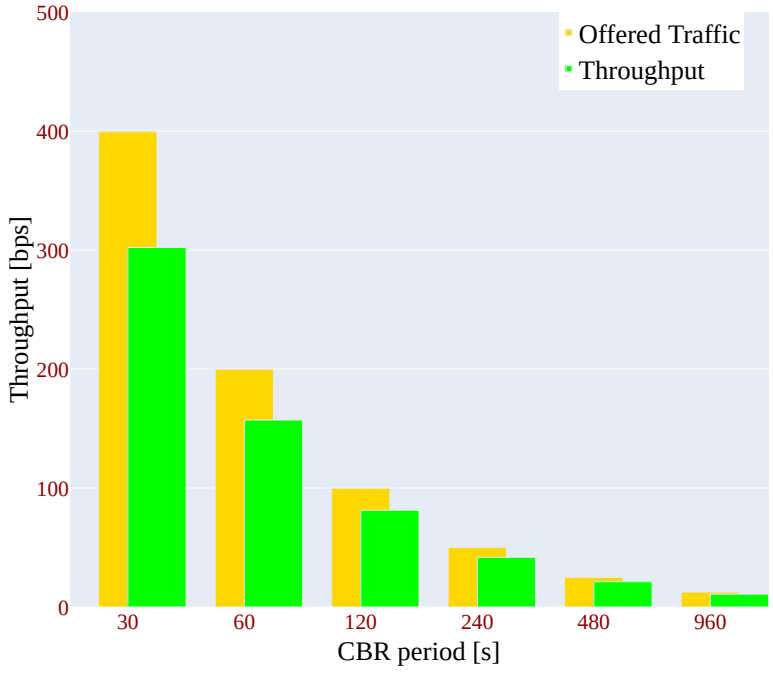
(b)



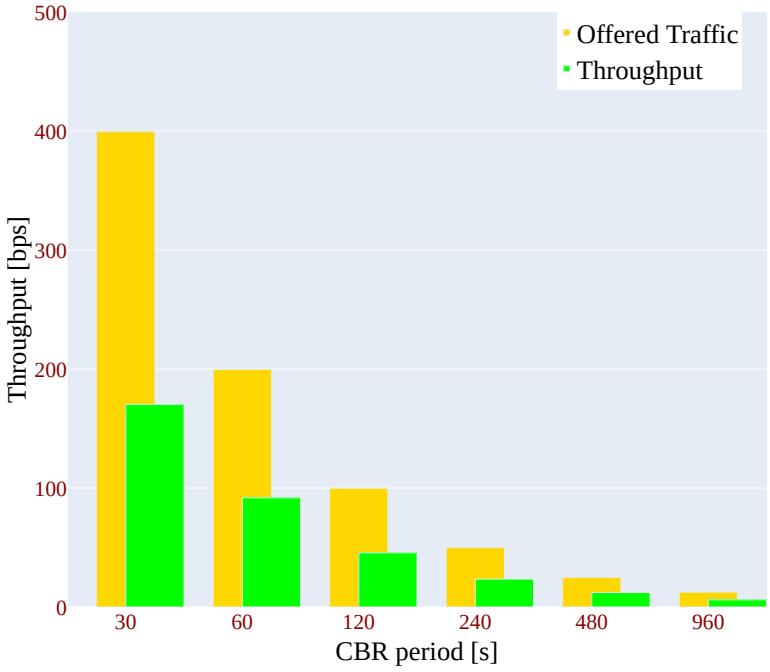
(c)

Figure 4.4. Generator nodes throughput analysis, 4 forwarders, no interferers, for 12 (a), 45 (b) and 60 (c) generators.

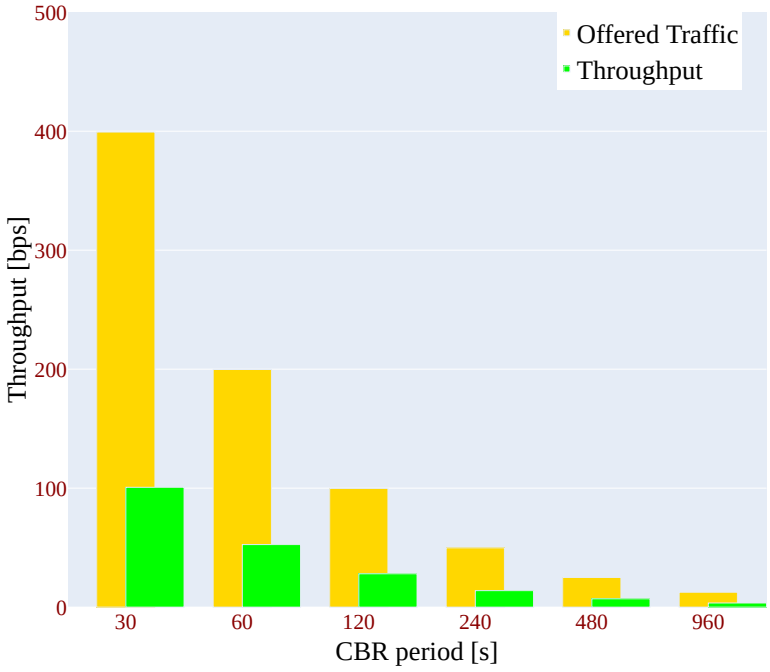
Furthermore, we may want to compare the average throughput of a forwarder with the average throughput of a generator. However, we must take into account the fact that, in the simulations, a forwarder is just a relay between the three underwater nodes connected to it and the LoRa gateway. These have the same CBR period of a generator node, and send the same amount of data, thus if we want a fair comparison between the two, a single forwarder's throughput should be compared with three different generators' throughputs. This is why in Figure 4.4 we report the offered traffic (in purple) and the throughput (in gray) against different CBR periods for 12, 45 and 60 generator nodes. Moreover, to understand the throughput difference between generators and forwarders (Figure 4.3), we have to account for the average probability of successful packet reception underwater and multiply this probability by the throughput at the gateway: the red markers in Figure 4.4 present what the generator throughputs would have been if they had to pass through the underwater channel, which are definitely closer to the actual forwarders throughput. Finally, it is useful to analyze the impact of other potential networks sharing the location of our deployment. To do this, we introduced the interferer nodes, which consist of generator nodes with a fixed CBR period of 30 s whose only purpose is to create noise. Figure 4.5 shows how the average throughput of both forwarders and generators computed at the gateway is deteriorated by the presence of the interferers with multiple CBR periods. In particular, we consider a network with 10 forwarders and 20 generators; if we focus on the lowest CBR period, we have that a number of 400 interferers reduces the actual throughput by almost 50%, and 800 interferers reduce it to 33%. Similar considerations hold for the other generation periods.



(a)



(b)



(c)

Figure 4.5. Global throughput degradation analysis, 10 forwarders, 20 generators, for 0 (a), 400 (b) and 800 (c) interferers.

We should then consider this as the worst case scenario.

In Figure 4.6 we plot the Packet Delivery Ratio for a network of 10 forwarders and 20 generators against the number of interferers for different generation periods. Higher periods guarantee a certain degree of robustness with respect to lower ones, but still the aggressiveness of the interferer nodes makes the PDR fall below acceptable levels quite soon when their number increases, and a number of only 100 interferers is already harmful for communication purposes. In such scenarios, we should take preventive measures, such as increasing the generation rates of the useful nodes or trying to decouple our network from the others as much as possible.

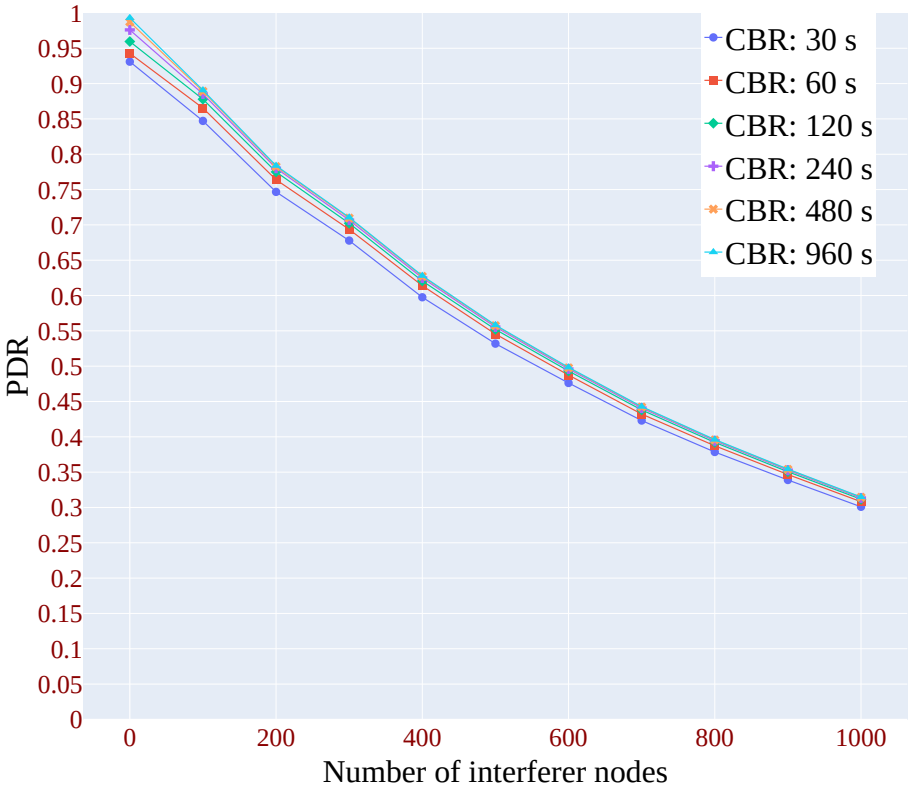


Figure 4.6. PDR for different CBR periods and number of interferer nodes, 10 forwarders, 20 generators, interferers CBR Period set to 30 s.

Chapter 5

Conclusions

In this Thesis we proposed a successful strategy to make data collection of water parameters more granular and efficient, using a hybrid network divided into an underwater section employing low-power devices using acoustic communication and an above-water section made of IoT nodes, which could forward the data to the cloud or a server for further processing.

We first presented a novel statistical model for better simulating the behavior of the underwater acoustic channel, and we proved its robustness by showing that its estimates match quite precisely actual data from real experiments (retrieved from ASUNA datasets). The model as been discussed in its two variants, respectively based on two- and three- state Markov chains, and, in addition to the realistic estimations it provides, it is neither significantly computationally demanding when implemented in a network simulator. This is because on the one hand for the simpler model the PER is calculated with a closed formula, on the other hand in the more complex model the PER is computable easily from the last PER computed during the simulation. The increase in accuracy with respect to the results is marginal between the two variants, with the three-state one being slightly better, but also the complexity does not grow meaningfully. In the future, it may be worth to analyze the tradeoff related to an increase in the number of state and the final accuracy in the estimates, and also how the developed model behaves in mobile networks, where a penalty could be associated to distance, speed or acceleration of the nodes. Indeed, these may trigger both a high Doppler effect and acoustic noise (due to eventual engines) [61].

Secondly, we proposed a solid solution to monitor the water parameters in coastal areas, which are more and more subject to the hazards originating from climate changes. In particular, the main aim of our proposal is to collect data in a coastal area with a remarkably increased granularity with respect to preexisting solutions, using energy efficient, inexpensive equipment. Our envisioned network is hybrid, since it is composed both by an underwater section using acoustic communication and from an above water part using LoRaWAN technology. We have selected a designated area representative of a challenging coastal environment, the Venice lagoon. Parameters of the network have been carefully selected according to the traffic requirements which have been chosen wisely and in compliance with the possible final application. A working, laboratory prototype of a node (eventually deployable underwater if equipped with proper waterproof gear) has been built and provided with a selection of relevant sensors to measure the most significant water parameters. Its functioning has been tested only in a labora-

tory environment and should be then put to test also in tougher conditions. Extensive simulations with meaningful deployments have been run in DESERT and ns3, in order to discover the parameters that would allow to obtain the best results according to the application and to evaluate the overall stability and performance of the network. All the results coming from DESERT simulations have been obtained relying also on the three-state HMM model presented in the first part of this work and have thus further demonstrated its high quality predictions.

Possible future work may include the enhancing of the sensor prototype we built, which has not been tested extensively and must be significantly improved before being ready to be used in the real world, the development of an ad-hoc cloud solution to store and manage the data collected by the sensor network (or the choice of an already existing solution), and actually deploy (at least part of) our envisioned network in the field, to accurately assess its performance in a challenging scenario.

Bibliography

[1] P. Casari, F. Campagnaro, E. Dubrovinskaya, R. Francescon, A. Dagan, S. Dahan, M. Zorzi, and R. Diamant, "ASUNA: A Topology Data Set for Underwater Network Emulation," *IEEE Journal of Oceanic Engineering*, vol. 46, pp. 307–318, Mar. 2021.

[2] E. Sozer, M. Stojanovic, and J. Proakis, "Underwater acoustic networks," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 72–83, 2000.

[3] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 2, pp. 125–136, 1996.

[4] M. Chitre, S. Shahabudeen, L. Freitag, and M. Stojanovic, "Recent advances in underwater acoustic communications & networking," *OCEANS 2008*, pp. 1–10, 2008.

[5] R. E. Williams and H. F. Battestin, "Coherent recombination of acoustic multipath signals propagated in the deep ocean," *The Journal of the Acoustical Society of America*, vol. 50, no. 6A, pp. 1433–1442, 1971.

[6] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 84–89, 2009.

[7] K. Scussel, J. Rice, and S. Merriam, "A new mfsk acoustic modem for operation in adverse underwater channels," in *Oceans '97. MTS/IEEE Conference Proceedings*, vol. 1, pp. 247–254 vol.1, 1997.

[8] D. Falconer, F. Adachi, and B. Gudmundson, "Time division multiple access methods for wireless personal communications," *IEEE Communications Magazine*, vol. 33, no. 1, pp. 50–57, 1995.

[9] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*. USA: Wiley-Interscience, 1995.

[10] J.-g. Huang, H. Wang, C.-b. He, Q.-f. Zhang, and L.-y. Jing, "Underwater acoustic communication and the general performance evaluation criteria," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 8, pp. 951–971, 2018.

[11] "EvoLogics Underwater Acoustic Modems." `https://evologics.de/acoustic-modems`. Last time accessed: Aug. 2022.

[12] "Develogic Subsea Systems." `http://www.develogic.de/`. Last time accessed: Aug. 2022.

[13] "Subnero M25M Series Modems." `https://subnero.com/products/modem.html`. Last time accessed: Apr. 2022.

[14] B.-C. Renner, J. Heitmann, and F. Steinmetz, "AHOI: Inexpensive, Low-power Communication and Localization for Underwater Sensor Networks and µAUVs," *ACM Transactions on Sensor Networks*, vol. 16, Jan. 2020.

[15] B. Benson, Y. Li, B. Faunce, K. Domond, D. Kimball, C. Schurgers, and R. Kastner, "Design of a Low-Cost Underwater Acoustic Modem," *IEEE Embedded Systems Letters*, vol. 2, pp. 58–61, May 2010.

[16] B. Sherlock, J. A. Neasham, and C. C. Tsimenidis, "Implementation of a spread-spectrum acoustic modem on an android mobile device," in *Proc. MTS/IEEE OCEANS, Aberdeen, UK*, 2017.

[17] E. Coccolo, F. Campagnaro, D. Tronchin, A. Montanari, R. Francescon, L. Vangelista, and M. Zorzi, "Underwater acoustic modem for a morphing distributed autonomous underwater vehicle (MODA)," in *Proc. MTS/IEEE OCEANS, Chennai, India*, Feb. 2022.

[18] "Popoto modem." `http://popotomodem.com/`. Last time accessed: Aug. 2022.

[19] "Water Linked Modem M64." `https://waterlinked.com/product/modem-m64/`. Last time accessed: Aug. 2022.

[20] M. Stojanovic, "On the relationship between capacity and distance in an underwater acoustic communication channel," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 11, pp. 34–43, Oct. 2007.

[21] M. Chitre, I. Topor, R. Bhatnagar, and V. Pallayil, "Variability in link performance of an underwater acoustic network," in *MTS/IEEE OCEANS*, (Bergen, Norway), jun 2013.

[22] M. Chitre, T.-B. Koay, G. Deane, and G. Chua, "Variability in shallow water communication performance near a busy shipping lane," in *Fifth Underwater Communications and Networking Conference (UComms)*, (Virtual), 2021.

[23] M. Porter *et al.*, "Bellhop code." `http://oalib.hlsresearch.com/Rays/index.html`. Last time accessed: Apr. 2022.

[24] B. Tomasi, P. Casari, M. Zorzi, G. Zappa, and K. McCoy, "Experimental study of the acoustic channel properties during subnet 2009," tech. rep., University of Padova, 2010.

[25] B. Tomasi, J. Preisig, and M. Zorzi, "On the predictability of underwater acoustic communications performance: The KAM11 data set as a case study," in *Proc. International Conference on Underwater Networks & Systems (WUWNet)*, (Seattle, WA, US), December 2011.

[26] P. A. van Walree and M. E. G. D. Colin, "In situ performance prediction of a coherent acoustic modem in a reverberant environment," *IEEE Journal of Oceanic Engineering*, vol. 47, pp. 236–254, Jan. 2022.

[27] R. Otnes, P. A. van Walree, H. Buen, and H. Song, "Underwater acoustic network simulation with lookup tables from physical-layer replay," *IEEE Journal of Oceanic Engineering*, vol. 40, pp. 822–840, Oct. 2015.

[28] V. Kalaiarasu, H. Vishnu, A. Mahmood, and M. Chitre, "Predicting underwater acoustic network variability using machine learning techniques," in *IEEE/MTS OCEANS*, (Anchorage, US), Sept 2017.

[29] C. Tapparello, P. Casari, G. Toso, I. Calabrese, R. Otnes, P. van Walree, M. Goetz, I. Nissen, and M. Zorzi, "Performance evaluation of forwarding protocols for the racun network," WUWNet '13, (New York, NY, USA), Association for Computing Machinery, 2013.

[30] A. Signori, F. Campagnaro, F. Steinmetz, B.-C. Renner, and M. Zorzi, "Data gathering from a multimodal dense underwater acoustic sensor network deployed in shallow fresh water scenarios," *Journal of Sensor and Actuator Networks*, vol. 8, no. 4, 2019.

[31] F. Campagnaro, F. Favaro, P. Casari, and M. Zorzi, "On the feasibility of fully wireless remote control for underwater vehicles," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, nov 2014.

[32] F. Campagnaro, R. Francescon, F. Guerra, F. Favaro, P. Casari, R. Diamant, and M. Zorzi, "The DESERT underwater framework v2: Improved capabilities and extension tools," in *Proc. Ucomms*, (Lerici, Italy), Sept. 2016.

[33] F. Campagnaro, A. Signori, R. Otnes, M. Goetz, D. Sotnik, A. Komulainen, I. Nissen, F. Favaro, F. Guerra, and M. Zorzi, "A simulation framework for smart adaptive long- and short-range acoustic networks," in *Proc. MTS/IEEE Oceans*, (Virtual Global Oceans San Diego - Porto), Sept. 2021.

[34] F. Pignieri, F. De Rango, F. Veltri, and S. Marano, "Markovian approach to model underwater acoustic channel: Techniques comparison," in *IEEE Military Communications Conference (MILCOM 2008)*, 2008.

[35] B. Tomasi, P. Casari, L. Finesso, G. Zappa, K. McCoy, and M. Zorzi, "On modeling JANUS packet errors over a shallow water acoustic channel using Markov and hidden Markov models," in *IEEE Military Communications Conference (MILCOM 2010)*, pp. 2406–2411, 2010.

[36] W. Turin and R. van Nobelen, "Hidden Markov modeling of flat fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1809–1817, Dec. 1998.

[37] "A shared underwater network emulation dataset." `https://sites.google.com/marsci.haifa.ac.il/asuna/`. Last time accessed: Apr. 2022.

[38] R. Diamant, G. N. Shirazi, and L. Lampe, "Robust spatial reuse scheduling in underwater acoustic communication networks," *IEEE Journal of Oceanic Engineering*, vol. 39, pp. 32–46, Jan. 2014.

[39] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*. Academic Press, third ed., 1999.

[40] "Design, simulate, emulate and realize test-beds for underwater network protocols." `http://desert-underwater.dei.unipd.it/`. Last time accessed: Apr. 2022.

[41] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "Overview of Cellular LPWAN Technologies for IoT Deployment: Sigfox, LoRaWAN, and NB-IoT," in *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pp. 197–202, 2018.

[42] "LoRa Alliance - Homepage." `https://lora-alliance.org/`. Last time accessed: Apr. 2022.

[43] M. Milenkovic, *Internet of Things: Concepts and System Design*, ch. Communications, pp. 92–94. Springer, Cham, May 2021.

[44] L. Vangelista, "Frequency shift chirp modulation: The LoRa modulation," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1818–1821, 2017.

[45] J. M. Marais, R. Malekian, and A. M. Abu-Mahfouz, "LoRa and LoRaWAN testbeds: a review," in *IEEE AFRICON*, pp. 1496–1501, 2017.

[46] M. Bor and U. Roedig, "LoRa Transmission Parameter Selection," in *13th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 27–34, 2017.

[47] I. Cappelli, A. Fort, M. Mugnaini, S. Parrino, and A. Pozzebon, "Underwater to above water lorawan networking: Theoretical analysis and field tests," *Measurement*, vol. 196, p. 111140, June 2022.

[48] "Arduino MKR WAN 1310." Last time accessed: Aug. 2022.

[49] "Gravity: Analog pH Sensor." Last time accessed: Aug. 2022.

[50] "Gravity: Analog Electrical Conductivity Sensor." Last time accessed: Aug. 2022.

[51] "Gravity: Analog Turbidity Sensor." Last time accessed: Aug. 2022.

[52] "GPS Module with Enclosure." Last time accessed: Aug. 2022.

[53] "Gravity: Waterproof DS18B20 Sensor." Last time accessed: Aug. 2022.

[54] "Gravity: Liquid Level Sensor." Last time accessed: Aug. 2022.

[55] D. Magrin, A. Signori, D. Tronchin, F. Campagnaro, and M. Zorzi, "Collaboration of LoRaWAN and Underwater Acoustic Communications in Sensor Data Collection Applications," in *Proc. MTS/IEEE Oceans*, (Global Oceans (Virtual)), Oct. 2020.

[56] "ns3 network simulator." `https://www.nsnam.org/`. Last time accessed: Aug. 2022.

[57] "DESERT Underwater web site." `http://nautilus.dei.unipd.it/desert-underwater`.

[58] "LoRaWAN ns-3 module." https://github.com/signetlabdei/lorawan. Last time accessed: Aug. 2022.

[59] F. Campagnaro, N. Toffolo, and M. Zorzi, "Modeling acoustic channel variability in underwater network simulators from real field experiment data," *Electronics*, vol. 11, no. 14, 2022.

[60] D. Magrin, D. Zhou, and M. Zorzi, "A simulation execution manager for ns-3: Encouraging reproducibility and simplifying statistical analysis of ns-3 simulations," in *Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, MSWIM '19, (New York, NY, USA), p. 121–125, Association for Computing Machinery, 2019.

[61] E. Coccolo, F. Campagnaro, A. Signori, F. Favaro, and M. Zorzi, "Implementation of AUV and ship noise for link quality evaluation in the desert underwater framework," in *Proc. ACM WUWNet*, (Shenzhen, China), Dec. 2018.

[62] F. Campagnaro, A. Signori, and M. Zorzi, "Wireless remote control for underwater vehicles," *Journal of Marine Science and Engineering*, vol. 8, no. 10, 2020.

[63] P. C. et al., "ASUNA: A Topology Data Set for Underwater Network Emulation," *IEEE Journal of Oceanic Engineering*, vol. 46, pp. 307–318, Mar. 2021.

[64] J. G. Ruiz, B. Soret, M. C. Aguayo-Torres, and J. T. Entrambasaguas, "On finite state Markov chains for Rayleigh channel modeling," in *Proc. 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace Electronic Systems Technology*, (Aalborg, Denmark), May 2009.

[65] A. Mahmood, H. Vishnu, and M. Chitre, "Model-based signal detection in snapping shrimp noise," in *IEEE Third Underwater Communications and Networking Conference (UComms)*, (Lerici, Italy), 2016.

[66] M. Chitre and K. Pelekanakis, "Channel variability measurements in an underwater acoustic network," in *Underwater Communications and Networking (UComms)*, (Sestri Levante, Italy), 2014.

[67] A. Caiti, K. Grythe, J. M. Hovem, S. M. Jesus, A. Lie, A. Munafò, T. A. Reinen, A. Silva, and F. Zabel, "Linking acoustic communications and network performance: Integration and experimentation of an underwater acoustic network," *IEEE Journal of Oceanic Engineering*, vol. 38, pp. 758–771, Sept. 2013.