ASUNA: A Topology Dataset for Underwater Network Emulation

Paolo Casari, Senior Member, IEEE, Filippo Campagnaro, Member, IEEE, Elizaveta Dubrovinskaya, Student Member, IEEE, Roberto Francescon, Amir Dagan, Shlomo Dahan
Michele Zorzi, Fellow, IEEE, Roee Diamant, Senior Member, IEEE

Abstract

We report the details of ASUNA, a freely shared dataset for underwater network emulation (ASUNA). ASUNA tackles the time-consuming and costly logistics of multiple underwater networking sea trials by providing a benchmark database of time-varying network topologies recorded across multiple sea experiments, thus facilitating experiment replay and network emulation. The ASUNA database currently includes 20 diverse, time-varying topology structures, multimodal communication technologies, and different link quality measurements. With the aim of becoming a standard benchmark, ASUNA is open to extensions as new data becomes available from the underwater communications community. We provide the details of ASUNA structure, the list of recorded topologies, as well as examples of how to use the database as part of an emulation system to test the performance of two scheduling protocols. We freely share the database and the emulation code both through a web server and via the Code Ocean repository.

Index Terms

Underwater acoustic communication networks; emulation system; underwater acoustic network benchmark; sea experiments; experiment replay; network performance validation.

P. Casari (email: paolo.casari@unitn.it) is with DISI, University of Trento, 38123 Povo (TN), Italy. F. Campagnaro (email: campagn1@dei.unipd.it) and M. Zorzi (email: zorzi@dei.unipd.it) are with the Department of Information Engineering, University of Padova, 35131 Padova, Italy. E. Dubrovinskaya (email: lisa.d@imdea.org) is with the IMDEA Networks Institute and University Carlos III of Madrid, 28918 Madrid, Spain. R. Francescon (email: roberto.francescon@wirelessandmore.it) is with Wireless and More, 35131 Padova, Italy. A. Dagan (email: adagan3@univ.haifa.ac.il), S. Dahan (email: sdahan3@univ.haifa.ac.il) and R. Diamant (email: roee.d@univ.haifa.ac.il) are with the Department of Marine Technologies, University of Haifa, 3498838 Haifa, Israel.
I. INTRODUCTION AND MOTIVATION FOR THIS WORK

Underwater communication devices have been steadily improving over time in terms of both reliability and bit rate [1], and can be arranged into underwater acoustic communication networks (UWANs) to support a broad variety of applications [2]. A multitude of protocols have been designed for UWANs to date, providing different functionalities at different layers of the ISO/OSI protocol stack [3]–[5].

Sea trials are considered a good option to test the performance of UWAN protocols. However, organizing and performing a sea experiment is usually time-consuming, effort-intensive, and implies high costs in terms of materials, rental of ship time, purchase, transport and deployment of underwater transceivers, etc. Moreover, sea experiment results are related to a local set of environment and channel conditions, which makes it difficult to extrapolate the results to different environments. Finally, a single experiment does not allow a fully fair comparison between algorithms.

As a result, simulations are often preferred when evaluating the performance of newly designed protocols against competing approaches in the literature. Simulations make it possible to approximately evaluate communication protocols and schemes by abstracting from specific hardware issues. However, on the one hand there are no statistical models of the underwater acoustic channel that are broadly agreed upon, so that a realistic simulation often has to rely on complex numerical propagation modeling (e.g., [6]); on the other hand a full-fledged evaluation must take into account the many practical issues that occur in actual underwater scenarios. This includes the time-dependency of the acoustic channel, the conditions of actual underwater environments, and possibly the behavior of hardware devices.

Hardware-in-the-loop systems are one of the means to improve the agreement between simulated protocol performance and actual performance at sea, at least in terms of the peculiarities of underwater transceivers. Examples of frameworks offering hardware-in-the-loop capabilities include DESERT Underwater [7], SUNSET [8], UNetStack [9], Aqua-Net-Mate [10], NETSIM [11], as well as the software-defined cognitive communications architecture presented in [12]. These capabilities are made possible by interchanging the procedures that simulate
underwater propagation and compute link budgets with software drivers for specific underwater modems. However, even the hardware-in-the-loop concept can reproduce actual underwater propagation and the variation thereof over time only to a limited extent.

The above discussion leads to the conclusion that, in the absence of both a fully detailed simulation model of an underwater acoustic communication system and of the resources to organize a sea experiment, a reliable performance evaluation method should preferably involve recordings from a real sea environment.

For the design and test of point-to-point underwater communication systems in realistic conditions, the community often resorts to publicly shared communication datasets in order to reproduce the broadest possible span of underwater channels (e.g., long or short delay spread, heavier or milder Doppler spread, single or multiple receivers, etc.). Examples include the measurements presented in [13], the SPACE08 and the KAM11 datasets, employed among others in [14]–[17]. More recently, the release of the Watermark benchmark [18] makes it possible to reproduce the distortion of acoustic waveforms transmitted through underwater channels that are either measured or stochastically replayed.

In this paper, we propose a similar solution for the testing of underwater network protocols, named ASUNA, for “A shared underwater network emulation dataset.” ASUNA is a collection of measurements from multiple sea experiments, and aims to be the first freely shared database that enables the replay of underwater acoustic networking trials, often referred to as emulation. To the best of our knowledge, this is the first attempt to assemble a dataset for the direct evaluation of the performance of network protocols. ASUNA provides a collection of time series of link quality indicators, collected over time during several sea experiments at different locations around Europe, Israel and West Africa. These experiments are representative of a broad set of conditions: different numbers of nodes, different deployments resulting in multiple network topologies, different transceivers, and multimodal setups (where communications are realized through a set of orthogonal technologies). Once this data has been loaded into a network emulator, the link quality time series can be used to reproduce the same realistic performance that could be experienced at the same location and time each experiment had been carried out. As a result, the user can evaluate networking solutions with a degree of accuracy that stands in between a
simulation and a full-fledged sea experiment, in a fully reproducible setting, without having to actually go to sea.

In total, ASUNA includes 22 network topologies from 7 different sea experiments, for a total of more than 10 hours of underwater data packet transmissions. We make the data available in an Octave/Matlab format, so that it can be easily manipulated, converted to other formats, as well as integrated into existing Octave/Matlab code. For each dataset, we document the experiment it is extracted from, so that the user knows the experiment’s location and time; the location of the nodes; the conditions of the water body at that time; and the types of link quality measurements available for that experiment. The metrics provided by the dataset so far include received signal strength, bit error ratios, and “1/01” indicators conveying whether a given packet would be received correctly or not if transmitted at a given time. Different metrics may be embedded in the future as additional datasets are added to the collection.

We hope that the community will find ASUNA useful and will contribute additional datasets to the collection. Along with the database, we provide a network emulation code as an example of how to use ASUNA. The emulator runs a simple time-division-multiple-access (TDMA) protocol over the recorded topologies. Yet, by no means is the usage of ASUNA confined to such a solution.

While there is some novelty in our approach and it has been recently endorsed that reproducible and interactive research results bear significant value for the underwater community [19], the focus of this technical communication is on the tool per se, rather than on novel results obtained through it. The remainder of this paper is organized as follows: Section II provides an account of related work; Section III describes the ASUNA dataset; Section IV discusses the emulator provided with the dataset and some results obtained with it; Section V concludes the paper.

II. RELATED METHODS

In terrestrial radio networks, it is customary to evaluate the performance of wireless networking protocols by means of simulations, supported by different types of channel models [20]–[22]. Initial studies on channel modeling for underwater networks followed the same approach. For example, [23] modeled packet errors from the SubNet09 campaign using Markov and hidden
Markov models. Typical statistical distributions of large-scale underwater channel gain [24]–[26] have been observed to be valid across a number of channel measurements.

Besides simulation, network performance can be evaluated through emulation or trace-based simulation. Emulation refers to the use of realistic networking hardware, or to the execution of actual applications on top of hardware components that reproduce the behavior of wireless networking equipment. For example, this implies running complex channel models in real-time in some dedicated hardware. Trace-based simulation [27]–[29], also described as channel replay-based, relies on the recording of the time series (or “traces”) of link quality metrics [30]. This makes it possible to exactly reproduce the same wireless channel conditions repeatedly, and to test different protocols in fully comparable scenarios. For tests that do not require to learn the channel evolution over time, the evaluation can be extended by suitably scrambling the measurements so that channel properties remain statistically coherent [31], [32].

In the underwater community, several works have tackled the reliable and validated reproduction of the communications performance measured during experimental campaigns. These studies mainly focused on the physical layer. For example, [13] proposed to collect underwater channel recordings in order to reproduce the impact of the acoustic channel on underwater modulation schemes. The collected dataset includes channel estimates from several sea experiments. More recently, Watermark [18] has been released as a benchmark for underwater modulation schemes. Watermark is based on the validated MIME tool, which enables both direct and stochastic underwater channel replay [16], [33]. In some cases, channel estimates can be directly obtained through deployed infrastructure that is shared with the community at large, typically for limited periods of time and under some form of collaboration agreement. This includes the NATO CMRE LOON [34], the equipment of Ocean Networks Canada [35], the SUNRISE testbed federation [36], as well as permanently online infrastructure such as the THEM O observatory [37].

Besides direct and stochastic channel replay, other methods have been considered to enable model-based channel reproduction. For example, in [38] the authors propose to evaluate the reliability of underwater communications through the multipath structure of previously measured underwater channels, which can be evaluated using numerical models rather than sea experiments.
Realistic channel simulations obtained through the Bellhop ray tracing software [6] have been incorporated in the World Ocean Simulation System (WOSS) [7], a framework that automatically retrieves the environmental information required by Bellhop in order to compute attenuation figures and channel impulse responses. A similar integration of models based on parabolic equations in network simulations is discussed in [39]. Like many other channel simulators, both Bellhop and a parabolic equation solver present the issue that their output is deterministic for fixed boundary conditions. This was addressed, e.g., in [25], which provides time-varying channel realizations as would result from the movement of the transmitter and receiver around their nominal locations. When numerical models or stochastic replay are not sufficient, hardware-in-the-loop systems offer one additional degree of realism by allowing network protocol code (typically written for simulations) to run on actual underwater transceivers. Examples of this approach include DESERT Underwater [40], SUNSET [8], UNetStack [9], Aqua-Net-Mate [10] and NETSIM [11].

Replicating a real underwater communication experiment in network simulations is often challenging and necessarily leads to approximations. Typical approaches include: placing nodes at random in an area and using acoustic models to predict the success of packet transmissions [41]–[45]; simulating node motion, especially in the presence of autonomous underwater vehicles (AUVs) or other types of mobile nodes [46]–[49]; letting nodes drift, e.g., by using water current models [50]–[52]; and injecting the acoustic noise generated by ships and AUVs navigating near the network deployment [53].

While the above methods approximate realistic scenarios to some degree, only in sea experiments can all the details of actual underwater communications be taken into account. Experiments with a large number of nodes were demonstrated by large organizations or collaborations. Relevant examples include the joint TNO/FFI tests on the NILUS node [1] (7 nodes); the collaborative experiments promoted by the NATO STO CMRE, such as CommsNet13 [54] (up to 9 nodes); the MISSION 2013 campaign [9] (10 nodes); the final sea trial of the RACUN project [55] (15 nodes); as well as the Jaffe lab sub-mesoscale ocean sampling experiment, featuring 5 static pingers and 13 passive drifters [56].

Besides their complex logistics and cost, underwater networking experiments still capture only
the local conditions of the underwater channel at a single location and time: such conditions are not easily extrapolated to different times and scenarios. Through ASUNA, we provide a number of experiment traces, each conveying recorded time series of link quality metrics for all links of several networking experiments. Our objective is to grow ASUNA into a rich and significant benchmark tool through contributions from the community: however, the experiments initially provided already represent a number of different conditions. ASUNA enables “network replay” in a form similar to [45] and [57], which employed previously recorded time series of the signal-to-noise ratio (SNR) or of successful packet receptions in order to test the performance of underwater scheduling and routing protocols, respectively. There are also similarities with the physical layer replay capabilities of the architecture in [12]. However, while the focus of the above approaches is on the performance evaluation of specific protocols or communication architectures, our objective here is to provide a growing collection of network communication traces. In doing so, we aim at making available a tool that remains positioned between pure simulation and pure experimentation, and that joins the repeatability of trace-based simulation with the rich representation of environments and contexts provided by a sea trial database.

III. DESCRIPTION OF THE DATASET

A. Overview and link reliability measurements

The ASUNA database is available for download at https://sites.google.com/marsci.haifa.ac.il/asuna/. ASUNA’s databases are basically constructed as time series of link reliability metrics opportunistically collected from UWAN experiments at sea. In each experiment, one or more network topologies were tested.

Link reliability signifies the integrity of the communications between adjacent nodes. It enables hard decisions about the existence of a link (e.g., by setting a threshold on the metric) or, alternatively, soft decisions (e.g., tying the bit error ratio to the probability of packet error). The link reliability is typically a time-varying property. This is especially true for underwater acoustic communications, where the channel impulse response and the ambient noise tend to change rapidly. While emulating physical layer reliability requires a fine time resolution (at least matching the symbol rate), the resolution constraint can be relaxed for the evaluation of...
underwater networks, where the most important aspect is typically the average (rather than instantaneous) link performance throughout the duration of a packet. In our experiments, we either (i) collected data on a per-packet rather than per-symbol basis, or (ii) relied on link metrics returned by the modems. The latter are derived either from a packet’s preamble, or by observing whether packets are successfully received. We remark that such phenomena as flickering (a condition by which a link appears and disappears at a fast rate in the network’s topology) are still present in our topologies at packet transmission time scales, and still enable the evaluation of adaptive protocols that specifically react to such phenomena.

We employ both physical layer and network layer metrics to characterize the link’s reliability. Depending on the experiment, we provide: bit error ratio (BER) values computed as the ratio of correctly received bits over the total number of bits in a received packet; received signal strength indicator (RSSI) values related to voltage readings at the receiver upon packet reception, or 1/0 flags that convey whether a link is available or not at a given time epoch. While these metrics can serve for experiment replay, future contributors of ASUNA are welcome to also record quality indices that are more specific to the setup of their experiment including, e.g., the packet error ratio (PER) or the link throughput.\footnote{Providing fine-grained information about the packet transmission and reception times as well as about multipath propagation would be very convenient and would convey additional details about acoustic propagation at the time of the experiment. Unfortunately such accurate information is not available for the current version of ASUNA. We still plan to include it for any future datasets we will integrate, provided that these datasets can demonstrate sufficiently accurate time reckoning and multipath measurements.} We remark that the datasets of ASUNA are opportunistically extracted from experiments originally designed to test specific communication protocol and schemes. As a consequence, the availability of link metrics depends on the logs collected from the experiment, and may vary across different sea trials. Moreover, the experiments were not necessarily focused on collision modeling. We leave the collection of collision-specific datasets to future extensions of ASUNA. In the meantime, it is still possible for ASUNA users to model collisions approximately by assuming that concurrently transmitted packets are always lost or that they are recovered with a given probability (e.g., as in the case of frequency-hopping schemes, where the recovery probability can be determined based on the hopping pattern).
B. Topology matrix information (TMI) structure

For each experiment, our database includes a description of the experiment’s setup, an Octave/ Matlab .mat file grouping link quality time series into a matrix for (called topology matrix information in the following, or TMI for short), and a reference to the publication(s) that convey the context of each experiment. The basic building block of each TMI is an instantaneous snapshot of the quality of all links. This can be seen as an $N \times N$ matrix, whose entry $(i, j)$ reflects the link quality between nodes $i$ and $j$ as measured from the experiment, and where $N$ is the number of nodes in the network.\(^2\)

The time variation of the TMI is captured by adding a time dimension to each topology matrix. The sampling time depends on the context of the experiment and on the configuration of the communication protocols. For example, for an experiment based on a time-division-multiple-access (TDMA) schedule, the topology information is obtained for each time frame. Conversely, in experiments focusing on the physical layer, we update the topology information once for every transmitted packet. Still, the sampling time is sufficiently frequent to enable the interpretation of the topology information as a continuous process.\(^3\)

During replay processes it is then possible to, e.g., check the quality of a link at the time of each transmission in order to determine which data packets are correctly received, and how many useful application bits they carried, so as to compute the goodput (defined as the rate of reception of useful information bits over time); alternatively, it is possible to provide the communicating nodes with a noisy version of the TMI to emulate some form of topology instability.

In some experiments, the time variation of the TMI was achieved through the dynamic relocation of one or more nodes in the same area. In this case, we provide link data for each topology separately in the same .mat file, with the understanding that the duration of the experiments may be different for each topology. The ASUNA dataset is generally obtained from static deployments. Some of these deployments include drifting nodes (e.g., the REP and

\(^2\) Note that the TMI may be asymmetric. This is the case when the SNR is location- or depth-dependent, and in scenarios involving near-far conditions, where interference blocks one end of the communication link.

\(^3\) We remark that the link sampling time is a feature of the data provided in the dataset, and depends on the structure of the experiment from which we derived the link quality measurements. For this reason, it is not possible to configure this parameter.
Haifa Harbor datasets), which leads to limited mobility. To improve the possibilities for the user to simulate some form of mobility, as well as to emulate underwater networking scenarios where abrupt link quality changes occur, we also provide a global time series that covers a whole experiment across all tested topologies. This is obtained by concatenating the link measurements of each TMI. In fact, between subsequent topologies in a given dataset, some links typically disappear, some new links appear, and those that persist experience significant quality changes. Additionally, we remark that mobility can be approximately emulated by rotating the position of the nodes throughout the locations indicated in each dataset.

In case several communication technologies are involved in an experiment, as is the case for multimodal network setups, a further dimension is added to the TMI. In this case, the time-varying TMI is provided per-technology. This makes it possible to have simultaneous or very close samples of the link quality perceived by different communication technologies. We remark that different technologies often have different transmission capabilities. For example, this is the case for the SC2R high-frequency (80-120 kHz) EvoLogics modem, which has a much higher nominal bit rate than the EvoLogics modem working in the 7-17 kHz band. Such different bit rates cause asynchronous channel sampling at unequal rates. Details about the sampling time are provided in the companion document of each dataset in ASUNA.

C. Analysis of TMIs

The resulting TMIs that create the heart of the database can be analyzed in different ways. For example, by setting a threshold over the link measurements, one may create an emulation system that avoids a physical layer and only uses realistic binary topologies to form time-varying communication links. This may become relevant when testing scheduling and routing protocols. The user can also treat the soft link quality measures to form a time-varying statistical model that generates links based on measured link reliability information. While some of our reported TMIs are small in terms of the number of nodes or short in terms of the testing time, the network size can be virtually increased by duplicating parts of it, and the time duration can be extended cyclically. In this manner, larger networks and longer deployment scenarios can be tested more
reliably than using models, although such an extension to the network cannot be considered as a replay.

An illustration of the emulation process is given in Fig. 1. The process begins with link quality data collection during a single sea experiment to form a matrix of time-varying TMIs. The experiment may include several arrangements of the network nodes into different topologies. The link quality data is used for network replay, where the time-varying link quality information determines the success of each data transmission. Similar to channel realizations used for channel replay [33], [58], the result is a reliable representation of the network performance in the sea conditions that occurred during the recorded network topology.
D. Structure and variety of the shared datasets

In this section, we describe the structure of the network TMIs currently available in ASUNA. When downloading ASUNA from the web site, TMIs come organized in separate folders. For a given TMI, call $N$ the number of nodes, $P$ the number of (physical layer) transmission technologies available to each node, and $T$ the total number of link quality sampling epochs. Normally, these epochs are separated by an interval $\Delta t = 1$ second, unless otherwise stated in the experiment description. The .mat files of the TMIs have the same structure, and contain the following data:

- a $\text{TopMat}$ matrix of size $T \times N \times N \times P$, where each entry $\text{TopMat}(t, i, j, p)$ (using Octave/Matlab notation) conveys the link quality for the link between nodes $i$ and $j$ through physical layer technology $p$ at time $t$;
- a $\text{LocMat}$ matrix of size $T \times N \times 3$, where the three entries $\text{LocMat}(t, i, 1:3)$ represent the two UTM coordinates and the depth of node $i$, respectively;
- a $\text{TechMat}$ matrix of size $T \times N \times P$, where each of the $k = 1, \ldots, P$ entries $\text{TechMat}(t, i, 1:P)$ is 1 if node $i$ has technology $k$ at epoch $t$, and 0 otherwise;
- an $\text{AdjMat}$ matrix of size $T \times N \times N$, where each entry $\text{AdjMat}(t, i, j)$ is 1 if nodes $i$ and $j$ are linked by any technology, at time $t$, and 0 otherwise.

A single experiment may contain measurements either for a single or for multiple TMIs. In the latter case, we provide the above matrices for each TMI separately, and name them, e.g., $\text{TopMat1}$, $\text{TopMat2}$, etc. We also provide four matrices resulting from the concatenation of all matrices over the time dimension. The latter are called $\text{FullTopMat}$, $\text{FullLocMat}$, $\text{FullTechMat}$ and $\text{FullAdjMat}$, respectively. This enables the emulation of abrupt link connectivity changes, as is often the case in UWANs. In particular, such changes may serve to emulate the performance of adaptive protocols.

A complete summary of the shared dataset is provided in Table I. The experiments from which the dataset has been retrieved were performed for a number of different purposes and applications, including the design of scheduling protocols, physical layer tests, and underwater communications security. As a result, each experiment has peculiarities which make it different.
<table>
<thead>
<tr>
<th>Location, time and coordinates</th>
<th>Topologies</th>
<th>Equipment</th>
<th>Means of measurement</th>
<th>Measurement rate</th>
<th>Total time</th>
<th>Protocol details</th>
<th>Original trial purpose</th>
<th>Interference-free?</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haifa harbor, Israel May 2009</td>
<td>6</td>
<td>Custom modems (ITC and Brüel&amp;Kjær transducers, offline processing)</td>
<td>BER from raw acoustic samples</td>
<td>Once every 5 s</td>
<td>6 hours</td>
<td>Spatial reuse TDMA $L_P$: 100 Bytes</td>
<td>Scheduling and PHY design</td>
<td>Yes</td>
<td>[45]</td>
</tr>
<tr>
<td>Garda lake, Italy December 2015</td>
<td>4</td>
<td>5 EvoLogics SC2R 18-34 modems; network stack on laptops</td>
<td>RSSI from modem logs</td>
<td>Less than once every 5 s</td>
<td>2 hours</td>
<td>TDMA, CSMA/CA $L_P$: 60 Bytes</td>
<td>Scheduling and handshake design</td>
<td>No</td>
<td>[59], [60]</td>
</tr>
<tr>
<td>Werbellin lake, Germany June 2016</td>
<td>5</td>
<td>5 EvoLogics SC2R 18-34, 3 SC2R 48-78, 2 SC2R 80-120 modems; network stack on laptops</td>
<td>Transmission success from modem logs</td>
<td>Less than once every 5 s</td>
<td>~1 hour</td>
<td>OMR, flooding $L_P$: 25 Bytes</td>
<td>Multimodal routing design</td>
<td>Yes</td>
<td>[61]</td>
</tr>
<tr>
<td>ALOMEX15, Spain, Morocco, West Sahara July 2015</td>
<td>1</td>
<td>3 EvoLogics SC2R 18-34, network stack on laptops</td>
<td>RSSI from modem logs</td>
<td>Less than once every 5 s</td>
<td>22 minutes</td>
<td>TDMA and continuous transmission $L_P$: 23 Bytes</td>
<td>Topology discovery</td>
<td>No</td>
<td>[62]</td>
</tr>
<tr>
<td>Hadera, Israel May 2017</td>
<td>1</td>
<td>3 software-defined EvoLogics 7-17 modems; 2 Cetacean RUDAR mk2 recorders; offline processing</td>
<td>RSSI from raw acoustic samples</td>
<td>Once every 20 s</td>
<td>10 minutes</td>
<td>TDMA $L_P$: 30 Bytes</td>
<td>Physical layer security</td>
<td>Yes</td>
<td>[63]</td>
</tr>
<tr>
<td>Hadera, Israel May 2017</td>
<td>2</td>
<td>4 software-defined EvoLogics 7-17 modems</td>
<td>RSSI from modem logs</td>
<td>Once every 5 s</td>
<td>45 minutes</td>
<td>TDMA $L_P$: 25 Bytes</td>
<td>Multimodal scheduling</td>
<td>Yes</td>
<td>[64]</td>
</tr>
<tr>
<td>REP16-A Sesimbra, Portugal July 2016</td>
<td>3</td>
<td>4 EvoLogics S2CR 18/34, laptops</td>
<td>Transmission success from modem logs</td>
<td>Every second</td>
<td>1 hour</td>
<td>Custom protocol $L_P$: 32 Bytes</td>
<td>Topology estimation</td>
<td>No</td>
<td>[65]</td>
</tr>
</tbody>
</table>
from others in our database, and contributes to increasing the coverage of a variety of scenarios. This is reflected in the list, which shows broad differences among the tests: from relatively large networks of 10 modems, to small link tests with 3 modems; from experiments of long duration (up to a few hours) to short experiments of a few tens of minutes; from tests including one type of modems to multimodal tests including multiple acoustic communication transceivers operating in orthogonal bands; and from tests involving commercial modems to tests that include custom modems and offline processing.

In Table I, we describe only the main points for each experiment. The full description is given in the document distributed with each dataset, as well as in related publications cited in each description and in the table. As the database is open to the community, we also welcome external datasets provided by other institutions, with the only constraints that the datasets should be adapted to match the format described above.

**IV. EXAMPLE OF RESULTS**

We now present the results of a network emulator built upon the ASUNA database. We remark that these are just meant to serve as an example, and that the applications are by no means limited by the scope of our results. In Section IV-A we describe the structure of the emulator, whereas in Section IV-B we provide its results.

**A. Structure of a network emulator**

Our example of emulator is a discrete-event system written in an Octave/Matlab-compatible code, and comes with all datasets currently shared. These datasets are already placed in the right subdirectory structure to make it possible to load them correctly in the simulator. In this way, the user can open the main file, TDMAsim.m, and run it upfront to obtain some first results. The emulation code is freely provided along with the dataset on the ASUNA web site, and the users may employ, extend or modify it to suit their purposes. The code has also been uploaded to the Code Ocean platform [66], from where the results provided below can be reproduced.

The baseline emulator implements an interference-free TDMA scheduling protocol, where each node is assigned an exclusive time slot to transmit a unicast packet to any of its neighbors. The
parameters of the protocol can be tuned via a configuration script named setGlobals.m. In the main file TDMAsim.m, a marked section instructs the user how to choose their desired dataset by commenting/uncommenting specific lines. After importing the data from the corresponding .mat files into the structures of the simulator, the emulation sets up the TDMA schedule and arranges a periodic computation of network metrics.

The TDMA schedule is computed based on the distances among the nodes as derived from the LocMat matrix. For a given sound propagation speed (system parameter), the emulator computes the time slot length as the sum of the packet duration (also a system parameter) and of a guard interval as long as the maximum propagation delay in the network. For each TDMA transmission, the emulator uses the instantaneous TMI in order to infer the one-hop neighbors of the transmitter (through the AdjMat matrix). The unique destination is then chosen at random out of this list. In case a multi-modal communication dataset is chosen, the emulator also checks which communication technologies are in use both by the transmitter and by its receiver (through the TechMat matrix) and chooses one of them at random. The transmission outcome is finally determined by comparing the link quality from matrix TopMat to a threshold (system parameter). In the provided code, such threshold is pre-set in order to make it easier for the user to immediately operate with the data, but can be changed in order to obtain different results.

At tunable intervals, the emulator collects relevant metrics for post-processing. This includes a count of the transmitted and correctly received packets, as well as the network throughput. The metrics are plotted at the end of the emulation, and the resulting figures are saved as images. Next, we show results obtained from our TDMA emulation.

B. Results

1) Haifa Harbor: We first discuss results obtained for the “Haifa Harbor” dataset. The experiment was carried out in Israel, and included four boats carrying custom modems. The boats moved to different locations in the harbor at designated times. Due to the structure of the harbor, no communication between docks was possible in the absence of line of sight. Hence, the change in the boats’ locations created a time-varying network topology. A map of the experiment
location is shown in Fig. 2a, and the formed topologies are illustrated in Fig. 2b. The recorded dataset includes the per-link time-varying BER measurements arranged in a single TMI and in per-topology TMIs. The experiment included roughly six hours of data collection.

In order to obtain the longest possible emulation, we resort to the FullTopMat matrix, which contains the concatenation of the datasets corresponding to each TMI. In our emulation, we consider a successful packet delivery only if the instantaneous BER value is less than $10^{-2}$. Considering this threshold, the packet delivery ratio (PDR) and the per-link throughput are shown in Fig. 3. Metrics are collected every 120 seconds and plotted against the collection epoch. Vertical dashed lines mark the instant where the switch between different subsequent topologies occurs, and the TMI enumeration fits the number of topologies in Fig. 2b.

We observe that the PDR changes over time due to both the topology configuration and the

\[4\]This value has been chosen for demonstration purposes. However, we note that this BER regime may be easily related to PER regimes depending on the employed modulation and coding scheme. For example, a BER of $10^{-2}$ yields a PER of about 0.5 for 64-bit, uncoded packets transmitted using BPSK. In the same conditions, applying a convolutional code of rate $1/3$ and soft Viterbi decoding would yield a BER of $10^{-5}$, which enables the transmission of 1024-bit packets with a PER of 0.01 [67, Section 8.2.8].
Fig. 3. Results for the full “Haifa Harbor” dataset. All topology data has been concatenated: dashed red lines indicate the transition between subsequent topologies.

link quality measurements. The former is mostly observed when there is a transition between TMIs, while the consequences of the latter are observed when the TMI remains the same. We also remark that the node deployment affects the throughput, as the maximum propagation delay in the network determines the TDMA slot length, and therefore the packet transmission rate. In all topologies, the maximum propagation delay is about 1 s (corresponding to a maximum distance of about 1500 m), except in topologies 1 and 4, where the maximum propagation delay is 0.88 s and 0.55 s, respectively. For example, in topology T4, this means that the TDMA frame has a significantly shorter duration, which accommodates about 45% more transmissions than in topologies 2, 3, 5, and 6. For this reason, the throughput is larger for topology T4, despite a similar or lower PDR than in topology T3.

2) Berlin Multimodal: We now discuss network emulation results based on the “Berlin Multimodal” dataset, which provides a set of simultaneous measurements from three different acoustic communication technologies. As reported in Table I, the communication technologies used in the experiment are the EvoLogics SC2R 18-34 kHz (5×), 48-78 kHz (3×) and the 80-120 kHz (2×) modems, respectively named LF, MF, and HF in the following, as a shorthand for low-frequency, medium-frequency, and high-frequency. The TDMA emulator assumes that the transmission rates of each modem are 4 kbit/s, 16 kbit/s, and 32 kbit/s, respectively. The setup
of the experiment and the tested topologies are shown in Fig. 4.

The results are given in Fig. 5. Each point along the curves corresponds to average values taken over windows of 30 s. The most significant difference between the TMIs is the performance of the HF modem, which requires a low-noise, short-distance link, in order to operate at its maximum efficiency. Since the distance between the only two nodes with an HF modem was smaller in topologies T3, T4, and T5 than in topologies T1 and T2, the HF throughput is much higher and stable for T3, T4, and T5. We also observe that the success ratio for the LF and
MF links is similar, and slightly lower for MF in topologies T4 and T5. Since the deployment includes a total of 3 MF and 5 LF modems, this explains the similar throughput achieved by LF and MF in Fig. 5b.

The area plot in Fig. 5c shows that the number of packets sent is about the same in each measurement window. The absolute values tend to remain stable over each window and depend on the connectivity of the sub-networks formed by each technology. For example, in topology T3, the nodes transmit fewer MF packets than in all other topologies. The reason is that, in topology T3, all nodes with MF also have LF. More specifically, we recall that in our tested TDMA scheduling protocol, a neighbor is chosen at random, and only then the transmission
technology is determined. Since there are more LF modems, a node with both LF and MF is likely to have additional neighbors, and thus it is less likely to transmit using MF in T3 than in any other topology.

Finally, we demonstrate the flexibility of ASUNA by testing the optimal multimodal scheduling (OMS) scheme in [64]. OMS is an adaptive TDMA-based algorithm that exploits multimodal links in order to schedule transmissions that obey a number of constraints. These include network topology structure, bounds to interference, and measures to favor multihop routing. OMS was already tested at sea via a dedicated experiment [64] (also part of ASUNA, see the second-to-last line of Table I), hence here we rather test OMS using the “Berlin Multimodal” dataset. This also enables a direct comparison against the baseline TDMA protocol considered above.

As before, we concatenate all topologies of the dataset, in order to obtain longer link quality time series, exhibiting significant connectivity changes across subsequent topologies. Figure 6 shows the average throughput per technology. We observe that the OMS protocol adapts well to the characteristics of the topology by allowing simultaneous transmissions over different technologies and by balancing channel access throughout the network. By setting a slot length of 2.5 s, it adapts the packet length to fill this slot length minus the maximum propagation

Fig. 6. Throughput for the OMS protocol tested over the “Berlin Multimodal” dataset. All topology data has been concatenated: dashed lines indicate the transition between subsequent topologies.
delay. This results in a slightly smaller number of transmissions being made, constantly equal to 12 packets per measurement interval of 30 seconds. However, OMS enables transmissions through multiple technologies at the same time, and additionally the above settings yield longer packets than for the baseline TDMA case of Fig. 5. As a result, the throughput achieved by all technologies is higher (see also Fig. 5c).

V. CONCLUSIONS

We presented ASUNA, a shared database containing recorded time-varying link quality measurements from various sea experiments. ASUNA serves as a tool to test underwater acoustic communication network algorithms through emulations or experiment replay. The ASUNA database includes an ensemble of time-varying link quality measures arranged as topology information matrices. The datasets cover different network configurations measured through a variety of acoustic communication devices, and using different network protocols. To demonstrate the use of ASUNA, we described the details and results of an emulation system built to test a time-division multiple-access scheduling protocol over all collected topology matrices. For a multimodal communications dataset, we also test the optimal multimodal scheduling approach in [64]. We freely share ASUNA as well as the emulation code with the underwater communications community, with the hope that ASUNA will constitute a benchmark to test underwater acoustic networking solutions including, but not limited to, scheduling, routing, and automatic repeat query schemes. ASUNA is open to future contributions. With the expansion of the database that would result, we believe that this benchmark has the potential to greatly contribute to establishing and standardizing UWAN research.

ACKNOWLEDGMENT

The authors express their gratitude to the institutions that collaborated to the organization and logistics of the sea trials, and gave their consent that the respective datasets be included in ASUNA. In particular, we thank the NATO STO Centre for Maritime Research and Experimentation (La Spezia, Italy) for the “ALOMEX15” and “REP16” datasets, Rafael Ltd. (Israel) for the “Haifa Harbor” dataset, the Israeli Electrical company for the “Hadera” datasets, and EvoLogics GmbH (Berlin, Germany) for the “Berlin Multimodal” dataset.
This work received support by the European Commission’s Horizon 2020 research and innovation programme under grant agreement no. 773753 (SYMBIOSIS).

REFERENCES


Paolo Casari received the Ph.D. in Information Engineering in 2008 from the University of Padova, Italy. He was on leave at the Massachusetts Institute of Technology in 2007, working on underwater communications and networks. He collaborated to several funded projects including CLAM (FP7), RACUN (EDA), as well as US ARO, ONR and NSF initiatives. He is the PI of the NATO SPS project ThreatDetect, and the scientific coordinator of the EU H2020 RECAP project. Previously, he was Technical Manager of the Italian NAUTILUS and WISEWAI projects. In 2015, he joined the IMDEA Networks Institute, Madrid, Spain, where he led the Ubiquitous Wireless Networks group. Starting late 2019, he joined the faculty at the Department of Information Engineering and Computer Science, University of Trento, Italy.

Dr. Casari regularly serves in the organizing committees of several international conferences. He also serves on the editorial boards of the IEEE TRANSACTIONS ON MOBILE COMPUTING and of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and has been guest editor of a special issue of IEEE ACCESS on “Underwater Acoustic Communications and Networking. He received two best paper awards. His research interests include many aspects of networked communications, such as channel modeling, network protocol design, localization, simulation, and experimental evaluations.

Filippo Campagnaro received the B.Sc. and the M.Sc. degrees from the University of Padova, Italy, in 2012 and 2014, respectively. In 2014, he joined the SIGNET Group, within the Department of Information Engineering of the same university, as a Research Engineer. In October 2019 he completed the Ph.D. program in Information Engineering and joined the same department as a post-doctoral researcher and lecturer of ICT master courses. His Ph.D. thesis involves the design and the evaluation of multimodal underwater optical and acoustic networks, with a particular focus on simulation and field experimentation (defence scheduled for February 2020).

Filippo has joined many sea trials with the NATO STO CMRE (La Spezia, Italy), EvoLogics GmbH (Berlin, Germany), ENEA (Rome, Italy), the IMDEA Network Institute (Madrid, Spain), and the University of Haifa (Israel). He is the Technical Manager of the MarTERA RoboVaaS project, and collaborates with the European Defence Agency project SALSA. His current work involves software development for the DESERT Underwater framework as well as the design of novel communication stacks and architectures, with special attention to cross-layer and multi-technology designs. His research interests revolve mainly around the design, analysis, implementation, and field evaluation of multimodal underwater networks.

Since 2017, he collaborates with Wireless and More srl, a spin-off company of the University of Padova, which he joined as a part-time employee in October 2019. His tasks in the company include the design and implementation of software modules for multimodal underwater networks.
Elizaveta Dubrovinskaya received her Diploma degree from the St. Petersburg State Transport University, Russia and her M.Sc. degree from the Christian Albrechts University of Kiel, Germany in 2008 and 2012, respectively. Her Master’s thesis was focused on the development of algorithms for a shallow-water acoustic localization. Currently, she is a board member of Teleone OU (a telecommunications startup in Tallinn, Estonia), and is working towards her Ph.D. degree at the IMDEA Networks Institute and Carlos III University of Madrid, Spain.

During her Ph.D., Elizaveta has been a visiting student at the Florida Atlantic University, where she worked on underwater LiDAR systems, and at the University of Haifa, Israel, where she worked on environment-aware acoustic localization algorithms and took part in multiple sea trials. The latter work received the Best Paper Award at the WPNC 2017 symposium. She is currently collaborating with the EU H2020 SYMBIOSIS project, focusing on acoustic array design, array processing, as well as on the implementation of an active sonar for marine life monitoring in collaboration with the University of Haifa and EvoLogics GmbH (Berlin, Germany). Her research interests revolve around many aspects of underwater acoustic networks, including algorithm design and field experiments for acoustic localization.

Roberto Francescon received his degree in Telecommunications Engineering from the University of Padova, Italy, in 2015. During his Master’s thesis he worked on vector quantization, motif extraction, and machine learning techniques for the understanding of a signal’s evolution over time. He got involved in the underwater communications field after graduation, by developing software modules for the DESERT Underwater framework. He joined many sea trials with the NATO STO CMRE (La Spezia, Italy), and EvoLogics GmbH (Berlin, Germany). He also collaborated to programming the framework that is running on the THEMO marine observatory, off the coast of Israel. He is currently a software developer at Wireless and More srl, a spin-off company of the University of Padova, where he focuses on the development of software supporting and managing new communication protocols and schemes for wireless networks.

Amir Dagan has been a programmer for the last 30 years, both as a professional and as a hobbyist. During the last 10 years, Amir has made CRM software, content management systems for web sites (including large PHP projects) as well as PHP/JS/CSS and HTML5 frontend projects. Amir has vast knowledge with PHP: he has built interfaces, used APIs, coded from scratch via the MVC method, as well as refactored PHP code by other programmers.
Shlomo Dahan works as an acoustics engineer for the THEMO project. He retired from the Israeli Navy as a manager for submarine systems (sonar, navigation and weapon systems). In 2013 received the “best project” award from the Israeli Navy’s Commander-in-chief.

Michele Zorzi (S’89–M’95–SM’98–F’07) received the Laurea and Ph.D. degrees in electrical engineering from the University of Padova in 1990 and 1994, respectively. From 1992 to 1993, he was on leave at the University of California at San Diego (UCSD). In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications, UCSD, in 1998, he joined the School of Engineering, University of Ferrara, Italy, where he became a Professor in 2000. Since November 2003, he has been on the faculty of the Information Engineering Department, University of Padova. His current research interests include performance evaluation in mobile communications systems, WSNa and Internet of Things, cognitive communications and networking, 5G mmWave cellular systems, and underwater communications and networks.

He was the Editor-in-Chief of IEEE WIRELESS COMMUNICATIONS from 2003 to 2005 and of the IEEE TRANSACTIONS ON COMMUNICATIONS from 2008 to 2011. He is currently the Editor-in-Chief of IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING. He served as a Member-at-Large for the Board of Governors of the IEEE Communications Society from 2009 to 2011, as its Director of Education from 2014 to 2015, and as its Director of Journals from 2020 to 2021.

Roee Diamant received his Ph.D. from the Department of Electrical and Computer Engineering, University of British Columbia, in 2013, and his B.Sc. and M.Sc. degrees from the Technion, Israel Institute of Technology, in 2002 and 2007, respectively. From 2001 to 2009, he worked as a project manager and systems engineer at Rafael Advanced Defense Systems, Israel, where he developed a commercial underwater modem with network capabilities. In 2009, he received the Israel Excellent Worker First Place Award from the Israeli Presidential Institute. In 2010, he received the NSERC Vanier Canada Graduate Scholarship. In 2015 and 2016, he was a visiting professor at the University of Padova, Italy.

Dr. Diamant has received three Best Paper awards, and serves as an associate editor for the IEEE JOURNAL OF OCEANIC ENGINEERING. Currently, he is the coordinator of the EU H2020 project SYMBIOSIS (BG-14 track), and leads the underwater Acoustic and Navigation Laboratory (ANL) as an assistant professor at the Department of Marine Technologies, University of Haifa, Israel. His research interests include underwater acoustic communications, underwater navigation, object identification, and classification.