Implementation of a Multi-modal Acoustic-Optic Underwater Network Protocol Stack

Filippo Campagnaro§, Federico Guerra§, Paolo Casari², Roeo Diamant‡, Michele Zorzi§

§Department of Information Engineering, University of Padova, via Gradenigo 6/B, 35131 Padova, Italy
®IMDEA Networks Institute, Madrid, Spain
‡Department of Marine Technology, University of Haifa, Israel
*Corresponding author. Email: campagn1@dei.unipd.it

Abstract—We describe the implementation of a protocol stack for multi-modal underwater networks, where multiple physical layer technologies are available to each node. This condition implies greater flexibility, by allowing each node to decide how to serve specific transmissions and traffic classes depending on system, quality-of-service, and application-specific policies implemented in a controller. In this paper, we describe the implementation details of an acoustic/optic multi-modal underwater network stack through the DESERT Underwater framework. Our system services various traffic types of different quality-of-service demands, and allows mobility. Extensive numerical results show that our multi-modal system offers greater throughput, robustness to mobility and traffic types, and provides much lower service delay.

Index Terms—Underwater networks; multi-modal acoustic/optic communications; multi-hop routing; mobility; simulation; NS-MIRACLE; DESERT Underwater

I. INTRODUCTION

Sharing information between submerged nodes at sea is required for any underwater application. This includes underwater surveying, command and control, and environmental monitoring. Since cables are heavy and their deployment is expensive, wireless communications are needed.

The technologies developed for underwater transmissions are often suited only to specific scenarios. The available technologies include optical underwater communications, which typically find their ideal application in very short-range communication scenarios in clear waters with little background light. Once a link is established, optic underwater communications can provide bit rates up to several Mbps [1], [2]. Radio-frequency and magneto-inductive communications cover similar ranges with lower bit rates, but do not require alignment efforts and do not depend on light conditions [3]–[5]. Low- and intermediate-frequency acoustic communication technologies [6], [7] are characterized by low bit rates, up to a few kbps, but their transmissions typically cover longer ranges, up to a few km. High-frequency acoustic communications have a comparatively shorter reach, but provide higher bit rates, on the order of a few tens of kbps [8], [9].

Current solutions tend to choose the communication technology for a given underwater application upfront: this limits the capability of the network to adaptively manage traffic demands. In addition, when the application requires significant flexibility in terms of coverage and link-level transmission performance, it is unlikely that a single communication technology can support all application requirements. To bridge these performance gaps, multi-modal systems have been proposed to allow an adaptive choice of communication technology according to the network demands and the channel conditions. These systems are the focus of this paper.

A multi-modal node is composed of communication subsystems exploiting different physical layer (PHY) technologies (e.g., acoustic and optic) and/or different implementations of the same PHY (e.g., high-frequency/short-range and low-frequency/long-range acoustics). Multi-modal systems exploit their inherent flexibility in a seamless manner via flexible user-defined policies to provide quality-of-service to different communication needs. Multi-modality involves an adaptive adjustment of the entire networking stack.

In this paper, we describe the implementation details of a network stack for multi-modal underwater communications. Our solution matches the communication requirements of the network with the capabilities of available PHY systems. To demonstrate these capabilities, we implement the first multi-modal network stack over the DESERT underwater framework [10]. The framework supports a smooth transition towards experiments with multi-modal communications over real underwater modems. Numerical results show that, compared to benchmark systems relying on a single communication technology, our multi-modal scheme provides higher throughput and lower service delay. The modular structure of DESERT Underwater allowed us to exploit several underwater network protocol implementations already available in the framework. In addition, several models of different PHY communication technologies are already available in DESERT. The simulator also supports the connection to real underwater modems, and thus will serve as a proof-of-concept experimentation engine in a future extension of this work.

The remainder of this paper is organized as follows: Section II presents related work; in Section III we detail the application scenario and the requirements to be supported by the network; Section IV describes the implementation of a multi-modal network stack in the DESERT Underwater framework [10]; Section VI presents the simulation scenario and results; finally, Section VII draws our concluding remarks.

II. RELATED WORK

Multi-modal systems in the context of underwater networks have been first introduced in [11]. By utilizing the energy
consumption difference between optic and acoustic communications, the authors proposed a data muling approach for data retrieval from a distributed underwater network. The setting includes broadcast of short packets over acoustic communications, and optic communications to convey bulk data once the mobile data mule has approached underwater sensors. Notably, both the acoustic and the optic modem implementations are custom. The work in [12] applies a different multi-modal system encompassing underwater acoustic communications and surface-level radio communications. Considering the consumed energy, the system automatically balances surfacing to employ radio communication, or acoustic transmission from the deep. Multi-modality is achieved in [13] via a mixture of different acoustic communication technologies. More recently, the authors in [14] implemented a similar concept on a custom reconfigurable underwater acoustic mode. The work include a demonstration to switch between two modulation types. Specifically, the he NATO standard JANUS and a higher-rate modulation format based on multilevel frequency shift keying.

The work in [15] uses acoustic communications to help manage routing in a clustered multi-hop optical network. The longer range acoustics connects cluster heads, thereby allowing sharing of cluster information for intra-cluster communications. In [16], hybrid acoustic/optic multi-modal networks are considered for the transmission of real-time video streams. Here, data streaming takes place through the optic channel, whereas acoustic communication is employed to send acknowledgment packets, transmit data while the optical devices are being aligned, and as a fallback solution in exceedingly turbid waters. A hybrid acoustic/optic communications is offered in [17] to coordinate and transfer information within swarms of autonomous underwater vehicles.

The above works employ specific multi-modal solutions for specific application. Yet, the changing channel conditions and traffic demands may require a more flexible implementation. In this work, we advocate that the flexibility of multi-modal communications should be leveraged in underwater network. To this end, we implement a multi-modal network stack that can be easily configurable to support changing service needs. We implemented the system on the free-access DESERT underwater framework [10]. The emulator allows flexible change of the multi-modal configuration to support serious communication technologies, traffic types, and service demands. In the following, we introduce the capabilities of our system over a test case underwater network, and give the details of our implementation.

III. APPLICATION SCENARIO AND REQUIREMENTS

Our system setting consists of a team of divers performing some underwater operations. The team includes several divers, one leader, and one remotely operated vehicle (ROV), each of which performs as a network node. The nodes work to pass information across the network to a surface vessel. The team has a broad range of requirements, typically leading to a diverse set of underwater operations and actions to be taken. First, the team must coordinate in order to carry out a set of prescribed tasks. For this, the leader will need to transmit short messages reliably. Second, large-size data packets need to be conveyed to the surface vessel to manage the progress of operations. Third, the status of each member of the team must be periodically monitored to avoid unsafe situations.

We assume that each network node is equipped with a predefined set of PHY technologies. In particular, the leader is equipped with underwater optics (OPT), short-range acoustics (SRA) and long-range acoustics (LRA), the ROV is equipped with OPT and LRA, and the remaining divers are equipped with SRA and LRA. The implemented system was made general to allow an easy change to this choice of technology allocation. In order to carry out and complete their tasks, the nodes exchange three types of messages, namely:

- **Health messages**: short status report packets delivered periodically to the ROV via the leader.
- **Control messages**: short packets, transmitted by the leader to any 1-hop neighbor, and conveying orders or commands (e.g., relocation).
- **Image packets**: long packets, transmitted by any node to the surface vessel through the team leader.

The transmission of each packet type is administered via the OPT, LRA, or SRA technologies depending on specific rules that represent the organization of the information flow in the network, and that manage the tradeoff between possibly competing requirements. An example for such competing service demands are control messages that should be transmitted reliably without interfering with the throughput of health messages. Another competing requirement is that image packets should be serviced with high throughput, while also allowing the implimely delivery of health and control messages. The implementation of a structured system to support the above listed requirements is the subject of the next section.

IV. IMPLEMENTATION OF A MULTI-MODAL MULTI-MESSAGE STACK IN DESERT UNDERWATER

A. Controller implementation details

The level of flexibility to be supported by the network required the implementation of several additional features at the level of both DESERT Underwater and its underlying engine NS-MIRACLE [18]. Since different message types can be serviced by various network protocols, cross-layer information is needed within the networking stack. The possibility of using different PHY technologies to service packets towards different destinations requires a dynamic management of the network stack. Based on the traffic demands and on information about the quality of the different PHY links (e.g., the network topology, link distances, and the past history), this control module funnels packets through dynamic choices of the available PHY and protocol options.

The main idea behind our implementation is that the policies that manage multi-modal transmissions should be message-driven. For this reason, a transmission control module should contain several queues, one for each type of message to be supported. Furthermore, the module should be able to wire the queues dynamically to the correct combination of
network protocols, link-level protocols, and PHY technologies. Implementing these functions requires additional cross-layer communication capabilities. These cross-layer messages enable the controller to easily query PHY-level metrics (e.g., the instantaneous signal strength perceived), as well as link-level metrics (the per-link packet error rate, both overall and relative to a given PHY). The retrieved metrics help make decisions related to which protocols and which PHY should be used to transmit a packet at any given time.

The features described above and the logic to manage packet transmissions have been implemented as part of a multi-modal multi-message (M^4) controller. It is worth remarking that the above functions substantially extend the support of DESERT Underwater for multiple PHYs on the same node. This includes the transmission, propagation and reception models for underwater acoustic and optical signals, and the automatic switch algorithms already presented as part of [19].

The M^4 controller has been implemented as an additional network layer within a DESERT add-on. This additional layer stores the packets received from the upper layers in different queues per traffic type. To each traffic type T, we associate a list ℓ(T) of lower layers to which it is possible to forward packets, and their management policy. A policy defines the behavior of the system. The user is free to define any number of policies depending on the application and on its requirements. We differentiate between policies that require channel probing and policies that allows immediate channel access. Physically, a probe is a short packet of prescribed size is sent to test the availability of a desired link. In this paper, we focus on two policies for the lower layers, which we name Fast and Robust:

- a lower layer that provides high bit rate and short range transmissions is managed through a Fast policy. The node can forward a packet to this layer only after probing it to verify its availability (e.g., an optical PHY would be managed via the Fast policy);
- a lower layer that can be safely assumed to be available with high probability is managed through a Robust policy. Packets can be sent immediately to this layer without the need to probe it (e.g., a long-range acoustic PHY would be managed via the Robust policy).

For any given traffic type T, the list ℓ(T) can include multiple Fast layers, but only one Robust layer. When a node receives a packet of type T, it checks ℓ(T) and acts as follows:

1) If T includes one or more Fast layers, the node probes these layers in the order they appear in T. If the probing procedure is successful, the layer reports this to the M^4 controller, which then forwards all packets of type T in its queue to this layer for transmission.

2) In case after the probing procedure no Fast layers was found available, the packet is sent through the Robust layer, as long as ℓ(T) lists any.

In case only Fast layers are listed in ℓ(T), the node loops over point 1) in accordance with the following rules.

- The node sends the first probe in unicast to the destination of the most recent packet in the queue, and starts a timer with expiration period equal to TTO seconds, where “TO” stands for timeout.
- if a timeout occurs, the node broadcasts the probe to all neighbor nodes, increases a failure counter N_F, and reschedules the timer after N_F × TTO.
- The node continues these actions until a response is received. If it does not receive answers to his probes for N_F^max times or more, it reschedules further attempts after N_F^max × TTO. The latter is performed without increasing the failure counter N_F further.

Each node specifies only one upper layer per packet type. Upon reception of a packet, the M^4 controller forwards a packet to this specified upper layer. An example to this configuration is shown in Fig. 1, where we depict the configuration of the M^4 controller for the application introduced in Section III.

V. SIMULATION SCENARIO

A. Configuration of the diver coordination application

The information flow for the diver operation support application in Section III is organized as follows:

- Health packets are periodically sent by the divers to the leader using LRA. The leader forwards these health messages as well as it’s own health to the ROV using either LRA or OPT (when possible).
- Control packets are sent by the leader to any of the other network nodes over LRA.
- Image packets are sent by the divers to the leader over SRA, and are conveyed by the leader to the ROV using OPT when available.

As health packets require a reliable resource allocation, they are transmitted via a Time-Division Multiple Access (TDMA) MAC protocol. Conversely, both control and image packet transmissions rely on a Channel-Sense Multiple Access (CSMA) MAC protocol, which improves the responsiveness of the system and does not bind transmissions to a specific time slot. A scheme of the implemented modules is shown in Fig. 1.

To demonstrate the system, we simulate an underwater network of one ROV and 5 divers, of which one is the leader. The position and mobility patterns of the network nodes are illustrated in Fig. 2. The ROV moves randomly at constant speed v_R in the ROV area, whereas the leader moves randomly at constant speed v_L along the vertical trajectory located between the ROV’s and the divers’ areas. At random intervals, both the leader and the ROV are forced to move to a “rendez-vous ROV” (RVR) position, indicated as a square in Fig. 2, in order to ensure the delivery of image packets through optic communications. At this location, their mutual distance is set to be about 6 m. Each of the four remaining divers has an assigned working area. For the divers, we distinguish between two different mobility patterns, respectively named MobA and MobB. In MobB, each diver can move at random at constant speed v_D within their own area. With the MobB pattern, at random intervals the divers return to a rendez-vous point RV_n, n = 1, . . . , 4, located in the central upper corner of their movement area, close to the leader line. This ensures a more frequent use of SRA for communications between the divers and the leader.
configuration, the optical transmission range is \( \approx \) ambient light noise via lookup tables [19]. In any event, we remark that our simulator supports the inclusion of overnight operations or modem with solar irradiance rejection, e.g., as outlined to 0.4 m\(^{-1}\) a coastal channel, with an optical attenuation coefficient equal of 20 dB to ensure correct reception, an optical wavelength of bandwidth, 1-Mbps of PHY bit rate, an SNR threshold is configured with a transmission power of 30 W, 100 kHz OPTICAL PHY LAYER the source. For OPT, the modem’s source level is set to 167 dB re \( S2CM \) HS [8] acoustic modem. For both SRA and LRA, the configuration reproduces the characteristics of the EvoLogics S2CR 18/34 acoustic modem [21] in instant messaging mode. For SRA, we employed a second instance and 16 kHz of bandwidth. This simulates the behavior of transmission range of about 3.5 km, carrier frequency 26 kHz and a bandwidth of 80 kHz. This configuration reproduces the characteristics of the EvoLogics S2CM HS [8] acoustic modem. For both SRA and LRA, the modem’s source level is set to 167 dB re \( \mu \) Pa\(^2\) at 1 m from the source. For OPT, the OPTICAL PHY LAYER module [19] is configured with a transmission power of 30 W, 100 kHz of bandwidth, 1-Mbps of PHY bit rate, an SNR threshold of 20 dB to ensure correct reception, an optical wavelength \( \lambda = 532 \) nm and a divergence angle \( \theta = 0.5 \) rad. We consider a coastal channel, with an optical attenuation coefficient equal to 0.4 m\(^{-1}\), and with no optical ambient light noise. In this configuration, the optical transmission range is \( \approx 12.5 \) m.

Health messages are generated periodically every 40 seconds, and comprise of \( N_H = 800 \) bits. Image and control messages are generated according to a Poisson process of rate 400 s and 40 s, respectively. The image message size is \( N_I = 80000 \) bits, and a control message is \( N_C = 800\) bit long. For system evaluation, we simulate the network for a total of \( 10^6 \) s. In the simulation, the nodes start from an initial position (the RV location in Fig. 2), and then move at random with uniform velocity inside the boundaries of their own area.

VI. RESULTS AND PERFORMANCE EVALUATIONS

In our simulations, we focus on evaluating the network performance improvements provided by multi-modality. To that end, we compare the performance of the multi-modal system presented in Section V against a benchmark system with a single PHY technology. For each traffic type we focus on different performance indications, in light with the application requirements. Namely, for health packets the objectives is to maximize throughput and minimize the packet delivery delay, while for the control messages and the images we want to maximize the packet delivery ratio. Finally, we will consider the impact of mobility on the performance of the multi-modal case by considering different movement area sizes and mobility patterns.

We start with Fig. 3, which presents the complementary cumulative distribution function (CCDF) of the delivery delay for health packets. We consider four cases: with and without multi-modality; and when the sum of the distance between the diver and the leader plus the distance between the leader and the ROV is either \((70 \pm 50) \) m or \((700 \pm 50) \) m. Only the 

![Figure 1. Structure of the M³ protocol stack for underwater communications. The controller decides which routing/MAC/PHY stack a message has to go through, based on message type and the policies.](image)

![Figure 2. Mobility of the three different entities: ROV, divers and leader. Each nodes moves at random within their indicated area, plus the ROV periodically returns to a rendez-vous point named RVR to facilitate data exchange with the leader via OPT (mobility pattern MobA). Under a second mobility pattern MobB, the divers also have a rendez-vous point RVR assigned to them, in order to enable more frequent SRA contacts with the leader.](image)
to their combination, the delivery delay in the presence of multi-modality is less than 2 s in 30% of the cases, whereas in case only LRA is available the delay is never less than 3.75 s. In the second case, when the typical distance between the divers and the ROV is larger, there are fewer chances to exploit the OPT and SRA systems. Anyway, multi-modality is exploited when possible, resulting in lower delivery delay in 5% of the cases with respect to a case with only LRA. However, when neither the OPT nor the SRA are in range, the packets are sent via LRA after probing the other channels. These delays are typically negligible, but justify the slightly higher probability of experiencing delays around 4 and 4.25 s with the multi-modal system.

The advantages of multi-modality in terms of throughput experienced by health packets are shown in Fig. 4, where we show how the throughput varies as different traffic types are serviced simultaneously, as a function of the divers’ movement area. The results show that multi-modality ensures a stable throughput for health messages even in the presence of large image data transmissions, and even over large movement areas. Conversely, when the nodes are equipped only with LRA, the throughput of health messages decreases significantly, up to 20% of its desired value in case image transmissions share the LRA technology. Although in the multi-modal system the health throughput is robust to the superposition to other types of traffic, it still marks a decreasing trend with increasing movement area. Conversely, better results are achieved with the MobB pattern, which ensures that the SRA and OPT systems are in range more often thanks to rendez-vous points. The corresponding performance is shown by the dot-dashed curve with triangle markers.

We proceed by analyzing the packet delivery ratio (PDR) for control messages in the presence of different traffic types as a function of the divers’ movement area size (Fig. 5). We recall that control messages are sent by the leader directly to other nodes via CSMA over LRA. Therefore, the delay is a comparatively less interesting metric in this case, whereas it is important that control packets are delivered reliably. We observe that the PDR of control packets is only slightly effected by the addition of health traffic, both with and without multi-modal capabilities. If image transmissions are added to the LRA-only system, however, the PDR decreases to a value about 0.2, whereas the multi-modal configuration is mostly unaffected. In any event, by allowing a more frequent use of the SRA subsystem via the mobility pattern MobB, the PDR of control packets increases to exactly 1.

The results discussed so far confirm that a non-multi-modal system becomes unstable in the presence of image traffic. Therefore, we will now discuss the PDR of image traffic only in the presence of multi-modal communications. Fig. 6 shows the PDR of image traffic against the diver’s movement area size, for different values of $N_I$. We observe that the PDR decreases for increasing movement area size in all configuration. This is because either the SRA is mostly out of range for large areas, or because probes are successfully exchanged while a point-to-point SRA link is close to its maximum range, and becomes unavailable while the transmission is in progress. The latter effect is mitigated by the usage of low image packet sizes $N_I$. When the mobility pattern MobB is employed, the nodes find more chances to operate image transmission through SRA, which substantially improves the PDR of image packets, especially in large movement areas.

**VII. CONCLUSIONS**

In this paper we presented an implementation of a multi-modal underwater network stack using the DESERT Underwater framework. The network exploits the simultaneous presence of different network protocols and PHY technologies to efficiently serve different traffic types with different performance requirements. The concept has been showcased by simulating an underwater diving operations scenario, where a leader diver coordinates the mission of an ROV and a group of divers, and at the same time it helps transferring data from the divers to the ROV. Our numerical results show that, compared to networks of nodes using a single communication technology, our multi-modal system is more
robust to mobility and to the simultaneous service of different traffic types, and delivers substantially better performance in terms of throughput and service delay. This performance gain compounds when the mobility patterns are engineered to favor the establishment of contacts among short-range higher-bit rate technologies.

**Acknowledments**

This work has been supported in part by the US Office of Naval Research under Grant no. N62909-14-1-N127 and by the European Commission through the 7th Framework Programme, AMAROUT II project, G.A. no. 291803.

**References**


Figure 5. Control packets PDR vs the divers’ moving area size. Comparison between the case with and without multi-modality, by adding different traffic types. The PDR is mostly unaffected by the increasing area size, especially in the presence of the mobility pattern MobB.

Figure 6. Packet Delivery Ratio (PDR) for image packets vs the size of the divers’ movement area for different values of the image packet length $N_l$. Multi-modality achieves better PDR values, especially in the presence of the MobB mobility pattern.