Interception of bio-mimicking underwater acoustic communications signals

Tamir Mishali, Paolo Casari, and Roee Diamant

***Abstract*—We describe a biomimicking interception scheme tailored to Underwater Acoustic Communications (UAC), which aims at separating authentic and biomimicking signals. Our interceptor leverages the expected stability of the biomimicking sources, as opposed to vocalizations by marine fauna, which are expected to move fast and rapidly change orientation. Consequently, the channel impulse response (CIR) of the link between a receiver and a biomimicking source is expected to be much more stable than those corresponding to actual vocalizations. We quantify this stability by testing the randomness of the representation of the CIRs. The latter are represented by two similarity metrics: the cross-correlation and the sample entropy between adjacent CIRs features. We offered two interception measures: 1) testing the similarity between a Gaussian distribution and the distribution of the similarity measures using the Kullback–Leibler divergence (KLD) criteria for quantification, and 2) the minimum number of clusters to effectively segment the similarity measures as a point cloud. Results from simulations for artificial signals mimicking dolphin whistles and the outcomes of a lake trial demonstrate the effectiveness of our biomimicking interceptor.**

***Index Terms*—Underwater acoustic communications; physical layer security; biomimicking; interception; dolphin whistles**

I. Introduction

Underwater Acoustic Communication (UAC) is used in applications such as data collection from oceanography sensors, operation coordination between scuba divers and between submerged vehicles, and for search-and-survey missions [2]. When used for security purposes, there is a need for secure communications to avoid disclosing the location of the transmitter.

Since the beginning of the "Internet of underwater things" terminology [3], there has been an increasing interest for smart, multi-purpose underwater networks of intelligent devices that can coordinate not only for general-purpose data exchange [4], [5], but also to accomplish detection tasks [6]–[8]. As underwater IoT devices become a key element of ocean connectivity, security issues arise [9], [10], and call for solutions that are specific to the underwater environment [11]. The above aspects are extremely relevant for the IoT community, as also shown by a recent special issue on IoT for smart oceans [12].

While several security schemes attempt adapting terrestrial network security schemes to underwater scenarios [13] or exploit the properties of underwater acoustic channels for authentication and privacy purposes [14]–[16], an entirely different line of work focuses on covert communication via Low Probability of Detection (LPD) signaling. This family of techniques differs from Low Probability of Interception (LPI) in that LPI aims to conceal the information bits, whereas in LPD the objective is to hide the very existence of communications from a possible interceptor.

Methods for LPD UAC often involve spreading the emitted signal power over frequency or time, such that the interceptor will observe the signal power to be less than its own noise floor. For instance, a popular method is to apply the spread spectrum technique across the frequency domain and despread the signal at a legitimate receiver. As a result, at some distance from the transmitter, the spectrum of the signal drops below the noise level and standard interception techniques like energy detection [17], [18] fail to detect it. Spreading is also achieved by convolving the modulated signal with a pseudo-random sequence, by using compressed signals such as chirps for modulation or via frequency hopping [17]. However, the relatively narrow bandwidth of underwater acoustic signals (usually around 10 kHz) and the limited frequency band available for transmission (usually from 5 kHz to 40 kHz for mid-range distances), both of which are due to the properties of the channel, makes interception easier. In particular, a possible interception technique can perform an exhaustive search within possible spreading sequences across the entire band, while keeping the computational complexity feasible. An alternative LPD communication that avoids these limitations is biomimicking.

Biomimicking is a general term for the playback or synthesis of real biological sounds to be used as the modulation signal for covert communications. The key idea is to transmit information-bearing signals that resemble bio-acoustic emissions. The interceptor can detect the communication signals, but is mistaken to consider these signals as coming from marine animals and therefore ignores them. As a result, different than the above LPD approaches, biomimicking allows a high source level, and the signal to noise ratio (SNR) at the receiver is likely to be high. On the other hand, the need to use biological signals for modulation limits the spectral efficiency of the communications to capacity values well below 1 bit/s/Hz [19].

Biomimicking communication is explicitly used in the underwater medium. Common biomimicking UAC methods

T. Mishali and R. Diamant are with the Department of Marine Technologies, University of Haifa, 3498838 Haifa, Israel. R. Diamant is also with the Faculty of Electrical Engineering and Computing, University of Zagreb, 10000 Zagreb, Croatia.

P. Casari is with DISI, University of Trento, 38123 Povo (TN), Italy, and with CNIT, 43124 Parma, Italy.

Corresponding author: Roee Diamant, email: roee.d@univ.haifa.ac.il.

Parts of this work were presented at the IEEE OCEANS 2023 Conference, Limerick [1]. This journal version extended the conference publication by adding: 1) extended survey, 2) use another similarity metric, 3) use of the stability metrics, 4) inclusion of attacker strategics and performance comparison, and 5) results of extended simulations and a field trial.

This work was sponsored in part by the NATO Science for Peace and Security Programme under grant no. G5884 (SAFE-UComm), and by the European Union's Horizon Europe programme under the UWIN-LABUST project (project no. 101086340).

involve a variety of biological vocalizations ranging from dolphin whistles and clicks to whale calls. The information is modulated, either by replaying an alphabet of recorded sounds, or via phase modulation over symbols that resemble bio-acoustic signals. Examples include using humpback whale song segments to function as symbols [20], hiding a carrier signal [21] or creating complex communication frames based on patterns of marine animals vocalizations [22], [23].

The existing literature on biomimicking interceptors presents a varied landscape. An interceptor designed for bionic binary orthogonal keying modulated signals (BBOKMSs) coded by Time-Frequency (TF) contour, as proposed in [24], leverages the statistical behavior of the contour slope to discriminate between authentic and BBOKMS whistles. However, its applicability is confined to a specific biomimicking modulation method. For the more practical case when the signal structure is unknown, alternative interceptors highlight the limitations of the transducer hardware for transmitting bio-acoustic signals, and use these limitations to distinguish between biological and biomimicking signals. These limitations include distortion of the biomimicking signals, which can be identified from lower entropy compared to real bio-acoustic signals [25], and a limited ability to mimic signals such as dolphin whistles, whose phase varies rapidly in time [26]. The accuracy of these interception approaches is reduced in high-reverberation environments, where channel-induced distortions make it difficult to distinguish between physical distortions and those caused by hardware limitations. Further, hardware limitations may be improved as the future design of underwater sensing technology improves.

In this paper, we present a new concept for biomimicking interception that relies on the dynamics of the source. We target a robust design, so that our interceptor work in diverse marine environments and be insensitive to the biomimicking modulation type and signal used. Our approach is based on the assumption that the biomimicking transmitting platform moves slower and smoother than an authentic biological source. For example, while a dolphin is expected to rapidly change its course, speed and depth [27], [28], an underwater vehicle or a human scuba diver is more likely to move along a straight trajectory. As a result, we expect the channel impulse responses (CIRs) between a biological source and the interceptor to vary significantly in time compared to the CIR variations of a biomimicking source. We therefore consider the stability of the CIR's as a measure of interception. For stability metrics, we propose the cross-correlation and the cross-sample entropy between CIRs. The former quantifies the relation between two CIRs while the latter statistically tests the information exchange between the CIRs. When considering the CIRs of signals received by the interceptor as point cloud, these two metrics are translated to similarity measures by calculating their degree of *randomness*. Here, high similarities within the point cloud members indicates a stable source and the signals are classified as biomimicking, whereas random-like similarities would correspond to real biological vocalizations. A heuristic solution on how to represent the CIR from the detected (unknown) signals as well as an analysis of possible attack strategies are also presented.

Our contribution is twofold:

1) A biomimicking interception method that builds on the CIR's stability as a classification metric. While current interception techniques that builds on the waveform structure, either the spectra-time characteristics or the phase, we avoid making assumptions on the biomimicking signal or on the communication hardware. Instead, we build on the statistical relation between the channel impulse response of closely received symbols to identify if the signal originates from a real biological source or a biomimicking one. The tools we use for similarity measures are correlation and sample entropy, while for detection we test how random the similarities are.
2) A clustering solution to determine whether a source is mobile or static. To identify the source type of the analyzed signal, we offer two approaches: 1) testing how close the empirical distribution of the similarity measures is to a Gaussian distribution, where a Gaussian distribution would mean weak channel relations and thus a real biological source, and 2) clustering the similarities between the channels and setting a threshold on the number of clusters, where a low number of clusters would refer to a less random ensemble and thus a biomimicking source. This way, we can make our interceptor insensitive to the signal transmitted, hardware used, or modulation types.

To explore the performance of our interceptor, we performed both numerical simulations and analysis of recordings from a field experiment. The simulations included a propagation model to generate realistic CIRs while changing the location of the transmitter to be either drifting (biomimicking source) or changing in a random walk fashion (real bio-acoustic source). The experiment included emissions of chirplet signals from a boat drifting or moving in a lake at different speeds. The results show a favorable trade-off between precision and recall for the classification between biomimicking and real signals, and reveal robustness to environmental conditions.

As opposed to the current literature that leverages the limitations of a transducer to recognize artificial sounds [25] or leverages time-dependencies to detect lack of phase discontinuities and reveal biomimicry [26], our approach in this paper focuses on a different characteristics of biological sound sources: mobility. By discriminating on whether the source of a biological signal is intentionally moving or drifting, and by using the CIR's stability for classification, we achieve a more general solution. For instance, we do not need to know the transmitted signal or hardware specification in advance.

The remainder of this paper is organized as follows. Section II presents related work. Section III describes our system model and assumptions. Section IV outlines the details of our interception methodology. Simulations and results from the field experiment are discussed in Sections V and VI respectively, and conclusions are drawn in Section VII.

II. Related work

Biomimicking UAC techniques focus on the playback or synthesis of biological sounds that are typical both of the

geographical region where the communication takes place and of animals found in the area. It is assumed that a potential interceptor will become used to receiving such bio-acoustic signals, and will not be able to identify the biomimicking signals as communication. The available literature focuses on the vocalizations of dolphins and whales, as these signals are frequently heard and at the same time complex enough to offer options for information modulation [29]. A good attribute of biomimicking is how much a synthetic signal resembles biological sounds. The method in [30] proposes an analytical signal model whose instantaneous frequency is non-linearly time-varying to match the signature of a dolphin's whistle. The encoding is based on the parameters of the signal, such as its duration, shape and bandwidth. At the receiver side, decoding requires to first obtain a maximum likelihood estimate of the signal's parameters, and then to map such parameters to information bits. Possible drawbacks are the atypical behavior in dolphin whistle patterns and different modulation applied to each biomimicking signal, whereas biological signals are expected to exhibit time correlation [31], [32].

Considering this, the method in [23] imitates a sequence of sperm whale pulses and encodes information on the inter-click intervals. The method in [22] augments the above approach in the context of a communication network. Based on killer whale pulses, the network establishes point-to-point communication links and assigns a unique communication address for each communication node. Yet, the repetitive nature of the communication may disclose it as being artificial. Another biomimicking approach is to use biological sounds to hide the communication signals, similar to steganography or watermarking. The method in [33] embeds the communication signal as a watermark inside a humpback whale call using a Discrete Cosine Transform (DCT). The watermark bits are set at positions known to the receiver. A similar concept is proposed in [21], where Direct Sequence Spread Spectrum (DSSS) replaces DCT to embed hidden bits inside the whale's call. In both schemes, the whale's call is used for channel estimation allowing the communication signal to be transmitted at a much lower SNR. Alternatively, the methods reported in [34] and [35] modulate the data over a dolphin's whistle contour. In [34], the data is modulated as Minimum Shift Keying (MSK). Utilizing the smoothness of MSK, the hidden signal does not stand out from the modulating whistles, and different than the methods in [21] and [33], transmission occur at high power. A similar approach is proposed in [35], where the data is modulated as a series of segmented chirp signals. While biomimicking UAC techniques are becoming increasingly popular, there are only a limited number of techniques that are geared for biomimicking interception.

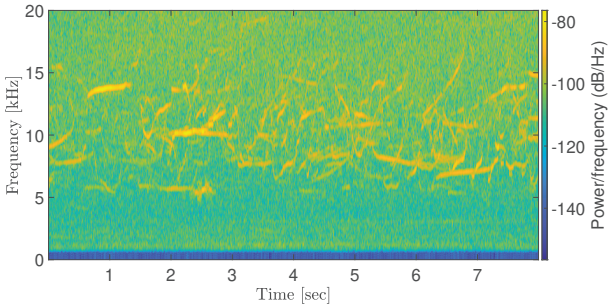
An interception technique geared for BBOKMS, shows that the slope distribution of a TF curve can effectively discriminate between BBOKMSs and real bio-acoustic signals [24]. The authors use the Gaussian kernel probability density estimation to model the slope distribution and extract characteristic parameters for recognition purposes. Another possible technique is reputation interception, where the detector tracks a belief that the signal detected is indeed of marine vocalization. An example for this is given in [36], where a trust model is designed based on changes in the environmental conditional as identified through the channel state information. Another option for interception is to locate the source using an acoustic array and compare its motion pattern to the expected one of e.g., a dolphin. However, this requires a complex structure from the interceptor and the location privacy of the source can be protected by push-back technologies such as proposed in [37] by synthetic multipath. More general biomimicking interceptors assume a measurable difference between the acoustic signals emitted by a man-made acoustic projector and the signals emitted by a marine animal such as a dolphin or a whale. In [25], the authors base their interceptor on the identification of distortions in the received biomimicking signal resulting from preprocessing (up or down sampling, filtering, etc.) and from large peak-to-peak hardware limitations. The distortions in the replayed signal are measured in reference to the original signal and evaluated by using joint entropy metrics. Similarly, the authors in [26] use the constraint of the damping factor of an acoustic projector to classify between real bio-acoustic and biomimicking signals based on the phase content of the signal. Although the above interceptors achieve accurate differentiation between biomimicking and real bio-acoustic signals, the results are limited to short ranges or to a non-reverberating sea environment, where channel distortions are small. However, in a complex environment, non-separable multipath may render phase differences and high entropy content also in the biomimicking received signal, making it difficult to identify between channel and hardware distortions.

To the best of the authors' knowledge, there is still a lack of contributions that apply deep neural networks (e.g., CNNs) to the detection of biomimicking sounds in practical and realistic settings. The main issue, in this respect is to collect a sufficiently large training dataset, that should possibly include both biomimicking signals and natural signals. While the former depends on experimental efforts to be carried out in a sufficient many locations and under sufficiently many sea conditions, the latter is much more difficult to obtain. In fact, high-quality, low-noise, annotated data on mammal vocalizations remains a scarce resource. Moreover, collecting such data entails expensive campaigns at sea, with long-term equipment deployment and human-in-the-loop post-processing by specialized personnel. Therefore, approaches based on signal processing and acoustic channel analysis like the one we propose in this paper, still constitute a more viable solution.

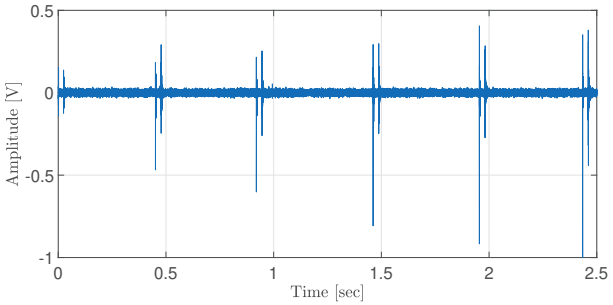
III. System model

Our system model considers two possible scenarios. The first includes a biological source that emits signals, e.g., a dolphin vocalizing whistle signals. The second consists of a stationary or mobile transmitter that communicates with its receiver through biomimicking signals. In both cases, the transmission is carried out at a high power, in order to achieve a high SNR, such that the interceptor is able to detect the signals. The goal of interception is binary classification: for a given sequence of biological-like signals, to distinguish between a real or a biomimicking source.

The interceptor is assumed to have a single hydrophone, such that directionality evaluation is not possible. Otherwise,



(a) A spectrogram containing multiple dolphin whistles.



(b) A time sequence of data including clicks of Sperm whales.

Fig. 1: Example for bio-acoustic vocalizations to demonstrate the potential in biomimiking. Dolphin whistles can be used to modulate communication by signal's type, and the delay between sperm whale clicks can be used for information encoding.

using an array of receivers, a direct calculation of the source's mobility is possible, which obviates the need to evaluate this pattern from the waveform as we do in our scheme. We argue that this is a practical assumption, as it is challenging to setup a receiver array. In particular, it requires a rigid platform to hold the hydrophones at fixed distances that are separated enough to meet the long wavelength requirement, and an accurate gyrocompass to compensate for any orientation changes of the array. Further, it is challenging to design an adaptive beamforming that can handle the wideband characteristics of the expected signal.

A. Main Assumptions

We design our interception scheme to be robust to the marine environment and insensitive to the emitted signals. Consequently, considering the many possible ways of biomimicking and the variation in dolphin whistles, we avoid assuming knowledge regarding the structure of the channel or the analytical form of the emitted signals and therefor refer from using learning schemes. We assume that both real and biomimicking communication involves a sequence of time-separated signals. This can be a group of dolphin whistles, as shown in Fig. 1a, or a sequence of whale clicks, as shown in Fig. 1b. In both cases, multiple bio-acoustic signals are observed, sometimes in the form of a *cocktail party*. In the following, our assumed signal for evaluation contains a series of suspected signals allowing stability analysis of a sequence of CIRs. The signals are assumed to be well received at the interceptor. As a worst-case scenario, we assume that the hardware of the biomimicking source is ideal and cannot be used to separate the two signal sources.

Our key assumption is that the dynamics of the biomimicking source are very different from those of a real marine animal. In particular, we assume that a biomimicking source is moving in a coherent fashion such that the CIRs of individual symbols within the signal sequence are statistically dependent. Contrary, a real bio-acoustic source does not move coherently, and the relation between the CIRs of adjacent vocalizations is small. This latter separation between the mobility of sources stems from the biomimicking platform limitations. In particular, a biomimicking source is constrained by the mechanical noises that can occur in the midst of biomimicking transmission. As a result, it is reasonable to assume that noisy systems such as propulsion and steering are shut down during communication so as to improve communications as well as not to reveal the source of the transmission, thereby reducing the maneuverability of the biomimicking source.

Further, to better decode the communication signals, the biomimicking source may prefer to move slowly. This is different for marine mammals that tend to communicate while in motion [38]. The dynamics of these animals, e.g., dolphins, seals or whales, are much faster and much less stable than a man-made system. For example, dolphins can swim at a speed of about 15 km/h and change their orientation rapidly while communicating [39], which is presumably well beyond the capabilities of a man-made system when it emits a communication signal. Therefore, we expect the CIRs for the biomimicking source-interceptor link to change much slower in time compared to the vocalizing marine animal-interceptor link. For CIR representation, we assume that the signal and its biomimicking version are both instantaneously narrowband chirplet signals and that the delay between the channel's taps are mostly outside the main lube of the instantaneous frequency, such that the multipath structure can be separated from the spectrogram of the received signal. For dolphin whistles, we support this assumption by the model of the signal [38].

B. Preliminaries

An underwater acoustic CIR refers to the channel's response to a sound wave traversing through the water medium. The channel's response is characterized by propagation and absorption loss as well as reflections from sea boundaries and volume scatterers. This response is a function of the frequency range, the sound speed profile, the local bathymetry, the wave height and the composition of the seabed, to name just a few parameters [40]. The CIR between an acoustic source and a receiver is commonly modeled by the tap delay line

$$h(t) = \sum_{m=1}^{M} A_m \delta(t - \tau_m) \,, \tag{1}$$

where $\{A_m, \tau_m;\ m = 1, ..., M\}$ are respectively the amplitudes and time-delays of the channel's taps. The values of A_m and τ_m vary with the change in position of the transmitter and receiver, and their dynamics are a key factor in determining the temporal stability of the CIR.

In our work, we use correlation and sample entropy as similarity metrics for CIRs. Sample entropy, initially proposed in [41], serves as a statistical measure to assess the irregularity or complexity of time series data. More than other entropy measures, sample entropy is less sensitive to the data length and provides robust estimates even with limited data. This makes sample entropy particularly advantageous for analyzing datasets with inconsistent characteristics. Formally, the sample entropy (SampEn) for a time series x with the length N and the tolerance threshold r is defined by

$$\mathrm{SampEn}(m, r, N) = -\log\left(\frac{A(m+1, r)}{A(m, r)}\right), \qquad (2)$$

where $A(m, r)$ represents the conditional probability that two sequences of length m having a maximum difference of r remain similar when their length increases to $m+1$. Parameters m and r signify the pattern length and tolerance respectively, influencing the sensitivity of the entropy calculation. This metric allows the characterization of the signal regularity while tolerating variations in data length and noise levels, making sample entropy a versatile tool for evaluating the complexity of signals measured with sensors. While other entropy or statistical measures can also be used to measure stability, we concentrate on sample entropy as it is able to capture how random a CIR compared to another CIR based on their temporal characteristics. The calculation is performed without available prior knowledge of the CIR characteristics, which suits our case since the environmental conditions for the interceptor are unknown.

The random walk (RW) mobility model is employed to obtain CIR instances between a mobile source and a fixed receiver location. The source initiates its RW at the 3D coordinate $\mathbf{P}^{(0)} = [X_0, Y_0, Z_0]^T$ and moves at a constant speed V throughout the duration of the RW. At each time step i of duration $T^{(i)}$, the source travels in the horizontal heading $\Phi^{(i)}$ and depth $D^{(i)}$. The RW parameters, including the initial position, speed, time steps, headings, and depths, are specified for each motion model. The source's 3D coordinate after the i-th step is given by

$$\mathbf{P}^{(i)} = \mathbf{P}^{(i-1)} + VT^{(i)} \begin{bmatrix} \cos(\phi^{(i)}) \\ \sin(\phi^{(i)}) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ D^{(i)} - D^{(i-1)} \end{bmatrix}. \qquad (3)$$

IV. Details of our biomimicking interceptor

A. Key Idea

Especially in an area with a complex bathymetric structure and depth-varying sound speed, the CIR is expected to change along with the variations in the transmitter's location. Changes are also expected due to the time-varying nature of the underwater CIR. However, for a relatively short interval between two signal emissions, there is still an expected dependency between the two corresponding CIRs. For a fast moving source, changes in the CIR will result with reduced statistical relation between the CIRs of symbols transmitted over a short interval — a packet. Thus, we target our interceptor to quantify similarities within a group of CIR's and determine by their stability if the source's motion dynamics are typical of a biomimicking transmitting platform or belong to an authentic animal. Due to the challenge of estimating the CIR without prior information about the structure of the transmitted signals, we avoid estimating directly the CIR and instead calculate the *propagation signature* (PS). The PS is a representation of the CIR, which, under the assumption of instantaneously narrowband signal as in chirplet signals, captures the multipath structure of the CIR but neglects the phase and absolute magnitude of the channel's taps. As described in details below, it is calculated by generating a template of the signal from its dominant TF contour of the signal, which is assumed to correspond to the direct path, and convolving it with the received signal. In the following, we refer to the PS rather than the CIR.

One option to determine if the signal originates from a real biological source or a biomimicking one is to use machine learning. The main challenge with training a machine learning model for this task is the variability of the channel impulse response which is hard to predict if no prior information is given on the bathymetry and sound speed profile. The key limiting factor is the lack of recordings to train for biomimicking signals, as we aim to detect all types of biomimicking communications — more so ones we never encountered before. Instead, we take a difference approach and perform the source type identification by statistical evaluation.

Our processing pipeline block diagram is shown in Fig. 2. We start with PS evaluation by following the TF contour of the detected signal and using it as an approximate template signal. This rather heuristic solution is chosen since the signal is expected at high SNR. For a set of N such evaluated PSs, we then calculate the sample entropy and cross-correlation as similarity metrics. This results with a point cloud of N^2 for each similarity metric, which is then used to calculate two stability measures: 1) the Kullback-Leibler divergence (KLD), D_{KL}, to quantify how similar the Gaussian distribution is to the distribution of the similarity metrics, and 2) a clustering solution to calculate the minimal number of states, $K_{\min}$, required to efficiently segment the point cloud of similarity measures. The former is used to measure the *randomness* in the signal, and the latter to test how diverse the similarity metrics are. The two stability measures are finally converged to make a decision about the type of source.

The approach procedure requires statistical data analysis, whose accuracy depends on the number of acquired symbols. While other types of entropy require distribution evaluation, the sample entropy one does not since it is communicated directly from the sequences compared. However, a distribution estimation is required for the calculation of the KLD to explore the match between the PS similarities and the Gaussian distribution. To this end, we calculate the distribution of the similarities numerically using an histogram, while that of the Gaussian distribution is calculated analytically using the method of moments. Since even for 10 symbols, the number of similarity comparison is 91 (the similarities can be non-symetric), we argue that the statistics suffice for distribution representation.

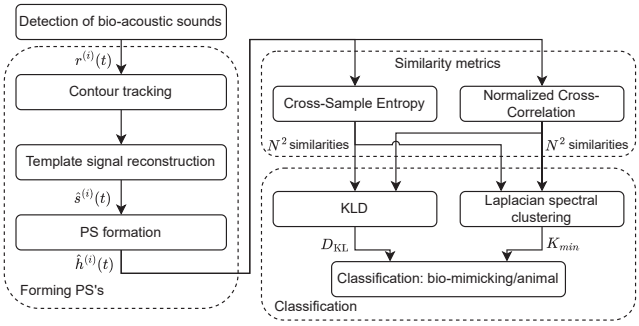


Fig. 2: Block diagram for our biomimicking interception scheme.

B. Classification into Real and Biomimicking Signals

We offer two complementary techniques to test stability between PSs. The first is based on the statistical evaluation of a point cloud of the similarity measure for a set of acquired PSs to find out whether their distribution follows a pattern that could indicate a biomimicking source. Given a sufficiently large enough number of PS pairs, we anticipate that the similarity metrics for i.i.d PSs, which presumably represents a real biological source, will follow a Gaussian distribution, while a different, more concentrated distribution is expected if the PSs are related. This observation can be translated into a null hypothesis test for the distribution of the PSs cross-correlation output. Our second stability test is based on clustering to measure the variation within the similarity metrics. This is motivated by the observation that the variation in the similarities of i.i.d PSs is lower than for unrelated PSs. As a result, changes in the minimal number of segments required to cluster the similarity data are expected between related and unrelated PSs.

For similarity evaluation, we consider two complementary metrics: 1) the normalized cross-correlation between pairs of PSs to quantify the similarities between the PSs structures, and 2) the cross-sample entropy to evaluate the statistical relationship between pairs of PSs. Formally, let $\rho_{i,j}^{\text{correlation}}$ and $\rho_{i,j}^{\text{entropy}}$ be the cross-correlation and the cross-sample entropy metric between the ith and the jth PSs, respectively. For a sequence of N estimated PSs, we obtain point clouds $\mathcal{S}^{\text{correlation}}$ of N^2 and $\mathcal{S}^{\text{entropy}}$ correlation and entropy similarity metrics, respectively. In the following, we describe both the distribution-based approach and the clustering-based one.

1) Probability-based biomimicking Interception: To evaluate the dissimilarity between a Gaussian distribution and the distribution of the similarity metrics, we employ the KLD metric. Denoted as $D_{\text{KL}}(P\|Q)$, the KLD quantifies the difference or 'distance' between two probability distributions $\mathbf{P}$ and $\mathbf{Q}$. Let, $\mathbf{P}$ represent the empirical distribution derived from the similarity set $\mathcal{S}$, and $\mathbf{Q}$ the Gaussian distribution fitted to $\mathcal{S}$. The latter distribution fitting is based on the method of moments to evaluate the mean and variance of the elements in $\mathcal{S}$. The KLD is expressed by

$$D_{\text{KL}}(P\|Q) = \sum_{x\in\mathcal{S}} P(x)\log\left(\frac{P(x)}{Q(x)}\right) \ . \tag{4}$$

2) Stability Testing by Clustering: Denote $a_{i,j}^{corr} = 1-\rho_{i,j}$ and $a_{i,j}^{entropy} = \rho_{i,j}$ as the cross-correlation and the entropy-based similarity metrics between PSs $h^{(i)}$ and $h^{(j)}$, respectively. Also denote an $N\times N$ weighted similarity matrix $\mathbf{W}$ whose (i,j)th element is

$$w_{i,j} = e^{-\alpha a_{i,j}} \ , \tag{5}$$

where α is a user-defined sparsification parameter that "stretches" the similarity histogram. The weight of each PS is accumulated in the diagonal matrix $\mathbf{D}$ whose ith element is defined by

$$d_i = \sum_{j=1}^{N} w_{i,j} \ , \tag{6}$$

Matrices $\mathbf{W}$ and $\mathbf{D}$ construct a symmetric graph Laplacian formalized by

$$\mathbf{L} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{D}-\mathbf{W})\mathbf{D}^{-\frac{1}{2}} \ , \tag{7}$$

Matrix $\mathbf{L}$ encapsulates the connectivity of the graph and represents the latent structures within the dataset. In other words, $L(i,j)$ represents how unique is the similarity of two data points (i,j) compared to the overall similarities in the dataset. Hence, clustering $\mathbf{L}$ depends on the stability of the PS, and the minimum number of feasible clusters, $K_{\min}$, is a measure of this stability. In particular, a small $K_{\min}$ reflects a low stable dataset.

To obtain $K_{\min}$, let $N\times 1$ binary vector $\mathbf{b}_k$ denote the clustering assignment such that $b_k[i]=1$ if the ith element in $S^{\text{correlation}}$ or S^{entropy} is assigned to the kth cluster or not. An element cannot be assigned to more than one cluster and each cluster must contain at least one element. The minimum number of clusters is determined by

$$\begin{aligned} K_{\min} &= \arg\min_{K} \frac{1}{K}\sum_{k=1}^{K}\mathbf{b}_k^T\mathbf{L}\mathbf{b}_k + \epsilon K \\ s.t.& \\ &b_i \in [0,1], \\ &\mathbf{b}_k^T\mathbf{b}_l = 0, \ \ \forall k\neq l, \\ &\sum_{i=1}^{N}\mathbf{b}_k[i] \geq 1, \ \ \forall k\in[1,K], \end{aligned} \tag{8}$$

where ϵK is a penalty to restrict a too high number of clusters. An example for the calculation of the utility function of (8) is given in Fig. 3 for three distribution types randomly generated at uniform, replacing $a_{i,j}$ in (5) and for $N=24$. The figure shows error bars for the standard deviation of the calculation for 50 random realizations. We observe that the more spread the data is the smaller the cost, and like-wise the smaller $K_{\min}$ is. This supports use $K_{\min}$ as a measure for stability of a given dataset. Note that $K_{\min}$ complements the KLD in (4) with the advantage of quantifying also non-linear stabilities, and that the calculation is possible even for a small number of N that does not allow for an accurate distribution analysis.

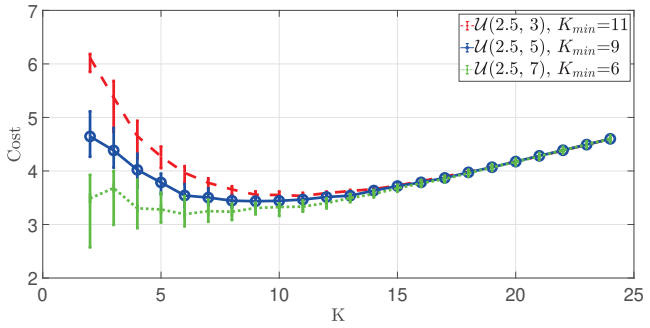


Fig. 3: Laplacian's cost function for each Uniform distribution parameters. The number of minimum clusters reduces as the sample range increases.

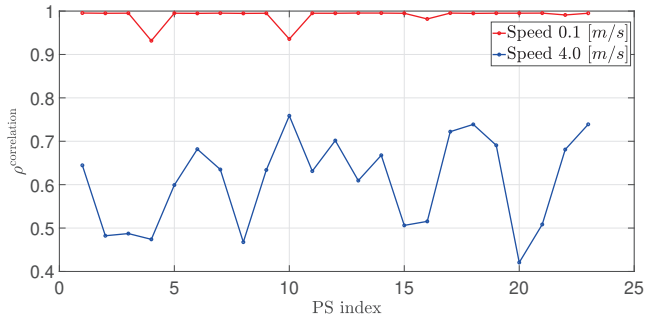


Fig. 4: Normalized cross-correlation between each pair of consecutive PSs, calculated for two source's dynamics.

C. Similarity Metrics for PS's pairs

Given two PSs, $h^{(i)}, h^{(j)}, i, j \in \{1, \ldots, N\}$, the cross-correlation metric is calculated by

$$\rho_{i,j}^{\text{correlation}} = \max_{l} \frac{\sum_{n} h^{(i)}[n] \cdot \overline{h^{(j)}[n-l]}}{\sqrt{\sum_{n} (h^{(i)})^2} \cdot \sqrt{\sum_{n} (h^{(j)})^2}}, \quad (9)$$

where n is the taps index defined over the overlapped indexes between the two PS[1]. An example for $\rho_{i,j}^{\text{correlation}}$ between pairs of consecutive PSs of a slow-moving source and a fast-moving one is shown in Fig. 4 (the simulation setup is described below). As expected, the former yields higher cross-correlation.

Our second similarity measure is sample entropy, which quantifies the regularity of PSs. The Cross-Sample Entropy (CSE) is defined by (2), where $A(m, r)(v||u)$ reflects the conditional probability between $h^{(i)}$ and $h^{(j)}$. An example of the CSE measure for the same simulated source as used for the cross-correlation example is given in Fig. 5. Upon comparing the two instances, it becomes evident that the faster target demonstrates higher CSE values in contrast to the slower target. Furthermore, it is noteworthy that the CSE values for the faster target exhibit considerable variability and fluctuations.

[1]If $h^{(i)}$ and $h^{(j)}$ are not of same size, the PSs can be zero-padded.

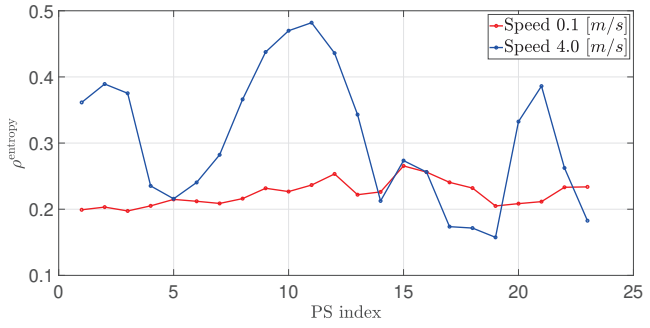


Fig. 5: Cross-sample entropy between each pair of consecutive PSs, calculated for two source's dynamics.

D. Forming PS's

A key enabling technique for our interceptor is the ability to generate the PS from a detected bio-acoustic-like signal of unknown structure. Here, we offer a heuristic solution motivated by the expected high SNR of the detected signal. Let

$$r^{(i)}(t) = h^{(i)}(t) * s^{(i)}(t) + n(t), \ \forall i \in \{1, 2, \ldots, U\} \quad (10)$$

be a model for the detected signal, where $s^{(i)}(t)$ is the ith bio-acoustic signal convolved with the ith CIR, $h^{(i)}(t)$, and $n(t)$ is the ambient noise. We represent $\hat{h}^{(i)}(t)$ by the matched filter between $r^{(i)}(t)$ and an estimate for $s^{(i)}(t)$, $\hat{s}^{(i)}(t)$. The latter is obtained by following the contour of the spectrogram of $r^{(i)}(t)$ and using an inverse transformation to obtain the time domain signal while filtering out noise. Clearly, this operation neglects the phase of $h^{(i)}(t)$ as well as its absolute magnitude. However, under the above assumption of a signal with instantaneous narrowband signal and a sparse CIR, the signal's template created from the TF contour would correspond to the direct path of the CIR. Correlating $r^i(t)$ with this template may thus capture the multipath structure of the CIR.

Obtaining the contour of the signal from the spectrogram is similar to a tracking problem. The tracker can be as simple as taking the frequency bin with the maximum value for each time interval [42], but this solution can be sensitive to noise transients from e.g., distant shipping cavitation. Instead, we perform the tracking by considering the frequency bins as 'states' and the time bins as 'observations'. The signal's spectrogram is normalized such that we refer to the (i, j)th bin of the spectrogram as a form of likelihood that a signal exists in that bin. As bio-acoustic signal is continuous, we expect that if a signal exists in the (i, j)th bin it will more likely exist in the band $(i \pm \delta, j + 1),\ 0 \leq \delta \leq \Delta$ than outside this band, where Δ is a bound on the signal's frequency changing rate. This is formalized in a hidden Markov model, where the emissions are the likelihood (the spectrogram normalized bins' value) and the transition probability is assumed uniform within the allowed frequency band and zero outside it. The solution for the tracking is obtained by the Viterbi algorithm (VA), which iteratively evaluates potential paths between the states while dynamically updating the most probable sequence. In the context of a bio-acoustic signal, this solution allows tracking the contour while avoiding confusion due to noise transients.

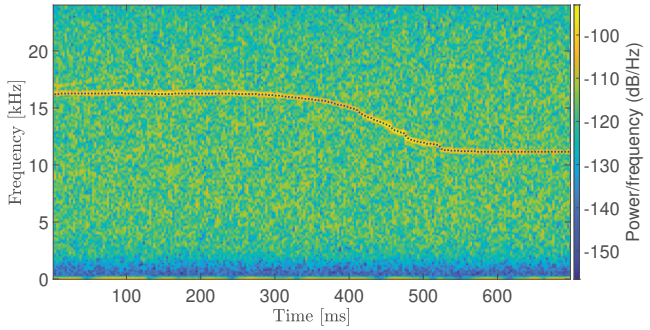


Fig. 6: A spectrogram of a real dolphin Whistle. Contour tracked by the VA is marked in red line. The instantaneous narrowband of the signal allows channel restoration.

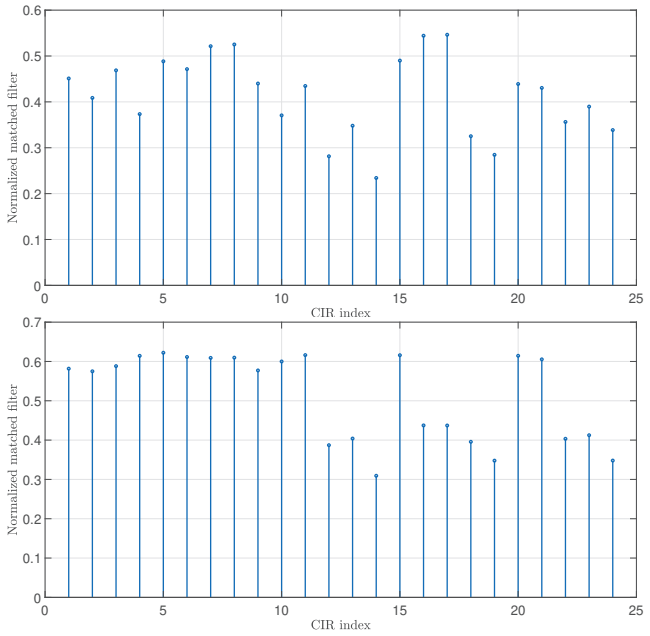


Fig. 7: Matched filter between the true simulated channel and the (Top) estimated channel obtained by using a clean template signal, (Bottom) PS obtained by using the reconstructed template signal from the spectrogram's contour.

An example for the VA contour tracking of a dolphin's whistle can be seen in Fig. 6.

We note that the above estimate of $\hat{s}^{(i)}(t)$ does not claim to provide a representation for the true CIR. This is because of the averaging within the process of making the spectrogram image, that largely distorts the phase. Instead, we argue that the PS of the CIR, holds enough information to quantify the similarities between pairs of CIRs. An example of this is shown in Fig. 7, where the normalized matched filter between a simulated sequence of channels is matched filter with two instances: one is a CIR estimate created by the clean analytic signal, and the other is a PS created from the above process of obtaining the spectrogram's contour. Simulated channel was formed by the Bellhop propagation model [43] for a diverse bathymetry. We observe that, in this case, the PS formed from the signal's contour actually provides a better match.

E. Attack strategy

A possible strategy to avoid an interception of biomimicking signals by our method is to convolve the emitted biomimicking signals with random channels prior to their transmission. At the interceptor side, this will yield an estimate of a stable PS (assuming a slowly drifting source) that is convolved with a random filter. As a result, the cross-correlation between the sequence of estimated PSs will decrease. As for the interceptor, we assume that the attacker doesn't possess any knowledge regarding the interceptor's location or the CIR between them. We note that this kind of attack is similar to the transmission of orthogonal biomimicking signals, for example in the biomimicking modulation scheme in [30]. Since we avoid any assumption regarding the correlation between the biomimicking signals, we find this defense technique weak. This is also shown by our results below.

Another possible defense against our interceptor can be the emission of signals whose structure is hard to construct. This technique attacks the ability to obtain the PS by following the contour of the signal. Possible signals are very short dolphin whistles or clicks that cover the entire examined spectrum while forming a steep spectrum curve. This, however, limits the ability of the source to transmit signals resembling bioacoustic vocalization as discussed in [25] and relies perhaps too much on the limitations of the interceptor to represent the signal.

F. Complexity Analysis

The computational complexity of the proposed interceptor is $\mathcal{O}(N^2L^2)$ for processing a group of N received signals, where each signal has a length of L samples. This calculation is explained in the following. The Viterbi algorithm is bounded by $\mathcal{O}(NN_TN_F^2)$, where N_T is the number of 'observations' and N_F is the number of 'states' [44]. The PS evaluation includes a matched filter, whose complexity is $\mathcal{O}(NL\log(L))$ using the Fast Fourier Transform (FFT) implementation [45]. In the stability calculations, the cross-sample entropy is more computationally intensive than the normalized cross-correlation. Its computational complexity is $\mathcal{O}(N^2L^2(m+(m+1)))$, where N^2 is the number of similarity calculations between all possible PS pairs and L^2 is the number of sub-sequences with a length of m whose difference is smaller than the tolerance r. Since m is small and constant, the complexity simplifies to $\mathcal{O}(N^2L^2)$. Finally, for the stability measurement the more time-consuming metric is Laplacian spectral clustering, which is $\mathcal{O}(N^3)$ [46]. Since $L >> \log(L)$ and $L >> N$, the tightest bound is given by $\mathcal{O}(N^2L^2)$.

V. Simulation

A. Simulation setup

Our simulation explores performance for various motion dynamics models for the biological source and the bio-mimic source. We consider dolphin whistles as the biological vocalization and target for biomimicking, due to their popularity in biomimicking protocols (e.g., [34] and [47]), and due to the wide distribution of dolphins. Two Monte-Carlo randomizations of RWs are generated: one for a dolphin and one

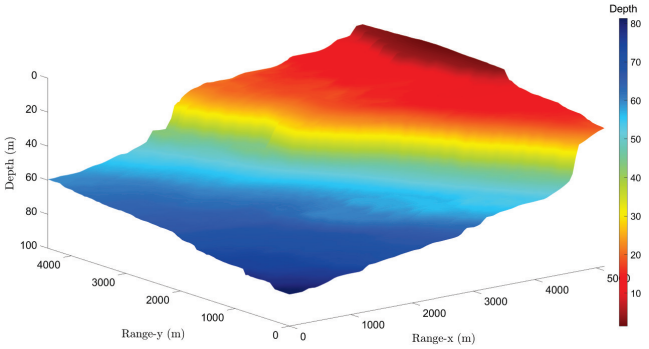


Fig. 8: San Diego bathymetric map.

for a biomimicking source. Each RW is characterized by a random duration, speed and heading. For the dolphin source, we randomize the duration by $\mathcal{N}(3,4)$ [seconds], the heading by $\mathcal{N}(0,(\pi/9)^2)$ [radians] and depth by $\mathcal{N}(3,2)$ [m]. These values are randomized for each step of the RW. The speed is randomized once for each RW by $\mathcal{U}(3.5,6.5)$ [m/s].These parameters fit the observation in [48] and [39]. For the biomimicking source, we consider a vessel performing a RW of duration similar to that of the dolphin, speed randomized by $\mathcal{U}(0.05,0.75)$ [m/s] and heading randomized by $\mathcal{U}(-\pi,\pi)$ [rad]. The depth of the biomimicking source remains stable at 2 [m]. Here, randomization is performed once for each RW. The result is a seemingly chaotic motion for the dolphin and a smooth trajectory for the vessel. The initial X,Y and Z 3D position coordinates of dolphin or vessel are drawn uniformly at random by $X\sim\mathcal{U}(0.5,2.5)$, $Y\sim\mathcal{U}(0.5,1.5)$ [km] and $Z\sim\mathcal{U}(1,10)$ [m]. The interceptor position is fixed at $(3500,2000,10)$ [m]. Each RW trajectory is obtained by applying the given parameters in (3)

For each RW step, we generate a CIR that reflects the channel between the current position of the source and the location of the interceptor. The modelling is performed by the Acoustics Toolbox (Ocean Acoustic Library (OALIB)) Bellhop3d model [43]. Due to its complex structure, we use the bathymetric map in Fig. 8, which corresponds to the coastal area north of San Diego [49].

For biomimicking we consider two types of signals: synthetic whistles, which closely resemble real dolphin whistles, and the playback of pre-recorded whistle sequences. The former is produced by applying the Remez filter over a time-varying frequency function. A similar procedure was applied in [26]. Examples for such signals are given in Fig. 9. For real dolphin signal, we use the DCLDE2011 database.

We identified three biomimicking interceptors benchmarks: [24], [25] and [26]. The first is geared to the specific case of binary orthogonal keying modulated signals, while the latter two are more general and rely on limitations of the projector hardware. We thus compare our method to the latter two methods. The approach in [25] employs entropy measures over the detected signal and in the following we refer to it as the *Entropy* benchmark. The method in [26] classifies the signals based on the randomness of their phase and in the following we refer to it as the *Phase* benchmark. Both benchmarks were implemented for performance comparison.

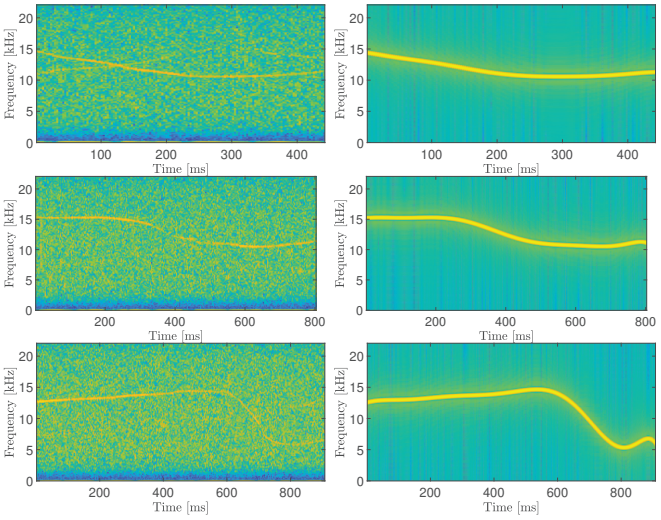


Fig. 9: Example for (Right) synthetic dolphin whistles generated from (Left) real dolphin whistles.

B. Simulation results

We measure the performance of our interceptor by the precision-recall trade-off. We define precision by

$$\text{precision}=\frac{N_{TP}}{N_{TP}+N_{FP}}\,, \tag{11}$$

where N_{TP} is the number of PSs groups that are correctly classified as biomimicking and N_{FP} is the number of PSs groups that are incorrectly classified as biomimicking. The recall is defined by

$$\text{recall}=\frac{N_{TP}}{N_{TP}+N_{FN}}\,, \tag{12}$$

where N_{FN} is the number of PSs groups that are misclassified as a real dolphin, although their true classification is a biomimicking. Several approaches have been explored in the existing literature for intercepting biomimicking UAC. However, none of these methods aligns with our specific scenario, characterized by an unknown biomimicking modulation structure and the use of an ideal transducer by the source. Therefor, we refrain from presenting a benchmark, as the existing interception methods do not cater to our setup.

Since changes of the PS are a function of both the source's speed and of the non-homogeneous of bathymetry and sound speed profile, we introduce a measure for the average PS change to quantify the *speed* of PS change:

$$\rho_{\text{avg}}=\frac{1}{N_{\text{rw}}}\sum_{i=1}^{N_{\text{rw}}}\left(\frac{1}{N(i)-1}\sum_{j=1}^{N(i)-1}\rho_{j,j+1}^{\text{correlation}}(i)\right)\,, \tag{13}$$

where $N(i)$ is the number of PS estimates within the i-th PS-group, $\rho_{j,j+1}^{\text{correlation}}(i)$ is the correlation between the j-th and $j+1$-th PSs in the i-th PS-group, and N_{rw} is the number of PS-groups.

For each stability measure, three configurations are tested: 1) stability measure computed individually over each of the two similarity metrics, 2) the union of the stability measure for each similarity metric, and 3) the intersection of the stability measure for the similarity metrics. A total of eight precision-

recall configurations are tested: four to each similarity metric. For the chosen configuration, four additional precision-recall curves were then computed, to explore the combination of the two stability measures.

We start by exploring the sensitivity of the interceptor to PS changes. Results for three values of ρ_{avg} for 100 RW randomizations using a synthetic dolphin whistle are presented in Fig. 10. Comparing the results for the different ρ_{avg} values, we note that performance hardly changes for relatively small changes in the PS of up to $\rho_{\text{avg}} \leq 0.07$. Significant decrease in performance is observed for $\rho_{\text{avg}} = 0.93$, which in our setup corresponds to a speed of 0.2 m/s. We observe that the most effective configurations are detection derived from KLD and the union of detections derived from KLD and $K_{\min}$ computations over the $\mathcal{S}^{\text{correlation}}$ similarity metric. In the following, we continue with these configurations. From Fig. 10 we observe some instabilities in the performance. This is because multiple methods are compared including: 1) correlation vs. sample entorpy, and 2) KLD vs. the $K_{\min}$ clustering. We use these results to choose the best ones.

The results for distinguishing a fast-moving source, "dolphin," from a slower source, "biomimicking" (BM), are shown in Fig. 11 using the stability measure KLD($\mathcal{S}^{correlation}$). To provide better control over the similarities of the channels, the "dolphin" signal is a simulated synthetic whistle running through a "fast" changing channel with $\rho_{\text{avg}} = 0.66$. The "BM Synthetic" signal is the same synthetic whistle as the "dolphin" signal after convolving with a "slowly" changing channel with $\rho_{\text{avg}} = 0.93$. The "BM Replay" is the recording of a real dolphin whistle after passing through a "slowly" changing channel of $\rho_{\text{avg}} = 0.95$. The "BM Countermeasure" is similar to the "BM Synthetic" signal, but its synthetic signal, $s^{(i)}(t)$, is first passed through a "fast" changing channel, $h^{(i)}_{\text{fast}}(t)$ with $\rho_{\text{avg}} = 0.66$ before it passes through the 'slow' channel $h^{(i)}_{\text{slow}}(t)$, for which $\rho_{\text{avg}} = 0.93$. That is, the resulting signal of the "BM Countermeasure" at the interceptor is

$$\left(s^{(i)}(t) * h^{(i)}_{\text{fast}}(t)\right) * h^{(i)}_{\text{slow}}(t) + n(t) \ , \tag{14}$$

where $n(t)$ is an additive Gaussian noise set so that the SNR is equal to that of the "BM Replay" signal, and the same applies to the noise added to the other "BM" and "dolphin" signals. We note the high precision and recall for the interception of the "BM Synthetic" and "BM Replay" signals, with a slight advantage for the interception of the latter. This is due to the small differences in the ρ_{avg} of the two signals. As expected, since it's effective ρ_{avg} is lower due to the use of $h^{(i)}_{\text{fast}}(t)$ in (14), the interception of the "BM Countermeasure" signal is more difficult than that of the "BM Synthetic" and "BM Replay" signals, and the interception performance decrease, arguably, still to a good interception level. The results in Fig. 11 are close for the playback and synthetic signals, while some degradation is observed for the countermeasure approach. However, the latter comes at the cost of distorting the transmitted signal, which may impact the communications performance.

The results in Fig. 11 are compared with the two benchmark methods *Entropy* and *Phase*. We find that the benchmarks achieve better results in separating the synthetic biomimicking signal from the Dolphin signal. The benchmarks struggle with the playback signal and fail with the countermeasure signal. We suspect that this is due to the complex underwater acoustic channel model, which makes it difficult to distinguish between projector- and channel-based distortions. For all three biomimicking signals, the Phase approach achieves better recall and precision rates than the entropy approach. This is because the latter is searching for changes in the waveform, which are highly distorted by the channel, while the former relies on identifying the "phase randomness," which is less channel-dependent. However, compared to the two benchmark methods, our proposed method achieves much better and more robust results. This is because our method does not depend directly on the channel, but rather on channel similarity measures.

VI. Field experiment

A. Experimental setup

As a proof of concept for the capabilities of our developed method to tell apart drifting and intentional motion, we performed a field experiment on the Garda lake, in northern Italy, in February 2023. We relied on EvoLogics SC2R software-defined modems working in the 18-34 kHz band, which were used as transmitters, and on one autonomous RTSys recorder equipped with a Colmar GP1280M boradband hydrophone, shown in Fig. 12. We used one fishing boat of size about 5 m×8 m. The boat left the eastern shore of the lake from the port of Lazise (45.5053°N, 10.7319°E).

As a first step, we deployed a receiver about 530 m west of the shore. The receiver was an autonomous RTSys recorder equipped with a Co.l.mar. GP1280M broadband low-noise hydrophone, set on continuous recording. The hydrophone was suspended at a depth of 5 m from the surface, attached to a buoy, and allowed to drift. The current slowly dragged the recorder towards the south-east, until the hydrophone reached a shallower area near the shore.

The experiment then proceeded in phases where we transmitted chirp signals while drifting or moving intentionally towards or away from the receiver. For the transmission, we employed the software-defined version of the EvoLogics S2CR mid-frequency modem, which operates in the 18-34 kHz band, and can transmit custom signals at a sampling rate of 250 ksamples/s. Each round of transmissions included 24 consecutive chirps spanning the whole 18-34 kHz bandwidth, and having a duration of 10 ms. A guard interval of 0.2 sec was set between subsequent chirp symbol. The depth of the transmitting transducer was 20 m for every transmission. We chose to transmit chirp signals rather than actual biological sounds such as dolphin whistles, as the latter are only available as noisy recordings. Instead, chirps worked as a clean source signal representing a form of synthetic whistle. The employment of short-duration chirps is advantageous for collecting a substantial dataset within the constraints of an experimental time frame. However, short signals have a lower processing gain, and thus we consider the use of short chirps in the experiment as a lower-bound scenario for the extraction of the PS.

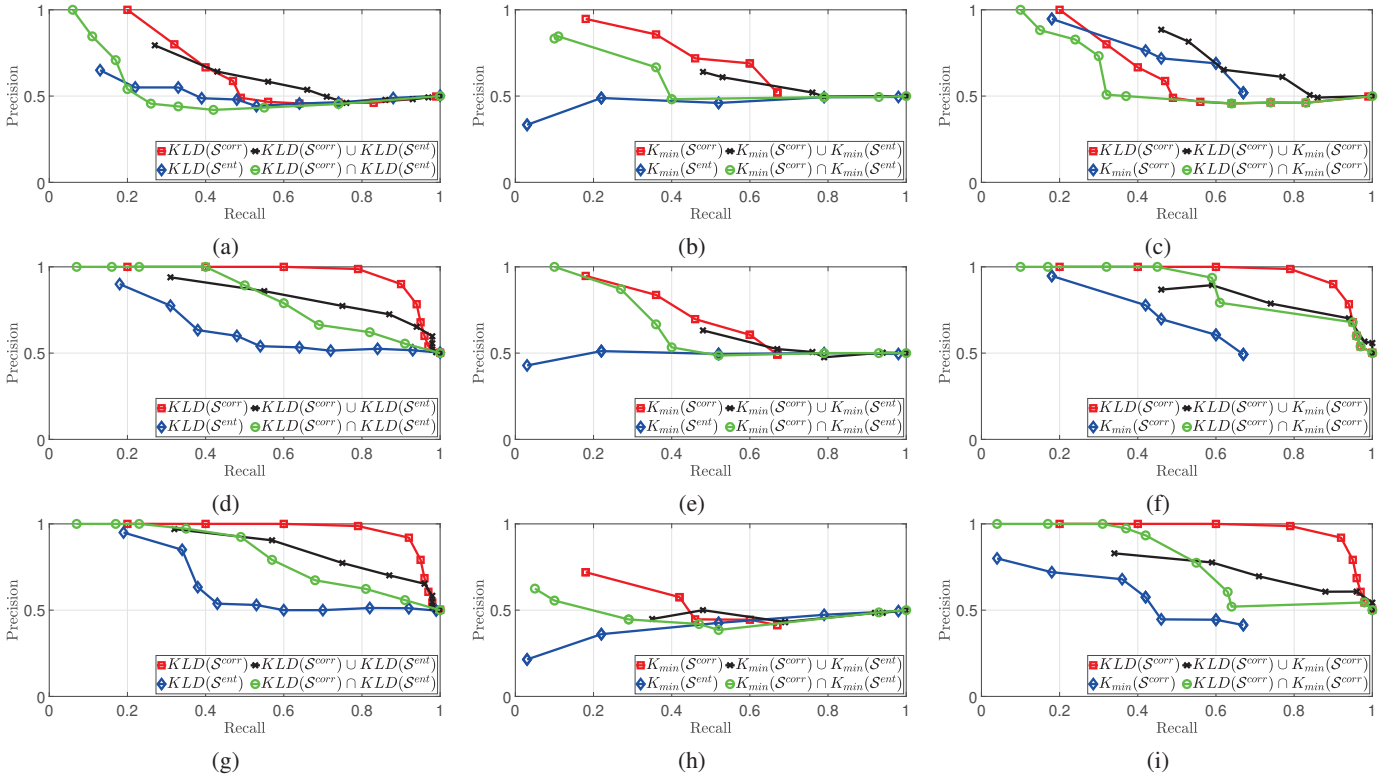


Fig. 10: Precision-recall curves illustrate (a, d, g) KLD and (b, e, h) K_{min} computations over combinations of $\mathcal{S}^{\text{correlation}}$ and $\mathcal{S}^{\text{entropy}}$, and (c, f, i) the best pair combinations derived from these metrics. Results presented for: (a, b, c) $\rho_{\text{avg}} = 0.93$, (d, e, f) $\rho_{\text{avg}} = 0.77$, and (g, h, i) $\rho_{\text{avg}} = 0.70$.

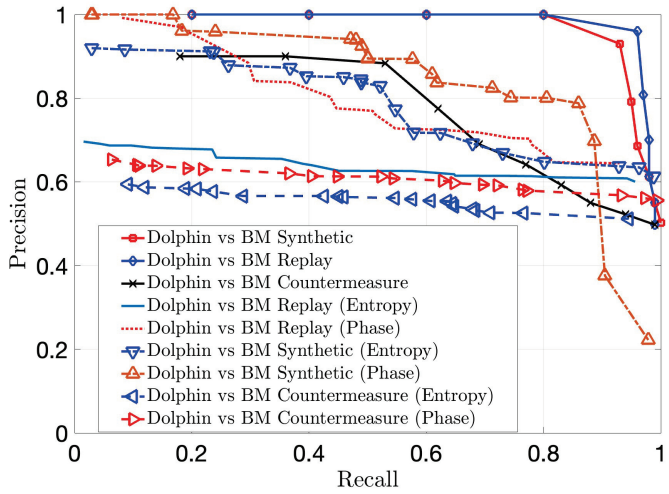


Fig. 11: Precision-recall tradeoff for discriminating a fast moving target ("dolphin") from a slow moving one (biomimicking - "BM") using the KLD($\mathcal{S}^{correlation}$) stability measure. ρ_{avg} for the "dolphin", "BM Synthetic", "BM Replay" and "BM Countermeasure" signals are 0.66, 0.93, 0.95, and a combination of 0.66 and 0.93, respectively. Results also shown for the Entropy and Phase benchmarks.

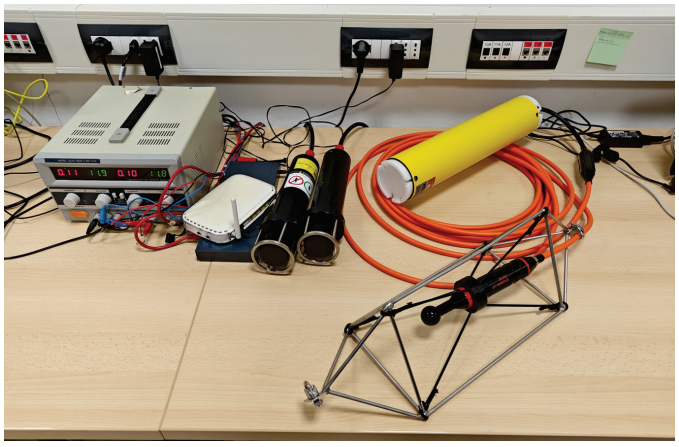


Fig. 12: Equipment used in the experiment. The recorder is the yellow casing with the hydrophone attached. The transmitter is one of the modems shown to the right of the recorder.

We performed the first chirp train transmission by positioning the boat 270 m north-east of the receiver (water depth: 50 m) and letting it drift. We repeated the procedure by repositioning the boat at different locations: 580 m west of the receiver (water depth 50 m), then 1430 m north-west of the receiver (water depth 60 m), and 520 m north-west of the receiver (water depth 35 m), respectively.

During the second phase of the experiment, we performed transmissions while the boat was dragging the transceiver at different speeds. The transceiver was attached to a weight of about 10 kg to keep it as vertical as possible in the water, although some lifting still occurred at the highest speeds. We

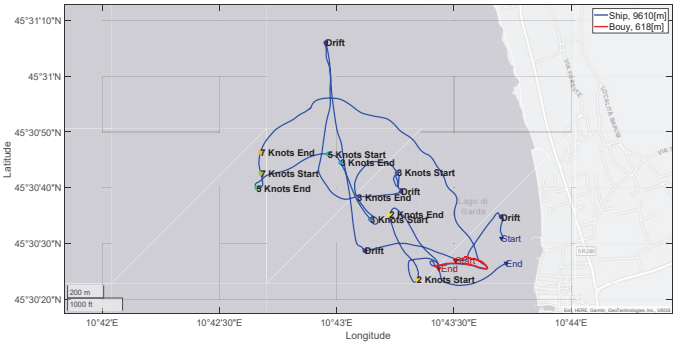


Fig. 13: Boat and buoy tracks with start and stop chirps transmissions specified.

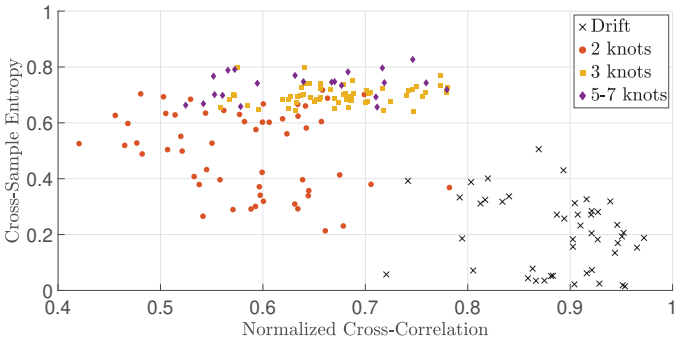


Fig. 14: Garda Lake scatter plot for PS features for four different boat speeds.

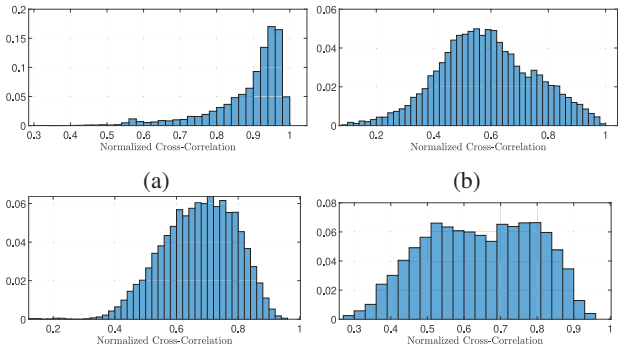


Fig. 15: Normalized histograms of $\mathcal{S}^{\text{correlation}}$ values for boat travel states: (a) drift, (b) 2 knots, (c) 3 knots and (d) 5-7 knots.

proceeded by positioning the boat 750 m north of the receiver and moving towards it at 2 knots until a distance of 500 m. We then moved away from the receiver towards the north-west at 3 knots; we then moved westwards at a speed of 5 knots for 450 m, and then north-eastwards at 7 knots for 500 m. For this experiment, the weight detached from the transceiver, and boat motion lifted it very close to the surface. At this point, the transmitter was 1.3 km away from the receiver, and we performed one final transmission sequence by letting the boat drift.

A total of $46, 55, 61$ and 22 chirp-trains were measured for drifting boat (< 1 knots), a boat moving at 2 knots, 3 knots and 5 and 7 knots. For KLD calculation, we estimated the distributions by means of histograms of 16 bins, and for the cross-sample entropy analysis we used a pattern length of $m = 2$ and a tolerance of $r = 0.2\sigma_{avg}$, where σ_{avg} is defined as the average of the two PSs standard deviations as $\sigma_{avg} = (\sigma_{\hat{h}^{(i)}} + \sigma_{\hat{h}^{(j)}})/2$.

We share the dataset of the experiment including the recorded audio signals, nodes locations and the channel representations, PS, in [50].

B. Experimental Result

A scatter plot for the cross-sample entropy and normalized cross-correlation showing the results for the tested mobility patterns is given in Fig. 14. Observing the results we see a good separation between results for drift and faster motion, while the separation between cases of speed ≥ 2 knots is much less obvious. For the case of drifting boat, the high normalized cross-correlation values and the low cross-sampled entropy values reflect on the statistical relation between the PSs. To comment on the resemblance to a Gaussian distribution of the similarity metrics, in Fig. 15 we show the normalized histograms of the normalized cross-correlation. We observe that the faster the vessel moves, the more *Gaussian-like* the distribution is, while for drifting motion the distribution is non-symmetric heavy-tail. Similar result is obtained for the cross-sample entropy similarity metric as shown in Fig. 16.

The results of the stability measures, $K_{\min}$, with parameters $K = 24$ and $\epsilon = 0.15$ are shown in Fig. 17 for the different motion patterns explored. Error bars show the spread of the results obtained per boat's speed. Performance for the cross-correlation similarity metric (see Fig. 17a) show clear separation between the drift motion and the mobile scenarios. The results show a larger $K_{\min}$ for the drifting motion. This result is consistent with the discussion in section IV-B2, and in particular Fig. 3, where we argue that similarity in the PS should lead to a large $K_{\min}$. A separation is not observed for the sample entropy metric (see Fig. 17b). This is due to the properties of the cross-sample entropy as its output provide information about the PSs fluctuations differ than normalized cross-correlation where its output is more suited to function as a distance measure. As a result, the clustering cost function is high as the negative penalty is low for most clustering

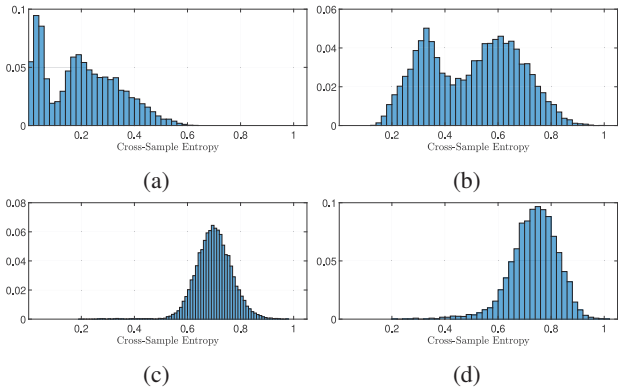


Fig. 16: Normalized histograms of $\mathcal{S}^{\text{entropy}}$ values for boat travel states: (a) drift, (b) 2 knots, (c) 3 knots and (d) 5-7 knots.

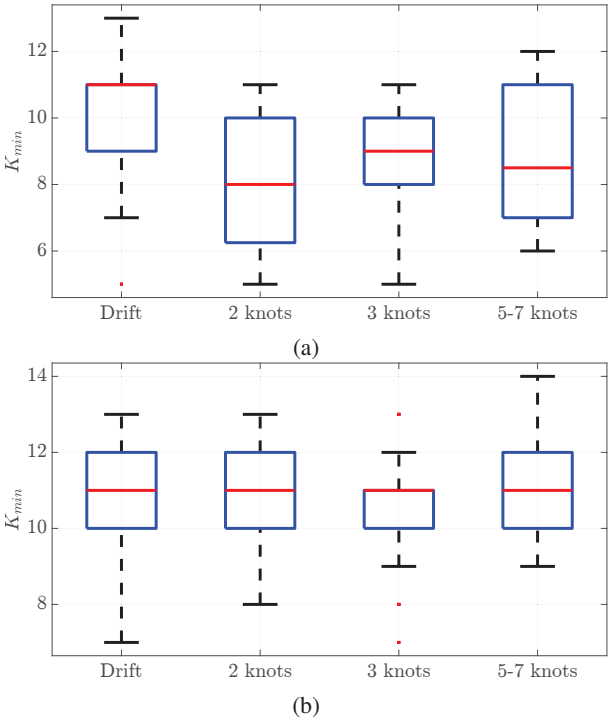


Fig. 17: Laplacian spectral clustering K_{min} for four different speeds calculated for (a) $\mathcal{S}^{\text{correlation}}$ and (b) $\mathcal{S}^{\text{entropy}}$. The line inside of each box is the sample median, the box shows the interquartile range (IQR), while the whiskers represent the interval spanned by samples that are closer to the median than 1.5 times the IQR.

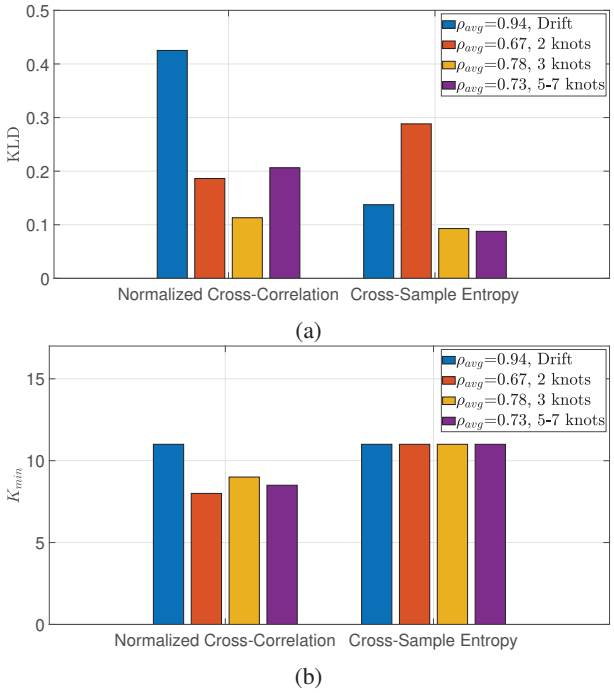


Fig. 18: (a) mean KLD and (b) median $K_{\min}$.

assignments. Also here a clear difference between the drift and the mobile scenarios is shown for the cross-correlation metric, but not for the sample entropy one.

We analyze performance for the experiment over the drift and the speeds by using a KLD and $K_{\min}$ pairs taken from $\rho_{\text{avg}} = 0.77$ in the simulation, chosen according to highest F1 score. Precision and recall pairs are shown for $\mathcal{S}^{\text{correlation}}$ in TABLE. I. Results for the KLD stability measure are shown in Fig. 18. Also here a clear difference between the drift and the mobile scenarios is shown for the cross-correlation metric, but not for the sample entropy one.

In our study, we explore the optimal amalgamation of similarity point clouds for each of the two stability measurements at different boat speeds. Analogous to our simulation outcomes shown in TABLE. I and Fig. 18, it is evident that $\mathcal{S}^{\text{correlation}}$ consistently provides the most favorable trade-offs between precision and recall for both similarity measurements. Upon identifying the best combinations, our findings indicate that at 5-7 knots, the combination of KLD($\mathcal{S}^{correlation}$) $\cap$ K_{min}($\mathcal{S}^{correlation}$) yields the most optimal trade-off between precision and recall. However, at 2 and 3 knots, KLD($\mathcal{S}^{correlation}$) stands out as the superior choice to achieve the best performance in this context.

C. Discussion

The outcome of the presented analysis shows that our interceptor is able to separate between dolphin and biomimicking sources. The separation is characterized with a favorable tradeoff between precision and recall regardless of the type of signal used in the biomimicking communication, synthetic or replay and regardless of the bathymetry of the environment. Moreover, separation is also available for countermeasures such as CIR insertion attack, when a channel with high ρ_{avg} is used. Despite its robustness in overcoming these primary challenges, it remains susceptible to limitations associated with the underwater medium. For instance, challenges arise from the low SNR in bio-acoustic signals often caused by a vast distance between the source and the receiver, man-made machine noise from ships, construction sites or oil rigs, or even natural disturbances such as adverse weather conditions. In addition, an environment with a particular bathymetry, such as an underwater mountain, can obstruct the direct path between source and receiver and cause distortions of the received signals. More specifically, there may be significant changes in sequential PSs even though the location of the source has changed only slightly. In contrast, our interceptor is vulnerable to the possibility of encountering simple PSs consisting of only one arrival, rendering our similarity features ineffective. In this study, a dolphin was chosen as a model for a real marine animal. The simulation is based on its general movement dynamics and bio-acoustics properties. However, when it comes to biomimicking communication using single wave bio-acoustic signals, there may be limitations in obtaining a rich and informative PS. In addition, the use of bio-acoustic signals originating from naturally slow-moving marine animals could lead to ambiguous separation in the interceptor's output. This research operates under the assumption that the bio-acoustic signal detector is ideal. However, in cases where the detector is not ideal, perhaps because it isolates only part of the signal or includes noise in the detected signal, we argue

TABLE I: Experiment (precision, recall) pairs for drift and different speeds

Drift vs speed	KLD($\mathcal{S}^{\text{corr}}$)	$K_{\min}(\mathcal{S}^{\text{corr}})$	KLD($\mathcal{S}^{\text{corr}}$) $\cup$ $K_{\min}(\mathcal{S}^{\text{corr}})$	KLD($\mathcal{S}^{\text{corr}}$) $\cap$ $K_{\min}(\mathcal{S}^{\text{corr}})$
2 Knots	(0.61, 0.80)	(0.47, 0.96)	(0.52, 1.00)	(0.52, 0.98)
3 Knots	(0.84, 0.80)	(0.43, 0.96)	(0.46, 1.00)	(0.64, 0.98)
5-7 Knots	(0.74, 0.80)	(0.67, 0.96)	(0.70, 1.00)	(0.73, 0.98)

that the resulting PS adequately represents the propagation features as long as the TF contour of the detected signal is widely distributed across the frequency domain and the VA can accurately track its contour. While challenges such as low SNR or distorted signals are inherent to all UAC, overcoming limitations in intercepting click-based or non-dolphin-based biomimicking communications, as well as those using interception countermeasure methods requires further research and exploration.

Future work will extend this work to expand the modeling of motion dynamics to a broader range of marine animals whose bio-acoustic sounds are used for biomimicking purposes, including whales and seals. As part of this extension, we will perform a comparative analysis with the established model of dolphin movement dynamics. In addition, we will investigate methods for obtaining a PS specifically tailored to short bursty bio-acoustic sounds, such as echolocation clicks and transients of snapping shrimps. Furthermore, we plan to explore additional PS similarity measurement metrics tailored to both complex and non-complex bathymetric environments. This research aims to improve our understanding of the group dynamics of PSs and thereby increase the clarity and precision of our interceptor's separation results.

VII. Conclusions

This paper presents a novel biomimicking interception scheme for underwater acoustic communications, offering a comprehensive methodology for classifying real and biomimicking signals. The approach leverages stability metrics such as cross-correlation and CSE, and introduces a clustering solution to determine source mobility. Through numerical simulations and the analysis of results from a field experiment, the effectiveness of the interceptor is demonstrated, showing a favorable trade-off between precision and recall for the classification of biomimicking and genuine signals. This work contributes to the advancement of covert communication techniques in underwater acoustic environments, and lays the foundation for further exploration of interception methods in this domain.

Acknowledgment

The authors extend their gratitude to Davide Eccher for his assistance during the execution of the experiments.

References

[1] T. Mishali and R. Diamant, "Interception of bio-mimicked underwater acoustic communications," in *Proc. MTS/IEEE OCEANS*. IEEE, 2023, pp. 1–4.

[2] M. Chaudhary, N. Goyal, A. Benslimane, L. K. Awasthi, A. Alwadain, and A. Singh, "Underwater wireless sensor networks: Enabling technologies for node deployment and data collection challenges," *IEEE Internet Things J.*, vol. 10, no. 4, pp. 3500–3524, 2023.

[3] M. C. Domingo, "An overview of the Internet of underwater things," *J. Network and Comput. Appl.*, vol. 35, no. 6, pp. 1879–1890, 2012.

[4] X. Deng, Y. Jiang, L. T. Yang, L. Yi, J. Chen, Y. Liu, and X. Li, "Learning-automata-based confident information coverage barriers for smart ocean Internet of things," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9919–9929, 2020.

[5] M. Han, J. Duan, S. Khairy, and L. X. Cai, "Enabling sustainable underwater IoT networks with energy harvesting: A decentralized reinforcement learning approach," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9953–9964, 2020.

[6] M. Alimadadi, M. Stojanovic, and P. Closas, "Object tracking in random access networks: A large-scale design," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9784–9792, 2020.

[7] Y. Li, L. Liu, W. Yu, Y. Wang, and X. Guan, "Noncooperative mobile target tracking using multiple AUVs in anchor-free environments," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9819–9833, 2020.

[8] D. Zhang, I. N'Doye, T. Ballal, T. Y. Al-Naffouri, M.-S. Alouini, and T.-M. Laleg-Kirati, "Localization and tracking control using hybrid acoustic–optical communication for autonomous underwater vehicles," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 10 048–10 060, 2020.

[9] A. G. Yisa, T. Dargahi, S. Belguith, and M. Hammoudeh, "Security challenges of Internet of underwater things: A systematic literature review," *Trans. Emerging Telecom. Tech.*, vol. 32, no. 3, p. e4203, 2021.

[10] A. Gerodimos, L. Maglaras, M. A. Ferrag, N. Ayres, and I. Kantzavelou, "IoT: Communication protocols and security threats," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 1–13, 2023.

[11] C. Hu, Y. Pu, F. Yang, R. Zhao, A. Alrawais, and T. Xiang, "Secure and efficient data collection and storage of IoT in smart ocean," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9980–9994, 2020.

[12] B. Lin, L. Zhao, H. A. Suraweera, T. H. Luan, D. Niyato, and D. T. Hoang, "Guest editorial special issue on Internet of things for smart ocean," *IEEE Internet Things J.*, vol. 7, no. 10, pp. 9675–9677, 2020.

[13] S. Jiang, "On securing underwater acoustic networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 729–752, 2019.

[14] R. Diamant, P. Casari, and S. Tomasin, "Cooperative authentication in underwater acoustic sensor networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 954–968, 2019.

[15] R. Diamant, S. Tomasin, F. Ardizzon, D. Eccher, and P. Casari, "Secret key generation from route propagation delays for underwater acoustic networks," *IEEE Trans. Inf. Forensics Security*, vol. 18, pp. 3318–3333, 2023.

[16] K. Pelekanakis, S. A. Yıldırım, G. Sklivanitis, R. Petroccia, J. Alves, and D. Pados, "Physical layer security against an informed eavesdropper in underwater acoustic channels: Feature extraction and quantization," in *Proc. UComms*, 2021, pp. 1–5.

[17] R. Diamant and L. Lampe, "Low probability of detection for underwater acoustic communication: A review," *IEEE Access*, vol. 6, pp. 19 099–19 112, 2018.

[18] T. Yang and W.-B. Yang, "Low probability of detection underwater acoustic communications for mobile platforms," in *Proc. MTS/IEEE OCEANS*, 2008, pp. 1–6.

[19] G. Qiao, M. Bilal, S. Liu, Z. Babar, and T. Ma, "Biologically inspired covert underwater acoustic communication–A review," *Physical Commun.*, vol. 30, pp. 107–114, 2018.

[20] M. Bilal, S. Liu, G. Qiao, L. Wan, and Y. Tao, "Bionic morse coding mimicking humpback whale song for covert underwater communication," *Applied Sciences*, vol. 10, no. 1, p. 186, 2019.

[21] S. Liu, G. Qiao, A. Ismail, B. Liu, and L. Zhang, "Covert underwater acoustic communication using whale noise masking on DSSS signal," in *Proc. MTS/IEEE OCEANS*, 2013, pp. 1–6.

[22] J. Jiajia, W. Xianquan, D. Fajie, F. Xiao, L. Chunyue, and S. Zhongbo, "A basic bio-inspired camouflage communication frame design and applications for secure underwater communication among military underwater platforms," *IEEE Access*, vol. 8, pp. 24 927–24 940, 2020.

[23] J. Jia-jia, W. Xian-quan, D. Fa-jie, F. Xiao, Y. Han, and H. Bo, "Bio-inspired steganography for secure underwater acoustic communications," *IEEE Commun. Mag.*, vol. 56, no. 10, pp. 156–162, 2018.

[24] Q. Yao, J. Jiang, G. Chen, Z. Li, Z. Yao, Y. Lu, X. Hou, X. Fu, and F. Duan, "Recognition method for underwater imitation whistle communication signals by slope distribution," *Applied Acoustics*, vol. 211, p. 109531, 2023.

[25] P. Casari, J. Neasham, G. Gubnitsky, D. Eccher, and R. Diamant, "Acoustic projectors make covert bioacoustic chirplet signals discoverable," *Scientific Reports*, vol. 13, 02 2023.

[26] I. Davidesco and R. Diamant, "Detection of dolphin whistle-like biomimicking signals by phase analysis," *IEEE Access*, vol. 10, pp. 36 868–36 876, 2022.

[27] H. Bailey and P. Thompson, "Quantitative analysis of bottlenose dolphin movement patterns and their relationship with foraging," *J. Animal Ecology*, vol. 75, no. 2, pp. 456–465, 2006.

[28] F. E. Fish and C. A. Hui, "Dolphin swimming–a review," *Mammal Review*, vol. 21, no. 4, pp. 181–195, 1991.

[29] R. P. Hodges, *Underwater acoustics: Analysis, design and performance of sonar*. John Wiley & Sons, 2011.

[30] A. ElMoslimany, M. Zhou, T. Duman, and A. Papandreou-Suppappola, "An underwater acoustic communication scheme exploiting biological sounds: A new communication scheme for underwater acoustic channels," *Wireless Commun. Mob. Comput.*, vol. 16, 10 2016.

[31] L. S. Sayigh, H. C. Esch, R. S. Wells, and V. M. Janik, "Facts about signature whistles of bottlenose dolphins, tursiops truncatus," *Animal Behaviour*, vol. 74, no. 6, pp. 1631–1642, 2007.

[32] J. R. Buck and P. L. Tyack, "A quantitative measure of similarity for tursiops truncatus signature whistles," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2497–2506, 1993.

[33] G. Qiao, M. Bilal, S. Liu, T. Ma, Y. Zhao, and B. Kong, "Symmetry oriented covert acoustic communication by mimicking humpback whale song," *Symmetry*, vol. 11, no. 6, p. 752, 2019.

[34] S. Liu, T. Ma, Q. Gang, and B. Kuang, "Bionic communication by dolphin whistle with continuous-phase based on MSK modulation," in *Proc. IEEE ICSPCC*, 2016, pp. 1–5.

[35] S. Liu, T. Ma, G. Qiao, L. Ma, and Y. Yin, "Biologically inspired covert underwater acoustic communication by mimicking dolphin whistles," *Applied Acoustics*, vol. 120, pp. 120–128, 05 2017.

[36] A. Signori, F. Campagnaro, I. Nissen, and M. Zorzi, "Channel-based trust model for security in underwater acoustic networks," *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20 479–20 491, 2022.

[37] H. Wang, G. Han, Y. Zhang, and L. Xie, "A push-based probabilistic method for source location privacy protection in underwater acoustic sensor networks," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 770–782, 2022.

[38] M. dos Santos, S. Louro, M. Couchinho, and C. Brito, "Whistles of bottlenose dolphins (*Tursiops truncatus*) in the Sado estuary, Portugal: Characteristics, production rates, and long-term contour stability," *Aquatic Mammals*, vol. 31, pp. 453–462, 12 2005.

[39] D. Houser, "A method for modeling marine mammal movement and behavior for environmental impact assessment," *IEEE J. Ocean. Eng.*, vol. 31, no. 1, pp. 76–81, 2006.

[40] R. Hodges, *Underwater Acoustics: Analysis, Design and Performance of Sonar*. Wiley, 2011.

[41] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American J. Physiol-Heart Circul. Physiol.*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[42] D. Gillespie, M. Caillat, J. Gordon, and P. White, "Automatic detection and classification of odontocete whistles," *J. Acoust. Soc. Am.*, vol. 134, no. 3, pp. 2427–2437, 09 2013.

[43] M. B. Porter, "Ocean acoustics library (oalib)," 2022, accessed on April 4, 2023. [Online]. Available: https://oalib-acoustics.org/models-and-software/acoustics-toolbox/

[44] T. Quach and M. Farooq, "Maximum likelihood track formation with the viterbi algorithm," in *Proc. IEEE CDC*, vol. 1, 1994, pp. 271–276.

[45] J.-C. Yoo and T. H. Han, "Fast normalized cross-correlation," *Circuits, systems and signal processing*, vol. 28, pp. 819–843, 2009.

[46] N. Tremblay and A. Loukas, "Approximating spectral clustering via sampling: a review," *Sampling Techniques for Supervised or Unsupervised Tasks*, pp. 129–183, 2020.

[47] S. Liu, G. Qiao, and A. Ismail, "Covert underwater acoustic communication using dolphin sounds," *J. Acoust. Soc. Am.*, vol. 133, pp. EL300–6, 04 2013.

[48] J. J. Rohr, F. E. Fish, and J. W. Gilpatrick Jr., "Maximum swim speeds of captive and free-ranging delphinids: Critical analysis of extraordinary performance," *Marine Mammal Science*, vol. 18, no. 1, pp. 1–19, 2002.

[49] D. Divins and D. Metzger, "NGDC coastal relief model," 2021. [Online]. Available: http://www.ngdc.noaa.gov/mgg/coastal/coastal.html

[50] T. Mishali, "Link to the raw audio data, estimated channels, and node locations," accessed: Apr. 24, 2024. [Online]. Available: https://drive.google.com/file/d/1vdLWXlHVNbq1j7toAtFIAHW1rocR-UA6/view