Robust channel comparison metrics for underwater physical layer authentication

Davide Eccher[1] and Paolo Casari[1]

[1]*DISI, University of Trento* and *CNIT*, Italy

Davide Eccher, DISI, Via Sommarive 9, 38123 Povo TN, Italy - davide.eccher-1@unitn.it.

Abstract: *Physical layer security enables authentication and privacy in communication systems by harvesting the randomness from wireless channel realizations. For this, it leverages unique channel features that legitimate parties (Alice and Bob) can reliably observe while the same features remain mostly secret to attackers (Eve). In underwater acoustic networks, the minimal correlation of acoustic channels in space provides a promising context for PLS. Our research focuses on implementing authentication via PLS using channel crafting. More specifically, we craft an artificial channel starting from a secret seed, and then precode the transmitted signal by convolving it with the crafted channel. The parameters of the channel impulse response (CIR), namely the delay and amplitude of each arrival, serve as the source of the seed for the crafting process. The receiver performs authentication by comparing the CIR extracted from the incoming signal with an expected CIR. Therefore, we need a mismatch metric that measures the matching of two channels, indicating that such channels have been generated by the two ends of the same link. Based on this metric, a decision algorithm can perform authentication. In this paper, we present the implementation of four mismatch metrics for CIRs, which account for differences that originate from measurement errors, node movement, noise, etc. These disturbances make the delay and amplitude values of the arrivals fluctuate, even to the point of disappearing. Such differences may lead to increased false alarms and missed detections, and therefore a convenient CIR mismatch metric should be robust against them. Specifically, we propose to reward strong channel similarities more than we penalize large differences.*

Keywords: *Underwater acoustic networks; physical layer security; authentication; simulation.*

1. INTRODUCTION

Security is a matter of increasing concern in the underwater communication field, especially with the growing adoption of underwater acoustic networks in such critical applications as surveillance, drone operations, or defense. Implementing security features in an underwater network is particularly challenging because of the high variability, low throughput, and long propagation delays of the acoustic channel. This makes the overhead introduced by typical key-exchange algorithms inconvenient in underwater networks. Lighter approaches have been explored in the literature, one of which is to exploit the information conveyed by the communication channel to implement security features. Some of these techniques involve generating secret keys, for example, by using the route propagation times [1], or by exploiting the channel characteristics [2] - [4]. These can be also used to perform target detection as in [5], and other approaches exploit position information for authentication, such as in [6] or in [7].

We focus on physical layer authentication in underwater acoustic networks, where the features of the estimated multipath components of the acoustic channel provide a secret shared between two legitimate parties. Acoustic multipath propagation has two desirable properties in this respect: reciprocity and very limited spatial correlation. For instance, Fig. 1 shows that multipath patterns as observed by two legitimate parties Alice and Bob the two ends of a simulated channel. The patterns match very well; in contrast, the the arrivals of the channel extracted by a third node (Eve) at a different location have considerably different amplitude and delay.

By leveraging matching multipath patterns, we generate a new artificial channel that is convolved with the signal sent to Alice. Alice then compares the received channel with the one she had crafted to perform the authentication decision. The channel crafted by Bob will be much more similar to the reference channel generated by Alice than any intruding transmission crafted by Eve, who will inevitably make mistakes when estimating multipath propagation between Alice and Bob. However, accepting Bob's transmission and rejecting Eve's requires a metric that measures channel mismatch every time a node receives a transmission.

A *maximum time-reversal resonating strength* metric is proposed in [8] for authentication. This score is computed for a received signal based on a dataset of previous communications, and was not originally developed to reflect quasi-similarities between bidirectional channels. In this paper, we propose four different channel mismatch metrics, and comment on their performance in a simulated scenario.

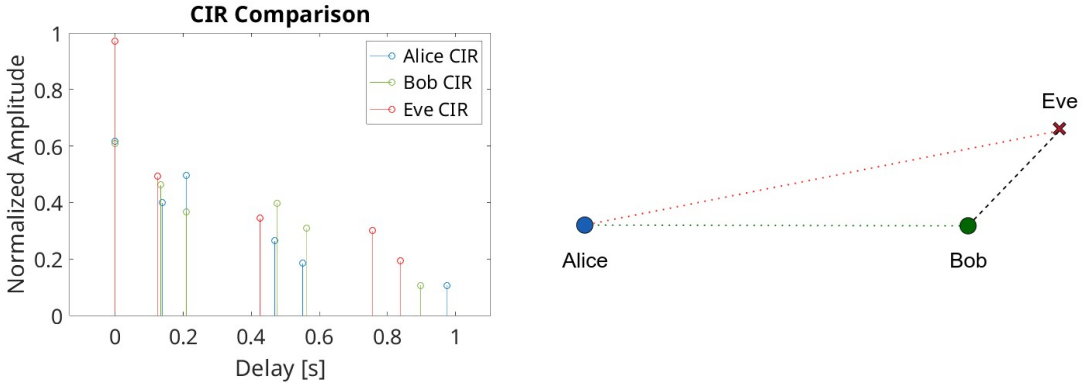

Figure 1: (Left) Comparison between CIRs of the AlicetoBob and BobtoAlice channels, against that of an attacker (Eve) arranged as in the topology on the right.

2. MISMATCH METRICS

The goal of a mismatch metric is to compare a received CIR against a reference one, and return a value that increases for increasing differences between the CIRs. Such differences may be due to the variability of underwater acoustic channels and can be observed from the different time of arrival and amplitude of their multipath components. Significant differences between CIRs may imply that an intruder is trying to impersonate a legitimate node. Hence, a mismatch metric should be able to detect these differences, while at the same time being robust against changes in the channels observed by Alice or Bob, which will also be subject to smaller but non-negligible fluctuations. Such fluctuation may include multipath components that even disappear temporarily due to, e.g., measurement errors, movement, noise, environmental changes, etc. An ideal metric should not penalize these changes too much over legitimate links in order to avoid false positives.

The metrics take as input both a reference CIR and a CIR to be compared against it. The CIRs are described in discrete terms, as arrays of delay-amplitude pairs, $\{(d_\mathrm{r}[i], a_\mathrm{r}[i])\}$, $i = 1, \ldots, N_\mathrm{r}$, and $\{(d_\mathrm{c}[j], a_\mathrm{c}[j])\}$, $j = 1, \ldots, N_\mathrm{c}$, where $d_\mathrm{r}[i]$ and $a_\mathrm{r}[i]$ represent the arrival delay and the complex amplitude of the ith multipath component of the reference CIR, $d_\mathrm{c}[j]$ and $a_\mathrm{c}[j]$ the same for the jth component of the CIR to be compared, whereas N_r and N_c are the number of components in the reference and compared CIR, respectively.

Before computing the mismatch metrics, we align the CIRs so that, for each channel, the first multipath arrival has delay $d[1] = 0$. Moreover, we join arrivals that are very close to each other. This kind of phenomenon is usually due to multiple reflections from extended surfaces or close-by scatterers. Small changes in the transmitting position cause small changes in the incidence angle that could create differences in the number of reflections and in the magnitude of constructive or destructive interference. Therefore, to mitigate this unpredictable behavior, we average the amplitudes of all components whose delays cluster within a range of 5 ms. After computing the mismatch metrics as described in the next subsections, we finally divide the obtained values by N_r. The above processing enables more comparable metric values across different scenarios or different time epochs. We now formalize the four metrics considered in this paper.

M_1: Full CIR comparison Here, we center a Gaussian kernel $g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}$ on each multipath component of the reference CIR, and use it to compute a mismatch between all pairs of multipath components on both CIRs, weighted with the reference amplitude. We exclude the first arrivals that, due to the initial alignment, match very well and may bias the metric. Formally, we compute

$$\widetilde{M_1} = \left(\sum_{i=2}^{N_\mathrm{r}} \sum_{j=2}^{N_\mathrm{c}} g\big(d_\mathrm{c}[j]; d_\mathrm{r}[i], \sigma_\mathrm{d}\big) \cdot g\big(a_\mathrm{c}[j]; a_\mathrm{r}[i], \sigma_\mathrm{a}\big) \cdot a_\mathrm{r}[i] \right)^{-1}, \tag{1}$$

where σ_a is the standard deviation of the amplitudes of the reference CIR's multipath components, and σ_d is computed such that a fraction Q of the area under the Gaussian kernel curve occurs within a delay interval proportional to the amplitude-weighed root mean-square delay difference

$$\hat{\tau} = \sqrt{\frac{\sum_{i=1}^{N_\mathrm{r}} (d_\mathrm{r}[i] - \bar{d}_\mathrm{r})^2 a_\mathrm{r}[i]}{\sum_{i=1}^{N_\mathrm{r}} a_\mathrm{r}[i]}}\,. \tag{2}$$

We divide this value by an additional parameter F to fine-tune the importance of delay differences over amplitude differences in the mismatch metric. Formally,

$$\sigma_{\mathrm{d}} = \frac{F\ \hat{\tau}}{\sqrt{2}\,\mathrm{erf}^{-1}(Q)}\ , \tag{3}$$

where $\mathrm{erf}(x) = \int_0^x \frac{2}{\sqrt{\pi}} \mathrm{e}^{-t^2}\mathrm{d}t$ is the Gaussian error function and $\bar{d}_{\mathrm{r}} = \frac{1}{N_{\mathrm{r}}}\sum_{i=1}^{N_{\mathrm{r}}} d_{\mathrm{r}}[i]$. This metric rewards multipath components with similar arrival delay more than components that are apart in time. Note that in (1) we take the reciprocal of the sum of the Gaussian kernel products, so that the mismatch decreases for CIRs with closely aligned arrivals.

For practical purposes, we impose a maximum mismatch value M_{max} to avoid that the metric approaches infinity for starkly different channels. Finally, we compute

$$M_1 = \widetilde{M_1}\ \cdot \max\left(1, N_{\mathrm{c}}/N_{\mathrm{r}}\right) \tag{4}$$

to balance channels with many more multipath components than the reference, which would otherwise be more likely to match reference multipath components and yield a low metric even if the channels are significantly different.

M_2: Optimal match This metric computes an optimal matching between each of the N_{r} components of the reference CIR and each of the components in the compared CIR, by minimizing the total delay difference between all matched components. The final value of the metric then depends on the delay and amplitude differences between matched pairs. Call $1 \leq C(i) \leq N_{\mathrm{c}}$ the index of the component of the compared CIR coupled to the ith component of the reference CIR, where the matching C can be any of the $N_{\mathrm{c}}!$ permutations of the indices from 1 to N_{c}. If one such permutation is $X(i)$, then

$$C = \arg\min_{X} \sum_{i=1}^{\min(N_{\mathrm{c}},N_{\mathrm{r}})} \left|d_{\mathrm{r}}[i] - d_{\mathrm{c}}[X(i)]\right| . \tag{5}$$

The optimal matching is computed via the Hungarian algorithm, which has complexity $O(N_{\mathrm{r}}^3)$. If $N_{\mathrm{c}} \neq N_{\mathrm{r}}$, only $\min(N_{\mathrm{c}}, N_{\mathrm{r}})$ matches are considered. However, if $N_{\mathrm{r}} > N_{\mathrm{c}}$, the remaining arrivals are matched with an artificial component having zero delay and amplitude: this prevents channels that are shorter than the reference from seeming accurate as a consequence of having a few matching components. After obtaining the optimal matching, we compute the metric as

$$M_2 = \sum_{i=1}^{N_{\mathrm{r}}} \left(g(0;0,\sigma_{\mathrm{d}}) - g(d_{\mathrm{c}}[C(i)];d_{\mathrm{r}}[i],\sigma_{\mathrm{d}})\right) \cdot \left(g(0;0,\sigma_{\mathrm{a}}) - g(a_{\mathrm{c}}[C(i)];a_{\mathrm{r}}[i],\sigma_{\mathrm{a}})\right) \cdot a_{\mathrm{r}}[i]\ . \tag{6}$$

Note that we use $g(0;0,\sigma) - g(x;\mu,\sigma)$, so that the higher the difference between the delay and amplitude of matched taps, the higher the metric value.

M_3: Interpolation difference This metric, denoted as M_3, oversamples the CIRs and then extends the duration of each arrival by an amount of time proportional to the normalized amplitude of the component (or until another arrival is reached). Such proportional amount is defined as the amplitude-weighed root mean-square delay difference $\hat{\tau}$ as defined in Eq. (2). Finally, the metric is computed by subtracting the two resulting CIRs and computing the average magnitude of the resulting difference. This approach is considered as a computationally inexpensive baseline.

M_4: Cumulative CIR difference Similar to the previous metric, this metric, denoted M_4, oversamples the CIRs and zero-pads the resulting CIR between subsequent multipath components. Then, it subtracts the interpolated CIRs and integrates the resulting differences. In this case, small errors in the delay values contribute slightly to the mean, whereas discrepancies in the amplitude values propagate to the cumulative sum and have a higher weight in the final mismatch value.

3. PERFORMANCE EVALUATION

We now set up a testing environment to evaluate the capability of the metrics to distinguish matching channels in an authentication context.

An attacker (Eve) tries to impersonate a legitimate node (Bob) at the receiver Alice, by sending a signal similar to the one that Bob would generate. We assume that Eve knows the propagation environment in detail, including the bathymetry, the sound speed profile, and Alice's location. Eve can also localize Bob with some error, and can simulate the acoustic channel between the estimated location of Bob and that of Alice. If the estimate of Bob's location is accurate, Eve will then obtain a channel similar to Bob's, which will likely yield a low mismatch metric. Conversely, we expect that in any sufficiently diverse environment, the discrepancies between Eve's estimated channel and Bob's real channel should increase as Eve's estimate of Bob's position is increasingly erroneous. Therefore, we evaluate the correlation between the computed mismatch metrics and the Eve's localization error for Bob's location. The objective of our evaluation is to identify a mismatch metric that yields low values on channels that should be quasi-reciprocal, while returning a higher value if channels are different, as a result of Eve localizing Bob at the wrong location. The latter aspect is crucial for the development of physical layer authentication schemes based on channel crafting (see also Section 1).

Simulation Environment To simulate the channels between two given locations, we use the Bellhop ray tracing software, with bathymetry and sound speed profile taken from the San Diego Bay Area. We analyze 10 scenarios with two different bathymetry shapes. For each scenario we place Alice and Bob at a random position, 500 to 1200 m apart from each other. We then generate about 1000 estimates of Bob's location in a circle of radius 700 meters around Bob. The depth of Alice, Bob, and all locations estimated by Eve have a random depth between 30 and 100 meters, as shown in Fig. 2. The parameters used by the mismatch metrics algorithms are set as $Q = 0.99$ and $F = 10$, see Eq. (3).

4. RESULTS

We proceed by presenting the results of our simulations. In Fig. 3, we show how the four proposed mismatch metrics change as a function of Eve's error on Bob's location estimate. The shaded area around the curves conveys the standard deviation of the metric as a function of distance. All metrics are normalized to the mismatch value for the true channel of Bob.

We observe that for distances close to 0, mismatch values approach 1, because Eve's estimated channel and Bob's actual channel almost coincide. As expected, the CIR measured in the Bob$\rightarrow$Alice direction is the channel having the lowest possible mismatch with the Alice$\rightarrow$Bob channel, compared to other CIRs measured at different locations around Bob's true one. The cases where the normalized mismatch is smaller than 1 mean that Eve generated a channel closer to the one generated by Alice than Bob's own channel. This is a rare occurrence, and is limited

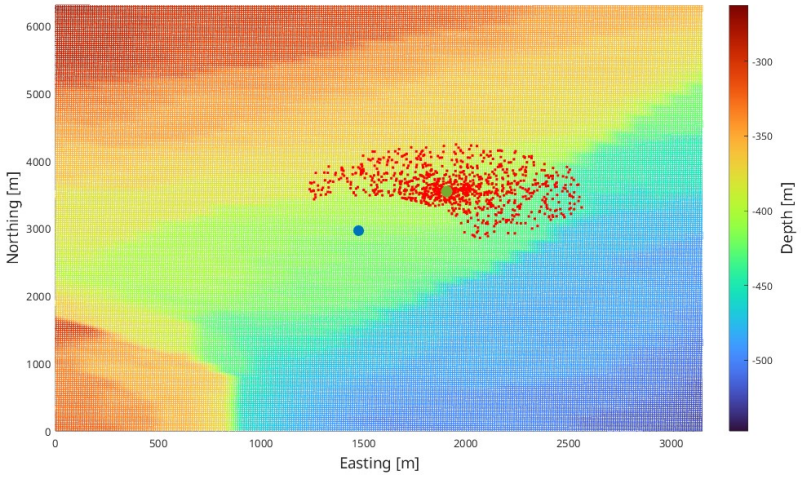

Figure 2: Simulation scenario and bathymetry. Alice: blue dot; Bob: green dot; Eve's estimates of Bob's location: red dots.

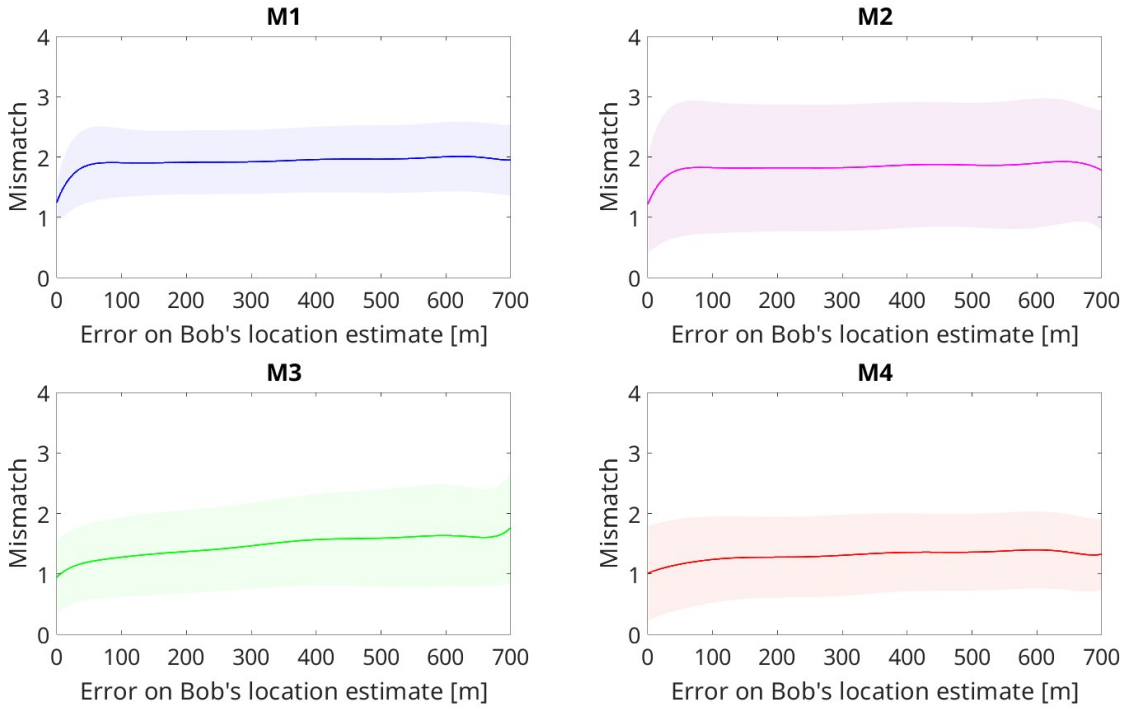

Figure 3: Channel mismatch as a function of distance between Bob and Eve's estimated position of Bob, normalized to the mismatch between Alice's and Bob's channels.

to those cases where Eve localizes Bob almost perfectly. From Fig. 3, we can also observe that spatial uncorrelation leads the mismatch metric to increase with increasing distance, until it tends to stabilize over an approximately constant value. This is because, beyond a certain distance, channels become uncorrelated. In our scenario, the uncorrelation distance is about 100 m from Bob.

Fig. 3 confirms that M_1 and M_2 yield similar results, and the best performance among the tested metrics. In particular, M_1 yields a smaller standard deviation in the results, as seen from the shaded areas around the solid curves. The baseline metrics M_3 and M_4 perform compara-

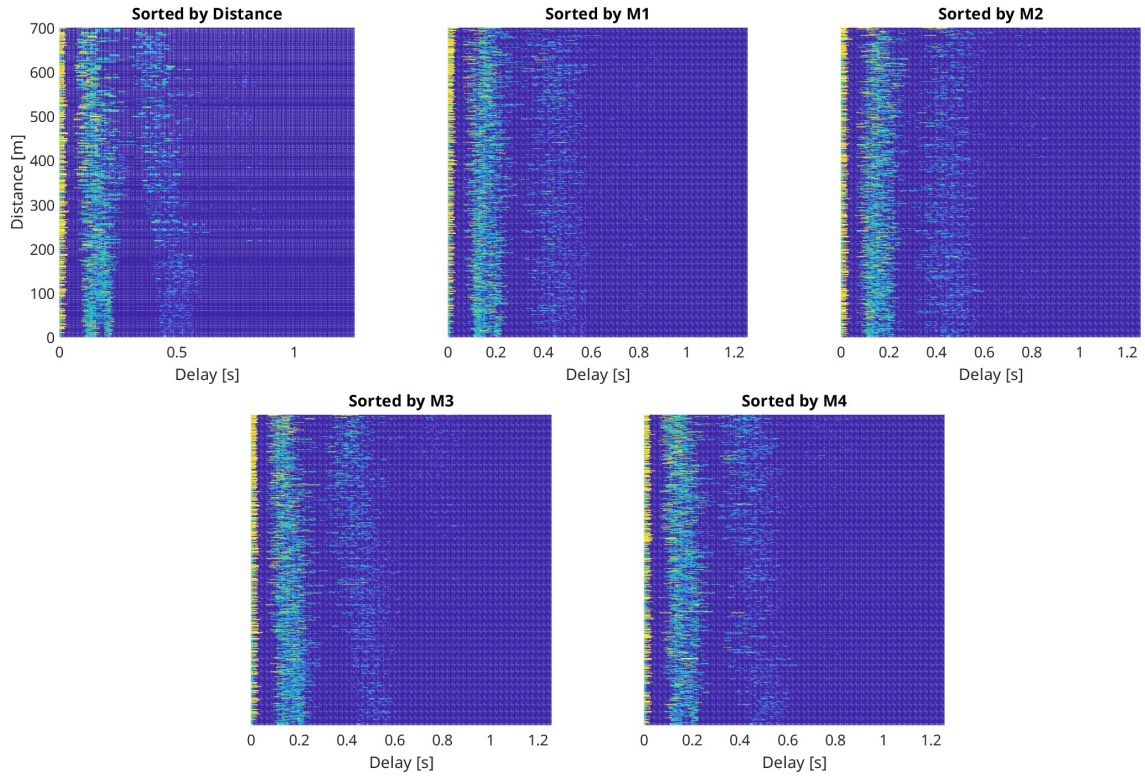

Figure 4: Channel impulse responses sorted by the actual distance between Eve's estimated position of Bob and Bob's true location (top-left), and by the four mismatch metrics (lowest mismatch at the bottom of each panel, highest mismatch at the top).

tively worse: M_4 barely differentiates between channels generated from locations close to Bob and far from it, whereas M_3 performs better only when Eve's error on Bob's position is large, hence channels are already starkly different.

Fig. 4 illustrates Eve's and Bob's CIRs, where the arrival delay of each multipath component is shown on the x-axis and amplitudes are color-coded. The channels are sorted along the y-axis in order of increasing localization error for Bob (first panel) and in order of increasing mismatch with respect to Alice's channel in the remaining panels.

The first graph shows that our assumption that closer nodes result in better-matching channels is confirmed by the simulations. In fact, we can observe that the first channels in the sorting share a significant number of components with similar delays and amplitudes. We also observe that, beyond a certain localization error, Eve's channels become almost completely uncorrelated from Bob's and Alice's, as confirmed by the metric values in Fig. 3.

Fig. 4 also provides a qualitative understanding of the behavior and implicit priorities of the different mismatch metrics. The best-performing metrics give greater weight to channels with similar arrival delays, and specifically to the first arrivals, which usually have higher amplitudes. In the case of metric M_2, the higher standard deviation of its output results in a more irregular channel sorting, where multipath arrivals exhibit a comparatively more scattered pattern. As M_3 gives the same weight to all arrivals, the channels with the lowest values of the metric are those where arrivals are positioned along the same delay span of the reference channel. However, as shown previously in Fig. 3, this weighing strategy does not necessarily imply a better capability to differentiate legitimate transmissions from impersonation attacks.

5. CONCLUSIONS

We presented four different metrics to measure the superimposability of underwater acoustic channel impulse responses. These metrics convey the mismatch between the channels observed at the two ends of the same link. Hence, a high value of the metric can serve as an indication of an impersonation attack, whereby an attacker is trying to "guess" a channel impulse response by localizing one of the two communicating parties, and by reproducing the channel that should be observed at the estimated location. The ultimate purpose of our work is to provide metrics that are robust when employed to make authentication decisions.

We have showed that the two proposed metrics M_1 and M_2 achieve better results than the baseline metrics M_3 and M_4, because they correlate better with the distance between Bob and Eve's estimate of Bob's location, and reach a plateau only when such localization error is sufficiently large to lead to uncorrelated channel impulse responses. Our simulations using the well-known ray tracing software Bellhop show that arrival delays form a more reliable basis than arrival amplitudes when matching two channels, hence a mismatch metric should prioritize delay information. Moreover, higher-amplitude arrivals are more reliably measured at both ends of a link and should have a bigger weight in the matching metric, whereas less powerful arrivals are not a good indication of reciprocity.

Acknowledgments The authors warmly thank EvoLogics for the support and discussion on this research topic. This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU – PNRR, Mission 4 Component 2, Investment 1.3 – PE RESTART Spoke 6 – Project EMBRACE (PE00000001, CUP E63C220020700069).

REFERENCES

[1] Diamant, Roee, et al. "Secret key generation from route propagation delays for underwater acoustic networks." *IEEE Trans. Inf. Forensics Security* **18**: 3318-3333 **2023**.

[2] Pan, Pan, et al. "A Secret key generation scheme exploiting spatio-temporal acoustic channel Characteristics for Underwater Sensor Networks." *IEEE Sensors J.* **2024**.

[3] Pelekanakis, Konstantinos, et al. "Physical layer security against an informed eavesdropper in underwater acoustic channels: Feature extraction and quantization." in *Proc. UComms* **2021**.

[4] Sklivanitis, George, et al. "Physical layer security against an informed eavesdropper in underwater acoustic channels: Reconciliation and privacy amplification." *Proc. UComms* **2021**.

[5] Morozs, Nils, Paul D. Mitchell, and Yuriy Zakharov. "Target detection using underwater acoustic networking." in *Proc. IEEE OCEANS* **2023**.

[6] Aman, Waqas, Saif Al-Kuwari, and Marwa Qaraqe. "Location-based physical layer authentication in underwater acoustic communication networks." in *Proc. IEEE VTC* **2023**.

[7] Khalid, Muhammad, Ruiqin Zhao, and Nauman Ahmed. "Physical layer authentication in line-of-sight underwater acoustic sensor networks." in *Proc. IEEE OCEANS* **2020**.

[8] Zhao, Ruiqin, et al. "Physical layer node authentication in underwater acoustic sensor networks using time-reversal." *IEEE Sensors J.* **22.4**: 3796-3809 **2022**.