Indoor 1D-Localization of Omnidirectional Power-Modulated Jammers: a Machine Learning Approach with Rapid Database Generation

Paul Monferran, Antonio Costanzo, Artur Nogueira de São José, Virginie Deniau, Jonathan Villain, Valeria Loscri, Christophe Gransart

Indoor 1D-Localization of Omnidirectional Power-Modulated Jammers: a Machine Learning Approach with Rapid Database Generation

Paul Monferran, Antonio Costanzo, Artur Nogueira de São José, Jonathan Villain, Virginie Deniau, Christophe Gransart, and Valeria Loscri

***Abstract*—The threat of power-modulated jammers to electronic systems has recently been reported in the literature. These are malicious devices that emit intentional electromagnetic interference whose power changes rapidly over time. Such dynamic power emissions make it hard for traditional localization algorithms to track the jammer position in indoor environments, especially if the shadowing effects of the objects and people nearby the monitoring antennas are strong. In this work, we propose an indoor jammer localization strategy based on machine learning. The machine learning models are built from simulations based on the shooting-and-bouncing rays technique to quickly generate the required databases and provide a parametric study. The simulation model is validated by comparison with the measurement performed in a real room in the presence of a commercial jammer. Decision tree algorithms lead to predictions with an accuracy of tens of centimeters for a constant-power jammer and a power-modulated jammer. This result significantly outperforms conventional trilateration approaches. Furthermore, a new machine learning feature based on power ratios was introduced and provided good predictions even if the jamming power is unknown by the machine learning model. In addition, the main limitations are evaluated according to the uncertainties between measurement and simulations, the learning dataset size and changes in the considered environment. Finally, the proposed framework is validated using measurement as input of the machine learning models.**

***Index Terms*—Jammer, Wireless Network, Indoor Environment, Geolocalization, Machine Learning Techniques.**

I. Introduction

Intentional electromagnetic interference (IEMI) is a broad concept that includes a wide range of interference sources resulting on a disturbance being coupled to a victim [1]. IEMI threats are increasingly studied in the field of electromagnetic compatibility (EMC) for several applications [2], [3], [4]. Jammers are part of IEMI and can cause disruption to electronic systems involving wireless communication link [5], [6], [7], [8], [9], [10], [11]. In particular, all wireless communication systems can fall victim to jamming attacks, whatever the communication protocol is. Thus, critical sectors such as companies, power plants, and transportation facilities involving wireless sensors must be particularly vigilant in protecting their communications networks from any form of attack or intrusion.

This kind of electromagnetic (EM) disturbances are not taken into account by classical EMC test standards which normally deal with well-known and well-behaved interference sources. As an example, we can mention the 80% amplitude-modulated (AM) interference waveform suggested by the IEC 61000-4-3 standard [12] which contrasts with typical jamming waveforms [7], [9]. Therefore, the interference control requires a special treatment when predicting jamming effects on wireless infrastructures. When active, the jammer disrupts electronic systems; the most effective mitigation is to locate and deactivate it.

In this article, we seek to geolocate the source of interference on the basis of radio frequency activity analysis. In the literature, various methods have been proposed for geolocating a jammer. In [13], the transmission power of the communication signal between the different nodes in the network is progressively increased until the signal can be received by all nodes. The jamming signal strength (JSS) values are then deduced, and the closest node to the jammer is identified (the node with the highest JSS value). In [14], the approach is based on increasing the power between each pair of nodes in the network and using a mathematical model involving the received signal strength to estimate the position of the jammer. Other methods involving data from the data link layer can be found in [15], [16]. However, all of them are based on geometric criteria assuming free-space propagation conditions. In a closed environment, there are many reflections, which can lead to significant errors in estimating the jammer position. On the other hand, methods using machine learning (ML) algorithms have been developed for geolocalization problems in closed environments [17], [18], [19]. However, these methods have never been applied to jammers in wireless networks. We believe that ML localization algorithms can be a potential tool to deal with a new class of jamming devices called *power-modulated jammers* [20]. Authors in [20] conduct a geolocalization analysis based on the linear least square (LLS) technique in order to locate both constant-power and power-modulated jammers. However, the authors do not take reflections, diffraction and real antenna patterns into

This work was supported by the University Gustave Eiffel and financed by the ANR project DEPOSIA.

Paul Monferran, Jonathan Villain, Virginie Deniau and Christophe Gransart are with the University Gustave Eiffel, 59650 Villeneuve-d'Ascq, France (e-mail: paul.monferran@univ-eiffel.fr; jonathan.villain@univ-eiffel.fr; christophe.gransart@univ-eiffel.fr; virginie.deniau@univ-eiffel.fr).

Antonio Costanzo is with the CERADE, ESAIP, 13090 Aix-en-Provence, France (e-mail: acostanzo@esaip.org).

Artur Nogueira de São José is with the University of Brasilia, Faculty of Technology, Department of Electrical Engineering, 70910-900, Brasilia, Brazil (e-mail: anog@ene.unb.br).

Valeria Loscri is with Inria, 59650 Villeneuve-d'Ascq, France (e-mail: valeria.loscri@inria.fr).

account, which are notably present in indoor environments.

The power-modulated jammer is indeed a potential threat for wireless communication systems. The power-modulated jammer proposed in [20] dynamically adjusts its transmission power to jam all communication nodes in a given area while minimizing its own energy consumption. Such dynamic power allocation makes it harder for most localization algorithms to track the jammer position. Furthermore, the shadowing effect due to the presence of walls and other obstacles also reduces the performance of the classical LLS localization algorithm, especially if fewer than five monitoring nodes are employed. Authors in [20] evaluate the LLS performance degradation by comparing its performance under power-modulated and constant-power jamming attacks. In such scenario, the localization error can be as high as 328 m with 3 monitoring nodes and a shadowing effect with a variance of σ = 3 dBm.

Our first attempt to solve this issue was reported in [21], where we compared the performances of four ML algorithms under different environmental noise levels and multipath conditions. This work was based on a radio frequency monitoring infrastructure, consisting of antennas and multi-port receivers distributed in the area to be protected. These multi-port receivers allow simultaneous measurements from different antennas distributed in an indoor environment. The main goal was to identify the jamming source position in 1D by performing ML on the radio frequency data collected by the different antennas. This proof of concept was based on data from a simulation model.

From the work initiated in [21], we propose in this article an end-to-end localization approach in 1D. The choice of 1D localization is motivated by both the use case and the aim of resource efficiency. Although 1D localization appears restricted to narrow environments such as aircraft or trains, it can also be effectively applied in buildings, where monitoring antennas are typically installed along corridors. Moreover, 2D and 3D approaches require extensive measurements at multiple jammer positions during the learning phase, unlike the 1D approach with equivalent discretization. Such measurement phase can quickly become prohibitively time-consuming and unrealistic.

In addition to the objective of reducing electromagnetic pollution of critical indoor infrastructure caused by IEMI sources such as jamming devices through geolocation, the entire proposed methodology follows an EMC framework. Due to the presence of certain unknown parameters, such as the transmit power of the jammer, and the need to obtain a rapid location response, a ML model that learns the propagation pattern of a jammer signal at multiple locations is relevant. To overcome the excessive time needed to collect ML training data, we propose an interference model based on different antennas and classical ray tracing EM simulations. Hence, a particular attention is paid to characterizing from measurement the EM environment, the antenna network, and the jammer signal in order to simulate real-world behavior sufficiently well. In addition, discrepancies between the simulation and measurements are also studied in order to assess the propagation of uncertainties to the localization results. Moreover, we explore the limitations of our ML model according to two different EM metrics - both based on the power received by the antenna network - which serve as input of it. Therefore, we contribute to different aspects of EMC through our work, ranging from interference control in critical EM environments to interference modeling and EMC measurement techniques.

This paper is the extended version of [21] where the main contributions were:

- the validation of a ray tracing simulation model which gives us the ability to generate a fast database and,
- the validation of the geolocalization methodology with a machine learning approach for a constant-power jammer with a known emitted power.

The main contributions of this article can be summarized as follows:

- we demonstrate, with simulation and measurement data, that the proposed methodology of [21] is not suitable for a real case as we usually do not know the emitted power of the jammer,
- a new methodology to resolve this crucial limitation by validating it using simulation and measurement data,
- we validate the methodology on a new case study regarding power-modulated jammers that significantly increases the difficulty of location prediction, and, which address the problem stated in [20], according to which the classical LLS location technique cannot deal with power-modulated jammers,
- we demonstrate that the methodology outperforms the classical methods even for the constant-power jammer,
- we evaluate how simulation inaccuracies propagate to ML predictions to assess the reliability of the methodology,
- we carry out a sensitivity analysis taken into account the learning database size, the robustness of the proposed method according to the choice of the test data, and the limitations due to an environment modification,
- we perform a time-cost analysis of the methodology that enables very rapid database generation, model learning and prediction.

In Sec. II, we resume the test setup, measurement process and simulation model of [21] on which we base our work, and we also introduce the localization challenge associated with the power-modulated jammer. Then, in Sec. III we describe our ML approach for jammer localization. Then, Sec. IV is dedicated to the geolocalization performances. Sec. V carries out the propagation of uncertainties and sensitivity analyses while Sec. VI addresses the practical aspects of deploying our methodology. Finally, we draw the conclusions in Sec. VII.

II. Related work: measurements and simulation model

A. Test setup and measurements

The same test setup than in [21] is taken into account in this article. Fig. 1 reminds this measurement setup. Three EM-6116 omnidirectional antennas vertically polarized [22] noted A, B and C are aligned along the x-axis where A and B are located at $z = 1$ m while C is located at $z = 1.5$ m as shown. Indeed, C is located at $z = 1.5$ m to allow the jammer (noted J) to be located at $z = 1$ m. This jammer

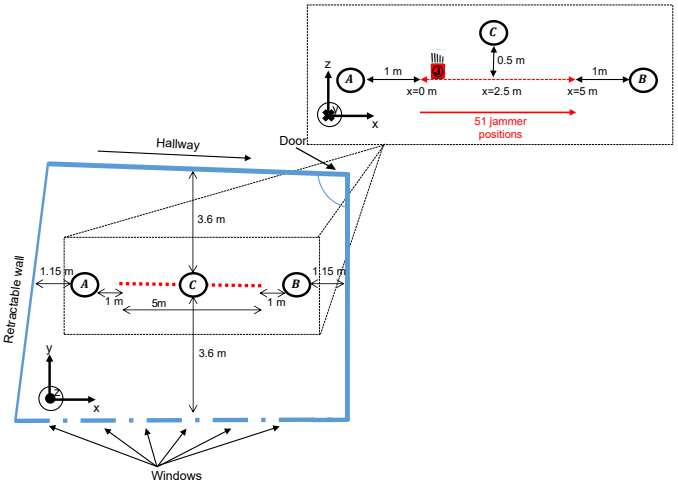


Fig. 1: Experimental threat model in the xy plan (at bottom) with a zoom (at top) on the xz plan.

Fig. 2: Picture of the commercial jammer.

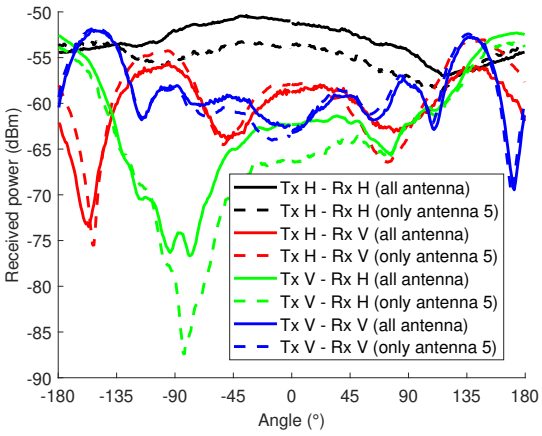


Fig. 3: Radiation pattern of the commercial jammer with all antennas connected and with only antenna 5 connected.

is actually located at 51 different positions along the x-axis which corresponds to a displacement of 10 cm over 5 m. To avoid near-field effects, the jammer is at least 50 cm away from antenna C and 1 m away from antennas A and B. The jammer power is considered strong enough for all three antennas to be in the jamming zone. The three antennas are passive and we consider only the received jamming signal. Hence, the received jamming signal is measured from the three antennas connected to a software-defined radio (SDR) receiver with coaxial cables. In addition, the commercial jammer is powered by an internal battery which lasts approximately five hours, working continuously. Besides, it is connected to five external antennas through SMA-type connectors as shown in Fig. 2. Each jammer antenna covers a certain frequency band

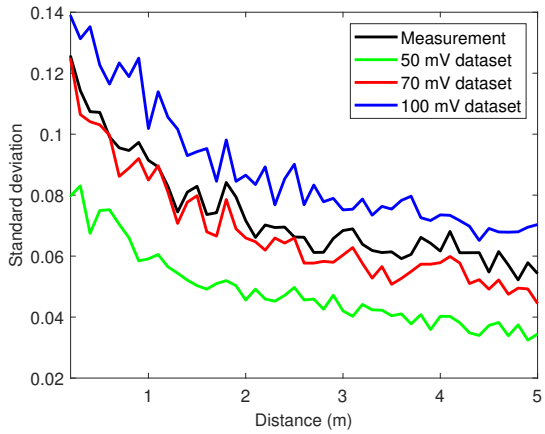


Fig. 4: Standard deviation of the mean power received by antenna A at each jammer position.

based on a sweep time of 5 µs. However, only one of these antennas, the one transmitting in the 2.5-2.7 GHz band, was activated during the tests (antenna 5). Such choice aimed at emulating jamming attacks targeting Wi-Fi services in the 2.4 GHz frequency band without actually disturbing the building's local network.

For practical considerations, commercial jammers are generally designed to radiate uniformly in all directions. To verify this claim, we measure the radiation pattern of our jammer in an anechoic chamber under two typical operating conditions. The results are shown in Fig. 3. In the first configuration, which is the most commonly used in practical applications, all five antennas remain connected. In the second case, only the antenna 5 is physically connected to the device, while the other four antennas are replaced with 50-ohm loads. In both scenarios, the only active antenna is the antenna 5 (see the antenna selection keys in Fig. 2). From the measurements, the typical radiation behavior of a monopole antenna can be observed, exhibiting an omnidirectional pattern in the plane transverse to the antenna orientation and an absolute null in the vertical plane. The differences between the two radiation patterns (*all antenna* vs *only antenna 5*) are minimal. The presence of the other four antennas slightly affects the received power due to minor out-of-band radiation from the out of band antennas. However, it does not alter the shape of the radiation pattern, either for the co-polar or cross-polar components. Indeed, the distances between the antennas and their physical dimensions do not allow for phase relationships that would affect the directivity of the system.

Knowing the jammer's radiation pattern, from the in-phase / quadratic (IQ) signal measured by the SDR card and written $I(t) + jQ(t)$, we extract the square mean μ, according to:

$$\mu = \frac{1}{N+1} \sum_{t=0}^{N} \left(I\left(t\right)^2 + Q\left(t\right)^2 \right). \quad (1)$$

where $N = 32767$ samples by acquisition. Therefore, Eq. 1 represents the mean power of a single IQ signal acquisition. More details about the signal acquisition can be found in [21].

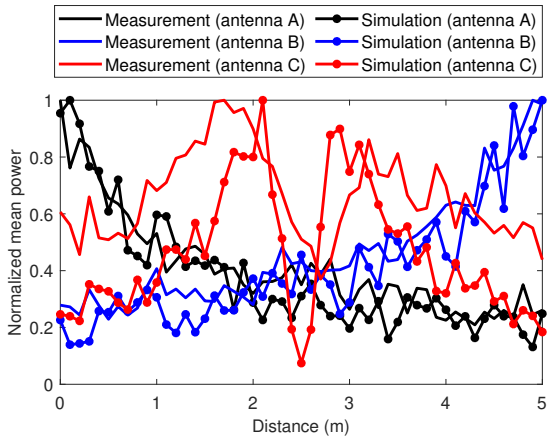


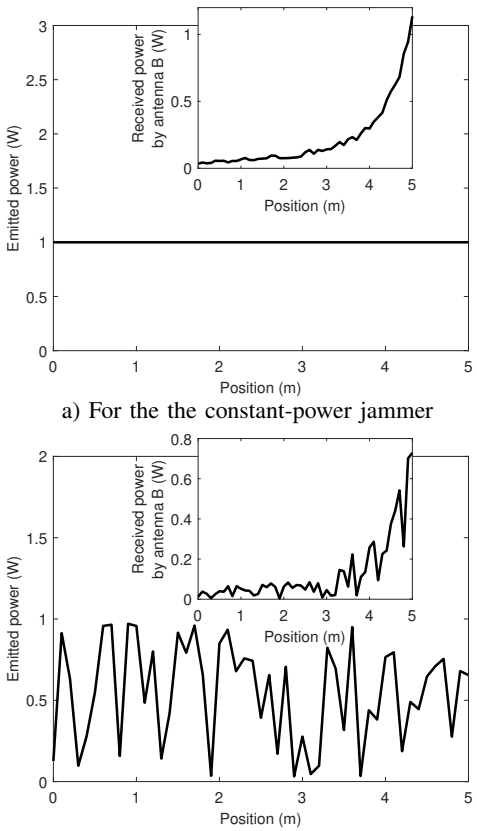
Fig. 5: Comparison between simulations with 70 mV noise level and measurement for the three antennas.

B. Simulation model

In the previous work [21], a shooting-and-bouncing rays (SBR) technique [23] is used to simulate the test setup depicted in the previous subsection. In this paper, we perform simulations with the same simulation model which include the addition of a noise source to the received jamming signals. All details of the simulation model can be found in [21]. In addition, we specify that a realistic monopole radiation pattern is employed for the receiving antennas according to [22]. Given that the near-to-far field transition of our antennas is estimated around 3.21 cm - according to the Fraunhofer distance [24] calculated from the antenna gain of [22] - only far-field radiation patterns are considered. Furthermore, the proposed simulation model has been validated in [21] by making a comparison between simulation and measurement. However, further analysis has shown that the 70 mV noise level from [21], reused in this paper as shown in Fig. 5, corresponds better to the measured power than the 50 mV noise level as shown in Fig. 4. For the sake of brevity, only the standard deviation for antenna A is shown, even though the behavior is identical for antennas B and C.

C. Simulated cases

The computational environment allows us to simulate the emissions of two classes of jamming devices: the classical frequency-sweeping jammer described in [7] and the power-modulated jammer which was recently proposed in [20]. Authors in [20] describe a jamming signal whose transmission power evolves with time. A comparison between the classical frequency-sweeping jammer (a *constant-power* jammer) and the power-modulated jammer is illustrated in Fig. 6 in terms of emission power. In this figure, we simulate the case where the jamming device is gradually moved from position $x = 0$ m to the position $x = 5$ m in a straight line. Fig. 6 illustrates the emitted power as well as the associated received power by antenna B, which is the quantity we actually measure and which will be used to find the position of the jammer. It can be noted that the received power may be slightly higher than the emitted power due to the addition of the noise source (compare emitted and received power levels in Fig. 6a for $x = 5m$.)



a) For the the constant-power jammer

b) For the power-modulated jammer

Fig. 6: Emitted power at each position for the constant-power jammer (at top) and the power-modulated jammer (at bottom). An inset shows the received power of antenna B with the 20 mV noise level. [21].

Indeed, it is a challenge to predict the position of a power-modulated jammer according to its emission power levels. This is because its emitting power is random. Consequently, one cannot easily correlate jamming power and jamming position, as suggested by the insets in Fig. 6. Hence, we define the power-modulated jamming signal, $j'(t)$, as a single frequency source with a variable emission power, as follows:

$$j'(t) = \sqrt{P_J(x)}.\exp j(f_q t), \tag{2}$$

where f_q is the angular frequency, which is a fixed value for the power-modulated jammer while it is a linear function of time for the constant-power jammer and

$$P_J(x) \sim U(]0, 1]) \tag{3}$$

is the emission power (in watts) as a function of the position x, within the indoor environment. Eq. 3 indicates that the emission power obeys a uniform law spanning from 0 W to 1 W, with the lower limit open to avoid the no-jamming case.

III. Machine Learning Approach

A. Database description

For each studied case, six different databases are studied to evaluate the sensitivity of the ML methods to the level of

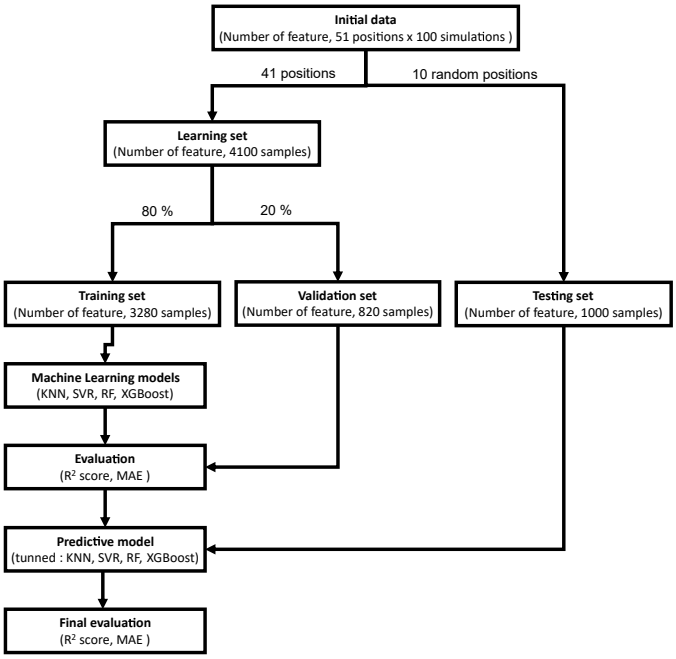


Fig. 7: Flowchart of the machine learning process.

noise as stated in Sec. II. Each database is constituted with three power measurements taken from each antenna (A,B,C) for each position of the jammer with a specific level of noise. The features, which are the three mean powers received by antennas, or, the three mean power ratios, are depicted in the Sec. III-C. For each database, 100 simulations are performed to take into account the variability according to the noise. Thus, we deal with these six databases of size 3 (the three features) $\times$ 51 (number of jammer positions) $\times$ 100 (number of simulations) as shown in Fig. 7. For each database, we would like to predict the jammer position using the considered features as input. Hence, we split the database according to the jammer positions as follows: the learning set is constituted with 41 positions while the remaining 10 positions are reserved only for final performance evaluation in the testing set (see Fig.7 of [21]). The choice of this 10 positions has been initially randomly chosen. For the sake of consistency, the same draw is used regardless of the database used. We note that some testing positions present specific characteristics. In particular, the position at $x = 5$ m is in the testing set, thus outside the range of the training set. In addition, the position $x = 2.4$ m is almost symmetric with the position $x = 2.6$ m according using the three features. This symmetry is clearly reinforced as noise levels increase. Moreover, some neighbor positions are in the testing set as for instance $x = 3.5$ and $x = 3.6$. Furthermore, the learning set of size 3 (the three considered features) $\times$ 4100 (41 jammer positions $\times$ 100 simulations) is split in a stratified fashion with 80% of these data for the training and 20% for the validation. As for the testing set, validation and training sets are the same whatever is the database used.

B. Machine learning methods

Based on previous results in [21], decision tree algorithms are better suited to our problem. Thus, we explore the performances of the following two well-known ML models in regression for each dataset:

- **Random forest**: Random forest (RF) [25], [26] is an algorithm based on the assembly of decision trees which aggregates their predictions by averaging. It can be applied to a wide range of prediction problems and have few parameters to tune. Thus, in our context, it is a must-have algorithm to take into account. Through tuning step, 100 trees have been chosen to be optimal.
- **XGBoost**: XGBoost [27] is a decision tree ensemble learning library that can be used for classification or regression. It uses gradient boosting to learn an ensemble of decision trees. XGBoost has made a name for itself by winning a large number of ML competitions, and particularly successful with tabular and structured data. As our input is tabular and corresponds to independent parameters with interpretable physical meaning, XGBoost is particularly suited to perform competitively.

C. Metrics

As mentioned previously, two different types of feature are used as input of our ML algorithms. On the one hand, we merely use the three mean powers received by each antenna calculated by the SBR model, respectively noted P_{A}, P_{B} and P_{C}. On the other hand, the three features are the three mean power ratios defined as follows:

$$P_{\mathrm{AB}} = \frac{P_{\mathrm{A}}}{P_{\mathrm{B}}}, \tag{4}$$

$$P_{\mathrm{AC}} = \frac{P_{\mathrm{A}}}{P_{\mathrm{C}}}, \tag{5}$$

$$P_{\mathrm{BC}} = \frac{P_{\mathrm{B}}}{P_{\mathrm{C}}}, \tag{6}$$

Indeed, we hope to increase the prediction accuracy for the power-modulated jammer localization using this second kind of metric instead of the first one, as it does not take into account the emitted power. Nevertheless, this second metric is more sensitive to a small variation of the power due to his ratio formulation, and consequently, more sensitive to noise.

Furthermore, to interpret the results, we use two performance metrics:

- the mean absolute error:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\hat{k}_i - k_i| \tag{7}$$

where $\hat{k}_i$ and k_i are, respectively, the predicted value and the ground truth of the i-th sample, where n is the total number of experiments. The ground truth is the position to predict.

- the coefficient of determination:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(k_i - \hat{k}_i)^2}{\sum_{i=1}^{n}(k_i - \bar{k})^2} \tag{8}$$

with $\bar{k}$ the mean value across samples.

The mean absolute error (MAE) gives an indication of how close the predictions are to the target values, and therefore of the model's practical potential. A MAE of 0 corresponds to perfect prediction. In contrast, the coefficient of determination (R^2) measures the proportion of the target variable's variance

TABLE I: Average coefficient of determination and mean absolute error for the constant-power jammer

Dataset	Model	Mean power features				Mean power ratio features			
		Validation test		Testing set		Validation test		Testing set	
		R^2	MAE (m)	R^2	MAE (m)	R^2	MAE (m)	R^2	MAE (m)
1 mV	RF	0.999	$2.073 \cdot 10^{-5}$	0.937	0.191	1.000	$2.439 \cdot 10^{-6}$	0.970	0.143
	XGBoost	1.000	$4.460 \cdot 10^{-4}$	**0.989**	**0.121**	1.000	$8.673 \cdot 10^{-4}$	**0.993**	**0.107**
20 mV	RF	0.997	0.037	**0.985**	**0.126**	0.997	0.047	0.983	0.123
	XGBoost	0.996	0.057	0.984	0.133	0.996	0.066	**0.990**	**0.104**
30 mV	RF	0.992	0.071	0.965	0.174	0.990	0.090	0.970	0.164
	XGBoost	0.990	0.104	**0.972**	**0.168**	0.990	0.104	**0.976**	**0.145**
50 mV	RF	0.975	0.151	**0.940**	**0.227**	0.968	0.173	**0.941**	**0.228**
	XGBoost	0.974	0.173	0.940	0.241	0.970	0.181	0.936	0.252
70 mV	RF	0.956	0.209	**0.901**	**0.289**	0.945	0.247	0.873	**0.326**
	XGBoost	0.957	0.223	0.892	0.322	0.951	0.243	**0.878**	0.338
100 mV	RF	0.906	0.318	0.788	**0.432**	0.890	0.353	0.742	**0.491**
	XGBoost	0.905	0.333	**0.796**	0.449	0.897	0.346	**0.767**	0.494

TABLE II: Average coefficient of determination and mean absolute error for the power-modulated jammer

Dataset	Model	Mean power features				Mean power ratio features			
		Validation test		Testing set		Validation test		Testing set	
		R^2	MAE (m)	R^2	MAE (m)	R^2	MAE (m)	R^2	MAE (m)
1 mV	RF	1.000	$3.658 \cdot 10^{-6}$	0.949	**0.221**	1.000	$2.963 \cdot 10^{-4}$	0.988	0.109
	XGBoost	1.000	0.002	**0.959**	0.223	1.000	$6.935 \cdot 10^{-4}$	**0.993**	**0.102**
20 mV	RF	0.993	0.035	0.967	0.188	0.997	0.115	0.944	0.215
	XGBoost	0.988	0.090	**0.977**	**0.146**	0.980	0.120	**0.973**	**0.170**
30 mV	RF	0.984	0.065	0.894	0.294	0.954	0.185	0.915	0.281
	XGBoost	0.972	0.154	**0.939**	**0.234**	0.957	0.191	**0.943**	**0.240**
50 mV	RF	0.955	0.172	0.880	0.317	0.910	0.281	0.889	0.317
	XGBoost	0.941	0.238	**0.899**	**0.276**	0.920	0.272	**0.906**	**0.310**
70 mV	RF	0.904	0.292	0.853	0.346	0.838	0.411	0.826	0.412
	XGBoost	0.898	0.331	**0.868**	**0.333**	0.861	0.390	**0.851**	**0.406**
100 mV	RF	0.810	0.446	0.718	0.560	0.723	0.518	0.723	0.518
	XGBoost	0.812	0.461	**0.775**	**0.445**	0.751	0.530	**0.763**	**0.492**

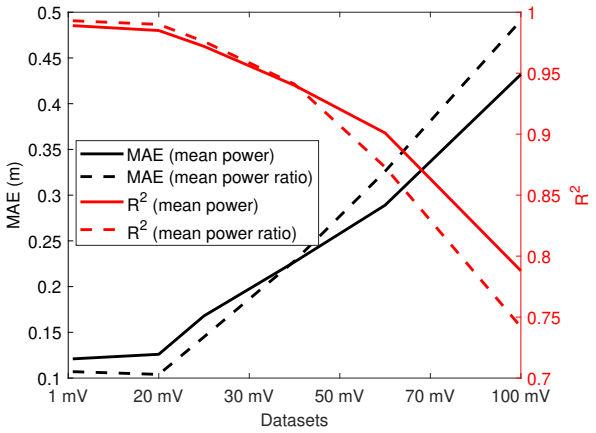
that can be explained by the model. In particular, a $R^2 = 1$ corresponds to a perfect prediction, while $R^2 = 0$ means that the model consistently predicts the mean value $\bar{k}$. If R^2 is negative, it means that we are not even predicting $\bar{k}$. The mean absolute error and the coefficient of determination serve different purposes and provide complementary information. The mean absolute error measures prediction accuracy by quantifying the average magnitude of errors between predicted and actual values. However, the coefficient of determination assesses the model's ability to capture information from the input variables.

IV. Geolocalization performances

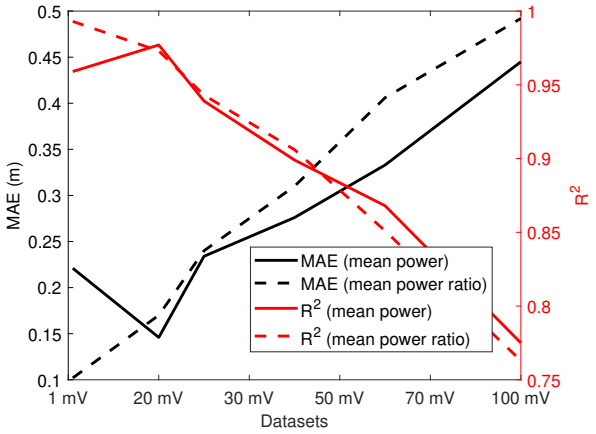
A. Overall performances

For each dataset, the results for each ML model are reported in Table I for the constant-power jammer and in Table II for the power-modulated jammer. Unsurprisingly, for both cases, the MAE usually increases with the level of noise, while R^2 decreases. Indeed, the higher the noise level, the more difficult it is to predict because the function to be interpolated becomes extremely discontinuous. We will focus on the testing set rather than the validation set, as the objective of the study is to geolocate a jammer in a position not included in the training set. Furthermore, with regard to the learning set boundaries ($0 \text{ m} \leq x \leq 4.9 \text{ m}$) and the discretization (41 points), we consider that a prediction with a MAE ≤ 0.2 m is a very good performance. Therefore, the ML models present good overall performances for a noise level below 50 mV. According to Table I and Table II, we claim that predictions can be very good if, during measurements, we can achieve a noise level below 30 mV.

Moreover, it is interesting to observe the behavior of the predictions according to the metrics used. In order to have a clearer representation, Fig. 8 compares the performances of the two metrics according to the considered datasets for the constant-power jammer and the power-modulated jammer using the XGBoost model. We note that in both cases (the constant-power jammer and the power-modulated jammer) the mean power ratio performs better up to a certain noise level and, after this threshold, the trend reverses. This behavior



a) For the the constant-power jammer



b) For the power-modulated jammer

Fig. 8: MAE and R^2 of the testing set according to the considered datasets for the constant-power jammer and the power-modulated jammer. Jamming power:]0,1] W (training set) and]0,1] W (testing set).

is due to the mathematical formulation of the mean power ratio which is more sensitive to a small variation. In other words, using the mean power ratio, the greater the noise, the greater the variation, the higher the variance of the input characteristics and therefore the more difficult the prediction. Furthermore, the considered threshold differs depending on the use case. Indeed, this threshold is higher for the constant-power jammer than for the power-modulated jammer because the latter by definition imposes an additive variation between jammer positions. Last but no least, with regard to the mean power metric performances for the power-modulated jammer at 1 mV and 20 mV, we also argue that adding a little noise could increase the prediction performance for a function that is still discontinuous to predict.

B. Comparison with trilateration

In order to further validate the effectiveness of the proposed ML-based technique, we performed a comparison with the classical trilateration approach based on the received signal strength indication (RSSI) [28]. This comparison is performed

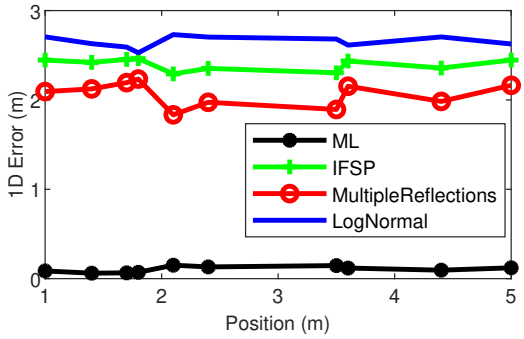


Fig. 9: Comparison between trilateration, considering different propagation models, and ML approach.

considering the 20 mV testing set of the constant-power jammer, according to the configuration described in Fig. 1.

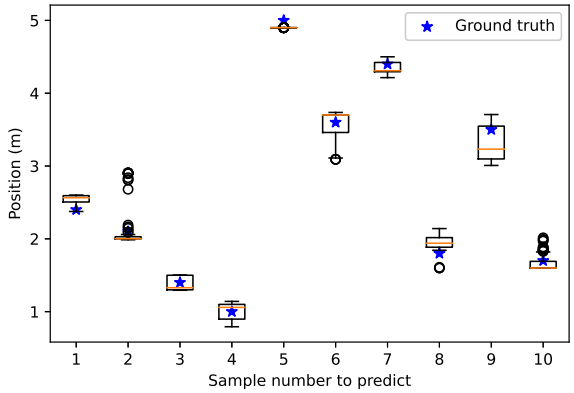
We consider, for comparison, three propagation models:

- indoor free space propagation (IFSP)
- multiple reflections (up to the 5th order and reflection coefficient equal to 0.7)
- log-normal model (with sigma equal to 3 dB).

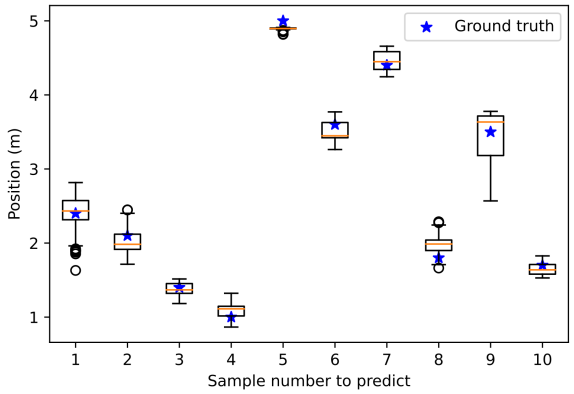
For all these propagation models, the reference power is taken one meter away of each receiving antenna, according to the 20 mV dataset. The link budget associated to each path, has been exploited for estimating the distances between the jammer and the three receivers. This operation has been repeated for each propagation model. These distances have been given in input to a classical trilateration algorithm for computing the position of the jammer. Then, for each position of the considered testing set and each associated simulation, we computed the distance error as the difference between the trilateration prediction and the ground truth. Averaging the 100 differences computed, we obtain an 1D error presented in Fig. 9. We note that the machine learning based approach significantly outperforms the RSSI based approach in terms of positioning error. This is due to the fact that classical techniques based on RSSI and trilateration do not take into account pattern variation in antennas and complex interaction of the signal in the environment.

C. Performances by position

To go into further details, Fig. 10 illustrates the predictions of the 20 mV testing set with a box-plot representation according to the best predictive model for the constant-power jammer (RF) and the power-modulated jammer (XGBoost). The x-axis corresponds to the 10 random positions to predict, reported in their real value in the y-axis. We observe some outliers predictions, as for instance, in the case of the constant-power jammer, the position $x = 2.1$ m which corresponds to the sample 2 to predict. However, for each position, we note that the median of the predicted values is very close to the ground truth. Thus, this result demonstrates the robustness of our ML algorithms, which perform well on average whatever the position. Nevertheless, the decision three algorithms are not able to interpolate as shown for the sample 5 which correspond to the position $x = 5$ m usually predicted at the known value $x = 4.9$ m.

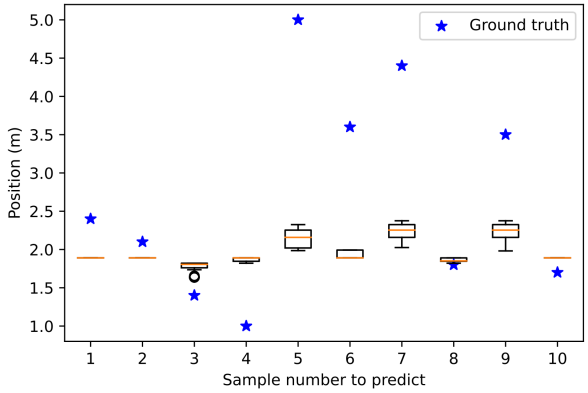


a) For the the constant-power jammer

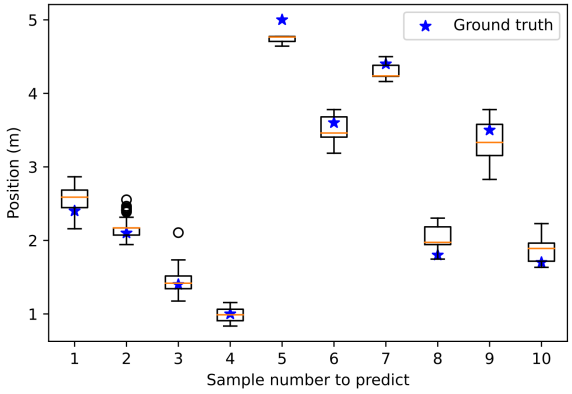


b) For the power-modulated jammer

Fig. 10: Box-plot of the predicted values (with the median represented by the red line) versus the ground truth using RF for the constant-power jammer and XGBoost for the power-modulated jammer both with the 20 mV noise testing set and the mean power features. Jamming power:]0,1] W (training set) and]0,1] W (testing set).



a) Using the mean power feature



b) Using the mean power ratio feature

Fig. 11: Box-plot of the predicted values (with the median represented by the red line) versus the ground truth using XGBoost for the power-modulated jammer with a 20 mV noise using the mean power features and the mean power ratio. Jamming power:]0,1] W (training set) and]0,10] W (testing set).

D. Feature metrics limitations

As a reminder, we have shown in the Sec. IV-A that the mean power ratio metric leads to better performance than the mean power metric up to a certain noise level. However, the mean power ratio has the advantage of being effective without knowing the power emitted. Fig. 11 typically illustrates this huge advantage. For both plots, the XGBoost model is trained from the 20 mV training set with power-modulated jamming power between 0 and 1 W. At top, the figure shows the predictions using the mean power for a new testing set with a 20 mV noise but with random emitted power between 0 and 10 W. It is clearly highlighted that the model is unable to predict such values. On the other hand, when we train an XGBoost model with the same training set but preprocessing the values with the mean power ratio, we obtain very good predictions on the new testing set as shown in the bottom plot. Indeed, by using the mean power ratio, a kind of normalization is carried out. In practice, it is essential to be able to predict a jammer without knowing its transmission power, because we do not know its power a priori. In this regard, the mean power ratio metric must therefore be used to achieve this requirement.

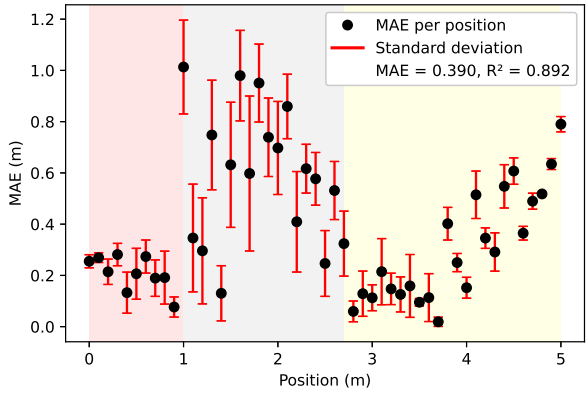


Fig. 12: MAE per position for measurement with the XGBoost model trained with the 70 mV using the mean power ratio.

TABLE III: MAE (in m), R^2, simulation time (in s) and MADMS according to the reflection order and the antenna model

Antenna model	1st reflection order	2nd reflection order	3rd reflection order
monopole	0.0600 / 0.997 / 12.921 / 0.0340	0.0550 / 0.997 / 13.014 / 0.0330	**0.0463 / 0.998 / 13.203 / 0.00123**
monopoleRadial	0.179 / 0.978 / 46.794 / 0.300	0.181 / 0.980 / 46.912 / 0.298	0.177 / 0.980 / 54.030 / 0.307
yagiUda	0.237 / 0.960 / 42.552 / 1.23	0.237 / 0.960 / 48.310 / 1.23	0.240 / 0.960 / 48.384 / 1.23
horn	0.650 / 0.683 / 130.382 / 0.570	0.650 / 0.683 / 171.921 / 0.570	0.650 / 0.683 / 180.230 / 0.570

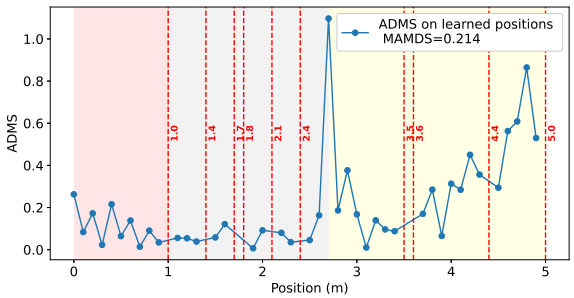


Fig. 13: Computed ADMS on learned positions. The red lines indicate the unlearned positions.

E. Methodology validation with measurement

During the measurement process, 100 acquisitions have been performed for each jammer position with the constant-power jammer of Fig. 2. As stated in Sec. II-B, the 70 mV dataset is the most realistic dataset compared to measurement. Thus, the trained XGBoost model with the 70 mV training set is used with the measured data as input. Fig. 12 shows the MAE and the standard deviation of the 100 measured data at each jammer position using the mean power ratio feature. We note that the prediction accuracy has an important variation according to the jammer position. As a consequence, we propose to compare the prediction results to the average difference between measurements and simulations per position noted ADMS and computed as follows:

$$\text{ADMS} = |\hat{P}_p^{sim} - \hat{P}_p^{mes}| \tag{9}$$

for each learned position p, with $\hat{P}_p^{sim}$ and $\hat{P}_p^{mes}$ the average feature computed as follows where $n_{sim} = n_{mes} = 100$:

$$\hat{P}_p^{sim} = \frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \left(\frac{P_{\text{AB}}^{sim_i} + P_{\text{AC}}^{sim_i} + P_{\text{BC}}^{sim_i}}{3} \right) \tag{10}$$

$$\hat{P}_p^{mes} = \frac{1}{n_{mes}} \sum_{i=1}^{n_{mes}} \left(\frac{P_{\text{AB}}^{mes_i} + P_{\text{AC}}^{mes_i} + P_{\text{BC}}^{mes_i}}{3} \right) \tag{11}$$

In addition, we also define the ADMS average across all positions, denoted MADMS.

By comparing Fig. 12 and Fig. 13, we can see that - except for the range [1-2.4] m - the predictions are better for the jammer positions that are best represented in simulation (low ADMS as in the pink area), whereas they are less accurate for positions where the discrepancies between simulation and measurement are significant (high ADMS as for the right side of the yellow area). Moreover, six of the ten unlearned test positions lie in the [1-2.4] m range (gray areas), which makes prediction more difficult in this range, not only for these positions but also for their neighbors. However, the results are quite consistent with the results of Table I. Over all the positions, the $R^2 = 0.892$ MAE $= 0.390$ m where we can expect at best a MAE $= \frac{41}{51} * 0.243 + \frac{10}{51} * 0.338 \simeq 0.262$ m with perfect fidelity between measurement and simulation. Thus, even with our lack of accuracy in the modelling, an average deviation of $0.390 - 0.260 = 0.128$ m could be acceptable for deployment. In addition, using the mean power feature results in a MAE $= 1.539$ m and a $R^2 = -0.621$, which once again underlines the mandatory use of the mean power ratio feature. Accordingly, the results support the validity of the methodology, with a manageable additional uncertainty that may be further decreased by refining the simulation model.

V. Propagation of Uncertainties and Sensitivity Analyses

A. Uncertainty analyses

In this subsection, we evaluate how uncertainties between measurement and simulation could propagate to ML predictions. In order to fully control the parameters, we perform the analyses in simulation. Only constant-power jammer is considered.

A.1 Antenna radiation pattern diagram and reflection order

Here, we evaluate the impact of two parameters, namely the radiation pattern of our three antennas and the reflection order used in the SBR simulation. To do this, we choose as a reference a simulation with three monopoles as antennas and taking into account up to the third order of reflection at 20 mV noise level. Therefore, we train and saved an XGBoost model on these data using the mean power ratio metric. As in Sec. IV-E, we computed the MAE, the R^2 and the MADMS on the 51 positions with fresh data from 100 simulations varying the order of reflection and the antenna model (all antennas are operating at 2.6 GHz with the default design values from the Matlab library [29]).

In order to gradually degrade the inferred pattern from the learned pattern, we consider different antenna models, namely:

- **Isotropic (for the jammer only)**: Theoretical spherical radiation pattern radiating uniformly in all directions (0 dBi).
- **Monopole**: Squashed torus pattern, omnidirectional in the horizontal plane; requires a conducting ground plane.
- **Radial Monopole**: A monopole implementation using a radial ground plane, providing a stable, self-contained omnidirectional pattern.
- **Yagi-Uda**: Highly directional pattern with a narrow beam and high gain (7-15+ dBi) in one specific direction.

TABLE IV: Propagation of uncertainties related to inaccuracies in the jammer model

Jammer model	MAE (m)	R^2	Simulation time (s)	MADMS
isotropic	0.0550	0.997	13.01	0.0330
monopole	0.0677	0.996	20.364	0.114
monopoleRadial	0.138	0.988	54.349	0.140
yagiUda	1.552	-0.654	49.475	214.994

- **Horn**: Highly directional, high-gain pattern with a characteristic flared waveguide structure that creates a focused beam at microwave frequencies.

The results are presented in Table III. First, we note that the MAE and MADMS values of the reference model are not zero, which is normal given that 1) the testing set is built from fresh data and 2) it takes into account 10 unlearned positions. Second, it is highlighted that the impact of the reflection order can be neglected, unlike the radiation pattern, which plays a major role in our predictions. The monopole antenna yields superior performance because its omnidirectional radiation pattern provides more consistent coverage of the jammer trajectory compared to the other two antennas. The Yagi antenna, being more directive, illuminates a smaller portion of the trajectory; however, its significant sidelobes contribute valuable spatial information. Conversely, the horn antenna is excessively directive, and in the absence of prior knowledge about the jammer's position, it delivers the poorest performance.

Moreover, the antenna model has a significant impact on the simulation time. In addition, we note that the MADMS cannot be used as an acceptance criterion to ensure good predictions. Indeed, it could be very high (as for the Yagi antenna for instance) but yet leading to good predictions. This could be explained by the fact that, due to the mathematical formula for the mean power ratio metric, the ADMS could exhibit a significant increase at certain positions, which would have a major impact on MADMS, but not on MAE, which is calculated here as an average across all positions. However, a very low MADMS should yield good predictions, but this is almost impossible to demonstrate and therefore difficult to define an acceptance threshold.

A.2 Jammer radiation pattern diagram

We explore here how uncertainties related to inaccuracies in the jammer model could propagate to ML predictions. We employ the same reference model as in Section V-A1, constructed from an isotropic jammer. To ensure consistency with the results of Sec. IV, 100 simulations are run with second-order reflections and monopole-modeled antennas across different jammer models. Table IV shows the results. It points out that modeling the jammer as a monopole does not differ greatly from modeling it as an isotropic antenna. However, attention should be paid to jammer modeling, which could lead to significant location errors. Clearly, the ML model is only able to correctly predict the position of the jammer for a propagation model that is sufficiently similar to the one that was learned. Therefore, when learning with an omnidirectional jammer, we cannot accurately locate a directional jammer, as shown in Table IV.

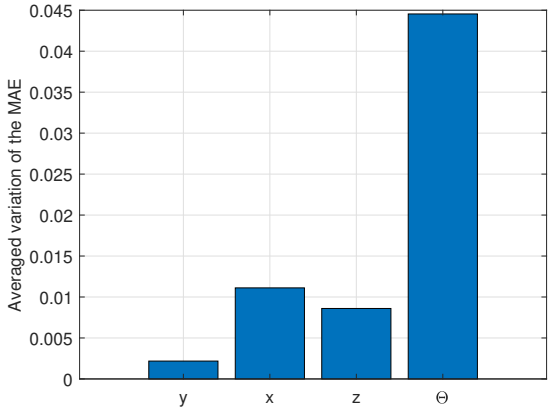


Fig. 14: Average variation of the MAE according to the antenna C angle and its displacement along x, y and z axes.

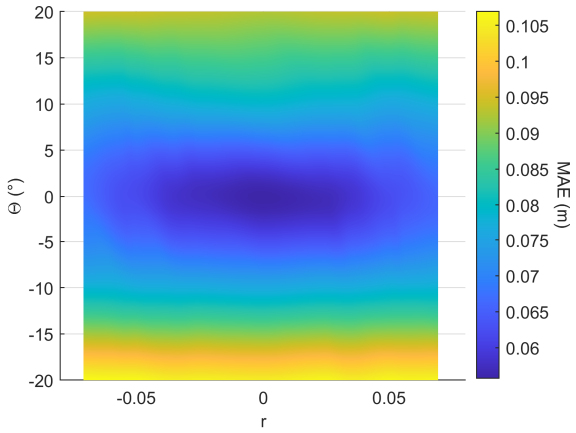
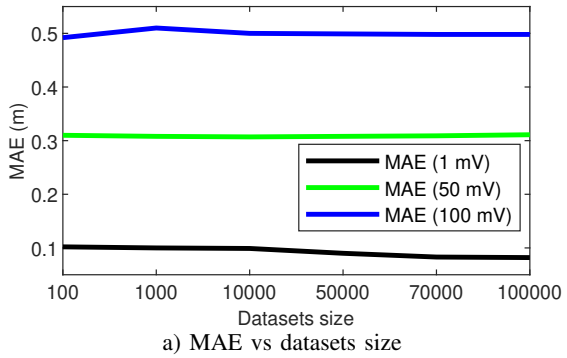


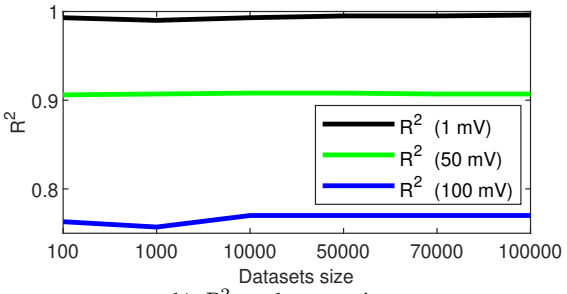
Fig. 15: Computed MAE according to the antenna C angle and its distance to the original learned position.

A.3 Antenna position

Now, we evaluate the impact of a misalignment between the antennas of the trained model and those on which we test geolocation. The reference model is the same as Sec. V-A1. To be consistent with all previous results, we perform 100 simulations taking into account only the second order of reflection and modeling the three antennas as monopoles, for different positions and angles of the antennas. For the sake of brevity, these variations only concern antenna C. Therefore, in order to remain within a possible range of uncertainty, we decide to move antenna C by ± 5 cm along the x,y and z axes for an angle Θ between the z axis and antenna C of $\pm 20°$ (positive if $x > 0$). For each configuration, the MAE is computed on the 51 positions. Fig. 14 shows that the most influential parameter is clearly the angle, while, as expected due to the position of antenna C, the variation along the y-axis is less influential than those of the x- and z-axes. Thus, we decide to explore the MAE variation according to to Θ and the distance $r = \pm\sqrt{x^2 + z^2}$ (positive if $x > 0$). This result is presented in Fig. 15. It is emphasized again that Θ has more impact than the position of the antenna, which has virtually no impact. Furthermore, as expected, the prediction becomes less accurate with the variation of Θ. In addition, we note that the MAE is almost symmetrical according to



a) MAE vs datasets size



b) R^2 vs datasets size

Fig. 16: MAE and R^2 of the testing set for the power-modulated jammer with the mean power ratio according to the number of simulations performed for the 1 mV, 50 mV and 100 mV.

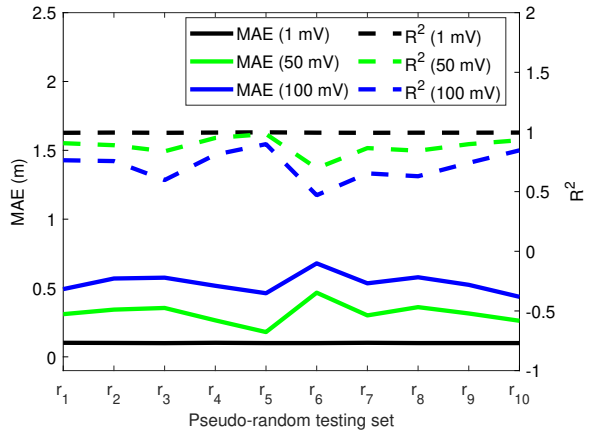


Fig. 17: MAE and R^2 of the testing set for the power-modulated jammer with the mean power ratio according to according to 10 different pseudo-random testing sets for the 1 mV, 50 mV and 100 mV.

the variation of Θ, with a slightly greater impact for negative values of Θ. This could be explained by the fact that the negative Θ has more impact at positions below 2.5 m, where a large number of unlearned positions are present (see Fig. 13). Finally, overall, the propagation of these uncertainties has a low impact according to MAE values.

B. Sensitivity analyses

B.1 Dataset size sensitivity

A first sensitivity analysis is performed based on the dataset size. The aim is to determine whether the model's performance can be improved with more training data, or whether a few pieces of data are sufficient. Fig. 16 illustrates the results according to six different dataset size built from 100 to 100000 simulations for three level of noise in the case of the power-modulated jammer with the mean power ratio. We note that the MAE and the R^2 are almost constant whatever the dataset size. For a dataset with a very low noise (1 mV), increasing the dataset size leads to a slow increase in performance, but not enough to be relevant. On the other hand, for a very noisy dataset (100 mV), increasing the dataset size could lead to better performances, for example when comparing the datasets of 100 and 1000 samples per position. This is probably due to the random initialization of the RF algorithm. In any case, the variations in performance are small enough to be considered irrelevant. Therefore, we can conclude that 100 simulations are sufficient to obtain relevant results.

B.2 Evaluated jammer position sensitivity

So far, the same 10 tested jammer positions have been evaluated. To validate the robustness of our methodology, we perform a sensitivity analysis with 10 testing sets noted r_i with r_1 taking the testing positions. The remaining r_i are built as follows:

- the combination of the 10 positions must be unique,
- the testing set cannot contain more than two successive positions,
- the testing set cannot contain two successive edge positions.

As in Sec. V-B1, the evaluated datasets consider the same three level of noise in the case of the power-modulated jammer with the mean power ratio using the XGBoost algorithm. The results are presented in Fig. 17. We note that the methodology is very robust when the noise is low, while some variation may occur when the noise increases. Indeed, noisier datasets are more sensitive to the position pattern chosen, due to the discontinuity of the function to be predicted. Furthermore, the pseudo-random draw can create a critical or optimal distribution for the learning model.

B.3 Environment modification sensitivity

In practice, the modeled room undergoes a number of changes over the course of the day (people moving around, this object moving around, objects being added, etc.). Thus, we decide here to analyze the sensitivity of the localization method to changes in the environment. To achieve this, we draw up a test plan with nine different configurations of the initial room presented in [21]. These configurations are illustrated in Fig. 18. From setup f to setup i, we only change the configuration of the tables and chairs, while people are also added from setup a to setup e. We retain the same simulation characteristics as those described in [21], except the human body, which is modeled in accordance with [30]. Moreover,

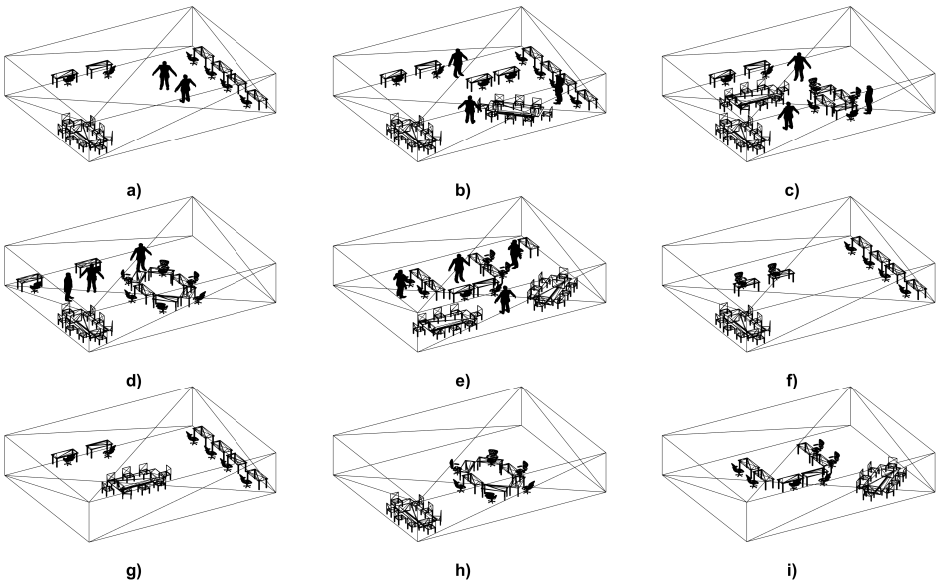
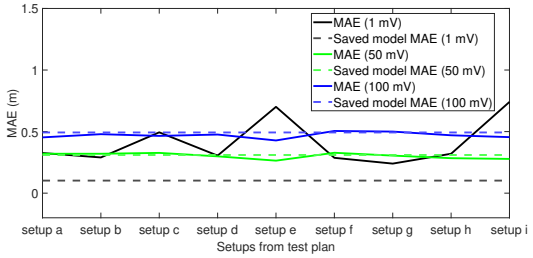
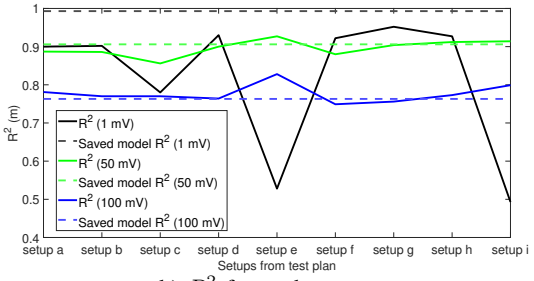


Fig. 18: Test plan for sensitivity analysis: a) setup a, b) setup b, c) setup c, d) setup d, e) setup e, f) setup f, g) setup g, h) setup h, i) setup i.



a) MAE for each setup



b) R^2 for each setup

Fig. 19: MAE and R^2 of the testing set for the power-modulated jammer with the mean power ratio according to each setup of the test plan for the 1 mV, 50 mV and 100 mV noise levels. Results performed with the initial XGBoost trained model (dash line correspond to its own performance).

we take care to avoid any collision between the antennas and the jammer with any objects.

From a machine learning point of view, the process is performed as follows:

- Only the case of the power-modulated jammer using the mean power ratio metric is studied for three level of noise (1 mV, 50 mV, 100 mV).
- The three XGBoost models corresponding to this case with the initial room are saved.
- For each setup of Fig. 18, the training set is built according for three levels of noise (1 mV, 50 mV, 100 mV).
- Each testing set is predicted, using the mean power ratio metric, by its corresponding XGBoost model trained from the data of the initial room.
- MAE and R^2 are computed for each testing test.

The results are presented in Fig. 19. The MAE and R^2 computed (solid lines) are compared to the results (dashed lines) of the testing test of the initial room, illustrated in Table II. This figure shows that, for the dataset with a high noise level (50 mv and 100 mV), overall performances is constant and consistent with the results in Table II. In the other hand, the quality of performance is strongly affected for the very low noise dataset (1mV). Indeed, a small variation due to changes in the environment has enough impact to disrupt the predictions of a model trained with very low noise. In contrast, when changes in the environment are included in the noise, the model's performance becomes virtually insensitive to these changes. In conclusion, the performance limitations are therefore linked to the inclusion of variation due to changes in the environment in the noise learned by the model.

VI. Discussion

Several results have already been presented, mainly from a theoretical point of view. This subsection focuses more on the practical domain. According to the previous results, the best practical case to achieve good predictions is to be in

TABLE V: Estimated training, prediction, and simulation times for different scenarios. Values in bold are measured.

Scenario	Scenario parameters			XGBoost computation time		Simulation time	
	Volume (m^3)	M	L	Training	Prediction	$M = 51$, $L = 3280$	M variable
Small Room	160	51	3 280	**0.25 s**	**22 μs**	**13 s**	**13 s**
Office Floor	480	153	12 240	0.93 s	22 μs	16 – 33 s	48 s
Entire Building	1 600	510	40 800	3.11 s	22 μs	19 – 134 s	190 s

moderately noisy environment. Indeed, if the noise is too high, the predictions lack precision, whereas if the noise is too low, the predictions are very sensitive to changes in the environment. Consequently, artificially adding noise in a very low-noise environment can increase the robustness of our models without affecting the desired performance too much. Moreover, it will also be interesting to study robustness as a function of the different machine learning models.

Additionally, only a few simulations suffice for satisfactory performance, yielding fast computation and learning; post-processing noise is applied to further minimize calculation time. We propose to explore the computational complexity of the proposed methodology for three scenarios including different environment sizes such as small rooms (cf. Fig. 1, office floors, and entire buildings. Their considered volumes are respectively $V_1 = 160$ m^3, $V_2 = 3V_1 = 480$ m^3 and $V_3 = 10V_1 = 1600$ m^3. Firstly, the computational complexity of the SBR model is analyzed under the following assumptions: (i) identical antenna and jammer models across scenarios, (ii) constant reflection order and number of antennas, and (iii) uniform mesh size. Therefore, the computational complexity C_{SBR} can be written:

$$C_{\text{SBR}}(V) \sim O\left(V^{\frac{2}{3}} \cdot \log\left(V^{\frac{1}{3}}\right) + M \cdot R \cdot \log\left(V^{\frac{1}{3}}\right)\right) \quad (12)$$

where V is the volume of the environment under study, M is the number of jammer positions, and R is the number of rays. It is important to note that Matlab automatically sets R based on the antenna model. In fact, R increases with antenna model complexity is to ensure the accuracy of the ray launching process. The SBR method approximates the continuous radiation of an antenna by a discrete set of rays. A simple, omnidirectional antenna (like an isotropic source or a basic monopole) has a smooth pattern that can be well-represented with a relatively low number of rays. However, a complex antenna with high directivity and intricate sidelobe structures (like a yagiUda or horn) requires a much denser sampling of rays to correctly capture its spatial power distribution. Once the antenna model is defined, we will assume that M and R do not vary depending on the scenario. Hence, knowing that on a personal device with an Intel i7-1185G7@3GHz processor, building the database with 100 simulations in the small room scenario requires $ST_{\text{ref}} = 13$ s (compared to one hour in measurement), we are able to estimate the simulation time ST for another volume V_i:

$$ST(V_i) \approx \begin{cases} \dfrac{\log\left(V_i^{\frac{1}{3}}\right)}{\log\left(V_1^{\frac{1}{3}}\right)} \cdot ST_{\text{ref}}, & \text{if } MR \gg V^{\frac{2}{3}}, \\ \dfrac{V_i^{\frac{2}{3}} \log\left(V_i^{\frac{1}{3}}\right)}{V_1^{\frac{2}{3}} \log\left(V_1^{\frac{1}{3}}\right)} \cdot ST_{\text{ref}}, & \text{if } MR \ll V^{\frac{2}{3}}. \end{cases} \quad (13)$$

Table V reports the estimated simulation time from Eq. 13 for scenarios with fixed and varying numbers of jammer positions M. Where M is varying only the case $MR \gg V^{\frac{2}{3}}$ has been estimated. It is clearly highlighted that the computational cost is negligible compared to making measurements and justifies the use of the ray-tracing model as a cost-effective approximation to the jammer location. In addition, the computational times of the XGBoost algorithm for training and single prediction are also considered. The XGBoost computational complexity C_{XGB} is:

$$C_{\text{XGB}} \sim \begin{cases} O(K \cdot F \cdot L \cdot d_{\max}), & \text{for training,} \\ O(K \cdot d_{\max}), & \text{for prediction,} \end{cases} \quad (14)$$

with $K = 300$ the number of trees , $F = 3$ the number of features, L the number of samples, and $d_{max} = 3$ the maximum depth of the threes. Obviously, we note that the prediction time depends solely on K and d_{max} while the training time linearly increase with L as shown in Table V. It is emphasized that training time and prediction time remain negligible compared to simulation time. However, the XGBoost model may require more and deeper trees to perform well in certain scenarios, which could result in longer computation times.

Furthermore, for specific practical applications, the question of how the proposed approach would deal with the complexity of locating the jammer in 2D or 3D could be raised. From the point of view of numerical computation, simulation already performs 3D calculations. Thus, only the number of points M requested will affect the number of iterations and therefore the calculation time, but not its complexity. However, the ML model will be slightly more complex, requiring 3 outputs (i.e. a set of x,y,z coordinates instead of one single x coordinate). Nevertheless, based on our experience, particularly with RF or XGBoost models, the addition of higher dimensions can slightly penalize complexity by making the trees a little deeper, but this remains quite negligible. Consequently, switching to 2D or 3D should not add any significant computation costs.

Finally, the proposed methodology makes it possible to rapidly design the best architecture for the monitoring network.

Indeed, we could perform an optimization with constraints (e.g., on the number of antennas, antenna location, antenna radiation pattern) to obtain the best localization performance.

VII. Conclusion

In this paper, we propose an original methodology, based on a machine learning approach, to localize jammers in indoor environment, which is very complicated because of the multiple reflections in such an environment. Locating jammers is the most effective way of resolving their impact on electronic devices. Indeed, the IEMI associated with jammers is often difficult to mitigate. Our methodology deals with two types of jammer: a constant-power jammer and a power-modulated jammer. Particular attention has been paid to the power-modulated jammer, for which it is not possible to obtain a precise location, even in free space according to [20]. We therefore propose a rigorous framework for accurately locating the jammer from a reliable, rapidly generated database. Indeed, the rapid database construction strategy is based on a SBR-based computational modelling of the environment under study. Such model provides with few seconds the jamming signal power received by a set of three monitoring antennas when the jammer is located nearby. As it is needed to create a database with several positions of the jammer for our machine learning algorithms, the use of such a simulation model avoids a time-consuming measurement process. This claim has been demonstrated through a multiscale analysis of the computational complexity of the whole methodology. Moreover, the advantage of simulation is that parametric studies can be set up quickly. From the simulation model of [21], different databases have been built with different noise levels for a constant-power jammer as well as a power-modulated jammer. Then, two learning algorithms were evaluated (RF and XGBoost) leading to predictions with accuracy of ten of centimeters for both the constant-power jammer and the power-modulated jammer. Compared with conventional trilateration approaches, the machine learning approach performs significantly better. Furthermore, we highlight the introduction of an input ML feature which is not based on the received jamming power itself, but on three power ratios obtained by comparing the mean powers received by each monitoring antenna. During our investigations, the power ratio feature showed a performance gain in comparison with the raw power feature (noted mean power metric) when the jamming power was not learned by the ML model. So, although the mean power ratio metric is more sensitive to noise than the mean power metric, it is still the best option when it comes to locating an unknown power jammer. In addition, the methodology was validated using the measurement as input to the ML model with an acceptable uncertainty, which could probably be further reduced by increasing the accuracy of the simulation model. To demonstrate this assumption, an analysis of uncertainty propagation was performed, leading in particular to the conclusion that the radiation patterns of the antenna network and the jammer must be simulated as accurately as possible. Moreover, we explored the limits of the proposed methodology in terms of dataset size and changes in the test environment. From these analyses, it has been demonstrated that few data is required to obtain accurate results, and that, as long as the effect of the change in the environment is included in the noise, this has little impact on the predictions. Finally, our ML algorithms can provide accurate predictions for relatively low noise levels using the received power of passive antennas in a very short time. In future works, we will focus on the integration into a real protocol, e.g. by exploiting commercial Wi-Fi transceivers. In addition, we will explore the performance of the methodology with other IEMI sources (e.g. high-power electromagnetic pulse). Moreover, we will investigate the performance of a neural network to enable extrapolation of predictions beyond the training set. Last but not least, we will also investigate the ability of the method to be accurate in 2D and 3D.

References

[1] W. A. Radasky, C. E. Baum, and M. W. Wik, "Introduction to the special issue on high-power electromagnetics (HPEM) and intentional electromagnetic interference (IEMI)," *IEEE Transactions on electromagnetic compatibility*, vol. 46, no. 3, pp. 314–321, 2004.

[2] T. Liang, G. Spadacini, F. Grassi, and S. A. Pignari, "Worst-case wideband radiated IEMI for unshielded and shielded cables: A statistical analysis of the main influencing parameters," *IEEE Transactions on Electromagnetic Compatibility*, vol. 61, no. 4, pp. 1244–1252, 2019.

[3] R. R. Tanuhardja, S. van de Beek, M. J. Bentum, and F. B. Leferink, "Vulnerability of terrestrial-trunked radio to intelligent intentional electromagnetic interference," *IEEE transactions on electromagnetic compatibility*, vol. 57, no. 3, pp. 454–460, 2015.

[4] N. Mora, I. D. Flintoft, L. Dawson, J. F. Dawson, F. Rachidi, M. Rubinstein, A. C. Marvin, P. Bertholet, and M. Nyffeler, "Experimental characterization of the response of an electrical and communication raceway to IEMI," *IEEE Transactions on Electromagnetic Compatibility*, vol. 58, no. 2, pp. 494–505, 2016.

[5] S. Mili, D. Sodoyer, V. Deniau, M. Heddebaut, H. Philippe, and F. Canavero, "Recognition process of jamming signals superimposed on GSM-R radiocommunications," in *2013 International Symposium on Electromagnetic Compatibility*, 2013, pp. 45–50.

[6] M. K. Hanawal, D. N. Nguyen, and M. Krunz, "Jamming attack on in-band full-duplex communications: Detection and countermeasures," in *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, 2016, pp. 1–9.

[7] A. N. De São José, V. Deniau, C. Gransart, T. Vantroys, A. Boé, and E. P. Simon, "Susceptibility of LoRa communications to intentional electromagnetic interference with different sweep periods," *Sensors*, vol. 22, no. 13, 2022.

[8] S. Zhao, Z. Lu, Z. Luo, and Y. Liu, "Orthogonality-sabotaging attacks against OFDMA-based wireless networks," in *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019, pp. 1603–1611.

[9] V. Deniau, C. Gransart, G. L. Romero, E. P. Simon, and J. Farah, "IEEE 802.11n communications in the presence of frequency-sweeping interference signals," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 5, pp. 1625–1633, 2017.

[10] R. Morales-Ferre, P. Richter, E. Falletti, A. de la Fuente, and E. S. Lohan, "A survey on coping with intentional interference in satellite navigation for manned and unmanned aircraft," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 249–291, 2020.

[11] D. Borio, F. Dovis, H. Kuusniemi, and L. Lo Presti, "Impact and detection of GNSS jammers on consumer grade satellite navigation receivers," *Proceedings of the IEEE*, vol. 104, no. 6, pp. 1233–1245, 2016.

[12] "Electromagnetic compatibility (EMC) — Part 4-3: Testing and measurement techniques — Radiated, radio-frequency, electromagnetic field immunity test," International Electrotechnical Commission, Geneva, Switzerland, Standard IEC 61000-4-3:2020, 2020.

[13] Q. Wang, X. Wei, J. Fan, T. Wang, and Q. Sun, "A step further of PDR-based jammer localization through dynamic power adaptation," in *11th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2015)*. IET, 2015.

[14] Y. S. Kim, F. Mokaya, E. Chen, and P. Tague, "All your jammers belong to us—localization of wireless sensors under jamming attack," in *2012 IEEE International Conference on Communications*, 2012, pp. 949–954.

[15] K. Pelechrinis, I. Koutsopoulos, I. Broustis, and S. V. Krishnamurthy, "Lightweight jammer localization in wireless networks: System design and implementation," in *GLOBECOM 2009-2009 IEEE Global Telecommunications Conference*. IEEE, 2009, pp. 1–6.
[16] Z. Liu, H. Liu, W. Xu, and Y. Chen, "Exploiting jamming-caused neighbor changes for jammer localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 3, pp. 547–555, 2011.
[17] T. Van Nguyen, Y. Jeong, H. Shin, and M. Z. Win, "Machine learning for wideband localization," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1357–1380, 2015.
[18] Y.-K. Cheng, H.-J. Chou, and R. Y. Chang, "Machine-learning indoor localization with access point selection and signal strength reconstruction," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. IEEE, 2016, pp. 1–5.
[19] A. H. Salamah, M. Tamazin, M. A. Sharkas, and M. Khedr, "An enhanced WiFi indoor localization system based on machine learning," in *2016 International conference on indoor positioning and indoor navigation (IPIN)*. IEEE, 2016, pp. 1–8.
[20] G. O. Pietro Tedeschi and R. D. Pietro, "Localization of a power-modulated jammer," *Sensors*, 2022.
[21] P. Monferran, A. Costanzo, A. N. de São José, V. Deniau, and C. Gransart, "Machine learning techniques for the geolocalization of jamming sources in indoor wireless networks," in *International Symposium and Exhibition on Electromagnetic Compatibility (EMC Europe 2024)*, 2024.
[22] EM-6116 *Omni-Directional Antenna*, ELECTRO-METRICS, Ref: 100513. [Online]. Available: https://em-antennas.com/product/omni-directional-antenna-em-6116-2-ghz-10-ghz/
[23] H. Ling, R.-C. Chou, and S.-W. Lee, "Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity," *IEEE Transactions on Antennas and Propagation*, vol. 37, no. 2, pp. 194–205, 1989.
[24] "Ieee standard for definitions of terms for antennas," *IEEE Std 145-2013 (Revision of IEEE Std 145-1993)*, pp. 1–50, 2014.
[25] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.
[26] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, pp. 197–227, 2016.
[27] T. Chen and C. Guestrin, "XGBoost," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, aug 2016.
[28] A. Costanzo and V. Loscri, "Error compensation in indoor positioning systems based on software defined visible light communication," *Physical Communication*, vol. 34, pp. 235–245, 2019.
[29] T. M. Inc. Antenna catalog. [Online]. Available: https://www.mathworks.com/help/antenna/antenna-catalog.html
[30] D. Kurup, W. Joseph, G. Vermeeren, and L. Martens, "In-body path loss model for homogeneous human tissues," *IEEE Transactions on Electromagnetic Compatibility*, vol. 54, no. 3, pp. 556–564, 2012.

Paul Monferran was born in Paris, France, in 1992. He received the master's degree in sensors, measurement and instrumentation from the University Pierre-et-Marie Curie, Paris, France, in 2015 and the Ph.D. degree in electronics from the XLIM Laboratory, University of Limoges, Limoges, France, in 2018. He is currently a Postdoctoral Fellow with the LEOST Laboratory, University Gustave Eiffel, Villeneuve d'Ascq, France. His current research interests include the electromagnetic characterization of several materials, lightning effects in aircraft for electromagnetic compatibility issues, and machine learning applied to EMC issues.

Antonio Costanzo received his Ph.D. degree in systems engineering and computer science at the Microwave Laboratory of the University of Calabria (Italy) in 2015. In 2017, he joined INRIA Lille-Nord Europe, France, as a Development Engineer in optical wireless networks. He is currently an assistant professor at ESAIP Engineering School (France). His current research interests include free space optics, signal processing, electromagnetic compatibility, and physical cybersecurity.

Artur Nogueira de São José received the Ph.D. degree in electrical engineering from the Federal University of Minas Gerais, Belo Horizonte, Brazil, in 2020. He is currently an Assistant Professor with the Department of Electrical Engineering, University of Brasilia, Brasilia, Brazil. His main research interests include electromagnetic compatibility, wireless communications, and cybersecurity.

Jonathan Villain obtained his PhD in applied mathematics from the Université de Bretagne-Sud, Vannes, France, in 2016 in a French mathematics laboratory (LMBA, Vannes) and a pharmaceutical laboratory (CERMN, Caen). His research focuses on machine learning and pattern recognition modeling in fields as varied as the study of pharmaceutical toxicity and electromagnetic security. Since 2018, he has applied his machine learning expertise in various projects (Shift2Rail, ELSAT2020, SECOURT, LoRa R, GLOCAT and DEPOSIA) studying electromagnetic disturbance classification models. He is currently attached to the Université Gustave Eiffel to continue his work on the Resilient Trust project.

Virginie Deniau received the M.S. and Ph.D. degrees in electronics from the University of Lille in 2000 and 2003, respectively. Since 2003, she is Researcher in electromagnetic compatibility (EMC) for the University Gustave Eiffel. She conducts works on electromagnetic compatibility (EMC) for land transport. Her research interests include EMC test facilities and methodologies, characterization and modeling of electromagnetic transport environments and the immunity test methodologies for embedded systems. Currently, she works in the hardening of land transport systems regarding cyberattacks, such electromagnetic attacks. She has participated in numerous national and European projects and she was scientific coordinator of the Eu project SECRET for SECurity of Railways against Electromagnetic aTtacks. She is vice-chair of the URSI Committee E (Electromagnetic Interference).

Christophe Gransart received the Ph.D. degree from the University of Lille, Villeneuve-d'Ascq, France, in 1995. He is a Senior Researcher with French Institute of Science and Technology for Transport, Development, and Networks, Villeneuve d'Ascq, with 15 years experience in participating in industrial and academic research projects dealing distributed systems and middleware for transportation systems, V2V and V2I communications, adaptive middleware and cybersecurity. He was involved in various national and European projects. The main competencies are computer science, distributed architecture design, and middleware expertise. He participated to FP6, FP7, H2020, Shift2Rail programs.

Valeria Loscri (M'03, SM'17) (http://researchers.lille.inria.fr/ loscri/home.html) is research director at Inria Lille (France). She received her MSc and PhD degrees in Computer Science in 2003 and 2007, respectively, from the University of Calabria and her HDR (Habilitation à diriger des recherches) in 2018 from Université de Lille (France). Her research interests focus on emerging technologies for new communication paradigms such as Visible Light Communication (VLC), mmWave, cyber security in wireless networks and cooperation and coexistence of wireless heterogeneous devices. In 2021 she has been nominated Women Stars in Computer Networking and Communications by the IEEE Communication Society. She has been shortlisted as Cyber Researcher by European Women Cyber Day - ECWD in 2024. She is Action Chair and Scientific Holder of BEiNG-WISE COST Action (https://beingwise.eu/) (since 2023). In 2024 she has been appointed as chair of the expert panel for fundamental research W&T5 Computer Science & Information Technology FWO. Since 2019, she is Scientific International Delegate for Inria Lille.