

Machine Learning For In-Region Location Verification In Wireless Networks

Alessandro Brighente, Francesco Formaggio, Giorgio Maria Di Nunzio, and Stefano Tomasin

Abstract

In-region location verification (IRLV) aims at verifying whether a user is inside an specific region, and in wireless networks it can exploit the features of the channel between the user to be verified and the set of trusted access points. As IRLV is an hypothesis testing problem we can resort to the Neyman-Pearson (N-P) theorem, when channel statistics are known. By removing the channel knowledge assumption we instead consider machine learning (ML) solutions based on either neural networks (NNs) or support vector machine (SVM) to learn channel features characteristics of the region of interest. We show that ML approaches are N-P-optimal for sufficiently complex machines and long enough training. For finite training ML turns out to be more accurate than then N-P applied on estimated channel statistics. Two specific security issues are then addressed. First, the attacker can choose the position from which the attack is done, thus changing its statistics and we exploit one-class classification, concluding that conventional ML solutions based on the auto-encoder and one-class SVM do not perform the generalized likelihood ratio test (GLRT), even under asymptotic conditions. Second, advanced attacks based on the use of ML techniques by the attacker are considered, in order to select the location from which the attack is performed and reduce the number of attacks before success. Numerical results are provided to support the results in a realistic cellular scenario, including shadowing and fading effects for the channel description.

Index Terms

In-region location verification, machine learning, neural network, auto-encoder, support vector machine.

I. INTRODUCTION

Location verification systems aim at verifying the location of the user, which can be used to implement location-based granting systems, e.g., location-based access control, media streaming, and social networking. Location information without verification gives ample opportunities to attack the service granting system, since the location information can be easily manipulated either by tampering the hardware/software reporting the location or by spoofing the global navigation satellite systems (GNSS) signal outside the user device. Location verification systems aim at verifying the position of devices in a mobile communication network, with applications in sensor networks [25], [6], [21], the Internet of Things [7], and geo-specific encryption solutions [14]. In the literature various solutions are available for location verification, that leverage the features of the wireless channel over which communication occurs. One approach provides the measurement of the distance between the user and other network nodes. An example is given by [23], where received signal strength (RSS) is exploited. Moreover, location verification is similar to the *user authentication* problem addressed at the physical layer, where channel measurements are processed to verify the identity of the message sender [12].

We focus here on in-region location verification (IRLV) that aims at verifying whether a user is inside an specific region of interest (ROI), by exploiting the features of the channel between the user to be verified and a set of trusted access points (APs) [25]. Among solutions presented in the literature, distance bounding techniques with rapid exchanges of packets between the verifier and the prover have been proposed in [5], also using radio-frequency and ultrasound signals [16], whereas solutions based on the use of anchor nodes and increasing transmit power by the sender have been proposed in [20]. More recently, a delay-based verification technique has been proposed in [2], leveraging geometric properties of triangles, which prevents an adversary from manipulating measured delays.

Some of the proposed techniques partially neglect wireless channel phenomena (such as shadowing and fading) that make problematic the estimation of distance measurements through the wireless communication signal ([5], [16], [20]). Other approaches instead assume specific channel statistics [14] that again may be not accurate due to changing environment conditions, their dependency on unknown parameters and the use of advanced signal processing techniques by the attacker. Indeed, if the statistics of channels of links to devices both inside and outside the

In Dec 2018 this work has been submitted to IEEE Journal on Selected Areas in Communications.

ROI are known to the network infrastructure, the optimal solution for IRLV is provided by the Neyman-Pearson (N-P) theorem [9], that provides the most powerful test for a given significance level.

However, as observed, the knowledge of the channel statistics may not be available. One solution in this case is to a) estimate the channel statistics and b) apply the N-P theorem on the estimated statistics. However, as confirmed also in this paper this approach may not be accurate. As an alternative, machine learning (ML) techniques can be exploited, since they are more flexible and exploit the available training data to optimize directly the decision process, rather than going through the two-step solution. In [22], no assumption is made on the channel model and logistic regression has been proposed as an alternative to hypothesis testing. In [19] the objective is to locate the user inside a building and a multi-class classification problem is solved via support vector machine (SVM). Note that neither [22] nor [19] analyze performance of their ML approaches against the theoretical optimum N-P criterion.

In this paper, by removing the channel knowledge assumption we instead consider ML solutions based on either neural networks (NNs) or SVM to learn channel features of the ROI. For NN solution we investigate two loss functions for the multy-layer perceptron (MLP) design: the cross entropy (CE) and the mean squared error (MSE), while for SVM we consider its least-squares (LS) implementation. We show that these ML approaches are N-P-optimal for sufficiently complex machines and long enough training. For finite training ML turns out to be more accurate than N-P applied on estimated channel statistics.

Two specific security issues are then addressed. First, the attacker can select the location for the attack and its statistics can be modified, thus we explore the one-class classification, both under the knowledge of legitimate channel statistics and by ML, concluding that conventional ML solutions based on both the auto encoder (AE) and one-class SVM do not coincide with the generalized likelihood ratio test (GLRT), even under asymptotic training conditions. Second, advanced attacks based on the use of ML techniques by the attacker are considered, to select the positions of attack. Numerical results are provided to support the results in a realistic network scenario, including shadowing and fading effects for the channel. We show that in a simple scenario a small number of neurons and a short training already provide close-to-optimal performance, and then assess the one-class IRLV approach, as well as the advanced attack strategy based on ML.

The contributions of this paper are here summarized:

- We propose a physical layer-based IRLV system which exploits ML techniques to perform hypothesis testing;
- We show that two-class classification ML techniques provide the most powerful test for a given significance level, as the N-P test;
- We compare different ML techniques and show how these can be exploited to implement both IRLVs and efficient attacks.

The rest of the paper is organized as follows. Section II introduces the system model for the IRLV problem, detailing the channel model and summarizing the N-P theorem under the knowledge of channel statistics. Section III describes the considered ML approaches for IRLV, namely the MLP NN and the SVM, with their design approaches based on MSE or CE for the NN, and on LS for SVM. In Section IV we propose the one-class classification approach, where only the channel statistics for locations inside the ROI are used for training: both the AE NN and the one-class least-square SVM (OCLSSVM) are considered and their performance is compared with the GLRT test. Advanced attack strategies, based on the use of ML by the attacker are proposed in Section V. Numerical results comparing the various techniques are shown and discussed in Section VI. Lastly, the main conclusions are outlined in Section VII.

The following notation is used throughout the paper: bold letters \mathbf{x} refer to vectors, whereas capital bold letters \mathbf{H} refer to matrices, $\mathbb{E}[x]$ denotes the expectation of random variable x , $(\cdot)^T$ denotes the transpose. $\mathbb{P}[A]$ denotes the probability of event A and $\ln(x)$ denotes the natural base logarithm of x .

II. SYSTEM MODEL

With reference to Fig. 1 we consider a cellular system with N_{AP} APs covering the area \mathcal{A} over a plane. We propose a IRLV system to determine if a user equipment (UE) is transmitting from within an *authorized* ROI \mathcal{A}_0 inside \mathcal{A} . The complementary region to the ROI is $\mathcal{A}_1 = \mathcal{A} \setminus \mathcal{A}_0$. The authentication process exploits the location dependency of the features of the channel between the UE and the APs. For the sake of a simpler exposition we consider a narrowband transmission and we choose the power attenuation as the feature on which we perform IRLV. Extensions of the technique to more elaborate channel features is straightforward and its evaluation is left for future study.

The exploitation of the channel features is obtained by letting the UE transmit a training signal with fixed power P_{tx} , known at the APs, from which the APs can estimate the received

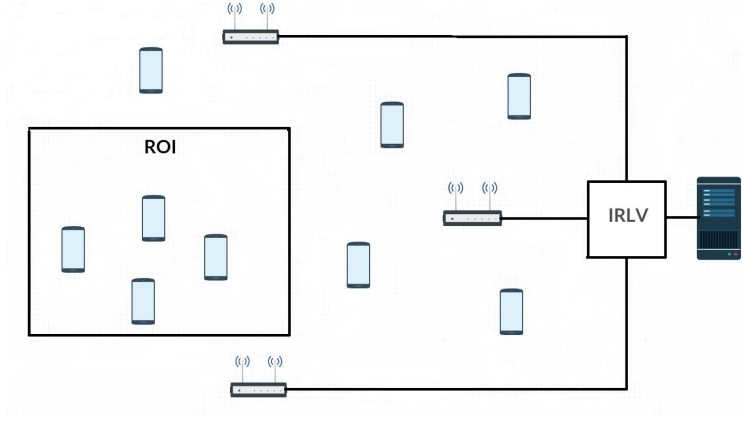


Fig. 1. The considered IRLV scenario.

power, thus obtaining a measure of the attenuation incurred by the channel. We assume that the attenuation estimation is perfect, i.e., not affected by noise or interference: this can be achieved by using a long enough training signal.

A. Channel Model

When the UE transmits with power P_{tx} , the received power at the n^{th} AP is

$$P_{\text{rc}}^{(n)} = \frac{P_{\text{tx}}}{a^{(n)}}, \quad (1)$$

where $a^{(n)}$ is the attenuation incurred over the channel between the UE and the AP n . The channel model for path-loss and shadowing is derived from [1]. The attenuation coefficient includes the effects of path-loss, shadowing and fading. In particular, by assuming Rayleigh model for fading we have

$$\sqrt{a^{(n)}} \sim \mathcal{N}(0, \sigma_{a,n}^2), \quad (2)$$

where $\sigma_{a,n}^2 = P_{\text{PL}}^{(n)} e^s$ accounts for the path-loss and shadowing components, $P_{\text{PL}}^{(n)}$ is the path-loss coefficient and $s \sim \mathcal{N}(0, \sigma_s^2)$ is the shadowing component.

Both path-loss and shadowing models are derived from [1]. In particular, let us denote as $\mathbf{x}_{\text{AP}}^{(n)} = (X_{\text{AP}}^{(n)}, Y_{\text{AP}}^{(n)})$ the position of AP $n = 1, \dots, N_{\text{AP}}$. For a UE located at $\mathbf{x}_{\text{UE}} = (X_u, Y_u)$, its distance from AP n is $L(\mathbf{x}_{\text{UE}}, \mathbf{x}_{\text{AP}}^{(n)}) = \sqrt{(X_{\text{AP}}^{(n)} - X_u)^2 + (Y_{\text{AP}}^{(n)} - Y_u)^2}$. For the path-loss we consider two scenarios: line of sight (LOS) and non-LOS. For a LOS link the path-loss in dB is modelled as

$$P_{\text{PL,LOS}}^{(n)} = 10\nu \log_{10} \left(\frac{f4\pi L(\mathbf{x}_{\text{UE}}, \mathbf{x}_{\text{AP}}^{(n)})}{c} \right), \quad (3)$$

where ν is the path-loss coefficient, f is the carrier frequency and c is the speed of light. For a non-LOS link the path-loss coefficient in dB is defined as

$$P_{\text{PL,NLOS}}^{(n)} = 40 \log_{10} \left(\frac{L(\mathbf{x}_{\text{UE}}, \mathbf{x}_{\text{AP}}^{(n)})}{10^3} \right) + 21 \log_{10} \left(\frac{f}{10^6} \right) + 80. \quad (4)$$

The shadowing parameter s is zero-mean Gaussian distributed with power σ_s^2 . Moreover, the shadowing parameters of two UEs located at positions \mathbf{x}_i \mathbf{x}_j and transmitting to the same AP have correlation $\sigma_s^2 e^{-\frac{L(\mathbf{x}_i, \mathbf{x}_j)}{d_c}}$, where d_c is the shadowing decorrelation distance.

Path-loss and shadowing are assumed to be time-invariant, while the fading component is independent at each attenuation estimate.

B. IRLV With Known Channel Statistics

The IRLV problem can be seen as an hypothesis testing between the two hypothesis (events):

- \mathcal{H}_0 : the UE is transmitting from area \mathcal{A}_0 ;
- \mathcal{H}_1 : the UE is transmitting from area \mathcal{A}_1 .

Note that this problem is also denoted as two-class classification. Given vector $\mathbf{a} = [a^{(1)}, \dots, a^{(N_{\text{AP}})}]$ collecting the attenuation estimates at all the APs, we aim at determining the most likely hypothesis, in order to perform IRLV. While a few measurements of path-loss would allow by triangulation to establish the exact position of the UE, both shadowing and fading make the IRLV more problematic and in general prone to errors. Let us indicate with $\hat{\mathcal{H}} \in \{\mathcal{H}_0, \mathcal{H}_1\}$ the decision taken at the APs on the two hypothesis, and let $\mathcal{H} \in \{\mathcal{H}_0, \mathcal{H}_1\}$ the ground true, i.e., the effective location of the UE. We have two possible errors: false alarms (FAs), which occur when the UE is classified as outside the ROI, while being inside it, and mis-detections (MDs), which occur when the UE is classified as inside the ROI, while being outside of it. We also indicate the FA probability as $P_{\text{FA}} = P(\hat{\mathcal{H}} = \mathcal{H}_1 | \mathcal{H} = \mathcal{H}_0)$ and the MD probability as $P_{\text{MD}} = P(\hat{\mathcal{H}} = \mathcal{H}_0 | \mathcal{H} = \mathcal{H}_1)$.

Now, let $p(\mathbf{a} | \mathcal{H}_i)$ be the probability of estimating the vector \mathbf{a} given that $\mathcal{H} = \mathcal{H}_i$. The log likelihood-ratio (LLR) for the considered hypothesis is defined as

$$\mathcal{M}(\mathbf{a}) = \ln \frac{p(\mathbf{a} | \mathcal{H}_0)}{p(\mathbf{a} | \mathcal{H}_1)}. \quad (5)$$

According to the N-P theorem, the most powerful test is obtained by comparing $\mathcal{M}(\mathbf{a})$ with a threshold value Λ , i.e., obtaining the test function

$$f^*(\mathbf{a}) = \begin{cases} -1 & \text{if } \mathcal{M}(\mathbf{a}) \geq \Lambda, \\ 1 & \text{if } \mathcal{M}(\mathbf{a}) < \Lambda. \end{cases} \quad (6)$$

This procedure provides the minimum MD probability for a given FA probability.

C. Example of N-P Test

We provide here an example of application of the N-P theorem to a simplified scenario. Consider the overall network area as a circle \mathcal{A} with radius R_{out} and assume a single AP ($N_{\text{AP}} = 1$) located at the center of \mathcal{A} . The ROI \mathcal{A}_0 is a rectangle of height H and length L and with nearest point to the center of \mathcal{A} at a distance R_{min} . The channel model includes only path-loss (without shadowing and fading), in a LOS scenario, therefore a is given by (3). We notice that the attenuation is a deterministic function of the distance R , hence we obtain R from the attenuation a as

$$R = \frac{c\sqrt[4]{a}}{f4\pi} \quad (7)$$

The probability that the UE is located at a distance $R \leq R_0$ in \mathcal{A}_0 is

$$F(R_0) = \mathbb{P}(R \leq R_0 | \mathcal{A}_0) = \frac{1}{|\mathcal{A}_0|} \int_{R_{\text{min}}}^{R_0} \rho \zeta(\rho) d\rho, \quad (8)$$

where $\zeta(\rho)$ denotes the angle of the circular sector located at distance ρ intersecting area \mathcal{A}_0 . By taking the derivative of (8) respect to R_0 we obtain the probability density function (PDF) of R_0 given that the UE is located in \mathcal{A}_0 , i.e.,

$$p_{R|\mathcal{A}_0}(R_0 | \mathcal{A}_0) = \frac{1}{|\mathcal{A}_0|} R_0 \zeta(R_0). \quad (9)$$

Following the same reasoning and considering that the length of the arc of circle with radius R_0 located in \mathcal{A}_1 is $2\pi - \zeta(R_0)$, we obtain the PDF of R given that the UE is located in \mathcal{A}_1 as

$$p_{R|\mathcal{A}_1}(R_0 | \mathcal{A}_1) = \frac{1}{|\mathcal{A}_1|} R_0 (2\pi - \zeta(R_0)), \quad (10)$$

From (7), (9) and (10) we obtain the LLR as a function of the UE's distance from the AP as

$$\mathcal{M}(a) = \frac{|\mathcal{A}_1| \zeta\left(\frac{c\sqrt[4]{a}}{f4\pi}\right)}{|\mathcal{A}_0| \left(2\pi - \zeta\left(\frac{c\sqrt[4]{a}}{f4\pi}\right)\right)}. \quad (11)$$

III. IRLV BY MACHINE LEARNING APPROACHES

The application of the N-P theorem requires the knowledge of the conditional channel statistics $p(\mathbf{a}|\mathcal{H}_i)$ at the APs, which can be hard to obtain, also because a-priori assumptions on their expressions may be quite unrealistic. Therefore, we propose to use a ML approach operating in two phases:

- *Learning phase*: the AP collects attenuation vectors from a trusted UE moving both inside and outside the ROI, and the UE reports its position to the APs. Therefore the APs can learn the behaviour of the attenuation in both regions \mathcal{A}_0 and \mathcal{A}_1 .
- *Exploitation phase*: the AP verifies the location of an un-trusted UE by the attenuation's estimate using the experience acquired in the learning phase.

In details, the learning phase works as follows. For training attenuation vectors $\mathbf{a}^{(i)}$, $i = 1, \dots, S$, collected during the learning phase, there are associated identification values t_i , $i = 1, \dots, S$, where $t_i = -1$ if the trusted UE is in region \mathcal{A}_0 and $t_i = 1$ if the trusted UE is in region \mathcal{A}_1 . Vector $\mathbf{t} = [t_1, \dots, t_S]$ collects the labels of all the attenuation vectors in the training phase. By these sets, the AP design a test function $\hat{t} = f(\mathbf{a}) \in \{-1, 1\}$ that provides a decision $\hat{\mathcal{H}}$ for each attenuation vector \mathbf{a} . Then in the exploitation phase the IRLV algorithm computes $\hat{t} = f(\mathbf{a})$ for the new attenuation vectors thus taking a decision between the two hypotheses. Note that our solution does not explicitly evaluate the PDF and the LLR, rather it directly implements the test function.

In the rest of this Section we briefly review the MLP NN and the SVM, describe the learning process and show that in asymptotic conditions (infinite training attenuation vectors and complex machines) both MLP and SVM functions approximate the LLR function.

A. Neural Networks

A NN is a $\mathbb{R}^N \rightarrow \mathbb{R}^O$ function mapping a set of N real values into O real values. A NN processes the input in Q stages, named layers, where the output of one layer is the input of the next layer. Layer 0 with input $\mathbf{y}^{(0)}$ is denoted *input layer*, while layer $Q - 1$ with output $\mathbf{y}^{(Q)}$ is denoted *output layer*, while intermediate layers are denoted *hidden layers*.

Layer $\ell = 0, \dots, Q - 1$ has $N^{(\ell)}$ outputs obtained by processing the inputs with $N^{(\ell-1)}$ functions named neurons. The output of the n^{th} neuron of the ℓ^{th} layer is

$$y_n^{(\ell+1)} = \psi^{(\ell)} \left(\mathbf{w}_n^{(\ell)} \mathbf{y}^{(\ell)} + b_n^{(\ell)} \right), \quad (12)$$

i.e., the mapping by the *activation function* $\psi^{(\ell)}(\cdot)$ of the weighted linear combination with weights $\mathbf{w}_n^{(\ell)}$ of the outputs $\mathbf{y}^{(\ell)}$ of the previous layer plus a bias $b_n^{(\ell)}$. We focus here on feedforward NNs, i.e., without loops between neurons' input and output, an architecture also known as MLP. For a in-depth description of NNs refer for example to [10]. While the activation functions are typically fixed, the vectors $\mathbf{w}_n^{(\ell)}$ must be properly chosen to perform the desired hypothesis testing.

In our setting, the input to the NN is the attenuation vector \mathbf{a} , thus $N = N_{AP}$ and the output layer has a single neuron ($O = 1$) providing as output the scalar $y_1^{(Q)}$. Let $\tilde{t}(\mathbf{a}) = y_1^{(Q-1)}$ be the output of the NN corresponding to the attenuation vector input \mathbf{a} . The test function performs a threshold of $\tilde{t}(\mathbf{a})$, i.e.,

$$f(\mathbf{a}) = \begin{cases} 1 & \tilde{t}(\mathbf{a}) > \lambda, \\ -1 & \tilde{t}(\mathbf{a}) \leq \lambda. \end{cases} \quad (13)$$

Different values of λ provide different values of FA and MD probabilities for the resulting IRLV test.

B. NN MSE Design

According to the MSE design criterion [10], during the learning phase the MLP parameters are designed in order to minimize the MSE

$$\Gamma = \sum_{i=1}^S |\tilde{t}(\mathbf{a}^{(i)}) - t_i|^2. \quad (14)$$

This is achieved by using the stochastic gradient descent algorithm [3].

We now prove the connection of MSE design criterion with the N-P theorem.

Theorem 1. *The MLP designed according to the MSE criterion with an infinite set of training points ($S \rightarrow \infty$) and thresholding (13) converges to a test function equivalent to the N-P test, thus is the most powerful test.*

Proof. It has been shown in [15] that a MLP trained via MSE implements a function that is the minimum MSE approximation of the Bayes optimal discriminant function

$$g_0(\mathbf{a}) = \mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}) - \mathbb{P}(\mathcal{H} = \mathcal{H}_1|\mathbf{a}). \quad (15)$$

By recalling that $\mathbb{P}(A|B) = \mathbb{P}(B|A)\mathbb{P}(A)/\mathbb{P}(B)$ for any events A and B , we can write

$$g_0(\mathbf{a}) = \frac{p(\mathbf{a}|\mathcal{H}_0)\mathbb{P}(\mathcal{H} = \mathcal{H}_0) - p(\mathbf{a}|\mathcal{H}_1)\mathbb{P}(\mathcal{H} = \mathcal{H}_1)}{\mathbb{P}(\mathbf{a})}, \quad (16)$$

which in turn can be written as

$$g_0(\mathbf{a}) = \frac{p(\mathbf{a}|\mathcal{H}_0)\mathbb{P}(\mathcal{H} = \mathcal{H}_0) - p(\mathbf{a}|\mathcal{H}_1)\mathbb{P}(\mathcal{H} = \mathcal{H}_1)}{p(\mathbf{a}|\mathcal{H}_0)\mathbb{P}(\mathcal{H} = \mathcal{H}_0) + p(\mathbf{a}|\mathcal{H}_1)\mathbb{P}(\mathcal{H} = \mathcal{H}_1)}. \quad (17)$$

The threshold function (13) imposes a threshold λ on $g_0(\mathbf{a})$ and reorganizing terms we obtain $f(\mathbf{a}) = -1$ when

$$\frac{p(\mathbf{a}|\mathcal{H}_0)}{p(\mathbf{a}|\mathcal{H}_1)} > \frac{1 + \lambda}{1 - \lambda} \frac{\mathbb{P}(\mathcal{H} = \mathcal{H}_1)}{\mathbb{P}(\mathcal{H} = \mathcal{H}_0)} = \lambda^*, \quad (18)$$

which is equivalent to the N-P criterion, apart from a fixed scaling of the threshold. \square

C. NN CE Design

Another design criterion for the NN is based on the CE [10]

$$\chi = - \sum_{i=1}^S t_i \ln(\tilde{t}(\mathbf{a}^{(i)})) + (1 - t_i) \ln(1 - \tilde{t}(\mathbf{a}^{(i)})). \quad (19)$$

We now prove the connection of CE design criterion with the N-P theorem.

Theorem 2. *The MLP designed according to the CE criterion with an infinite set of training points ($S \rightarrow \infty$) and thresholding (13) converges to a test function equivalent to the N-P test, thus is the most powerful test.*

Proof. The probability of being in hypothesis \mathcal{H}_1 given that the attenuation vector is \mathbf{a} satisfies

$$\mathbb{P}(\mathcal{H} = \mathcal{H}_1|\mathbf{a}) = 1 - \mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}). \quad (20)$$

When training is performed with CE loss function the output of the MLP is the minimum MSE approximation of the probability $\mathbb{P}(\mathcal{H}_0|\mathbf{a}^{(i)})$ of being in hypothesis \mathcal{H}_0 given that the attenuation vector is \mathbf{a} [3], i.e.,

$$\tilde{t}(\mathbf{a}) \approx \mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}), \quad (21)$$

where the approximation is in the MSE sense.

Now, by using the threshold function (13) we have

$$\mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}) \approx \tilde{t}(\mathbf{a}) \gtrsim \lambda, \quad (22)$$

which can be rewritten as

$$2\mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}) - 1 \gtrsim \hat{\lambda} \quad (23)$$

$$\mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}) - (1 - \mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a})) \gtrsim \hat{\lambda} \quad (24)$$

$$\mathbb{P}(\mathcal{H} = \mathcal{H}_0|\mathbf{a}) - \mathbb{P}(\mathcal{H} = \mathcal{H}_1|\mathbf{a}) \gtrsim \hat{\lambda}. \quad (25)$$

The using (13) on the output of the NN designed with the CE criterion we obtain at convergence a function performing (25), which in turn coincides (apart for a different threshold value) with the function performed by the NN trained with the MSE criterion, as shown in (15). Therefore, from Theorem 1 we conclude that also the CE design criterion provides a test function equivalent the N-P test function. \square

D. Support Vector Machine

A SVM [3] is a supervised learning model that can be used for classification and regression. We focus here on binary classification to solve the IRLV problem. The SVM solution comprises the $\tilde{t}(\mathbf{a}) : \mathbb{R}^{N_{AP}} \rightarrow \mathbb{R}$ function

$$\tilde{t}(\mathbf{a}) = \mathbf{w}^T \phi(\mathbf{a}) + b, \quad (26)$$

where $\phi : \mathbb{R}^{N_{AP}} \rightarrow \mathbb{R}^K$ is a feature-space transformation function, $\mathbf{w} \in \mathbb{R}^K$ is the weight vector and b is a bias parameter. The test function is again provided by (13) where now $\tilde{t}(\mathbf{a})$ is given by (26). Note that in the conventional SVM formulation we have $\lambda = 0$, while here λ is chosen according to the desired FA and MD probabilities.

While the feature-space transformation function is typically fixed [10], vector \mathbf{w} must be properly chosen to perform the desired hypothesis testing.

E. SVM LS Design

For the the choice of the SVM parameters we consider the least squares SVM (LS-SVM) approach [18]. Learning for LS-SVM is performed by solving the following optimization problem

$$\min_{\mathbf{w}, b} \quad \omega(\mathbf{w}, b) \triangleq \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \frac{1}{2} \sum_{i=1}^S e_i^2 \quad (27a)$$

$$e_i = t_i[\mathbf{w}^T \phi(\mathbf{a}^{(i)}) + b] - 1 \quad i = 1, \dots, S, \quad (27b)$$

where C is a hyper-parameter. In conventional SVM, variables e_i are constrained to be non-negative and appear in the objective function without squaring. Inequalities in the constraints translates into a quadratic programming problem, while equalities constraints in LS-SVM yield a linear system of equations in the optimization values. In [24] it is shown that SVM and LS-SVM are equivalent under mild conditions.

From constraints (27) and the fact that $t_i = \pm 1$ we have

$$e_i^2 = (1 - t_i \tilde{t}(\mathbf{a}^{(i)}))^2 = (t_i - \tilde{t}(\mathbf{a}^{(i)}))^2, \quad (28)$$

that is the squared error between the soft output of the LS-SVM $\tilde{t}(\mathbf{a}^{(i)})$ and the correct training label t_i .

We now prove the equivalence between the LS-SVM and N-P classifiers. Let us first consider the following lemma that establishes the convergence of the learning phase of SVM, as $S \rightarrow \infty$.

Lemma 1. *For large number of training samples $\mathbf{a}^{(i)}$ taken with a given static probability distribution from a finite alphabet \mathcal{C} , i.e., for $S \rightarrow \infty$, the vector \mathbf{w} of the LS-SVM converges in probability to a vector of finite norm $\|\mathbf{w}\|_2 = \mathbf{w}^T \mathbf{w}$.*

Proof. See the Appendix. □

We are now ready to prove the the following theorem establishing the optimality of the SVM solution, as it provides the most powerful N-P test, for a given FA probability.

Theorem 3. *Consider a LS-SVM with perfect training, i.e., the training reaches a global minimum of $\omega(\mathbf{w}, b)$ given an infinite number of training points $\mathbf{a}^{(i)}$ drawn from the finite alphabet \mathcal{C} . Then the test function obtained by training the LS-SVM and by thresholding the soft output (26) converges to the N-P test, thus is the most powerful test.*

Proof. From (27a) consider

$$\lim_{S \rightarrow +\infty} \frac{1}{S} \omega(\mathbf{w}, b) = \frac{C}{2} \lim_{S \rightarrow +\infty} \frac{1}{S} \sum_{i=1}^S e_i^2 = \frac{C}{2} E_t(\mathbf{w}, b), \quad (29)$$

where $E_t(\mathbf{w}, b) = \mathbb{E} \left[(t_i - \tilde{t}(\mathbf{a}^{(i)}))^2 \right]$ is the expected value carried out with respect to the training points $\mathbf{a}^{(i)}$, as S goes to infinity. The first equality in (29) comes from Lemma 1: since \mathbf{w} converges to a finite norm, we can write

$$\lim_{S \rightarrow \infty} \frac{1}{S} \mathbf{w} \mathbf{w}^T = 0. \quad (30)$$

The last equality comes from the strong law of large numbers. In the limit, the optimization problem (27) is equivalent to

$$\min_{\mathbf{w}, b} E_t(\mathbf{w}, b), \quad (31)$$

where we dropped constraints (27) by using (28). The optimization problem is the same as in the NN case and from [15], with the couple (\mathbf{w}^*, b^*) minimizing (31) and parametrizing (26) we have

$$\tilde{t}(\mathbf{a}^{(i)}) \approx \mathbb{P}(\mathcal{H}_0|\mathbf{a}^{(i)}) - \mathbb{P}(\mathcal{H}_1|\mathbf{a}^{(i)}). \quad (32)$$

Lastly, we exploit Theorem 1 to conclude the N-P-optimality of the LS-SVM. \square

In summary, we have proven that both NN (with CE and MSE design) and SVM (with LS design) converge to the N-P test function as the training set size S goes to infinity, thus establishing their asymptotic optimality and their relation to theory of power powerful hypothesis testing.

IV. IRLV BY ONE-CLASS CLASSIFICATION

In practice having learning points in both regions \mathcal{A}_0 and \mathcal{A}_1 may be difficult since *a)* region \mathcal{A}_1 may be quite wide and not necessarily well defined (being simply the complementary region of \mathcal{A}_0) and *b)* the attacker may use multiple antennas and by beamforming can induce attenuation estimates that not necessarily correspond to points in the region \mathcal{A}_1 . Therefore we consider here methods that rely only on the statistics and data collected from within the region \mathcal{A}_0 . This problem can also be denoted as one-class classification since we know only samples taken from one of the classes of the classifier.

When only the channel statistics from within \mathcal{A}_0 are known a-priori, the LLR (5) can not be used as discriminant function, as $p(\mathbf{a}|\mathcal{H}_1)$ is not known. In this case we can resort to the GLRT [13], which, although not proven to be optimal, is a meaningful generalization of the N-P test. Hence the test function becomes

$$f^*(\mathbf{a}) = \begin{cases} -1 & \text{if } p(\mathbf{a}|\mathcal{H}_0) \geq \Lambda \\ 1 & \text{if } p(\mathbf{a}|\mathcal{H}_0) < \Lambda. \end{cases} \quad (33)$$

In the following we address the problem of one-class classification implemented via both NN and SVM. In particular, during the learning phase the network collects only samples obtained from region \mathcal{A}_0 . Two approaches are considered: the AE, using a NN, and the OCLSSVM.

A. Auto Encoder NN

When attenuation statistics conditioned to the two hypotheses are not available, we can apply ML solutions, as for the case of two-class classification of Section III.

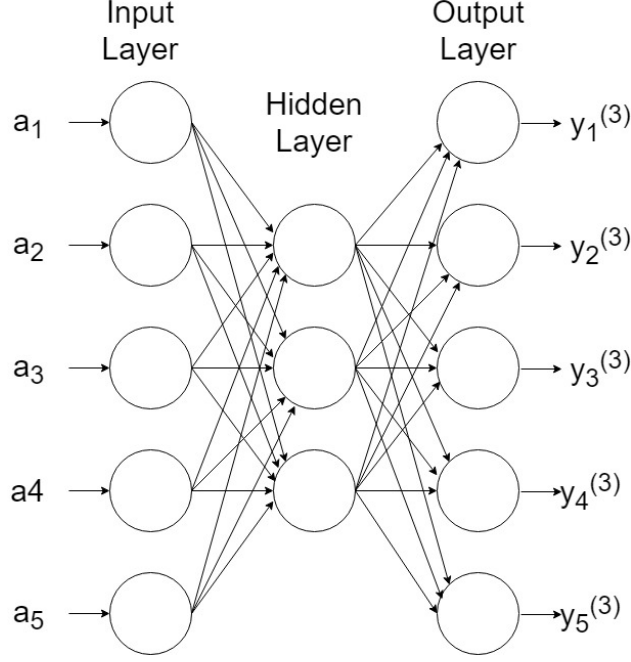


Fig. 2. Example of AE architecture with 5 input (and output) values and one hidden layer with 3 neurons.

In particular we consider here the AE [11], i.e., a NN that is trained to: *a*) convert the high-dimensional input vector into a low-dimensional vector in the hidden layer; *b*) reconstruct the high-dimensional input vector at the output layer from the low-dimensional vector. Therefore the size of the hidden layer is smaller than the size of the input layer [4]. The capacity of the auto-encoder to replicate only certain values at the output is due to the fact that the hidden layer is able to extract the features of the training set. The simplest architecture for AE is a feedforward NN with three layers, i.e., a size N input layer, a size M hidden layer and a size N output layer, as shown in Fig. 2. In this case it is convenient to use a linear activation function when mapping the hidden layer to the output layer [10]. Note that the output of the AE is a vector of the same size of the AE input, and the one-class classification is obtained by computing the reconstruction error between the input and the output of the AE and comparing its absolute value with a chosen threshold.

For our IRLV problem, we train the AE with attenuation vectors $\mathbf{a}^{(i)}$ taken only when the trusted UE is in ROI \mathcal{A}_0 . Then, by letting $\mathbf{y}^{(L)}$ be the output vector of the AE for the attenuation input \mathbf{a} , the MSE of the reconstruction is

$$\epsilon(\mathbf{a}) = \frac{1}{N} \sum_{n=1}^N |a_n - y_n^{(L)}(\mathbf{a})|^2. \quad (34)$$

Finally, the IRLV test function is

$$f(\mathbf{a}) = \begin{cases} 1 & \text{if } \epsilon(\mathbf{a}) \geq \Lambda, \\ -1 & \text{if } \epsilon(\mathbf{a}) < \Lambda. \end{cases} \quad (35)$$

About the test power of the AE, we observe that it can be seen as a quantization (or compression process) that quantizes an N -dimensional signal into an M -dimensional signal. In order to minimize the MSE of the reconstruction error, inputs with higher probability will have smaller quantization regions. Moreover, as the number of quantization points goes to infinity (since the quantization indices are in the continuous M -dimensional space) all points in the same quantization region will have approximately the same probability. However, quantization error for points within each region will be different for each points, in particular being zero for the quantization point and higher at the edges of the quantization region. Thus we can conclude that even with infinite training and an infinite number of neurons the AE can not provide as output the PDF of the input, as required by the optimal decision rule (6). On the other hand, input points with a smaller PDF belong to larger quantization regions for which the reconstruction error is *on average* larger, thus on average the output provided by the AE is monotonically decreasing with the PDF of the input point.

B. One-Class LS-SVM

Similarly, we can resort to SVM to perform the one-class classification in IRLV. We focus in particular on the OCLSSVM, first introduced in [8] as an extension of the one-class SVM [17].

The only difference with respect to the SVM introduced in Section III is that the training optimization problem is now

$$\min_{\mathbf{w}, b} \quad \omega(\mathbf{w}, b) \triangleq \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{C}{2} \sum_{i=1}^S e_i^2 + b \quad (36a)$$

$$\text{subject to } -b - \mathbf{w}^T \phi(\mathbf{a}^{(i)}) = e_i, \quad i = 1, \dots, S, \quad (36b)$$

where C is a hyper-parameter. Note that in the one-class case, the bias parameter b appears also in the objective function.

Also in this case we observe that the one-class SVM is suboptimal, as it does not provides a monotone function of the PDF of the input points. In fact, it is not necessary that training points further away from the bound of the classification region are less probability and the direction \mathbf{w} is obtained by an ensemble elaboration of the PDF of the input rather than been a point-wise

function of it. Still, by resorting to the Chernoff bound we can conclude that by minimizing the MSE we also minimize the upper bound the probability of false alarm, thus although not optimal, the optimization process goes in the right direction.

C. ML-Based Attack Strategies

ML techniques can be also exploited in order to perform more effective attacks. In particular the attacker *a)* obtains estimates of the attenuation vectors of its channel to all the APs, and *b)* moves around in the area \mathcal{A}_1 and performs attacks. Point *a)* is possible if APs transmit training signals to the UE, so that it can estimate the channel characteristics. We also assume that the attacker has means to determine if its attack has been successful, e.g., by receiving the services reserved to UEs in the ROI.

Although it is possible to perform multiple attacks, the purpose of the attacker is to be successful with the minimum number of attacks in order not to let the network detect to be under a series of attacks and take countermeasures (e.g., activating additional IRLV techniques or switching off the service). Moreover, a smaller number of attacks reduces the resource consumption by the attacker and provides faster access to the services. In order to make attacks more efficient we propose that the attacker moves in the ROI-complementary area \mathcal{A}_1 , measures the attenuation vectors at various position and then decides whether to attack or not, according to its previous experience of failed attacks. We denote this attack as *selective ML attack*. As soon as an attack is successful the procedure is stopped, therefore the experience on which the attacker can base its decision comprises only failed attacks.

In details, the attacker uses the attenuation vectors of (failed) attacks to train a one-class classifier. Then, when reaching a new position it feeds the attenuation vector to the classifier and if it is classified as belonging to the same class of training, no attacks is performed since the position is deemed to be useless. Otherwise, if the attenuation vector is not recognized as belonging to the class of collected points, it is evaluated as a potential successful attack and the attacker sends a message claiming to be in \mathcal{A}_0 to the network. Upon success the procedure is stopped, while upon failure the attenuation vector is fed as additional training to the machine, so that the classifier becomes more accurate.

For the one-class classifier we use either the AE or the SVM, as described in the previous section.

V. NUMERICAL RESULTS

In this section we present numerical results obtained using the proposed IRLV and attack solutions for the scenario described in Section II. In particular, for the NN for two-class classification has one hidden layer. The activation function for the input layer 0 is the identity function, for the hidden layer is the sigmoid function

$$\psi^{(1)}(x) = \frac{1}{1 - e^{-x}}, \quad (37)$$

and for the output layer is

$$\psi^{(1)}(x) = \tanh^{-1}(x) = \frac{1}{2} \left(\frac{1+x}{1-x} \right). \quad (38)$$

In order to build the training set we consider n_x spatial points \mathbf{x}_{UE} , each one representing the coordinates of the position of a UE, uniformly distributed over the area \mathcal{A} .

About the channel model we consider a unitary transmitting power for each user and a carrier frequency of 2.12 GHz. The path-loss coefficient is $\nu = 3$, the shadowing power is $\sigma_s^2 = 3.39$ and the shadowing decorrelation distance is $d_c = 75$ m.

A. Two-class IRLV With Single AP

We consider here the IRLV system using a single AP. Two channel models are considered: a simple LOS channel model as discussed in Section II-C and a no-LOS model including both path-loss and shadowing.

LOS Model: We consider an overall circular area \mathcal{A} with radius $R_{\text{out}} = 40$ m, a square authentic area \mathcal{A}_0 of $L = H = 25$ m located inside \mathcal{A} , with upper left corner at a distance of $R_{\text{min}} = 4$ m from the center of \mathcal{A} . In the simplified scenario of Section II-C we have a close-form expression of the N-P test function. For the ML approaches we consider $S = 10^6$ training points and various numbers N_h of the hidden layer neurons.

Fig. 3 shows the FA versus (vs.) the MD probabilities – the so-called ROC – obtained with various IRLV techniques. In particular we compare the N-P test function with LS-SVM and MLP designed according both MSE and CE criteria, and denoted as NN-MSE and NN-CE, respectively.

We first note that both NN-CE and NN-MSE obtain the same performance for the same number of neurons N_h , as expected as they both are equivalent to N-P testing in the asymptotic regime. When compared to the N-P approach, we note that as the number of hidden layer neurons grows

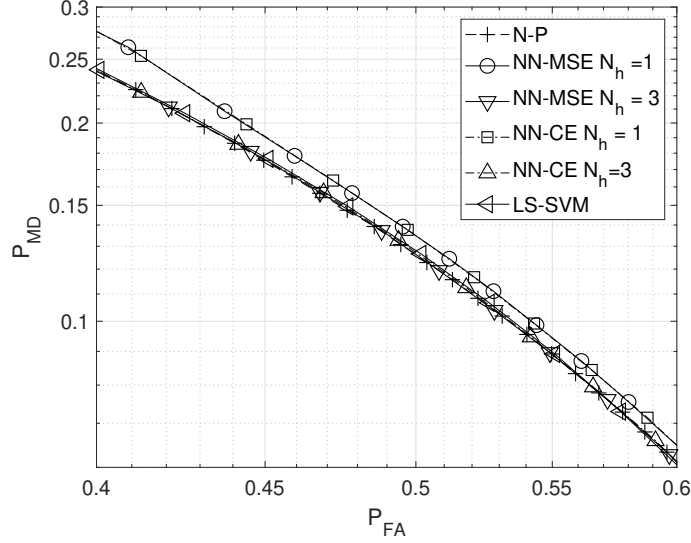


Fig. 3. ROC of IRLVs methods in the scenario of Section II-C: N-P test function, LS-SVM, MSE-NN and CE-NN with N_h hidden layer neurons.

the performance of the NN-based IRLV approaches the optimal N-P test, and the equivalence is reached already with $N_h = 3$ neurons. Moreover, LS-SVM attains the same performance of the NN solution, confirming that also SVM is optimal in the N-P sense.

In general, the achieved (P_{FA}, P_{MD}) probability couples show quite high values: this is due to the fact that a single AP is used for IRLV, and there are many points both outside and inside the ROI that are at the same distance from the AP. Since the attenuation is a function only of this distance, there is a huge ambiguity in the attenuation and the hypothesis testing in quite unreliable.

No-LOS Scenario: We now consider a more general channel model including shadowing and no-LOS and more general shapes of \mathcal{A} and \mathcal{A}_0 . With reference to Fig. 4 we focus on a scenario in which the AP is located at the map center. Moreover, we have a Manhattan grid with two streets crossing at the center of a square region, and the ROI is located in the south-west of building. Along the the streets LOS propagation conditions are present, while no-LOS propagation conditions are present in the rest of the area. Fig. 4 shows a realization of the attenuation map including path-loss and shadowing. We clearly see LOS propagation conditions along the streets and no-LOS propagation conditions in buildings.

Again, using a single AP ($N_{AP} = 1$) we do not expect to be able to achieve simultaneously

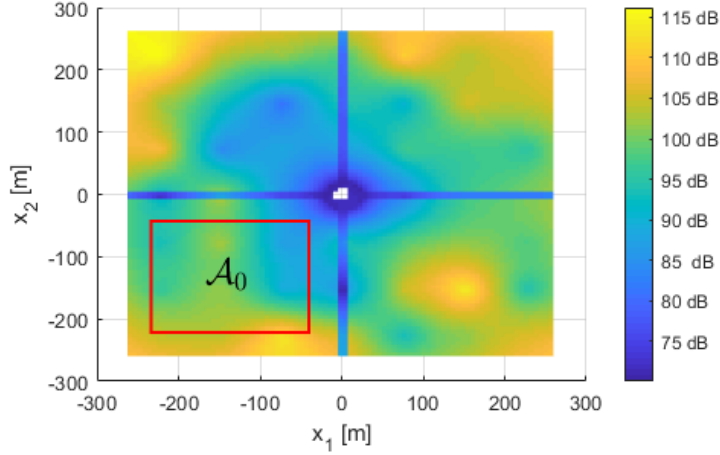


Fig. 4. Example of attenuation map including path-loss and shadowing in the non-LOS scenario.

low MD and FA probabilities. Indeed, when compared to the LOS scenario the presence of shadowing increases the ambiguity on the attenuation inside and outside the region. However in this simple scenario we still compare ML and N-P IRLV approaches.

Since no close-form expression of the LLR is available in this case, we start from the collected data in the learning phase and quantize the attenuation values with a large alphabet and estimate the sampled PDF for the quantized attenuation to be used in the LLR computation. For the considered scenario with path-loss and shadowing we use $4.46 \cdot 10^6$ training points in the area and a uniform quantizer for the attenuation was considered with 300 quantization values. Instead, for training both MLP and SVM we use only 10^3 points.

Fig. 5 shows the ROC of N-P, NN-CE and LS-SVM obtained by averaging over many attenuation maps with different shadowing. The N-P test function has been obtained from the sampled PDF, as just described. We notice that both NN-CE and LS-SVM outperform the N-P test. This means that even for such a huge vale of samples used to estimate the PDF, we still we have a performance degradation with respect to perfect knowledge of the statistics, which would perform all other methods. On the other hand a relatively smaller number of training points for ML methods already provide more powerful test functions, showing the effectiveness of ML techniques when no channel statistics is available a-priori. Therefore in the following we drop the N-P method and we consider long training and many neurons for the ML in order to achieve close-to-optimal performance.

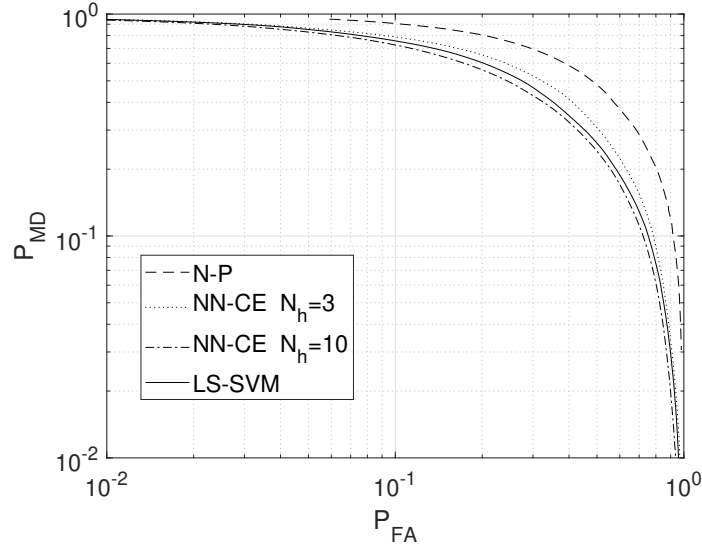


Fig. 5. ROC for a scenario with path-loss and shadowing for N-P, SVM and MSE-NN and hidden layer N_h neurons.

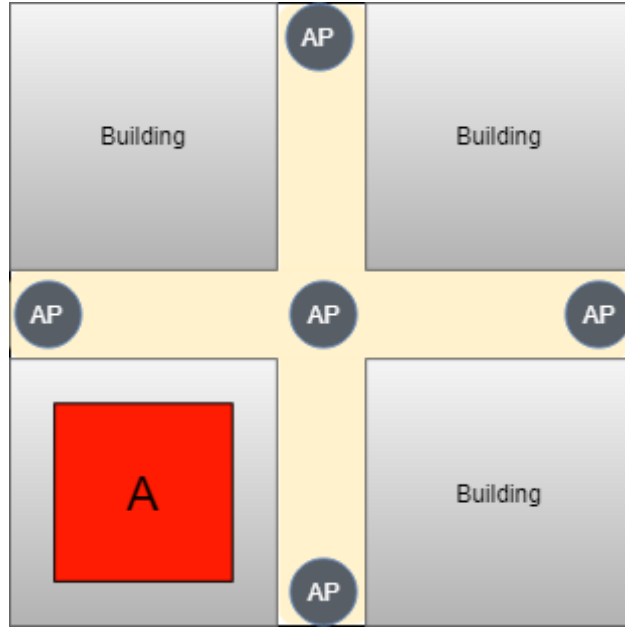


Fig. 6. Scenario with $N_A = 5$ APs, the side of the square representing area \mathcal{A} has a length of 525 m, while the side of area \mathcal{A}_0 has a length of 255 m.

B. Two-class IRLV With Multiple APs

We consider the network of Fig. 6, with $N_{ap} = 5$ APs used for IRLV and area \mathcal{A} is square with side length of 525 m, while the ROI \mathcal{A}_0 is a square with side length 255 m. We include

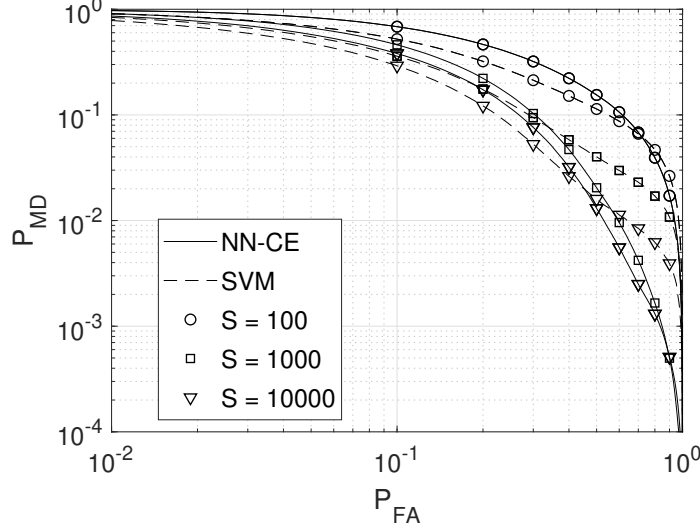


Fig. 7. ROC of SVM and NN-CE ($N_h = 10$) for different numbers of training set size S .

also fading into the channel model, as introduced in Section II.

Fig. 7 shows the ROC for NN-CE (with $N_h = 10$) and LS-SVM IRLV methods and different values of the training set size S . We observe that for the same FA probability we achieve a lower MD probability as the size of the set S increases, and that both methods performs in a similar way at high S , as we know that they both converge to the optimal N-P solution, not reported here for the difficulty of obtaining the LLR, as discussed above. Moreover, thanks to the presence of multiple APs, we can better distinguish attenuation vectors associated to UE positions inside and outside the ROI with respect to the scenario with a single AP analyzed before. Still, for security purposes we would prefer even lower values of FA and MD probabilities, that can be achieved by either further increasing the APs or consider other channel features, e.g., its wideband impulse response.

We now investigate the choice of the points used for training of the ML approaches in the presence of fading. First note that for a given UE positions, the attenuation takes various values over time, according to the instantaneous fading realization. Therefore we may wonder if it is better to collect multiple attenuation vectors for each explored location in the training phase. Let n_x be the number of explored locations by the UE, and k_f the number of collected fading realizations per location. Overall we obtain $S = n_x \cdot k_f$ training attenuation vectors. For example, Fig. 7 is obtained with $k_f = 1$ as each training point is associated to a different uniform randomly

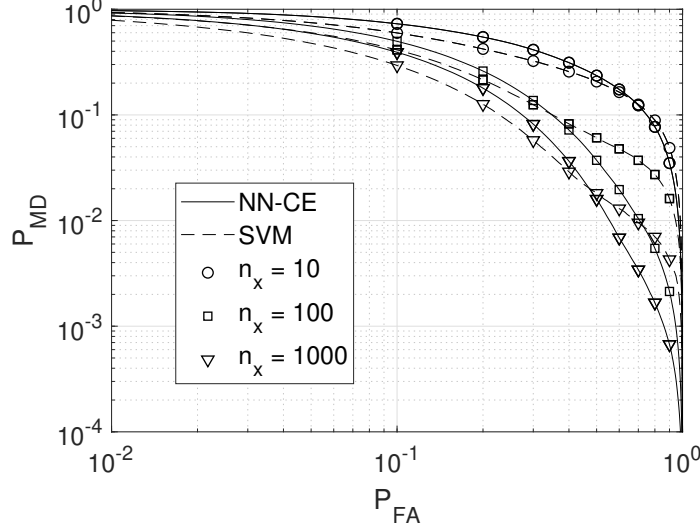


Fig. 8. ROC of SVM and NN-CE ($N_h = 10$) for different numbers of collected locations (n_x) and $k_f = 10$ fading realization per location.

generated location. Fig. 8 instead shows the ROC for the same values of S of Fig. 7, but with $k_f = 10$ fading realizations per location (thus $n_x = S/10$). Comparing the two figures we note that for small values of S , performance get slightly worse as k_f grows. For a large enough training set size S , different values k_f provide approximately the same performance. Therefore we can conclude that, for large training sets ML algorithms are also robust to fading irrespective of the number of collected fading realization per location. However, in practical situations where the points for training may be limited, it may be advantageous to collect more fading realization per location.

C. One-Class IRLV With Multiple APs

We now focus on the one-class IRLV solutions, where the training points come only from inside the ROI \mathcal{A}_0 . We first consider a scenario with path-loss and shadowing only, thus without fading.

Fig. 9 shows the ROC for both OCLSSVM and AE with $N_h \in [1, 5]$ and $S = 10^4$ training vectors. We see that by increasing N_h the performance of AE-based IRLV does not improve: indeed, the optimum ROC is obtained for $N_h = 2$. This is due to the fact that AE compresses the attenuation vectors and hence best performance are achieved when it extracts the optimal number of features from the input. As we have seen, the one-class solutions are not optimal

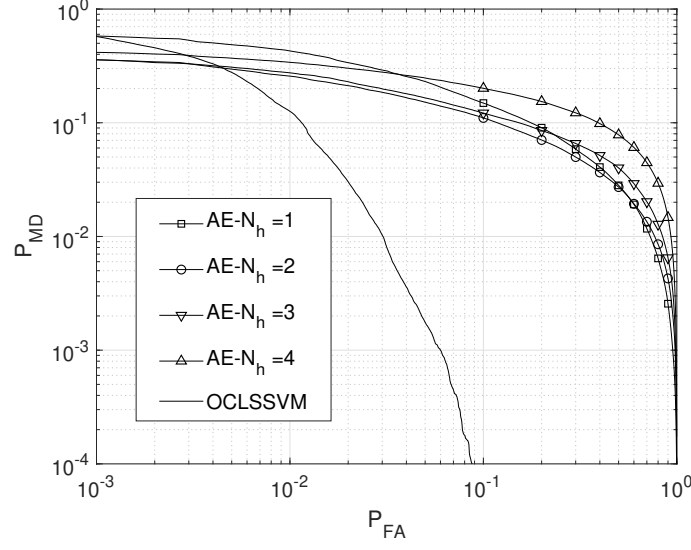


Fig. 9. ROC for one-class IRLVs in a scenario with path-loss, shadowing and without fading.

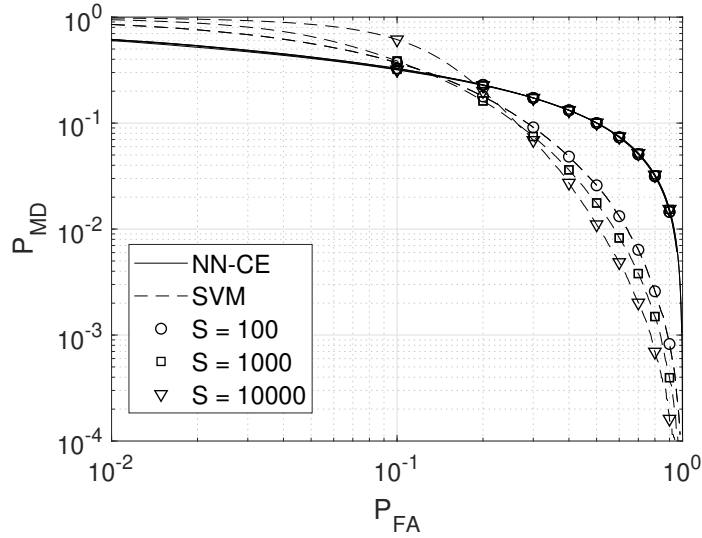


Fig. 10. ROC for one-class IRLVs in a scenario with path-loss, shadowing and fading with different training set size and $k_f = 1$ fading realizations, AE with $N_h = 2$.

in general, and we clearly see that OCLSSVM is significantly more powerful than AE, as the obtained ROC achieves a lower P_{MD} for the same P_{FA} .

We now consider the effects of fading and the choice of the training points. Note that for one-class IRLV the training set collects only attenuation vectors from UE located in \mathcal{A}_0 , whereas

the testing set comprises attenuation vectors of UEs both inside and outside \mathcal{A}_0 . We consider $S = n_x \cdot k_f$ training points with n_x locations and k_f fading realizations per location. Figs 10 and 11 show the ROC for one-class IRLVs systems for $k_f = 1$ and 10, respectively. In particular OCLSSVM and AE with $N_h = 2$ are considered. We first notice that the AE is less sensitive to the training set size S . Furthermore we note that for $P_{FA} > 10^{-1}$ the OCLSSVM attains a lower P_{MD} . Moreover, we note that AE is less sensitive to fading, as error probabilities are very similar in both figures. Instead we notice that OCLSSVM is more sensitive to fading for small values of S , while for a large S , performance get close for both values of k_f . Furthermore we notice that for small S it is better to use one fading realization per space point ($k_k = 1$) in building the training set for the OCLSSVM. This is different from what we observed for two-class classification, where taking more fading measures provided an advantage.

We can also compare Figs 8 and 10 and observe that in the considered scenario the two-class IRLVs outperform the one-class IRLV, the first achieve a lower P_{MD} for the same P_{FA} , although the difference between the performance of the two methods is small. This result is expected, since two-class IRLV also exploits the (estimated) statistics of attacks, which are instead not exploited by the one-class solution. Note that in this case the attacker is behaving as expected by the network in the training phase, so that two-class approach is more powerful in the hypothesis testing, actually is optimum as we shown its convergence to the N-P test performance.

D. ML Attack Strategies

We consider now the ML attack strategies described in Section IV-C for the scenario of Fig. 6 and channels including path-loss, shadowing and fading. The IRLV is also implemented with one-class classifiers.

We compare the ML attack strategies with uniformly random attacks, wherein attacks are launched by uniform randomly selected positions in the ROI-complementary area \mathcal{A}_1 . Moreover, as a naive enhancement we consider a *random-border* attack, wherein the attacker moves only along a strip of 2.5 m the border between areas \mathcal{A}_0 and \mathcal{A}_1 , as we expect that being closer to the ROI increases the chances of a successful attack. However in this case we assume that the ROI \mathcal{A}_0 coincides with the whole south-west building of Fig. 5. Therefore, on the border between \mathcal{A}_0 and \mathcal{A}_1 the UE is the streets, having a LOS propagation characteristic, whereas region \mathcal{A}_0 has a no-LOS propagation characteristic.

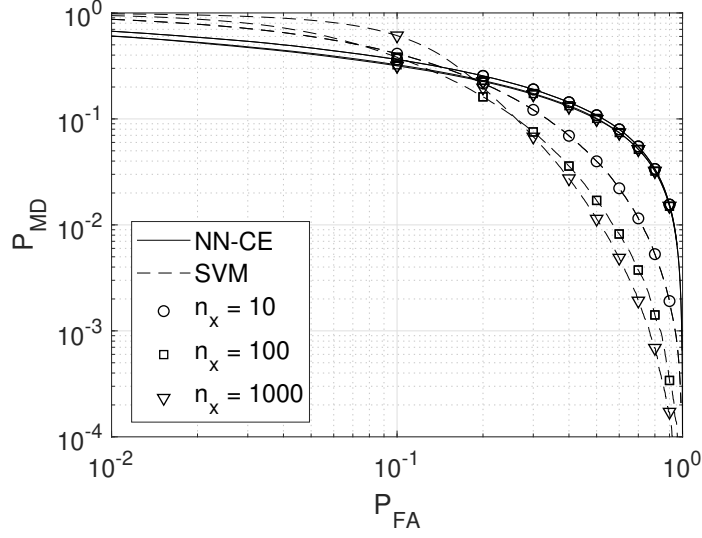


Fig. 11. ROC for one-class IRLVs in a scenario with path-loss, shadowing and fading with different training set size and $k_f = 10$ fading realizations, AE with $N_h = 2$.

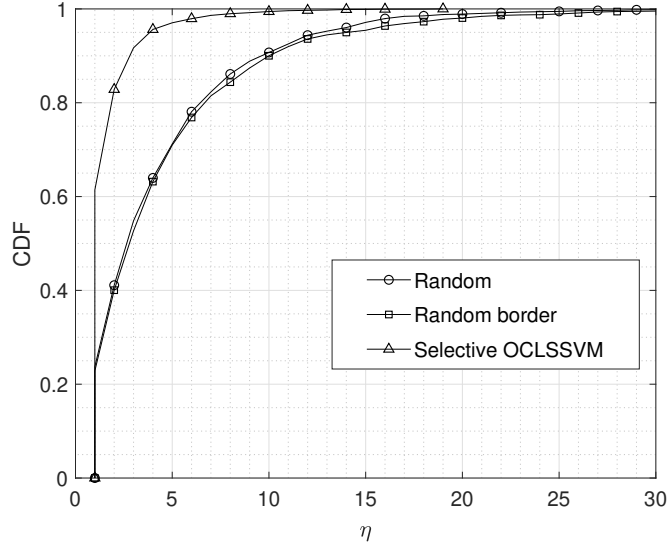


Fig. 12. CDF of the time of first successful attack η for various attack strategies. Both the selective ML attack and IRLV are based on OCLSSVM and $P_{FA} = 10^{-2}$.

The same one-class ML algorithm is implemented both for attack and IRLV, although the attacker training set includes only points in \mathcal{A}_1 , while the AP network training set includes only points in \mathcal{A}_0 . For IRLV we set $k_f = 10$.

Figs 12 and 13 show the CDF of the number of of the first successful attack η for the

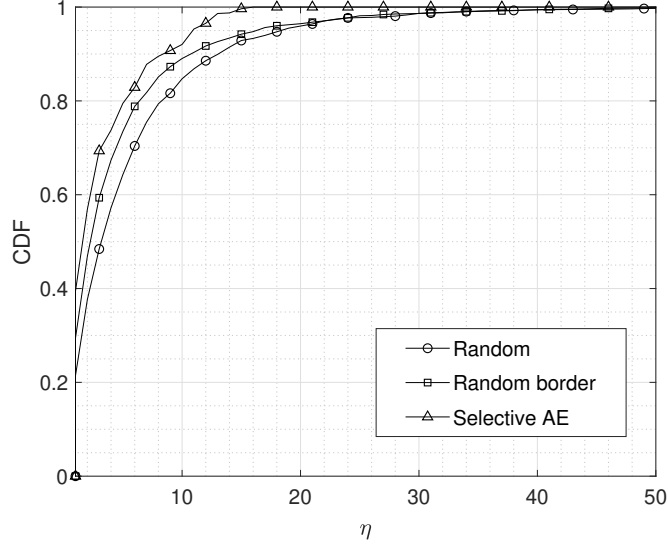


Fig. 13. CDF of the time of first successful attack η for various attack strategies. Both the selective ML attacks and IRLV are based on AE and $P_{FA} = 0.5$.

random, random-border and selective ML attack strategies. Fig. 12 presents results for IRLV with $P_{FA} = 10^{-2}$ and ML approaches based on OCLSSVM. We note that the selective-ML attack is faster than pure random attack, which in turns perform similarly to the border-random attack. Fig. 13 instead shows results for IRLV $P_{FA} = 0.5$ and ML approaches based on AEs. We observe that the random-border attack outperforms now the pure random attack, while still being less-powerful than selective ML.

We can conclude that the ML selective attack is clearly reducing the time to reach a success with respect to random attacks. Moreover, AE is more robust to attacks than OCLSSVM when used for IRLV, as for AE the first successful attack as a similar statistics to random attacks while for OCLSSVM the ML selective strategy is successful much earlier than the random strategies.

VI. CONCLUSIONS

In this paper we have proposed innovative solutions for the IRLV in wireless networks that exploit the features of the channels between the UE whose location must be verified by a trusted network of APs. By observing that in typical situations the channel statistics are not available for IRLV, we have proposed ML-based solutions, operating with both one- and two-class classification, i.e., with and without a-priori assumptions on attack statistics. For two-class

classification we have proved that both NN and SVM solutions are the most powerful tests for a given sensitivity, i.e., they are equivalent to the N-P test. Instead, for one-class classification both AE and SVM solutions are not equivalent to the GLRT. From numerical results we conclude that ML converges to optimal performance with much smaller training set sizes than N-P test implemented with estimated channel statistics. We have also investigated how to collect the training points for training in order to be robust against the channel fading.

APPENDIX

Given a finite attenuation vector alphabet $\mathcal{C} = \{\alpha_1, \dots, \alpha_M\}$ of M elements, with $\mathbf{a}^{(i)} \in \mathcal{C}$, we indicate with $p_{\mathbf{a}^{(i)}, t_i}(\alpha_j, t)$, with $t \in \{-1, 1\}$, the joint probability of input vector $\mathbf{a}^{(i)}$ and corresponding output t_i , $i = 1, \dots, S$.

By the Glivenko–Cantelli theorem we have that with probability 1 as $S \rightarrow \infty$ there are $Sp_{\mathbf{a}^{(i)}, t_i}(\alpha_j, t)$ training vectors α_j with associated true value t in any training sequence. All these training points will have the same value e_i , from (27b), that will appear $Sp_{\mathbf{a}^{(i)}, y_i}(\alpha_j, t)$ times in the sum $\sum_{i=1}^S e_i^2$. Note that in the training ensemble there could be two equal instances $\mathbf{a}^{(m)} = \mathbf{a}^{(n)} = \alpha_j$, but with different labels $t_m \neq t_n$. Therefore, for $\mathbf{a}^i = \alpha_j$ we can have two possible values for e_i , depending on y_i , and we denote them with $e_{j,1}$ and $e_{j,-1}$. This translates in only $2M$ *distinct* constraints of the type (27b). Asymptotically, for $S \rightarrow \infty$, problem (27) becomes

$$\min_{\mathbf{w}, e} f'_l \triangleq \frac{1}{2} \mathbf{w}^T \mathbf{w} + CS \frac{1}{2} \sum_{j=1}^M [p_{\mathbf{a}^{(i)}, t_i}(\alpha_j, 1) e_{j,1}^2 + p_{\mathbf{a}^{(i)}, y_i}(\alpha_j, -1) e_{j,-1}^2] \quad (39a)$$

subject to

$$[\mathbf{w}^T \phi(\alpha_j) + b] = 1 - e_{j,1} \quad j = 1, \dots, M. \quad (39b)$$

$$-[\mathbf{w}^T \phi(\alpha_j) + b] = 1 - e_{j,-1} \quad j = 1, \dots, M. \quad (39c)$$

whose solution provides the convergence value (in probability) of vector \mathbf{w} . We write the Lagrangian

$$\mathcal{L}_1 = f'_l - \sum_{k=1}^M v_k [\mathbf{w}^T \phi(\alpha_j) + b - 1 + e_{j,1}] - \sum_{k=1}^M u_k [-\mathbf{w}^T \phi(\alpha_j) - b - 1 + e_{j,-1}], \quad (40)$$

where $\{u_k, v_k\}_{k=1}^M$ are the Lagrangian multipliers. Let us set to zero the derivatives with respect to $\{\mathbf{w}, b, e_{j,1}, e_{j,-1}, v_j, u_j\}$

$$\frac{\partial \mathcal{L}_1}{\partial \mathbf{w}} : \quad \mathbf{w} = \sum_{k=1}^M (u_k - v_k) \phi(\alpha_k), \quad (41a)$$

$$\frac{\partial \mathcal{L}_1}{\partial b} : \sum_{k=1}^M (u_k - v_k) = 0, \quad (41b)$$

$$\frac{\partial \mathcal{L}_1}{\partial e_{j,1}} : v_j = CSp_{\mathbf{a}^{(i)}, t_i}(\boldsymbol{\alpha}_j, 1)e_{j,1} \quad j = 1 \dots M, \quad (41c)$$

$$\frac{\partial \mathcal{L}_1}{\partial e_{j,-1}} : u_j = CSp_{\mathbf{a}^{(i)}, t_i}(\boldsymbol{\alpha}_j, -1)e_{j,-1} \quad j = 1 \dots M, \quad (41d)$$

$$\frac{\partial \mathcal{L}_1}{\partial v_j} : \mathbf{w}^T \phi(\boldsymbol{\alpha}_j) + b - 1 + e_{j,1} = 0 \quad j = 1 \dots M, \quad (41e)$$

$$\frac{\partial \mathcal{L}_1}{\partial u_j} : -\mathbf{w}^T \phi(\boldsymbol{\alpha}_j) - b - 1 + e_{j,-1} = 0 \quad j = 1 \dots M. \quad (41f)$$

Substituting (41a), (41c) and (41d) in (41e) and (41f) we get the system of equations

$$\sum_{k=1}^M (u_k - v_k)k(\phi(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_j)) + b - 1 + \frac{v_j}{CSp_{\mathbf{a}^{(i)}, t_i}(\boldsymbol{\alpha}_j, 1)} = 0 \quad j = 1 \dots M \quad (42a)$$

$$-\sum_{k=1}^M (u_k - v_k)k(\phi(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_j)) - b - 1 + \frac{v_j}{CSp_{\mathbf{a}^{(i)}, t_i}(\boldsymbol{\alpha}_j, -1)} = 0 \quad j = 1, \dots, M \quad (42b)$$

$$\sum_{k=1}^M (u_k - v_k) = 0. \quad (42c)$$

(42) is a system with $2M + 1$ equations, linear in the $2M + 1$ unknowns $\{u_k, v_k, b\}_{k=1}^{k=M}$ and therefore has finite solution. In particular, using (41a), we have

$$\mathbf{w}^T \mathbf{w} = \sum_{k=1}^M \sum_{h=1}^M k(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_h)(v_k v_h + u_k u_h - 2v_k u_h), \quad (43)$$

where we used the fact that the kernel function

$$k(\boldsymbol{\alpha}_k, \boldsymbol{\alpha}_h) \triangleq \phi(\boldsymbol{\alpha}_k)\phi(\boldsymbol{\alpha}_h)^T \quad (44)$$

is symmetric with respect to its inputs.

We conclude that \mathbf{w} has a finite norm since the right hand side of (43) is a finite sum.

REFERENCES

- [1] LTE; evolved universal terrestrial radio access (E-UTRA); radio frequency (RF) system scenarios. Tr 36.942 version 15.0.0 release 15, 3GPP, Jul 2018.
- [2] A. Abdou, A. Matrawy, and P. C. van Oorschot. CPV: Delay-based location verification for the internet. *IEEE Trans. on Dependable and Secure Computing*, 14(2):130–144, March 2017.
- [3] Christopher M. Bishop. *Pattern Recognition And Machine Learning*. Springer, 2006.
- [4] H. Bourlard and Y. Kamp. Auto-association by multilayer perceptrons and singular value decomposition. *Biological Cybernetics*, 59:291–294, sep. 1988.
- [5] Stefan Brands and David Chaum. Distance-bounding protocols. In *Proc. Workshop on the Theory and Application of of Cryptographic Techniques*, pages 344–359. Springer, July 1993.
- [6] G. Caparra, M. Centenaro, N. Laurenti, and S. Tomasin. Optimization of anchor nodes’ usage for location verification systems. In *Proc. 2017 International Conf. on Localization and GNSS (ICL-GNSS)*, pages 1–6, June 2017.
- [7] L. Chen, S. Thombre, K. Järvinen, E. S. Lohan, A. Alén-Savikko, H. Leppäkoski, M. Z. H. Bhuiyan, S. Bu-Pasha, G. N. Ferrara, S. Honkala, J. Lindqvist, L. Ruotsalainen, P. Korpisaari, and H. Kuusniemi. Robustness, security and privacy in location-based services for future IoT: A survey. *IEEE Access*, 5:8956–8977, April 2017.
- [8] Young-Sik Choi. Least squares one-class support vector machine. *Pattern Recognition Letters*, 30(13):1236–1240, 2009.
- [9] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
- [10] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*. MIT press, Cambridge, 2016.
- [11] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, Jul. 2006.
- [12] E. Jorswieck, S. Tomasin, and A. Sezgin. Broadcasting into the uncertainty: Authentication and confidentiality by physical-layer processing. *Proc. of the IEEE*, 103(10):1702–1724, Oct 2015.
- [13] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.
- [14] E. A. Quaglia and S. Tomasin. Geo-specific encryption through implicitly authenticated location for 5G wireless systems. In *Proc. 2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, pages 1–6, July 2016.
- [15] Dennis W Ruck, Steven K Rogers, Matthew Kabrisky, Mark E Oxley, and Bruce W Suter. The multilayer perceptron as an approximation to a bayes optimal discriminant function. *IEEE Trans on Neural Networks*, 1(4):296–298, Dec 1990.
- [16] Naveen Sastry, Umesh Shankar, and David Wagner. Secure verification of location claims. In *Proc. of the 2nd ACM workshop on Wireless security*, pages 1–10. ACM, September 2003.
- [17] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alex J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471, Jul. 2001.
- [18] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, Jun. 1999.
- [19] Ye Tian, Bruce Denby, Iness Ahriz, Pierre Roussel, and Gérard Dreyfus. Robust indoor localization and tracking using gsm fingerprints. *EURASIP Journal on Wireless Comm and Networking*, 2015(1):157, June 2015.
- [20] A. Vora and M. Nesterenko. Secure location verification using radio broadcast. *IEEE Trans. on Dependable and Secure Computing*, 3(4):377–385, Oct 2006.
- [21] Yawen Wei and Yong Guan. Lightweight location verification algorithms for wireless sensor networks. *IEEE Trans on Parallel and Distributed Systems*, 24(5):938–950, May 2013.

- [22] Liang Xiao, Xiaoyue Wan, and Zhu Han. Phy-layer authentication with multiple landmarks with reduced overhead. *IEEE Trans on Wireless Comm*, 17(3):1676–1687, December 2018.
- [23] Shihao Yan, Ido Nevat, Gareth W Peters, and Robert Malaney. Location verification systems under spatially correlated shadowing. *IEEE Trans on Wireless Comm*, 15(6):4132–4144, February 2016.
- [24] Jieping Ye and Tao Xiong. Svm versus least squares svm. pages 644–651, Apr. 2007.
- [25] Yingpei Zeng, Jiannong Cao, Jue Hong, Shigeng Zhang, and Li Xie. Secure localization and location verification in wireless sensor networks: a survey. *The Journal of Supercomputing*, 64(3):685–701, November 2013.