Hybrid Physical Layer Security for Passive RFID Communication

A. Gouissem[1], K. Abualsaud[1], E. Yaacoub[1], T. Khattab[2], M. Guizani[1]
[1] Computer Science and Engineering, Qatar University, Doha, Qatar,
[2] Electrical Engineering, Qatar University, Doha, Qatar.
$gouissem.ala, k.abualsaud@qu.edu.qa; eliasy, tkhattab, mguizani@ieee.org$

***Abstract*—Thanks to its low cost, small weight and energy efficiency, passive radio frequency identification (RFID) backscatter communications systems have attracted a lot of attention in several application fields. However, such devices have limited computational capabilities and resources which makes them unable to incorporate traditional security protocols and are therefore vulnerable to several types of attacks including cloning and counterfeiting. Therefore, in this paper, a novel hybrid RFID tags identification and malicious devices detection system is proposed by exploiting the estimated tags locations and manufacturing imperfections. In particular, an iterative approach is proposed to estimate the minimum power response at each frequency of the tag in addition to its location. The conducted simulation results show the efficiency of this technique in detecting all the malicious tags and classify the legitimate ones under different network configurations.**

***Index*— RFID fingerprinting, tags localization, physical layer security, tags classification.**

I. Introduction

The Radio frequency identification (RFID) technology has recently attracted an increasing attention by both researchers and industrialists due its large range of applications such as commercial, medical, transportation, environmental and localization applications [1]. In particular, due to their limited sizes, weight, cost and power consumption, RFID systems are used in the design of body wireless sensor networks that connect implanted devices with the medical facility network. This allows for examples easy patients vital signs monitoring and fast medical information retrieval [2].

In fact, the RFID is based on the backscatter communication technology where a reader transmits a signal to power and communicate with the tags. There are different types of tags, but passive tags that simply reflect the received signal from the reader are the most used due their low cost and power efficiency [3]. Such tags have limited computational capabilities and resources that make the implementation of the traditional cryptographic protocols to secure the tag challenging. This makes the security issue one of the biggest challenges in the design and incorporation of RFID tags, especially in health monitoring networks that require a high level of security. For example, the RFID tags data can be easily read for cloning, emulating or counterfeiting. Also, even if lightweight cryptographic protocols are implemented, it has been shown that malicious users with sufficient resources might be able to access the secured data [4].

This motivated several researchers to investigate the design of physical layer security systems that can protect the RFID networks [2], [5], [6]. This can be done by exploiting the channel characteristics to inject an artificial noise signal to jam the eavesdroppers [6]. Directional jamming towards the eavesdroppers is also investigated in [7]. The communication secrecy is further improved by making use of tags with multiple antennas in [5]. Beam steering is also used in [8], [9] to make sure that most of the reader signal power is directed towards the legitimate tags. Directional modulation is also used in [10] to secure the data from being leaked to the eavesdroppers. The authors in [2] also made use of beam steering in the tag side and noise from the reader side.

Several works have also investigated the use of the manufacturing characteristics of the tags to fingerprint them. For example, after analyzing the reflected signal by 100 tags from two manufacturers, the authors in [4] concluded that each tag had a unique minimum power response at each frequency. Also, by analyzing real-world environmental variations, the authors in [11] showed that that identically programmed RFID tags can be distinguished using wavelet fingerprinting techniques. Three different identification features are also investigated in [12]. A physical layer RFID identification system named GenePrint is presented in [13] for UHF passive tags.

However, most of the above mentioned RFID fingerprinting techniques are sensitive to the location of the tag and its direction of communication with the reader. Therefore, if the position of the tag is not known, it might be challenging to accurately extract the fingerprinting features. Therefore, in this paper, a novel hybrid RFID physical layer security technique is proposed to jointly estimate the location of the tags and their power sensitivity. By knowing the locations and the power sensitivities of the legitimate tags, the proposed scheme allows the reader to identify any potential intruder to the network.

The main contributions of this paper are summarized as follows:

- Design of an algorithm that can estimate the normalized tag sensitivity independently from its location.
- Location estimation of the tags.
- Design of an intruders detection and tags identification mechanism.
- Simulations are conducted to validate the accuracy of the proposed scheme in identifying and classifying malicious and legitimate tags.

The rest of this paper is organized as follows. The system model is presented in Section II. Sections III and IV investigate the proposed tags RF identification and localisation technique. Section V investigates the designed authentication and classi-

978-1-7281-7624-6/20/$31.00 ©2020 IEEE
DOI 10.1109/CSCI51800.2020.00033

fication scheme. The simulation results are then presented in Section VI to validate and verify the findings of the paper. Finally, Section VII concludes the paper.

II. System Model

A. Network Model

As depicted in Fig. 1, the adopted system model consists of one RFID reader (denoted by R) communicating with one out of N_T passive RFID tags (denoted by $T_i,\ i = 1..N_T$) with known positions. The communication is performed using Orthogonal Frequency Division Multiplexing (OFDM) transmission over N_S subcarriers. One or multiple malicious passive RFID tags (denoted by T_M) are also assumed to be present in the network aiming to send false data to the reader. These malicious tags can be for example intruders trying to retrieve personal patients medical files or report false health monitoring data. Also, all the nodes are assumed to be equipped with single antennas. In particular, the reader is assumed to communicate with an unknown tag T with unknown position and tries to identify whether it is one of the N_T legitimate tags or if it is an intruder device.

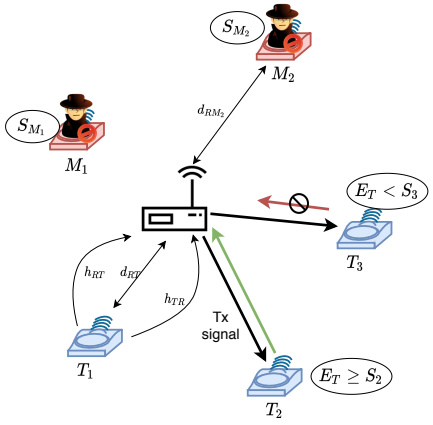

Fig. 1: System model

The objective of this paper is, therefore, the design of low complexity physical layer security techniques that can protect the RFID network from malicious users. In particular, an iterative algorithm is proposed to estimate the Tag location and RF fingerprint in order to differentiate between legitimate and malicious users.

B. Communication Model

Let $y^n_{AB}[t]$ denote the signal transmitted by the node A and received by the B at time t on a frequency subcarrier n, where A and B designate either R, T_i or T_M. Also, at time t and on a subcarrier n, let $h^n_{AB}[t]$ and $n^n_A[t]$ denote the channel response from the node A to the node B, and the noise at the device A, respectively. All the channel links $h^n_{AB}[t]$ and the noise signals $n^n_A[t]$ are assumed to follow zero mean complex Gaussian distributions with variances $(\sigma^n_{AB})^2$ and N^A_0, respectively.

Since a passive RFID model is adopted in this paper, the RFID tag simply reflects the unmodulated wave signal received from the reader. Therefore, by modeling the effect of the additive white Gaussian noise (AWGN) channel from the reader to a tag T and by taking into consideration the path-loss effect, the received signal at a tag T at time instant t and subcarrier n is given by [14]

$$y^n_{RT}[t] = \sqrt{P_R\ \alpha_{RT}}\ h^n_{RT}[t]\ X_n[t] + n^n_T[t], \tag{1}$$

where, the terms P_R, and $n^n_T[t]$ denote the power of the transmitted signal by the reader and the unmodulated wave signal at time t and subcarrier n, respectively. Also, α^n_{RT} denotes the free space path loss from R to T at a subcarrier n given by [14]

$$\alpha^n_{RT} = \left(\frac{\lambda_n}{4\pi d_{RT}}\right)^2. \tag{2}$$

Consequently, the instantaneous received signal power at the tag at time t and subcarrier n is given by

$$E^n_T[t] \triangleq E\left(|y^n_{RT}[t]|^2\right) = P_R\ \alpha_{RT}|h^n_{RT}[t]|^2|X_n[t]|^2 + N^T_0. \tag{3}$$

Depending on the manufacturing characteristics of each RFID tag T, there is a sensitivity threshold that corresponds to the minimum received power required for the tag to power its circuits [4], [14]. Furthermore, as shown in [4], this parameter is unique for each tag and is different for each frequency band. This uniqueness property has been verified even for tags designed by the same manufacturer [4].

Let the sensitivity of a tag T_i at subcarrier n denoted by S^n_i. If the received power at the tag $E^n_T[t]$ is below S^n_i, the Tag will not have enough power to transmit back a modulated signal to the reader. Consequently, at time instant t, a decision variable $u_n[t]$ that corresponds to whether the tag T_i will transmit a signal or not is defined by

$$u_n[t] = \begin{cases} 1 & \text{if } E^n_T[t] \geq S^n_i \\ 0 & \text{otherwise.} \end{cases}, \tag{4}$$

Consequently, the received reflected signal at the reader from the tag T at time t on subcarrier n becomes expressed by

$$y_{TR} = u_n[t]\ \alpha^n_{RT} h^n_{TR}[t]\ y^n_{RT}[t]\ s_n[t] + n_R[t], \tag{5}$$

where, $s_n[t]$ denotes the unmodulated data at the tag at time t and subcarrier n.

III. Tags RF Identification

A. introduction

Due to the low cost and small size of the passive RFID tags, limited computational capabilities and resources are available in such devices. These limitations constrain the use of conventional encryption algorithms and security protocols to prevent cloning and counterfeiting of an RFID tag. Therefore, this section presents a technique that creates an electronic fingerprint of each tag based on its minimum power responses measured at multiple frequencies denoted in the sequel as sensitivity.

B. Problem Formulation

In order to measure the sensitivity of the RFID devices, the instantaneous channel gains denoted by $\gamma_{RT}^{n,t} = |h_{RT}^n[t]|^2$ and $\gamma_{TR}^{n,t} = |h_{TR}^n[t]|^2$ are assumed to be known by the reader and that they can be perfectly estimated. This assumption is realistic in case the coherence time is large enough (i.e. $\gamma_{TR}^{n,t}$ and $\gamma_{RT}^{n,t}$ remain constant for a long duration of time) and in case the channel symmetry property applies to the investigated system. In particular, the channel phase is not required and only the channel response amplitude is needed to apply the tag identification as it will be detailed later.

The objective in the sequel is the design of an algorithm that allows the reader to estimate both d_{RT} and the sensitivity $(S_i^n = 1..N_S)$ which are unique for each tag allowing the identification of any potential intruder to the network.

Instead of sending all the signals from the readers with the same power P_R, the transmission power denoted in the sequel by $P_R^n[t]$, is made variable for each subcarrier n and time t. In particular, by sending a signal with power $P_R^n[t]$ at subcarrier n and time t and receiving a response from the investigated tag, the receiver can decide that $E_T^n[t] \geq S_i^n$. Similar, when the reader does not hear any response at that particular subcarrier, it can decide that $E_T^n[t] < S_i^n$.

However, by only changing the transmission power, the reader cannot estimate at this stage the received power at the tag $E_T^n[t]$ even by knowing the forward channel gain since the position of the tag is still unknown and so does the path loss factor α_{RT}^n. Consequently, the parameter $\beta_n[t]$ is defined as the received forward channel gain with a normalized effect of the path-loss at a reference distance d_0 and is expressed as follows

$$\beta_n[t] = P_R^n[t]\ \alpha_{RT_0}^n\ |h_{RT}^n[t]|^2 |X_n[t]|^2, \tag{6}$$

where $\alpha_{RT_0}^n$ denotes the path loss factor at a distance d_0 from the reader and at subcacrrier n. Consequently, the received energy at the tag becomes

$$E_T^n[t] = \frac{d_0^2}{d_{RT}^2}\beta_n[t] + N_0^T. \tag{7}$$

To identify the vector $S_i = [S_1^i, ..S_{N_S}^i]$, the first step is to identify the vector $\beta^* = [\beta_1^*, ..\beta_{N_S}^*]$ that verifies

$$\frac{d_0^2}{d_{RT}^2}\beta_n^* + N_0^T = S_n^i\ \forall n \in 1..N_S. \tag{8}$$

C. Iterative Tag Identification

As it can be seen in Eq. (7), the new definition of the normalized received forward channel gain β_n makes sure that the energy received at the tag is a function of β_n only, which can be accurately linked to the transmitted power P_R^n. Therefore, β_n^* can be estimated even without knowing the position of the tag itself.

In particular, as detailed in Alg. 1, the tag normalized fingerprint vector β_n^* is obtained using the proposed Iterative Tag Identification (ITI) algorithm without any previous knowledge of the position of the tag. This is done by tuning the transmission power for each subcarrier up and down till reaching the required value for each subcarrier to reflect the signal.

Algorithm 1 Iterative Tag Identification(ITI)

1: Input: $t_0, Max_{\beta_{err}}, Itr_{max}$ and $\beta_0^1, ..\beta_0^{N_S}$
2: Ouput:$[\beta_1, .., \beta_{N_S}]$
3: $Itr \leftarrow 0$
4: $\beta_{err} \leftarrow +\infty$
5: $\beta_{min}^n \leftarrow \beta_0^n\ \forall n = 1..N_S$
6: $\beta_{max}^n \leftarrow \beta_0^n\ \forall n = 1..N_S$
7: $\beta_n \leftarrow \beta_0^n\ \forall n = 1..N_S$
8: **while** $\beta_{err} > Max_{\beta_{err}}$ And $Itr \leq Itr_{max}$ **do**
9: **for** $n = 1 : N_S$ **do**
10: $P_R^n[t] \leftarrow \frac{\beta_n \left(4\pi d_0\right)^2}{\lambda_n^2 \left|h_{RT}^n\left[t_0+\Delta_t\ Itr\right]\right|^2 \left|X_n[t]\right|^2}$
11: **end for**
12: Transmit the signal with power vector $[P_R^1, .., P_R^{N_S}]$ at time $t_0 + \Delta_t\ Itr$
13: Collect $U_n[t_0 + \Delta_t\ Itr],\ \forall n = 1..N_S$
14: **for** $n = 1 : N_S$ **do**
15: **if** $U_n == 0$ **then**
16: **if** $\beta_n == \beta_{max}^n$ **then**
17: $\beta_{max}^n \leftarrow (1+\tau)\ \beta_{max}^n$
18: $\beta_{min}^n \leftarrow \beta_n$
19: $\beta_n \leftarrow \beta_{max}^n$
20: **else**
21: $\beta_{min}^n \leftarrow \beta_n$
22: $\beta_n \leftarrow \frac{\beta_{min}^n + \beta_{max}^n}{2}$
23: **end if**
24: **else**
25: **if** $\beta_n == \beta_{min}^n$ **then**
26: $\beta_{min}^n \leftarrow (1-\tau)\ \beta_{min}^n$
27: $\beta_{max}^n \leftarrow \beta_n$
28: $\beta_n \leftarrow \beta_{min}^n$
29: **else**
30: $\beta_{max}^n \leftarrow \beta_n$
31: $\beta_n \leftarrow \frac{\beta_{min}^n + \beta_{max}^n}{2}$
32: **end if**
33: **end if**
34: **end for**
35: $\beta_{err} \leftarrow \sum_{n=1}^{N_S} |\beta_{max}^n - \beta_{min}^n|^2$
36: $Itr \leftarrow Itr + 1$
37: **end while**
38: $Return\Big([\beta_1, .., \beta_{N_S}]\Big)$

The ITI algorithm starts by initializing the initial iteration to zero (Line 3), initial convergence error to $+\infty$ (Line 4), and initial lookup range denoted by $[\beta_{min}^n, \beta_{max}^n]$ for each subcarrier n to an approximated expected value of β_n denoted by β_0^n. Note that β_0^n is used just to start the lookup close to the actual solution and can be either be used as the expected value for of β_n or and average of β_n for different tags from different manufacturers. Note also that β_0^n depend not only on the devices sensitivities but also on their locations, therefore, this initial value is computed assuming the tag is at distance d_0 from the tag.

As detailed from Line 9 to Line 11, the transmit power P_R^n for each subcarrier is computed based on the actual values of β_n given the perfect knowledge of the channel state information (CSI). To guarantee that the transmission is performed using the computed normalized received forward channel gain β_n, the transmit power P_R^n is set to

$$P_R^n = \frac{\beta_n \big(4\pi d_0\big)^2}{\lambda_n^2 \Big| h_{RT}^n \big[t_0 + \Delta_t \ Itr\big] \Big|^2 \Big| X_n[t] \Big|^2}. \tag{9}$$

Once the desired transmission powers are computed, the transmission is performed at time $t = t_0 + \Delta_t \ Itr$, where t_0 and Δ_t denote the initial transmission time and the delay between consecutive transmissions. Also, the reflected signal is collected to identify whether there was a reflection ($U_n = 1$) or not ($U_n = 0$) for each subcarrier n (Line 12 and Line 13).

Once the reader knows U_n, it can decide for each subcarrier n on the required update of the lookup range of β_n. In particular, if U_n is equal to zero, then the tag needs higher power to respond at subcarrier n. Therefore, the next β_n should be higher than the current one. Consequently, if $\beta_n = \beta_{max}^n$ (i.e. No value of β_n that guarantees signal reflection has been found yet), then the minimum β_{min}^n is set to β_n and β_{max}^n is increased by a rate τ (Line 17 to 19). If β_n and β_{max}^n are different, the algorithm starts iteratively splitting the lookup interval into smaller intervals by setting β_{min}^n to β_n and β_n to $\frac{\beta_{max}^n + \beta_{min}^n}{2}$ (Line 21 to 22).

Similar, if U_n is equal to one, then the tag needs lower power to respond at subcarrier n. Therefore, if β_n and β_{max}^n are the different, the lookup interval is split into smaller intervals by setting β_{max}^n to β_n and β_n to $\frac{\beta_{max}^n + \beta_{min}^n}{2}$ (Line 30 to 31). Otherwise, β_{min}^n is reduced by a factor τ (Line 26 to 28).

The iteration is then incremented and the convergence errors β_{err} is computed as the sum of the squared error between all the β_{max}^n and β_{min}^n. Finally, this operation is repeated until either β_{err} goes below a predefined threshold $Max_{\beta_{err}}$ or the number of iterations exceeded a predefined threshold Itr_{max}.

Consequently, after reaching the stopping criteria, the algorithm ITI return the vector β_n^* that satisfies the condition in Eq. (8).

IV. Tags Localisation

The distance between the reader and the investigated tag needs to be accurately estimated for two main reasons. First, the sensitivity estimation can be extracted from β^* only if this distance is known (see Eq. (8)). Also, by knowing the positions of the legitimate tags and by comparing them with the estimated one, the reader can add another layer of security by making sure that the investigated tag is in a legitimate location.

The distance can be estimated using a statistical approach by computing the distance out of the average received signal. However, such approach might lack estimation accuracy especially if the ITI algorithm converged in just few iterations. In particular, both the modulation signal $X_n[t]$, and the channel gains are known by the user. Therefore, not including them in the distance estimation process is a waste of information and consequently, the statistical estimation shall be used only in case the reader has limited computation capabilities or after making sure the ITI converged in a large number of iterations.

During each iteration of Alg. 1, depending on the computed $P_R^n[t]$, instead of considering the average received power, the instantaneous received power at time t and subcarrier n at the tag is computed as

$$E_R^n[t] = u_n[t]\alpha_{RT}^n |h_{TR}^n[t]|^2 \big(P_R^n[t]\alpha_{RT}^n |h_{RT}^n[t]|^2 |X_n[t]|^2 + N_0^T\big)\sigma_s^2 + N_0^R. \tag{10}$$

Therefore, by averaging the received energy over only the time slots and subcarriers with a signal reflected from the Tag, the average received energy per subcarrier n becomes

$$\bar{E}_R^n = \sum_t \Big(\frac{\alpha_{RT}^n}{N_t^n} |h_{TR}^n[t]|^2 \big(P_R^n[t]\alpha_{RT}^n |h_{RT}^n[t]|^2 |X_n[t]|^2 + N_0^T\big)\sigma_s^2 \Big) + N_0^R. \tag{11}$$

Consequently, by summing the energy of all the subcarriers, the total received energy at the reader becomes

$$\bar{E}_R = \frac{d_0^4 \sigma_s^2}{d_{RT}^4} \sum_{n=1}^{N_S} \frac{\big(\alpha_{RT_0}^n\big)^2}{N_t^n} \sum_t P_R^n[t] |h_{TR}^n[t]|^2 |h_{RT}^n[t]|^2 |X_n[t]|^2 + \frac{d_0^2 \sigma_s^2}{d_{RT}^2} \sum_{n=1}^{N_S} \frac{\alpha_{RT_0}^n}{N_t^n} \sum_t |h_{TR}^n[t]|^2 N_0^T + N_S N_0^R. \tag{12}$$

Consequently, the distance d_{RT} can be estimated as follows

$$\hat{d}_{RT} = \sqrt{\frac{2a d_0^2}{-b \pm \sqrt{b^2 - 4ac}}}, \tag{13}$$

where

$$\begin{aligned} a &= \sigma_s^2 \sum_{n=1}^{N_S} \frac{\big(\alpha_{RT_0}^n\big)^2}{N_t^n} \sum_t |h_{TR}^n[t]|^2 P_R |h_{RT}^n[t]|^2 |X_n[t]|^2, \\ b &= \sigma_s^2 \sum_{n=1}^{N_S} \frac{\alpha_{RT_0}^n}{N_t^n} \sum_t |h_{TR}^n[t]|^2 N_0^T, \\ c &= N_S N_0^R - \bar{E}_R. \end{aligned} \tag{14}$$

Note that estimating d_{RT} does not provide the reader with the exact location of the tag. In particular, the distance between the reader and the tag is used here as an extra layer of security. Therefore, even if the eavesdropper and the tag are at different locations and with the same distance to the reader, the reader shall be able to recognize them using the proposed RF fingerprinting scheme.

Also, note that in case the reader needs to accurately estimate the location of the tag, three readers should be used in the network to triangulate the location of the tag using the same proposed technique in this paper.

V. Authentication Scheme

A. Trustworthiness and Identification

Once the forward channel gain with a normalized effect of the path-loss $\hat{\beta}_n^*$ is accurately estimated for all the subcarriers using Alg. 1 and $\hat{d}_{RT}$ is computed using using the power of the received signal, the tag sensitivity at each subcarrier n can be estimated by using the expression in Eq. 8.

$$\hat{S}_n = \frac{d_0^2}{\hat{d}_{RT}^2}\hat{\beta}_n^* + N_0 = \ \forall n \in 1..N_S. \tag{15}$$

Consequently, given that a tag T_i is known to have a sensitivity vector $[S_1^i, .., S_{N_S}^i]$, the authentication error ξ_i for a tag T authenticating as tag T_i is defined as the weighted sum of the euclidean distance between the known and estimated sensitivity vector in addition to the distance error. i.e.

$$\xi_i = \theta \sum_{n=1}^{N_S} \left|\hat{S}_n - S_n^i\right|^2 + (1-\theta)\left|\hat{d}_{RT} - d_{RT}\right|^2, \tag{16}$$

where $\theta \in [0,\ 1]$ is a parameter defined to control the effect of the sensitivity and the distance on the error computation. A trustworthiness threshold ξ_{th} is then defined as the maximum error ξ_i that can be accepted by the reader to trust the tag. Therefore, the reader decides that the tag belongs to a malicious user if the condition $\mathfrak{M}$ is satisfied

$$\mathfrak{M} : \min_{i=1..N_T} (\mathfrak{E}_i) > \epsilon_{\mathfrak{E}}. \tag{17}$$

In case the tag is not recognized as a malicious device (the condition $\mathfrak{M}$ is not satisfied), the reader also identifies the communicating tag i^* as follows

$$i^* = \underset{i=1..N_T}{argmin}(\mathfrak{E}_i) \tag{18}$$

B. Authentication Acceleration

In case the reference distance d_0 is very far from the actual tag distance, the proposed authentication scheme might take a large number of iterations to converge since the initial search range is far form the actual value. Therefore, to accelerate the search, the ITI algorithm is slightly modified so that to update the value of β_n according to the estimated distance at the second iteration.

In fact, after finishing the first iteration of ITI algorithm, the distance $\hat{d}_{RT}$ is estimated as detailed in Eq. (13). The reference distance d_0 is then updated to the estimated value and β_0^n is also updated accordingly. Consequently, in the second iteration of the ITI algorithm β_n, β_{max}^n and β_{min}^n are forced to take the updated value of β_0^n.

This would accelerate the convergence of the authentication algorithm as it helps the lookup to be performed in a region close to the actual values.

VI. Simulation Results

This section presents the results of the simulations conducted to verify and validate the efficiency of the proposed authentication scheme. The noise variances at the reader and the tags N_0^R and N_0^T are set to $-110dBm$. All the channel links have normalized variances set to $0dBm$ since the effect of the pathloss is investigated separately. Different distances from the tags to the reader are tested through the different simulations ranging from $0.5m$ to $5m$.

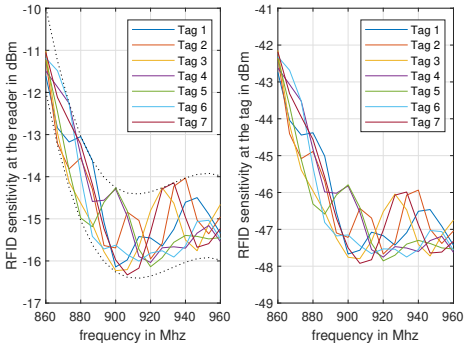

Fig. 2: Tags sensitivity random generation.

To have realistic simulations, the tags sensitivity are generated randomly so that to be close to the values reported in [4]. In particular, as it can be seen in Fig. 2.a, the range of the sensitivity values for each frequency band is defined by analyzing the values in Fig. 2 in [4]. A random frequency correlated signal is then generated in this range to model the sensitivity of the tags at each particular frequency. Note that the sensitivity level ranging from $-17dBm$ and $-10dBm$ are measured in [4] at the reader side assuming a distance $d_{RT} = 1m$. Consequently, the sensitivity at the tag is obtained by multiplying this parameter with α_{RT}^n as in Fig. 2.b.

Fig. 3 analyzes the convergence of the proposed ITI algorithm with and without the acceleration part. In this figure, tag T_1 is assumed to communicate with a reader that knows the characteristics of 4 legitimate tags T_1, T_2, T_3 and T_4. The distance from T_1 to the reader is set to $d_{RT} = 4m$ while d_0 is equal to $1m$. Therefore, the objective of authentication scheme should be to minimize ξ_1 (similarity error between the communication tag and T_1) and maximize the errors ξ_2, ξ_3 and ξ_4 (similarity error between the communication tag and T_2, T_3 and T_4) as fast as possible.

First, Fig. 3.a shows that either the ITI acceleration is used or not, the ITI algorithm makes the authentication error converge to zero for the actual communicating tag (T_1) compared to the other tags. i.e. the reader is able to recognize that it is communicating with tag T_1. Also, it can be seen that by using the acceleration part, the ITI algorithm converges in only 14 iterations compared to 83 without acceleration.

In particular, as it can be seen in 3.c, the initial lookup range for β_n is far from the actual one due to the difference between d_0 and d_{RT}. Therefore, when no acceleration is performed, the ITI algorithm slowly identified the correct range of β_n after around 60 iterations. However, when the acceleration step is applied, the β_n lookup range became accurate from the first iterations. Therefore, the convergence error managed to reach the ,maximum error threshold in just few iterations as it can be seen in Fig. 3.b.

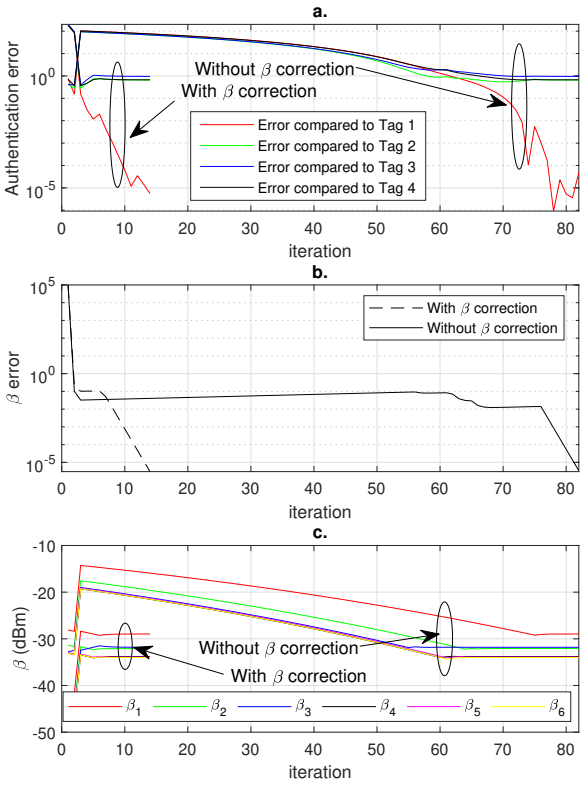

Fig. 3: ITI algorithm convergence.

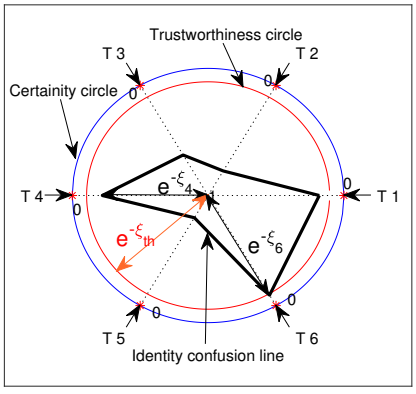

Fig. 4: Authentication confusion map.

To visualize the efficiency of the authentication process, the authentication confusion map is presented in Fig. 4. The objective of this map is to visualize all the authentication errors ξ_i, how much the device is trustworthy and the tag identification confusion. To do so, a circle of radius 1 (blue circle) is first plotted to designate the certainty circle, i.e. where the error ξ_i is equal to zero and the reader is 100% sure of the identity of the tag. Each tag is then assigned a location in the certainty tag. The identity confusion line is then plotted so that the distance from the centre of the circle to the line in the direction of tag T_i is equal to $e^{-\xi_i}$. Consequently, the closer is the confusion line at the direction of tag T_i from the certainty circle, the smaller is the error and the more certain is the reader that it is T_i. Similar, the closer is the confusion line from the circle center, the bigger is ξ_i and the more certain is the reader that the device is not T_i. Also, the circle with radius $e^{-\xi_{th}}$ is defined as the trustworthiness circle. i.e. Any device that has at least one point in its confusion line between the trustworthiness and certainty circles (i.e. $\underset{i=1..N_T}{max}(e^{-\xi_i}) > e^{-\xi_{th}}$) is considered a trusted device.

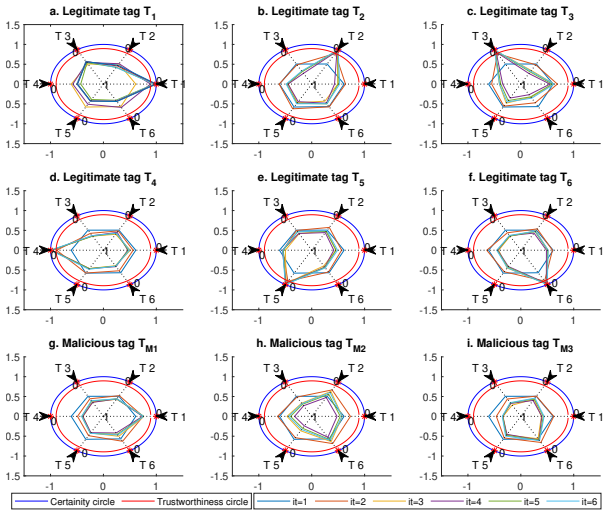

Fig. 5: Authentication confusion map for malicious and legitimate tags over time.

Fig. 5 presents the authentication confusion map for a reader communicating with 6 legitimate tags and 3 cloned malicious tags during 6 iterations of the ITI algorithm. The number of communication subcarriers are set to 16 and the distance between the reader and all the tags is uniformly distributed between $1m$ and $2m$. Each subplot of Fig. 5 presents the confusion map for one tag when communicating with the reader. For example, in Fig. 5.a, the reader is communicating with the legitimate tag T_1. It can be seen that after only 4 iterations, the reader managed to recognize the tag and make it pass the trustworthiness circle. Also, after 6 iteration the map almost converged to the certainty circle at the direction of T_1 which means that the reader is certain about the identity of the tag. Similar for all the legitimate tags, the reader managed to make them all pass the trustworthiness circle after 2 to 4 iterations and it accurately classified all the investigated tags. Also, Fig. 5.(g-i) present the authentication confusion maps for three different malicious tags. It can be seen that none of them managed to cross the trustworthiness circle and

the reader detected accurately that they are intruders to the investigated network.

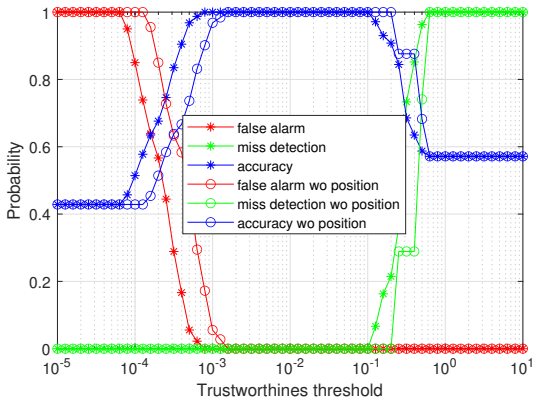

Fig. 6: Tags identification performance.

Fig. 6 presents the identification performance of the tags in terms of false alarm, miss detection and accuracy probabilities as a function of the trustworthiness threshold. The false alarm probability denotes the rate of legitimate users wrongly identified as malicious tags. The miss detection probability denotes the rate of malicious users wrongly identified as legitimate tags. Finally, the classification accuracy is defined as the rate of tags correctly classified into the corresponding tag identity out of all the investigated legitimate and malicious tags.

First, it can be seen that with the adequate choice of the trustworthiness threshold (around 10^{-2}), the proposed system detected all the malicious users and classified all the legitimate users with a perfect accuracy and without any false alarm or miss detection even after running the simulation with 1000 trials. It can be seen also that when a small trustworthiness threshold is used (The trustworthiness circle is close to the center of the confusion map), the false alarm probability rises and the reader can wrongly identify legitimate users as malicious ones. Similar, when a large trustworthiness threshold is used (The trustworthiness circle is close to the certainty center of the confusion map), the miss detection probability rises and the reader can wrongly identify malicious users as legitimate users because it trusts tags even with very high authentication errors. Also, note that including the distance error in the authentication decision ($\theta = 0.5$ compared to $\theta = 1$) improves the accuracy and false alarm performances of the system for low trustworthiness threshold. This would also be beneficial in case the channel gains cannot be accurately estimated which would increase the classification errors even with the adequate thresholds unless the distance error is used to strengthen the classification decision.

VII. Conclusion

By exploiting the physical characteristics of the RFID tags due to manufacturing imperfections in addition to the legitimate tags locations information, a novel hybrid RFID tags fingerprinting scheme is proposed to identify malicious and cloned tags and classify the legitimate ones. The RF fingerprinting is performed based on a normalized minimum power response for each tag that is estimated without the need to know the tag location. The tag location is then estimated to strengthen the accuracy of the authentication error expression and enhance the accuracy of the identification. Simulation results show that when the adequate system configuration parameters are adopted, the proposed scheme can result in perfect classification and malicious users identification performance.

Acknowledgment

This publication was made possible by the NPRP award [NPRP 10-1205-160012] from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

[1] E. Elbasani, P. Siriporn, and J. S. Choi, "A Survey on RFID in Industry 4.0," in *Internet of Things for Industry 4.0*. Springer, 2020, pp. 1–16.

[2] G. Essam, H. Shehata, T. Khattab, K. Abualsaud, and M. Guizani, "Novel hybrid physical layer security technique in RFID systems," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, 2019, pp. 1299–1304.

[3] A. Alwadi, A. Gawanmeh, S. Parvin, and J. N. Al-Karaki, "Smart solutions for RFID based inventory management systems: A survey," *Scalable Computing: Practice and Experience*, vol. 18, no. 4, pp. 347–360, 2017.

[4] S. C. G. Periaswamy, D. R. Thompson, and J. Di, "Fingerprinting RFID tags," *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 6, pp. 938–943, 2010.

[5] Q. Yang, H.-M. Wang, Y. Zhang, and Z. Han, "Physical layer security in MIMO backscatter wireless systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7547–7560, 2016.

[6] J. M. Hamamreh, H. M. Furqan, and H. Arslan, "Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1773–1828, 2018.

[7] Y. Huo, Y. Tian, L. Ma, X. Cheng, and T. Jing, "Jamming strategies for physical layer security," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 148–153, 2017.

[8] T. Hong, M.-Z. Song, and Y. Liu, "Dual-beam directional modulation technique for physical-layer secure communication," *IEEE Antennas and Wireless Propagation Letters*, vol. 10, pp. 1417–1420, 2011.

[9] Y. Ding and V. F. Fusco, "A vector approach for the analysis and synthesis of directional modulation transmitters," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 361–370, 2013.

[10] M. P. Daly and J. T. Bernhard, "Directional modulation technique for phased arrays," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 9, pp. 2633–2640, 2009.

[11] M. K. Hinders, *Intelligent Feature Selection for Machine Learning Using the Dynamic Wavelet Fingerprint*. Springer Nature, 2020.

[12] D. Zanetti, B. Danev, and S. apkun, "Physical-layer identification of UHF RFID tags," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, 2010, pp. 353–364.

[13] J. Han, C. Qian, P. Yang, D. Ma, Z. Jiang, W. Xi, and J. Zhao, "GenePrint: Generic and accurate physical-layer identification for UHF RFID tags," *IEEE/ACM Transactions on Networking*, vol. 24, no. 2, pp. 846–858, 2015.

[14] M. A. Al-Jarrah, A. Al-Dweik, E. Alsusa, and E. Damiani, "RFID reader localization using hard decisions with error concealment," *IEEE Sensors Journal*, vol. 19, no. 17, pp. 7534–7542, 2019.