Adversarial Attacks on CFO-Based Continuous Physical Layer Authentication: A Game Theoretic Study

Serkan Sarıtaş, Henrik Forssell, Ragnar Thobaben, Henrik Sandberg, and György Dán
School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, SE-10044, Stockholm, Sweden,
{saritas, hefo, ragnart, hsan, gyuri}@kth.se

***Abstract*—5G and beyond 5G low power wireless networks make Internet of Things (IoT) and Cyber-Physical Systems (CPS) applications capable of serving massive amounts of devices and machines. Due to the broadcast nature of wireless networks, it is crucial to secure the communication between these devices and machines from spoofing and interception attacks. This paper is concerned with the security of carrier frequency offset (CFO) based continuous physical layer authentication. The interaction between an attacker and a defender is modeled as a dynamic discrete leader-follower game with imperfect information. In the considered model, a legitimate user (Alice) communicates with the defender/operator (Bob) and is authorized by her CFO continuously. The attacker (Eve), by listening/eavesdropping the communication between Alice and Bob, tries to learn the CFO characteristics of Alice and aims to inject malicious packets to Bob by impersonating Alice. First, by showing that the optimal attacker strategy is a threshold policy, an optimization problem of the attacker with exponentially growing action space is reduced to a tractable integer optimization problem with a single parameter, then the corresponding defender cost is derived. Extensive simulations illustrate the characteristics of optimal strategies/utilities of the players depending on the actions, and show that the defender's optimal false positive rate causes attack success probabilities to be in the order of 0.99. The results show the importance of the parameters while finding the balance between system security and efficiency.**

***Index Terms*—physical layer authentication, carrier frequency offset, continuous authentication, adversarial learning, threshold policy, binary hypothesis testing**

I. Introduction

Wireless networks provide cheap, scalable, efficient and convenient means of communication between applications, services and infrastructures. With 5G and beyond 5G low power wireless networks, efficient communication will become feasible across various verticals with heterogeneous devices and machines, e.g., Internet of things (IoT) systems allow any devices to create, process and share data within industrial facilities as well as user spaces, such as workplaces and homes. However, the broadcast nature of radio signal propagation makes wireless communication vulnerable to spoofing and interception attacks [1]. Since conventional cryptographic authentication is often infeasible due to computational or latency constraints in the IoT, continuous physical layer authentication is emerging as a potential solution.

This work is supported in part by the Swedish Civil Contingencies Agency (MSB) through the CERCES-2 project.

A. Related Work

Physical layer authentication can be accomplished by using channel-based and radio frequency (RF) fingerprint-based schemes [2]. The clock skew [3] and the carrier frequency offset (CFO) [4] are among the device specific parameters whereas channel state information (CSI) [5], power spectral density [6] and received signal strength [7] are channel specific and location dependent parameters that can be used for authentication. [2] provides a comprehensive survey on physical layer authentication in wireless communication networks.

Recently, there have been numerous works on continuous physical layer authentication. In [4], CFO is used for authentication, and adaptive thresholds are derived based on the received signal-to-noise ratio (SNR), then the study is extended for mobile systems with time-varying CFO using Kalman filtering in [8]. In [5] multi-feature CSI-based device fingerprint is proposed and used as a basis for device-to-device continuous authentication scheme for the IoT. In [9], temporal channel variations in the dimensions of amplitude and multi-path time delay spread are integrated and represented by two one-bit quantizers. In [10], the optimal attack strategy is derived when there is some degree of correlation between channels. In [11], a spoofing detection problem is considered in which the defender chooses the threshold of the hypothesis test based on the channel gain while the attacker chooses its attack rate, and the optimal test threshold under varying environment parameters is obtained by using reinforcement learning. Another continuous physical-layer authentication technique with time-varying parameters based on an adaptive Orthogonal Frequency Division Multiplexing (OFDM) platform is proposed in [12].

In [11], a game theoretic approach is followed but with different setups/assumptions. We consider a setup similar to [4] but we consider an adversarial attacker that learns the characteristics of the legitimate device/user and attacks accordingly.

B. Contributions

(i) The carrier frequency offset (CFO) based continuous physical layer authentication security problem is modeled as a dynamic discrete leader-follower game with imperfect information between the attacker and the defender.

(ii) It is shown that the optimal attacker strategy has a threshold structure and an exponentially growing action space of the attacker can be reduced to an integer-valued scalar space, which is proper for exhaustive search.
(iii) The optimization problem of the defender is formulated as balancing between the system efficiency and security, and corresponding numerical results are provided.

II. Problem Definition

A. System Model

We consider a system with three single-antenna nodes as depicted in Fig. 1. Alice (legitimate user) communicates with Bob (defender/operator/data server) over a time-slotted channel, while the objective of Eve (attacker) is to impersonate Alice. Bob employs CFO as a decision metric for binary hypothesis testing based continuous authentication.

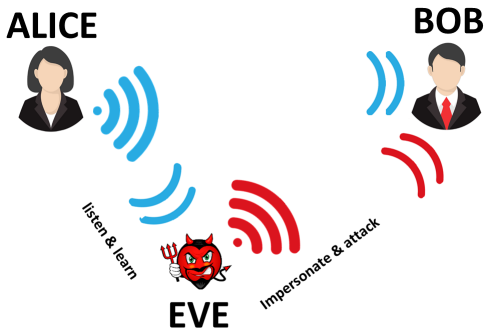


Fig. 1: Illustration of attack on physical layer authentication.

Before starting communication, Alice sends a training sequence to Bob, which Bob uses to generate an estimate of the CFO between Alice and himself. As derived in [4], the estimate can be approximated as a Gaussian random variable, that is

$$\hat{\epsilon}_{AB}^{B} \sim \mathcal{N}\left(\epsilon_{AB}, \frac{1}{4\pi^2 L_s^3 (N_s-1)\gamma_{AB}}\right), \tag{1}$$

where $\hat{\epsilon}_{AB}^{B}$ represents Bob's estimate of the CFO between Alice and himself, ϵ_{AB} denotes the true CFO between Alice and Bob normalized by the sampling rate f_s of analog-to-digital conversion (i.e. $\epsilon = \Delta f/f_s = \Delta f T_s$), N_s is the number of the identical training sequences transmitted, L_s is the length of one training sequence, i.e., the total length of the training sequence is $N_s L_s$, $\gamma_{AB} = P_{r_{AB}}/\sigma^2_{n_{AB}}$ and $P_{r_{AB}}$ are the received SNR and the received signal power when the transmission occurs from Alice to Bob, respectively, and $\sigma^2_{n_{AB}}$ is the variance of the additive zero mean complex Gaussian noise of the channel between Alice and Bob.

We consider that time is slotted, and every time-slot Alice communicates with Bob by transmitting data sequences. During every time-slot, after receiving data, Bob authenticates the sender by comparing the CFO of the received data with the CFO estimate of Alice generated from training sequences.

B. Communication Model

We assume a free-space path-loss communication model, i.e., $P_r = P_t G_t G_r \frac{\lambda^2}{(4\pi d)^2}$, where P_t and P_r are transmitted and received powers, respectively, G_t and G_r are the transmitter and receiver antenna gains, respectively, λ is the wavelength, and d is the distance between the transmitter and the receiver. The antennas are assumed to be isotropic and have no directivity, i.e., $G_t = G_r = 1$, and all players use the same carrier frequency, i.e., λ is constant for the whole setup. Thus, we can define $C \triangleq \frac{\lambda^2}{(4\pi)^2}$ so that $P_r = P_t C \frac{1}{d^2}$. Furthermore, a static noise is assumed in the environment, i.e., σ_n^2 is constant everywhere. For our setup, P_{t_A} and P_{t_E} denote the transmission signal powers of Alice and Eve, respectively, d_{XY} denotes the distance between X and Y, and γ_{XY} the received SNR when the message is transmitted by X and received by Y, where $\{X, Y\} \in \{A(\text{lice}), B(\text{ob}), E(\text{ve})\}$.

Remark 2.1: Since $\gamma_{AB} = \gamma_{AE}\frac{d_{AE}^2}{d_{AB}^2}$, Eve can calculate Bob's received SNR when she receives a message from Alice. Thus, she can adjust her transmission power P_{t_E} such that Bob's received SNRs γ_{AB} and γ_{EB} are equal, i.e., Bob cannot distinguish the origin of the messages from SNRs.

C. Attacker Model

We consider an attacker Eve that is capable of eavesdropping the communication between Alice and Bob and transmitting malicious sequences to Bob. In every time-slot, the attacker has two actions: listening and attacking. By listening, the attacker learns the CFO characteristics of Alice whereas attacking means transmitting malicious packets to Bob by impersonating Alice. If the attack is successful, i.e., if the packets from Eve are recognized as from Alice, Eve gets a reward r, with a discount factor $\beta \in (0,1)$ for future rewards, and Bob gets a penalty r. While designing an optimal strategy to maximize her discounted reward, Eve must consider her total energy constraint, i.e., she can spend at most Ξ energy while listening and attacking. While listening (to Alice), Eve spends $\kappa_l(\gamma_{AE})$ amount of energy in one time-slot, where $\kappa_l(\gamma_{AE})$ is a non-increasing function of γ_{AE}, and while attacking (i.e., transmitting packets to Bob), Eve spends $\kappa_a(P_{t_E})$ amount of energy in one time-slot, where $\kappa_a(P_{t_E})$ is a monotonic increasing function of P_{t_E}. In view of the above, we make the following assumptions:

Assumption 2.1:

(i) The locations are fixed and Eve knows the locations of Alice and Bob. On the contrary, Bob knows only the location of Alice and has a prior belief on the possible locations of Eve.
(ii) Alice uses constant transmission power P_{t_A}, and due to Remark 2.1, Eve also sends messages with constant transmission power P_{t_E}.
(iii) As a consequence of the above assumptions, γ_{AE} and P_{t_E} are constants, resulting in constant listening and transmission energy $\kappa_l \triangleq \kappa_l(\gamma_{AE})$ and $\kappa_a \triangleq \kappa_a(P_{t_E})$, respectively. Furthermore, transmission requires more energy than listening, i.e., $\kappa_a > \kappa_l$.

D. Attacker's CFO Manipulation

Since ϵ_{AB} is the CFO between Alice and Bob, we have $\epsilon_{AB} = \epsilon_{AE} + \epsilon_{EB}$, where ϵ_{XY} denotes the CFO between X and Y where $\{X, Y\} \in \{A(\text{lice}), B(\text{ob}), E(\text{ve})\}$. Then, by generating an estimate of ϵ_{AE}, Eve is able to impersonate Alice by shifting her carrier frequency accordingly, i.e., $\hat{\epsilon}^B_{AB} = \hat{\epsilon}^E_{AE} + \hat{\epsilon}^B_{EB}$, where $\hat{\epsilon}^E_{AE}$ represents Eve's estimate of the CFO between Alice and her, $\hat{\epsilon}^B_{EB}$ represents Bob's estimate of the original CFO between Eve and him, and $\hat{\epsilon}^B_{AB}$ represents the overall CFO (which is manipulated by Eve) perceived by Bob from Eve. The goal of Bob is to distinguish the original messages from Alice with CFO $\hat{\epsilon}^B_{AB}$ and the malicious messages with CFO $\hat{\epsilon}^B_{AB}$, transmitted by Eve.

Assume that the lengths of the training sequences and data sequences are the same, i.e., they are equal to L_s, and Eve eavesdrops/overhears the sequences from Alice (which is composed of N_s training/data sequences with length L_s of each) x times. Then, Eve can generate a CFO estimate

$$\hat{\epsilon}^E_{AE} \sim \mathcal{N}\left(\epsilon_{AE}, \frac{1}{4\pi^2 L_s^3 x (N_s - 1)\gamma_{AE}}\right). \tag{2}$$

Notice that as the number x of captured sequences increases, the variance of Eve's CFO estimates becomes smaller, i.e., Eve's estimate gets better.

Thus, the problem faced by Bob is to distinguish the CFO of the legitimate messages with distribution

$$\hat{\epsilon}^B_{AB} \sim \mathcal{N}\Big(\epsilon_{AB}, \underbrace{\frac{1}{4\pi^2 L_s^3 (N_s - 1)\gamma_{AB}}}_{\sigma_A^2}\Big), \tag{3}$$

from the CFO of the malicious/adversarial messages with distribution

$$\hat{\epsilon}^B_{AB} \sim \mathcal{N}\Bigg(\epsilon_{AB}, \underbrace{\frac{1}{4\pi^2 L_s^3 (N_s - 1)}\left(\frac{1}{x\gamma_{AE}} + \frac{1}{\gamma_{EB}}\right)}_{\sigma_E^2(x)}\Bigg). \tag{4}$$

E. Defender's Decision Rule

The decision of the defender on whether accepting or rejecting the messages with the CFO ϵ can be expressed as the following binary hypothesis testing problem:

$$\begin{aligned} \mathcal{H}_0 &: \epsilon \sim \mathcal{N}(\epsilon_{AB}, \sigma_A^2) \\ \mathcal{H}_1 &: \epsilon \sim \mathcal{N}(\epsilon_{AB}, \sigma_E^2(x)) \end{aligned}. \tag{5}$$

Here, the defender knows only the first hypothesis, i.e., from the perspective of the defender, the problem reduces to decide whether the CFO of the received sequence belongs to the hypothesis $\mathcal{H}_0$ (i.e., the sequence is coming from the legitimate source) or not. For this purpose, we assume that the defender applies a distance-based decision rule. In particular, the sequences with the CFO ϵ satisfying $|\epsilon - \epsilon_{AB}| \leq \tau$ are accepted, where τ is a parameter to be determined by the defender that satisfies the desired false positive (FP) rate η, i.e., $\Pr(|\epsilon - \epsilon_{AB}| \geq \tau | \mathcal{H}_0) = \eta \Rightarrow \tau = -\sigma_A \Phi^{-1}\left(\frac{\eta}{2}\right)$, where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution.

Meanwhile, the attacker is aware of both hypotheses. In particular, by knowing the defender's decision rule and having a complete knowledge of $\sigma_E(x)$, the attacker can calculate the probability of a successful attack. In particular, for given[1] η and $\sigma_A^2 \leq \sigma_E^2(x)$, the attack detection probability $\Pr(|\epsilon - \epsilon_{AB}| \geq \tau | \mathcal{H}_1)$, or equivalently, the Receiver Operating Characteristic (ROC) curve[2] (i.e., the true positive (detection) rate) can be characterized as

$$\text{ROC}(\eta, x) = \Pr(|\epsilon - \epsilon_{AB}| \geq \tau | \mathcal{H}_1) = 2\Phi\left(\tfrac{\sigma_A}{\sigma_E(x)} \Phi^{-1}\left(\tfrac{\eta}{2}\right)\right). \tag{6}$$

Note that $\text{ROC}(\eta, x)$ is an increasing function of $\sigma_E(x)$ (and a decreasing function of x), which implies that as x increases, or equivalently, as Eve listens more, the attack success probability

$$\alpha(\eta, x) \triangleq 1 - \text{ROC}(\eta, x)$$

also increases for a given FP rate η.

The defender is aware of the existence of the attacker; however, he does not know the exact location of her (Assumption 2.1), and he gets a penalty r if he cannot detect the attack. Besides the attack cost, the defender should also consider the delay cost caused by false alarms. The goal of the defender is to minimize his total cost consisting of the attack and delay costs by adjusting the FP rate η, i.e., he tries to find a balance between system security and efficiency.

F. Game Formulation

Our focus is on the interaction between the operator (Bob) and the attacker (Eve), which we model as a dynamic discrete Stackelberg game with imperfect information. In the game, the defender is the leader, and chooses a defense strategy η, which is known to the attacker. As a leader, the defender's goal is to minimize his total cost by anticipating the follower's best response, whereas the attacker, as a follower, aims to maximize her discounted reward.

III. Optimal Attack Strategy

A. States and State Transitions

In every time-slot, the attacker decides whether to listen or to attack. While determining an optimal strategy, the attacker keeps track of x (total number of listening time-slots) and Ξ_{rem} (remaining energy). When the game starts, i.e., at the zero-th time-slot, the state is $S(0) = (0, \Xi)$. Let us consider that at the beginning of the t-th time-slot Eve has been listening to Alice for x time-slots, and the remaining energy of Eve is Ξ_{rem}, i.e., the state of the attacker is $S(t-1) = (x, \Xi_{\text{rem}})$. If the attacker listens, the next state will be $S(t) = (x+1, \Xi_{\text{rem}} - \kappa_l)$. If the attacker attacks (transmits a sequence), the next state will be $S(t) = (x, \Xi_{\text{rem}} - \kappa_a)$, and the attacker gains a reward r with attack success probability[3] $\alpha(\eta, x) \triangleq 1 - \text{ROC}(\eta, x)$.

[1] Since $\gamma_{EB} = \gamma_{AB}$ is assumed, $\sigma_A^2 \leq \sigma_E^2(x)$ holds.

[2] Unless otherwise stated, $\text{ROC}(\eta)$ stands for $\text{ROC}(\eta, \sigma_E(x))$ to avoid complex notations, i.e., ROC curve is a function of FP rate η and $\sigma_E(x)$. Similarly, we generally omit the relation between $\sigma_E(x)$ and x, and prefer σ_E for simplicity/readability.

[3] Although the attack success probability $\alpha(\eta, x)$ is a function of FP rate η and number-of-listening time-slots x, since η is known by the attacker, we will prefer $\alpha(x)$ from the attacker's perspective.

B. Attacker Reward as a Dynamic Programming Recursion

We will characterize the reward of the attacker using a dynamic programming approach. At the beginning, Eve has no observation about Alice and has full energy, the total attacker reward is expressed as $J(0,P)$. It can be seen that, the attacker reward is an increasing function in both of its parameters, i.e., the attacker gains more if she has more energy, and the success probability of attack increases as Eve's amount of observation about Alice increases.

Let the attacker's current state be (x,Ξ_{rem}). If Eve chooses to listen for the next time-slot, since the future rewards are discounted by β, the recursive relation of attacker's reward is

$$J(x,\Xi_{\text{rem}}) = \beta\, J(x+1,\Xi_{\text{rem}}-\kappa_l)\,.$$

Similarly, if Eve decides to attack for the next time-slot, the recursive relation of the attacker's reward is

$$J(x,\Xi_{\text{rem}}) = \beta\ \left(\alpha(x)r + J(x,\Xi_{\text{rem}}-\kappa_a)\right)\ .$$

Thus, based on the attacker's actions, we have

$$\begin{aligned} J(x,\Xi_{\text{rem}}) = \max\Big\{&\beta J(x+1,\Xi_{\text{rem}}-\kappa_l)\mathbb{1}_{\{\Xi_{\text{rem}}\geq\kappa_l\}},\\ &\beta\left(\alpha(x)r + J(x,\Xi_{\text{rem}}-\kappa_a)\right)\mathbb{1}_{\{\Xi_{\text{rem}}\geq\kappa_a\}}\Big\}, \end{aligned} \tag{7}$$

where $\mathbb{1}_{\{D\}}$ represents the indicator function of an event D.

C. Characterization of Optimal Attacker Strategy

In this part, we first investigate the form of an optimal attacker strategy, then state the integer optimization problem of the attacker over the reduced strategy space.

Theorem 3.1: An optimal attacker strategy is in the form of listening (L) during x consecutive time-slots followed by attacking (A) during y consecutive time-slots where x and y are non-negative integers. Equivalently, an optimal attacker strategy is in the form of $\{\text{LL...LLAA...AA}\}$ with $x\geq 0$ times (L) and $y\geq 0$ times (A).

Proof: To prove this theorem, we first investigate the last action(s) of the attacker, then the possible transitions between (L) and (A) actions. Before doing so, we make the following observations on $\alpha(x)$.

Lemma 3.1: $\alpha(x)$ is a monotone increasing concave function of x with $\alpha(0)=0$ and $\alpha(\infty)=1-\eta$. Furthermore, $\frac{\alpha(x+1)}{\alpha(x)}$ is a monotone decreasing function of x.

Corollary 3.1: If $\frac{\alpha(1)}{\alpha(0)} > \frac{1}{\beta}$, then there exists a critical value $\widetilde{x}>0$ such that $\frac{\alpha(\widetilde{x}+1)}{\alpha(\widetilde{x})} = \frac{1}{\beta}$. Otherwise, i.e., if $\frac{\alpha(1)}{\alpha(0)} \leq \frac{1}{\beta}$, it is defined as $\widetilde{x}=0$.

For any given state $S(t)=(x,\Xi_{\text{rem}})$, the attacker chooses the action (i.e., listening (L) or attacking (A)) with a higher reward in (7). Regarding the terminal conditions, we have the following results.

Observation 3.1:

(i) The last action of the attacker must be attacking (A) since listening (L) does not give any additional reward, it only increases the probability of a successful attack.

(ii) If the remaining energy of the attacker is less than κ_a, then the attacker cannot gain any reward, i.e., $J(x,\Xi_{\text{rem}})=0$ if $\Xi_{\text{rem}}<\kappa_a$.

(iii) Thus, the last action of the attacker must be attacking (A) and chosen when $\kappa_a \leq \Xi_{\text{rem}} < 2\kappa_a$.

Lemma 3.2: Let the remaining energy Ξ_{rem} of the attacker be $\kappa_a \leq \Xi_{\text{rem}} < 2\kappa_a$. If $x<\widetilde{x}$, then the attacker prefers (L) over (A).

Proof: If the attacker chooses (L), she will get a reward

$$\begin{aligned} &J(x,\Xi_{\text{rem}}) = \beta\, J(x+1,\Xi_{\text{rem}}-\kappa_l)\\ &\overset{(a)}{\geq} \beta\left(\beta\left(\alpha(x+1)r + J(x+1,\Xi_{\text{rem}}-\kappa_l-\kappa_a)\right)\right)\\ &= \beta^2\,\alpha(x+1)r \overset{(b)}{>} \beta\,\alpha(x)\,r = \beta\left(\alpha(x)r + J(x,\Xi_{\text{rem}}-\kappa_a)\right), \end{aligned}$$

where (a) follows from (7), (b) holds since $\frac{\alpha(x+1)}{\alpha(x)} > \frac{1}{\beta}$ for $x<\widetilde{x}$ by Lemma 3.1, and the equalities hold since $\Xi_{\text{rem}}<2\kappa_a$, we have $J(x,\Xi_{\text{rem}}-\kappa_l-\kappa_a) = J(x,\Xi_{\text{rem}}-\kappa_a) = 0$ by Observation 3.1. Thus the listening reward $\beta\, J(x+1,\Xi_{\text{rem}}-\kappa_l)$ is greater than the attacking reward $\beta\left(\alpha(x)r + J(x,\Xi_{\text{rem}}-\kappa_a)\right)$. ∎

Thus we have that the last action must be (A) by Observation 3.1, and at some point (for the case when $\widetilde{x}>0$), there must be a (L) action by Lemma 3.2. Thus, there must be some transitions between (L) and (A), which we investigate next.

Lemma 3.3: Consider the actions chosen in two consecutive time-slots, listening-first-then-attacking (LA) and attacking-first-then-listening (AL) actions. If $x<\widetilde{x}$, the attacker prefers (LA); otherwise, i.e., if $x\geq\widetilde{x}$, the attacker prefers (AL).

Proof: Let the attacker state be (x,Ξ_{rem}). If the attacker first listens then attacks (LA), her reward is

$$\begin{aligned} J(x,\Xi_{\text{rem}}) &= \beta\, J(x+1,\Xi_{\text{rem}}-\kappa_l)\\ &= \beta^2\ \left(\alpha(x+1)r + J(x+1,\Xi_{\text{rem}}-\kappa_a-\kappa_l)\right)\ . \end{aligned}$$

If the attacker first attacks then listens (AL), her reward is

$$\begin{aligned} J(x,\Xi_{\text{rem}}) &= \beta\ \left(\alpha(x)r + J(x,\Xi_{\text{rem}}-\kappa_a)\right)\\ &= \beta^2\ \left(\frac{\alpha(x)}{\beta}r + J(x+1,\Xi_{\text{rem}}-\kappa_a-\kappa_l)\right)\ . \end{aligned}$$

Then, the decision rule between (LA) and (AL) becomes

$$\frac{\alpha(x+1)}{\alpha(x)} \underset{\text{AL}}{\overset{\text{LA}}{\gtrless}} \frac{1}{\beta} = \frac{\alpha(\widetilde{x}+1)}{\alpha(\widetilde{x})} \Rightarrow x \underset{\text{AL}}{\overset{\text{LA}}{\lessgtr}} \widetilde{x}\,. \tag{8}$$

∎

Corollary 3.2:

(i) If $\widetilde{x}>0$, the first action of the attacker is (L) and the last action is (A), where a single transition from (L) to (A) occurs when $x<\widetilde{x}$.

(ii) If $\widetilde{x}=0$, i.e., if $\frac{\alpha(1)}{\alpha(0)} \leq \frac{1}{\beta}$ holds, the attacker always chooses (A) without any (L).

As stated in Corollary 3.2, for both cases ($\widetilde{x}>0$ and $\widetilde{x}=0$), an optimal attacker has $x\geq 0$ consecutive (L) actions followed by $y\geq 0$ consecutive (A) actions, which proves the theorem. ∎

Remark 3.1: Since there is not enough energy for (A) when $\Xi<\kappa_a$, y should be equal to zero, which results in a zero attacker reward, i.e., $J(0,\Xi)=0$. Thus, it can be assumed that $\Xi\geq\kappa_a$ and $y\geq 1$.

An optimal attacker strategy with $x \geq 0$ consecutive (L) and then $y \geq 1$ consecutive (A) results in the reward[4]

$$J(0,\Xi) = \mathcal{R}(x,y) \triangleq \beta^x \alpha(x) r \sum_{i=1}^{y} \beta^i \,. \tag{9}$$

Thus, the optimal attacker strategy is found by solving the integer optimization problem

$$\max_{x\geq 0,\, y\geq 1} \mathcal{R}(x,y) \text{ s.t. } x\kappa_l + y\kappa_a \leq \Xi \,. \tag{10}$$

Lemma 3.4: Let the integers x and y satisfy $x\kappa_l + (y+1)\kappa_a \leq \Xi$, then consider three strategies $\mathcal{S}_1 = (x,y)$, $\mathcal{S}_2 = (x,y+1)$, and $\mathcal{S}_3 = (x+1,y)$.

(i) $\mathcal{S}_2$ is always preferred over $\mathcal{S}_1$,

(ii) $\mathcal{S}_3$ is preferred over $\mathcal{S}_1$ if $x \leq \widetilde{x}$, and vice versa.

Thus, the attacker utilizes her energy Ξ as much as possible by listening and/or attacking more.

Based on Lemma 3.4, it is sufficient to consider the cases $0 \leq x \leq \lfloor \widetilde{x} \rfloor$ and the corresponding $\left\lceil \frac{\Xi - \lfloor \widetilde{x} \rfloor \kappa_l}{\kappa_a} \right\rceil \leq y \leq \left\lfloor \frac{\Xi}{\kappa_a} \right\rfloor$, where $\lceil \cdot \rceil$ and are $\lfloor \cdot \rfloor$ the ceiling and floor functions, respectively. To be more precise, for a given number of (A) actions, it is possible to evaluate the corresponding number of (L) actions. In particular, let us define $\chi(y) \triangleq \frac{\Xi - y\kappa_a}{\kappa_l}$ and $\underline{\chi}(y) \triangleq \lfloor \chi(y) \rfloor$. Thus, the two-dimensional integer optimization problem of the attacker in (10) reduces to following integer problem:

$$\max_{y} \mathcal{R}(\underline{\chi}(y), y) \text{ s.t. } \left\lceil \frac{\Xi - \lfloor \widetilde{x} \rfloor \kappa_l}{\kappa_a} \right\rceil \leq y \leq \left\lfloor \frac{\Xi}{\kappa_a} \right\rfloor . \tag{11}$$

Note that the size of the one-dimensional feasible region in (11) is $\left\lfloor \frac{\Xi}{\kappa_a} \right\rfloor - \left\lceil \frac{\Xi - \lfloor \widetilde{x} \rfloor \kappa_l}{\kappa_a} \right\rceil + 1$, that is very small compared to the size of the two-dimensional search space in (10), which is $\mathbf{card}(\{(x,y) : x\kappa_l + y\kappa_a \leq \Xi\}) \approx \frac{1}{2}(\lfloor \frac{\Xi}{\kappa_l} \rfloor + 1)(\lfloor \frac{\Xi}{\kappa_a} \rfloor + 1)$, and significantly smaller than the size of the exponentially growing original domain of strategies in (7), which is $\sum_{(x,y):x\kappa_l + y\kappa_a \leq \Xi} \frac{(x+y)!}{x!y!}$. Here, in (11), the optimal attacker strategy can be found by a simple exhaustive search.

Remark 3.2: Since $\mathcal{R}(\underline{\chi}(y+1), y+1) + \mathcal{R}(\underline{\chi}(y-1), y-1) < 2\mathcal{R}(\underline{\chi}(y), y)$ does not always hold, i.e., $\mathcal{R}(\underline{\chi}(y), y)$ may not be a concave function of y, the optimization problem in (11) may have more than one local maxima.

IV. Optimization Problem of Defender

The defender sets a FP rate η to minimize his total cost. For larger FP rates, the true detection probability $\text{ROC}(\eta, x)$ of the defender's decision rule increases, i.e., the attack success probability $\alpha(\eta, x)$ reduces, which results in a smaller attack cost. On the other hand, increasing the FP rate reduces the system performance and increases the average delay and the corresponding cost. Therefore, the optimization of the FP rate requires a careful analysis considering both effects. However, the defender cannot do a simple analysis on the attack and delay cost based on only the FP rate. Even though the defender can anticipate the best response of the attacker as a leader for a given attacker location, due to the uncertainty on the attacker's location, the defender must take the expectation of the attack costs based on his prior on the location of the attacker.

In particular, for any FP rate η and for any given attacker location (i.e., for given d_{AE} and d_{EB}), letting y^* be the maximizer of (11), the attack cost is $C_A(\eta, d_{AE}, d_{EB}) \triangleq \alpha(\underline{\chi}(y^*)) r y^*$, that can be found by setting the discount factor $\beta = 1$ in (9). Note that y^* depends on η, d_{AE} and d_{EB}, and the expected (average) attack cost for a FP rate η is $\text{E}C_A(\eta) = \mathbb{E}_{d_{AE}, d_{EB}}[C_A(\eta, d_{AE}, d_{EB})]$, where $\mathbb{E}_M$ denotes the expectation over the random variable M.

After presenting the attack cost, now we can investigate the delay cost of the defender. Let m' be the number of time-slots to transmit m sequences on average, we have $m = (1-\eta)m'$, which results in the expected delay per time-slot $\frac{m'-m}{m} = \frac{\eta}{1-\eta}$. Then, we define the defender's delay cost $C_D(\eta)$ as $C_D(\eta) = N\kappa_\eta \frac{\eta}{1-\eta}$, where N is the length of the analysis window[5], κ_η is a cost coefficient for the delay.

The total cost $C(\eta)$ of the defender is the sum of the expected attack cost and the delay cost as follows:

$$C(\eta) = \mathbb{E}_{d_{AE}, d_{EB}}[C_A(\eta, d_{AE}, d_{EB})] + N\kappa_\eta \frac{\eta}{1-\eta} \,. \tag{12}$$

Here, if $N\kappa_\eta$ is large enough, the delay cost will be dominant and the defender will prefer smaller η to reduce the delay cost, i.e., $C(\eta)$ will be an increasing function of η. On the other hand, if $N\kappa_\eta$ is small enough, the attack cost will be dominant and the defender will prefer larger η to reduce the attack cost, i.e., $C(\eta)$ will be a decreasing function of η. For intermediate values of $N\kappa_\eta$, the total cost $C(\eta)$ can be of any behavior, thus the defender will select an optimal η accordingly.

V. Simulation Results

In the following we illustrate the optimal strategy/cost of the attacker with respect to her location and corresponding expected cost of the defender. For the simulations, we use the parameters[6] in Table I. For reliable wireless communication and continuous authentication security, we consider $\eta \leq 0.01$. Furthermore, we consider a 2-D uv-plane (in meters) in which Alice's location is $(0,0)$, which is indicated as a green dot, and Bob's location is $(10,0)$, which is indicated as a blue dot. From Bob's perspective, Eve can be at any point in the region $[2,8] \times [2,5]$ with uniform distribution, which is shaded as red.

Table I: Default parameters.

$\frac{1}{4\pi^2 L_s^3 (N_s - 1)}$	γ_{AB}	Ξ	κ_l	κ_a	β	r	N
$3 \cdot 10^{-9}$	37dB	2000	0.5	$1.5 + 0.05 d_{EB}^2$	0.9999	1	2000

For any particular location and FP rate η, the defender, as a leader, can calculate the optimal strategy of the attacker

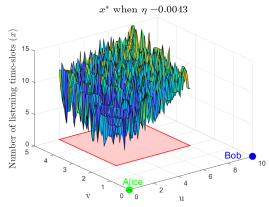
[4]From now on, the attacker strategy will be searched within the reduced action space as described in Theorem 3.1, and to describe the attacker reward, $\mathcal{R}(x,y)$ will be used instead of $J(0,\Xi)$.

[5]We define the goal of the defender is to minimize his total cost over N time-slots, which is assumed to be greater than the length of the longest attack, i.e., the sum of the numbers of listening and attacking time-slots.

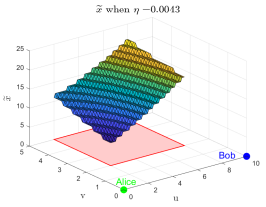
[6]Distances are normalized to a unit distance, and hence, the energies are also normalized and scale with the unit distance.

(Figure 2-((a)-(d))). Then, based on his prior on the attacker's location, he can derive his overall cost with respect to η (Figure 2-(e)). For $\kappa_\eta = 0.4$, the optimum false alarm rate that minimizes the overall defender cost is $\eta = 0.0043$, which is the FP rate used in Figure 2-((a)-(d)).

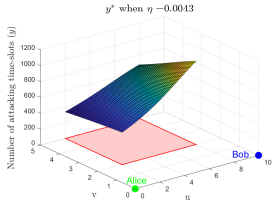
As the distance between the attacker and the defender decreases, the attacker is able to attack longer (Figure 2-(c)) and gets a greater reward (Figure 2-(d)) since the transmission cost reduces with the distance. Moreover, as the distance between the attacker and the user increases, the learning rate of the attacker by listening reduces, thus the attacker needs to listen for more time slots to achieve higher attack success rate (Figure 2-(b)). However, as Figure 2-(a) shows, the behavior of the optimal number of listening is fluctuating due to the energy constraint and the higher ratio between the transmission and listening energies. The corresponding attack success probabilities are in the order of 0.99.



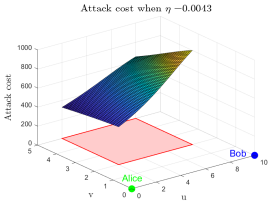
(a) Optimal number of listening vs. attacker location.



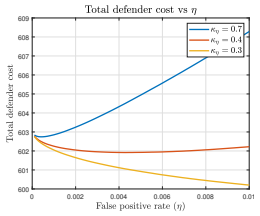
(b) Threshold ($\widetilde{x}$) vs. attacker location.



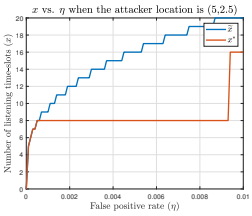
(c) Optimal number of attacking vs. attacker location.



(d) Optimal attack cost vs. attacker location.



(e) Total defender cost vs. η.



(f) Number of listening vs. η.

Fig. 2: (a)-(d) Optimal attack strategies/cost vs. attacker location when $\eta = 0.0043$. (e)-(f) Defender cost and number of (L) vs. η.

As described in Section IV, the defender tries to keep a balance between attack cost and delay cost by adjusting the FP rate, and this balance highly depends on the delay cost coefficient κ_η. A corresponding analysis is shown in Figure 2-(e). The overall defender cost decreases for small κ_η whereas it increases for large κ_η. Furthermore, as it can be seen in Figure 2-(f), as the FP rate increases, since the detection probability will also increase, the attacker needs more observation in order to achieve a successful attack with high probability.

VI. Conclusions and Extensions

We considered a CFO based continuous physical layer authentication security problem as a dynamic discrete leader-follower game with imperfect information between an attacker and a defender. After deriving a backward recursion for the optimal attacker reward, we characterize the optimal attacker strategy. To do this, we first reduce the exponentially growing action space to a two dimensional integer optimization problem, then to a one dimensional, possibly non-convex optimization problem, which can be easily solved by a simple exhaustive search. Then, based on the optimal strategy of the attacker, we expressed the expected cost of the defender consisting of the delay cost due to false alarms and attack cost, considering the imperfect information of the defender about the location of the attacker. Extensive simulations illustrate the characteristics of optimal strategies/utilities of the players depending on the actions and locations.

Our model has many possible interesting extensions. Of particular interest are the case when the feature vectors are multi-dimensional (i.e., besides the CFO, other device/channel specific properties can be utilized), when the defender applies moving target defense strategies (e.g., frequency hoping techniques), and the case of dynamic/moving players.

References

[1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[2] L. Bai, L. Zhu, J. Liu, J. Choi, and W. Zhang, "Physical layer authentication in wireless communication networks: A survey," *Journal of Comm. and Information Networks*, vol. 5, no. 3, pp. 237–264, 2020.

[3] K. O. Saputra, W. Teng, and Y. Chu, "A clock skew replication attack detection approach utilizing the resolution of system time," in *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, vol. 3, 2015, pp. 211–214.

[4] W. Hou, X. Wang, and J. Chouinard, "Physical layer authentication in OFDM systems based on hypothesis testing of CFO estimates," in *IEEE ICC*, 2012, pp. 3559–3563.

[5] B. Yu, C. Yang, and J. Ma, "Continuous authentication for the Internet of Things using channel state information," in *IEEE GlobeCom*, 2019, pp. 1–6.

[6] J. K. Tugnait, "Wireless user authentication via comparison of power spectral densities," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1791–1802, 2013.

[7] M. Mohamed and M. Cheffena, "Received signal strength based gait authentication," *IEEE Sensors Journal*, vol. 18, no. 16, pp. 6727–6734, 2018.

[8] W. Hou, X. Wang, J. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Trans. on Communications*, vol. 62, no. 5, pp. 1658–1667, 2014.

[9] J. Liu and X. Wang, "Physical layer authentication enhancement using two-dimensional channel quantization," *IEEE Trans. on Wireless Communications*, vol. 15, no. 6, pp. 4171–4182, 2016.

[10] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Trans. on Wireless Communications*, vol. 11, no. 7, pp. 2564–2573, 2012.

[11] L. Xiao, T. Chen, G. Han, W. Zhuang, and L. Sun, "Game theoretic study on channel-based authentication in MIMO systems," *IEEE Trans. on Vehicular Technology*, vol. 66, no. 8, pp. 7474–7484, 2017.

[12] X. Wang, F. J. Liu, D. Fan, H. Tang, and P. C. Mason, "Continuous physical layer authentication using a novel adaptive OFDM system," in *IEEE ICC*, 2011, pp. 1–5.