A Study on Poisoning Attack on Channel-based Physical-Layer Authentication in WLAN Systems

Assan Ralph Kumah, Gyuho Lee[+], Hyeonseon Min, Taehoon Kim[+*], and Inkyu Bang[*]

Department of Intelligence Media Engineering, Hanbat National University

[+]Department of Computer Engineering, Hanbat National University

{ralphassan, ghlee, hsmin}@edu.hanbat.ac.kr, {ikbang, thkim}@hanbat.ac.kr

Abstract

In this paper, we investigate a poisoning attack on a channel state information (CSI) based physical-layer authentication algorithm during the training phase. We show experimental results using a software-defined radio (SRD) based testbed. Furthermore, we introduce a possible candidate to identify and mitigate the poisoning attack.

I. Introduction

The rapid increase in Internet of Things (IoT) technologies has made wireless networks ubiquitous. However, it raises security concerns due to the shared nature of the wireless medium and thus there is increasing demand for higher-performing authentication schemes. Physical-layer authentication (PLA) utilizes physical-layer attributes like channel impulse response (CIR), channel frequency response (CFR), received signal strength (RSS), channel state information (CSI), and radio frequency fingerprint (RFF) to authenticate wireless transmitters.

There have been several studies related to PLA. The authors of [1] investigated three types of PLA techniques based on channel information, radio frequency, and identity watermarks. Liu et al. proposed a PLA scheme to consider user profiles built from CSI in a stationary environment and investigated the temporal correlation of CSI for mobile user authentication [2].

In this paper, we investigate a poisoning attack on a channel state information (CSI) based physical-layer authentication algorithm during the training phase. Our main contribution is to perform a software-defined radio (SRD) based experiment. Further, we discuss a possible candidate to identify and mitigate the poisoning attack.

II. System Model

In this section, we introduce the basic setup for our system model, describe the threat model, and define our performance metrics.

The system model consists of a stationary receiver (Alice) and N legitimate transmitters (Bobs) equipped with a single antenna differently located as in Fig. 1. We also consider one illegitimate transmitter (Mallory) in the experimental setup. The objective of the malicious user is to contaminate the training data used by the machine learning model during its profiling phase for authorized users' wireless fingerprints. We consider IEEE 802.11 based wireless local area networks (WLAN) using GNU-radio and USRP as in Fig. 2.

* Corresponding Authors: Taehoon Kim and Inkyu Bang

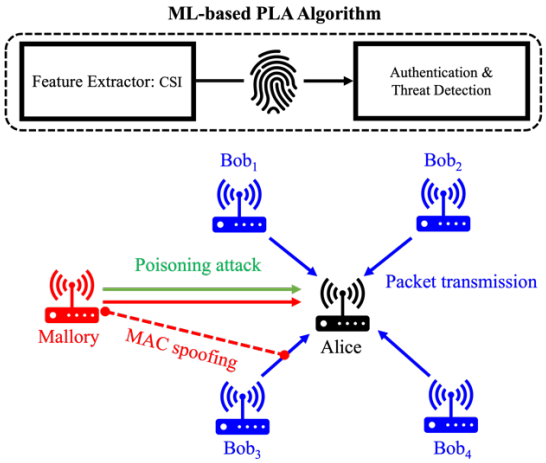

Fig. 1. An example of system model with $N = 4$

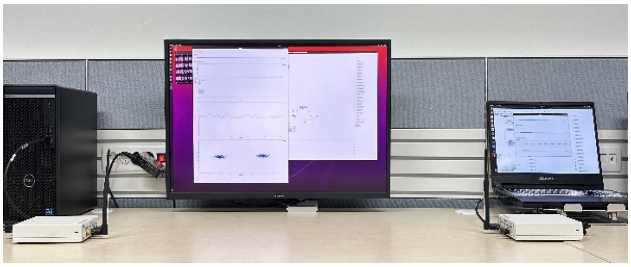

Fig. 2. Experimental setup

In IEEE 802.11 based WLAN system (e.g., Wi-Fi), the CSI comprises a vector of 52 complex values, each representing the channel response over one OFDM subcarrier. We average out the 52 CSI values for each packet for preprocessing. We employ Support Vector Classification (SVC) to train a model that uses CSI as input features and MAC addresses as the class labels for identifying unique devices.

We measure the quality of our machine-learning based authenticator using the following metrics.

(1) Location Discrimination Accuracy (LDA): The system's ability to correctly identify authorized users' locations based on trained fingerprints.

(2) Location Misclassification Rate (LMR): Probability of a legitimate packet from one location being incorrectly classified as originating from another location.

III. Poisoning Attack

Mallory's approach involves actively poisoning the training data by impersonating legitimate users and injecting false CSI samples into the authentication system as shown in Fig. 3. Specifically, the attacker spoofs the MAC addresses of an authorized device and appends it to the false CSI values. Consequently, the legitimate receiver obtains a mixture of CSI samples associated with a single MAC address, originating from two distinct devices at different physical locations. This contaminated training set compromises the machine learning model's ability to accurately characterize the unique wireless fingerprints of legitimate users during the profiling phase. By strategically poisoning the training data in this manner, the attacker can potentially cause the authentication model to become confused and misclassify malicious devices as authorized entities during the operational phase.

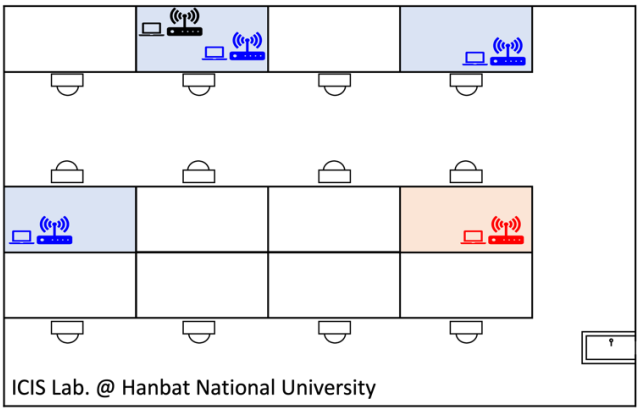

Fig. 3. Locations of Alice, Bobs, Mallory

Remark: The K-means algorithm can be one of the candidate techniques to mitigate the poisoning attack on the PLA training phase. For example, we set $K = 2$, resulting in two clusters. We can set a threshold value about the distance between the two centroids to detect the poisoning attack. Detailed analysis and experiment remain for future work.

IV. Numerical Results

In this section, we provide numerical results to evaluate the performance of our SVM-based PLA algorithm based on the performance metrics in two different scenarios: Normal conditions and the presence of a poisoning attacker. The machine learning PLA achieves an accuracy of 0.98 (98%) under normal conditions. However, we observe a decrease in accuracy to 0.79 (79%) in the case of the poisoning attack as shown in Tables 1 and 2.

Table 1. Confusion matrix (No attack)

Location	1	2	3
1	**99.91%**	0.09%	0.00%
2	0.00%	**96.40%**	3.60%
3	0.00%	2.30%	**97.70%**

Table 2. Confusion matrix (Poisoning attack)

Location	1	2	3
1	90.80%	0.18%	9.02%
2	0.00%	96.91%	3.09%
3	**54.48%**	**2.57%**	**38.96%**

Table 3 shows the effect of the attack on each location in terms of misclassification rate. Note that the CSI sample of the spoofer is very similar to the CSI measurement at location 3. As a result, the SVM-PLA finds it difficult to classify CSI samples from location 1 and location 3 correctly.

Table 3. Misclassification rate

Location	1	2	3
No Attack	0.0001	0.036	**0.028**
Poisoning Attack	0.100	0.036	**0.700**

V. Conclusion

In this paper, we investigated the effect of the poisoning attack on a CSI-based SVM-PLA scheme. Additionally, a possible candidate to mitigate the attack is discussed. Future works include extending the work for mobile users and finding appropriate techniques to detect co-located poisoning attacks.

ACKNOWLEDGEMENT

This research was partly supported by the Institute of Information and communications Technology Planning and Evaluation (IITP) grant funded by the Korea Government (MSIT), (No. 2021-0-00796, Research on Foundational Technologies for 6G Autonomous Security-by-Design to Guarantee Constant Quality of Security, 50%) and the MSIT(Ministry of Science, ICT), Korea, under the National Program for Excellence in SW), supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) in 2024 (2022-0-01068, 50%).

REFERENCES

[1] L. Bai, et al., "Physical layer authentication in wireless communication networks: A survey," in *Journal of Communications and Information Networks*, vol. 5, no. 3, pp. 237-264, Sept. 2020

[2] H. Liu, et al., "Authenticating Users Through Fine-Grained Channel Information," in *IEEE Transactions on Mobile Computing*, vol. 17, no. 2, pp. 251-264, 1 Feb. 2018