

PRAPD: A novel received signal strength–based approach for practical rogue access point detection

International Journal of Distributed
Sensor Networks
2018, Vol. 14(8)
© The Author(s) 2018
DOI: 10.1177/1550147718795838
journals.sagepub.com/home/dsn
 SAGE

Wenjia Wu¹ , Xiaolin Gu¹, Kai Dong¹, Xiaomin Shi² and Ming Yang¹

Abstract

Rogue access point attack is one of the most important security threats for wireless local networks and has attracted great attention from both academia and industry. Utilizing received signal strength information is an effective solution to detect rogue access points. However, the received signal strength information is formed by multi-dimensional received signal strength vectors that are collected by multiple sniffers, and these received signal strength vectors are inevitably lacking in some dimensions due to the limited wireless transmission range and link instability. This will result in high false alarm rate for rogue access point detection. To solve this issue, we propose a received signal strength–based practical rogue access point detection approach, considering missing received signal strength values in received signal strength vectors collected in practical environment. First, we present a preprocessing scheme for received signal strength vectors, eliminating missing values by means of data filling, filtering, and averaging. Then, we perform clustering analysis on the received signal strength vectors, where we design a distance measurement method that dynamically uses partial components in received signal strength vectors to minimize the distance deviation due to missing values. Finally, we conduct the experiments to evaluate the performance of the practical rogue access point detection. The results demonstrate that the practical rogue access point detection can significantly reduce the false alarm rate while ensuring a high detection rate.

Keywords

Wireless security, rogue access point detection, received signal strength, clustering analysis

Date received: 15 April 2018; accepted: 30 July 2018

Handling Editor: Xinwen Fu

Introduction

Nowadays, the IEEE 802.11 wireless local network (WLAN) is becoming an extremely popular wireless technology for various scenarios, such as campuses, homes, enterprise environments, and public spaces.^{1,2} With the wide deployment of WLANs, the issues of security and privacy have been increasingly emerging.³ Due to the openness of the wireless transmission medium, a variety of attacks can be launched easily. Among these attacks, rogue access point (AP) attacks have attracted more and more attention, and the rogue AP is defined as an illegal AP that is not deployed by the WLAN administrator.⁴

An adversary can set up a rogue AP with the same service set identifier (SSID) of the legitimate APs, and attract users to connect with it, so as to passively obtain user privacy information or actively perform

¹School of Computer Science & Engineering, Southeast University, Nanjing, China

²China Information Consulting & Designing Institute Co. Ltd, Nanjing, China

Corresponding author:

Wenjia Wu, School of Computer Science & Engineering, Southeast University, Nanjing 211189, China.
Email: wjwu@seu.edu.cn



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License

(<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

man-in-the-middle (MITM) attack.⁵ Moreover, in order to avoid being caught, adversaries usually capture the MAC address information of legitimate APs through passive monitoring and then modify the MAC address of the rogue AP by simply issuing an *ifconfig* command to masquerade as a legitimate AP. At present, the hardware cost of a rogue AP is low, and its software installation is also very convenient, which brings serious security threats to WLANs, especially public WLANs. The CCTV “3.15” evening party in 2015 reported the whole process of hackers using rogue APs to steal users’ photos, email accounts, passwords, and other private information. In recent years, some cases of personal property loss have occurred frequently, and bank accounts and passwords have been stolen because of misuse of rogue APs in public WLANs. Within the enterprise, the rogue AP attack has become an effective approach for hackers to invade internal networks, which poses a great threat to intranet security. Moreover, the performance of enterprise WLANs is being significantly impacted by the ever-increasing rogue APs due to the carrier sense interference and hidden terminal interference.⁶ Because the vast majority of existing user devices lack the recognition and authentication mechanisms for the APs, it is impossible to effectively distinguish legitimate APs from rogue APs. Therefore, it is necessary to study the rogue AP detection technology, which is the premise and basis for locating and troubleshooting rogue APs.

The existing solutions of rogue AP detection mainly include the following three aspects, that is, user-side detection, wired-side detection, and wireless-side detection. Through using user-side detection methods,^{4,7-9} a user device can determine whether its associated AP is a rogue AP. Although these solutions are lightweight and low cost, the user devices need to be customized or additional software should be installed. Thus, it is hard to be widely applied for user-side detection, while the detection from the perspective of network management has more advantages. Among them, the wired-side detection methods¹⁰⁻¹² assume that the adversary will use a specific wired backbone network to pass victims’ data to the Internet and achieve rogue AP detection by monitoring the backbone network and looking for the traffic that appears to come from rogue APs. However, this assumption limits the application scope because the adversary may use the same wired backbone network with legitimate APs, or use a different way, such as 4G long term evolution (LTE) and other mobile communication networks. Besides wired-side detection, another approach is to detect rogue APs in the wireless side,¹³⁻²² which directly senses the signals of rogue APs in the air and detects them according to the characteristics of the signals, such as clock skew and received signal strength (RSS). An RSS value is the signal strength of a received frame measured at the sniffer, and the signal strength of

the frame measured at multiple sniffers can constitute an RSS vector. Because RSS is correlated to the transmission power, the transmitter-receiver distance, and the environment, the RSS vectors of frames at one location will differ from that at another location. Therefore, the RSS vectors can be used to effectively detect the rogue APs that have different locations than legitimate APs, where a multi-dimensional RSS vector is aggregated from distributed RSS measurements.

However, for a practical environment, we have found that there exist a great many missing values in RSS vectors, due to the limited wireless transmission range and unreliable wireless links. If the sniffer is out of the frame’s transmission range, the corresponding RSS value is missing. If the frame is lost at the sniffer, the corresponding RSS value is also missing. The missing values in RSS vectors will affect the effectiveness of rogue AP detection, resulting in a higher false alarm rate.

To solve this issue, we propose a novel RSS-based approach for practical rogue access point detection (PRAPD), reducing the effect of missing RSS values on detection performance as much as possible. The main contributions of this article are as follows:

- (1) We present a preprocessing scheme to eliminate the missing values in the collected RSS vectors by means of data filling, filtering, and averaging.
- (2) We use k-medoid algorithm to perform clustering analysis on the RSS vectors, where we design a distance metric method that dynamically uses partial components in RSS vectors to minimize the distance deviation caused by the missing values.
- (3) We conduct the experiments in the practical environment and evaluate the performance of the proposed approach PRAPD. The results show that the PRAPD can effectively reduce the false alarm rate while ensuring a high detection rate.

The rest of the article is organized as follows: section “Related work” briefly reviews the existing work on rogue AP detection. Section “Rogue AP attack” describes the attack model we address in this article. In section “Design of PRAPD,” we present the rogue AP detection approach, mainly including data preprocessing and clustering analysis. In section “Evaluation,” the experiments and results are presented and analyzed. Finally, section “Conclusion” concludes this article.

Related work

The rogue AP attack is a serious security threat in WLANs and has attracted significant attention from

基
理

缺
失
值
原
因

多
维
度
处
理
空
值

聚
类
最
小
化
距
离
偏
差

不
能
大
范
围
使
用

不
能
大
范
围
使
用

不
能
大
范
围
使
用

不
能
大
范
围
使
用

both industry and academia.⁵ At present, the function of rogue AP detection has been integrated in some WLAN network management systems, such as AirWave.²³ In addition, there are some systems that deploy sniffer nodes to achieve rogue AP detection, such as AirMagnet.²⁴

Rogue AP detection has also caught the attention of researchers for many years. The traditional detection method can collect the SSID, MAC, and other parameters of the legitimate AP in advance to form a white list, then capture 802.11 frames, and detect rogue APs by MAC address filtering. Although this method is simple and efficient, an attacker can easily evade the detection by modifying the MAC address of the rogue AP so that it is the same with a legitimate AP.²⁵ Because the MAC address can be spoofed, researchers have begun to explore other effective features for rogue AP detection. In this article, we classify the existing approaches of rogue AP detection into three main categories: user-side detection, wired-side detection, and wireless-side detection.

User-side detection

Several approaches have been proposed to implement low-cost and lightweight rogue AP detection from the perspective of users.

Han et al.⁴ considered a category of rogue APs that are equipped with dual wireless interfaces (an interface is connected to a legitimate AP, while the other interface is pretended to be a legitimate AP to induce users) and designed a timing-based scheme that allows users to avoid connecting to rogue APs. The user-centric scheme employs the round trip time between the user and the DNS server to independently determine whether the associated AP is a rogue AP. For detecting the same category of rogue APs, Yang et al.⁷ analyzed the interpacket arrival time (IAT) as the detection feature, and the IAT is the time interval between two consecutive packets from the same device (the remote sever or the associated AP) to the user device. On this basis, the authors proposed detection algorithms that utilized the IAT as the detection feature, considered the influencing factors of RSS and network saturation, and employed sequential probability ratio test technology to achieve rogue AP detection.

Nakhila et al.⁸ proposed a comprehensive real-time user-side method to detect both types of rogue APs in parallel by creating two virtual wireless clients (VWCs). For a rogue AP connecting to a legitimate AP, a VWC monitored multiple channels in random order looking for specific data packets sent by a server on the Internet, and the rogue AP would be detected if duplicated data or no data were captured. For a rogue AP connecting to 4G LTE and other mobile communication networks, the second VWC would detect it when

the wireless network used two different gateways by switching from one AP to another in the middle of a secure connection. Gonzales et al.⁹ presented a context-leashing strategy for rogue AP detection, where users compared the current context with the previously learned context for the AP, and determined whether it was a rogue AP.

These approaches provide technical solutions to prevent rogue AP attack from the user's point of view and have the advantages of lightweight and low cost. However, users are required to customize their devices or install additional software, and it is difficult to be widely applied.

Wired-side detection

Wired-side detection is a category of rogue AP detection technologies from the perspective of network administrators. Because WLANs generally use a wired network as the backhaul network to connect to the Internet, network traffic can be captured at the gateway or at the mirror port of the switch, and then traffic analysis is performed to detect rogue APs.

Beyah et al.¹⁰ found that a wireless link in a network path of multiple links would cause a more random and temporally different spreading of packets. On this basis, the characteristic of inter-packet spacing was used to detect unwanted wireless traffic on the switch port and decides whether there exists a rogue AP. Wei et al.¹¹ demonstrated that the inter-arrival time of TCP ACK pairs could effectively differentiate wired and wireless connections and designed two online algorithms to detect rogue APs. Both algorithms used sequential hypothesis test technique and took the inter-ACK times as the input. Burns et al.¹² considered the relay-based rogue AP attack, that is, relaying the traffic through a legitimate AP, and set up a remote server to detect rogue APs by analyzing the user-server and server-user traceroute results. Although a rogue AP could hide the existence of the legitimate AP by tampering the user-server traceroute results, it could not prevent the server discovering the rogue AP from the server-user traceroute results.

It can be seen that this category of detection methods utilizes the characteristics of the traffic in the wired network to perform rogue AP detection and have the advantages of easy data collection, high detection efficiency, and independence of the signal range of rogue APs. However, these methods can only apply to rogue APs that connect to the Internet through a wired network that can be controlled by the administrator and cannot work if rogue APs use 4G LTE networks to connect to the Internet.

Wireless-side detection

Wireless-side detection is another category of rogue AP detection technologies from the perspective of network

三类
检测方式

特点

RTT

局限

administrators. It aims to capture wireless signals and extract some features to achieve rogue AP detection.

Jana and Kasera¹³ used the clock skew as AP's fingerprint to detect rogue APs, and the clock skew was estimated using the time synchronization function (TSF) timestamps in the 802.11 beacon/probe response frames. Lanze et al.¹⁴ considered the effect of temperature on clock skew and proposed a corresponding detection method. Jang et al.¹⁵ developed a rogue AP detection mechanism that used the feature of channel interference. Guo and Chiueh¹⁶ proposed a detection algorithm that leveraged the sequence number field in the link-layer header of IEEE 802.11 frame. In addition to these features, the RSS is the most commonly used location-dependent feature and has also attracted the attention of researchers.²⁶

Sheng et al.¹⁷ discovered that the RSS values followed a mixture of multiple Gaussian distributions due to antenna diversity and then proposed an approach based on Gaussian mixture models, building RSS profiles for rogue AP detection. Chen et al.¹⁸ used the spatial correlation of RSS and detected rogue APs by performing clustering analysis in RSS. Furthermore, Yang et al.¹⁹ considered that multiple rogue APs were existing in the network and proposed the corresponding solutions. Alotaibi and Elleithy²⁰ also utilized RSS and proposed an approach based on the random forests ensemble method to detect rogue APs. Zhou et al.²¹ proposed a crowdsensing-based approach to detect rogue APs without specialized hardware requirement. The authors designed a grid-based profiling method to build RSS profile with crowdsensing collections and presented a matching algorithm to detect abnormal samples based on the majority voting. Qu et al.²² considered a rogue AP that was set up in moving vehicles and proposed a detection algorithm based on RSS. The algorithm used RSS to estimate the distance between the rogue AP and the detector and compared it to the distance calculated by the fake GPS location of the rogue AP.

As discussed above, the spatial correlation of RSS can be used to effectively detect a rogue AP that should be at a different location from the legitimate AP, and several RSS-based approaches have been proposed. However, few of them considers the missing RSS values that will significantly affect the effectiveness of rogue AP detection. Hence, in this article, we propose a novel RSS-based approach for PRAPD, reducing the effect of missing RSS values on detection performance as much as possible.

Rogue AP attack

To demonstrate the efficacy of the proposed approach for rogue AP detection, we present the description of

the rogue AP attack and related assumptions as follows.

The adversary can use an off-the-shelf device to impersonate a legitimate AP, such as an OpenWrt wireless router or a Linux laptop running the hostapd software. In this article, we assume that the rogue AP attack can be implemented by an adversary with the capability to mimic the configurations of the legitimate WLAN. Specifically, the SSID, BSSID (an identifier that uniquely identifies an access point, and corresponds to the MAC address of the access point), and other configurations of a rogue AP should be exactly the same as the corresponding legitimate AP.

We assume that the adversary sets up a rogue AP after finding a legitimate AP as the target, and the legitimate AP and the rogue AP coexist in the WLAN during the process of the attack, as shown in Figure 1. The rogue AP can increase its signal strength to lure users to connect. We also assume that the rogue AP is deployed in a different location from the legitimate AP because an unknown device placed near the legitimate AP will easily attract administrators' attention. In addition, because wireless-side detection is done in this article, we do not care how the rogue AP connects to the Internet. The rogue AP can relay to legitimate APs or use 4G LTE networks.

Design of PRAPD

To detect a rogue AP coexisting with the mimicked legitimate AP, we design the approach PRAPD that senses and identifies the changes of RSS information when a rogue AP is set up. In this section, we discuss the overview of our approach and present details about data preprocess and clustering analysis.

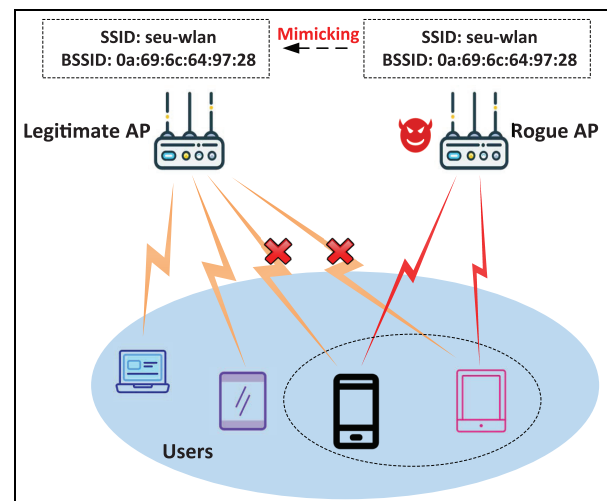


Figure 1. Attack model of rogue APs.

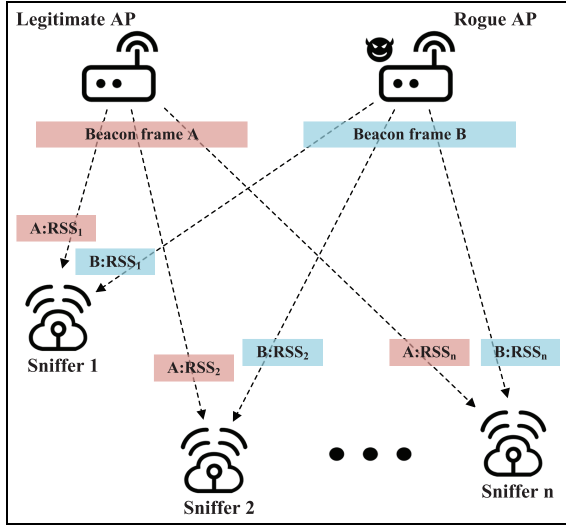


Figure 2. RSS collection by n sniffers.

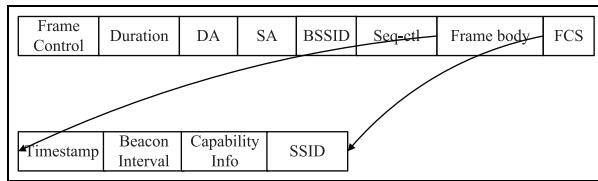


Figure 3. 802.11 beacon frame.²⁷

Overview

In PRAPD, an RSS value is the signal strength of an AP's beacon frame captured by a sniffer. The RSS value is closely related to the AP's physical location and is determined by the distance to the sniffer. We deploy n sniffers that, respectively, capture APs' beacon frames and collect the corresponding RSS values, as shown in Figure 2. Since the locations of these sniffers are distinctive in physical space, the RSS values for a beacon frame are also usually different.

Then, the RSS values are aggregated and processed by a central server. It can be seen from the structure of the beacon frame in Figure 3 that there is a timestamp field. The sending timestamp will be inserted into this field when the frame is ready to send, and we can use the timestamp to identify a beacon frame, and aggregate its RSS values collected by multiple sniffers. For a beacon frame, the aggregated RSS vector is denoted as $S = \{s_1, s_2, \dots, s_n\}$, where the component s_i is the RSS value collected by sniffer i . Thus, each RSS vector can be mapped to a point in an n -dimensional signal space.

For each BSSID, the RSS vectors will be close to each other in signal space when no rogue AP exists because the corresponding frames are sent from the same location. However, when there is a rogue AP using this BSSID, there are two APs at different

physical locations claiming the BSSID. As a result, the RSS vectors from the legitimate AP will be mixed with that from the rogue AP, and these RSS vectors from the two different locations in the physical space should form two clusters in signal space. Hence, we can conduct clustering analysis on the RSS vectors from each BSSID in order to detect rogue APs.

As shown in Figure 4, our approach consists of two phases. During the offline profiling phase, we collect RSS vectors multiple times for each legitimate AP and obtain the distance between two centroids in signal space after clustering analysis. Then, we use the distribution of the distance information to determine the threshold α . For each legitimate AP, its threshold is maintained in the server. During the online detection phase, we collect RSS vectors for each BSSID and perform clustering analysis to determine whether a rogue AP exists. In particular, we present the details of data preprocess and clustering analysis.

Data preprocess

In the practical environment, there are many missing values in RSS vectors, as shown in Table 1. This is because that the wireless transmission range is limited or wireless links are unreliable. Specifically, the sniffer cannot receive an AP's frame if it is out of the transmission range of the AP, and there is another case that the sniffer may lose a frame due to unreliable wireless links. In this article, we deal with these missing values by data filling, filtering, and averaging.

We first use a constant (-100) that is smaller than any of the measured values and fill the RSS vectors to eliminate these missing values.²⁸ The data filling can well address the issue of missing values caused by AP's limited transmission range, such as the missing values of the component s_4 in Table 1. However, this is not suitable for the missing values of components s_1 and s_2 that are caused by frame lost. As shown in Table 2, the distance between RSS vectors is obviously exaggerated after data filling, but in fact these signals are transmitted from the same location.

In order to solve this issue, we perform data filtering and data averaging in sequence:

- **Data filtering.** For each component, we think that the missing values are caused by frame lost, if there are only a small amount of missing values in the component, that is, the proportion of vectors missing in the component is less than β ($\beta = 0.15$). We can remove the corresponding vectors in the sample.
- **Data averaging.** After data filtering, we average the RSS values in each component and replace the constant value (-100) with the average.

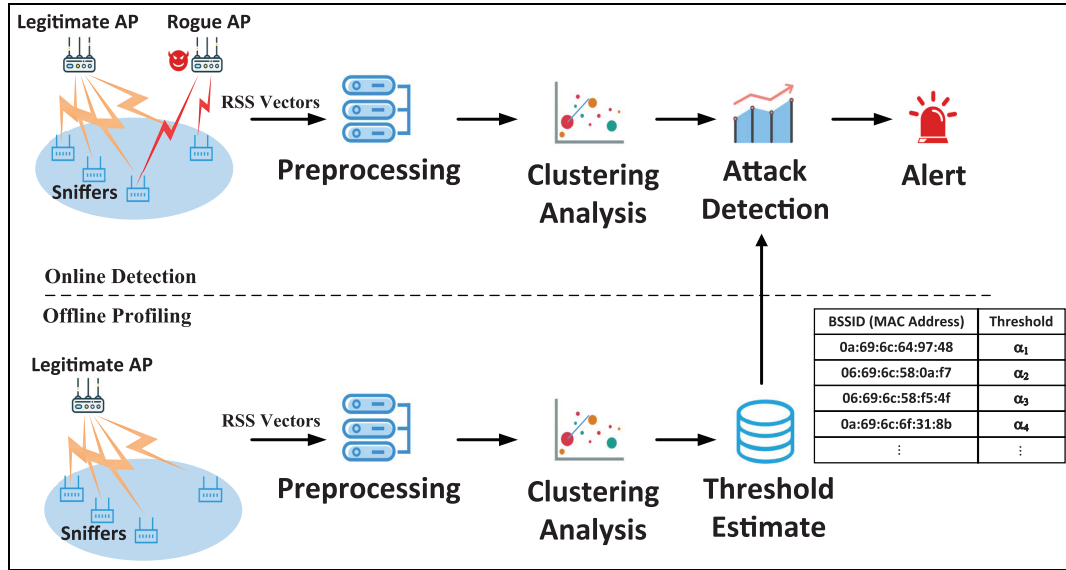


Figure 4. Overview of the proposed approach PRAPD.

Table 1. Fragment of RSS vectors from a legitimate AP (dBm).

ID	Timestamp	s_1	s_2	s_3	s_4	s_5	s_6
1	00001082959ce180	-40	-61	-67	-	-61	-53
2	0000108295b45180	-	-	-69	-	-63	-47
3	0000108295b5e180	-38	-63	-69	-	-65	-49
4	0000108295cee180	-	-	-69	-	-61	-49
5	0000108295d39180	-40	-63	-69	-	-61	-49
6	0000108295e7e180	-42	-61	-69	-83	-61	-49
7	0000108295efb18d	-	-	-	-	-59	-49
8	0000108295f91189	-40	-	-69	-	-61	-49
9	00001082962fc180	-	-	-69	-	-61	-47
10	00001082964a5180	-	-	-69	-	-	-
11	000010829677a180	-	-	-	-	-61	-51
12	000010829680b180	-40	-	-69	-83	-61	-49

AP: access point; RSS: received signal strength.

Table 2. RSS vectors after data filling (dBm).

ID	s_1	s_2	s_3	s_4	s_5	s_6
1	-40	-61	-67	-100	-61	-53
2	-100	-100	-69	-100	-63	-47
3	-38	-63	-69	-100	-65	-49
4	-100	-100	-69	-100	-61	-49
5	-40	-63	-69	-100	-61	-49

RSS: received signal strength.

Clustering analysis

In the section, we use the k -medoid algorithm ($k = 2$) to perform clustering analysis on RSS vectors because this clustering algorithm is robust in the existence of noise and outliers.²⁹

Because we only consider the RSS vectors from one or two APs (offline profiling or online detection) in clustering analysis, there exist a large number of RSS vectors similar to each other. Since the measured RSS value is a discrete integer, we find that multiple RSS vectors will correspond to a point in the signal

space. In order to improve the efficiency of clustering analysis, we can scan the RSS vectors, obtain a set of RSS vectors that differ from each other, and count the number of occurrences for each RSS vector. The set of RSS vectors is denoted by $\mathbb{S} = \{S_1, S_2, S_3, \dots, S_m\}$, and the number of S_j 's occurrences is denoted by w_j .

It is of great importance for clustering analysis to define the distance between RSS vectors. The traditional Euclidean distance considers all the components of the RSS vectors and can be easily exaggerated due to the missing values. Therefore, we define a new distance function that dynamically considers partial components of the RSS vectors, that is

$$D_{p,q}^B = \sqrt{\frac{\sum_{b \in BL} (S_p(s_b) - S_q(s_b))^2}{B}} \quad (1)$$

where B ($B \leq n$) is the parameter as the number of components considered in this distance function, BL is the list of top B components sorted in ascending order of $|S_p(s_i) - S_q(s_i)|$, and $S_p(s_i)$ is the component s_i of the RSS vector S_p .

As described in Algorithm 1, we first initialize two medoids S_u and S_v by randomly selecting two RSS vectors from \mathbb{S} . Then, we iteratively partition the set of RSS vectors \mathbb{S} into two parts (C_1 and C_2) and update the two medoids S_u and S_v . The algorithm terminates when the number of iterations reaches T or the two medoids are the same as the previous iteration and return the final two medoids. For the set partition, an RSS vector S_j joins cluster C_1 or C_2 , according to the distance $D_{j,u}^B$ and $D_{j,v}^B$. For medoid selection, we consider that each RSS vector in \mathbb{S} corresponds multiple vectors in raw RSS vectors and designs a new selection metric as follows

Algorithm 1: k -medoid algorithm ($k = 2$)

Input: $\mathbb{S} = \{S_1, S_2, S_3, \dots, S_m\}$, B , T

Output: Two medoids S_u and S_v

```

1 Initialize  $S_u$  and  $S_v$  from  $\mathbb{S}$  randomly
2  $t \leftarrow 1$ 
3 while true do
4    $C_1 \leftarrow \emptyset, C_2 \leftarrow \emptyset$ 
5   for each  $S_j \in \mathbb{S}$  do
6     Calculate  $D_{j,u}^B$  and  $D_{j,v}^B$ 
7     if  $D_{j,u}^B \leq D_{j,v}^B$  then
8        $C_1 \leftarrow C_1 \cup \{S_j\}$ 
9     else
10       $C_2 \leftarrow C_2 \cup \{S_j\}$ 
11   Select new medoid  $S_{nu}$  from cluster  $C_1$ 
12   Select new medoid  $S_{nv}$  from cluster  $C_2$ 
13   if  $t = T$  or  $(S_u = S_{nu}$  and  $S_v = S_{nv})$  then
14     return  $S_u, S_v$ 
15    $S_u \leftarrow S_{nu}, S_v \leftarrow S_{nv}$ 
16    $t \leftarrow t + 1$ 
```

$$S_{nu} = \arg \min_{S_p \in C_1} \sum_{S_q \in C_1} w_q D_{p,q}^B \quad (2)$$

where S_{nu} is a new medoid from cluster C_1 .

After clustering analysis, we can calculate the distance between the two medoids. In the offline profile phase, for each BSSID, we perform clustering analysis multiple times on different groups of RSS vectors that are collected from the network scenario with no rogue AP and obtain the distribution of medoid distance. We can use the distribution information to determine the distance threshold α .¹⁸ In the online detection phase, for each BSSID, we perform clustering analysis on RSS vectors that are collected from the current network scenario and calculate the medoid distance D_c . Our condition for declaring the existence of a rogue AP using this BSSID is $D_c > \alpha$.

Evaluation

We conduct experiments to evaluate the performance of our approach PRAPD. In this section, we describe our evaluation methodology and results, including experimental setup, performance metric, impacts of medoid distance threshold, varying number of components used, and varying locations of the rogue AP.

Experimental setup

First, we set up a 15 m \times 7 m environment at Computer Building of Jiulonghu Campus of Southeast University, as shown in Figure 5.

In our experiments, the implementation of the proposed approach PRAPD is described as follows:

- *Legitimate AP and rogue AP.* Two NETGEAR WNDR3800 routers running OpenWrt are used as the legitimate AP and the rogue AP, respectively. We modify the BSSID of the rogue AP to be the same of the legitimate AP.
- *Sniffers.* We deploy six sniffers, that is, Raspberry Pi equipped with wireless adapter Ralink RT5370. We configure the wireless interface to be monitor mode and use the libpcap library to capture 802.11 beacon frames. Sniffer deployment is presented in Figure 5.
- *Central server.* We use an HP PC as a central server that aggregates the collected RSS values and perform clustering analysis to detect rogue APs.

We collect 300 groups of RSS vectors from the scenarios without no rogue AP, and each group contains 1000 RSS vectors. For each rogue AP location, we, respectively, collect 150 groups of RSS vectors from the scenarios with the rogue AP.

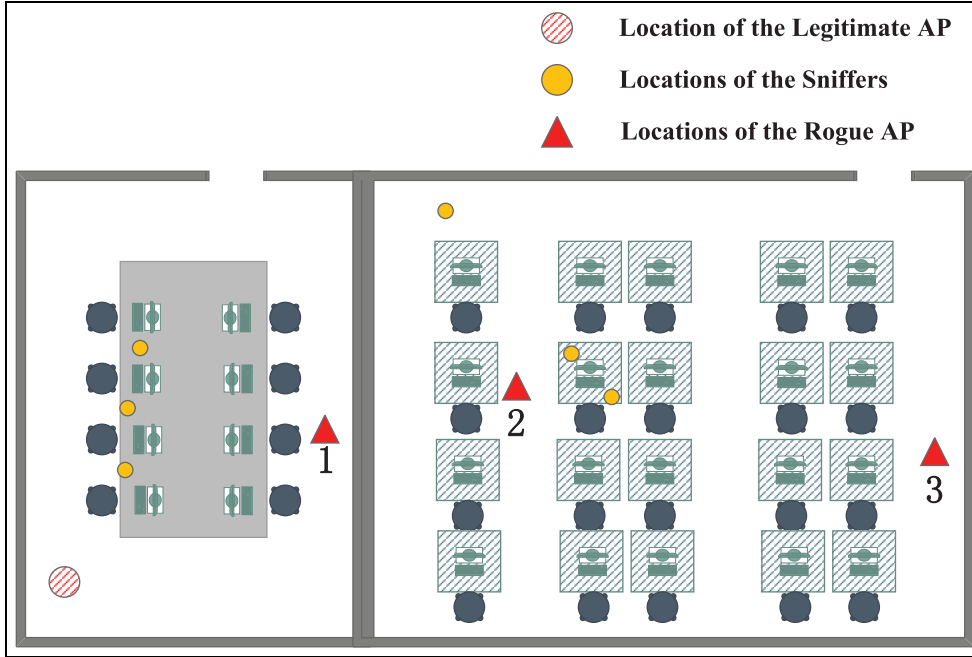


Figure 5. Experimental environment.

Performance metric

In this work, we use the following two metrics to evaluate the performance of our proposed approach:

- *Detection rate.* For the RSS samples (we use a group of RSS vectors as a sample) collected from the scenarios with a rogue AP, the detection rate is the proportion of the samples through which our approach detects the existence of the rogue AP correctly.
- *False alarm rate.* For the RSS samples collected from the scenarios with no rogue AP, the false alarm rate is the proportion of the samples through which our approach incorrectly determines the existence of a rogue AP.

Impacts of medoid distance threshold

Here, we evaluate the performance of the proposed approach affected by threshold estimation. We set the number of components used (B) to be 5. The rogue AP is deployed at location 2 in Figure 5.

We use 150 RSS samples (with no rogue AP) to estimate the medoid distance threshold. According to the distribution of the medoid distance, we obtain a quantile as the threshold that corresponds to the probability λ ($\lambda = 0.9$), and the initial threshold is 2.0 dBm. Due to the unreliability of wireless signals, we can appropriately relax this threshold and consider the threshold to be 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0, respectively. Under these threshold, we test our proposed approach

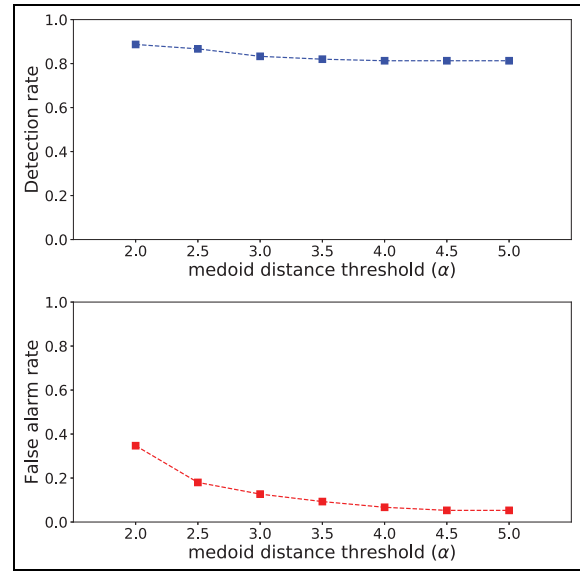


Figure 6. Impacts of medoid distance threshold.

through other 150 RSS samples (with no rogue AP) and the 150 RSS samples (with the rogue AP). The results are presented in Figure 6.

It can be seen from Figure 6 that it is able to find a trade-off between the detection rate and false alarm rate through adjusting the medoid distance threshold.

Varying number of components used

Next, we evaluate the performance of the proposed approach varying number of components used (B). We

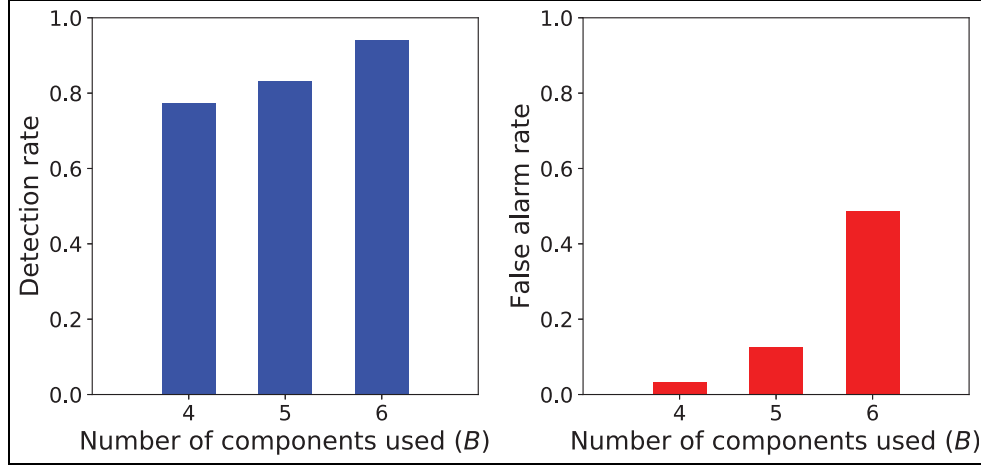


Figure 7. Varying number of components used (B).

set the medoid distance threshold α to be 3.0 dBm. The rogue AP is deployed at location 2 in Figure 5. The parameter B is set to be 4, 5, and 6. We use the 150 RSS samples (with no rogue AP) and the 150 RSS samples (with the rogue AP) to evaluate our proposed approach.

Figure 7 shows the results. From this figure, we can see that the false alarm rate is high when all the components are used ($B = 6$) because the components with missing values are involved in distance measurement and the medoid distance is exaggerated. We also find that the false alarm rate is effectively decreased without significantly reducing the detection rate when partial components are used ($B = 5$ or $B = 4$).

Varying locations of the rogue AP

Finally, we evaluate the performance of the proposed approach varying locations of the rogue AP. We set the number of components used B to be 5 and set the medoid distance threshold α to be 3.0 dBm. We choose three locations for rogue AP deployment, that is, location 1, location 2, and location 3 in Figure 5, where location 1 is 4.7 m away from the legitimate AP, location 2 is 8.2 m away, and location 3 is 15.4 m away. For each rogue AP location, we use the corresponding 150 RSS samples to validate our proposed approach.

Figure 8 shows the results. We can see from this figure that our proposed approach is effective when the rogue AP is deployed in any of these three locations. Moreover, the results show that the detection rate is slowly decreasing when the rogue AP is near to the legitimate AP and stays above 0.80 even if the rogue AP is only 4.7 m away from the legitimate AP.

Conclusion

We investigated the RSS-based rogue AP detection approach in practical environments, while considering

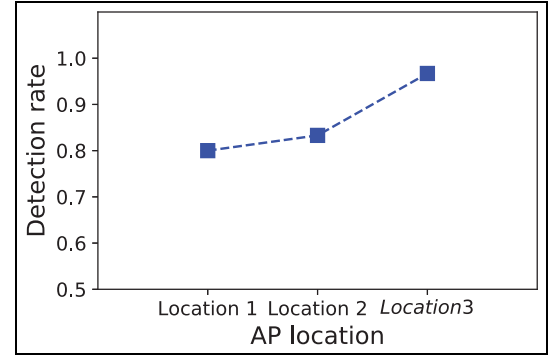


Figure 8. Varying locations of the rogue AP.

missing RSS values in the collected RSS vectors. First, we presented a data preprocessing scheme to eliminate the missing values in the collected RSS vectors by means of data filling, filtering, and averaging. Then, we utilized the k -medoid algorithm to perform clustering analysis on the RSS vectors, where we designed a distance metric method that dynamically used partial components in RSS vectors to minimize the distance deviation caused by the missing values. Finally, we conducted the experiments in the practical environment and evaluated the performance of the proposed approach PRAPD. The results demonstrated that the PRAPD can effectively reduce the false alarm rate while ensuring a high detection rate.

In the future, we plan to extend the proposed approach for large-scale scenarios. Specifically, we will investigate how to properly deploy sniffers and explore how to deal with high-dimensional RSS data.


Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by National Key R&D Program of China (Grant No. 2017YFB1003000), National Natural Science Foundation of China (Grant Nos 61632008, 61572130, 61502100, 61602111, 61532013, and 61320106007), Jiangsu Provincial Natural Science Foundation of China (Grant Nos BK20150637 and BK20150628), Jiangsu Provincial Scientific and Technological Achievements Transfer Fund (Grant No. BA2016052), Jiangsu Provincial Key Laboratory of Network and Information Security (Grant No. BM2003201), and Key Laboratory of Computer Network and Information Integration of Ministry of Education of China (Grant No. 93K-9).

ORCID iD

Wenjia Wu  <https://orcid.org/0000-0002-4437-0850>

References

1. Sun W, Lee O, Shin Y, et al. Wi-Fi could be much more. *IEEE Commun Mag* 2014; 52(11): 22–29.
2. Bellalta B, Bononi L, Bruno R, et al. Next generation IEEE 802.11 wireless local area networks: current status, future directions and open challenges. *Comput Commun* 2016; 75: 1–25.
3. Zou Y, Zhu J, Wang X, et al. A survey on wireless security: technical challenges, recent advances, and future trends. *Proc IEEE* 2016; 104(9): 1727–1765.
4. Han H, Sheng B, Tan CC, et al. A timing-based scheme for rogue AP detection. *IEEE T Parall Distr* 2011; 22(11): 1912–1925.
5. Beyah RA and Venkataraman A. Rogue-access-point detection: challenges, solutions, and future directions. *IEEE Secur Priv* 2011; 9(5): 56–61.
6. Sui K, Zhao Y, Pei D, et al. How bad are the rogues' impact on enterprise 802.11 network performance? In: *Proceedings of the conference on computer communications*, Kowloon, Hong Kong, 26 April–1 May 2015, pp.361–369. New York: IEEE.
7. Yang C, Song Y and Gu G. Active user-side evil twin access point detection using statistical techniques. *IEEE T Inf Foren Sec* 2012; 7(5): 1638–1651.
8. Nakhila O, Amjad MF, Dondyk E, et al. Gateway independent user-side wi-fi evil twin attack detection using virtual wireless clients. *Comput Secur* 2018; 74: 41–54.
9. Gonzales H, Bauer K, Lindqvist J, et al. Practical defenses for evil twin attacks in 802.11. In: *Proceedings of the global communications conference*, Miami, FL, 6–10 December 2010, pp.1–6. New York: IEEE.
10. Beyah RA, Kangude S, Yu G, et al. Rogue access point detection using temporal traffic characteristics. In: *Proceedings of the 47th annual global telecommunications conference*, Dallas, TX, 29 November–3 December 2004, vol. 4, pp.2271–2275. New York: IEEE.
11. Wei W, Suh K, Wang B, et al. Passive online rogue access point detection using sequential hypothesis testing with TCP ACK-pairs. In: *Proceedings of the internet measurement conference*, San Diego, CA, 24–26 October 2007, pp.365–378. New York: ACM.
12. Burns A, Wu L, Du X, et al. A novel traceroute-based detection scheme for Wi-Fi evil twin attacks. In: *Proceedings of the global communications conference*, Singapore, 4–8 December 2017, pp.1–6. New York: IEEE.
13. Jana S and Kasera SK. On fast and accurate detection of unauthorized wireless access points using clock skews. In: *Proceedings of the 14th annual international conference on mobile computing and networking*, San Francisco, CA, 14–19 September 2008, pp.104–115. New York: IEEE.
14. Lanze F, Panchenko A, Braatz B, et al. Letting the puss in boots sweat: detecting fake access points using dependency of clock skews on temperature. In: *Proceedings of the 9th symposium on information, computer and communications security*, Kyoto, Japan, 4–6 June 2014, pp.3–14. New York: ACM.
15. Jang R, Kang J, Mohaisen A, et al. Rogue access point detector using characteristics of channel overlapping in 802.11n. In: *Proceedings of the 37th international conference on distributed computing systems*, Atlanta, GA, 5–8 June 2017, pp.2515–2520. New York: IEEE.
16. Guo F and Chiueh T. Sequence number-based MAC address spoof detection. In: *Proceedings of the 8th international workshop on recent advances in intrusion detection*, Seattle, WA, 7–9 September 2005, pp.309–329. New York: ACM.
17. Sheng Y, Tan K, Chen G, et al. Detecting 802.11 MAC layer spoofing using received signal strength. In: *Proceedings of the 27th conference on computer communications*, Phoenix, AZ, 15–17 April 2008, pp.1768–1776. New York: IEEE.
18. Chen Y, Trappe W and Martin RP. Detecting and localizing wireless spoofing attacks. In: *Proceedings of the 4th annual communications society conference on sensor, mesh and ad hoc communications and networks*, San Diego, CA, 18–21 June 2007. New York: IEEE.
19. Yang J, Chen Y, Trappe W, et al. Detection and localization of multiple spoofing attackers in wireless networks. *IEEE T Parall Distr* 2013; 24(1): 44–58.
20. Alotaibi B and Elleithy K. A new MAC address spoofing detection technique based on random forests. *Sensors* 2016; 16(3): 281.
21. Zhou T, Cai Z, Xiao B, et al. Detecting rogue AP with the crowd wisdom. In: *Proceedings of the 37th international conference on distributed computing systems*, Atlanta, GA, 5–8 June 2017, pp.2327–2332. New York: IEEE.
22. Qu H, Guo L, Zhang W, et al. Rogue access point detection in vehicular environments. In: *Proceedings of the international conference on wireless algorithms, systems, and applications*, Qufu, Shandong, China, 12–15 August 2015, pp.446–456. Cham: Springer.
23. AirWave. Aruba AirWave network management, <http://www.arubanetworks.com/products/networking/management/airwave/>
24. AirMagnet. AirMagnet enterprise, <http://enterprise-cn.netscout.com/enterprise-network/wireless-network/AirMagnet-Enterprise>
25. Adya A, Bahl P, Chandra R, et al. Architecture and techniques for diagnosing faults in IEEE 802.11 infrastructure networks. In: *Proceedings of the 10th annual international*

- conference on mobile computing and networking*, Philadelphia, PA, 26 September–1 October 2004. New York: ACM.
26. Xu Q, Zheng R, Saad W, et al. Device fingerprinting in wireless networks: challenges and opportunities. *IEEE Commun Surv Tut* 2016; 18(1): 94–104.
 27. Gast MS. *802.11 wireless networks: the definitive guide*. 2nd ed. Sebastopol, CA: O'Reilly Media, Inc, 2005.
 28. Zuo J, Liu S, Xia H, et al. Multi-phase fingerprint map based on interpolation for indoor localization using iBeacons. *IEEE Sens J* 2018; 18(8): 3351–3359.
 29. Kaufman L and Rousseeuw PJ. *Finding groups in data: an introduction to cluster analysis* (Wiley series in probability and statistics), vol. 344. Hoboken, NY: John Wiley & Sons, 2009.