Wang, Pu; Yan, Zheng; Zeng, Kai

BCAuth: Physical Layer Enhanced Authentication and Attack Tracing for Backscatter Communications

Published in:
IEEE Transactions on Information Forensics and Security

DOI:
10.1109/TIFS.2022.3195407

Published: 01/01/2022

Document Version
Publisher's PDF, also known as Version of record

BCAuth: Physical Layer Enhanced Authentication and Attack Tracing for Backscatter Communications

Pu Wang, *Student Member, IEEE*, Zheng Yan, *Senior Member, IEEE*, and Kai Zeng, *Member, IEEE*

***Abstract*— Backscatter communication (BC) enables ultra-low-power communications and allows devices to harvest energy simultaneously. But its practical deployment faces severe security threats caused by its nature of openness and broadcast. Authenticating backscatter devices (BDs) is treated as the first line of defense. However, complex cryptographic approaches are not desirable due to the limited computation capability of BDs. Existing physical layer authentication schemes cannot effectively support BD mobility, multiple attacker identification and attacker location tracing in an integrated way. To tackle these problems, this paper proposes BCAuth, a multi-stage authentication and attack tracing scheme based on the physical spatial information of BDs to realize enhanced BD authentication security for both static and mobile BDs. After initial authentication based on BD identity with its position information registration, preemptive authentication and re-authentication are performed according to spatial correlation of backscattered signal source locations associated with the BD. By exploiting clustering-based analysis on spacial information, BCAuth is capable of determining the number of attackers and localizing their positions. In addition, we propose a reciprocal channel-based method for BD re-authentication with better authentication performance than the clustering-based method for mobile BDs when the BDs is able to measure received signal strength (RSS), which also enables mutual authentication. We theoretically analyze BCAuth security and conduct extensive numerical simulations with various settings to show its desirable performance.**

***Index Terms*— Backscatter communication, physical layer security, device authentication, attack detection, positioning.**

I. Introduction

BACKSCATTER communication (BC), which enables device energy harvesting and ultra-low-power communications, has emerged as a cutting-edge wireless technology to enable sustainable Internet of Things (IoT) applications [1]. Compared with traditional wireless devices with power-hungry radio frequency (RF) functionalities, backscatter devices (BDs) have no active RF components, but can conduct ultra-low-power communications by backscattering RF signals from an access point (AP) [2]. Another particular feature of BDs is they can harvest energy from RF signals to power their circuits even without battery energy [3]–[6]. As a result, BDs can be made by low-cost hardware with extremely low power consumption, which facilitates their large-scale deployment in specific IoT applications, such as implantable medical devices and industrial sensor systems in a restrictive environment [1], [7], [8]. But due to the openness and broadcast nature of backscattering, BC systems face various security threats and vulnerabilities, i.e., BD identity impersonation and wireless spoofing attacks [9]–[11].

Device authentication is an essential manner to ensure fundamental BC security for preventing BD identity impersonation and wireless spoofing attacks. It aims to validate whether a BD is indeed a legitimate device. Conventional authentication schemes are mainly based on cryptographic mechanisms by pre-assigning secret keys and identities for device identification and authorization [12], [13]. These schemes work well for devices with powerful capabilities, e.g., smartphones. But due to the finite energy and limited computation capability of BDs, it is hard for BDs to employ cryptographic algorithms [12]. For example, only simple hash functions are used to protect the identity of tags in radio-frequency identification (RFID) systems [13], [14] due to their limited resources. But applying simple cryptographic authentication protocols makes secret identities easily sniffed by an adversary which can masquerade as a legitimate tag to intrude a system by raising identity-based attacks (IBA) [15]. Moreover, as the scale of BD deployment increases, it becomes quite difficult to effectively distribute and manage keys to realize cryptographic authentication. Physical layer authentication (PLA) can achieve effective and light-weight device identification and authentication without any encryption operations at BDs, which has been validated by many existing works [16]–[18]. Specific features of RF signals or device hardware are extracted by a verifier for authentication without requiring extra operation from devices [17], [19], [20]. It is suitable for resource-limited BDs due to low computation cost and trivial energy consumption.

However, the literature still lacks an effective physical layer BD authentication scheme for securing BC systems. For example, Qiu *et al.* [21] proposed a deep learning-based authentication approach to learn and track the variations of

Manuscript received 18 November 2021; revised 22 March 2022 and 18 June 2022; accepted 11 July 2022. Date of publication 1 August 2022; date of current version 11 August 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62072351; in part by the Academy of Finland under Grant 345072, Grant 308087, Grant 335262, and Grant 350464; in part by the Open Research Project of Zhejiang Laboratory under Grant 2021PD0AB01; in part by the 111 Project under Grant B16037; and in part by the U.S. National Science of Foundation through Networking Technology and Systems (NeTS) Program under Project 2131507. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Hossein Pishro-Nik. *(Corresponding author: Zheng Yan.)*

Pu Wang is with the State Key Laboratory of ISN, School of Cyber Engineering, Xidian University, Xi'an, Shaanxi 710026, China (e-mail: wangpu@stu.xidian.edu.cn).

Zheng Yan is with the State Key Laboratory of ISN, School of Cyber Engineering, Xidian University, Xi'an 710026, China, and also with the Department of Communications and Networking, Aalto University, 02150 Espoo, Finland (e-mail: zheng.yan@aalto.fi).

Kai Zeng is with the Department of Electrical and Computer Engineering, the Department of Cyber Security Engineering, and the Department of Computer Science, George Mason University, Fairfax, VA 22030 USA (e-mail: kzeng2@gmu.edu).

Digital Object Identifier 10.1109/TIFS.2022.3195407

channel characteristics in order to enhance authentication security, but this approach requires data for training a proper learning model, which is not context-adaptive. Whether it can be applied into mobile BD authentication and detect attackers requests additional investigation. Geneprint [22] leverages the internal similarity of backscattered signals to extract BD hardware features for authentication in RFID systems. Mehmood *et al.* [23] used the non-reciprocity of the residual channel between a reader and an RFID tag as the fingerprint of the tag to defend fault tag attacks. However, the existing works about RFID tag authentication including the above [17], [22], [24], [25] assume a sophisticated signal analyzer applied by an RFID reader to extract specific features. Such an analyzer brings high hardware costs, making it infeasible for general BC systems. Besides, such features as applied are sensitive and hard to be stably measured in dynamic environments. These limitations make existing approaches infeasible to be applied to mobile BC systems, where BDs could be static or mobile. More seriously, the hardware/software-based features of backscattered signals are vulnerable to signal counterfeit, signal relay and replay attacks, since attackers can learn constant signal features or record RF signals to mimic a legitimate BD [26]. But few existing schemes offer attack tracing and attacker number identification. All above related work cannot support mutual authentication between a BD and an AP. Therefore, it is highly expected to develop a novel physical layer BD authentication scheme that can effectively enhance authentication security by supporting BD mobility, tracing the location of attackers and identifying their number. Mutual authentication is also preferred at the same time if BD has sufficient capability to identity a fake AP.

But how to employ suitable physical layer features in BC systems to offer such effective authentication as expected above is not a trivial problem. We are facing a number of challenges. First, it is hard to perfectly estimate or measure the features or fingerprints of physical channel, considering environmental interference. Thus, imperfect channel estimation greatly impacts authentication accuracy. Second, it is hard to authenticate mobile BDs. Prior arts about physical layer authentication mainly focus on static BD authentication [20], [21], [23], [27]. Seldom, they consider BD mobility, especially when a BD could move across the coverage of multiple APs. Third, it is not an easy job to locate the concrete positions of the sources of attackers and identity their concrete number. Existing schemes [22] based on channel features or fingerprints cannot realize this goal since they do not collect any information about locations and analyze it. In particular, counting the number of potential attackers need special efforts due to the mixture of signals from legitimated devices and attackers.

To tackle the above challenges, in this paper, we propose BCAuth, a multi-stage authentication and attack tracing scheme based on the physical spatial information of BDs. We set up a hierarchical architecture of the BC system where a server controls APs to simultaneously transmit RF signals to power BDs and receive information from them. In such a hierarchical BC system, the server and APs cooperatively authenticate a BD, when the BD responds to the APs. We employ preemptive authentication at the AP and re-authentication at the server to offer enhanced authentication security for both static and mobile BDs. The AP first authenticates BD based on BD identity with BD position registration in a prior stage, different from some existing schemes [19], [22] that only apply physical layer features. Without any need to extract specific features, e.g., average baseband power [22], [24] that requests specific hardware, positioning information can be calculated from the received signal strength (RSS) and the angle of arrival (AoA) of backscattered signals, which are easily estimated in BC systems with high accuracy based on the advance of prior-arts [28]. In addition, we randomize the power of carrier signals from AP to randomize the RSS values of backscattered signals, so as to make RSS values unpredictable and unforgeable. Except for identity impersonation or spoofing attacks, our scheme can prevent signal counterfeit and replay attacks, as well as relay attacks in the physical layer, due to the randomness of RSS values.

In the following stages of authentication, the server exploits a clustering-based algorithm to analyze the spatial correlation of estimated locations associated with the BD to perform re-authentication. Accurate authentication can be ensured based on correlated locations provided by multiple APs and meanwhile a mobile BD can be authenticated based on BD trajectory estimation and matching by using machine learning, concretely location clustering [28]. In particular, BCAuth can also conduct accurate attack detection through spacial information analysis by clustering multiple locations provided by the APs after the preemptive authentication. Besides, in case that the BD has the capability of recording RSS values of downlink signals from the AP, a reciprocal channel-based method can be proposed to perform BD re-authentication with the returned RSS records from the BD at the server. This method can improve the performance of mobile BD authentication by comparing the RSS records reported by BD with the ones saved by AP. Mutual authentication can also be realized if the BD is capable of measuring RSS and calculating the correlation of its recorded RSS values of AP with the ones provided by AP [29], [30]. In case authentication failure, BCAuth can further determine the number of attackers through clustering-based analysis on the location information related to the BD and then locate potential attackers, as well as judge the number of attackers.

Specifically, the main contributions of this paper are summarized as below:

- We propose BCAuth, a multi-stage authentication and attack tracing scheme in BC systems based on physical spatial information of BDs. BCAuth is the first work to provide enhanced authentication security for both static and mobile BDs and is capable of determining the number of attackers and tracing their locations. Except for identity impersonation attacks, BCAuth can also detect signal counterfeit attacks, replay attacks, and relay attacks.
- We propose a re-authentication method by exploiting spatial correlation of estimated locations associated with the BD. We also design another reciprocal channel-based method for re-authentication in case that the BD is capable of measuring the RSS of downlink signals to

TABLE I
COMPARISON BETWEEN EXISTING WORKS AND BCAUTH

References	BD authentication	Mobility support	Attacker location tracing	Attacker number identification	Mutual authentication	ML Model training avoidance	Signal analyzer independence	Resilience to imperfect channel state estimation	Detected attacks
[38]	○	○	○	○	-	●	○	○	ISA
[39]	○	○	○	○	-	●	○	○	ISA, RpA
[21]	●	○	○	○	○	○	-	○	ISA
[37]	●	○	○	○	○	○	○	○	ISA
[22]	●	○	○	○	○	-	○	○	ISA
[24]	●	●	○	○	○	-	○	○	ISA
[25]	●	○	○	○	○	-	○	○	ISA
[17]	●	○	○	○	○	-	○	○	ISA
[23]	●	○	○	○	○	-	○	○	ISA
[27]	●	●	○	○	○	○	○	○	ISA
[20]	●	○	○	○	○	●	●	◒	ISA
BCAuth	●	●	●	●	◒	●	●	◒	ISA, SCA, RpA, RlA

○: not supported; ◒: conditionally supported ; ●: supported; -: not related.
ISA: identity impersonation/spoofing attacks; SCA: signal counterfeit attack; RpA: replay attack; RlA: relay attack.

improve authentication performance for mobile BDs. With the capability of RSS measurement at BD, mutual authentication between BD and AP can also be achieved.
- We conduct both theoretical analysis and extensive numerical simulations under different parameters to show the desirable security and performance of BCAuth for BD authentication and attack tracing.

The rest of the paper is organized as follows. We review the literature and qualitatively compare BCAuth with a number of related physical layer authentication schemes in Section II. Section III introduces the system and security models of BCAuth and overviews the overall procedure of BCAuth. In Section IV, we present the detailed design of BCAuth, including authentication initialization, preemptive authentication, re-authentication and attack tracing. We conduct security analysis in Section V, followed by performance evaluation results based on numerical simulations in Section VI. Finally, we conclude the paper and indicate our future work in the last section.

II. RELATED WORKS

In this section, we review the existing works of device authentication based on cryptography and physical layer features, respectively. We also compare existing related physical layer authentication schemes with BCAuth in Table I.

A. Cryptographic Authentication

Existing authentication schemes applied in BC systems in practice mainly focus on RFID systems, some of which highly depend on cryptographic technologies [31], [32]. For instance, in the RFID systems that employ EPCgloable standards, each RFID tag stores a unique and random number allocated by a back-end server as an identity (ID) for authentication [33]. But it can be easily intercepted by adversaries, which could induce serious threats to the BC systems, such as tracking, IBA, replay attack and device cloning. To relieve these vulnerabilities, a pre-shared encryption key is applied to secure BD identity or generate a random code, such as encrypted ID and Electronic Product Code (EPC) [34]. But this kind of methods still cannot prevent some interception adversaries with powerful capabilities, and it faces serious challenges regarding key distribution and update.

Due to limited supply energy and computational resource of BDs, light-weight authentication schemes are proposed since it is hard to employ complex cryptographic algorithms at BDs. Some protocols only employ XOR or simple hash functions to hide ID information with a pre-shared key to reduce complexity and cost, e.g., the EPCglobal C1G2 and YA-TRAP protocol [33], [35]. However, XOR and hash functions cannot provide sufficient security since an attacker can easily intercept the ID information by sniffing initial communications [36]. Therefore, only applying cryptography is infeasible to authenticate BDs with sufficient security to avoid authentication vulnerabilities and threats, especially for resource-limited BDs.

B. Physical Layer Authentication

A number of approaches have been proposed to authenticate a BD by extracting the physical layer features of its backscattered signals or BD hardware as its fingerprints [17], [21], [22], [24], [25], either using machine learning [21], [37] or fingerprint matching [22], or both [27].

Reference [38] proposes a channel-based physical layer authentication (PLA) enhancement scheme by exploiting the inherent two-dimensional properties of multipath fading channels, such as channel amplitude and multipath time delay spread. Reference [39] introduces a PLA scheme by applying tagged signals as a proof of authentication, which is generated by a secret key shared between a sender and an authenticator. Both schemes work in a non-BC scenario, not for BD authentication. That is their working scenarios are different from

ours. These two schemes cannot support mobility, nor trace attackers and identify their number. Mutual authentication is not mentioned. The scheme in [38] can only resist identity impersonation/spoofing attacks, while the scheme in [39] can resist replay and impersonation attacks to some extent. Since both schemes do not apply ML, thus there is no need to train an ML model. Both request signal analysis and are not resilient to imperfect channel state estimation.

Some researchers apply machine learning for authentication based on physical layer features. Qiu *et al.* [21] proposed a deep learning-based authentication approach to learn and track the variations of channel characteristics in order to enhance authentication security and detect spoofing attacks. Liu *et al.* [37] used channel state information (CSI), available from off-the-shelf WiFi devices to conduct fine-grained user authentication and detect a spoofer. They applied machine learning based user authentication techniques (e.g., Support Vector Machine) to distinguish two users with similar signal fingerprints. It requests labelled data to train a machine learning (ML) model and depends on a signal analyzer for CSI analysis. Commonly, the above two approaches require data to train a learning model. In practice, it is normally hard and costly to collect sufficient data and label them to train an effective model that can be applied into various contexts. Whether they can work for mobile BD authentication needs to be further investigated. Mutual authentication is not considered in these two works. It is impossible for them to track the locations of attackers and identify their number.

BD authentication in RFID systems has been studied in the literature, but current schemes still suffer from some shortcomings. Geneprint [22] leverages the internal similarity of backscattered signals, concretely the covariance among two consecutive sequences of RN16 preamble, as fingerprints for tag authentication in RFID systems. Wang *et al.* [24] proposed Hu-Fu for RFID tag identification by utilizing the unique inductive coupling feature of two adjacent tags as identification information. Danev *et al.* [25] proposed an authentication scheme by extracting the modulation shape and spectral features of backscattered signals. Zanetti *et al.* [17] proposed a fingerprinting scheme that measures average baseband power (ABP) and time interval error (TIE) from the fixed RN16 preamble as tag identification information. Mehmood *et al.* [23] used the non-reciprocity of the residual channel between a reader and an RFID tag as the fingerprint of the tag to defend fault tag attacks. However, such features, e.g., modulation shape and TIE, are sensitive to environments, thus unstable for device authentication when BD is moving. The above schemes easily suffer from imperfect channel estimation. They are also vulnerable to signal counterfeit, replay and relay attacks although they can resist identity impersonation attacks. This is because attackers can easily learn the constant RF signal features of a genuine device or record its RF signals to mimic a legitimate device [26]. Besides, using a sophisticated signal analyzer usually increases hardware cost, making it infeasible as a general solution for BC systems. Unfortunately, all above schemes cannot support BD mobility, attacker tracing, and attacker number identification, as well as mutual authentication.

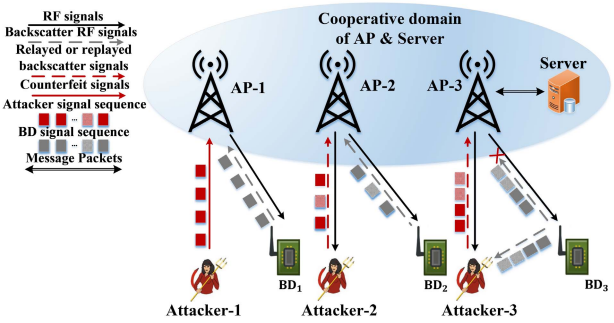


Fig. 1. System model and security model.

There are other works in the literature, which utilize backscatter or spacial information for device authentication [20], [27], [40]. For example, ShieldScatter [27] is an IoT device authentication scheme by using multiple backscatter tags attached to AP to intentionally create fine-grained multipath signatures for helping AP identify a device and detect impersonation attackers. Both machine learning and fingerprint matching are applied in authentication. It can work in both static and dynamic environments. For authentication, model training is needed due to the use of support vector machine (SVM). Wang [20] proposed an authentication scheme for mobile wireless sensor networks, assisted by physical layer features. It explores the reciprocity and spatial uncorrelation of the wireless channel to verify a transmitter and judge whether all messages are from a same sender and detect spoofing attacks. This scheme can overcome imperfect channel estimation to some extent. But device mobility is not supported by this scheme. Mutual authentication is not mentioned in the above two works. They can detect an attacker, but cannot trace its location and identify the number of attackers.

Table I compares the above reviewed works with BCAuth in terms of mobility support, attacker location tracing, attacker number identification, mutual authentication, detected attacks, ML model training avoidance, and signal analyzer independence, as well as resilience to imperfect channel state estimation. We can see that the literature still lacks an advanced BD authentication scheme to support mobility, offer attacker tracing, identify attacker number, and overcome imperfect channel estimation to some good extent. BCAuth shows great advance with regard to the above merits and functionalities. It can detect not only impersonation or spoofing attacks, but also signal counterfeit, replay and relay attacks. Thus, it also shows great advance on attack resistance. In addition, it does not request data collection for ML model training, neither depends on an expensive signal analyzer. Thus, it is also economic and easy to be deployed in practice.

III. BCAuth Overview

In this section, we first introduce the system model and security model of BCAuth, and then overview the procedure of BD authentication and attack tracing.

A. System Model

Fig. 1 shows the system model of BCAuth. We consider a hierarchical BC system with three types of parties in BCAuth:

a server, access points (APs) and backscatter devices (BDs). Multiple APs are set up in fixed positions and are controlled by the server. The server controls each AP to simultaneously transmit RF signals to power BDs within its coverage and receive information from the BDs [8], [41], [42]. BDs and potential attackers could be static and mobile. All BDs within the coverage of one AP alternately perform backscatter transmission in a time-division manner, as designed in prior arts [2], [4], [5], [43]. In each slot of time, the AP selects only one BD by signaling, making the RF signals backscattered only from this BD without mutual interference at the receiver AP. This setting is the same as our previous work and proved feasible [2]. The selected BD responds to the AP via backscattering and then communicates with the server. The server cooperates with the AP to verify the identity of the BD in each message responded by the BD to prevent receiving illegal messages from attackers. Herein, APs can estimate the positions of BDs based on the backscattered signals by probing their RSS and AoA, and/or measuring Time of Arrival and AoA according to some trustworthy method as we have previously explored [28]. And then APs help the server authenticating the BDs within its coverage when the BD transmits message packets. With this method, we can reduce authentication cost at the server when a large-scale of BDs are deployed in the BC system with the support of multiple APs since this method does not rely on any expensive signal analyzer. Additionally, we can achieve accurate positioning based on truth discovery from the positions provided by multiple APs [28] in order to reduce the impact of estimation errors of RSS and AoA, and subsequently improve the accuracy of authentication (Refer to our experimental results shown in Fig. 8 and Fig. 11. in Section VI.)

BCAuth solves the following fundamental problems in the above modeled BC system. First, the server with the help of the AP can validate whether the replies of a BD are from a legitimate BD or an attacker. Second, BCAuth is capable of adapting application scenarios to authenticate static or dynamic BDs. Third, in case authentication failure due to attacks, the server can conduct further analysis to detect the number of attackers and trace their locations.

B. Security Model

We assume that the server is trusted. APs work as BCAuth design. They are cooperative with each other and with the server. APs communicate with the server in a secure way by applying some existing technique, such as OpenSSL (https://www.openssl.org/). Each AP can obtain a synchronized timestamp (e.g., from public GPS signals or a reliable time system [41]). Meanwhile, a timestamp validation algorithm [42] is applied by the server to verify the truth of the timestamp attached to each AP. The server controls all APs to work in a synchronized way, i.e., sending RF signals following the same time slot arrangement. During initialization, we assume that the source AP sends a synchrony message to all BD and APs, or that a BD injects some constant preamble symbol in each message when it backscatters signals for information transmission to realize synchronization at all APs. Since the BD is a passive device to only reflect AP's signals, it relies on the AP to direct its communication. Once APs are synchronized, it is easy to let BD follow the pace of APs based on their synchronization messages or handshaking messages. Regarding the working scenario of BCAuth, we assume that it is applied into such a field as a warehouse where BDs are attached to goods and an apartment where BDs are embedded into different objects. In such scenarios, we model the relatively complex channel between APs and BDs as the Rayleigh channel model because there could be some obstacles blocking the main channel between the AP and the BD.

We assume a powerful attacker that can eavesdrop on any communications between the APs and legitimate BDs to obtain the identity information of BDs. Then, the attacker can masquerade as genuine BDs by modifying its identity to carry out identity impersonation attacks. Besides, in the physical layer, the attacker can intercept the features of the RF signals from a genuine BD and then perform wireless impersonation attacks. Concretely, we define a counterfeit attacker, i.e., Attacker-1 in Fig. 1, who can eavesdrop on the backscattered signals of a genuine BD and record corresponding features, i.e., RSS and AoA. Then, it similarly backscatters the RF signal to imitate the behavior of the genuine BD to make the RF signals arriving at the AP the same as the previous ones of genuine BD. In addition, a replay attacker could record any RF signals backscattered by BDs and replay the identical signals of prior communication to the AP, i.e., Attacker-2 in Fig. 1. Besides, a signal relay attacker just relays the backscattered signal of a legitimate BD to the AP when the BD is backscattering signals, i.e., Attacker-3 in Fig. 1. And, there could be a more powerful signal relay attacker that is able to block the backscattered signal from the BD to the AP. Based on the attacking techniques above, the attacker intends to insert fake messages to deceive the server when the server is communicating with the genuine BDs. Thus, some or all of the received messages of the server might be sent from an attacker, as shown in Fig. 1.

Besides, since BDs could be static or mobile, the attacker could behave adaptively as static or mobile. We assume that attackers normally take attacking actions that are most likely to succeed. For example, the attacker chooses to stay static when BDs are static, otherwise hardly for it to launch counterfeit and replay attacks. For a mobile BD, we mainly consider such a case that there is only one attacker that is close to a legitimate BD and follows its moving trajectory. But in practice, an attacker cannot be very close (less than a few multiples of wavelength) to any legitimate BDs in a physical space.

C. BCAuth Overview

BCAuth performs BD authentication and attack detection by analyzing the physical spatial information of BDs, i.e., positioning information. RSS and AoA values are widely available in wireless communication systems and highly correlated with the physical location of a wireless device. Location can be calculated with several ways, for example based on RSS

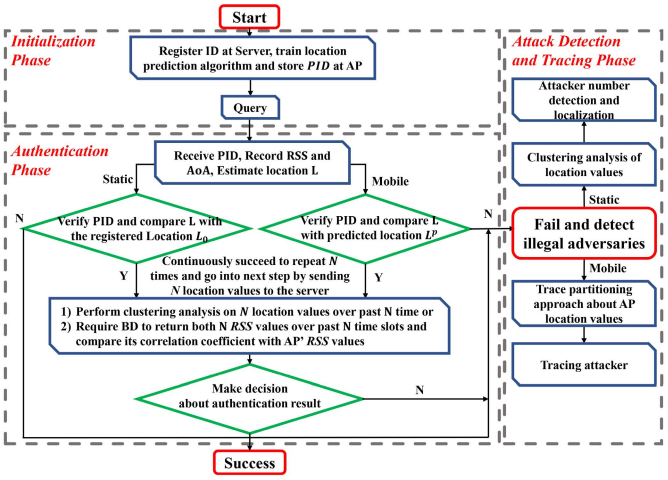


Fig. 2. BCAuth overview.

and AoA, or based on time of signal arrival and AoA, etc. RSS and AoA are independent with each other. Different ways of location calculation and multiple positioning results reported by different APs benefit accurate positioning [28], which helps finding a ground truth. In this paper, we use RSS and AoA to calculate the location of a signal source, as an illustration. Thus, the accurate positioning information estimated from the RSS and AoA values represents a mean to distinguish different BDs. Therefore, the main idea of the proposed BCAuth is to use the positioning information of a BD to perform authentication and attack detection. As shown in Fig. 2, the proposed BCAuth mainly consists of three parts: a) Initialization phase, b) Authentication phase, c) Attack detection and tracing phase.

1) Initialization Phase: In this phase, a new BD need to register at the server with their identity information (ID). This BD sends its ID via backscatter communication to a AP that it belongs to and the server. Except for receiving ID, the AP records the corresponding RSS and AoA values of backscattered signals. Then, the AP will calculate the location information L_0^r and store the vector $\{ID, RSS_0, AoA_0, L_0^r\}$ for subsequent BD authentication. If BDs are mobile, the AP build a position prediction algorithm based on physics-based approaches [44], [45] or RSS fingerprint-based approaches [46], [47] by collecting trajectory information of the BD over a period of time [44], [45]. The initialization is needed when every new BD enters the coverage of an AP controlled by the server and joins the system for the first time.

2) Authentication Phase: This phase is composed of two authentication stages to successively validate whether a message is responded by a legitimate BD or an attacker. When the AP receives a message packet from a BD, the AP measures the RSS and AoA values of its backscattered signals. Based on the received ID, RSS and AoA information, the server and the AP can perform multi-stage authentication to enhance BD authentication security.

Concretely, in the first stage, the AP authenticates the BD via the received ID and its positioning information L. In case of a static BD, the AP directly compares the vector $\{ID, L\}$ with registered ones $\{ID, L_0^r\}$ to determine whether it is legitimate. In case of a mobile BD, the AP can predict BD's location L^p based on its previous location trajectory and compare $\{ID, L\}$ with $\{ID, L^p\}$ to validate the BD. In the second stage, after receiving N messages of the BD relayed from the AP for a certain period, i.e., N time slots, the server re-authenticates BD based on a number of its N positioning values. In BCAuth, the server utilizes a clustering algorithm to analyze spatial correlation of N positioning values to detect potential attacks, so as to re-authenticate the BD regarding its N messages. Moreover, in case that the BD can measure RSS by itself, it can send N RSS values measured in the past N time slots to the server for BD re-authentication. Concretely, based on the reciprocity of a backscatter channel between the BD and the AP, the server can validate the BD by comparing the two sets of N RSS values provided by the BD and recorded by the AP, respectively. This technique can be viewed as an upgraded method to replace the clustering-based re-authentication in the mobile BD case in order to achieve better performance. With the capability of RSS measurement, the BD can also authenticate AP by calculating the correlation of its recorded RSS values of AP with the ones provided by AP. For this purpose, the AP need to send its RSS values to BD. Thus, mutual authentication can also be realized.

3) Attack Detection and Tracing Phase: In case BD authentication fails and there are potential attackers, the server can utilize the recorded positioning values to detect the number of attackers and track their locations. In case of static BDs, first of all, a clustering algorithm run by the server is used to cluster the location values, so as to determine the number of attackers. Then, the server utilizes the medoids or the average values of different RSS clusters to localize the positions of attackers. In case of mobile BDs, our design can filter out abnormal positioning values and try to trace the motion trajectory of potential moving attackers.

IV. BCAuth Design

This section describes the details of BCAuth design, including initialization, multi-stage authentication, as well as attack detection and tracing.

A. Initialization

When deploying the BC system in practice, all BDs need to register at the server with their real identities. The server firstly starts an inventory round by controlling all APs to send an identification query. Based on the slotted ALOHA protocol with multiple BDs [48], a BD in a selected slot responds with a random number RN and $metaID = h(ID||k)$, where $h(ID||k)$ is the hash value of its real identity ID and a secret key k, which is preset by an upper security protocol. The server authenticates the BD by searching its database to find the corresponding record $metaID$, because the information (e.g., ID, k and $metaID$) of all legal BDs has been stored in the server ahead of time. If there is, it sends an acknowledgement message to the AP, including $h(RN)$, $h(ID)$ and $h(k)$. $h(ID)$ is the hash value of ID. Otherwise, it returns an authentication failure message to the AP, which no longer receives messages from the underlying BD. After receiving $h(ID)$ and $h(k)$ from

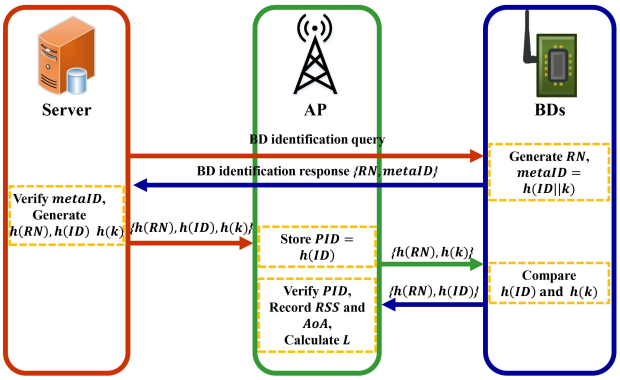


Fig. 3. Initialization phase for BD attachment.

the server, the AP retains $PID = h(ID)$ as the pseudonym of BD and sends $h(RN)$ and $h(k)$ to the BD. The BD compares $h(k)$ with the hash value of the secret key it stores. If the values match, it authenticates the server and sends $h(ID)$ to the AP. The AP compares PID from the server with the one from the BD to authenticate the BD. If they match, the AP records the RSS and AoA values of the backscattered signals and calculates its location L_0 as a spatial feature vector $\{RSS_0, AoA_0, L_0\}$. If the BD is mobile, the AP needs to obtain multiple locations to setup a location prediction algorithm based on physics-based approaches [44], [45] or RSS fingerprint-based approaches [46], [47]. The procedure of the initialization phase for BD identification and attachment at AP is shown in Fig. 3.

B. Preemptive Authentication at AP

With such stored spatial information, AP can directly perform preemptive authentication of BDs with two factors, consisting of PID and positioning information, when they subsequently transmit messages to the AP. As in existing protocols of BC systems, AP firstly broadcasts a query command to select a BD and then transmits a carrier signal to bear information from the BD via backscatter communications. We propose to design the carrier signal, assumed in time slot t_j, with a random transmission power P_{t_j} in our scheme. The different random power P_{t_j} selected in each slot can make the received RSS value unpredictable in subsequent slots, so as to prevent RF signal counterfeit and replay attacks. This is because the random power can make RSS at BDs unable to form any patterns, as RSS is mainly correlated to the transmission power and the distance between the BD and the AP. The signal power received by the BD is as follows,

$$P_{t_j}^{BD} = P_{t_j} h(d) + P_w^{BD}, \tag{1}$$

where $h(d)$ is the power attenuation function of the downlink channel with a parameter d, which is the distance between the BD and the AP. P_w^{BD} is the power of environmental noise at the BD.

After receiving the query command, the selected BD backscatters the carrier signal and transmits its pseudonym ID and other information in a message packet. The power of the backscattered signal received by the AP is as follows,

$$P_{t_j}^{AP} = \kappa P_{t_j} h(d) h^b() + \kappa P_w^{BD} h(d) + P_w^{AP}, \tag{2}$$

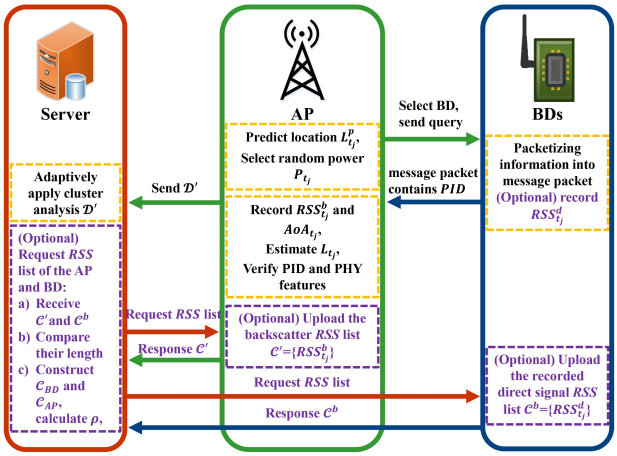


Fig. 4. Multi-stage authentication procedure.

where $h^b(d)$ is the power attenuation function of the backscatter channel, κ is the backscatter coefficient of the BD, and P_w^{AP} is the power of environmental noise at the AP. We measure the received signal power $P_{t_j}^{AP}$ as the RSS value $RSS_{t_j}^b$, and calculate the distance d with the RSS value based on function $h(d)$ by assuming $h(d)$ is equal to $h^b(d)$ because of the channel reciprocity between the BD and the AP. Besides, we do not consider the P_w^{BD} at the AP since it suffers the attenuation of backscatter channel $h^b(d)$ and the backscatter coefficient κ.

The AP obtains the pseudonym ID from the backscattered message and records the RSS value, $RSS_{t_j}^b$, and the AoA value, $AoA_{t_j}^b$ of the backscattered signals. Then, the AP calculates the location $L_{t_j} = (x, y)$ of the BD as follows:

$$x = x_{AP} + d \sin AoA_{t_j}^b, \tag{3}$$

$$y = y_{AP} + d \cos AoA_{t_j}^b, \tag{4}$$

$$d = \frac{1}{2}\sqrt{\kappa P_{t_j} / 4\pi RSS_{t_j}^b}, \tag{5}$$

where (x_{AP}, y_{AP}) is the location of the AP. With PID and location information L_{t_j}, the AP performs a preemptive authentication to verify whether this message packet is from a legitimate BD. With the calculated location and location history, AP determines whether the BD is static or mobile and adaptively adopts different detailed techniques for authentication. Concretely, based on the RSS and AoA of BD backscattered signals, the location of BD can be calculated according to Eq. (3)-(5). If the calculated locations of the BD with the same PID at different time slots are changed, we can judge that the BD is moving. Otherwise, it is stationary. If the BD is static, the AP compares $\{PID, L_{t_j}\}$ with the registered one $\{PID, L_0\}$. If the BD is mobile, the AP compares $\{PID, L_{t_j}\}$ with $\{PID, L_{t_j}^p\}$, which consists of the predicted location of BD, to authenticate the device. This process can be iteratively used to authenticate the BD regarding each transmission message packet.

If succeeding over past N time slots, the AP sends the location value list $\mathcal{D}' = \{L_i, L_{i+1}, \ldots, L_{i+N-1}\}$ with N values to the server. The server can exploit the spatial correlation of N location values to re-authenticate the BD with a clustering-based method. In this way, it can also reduce the server's

burden without verifying the identity of N message packets one by one. In some specific case, if the BD has the ability to measure the RSS of the incident carrier signal from the AP, it records the RSS value RSS_{t_j} of the carrier signal. In such a case, the server can require the BD to return its RSS value list $\mathcal{C}' = \{RSS_i, RSS_{i+1}, \ldots, RSS_{i+N-1}\}$ over past N time slots. With RSS list $\mathcal{C}^b = \{RSS_i^b, RSS_{i+1}^b, \ldots, RSS_{i+N-1}^b\}$ uploaded from the AP, the server can exploit these two RSS lists to re-authenticate the BD based on the reciprocity of the backscatter channel, and vice versa. Thus, mutual authentication can be achieved. Fig. 4 illustrates the detailed procedure of multi-stage authentication, which consists of preemptive authentication at the AP and the re-authentication at the server.

C. Clustering-Based Re-Authentication

After N time slots, each of which the AP receives a message from a BD, the AP uploads N location values to the server for re-authenticating the BD. The server can determine the number N by considering the trade-off between authentication accuracy and time. Due to the previous preemptive authentication, the server only needs to conduct a one-time verification on the identity of N message packets. This method is able to reduce the authentication overhead at the server by distributing authentication tasks to APs with high scalability.

1) Authenticating Static BDs: In the case of a static BD, the sever conducts spatial correlation analysis on the set of location values by applying clustering algorithms, such as K-means and Partitioning Around Mediods (PAM) method [49]. The calculated location values are highly related to the practical physical location, though they are affected by estimation bias. If there is no attack, for each BD with the same PID, all location values are close to each other. They fluctuate around a mean value, because the RSS and AoA of backscattered signals are affected by noise and environments. And the distances among different clusters are small if dividing N locations into several clusters. However, if under an attack, there is more than one device at a different physical location claiming the same PID. The location values claiming the legitimate BD are mixed with the values related to at least one different location. Normally, the location values of the legitimate BD belong to the same cluster, while the location values of attackers located in different physical locations should fall into different clusters. And the distances among different clusters, especially the distance between the cluster of the legitimate BD and the cluster of the attacker, are larger than the distances when there are no any attacks.

This observation suggests that the server can conduct clustering analysis on the calculated locations claiming the same PID to achieve one-time authentication on the BD identity of N message packets. If there are N location values related to the same PID, the clustering algorithm partitions N values into K (i.e., $K = 2$) disjoint subsets S_j. In our simulation, we use the k-means algorithm and set $K = 2$ during clustering by classifying the locations into two classes: normal locations of the legitimate BD and abnormal locations of potential attackers. BCAuth provides a generic BD authentication framework for BC. Obviously, more advanced clustering algorithms (e.g., the algorithms that can automatically decide the number of clusters) can be applied to further improve BCAuth performance. Each subset S_j contains contains N_j values, so as to minimize the sum-of-squares criterion to obtain an optimal clustering result as follows:

$$J_{min} = \sum_{j=1}^{K} \sum_{L_m \in S_j} ||L_m - \mu_j||^2, \tag{6}$$

where L_m is a location value representing the m-th value and μ_j is the geometric centroid of subset S_j.

Besides, we formulate the re-authentication as a statistical significance test, where the null hypothesis is

$$\mathcal{H}_0 : \textit{success} \text{ (no attack)}. \tag{7}$$

In this significance test, a test statistic $\mathbf{T}$ is used to evaluate whether the observed data belongs to the null hypothesis or not. If the observed test statistic $\mathbf{T}^{obs}$ differs significantly from the hypothesized values, the null hypothesis is rejected and we claim that the re-authentication fails and there is a presence of attacks.

Based on the null hypothesis, we thus choose the distance between any two cluster centroids as the test statistic $\mathbf{T}^{obs}$ for re-authenticating BD and detecting potential attacks,

$$D_{i,j} = ||\mu_i - \mu_j||, \tag{8}$$

where $i, j \in \{1, 2, \ldots, K\}$. Under the conditions of *success* (no attack), the distances between the centroids of different clusters should be small. This is because all positioning values are in high proximity to the real physical location. Otherwise, there is more than one device at a different location claiming the same pseudonym ID. As a result, the distance $D_{i,j}$ becomes large as the clusters are associated with different locations. Therefore, we utilize the distance D_{t_j} to validate the BD regarding its N messages at the server at once.

Next, we use empirical training from a collected data set to determine a threshold for defining the critical region for a significance testing. Appropriately setting a threshold τ can ensure the accuracy of re-authentication and its reliability, considering false authentication or false acceptation. In the phase of initialization, we can obtain multiple calculated locations of a BD in a known location and obtain the distances between any two centroids to determine the threshold. In the phase of authentication, based on the positioning values of a BD with PID, we can calculate an observed value $D_{i,j}^{obs}$. Our condition for declaring re-authentication failure and attack is:

$$D_{i,j}^{obs} > \tau, \exists i, j \in \{1, 2, \ldots, K\}. \tag{9}$$

2) Authenticating Mobile BDs: In the case of a mobile BD, because the BD is moving around, the distribution of estimated positions is highly dependent on the movement pattern of the BD. It is prohibitive to directly apply clustering algorithms to analyze all position values even with the knowledge of the movement pattern of the BD. Based on the position values, the server exploits a partitioning approach to separate the positioning values into two classes in each time interval. Then, these two classes are used to reconstruct two location trajectories over all time intervals. In a *success* (no attack) situation, two trajectories will be correlated and directly reflect the movement pattern of the BD, whereas they are uncorrelated

under potential attacks as the movement trajectories of the victim BD and the attacker are different. Therefore, by examining the degree of correlation of the location trajectories in physical space, the server can determine whether there is an attacker present in the system.

We denote all position values sent from the AP related to one BD within a time window T_N is $\mathcal{L}$. We equally divide all values into M non-overlapping time intervals. Denote positioning values in m-th time interval as $\mathcal{L}_m$, while $\mathcal{L}$ can be represented as $\{\mathcal{L}_1, \mathcal{L}_2, \ldots, \mathcal{L}_M\}$. Within each time interval, the positioning values are partitioned into two classes $\mathcal{L}_{a,m}$ and $\mathcal{L}_{b,m}$. We assume that one belongs to the victim BD and the other belongs to a potential attacker. Similarly, we exploited K-means with $K = 2$ as the location partitioning approach to partition the location values $\mathcal{L}_m$ in each time interval. Identically, we need to analyze how to obtain the optimal threshold that can minimize the partitioning error of the positioning value.

From two partitioned classes in each time interval, we further reconstruct two location trajectories, $V_1 = \{v_{1,1}, \ldots, v_{1,M}\}$ and $V_2 = \{v_{2,1}, \ldots, v_{2,M}\}$, over the whole time window T_N. If under attacks, one trajectory is associated with the victim BD and the other associated with the potential attacker. These two reconstructed trajectories are next used to authenticate the identity of all packets by measuring their similarity via a correlation coefficient. Since each time interval is small, we can directly use the average value, represented as $\bar{l}_{a,m}$ and $\bar{l}_{b,m}$, of each location class $\mathcal{L}_{a,m}$ and $\mathcal{L}_{b,m}$ in a time interval for trajectory reconstruction. Then, in the m-th time interval, trace reconstruction needs to determine whether to assign $\bar{l}_{a,m}$ to $v_{1,m}$ and $\bar{l}_{b,m}$ to $v_{2,m}$ or on the contrary.

Since we do not have prior knowledge of the movement patterns of BDs except for past position values, we cannot directly construct the trajectory of a moving BD. However, there is a temporal spatial constraint presented in location trajectory that the position values within few consecutive time intervals are correlated [50]. Though, the location trajectory in the whole time window T_N may not follow any form of curves, the location trajectory within several small time intervals can be modeled to follow a conic curve [50]. In the m-th time interval, we can use some past position values based on conic curve fitting to predict the location values in the $(m+1)$-th time interval. We then compare the predicted values $l^p_{a,m+1}$ and $l^p_{b,m+1}$ with $\bar{l}_{a,m+1}$ and $\bar{l}_{b,m+1}$ and decide how to assign $\bar{l}_{a,m+1}$ and $\bar{l}_{b,m+1}$ to $v_{1,m+1}$ and $v_{2,m+1}$, respectively.

We develop the prediction algorithm to predict the location values during the location trajectory reconstruction. The algorithm utilizes the determined location values in the last L time intervals ranging from $(m-L+1)$-th to m-th time interval to perform conic curve fitting and predict the location values $l^p_{a,m+1}$ and $l^p_{b,m+1}$ in the $(m+1)$-th time interval:

$$l^p_{a,m+1} = k_{a,1,m} + k_{a,2,m}(m+1) + k_{a,3,m}(m+1)^2, \quad (10)$$

and

$$l^p_{b,m+1} = k_{b,1,m} + k_{b,2,m}(m+1) + k_{b,3,m}(m+1)^2, \quad (11)$$

where the coefficients $\{k_{a,1,m}, k_{a,2,m}, k_{a,3,m}\}$ and $\{k_{b,1,m}, k_{b,2,m}, k_{b,3,m}\}$ are determined by the latest L location values $\{v_{1,m-L+1}, \ldots, v_{1,m}\}$ and $\{v_{2,m-L+1}, \ldots, v_{2,m}\}$ according to the least-squares polynomial approximation [50]. L is a variable, which is set as $L = 4$ in our study based on experimental results. Setting L as 4 can obtain sufficient authentication performance with relatively low time complexity for mobile BD authentication. Besides, it is suggested to set it with a small value based on [51] when constructing the trajectory of a moving object with a relatively low velocity.

We further define the prediction error as:

$$p_{e,1} = (l^p_{a,m+1} - \bar{l}_{a,m+1})^2 + (l^p_{b,m+1} - \bar{l}_{b,m+1}), \quad (12)$$

and

$$p_{e,2} = (l^p_{a,m+1} - \bar{l}_{b,m+1})^2 + (l^p_{b,m+1} - \bar{l}_{a,m+1}). \quad (13)$$

If $p_{e,1} \leq p_{e,2}$, we assign $\bar{l}_{a,m+1}$ to $v_{1,m+1}$ and $\bar{l}_{b,m+1}$ to $v_{2,m+1}$. Otherwise, we assign $\bar{l}_{a,m+1}$ to $v_{2,m+1}$ and $\bar{l}_{b,m+1}$ to $v_{1,m+1}$. In the initial setup, when $m = 1$, we set $v_{1,1} = \bar{l}_{a,1}$, $v_{2,1} = \bar{l}_{b,1}$ and $l^p_{a,2} = v_{1,1}$, $l^p_{b,2} = v_{2,1}$. When $1 < m < L$, we use the first m location values of V_1 and V_2 to fit conic curves and then predict and determine the $(m+1)$-th location values. Therefore, we finish reconstructing two location trajectories whose similarity is exploited to perform re-authentication of the BD regarding its all message packets.

In a *success* (no attack) situation, the location trajectories V_1 and V_2 are correlated and belong to one legitimate BD, whereas they are uncorrelated under potential attacks as they belong to the victim BD and the potential attacker, respectively. Then, we measure the correlation coefficient to capture their similarity as below [52], [53],

$$r = \frac{\sum_{i=1}^{n} (v_{1,i} - \bar{v}_1)(v_{2,i} - \bar{v}_2)}{(n-1)\delta_1\delta_2}, \quad (14)$$

where $v_{1,i} \in V_1, v_{2,i} \in V_2, \bar{v}_1$ and $\bar{v}_2$ are the means of V_1 and V_2, δ_1 and δ_2 are the standard deviations of V_1 and V_2, respectively. $0 \leq |r| \leq 1$. Similarly, we exploit the similarity r as the test statistic for re-authentication and potential attack detection. Thus, only $|r|$ near 1 indicates V_1 and V_2 trend to change together with high positive linearity, which means they are from the same BD. Thus, the server re-authenticates the legitimate BD about the location values associated to its N message packets uploaded from the AP.

D. Reciprocal Channel Variation-Based Authentication

If the BDs have the ability to measure the RSS value of the carrier signal transmitted by AP, the server can require the BD to feedback the list of a number of N RSS records, $\mathcal{C}' = \{RSS'_{t_{j-N+1}}, \ldots, RSS'_{t_{j-1}}, RSS'_{t_j}\}$, over a past certain period of time. Then, with the RSS record list $\mathcal{C} = \{RSS_{t_{j-N+1}}, \ldots, RSS_{t_{j-1}}, RSS_{t_j}\}$ uploaded from the AP, the server can validate the BD identity of its N message packets at once. It can perform re-authentication by comparing these two RSS lists [15] by replacing the clustering-based method as described above, especially for improving authentication performance in the case of mobile BDs. This is because the increase of BD moving speed could worsen the authentication

performance of the clustering-based method. The authentication performance can be partly improved by increasing the number of the coefficients of the conic curve to obtain more accurate trajectory, but also inducing high complexity of reconstruction for the server. With the development of BC technology, BDs will be able to measure the RSS values of incident signals when they are harvesting RF signal energy [54]. Thus, the BDs can measure the RSS values of the carrier signals from the AP and send them back to the server for device re-authentication. And identically, the AP can upload the RSS values of the backscattered signals from the BD to the server.

After receiving the two RSS record lists, the server first compares the length of $\mathcal{C}'$ and the length of $\mathcal{C}$. If they are not equal, AP directly outputs an authentication failure. This is because if the length of $\mathcal{C}$ is larger than the length of $\mathcal{C}'$, it means some transmissions are probably from attackers that impersonate the legitimate BD to communicate with the AP. Otherwise, the server then constructs two corresponding vectors as follows $\mathcal{C}_{BD} = \{2RSS'_{t_{j-N+1}}, \ldots, 2RSS'_{t_j}\}$ and $\mathcal{C}_{AP} = \{P_{t_{j-N+1}} + RSS_{t_{j-N}}, \ldots, P_{t_j} + RSS_{t_j}\}$ based on $\mathcal{C}'$ and $\mathcal{C}$, respectively. This construction is performed to eliminate the impact of the random transmission power P_{t_j} on two RSS lists. Because of the reciprocity, the channel attenuation $h_{a,b}$ from the AP to the BD is equal to the channel attenuation $h_{b,a}$ from the BD to the AP. We denote the RSS received at the BD is $RSS'_{t_j} = P_{t_j} - h_{a,b}$ and the RSS at the AP is $RSS_{t_j} = P_{t_j} - h_{a,b} - h_{b,a} = P_{t_j} - 2h_{a,b}$ with round-trip channel attenuation. Thus, the corresponding elements in two calculated vectors about the RSS values are equal to each other, i.e., $2RSS'_{t_j} = P_{t_j} + RSS_{t_j}$. If no attacks, the two constructed RSS vectors are highly correlated with each other, though the RSS values are affected by environmental noises.

If there is an attacker that sends fake messages to AP by impersonating the legitimate BD, the location of the attacker is different from the one of the BD. The attacking channel from the attacker to the AP is uncorrelated with the legal channel between the legitimate BD and the AP. This is because the well-known Jakes model [55] states that there is a rapid decorrelation between the attacking channel and the legal channel over a distance of over half a wavelength, and even independent if the distance is over several wavelength. Thus, the server can measure the correlation coefficient ρ_{AP} of the two calculated RSS vectors for one-time re-authenticating the BD regarding its N message packets [15], [56]. If the correlation coefficient is larger than a threshold $\rho_{AP} > \tau_{rss}$ set based on previous data, the server assumes successful re-authentication without any attacks and considers the BD as a legitimate BD. Otherwise, it outputs authentication failure and raises an alarm about potential attacks. The gist behind the method is that if there is no attack, the RSS vectors of the BD and AP should be highly correlated according to the reciprocity of the wireless channel. If there is an attack, the RSS vector of AP is a mixture of the RSS value of the legitimate BD and the RSS value of the attacker, which seriously degrade their correlation.

Likewise, the BD can authenticate the AP by requesting it to feedback its recorded RSS list. Thus, mutual authentication can be achieved between the BD and the AP. Since the BD does not have any knowledge of the random signal power P_{t_j}, it requests the AP to transmit the RSS value vector $\mathcal{C}'_{AP} = \{P_{t_{j-N+1}} + RSS_{t_{j-N+1}}, \ldots, P_{t_j} + RSS_{t_j}\}$. Then, it constructs its own RSS value vector $\mathcal{C}'_{BD} = \{2RSS'_{t_{j-N+1}}, \ldots, 2RSS'_{t_j}\}$. The BD firstly compares the length of $\mathcal{C}'_{AP}$ with the length of $\mathcal{C}'_{BD}$. If they are equal, the BD continuously calculates the correlation coefficient ρ_{BD} of these two vectors to authenticate the AP. If the coefficient is lower than a pre-set threshold, the BD outputs authentication failure and raises an alarm on the AP.

E. Attack Detection and Tracing

If the authentication of BD fails at the server, this implies that some adversaries are implementing attacks, such as counterfeit attack and replay attack. In this subsection, we conduct further analysis on the location values estimated by AP to determine the number of attackers and localize their positions. With the security assumption that there could be various attack actions, including static and mobile attackers, we only consider the cases where the attacks are most likely to succeed. Thus, in the static BD scenario, we only consider the case that attackers are static but with various amounts. And in the mobile BD scenario, we only consider there is a moving attacker that is close to a legitimate BD and follows its moving trajectory.

In the static case, since there could be several attackers, the locations with the same ID may be mixed with locations of both the legitimate BD and a number of attackers from different physical positions. Averaging location list L cannot differentiate the location of the legitimate BD and attackers, thus not feasible for localizing attackers. As we do not know how many attackers use the same ID to launch attacks, we need firstly to determine the number of attackers. With the N locations, discovering the attacker amount is a multi-class detection problem and is similar to determine how many clusters existing in the location set. Thus, we can directly utilize the existing methods, such as the SILENCE algorithm [57] and Silhouette Plot [58], over locations L to obtain the number of clusters. In our paper, we utilize the SILENCE algorithm to detect the attacker amount, since it takes the advantage of both the Silhouette Plot [58] and System Evolution [59] to eliminate the effect of location variations and outliers. With the number of clusters over N locations, we can determine the number of attackers and then estimate the location of each cluster with the mean of each cluster. All returned locations above include the location estimation of the legitimate BD and the potential attackers. To this end, we first determine the normal location of the legitimate BD via their relationship with the registered location and then determine the location of all potential attackers. Then, the server raise alarms with essential security information, including the location of the legitimate BD and the number of potential attackers with their estimated locations.

Besides, in the mobile case, we obtain two returned trajectories from the re-authentication process. However, even though we assume there is only one mobile attacker, the server cannot directly identify which trajectory is from the legitimate BD and which is from the attacker. Thus, we first perform

clustering to analyze the location values, i.e., $\mathcal{L}_m$, in each time slot and partition $\mathcal{L}_m$ into two classes $\mathcal{L}_{a,m}$ and $\mathcal{L}_{b,m}$. Then, we similarly apply the conic curve fitting method with previous location values of the legitimate BD to predict which class belongs to the legitimate BD. In this way, we can obtain the location of the moving attacker in m-th time slot, and so forth, until we get the locations of the attacker in all time slots. With all locations in every time slot, we can trace the moving trajectory of the attacker. Then, the server can raise an alarm by announcing the trajectories of both the legitimate BD and the attacker.

V. Security Analysis

In this section, we theoretically analyze BCAuth security with regard to identity impersonation attack, signal counterfeit attack, signal replay attack and signal relay attack. In addition, we further justify the rationality of BCAuth for static and mobile BD authentication.

A. Identity Impersonation Attack

Regarding the identity impersonation attack, it is assumed that an attacker knows which authentication scheme is used. In conventional communication systems, i.e., IEEE 802.11, an attacker can intercept PID of a legitimate BD and masquerade as the legitimate BD by transmitting the PID to the AP. But, BCAuth not only exploits the PID of the BD for device authentication, but also utilizes the position information as an additional factor to realize multi-stage authentication and re-authentication. Based on the equations 1 2, we can theoretically measure the distance d between the BD and the AP with $h^2(d) = RSS^b_{t_j}/\kappa P_{t_j}$ and the angle of backscatter signal $AoA^b_{t_j}$. Then, we can calculate the position of the BD with the above two parameters based on the equations (3)(4)(5) as the authentication proof of the BD. In practice, the attacker cannot be very close (i.e., less than a few multiples of wavelength) to any legitimate BDs in a physical space. If an attacker impersonates a legitimate BD to backscatter RF signals, the AP calculates a distinct location of the attacker. This is because the AP obtains $h^2(d) = RSS^b_{t_j}/\kappa P_{t_j}$ from the abnormal RSS value $RSS^b_{t_j}$, which is not correct regarding the legal BD.

Thus, the authentication attempts from the attacker only with the PID cannot succeed at AP. Besides, the server can perform attack detection by re-authenticating the BD with high accuracy. Therefore, with these two stages of authentication, BCAuth can provide enhanced authentication security to resist the identity impersonation attack. It does not introduce any additional overhead to the BD for authentication, since the BD only needs to respond with its identity like existing BC systems. Therefore, BCAuth is compatible with existing cryptography-based schemes in BC systems but with enhanced security by exploiting the advantages of both physical-layer authentication and upper-layer authentication.

B. Signal Counterfeit Attack

In the counterfeit or forge attack, an attacker tries to imitate the behavior of the legitimate BD to make the RF signals arriving at the AP the same as the ones of the BD. But the attacker cannot be very close (less than a few multiples of the wavelength) to any legitimate BDs in a physical space. Due to the independence of Rayleigh fading channel, the RF signals backscattered from the attacker have different and uncorrelated features, including RSS and AoA, from the ones of those from the legitimate BD. Besides, even the attacker observes the backscattered signal in a previous slot, it cannot know the random power of the backscattered signal in next slot. If the attacker backscatters signals according to the previous backscatter behavior of the BD, the AP obtains different position information regarding RSS and AoA, which are affected by channel environment and transmission power. Thus, the AP can detect the signal counterfeit attack by performing two-factor authentication even though the attacker uses the real PID of an legitimate BD.

C. Signal Replay Attack

In the signal replay attack, an attacker's goal is to record signals from a target BD in a digital form, and then re-transmit (or replay) these signals towards passing authentication. The attacker does not modify the captured signals, that is, the analog signal and the data payload are preserved. In the BC systems, the AP receives the information backscattered from a selected BD only when it is broadcasting the carrier signal. In BCAuth, the carrier signal is designed with a random signal power P_{t_j}, which is different in each time slot. Even a strong relay attacker staying in the same location as the legitimate BD re-transmits the recorded RF signals of a past time slot for authentication, the AP can easily identify and detect the signal replay attacks. This is because the AP calculates a completely unrelated location from the measured RSS value, which is affected by the random power assigned in the current slot.

D. Signal Relay Attack

Compared with the signal replay attacks, the signal relay attack just relays the backscattered signal of the legitimate BD to the AP. The AP can directly detect the signal relay attack, since the AP receives the signals including both the relayed signal and the backscattered signal from the BD. Besides, we assume that the attacker can work hard to block the backscattered signal from the BD. But the relay attacker keeps a certain distance from the BD, which makes the received signal at the AP spreading through a different cascade channel from the legitimate BD to the attacker and then to the AP. Thus, the AP obtains different position information of the relay attacker due to different RSS values of the backscattered signals, which makes the attack fail. Therefore, the signal relay attacker cannot succeed due to the preemptive authentication and re-authentication by checking BD location.

E. BD Authentication

Static BD: If a BD is static, through positioning, we can easily judge if the BD is the previously authenticated one by matching the currently calculated position of the BD with the one initially registered. Therefore, on the basis of the

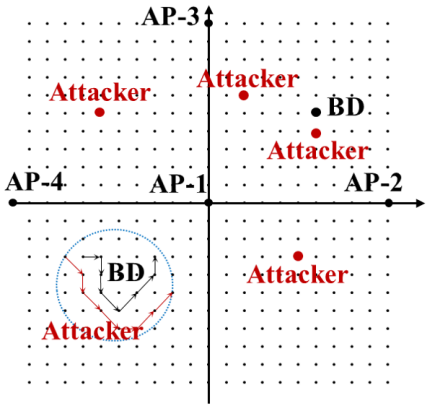


Fig. 5. Simulation setting.

initialization and preemptive authentication of BCAuth, it is capable of BCAuth to authenticate a static BD.

Mobile BD: If a BD is moving, BCAuth can judge if a moving BD is in a position as expectation through positioning and trajectory prediction based on location clustering, trajectory reconstruction, trajectories correlation, and conic curve fitting. Thus, BCAuth is capable of authenticating a moving BD.

VI. Performance Evaluation

In this section, we evaluate the performance of BCAuth in terms of different scenarios under different parameters with extensive simulations.

A. Simulation Setting

In our simulations, we consider a BC system consisting of several APs, a static BD surrounded by several attackers located around the BD within a circle with a radius of 0.5m, and a mobile BD followed by a mobile attacker, both moving within a defined circular region. As shown in Fig. 5, APs with multiple antennas are located at fixed positions to transmit carrier signals to power BDs. Concretely, AP-1 locates at (0,0) as the main AP for the one AP case. In the mobile BD case, we assume there are only one BD and one attacker, whose moving trajectories are emulated with a random way point (RWP) mobility pattern [60]. The attacker masquerades as an legitimate BD with intercepted PID information to inject illegal messages, when the legitimate BDs are communicating the AP.

We assume that BCAuth is applied into a field where there could be some obstacles blocking the main channel between the AP and the BD, such as a warehouse where BDs are attached to goods and an apartment where BDs are embedded into different objects. Thus, all related channels are modeled by an independent complex Gaussian random variable (Rayleigh fading) with its average power that follows a profile [61] and the channel reciprocity holds in each time block within coherence time. The forward and backward channel gains are set as $10^{-2}d^{-2}$ referring to our previous work [2] and [61], where d denotes the distance between the BD and the AP. The received noise power at APs is set to be $\sigma^2 = -30$dBm, and the maximum average power of the RF signal is assumed as P=1dBm. For each authentication simulation, the AP sends the carrier signals with random power to the BD, decodes the BD's identity information from backscattered messages. At the same time, the AP measures the features of the BD's backscattered signals, including RSS and AoA, and estimates the BD's location information for subsequent authentication. Similarly, we assume the attackers also perform backscattering to inject illegal messages, and the AP also measures related features of backscattered signals from the attackers.

B. Evaluation Metrics

After the preemptive authentication at AP and the clustering-based or reciprocal channel-based authentication at the server, an authentication result can be obtained with either "Accept" or "Reject". We use the following metrics to evaluate the performance of authentication: **1) Authentication Rate**, defined as the true positive rate that a legitimate BD is truly accepted by BCAuth and **2) False Acceptance Rate (FAR)**, defined as false positive rate, which is the rate that attackers are accepted by BCAuth. Then, we use a **Receiver Operating Characteristic (ROC)** curve to denote the trade-off between the authentication rate and FAR by varying authentication thresholds.

For the statistical characterization of detection performance, we consider the percentage that the number of attackers can be accurately detected over all possible testing attempts with a mixed number of attackers. Associated with a specific number of attackers, i, we define the **Hit Rate** $HR_i = \frac{N_{true}}{N_i^i}$, where N_{true} is the amount of the true positive detection for the attacker number i and N_i^i is the total amount of testing attempts for asserting attacker number i. And we define the false positive rate $FP_i = \frac{N_{false}}{N_i^{ni}}$, where N_{false} is the amount of false detection for the attacker amount i and N_i^{ni} is the total amount of testing attempts for asserting other attacker number. And the **Precision** is defined as $Precision_i = N_{true}/(N_{true} + N_{false})$. **F-measure** is originated from information retrieval and measures the accuracy of detection by considering both the Hit Rate and the Precision [57], [62]

$$F - measure_i = \frac{2}{\frac{1}{Precision_i} + \frac{1}{HR_i}}. \tag{15}$$

C. Simulation Results

In this part, we first evaluate authentication performance at the AP in the preemptive authentication step with numerical results. Then, we test the performance at the server for BD re-authentication with N location values, as well as the performance of detecting attacker amount in the static BD case.

1) Performance of Preemptive Authentication: In the case of a single BD without any attackers, Fig. 6 illustrates the impact of distances between BD and AP over authentication performance. As the distance increases, the authentication rate gradually decreases. This is because the increased distance reduces the signal-noise ratio (SNR) of the backscattered signal at the AP, which induces a greater error of location

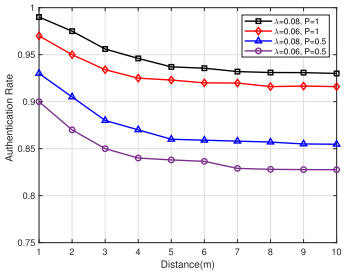


Fig. 6. Authentication rate vs. the distance between BD and AP (static).

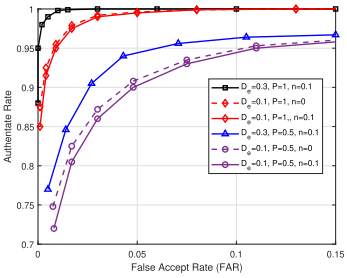


Fig. 7. ROC under different signal powers and distances between BD and an attacker.

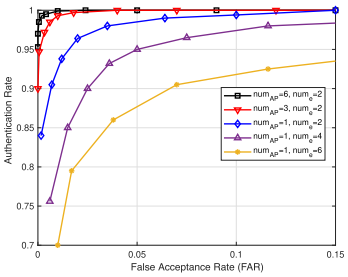


Fig. 8. ROC with different numbers of APs and density of attackers.

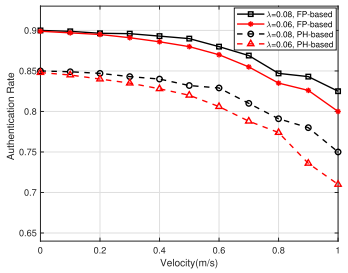


Fig. 9. Authentication rate vs. velocity of a moving BD.

estimation with RSS and AoA information. Likewise, if the AP reduces the average power of its carrier signals, the authentication rate also decreases. Thus, we mainly consider the impact of SNR on the measurement of RSS by setting different transmission power and the distance between BDs and APs. Besides, by setting a stricter threshold for the residual between the estimated location and the registered location, the authentication rate also drops.

Fig. 7 shows the ROC performance when there is one single BD followed by an attacker, under different distances (D_e) between the BD and the attacker, different average powers (P) of carrier signals, and different noises (n). From Fig. 7, we observe that when FAR grows, the authentication rate also grows in each curve. Therefore, in the preemptive authentication, by choosing an appropriate threshold, we are able to authenticate the BD with an expected authentication rate. Besides, as the AP reduces the average power of its carrier signals, both the authentication rate and FAR become worse. We also find that the higher the SNR, the better the authentication rate is. In particular, if satisfying a low FAR with the signal power $P = 0.5$, the authentication rate is also low, which implies that a legitimate BD has a high probability to be wrongly authenticated by the AP. Besides, by comparing the ROCs when $D_e = 0.1$ and $D_e = 0.3$, we find that the closer the attacker is to the BD, the higher the risk that the attacker is falsely authenticated by the AP. But we can improve authentication performance by increasing the average signal power at the AP. Theoretically, we can obtain $RSS^b_{T_j} = \kappa P_{t_j} h^2(d)$ without considering environmental noise to calculate the distance with $h^2(d) = RSS^b_{t_j}/\kappa P_{t_j}$, so as to perform authentication with the estimated position of the BD. In Fig. 7, without considering noise (i.e., $n = 0$) and estimation error, we also illustrate the ideal performance with the estimation equations (1)-(5) in Fig. 7. As observed, the ideal performance is better than the one with noises even with the same D_e, demonstrating an inevitable estimation error makes performance slightly declined.

Fig. 8 presents the ROC performance under different numbers of APs and different densities of attackers around the BD. From our numerical results, we can see that increasing the number of APs is able to improve authentication performance, as each AP estimates the locations of both the BD and the attackers and then sends them to the AP-1 to reduce the estimation error of their locations. Besides, Fig. 8 also compares ROC performance with different number of attackers that randomly locate around the BD within a circle with a radius of 0.5m. In Fig. 8, we can see that BCAuth achieves an authentication rate $>$ 90% against FAR $<$ 7% under $num_e = 6$. Correspondingly, the larger the attacker number is, the lower the ROC performance. Because it increases the probability that the attackers locate closer to the BD, which could cause high FAR, as shown in Fig. 7.

We exploit two location prediction algorithms, a physics-based (PH-based) approach and a RSS fingerprinting-based (FP-based) approach, to predict the location of the mobile BD in the next time slot. Then, under different thresholds, the predicted locations are compared with the estimated location calculated with the RSS and AoA information for authentication. Fig. 9 shows the authentication rate under different velocities of the moving BD. With the increase of the velocity, which induces the difficulty of location prediction and increases prediction deviation, the authenticate rate degrades apparently. But the authentication rate can be improved by setting a loose threshold. Besides, compared to the PH-based approach, the FP-based approach provides a higher authentication rate. However, the FP-based approach needs to collect all RSS values and historical trajectory information to build a fingerprint map, which incurs considerable cost and complexity, especially at the BD. On the contrary, the PH-based approach only stores location information and motion states in the past period of time to predict the next location with comparatively low complexity. It has no additional requirement on BDs.

In addition, we analyze the ROC performance of the moving BD followed by a moving attacker as shown in Fig. 5. As shown in Fig. 9, the increase of moving velocity induces a

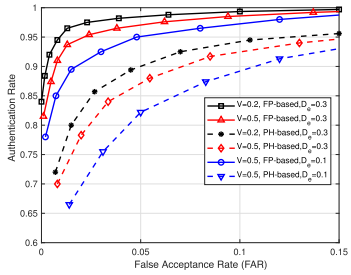


Fig. 10. ROC under different distances, velocities and methods.

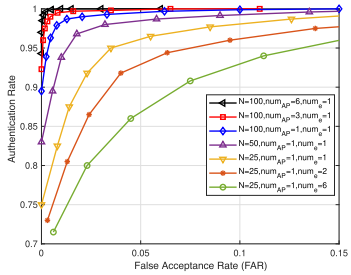


Fig. 11. ROC under different numbers of locations and APs, and attackers.

large prediction deviation, making authentication performance worse. From Fig. 10, we can observe that increasing velocity of the BD obviously decreases the ROC performance, including the decline of authentication rate and the rise of FAR. Besides, we compare the ROC performance under different distances D_e between the BD and the attacker. The smaller the distance D_e is, the worse the ROC performance is, which implies that the attacker is more likely to be falsely accepted by the AP. The same as the authentication rate, the ROC performance of the PH-based approach is inferior to the FP-based approach even with comparatively lower velocity and larger distance D_e.

2) Performance of Re-Authentication: Fig. 11 presents the ROC performance under different parameters for clustering-based authentication in the static BD case. We observe that in the single AP case with an attacker, when the number of location values is larger than 50, all authentication rates are above 94% and the FARs are larger than 1%. Furthermore, with 100 locations, BCAuth can achieve 89.8% authentication rate even with 0% FAR. The performance with only 25 locations is worse than the 100 locations and 50 locations, and its authentication rate is above 90% but the FAR is below 2%. These results show that the server can well authenticate the BD with enough location values by exploiting the clustering-based method. However, accumulating enough location values at the server takes time, which could cause a delay on potential attack detection and BD re-authentication. Note that the locations values are clustered into 2 clusters to perform authentication, the dimension of the RSS value is 1, which does not affect the performance of the re-authentication scheme.

Furthermore, we study whether increasing the number of APs improves ROC performance. Every AP receives the signals backscattered from the BD when BD performs communications with its main AP via backscattering, i.e., AP-1. Every AP measures the features of the backscattered signals and send them to the server. The server estimates the corresponding

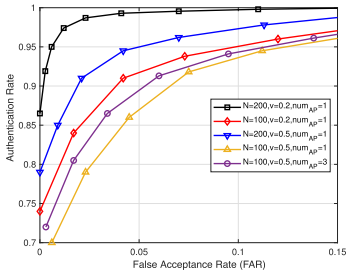


Fig. 12. ROC under different numbers of locations and APs, and velocities.

locations of BD based on these feedback features and then jointly considers the locations reported from AP-1 to perform clustering-based re-authentication. From Fig. 11, we can observe that the authentication rate increases with decreased FAR when the number of APs is increased. Particularly, when the number of locations is below 100, the authentication rate increases from 89% to 93% and further to 95% with FAR 0% when the number of APs increases from 1 to 3 and to 6. This result suggests a practical deployment solution with multiple APs for BD authentication in a real BC systems, e.g., warehouse systems.

Besides, we analyze the impact of the attacker number on the ROC performance as shown in Fig. 11. We select different numbers of attackers in a circle with the BD as its center and a radius of 0.5 meter. As the number of attackers increases with more attackers distributed in the circle, the ROC performance of the clustering-based authentication decreases. But the cost of attacking also increases since it becomes easier to discover the attackers. Moreover, we can deploy more APs to enhance the ROC performance.

In the mobile BD case with a following attacker, we study the ROC performance under different location values, velocities and AP amounts, as shown in Fig. 12. Firstly, we observe that the ROC performance with a large number of location values outperforms the one with a small number of location values. This is because more location information can be utilized to construct two location trajectories V_1 and V_2. With longer trajectories V_1 and V_2, the similarity can be calculated more accurately, so as to improve the ROC performance with a higher authentication rate and lower FAR. However, similar to the performance of mobile BD in preemptive authentication, ROC performance decreases when the velocity of the BD increases. For example, with the same FAR 0%, the authentication rate decreases from 86% to 79% when the velocity of the BD increases from 0.2m/s to 0.5m/s. But its ROC performance can be improved by increasing the number of APs. With the same location amount $N = 100$ and velocity $v = 0.5$m/s, deploying two additional APs can enhance the authentication rate of the legitimate BD and the true rejection of the attacker, i.e., the authentication rate increases from 87% to 89.5% with the same FAR 5% in Fig. 12.

3) Performance of Reciprocal Channel Variation-Based Authentication: For authenticating a static BD using the reciprocal channel variation-based (RSS-based) authentication, we collect RSS values during the initialization phase and then calculate the correlation coefficient between the AP's RSS list and the BD's RSS list as the initial threshold τ_{rss}. For

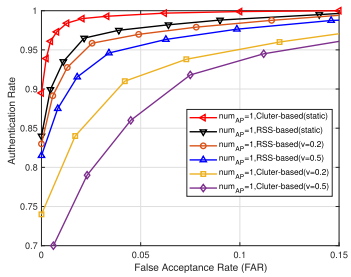


Fig. 13. ROC of re-authentication with different methods (AP=1).

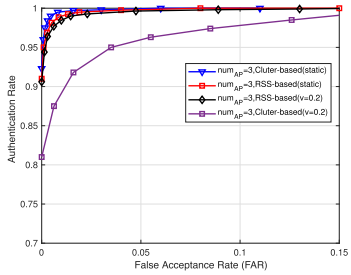


Fig. 14. ROC of re-authentication with different methods (AP=3).

authenticating a mobile BD, after obtaining the correlation coefficient as the threshold τ_{rss} in the initialization phase, the RSS values collected by both the AP and the BD for an underlying round of authentication are used to update the threshold after each successful authentication.

In our simulation, we collect 100 sets of RSS and AoA of the backscattered signals from 100 backscattered message packets from the BD and estimate the corresponding location values of packet sources in both the static BD case and the mobile BD case with $v = 0.5$m/s. These location values are used to authenticate the BD regarding its 100 message packets at one time with the clustering-based authentication method. Besides, we assume the BD can measure the RSS of the downlink signal for these 100 packets and send them to the server, in both static and mobile cases. The server performs the RSS-based authentication by comparing BD's 100 RSS values with AP's 100 RSS values. The initial threshold is set with the correlation coefficient calculated with two RSS lists with 50 values in the initialization phase. The threshold is changed to obtain an ROC curve with different authenticate rates and FARs. Fig. 13 shows the ROCs of the two methods in the static and mobile BD cases. We can see that the ROC performance of the RSS-based method is only slightly inferior to that of the clustering-based method in the static BD case. But the ROC performance of the RSS-based method are much better than that of the clustering-based method in the mobile BD case. Therefore, the RSS-based method is suitable in general when we do not know whether the BD is static or mobile. There is no doubt that the ROC performance of the clustering-based method in the mobile BD case is much worse than that in the static BD case, and it is getting worse with the increase of BD velocity. There is similar impact of the velocity on the ROC performance of the RSS-based method, i.e., the ROC performance decreases when the velocity of a mobile BD increases. However, its ROC performance only suffers from a slight decline as shown in Fig. 13. For example, under the same FAR 0%, the authentication rate reduces from 84% to 83% and further to 81.5% when the velocity of the BD increases from 0m/s to 0.2 m/s and to 0.5m/s.

TABLE II
PERFORMANCE OF DETERMINING THE NUMBER OF ATTACKERS

Number of attackers	1	2	4	6
Hit rate	99.89%	99.46%	89.88%	80.89%
Precision	99.65%	91.59%	85.88%	94.49%
F-measure	98.40%	94.37%	87.58%	87.89%

As tested already, increasing the number of APs can enhance the ROC performance of the clustering-based method. We also deploy 2 additional APs to communicate with BD via backscattering. Likewise, each AP communicates with the BD and measure related RSS values. Meanwhile, the BD measures the RSS values of the downlink signals from the corresponding APs. Fig. 14 shows that the ROC performance with three APs gets better compared to the ROC performance with one AP in Fig. 13. Specifically, under the FAR 0%, the authentication rate of the RSS-based method in the case of static BD is 90% with three APs compared to 84% with one AP.

4) Performance of Attacker Amount Detection: Table II provides the experimental results of Hit rate, Precision, and F-measure when the attacker number $i = \{1, 2, 4, 6\}$. E.g., $i = 2$ means that there are two attackers to masquerade a legitimate BD by injecting illegal messages. We carried out our test for 1000 times in each case and exploited the SILENCE mechanism to determine the amount of attackers and their locations. There is no doubt that when the number of attackers equals to 1, i.e., one attacker masquerades the same identity of BD in the BC system, the SILENCE mechanism achieves the highest Hit Rate, above 99%, and the highest F-measure, over 98%. Likewise, we found that BCAuth also provides good performance with 99.46% Hit Rate and 94.97% F-measure when there are 2 attackers. In the case of 6 attackers, BCAuth can also provide a good Precision with 94.49%, which indicates that the detection of attacker amount is comparatively accurate, but it only achieves 80.89% Hit Rate. But, the Precision in the case of 4 attackers is lower than that of other cases. This is because BCAuth could make a mistake to claim 4 attackers in the cases of 2 and 6 attackers. From our results, we can also see that the Hit Rate reduces with the increase of the attacker amount. But the more attackers, the more likely for them to be exposed to the physical world.

Based on the qualitative comparison between BCAuth and other related works. We found that BCAuth is the first physical layer BD authentication scheme that can effectively enhance authentication security by supporting BD mobility, tracing the locations of attackers and identifying their number, as well as providing conditional mutual authentication if BD is capable of measuring RSS. It can detect not only impersonation/spoofing attacks, but also signal counterfeit, replay and relay attacks, showing great advance on attack resistance. With our best efforts, we compare its authentication rate with existing works [17], [24], both of which are applied in RFID systems. We set parameters in our simulation as distance $d = 3$m between the BD and the AP, signal power $P = 1$, the threshold $\lambda = 0.06$ in the static BD case, and the threshold

TABLE III
COMPARISON OF AUTHENTICATION RATE

	[17]	[24]	BCAuth		
			Pre-auth	Re-auth (cluster-based)	Re-auth (RSS-based)
Static	98.7%	95%	85%	94%	90%
Mobile	36.24%	90%	80%	86%	88%

$\tau = 0.03$ (Cluster-based) and $\tau_{rss} = 0.04$ (RSS-based) in the mobile BD case. From Table III, we can see the authenticate rates of three schemes in both static and mobile BD cases. Although the scheme in [24] has a bit better authentication rate than BCAuth, it needs a complex signal analyzer to obtain a specific inductive coupling feature. Although the scheme in [17] performs better than BCAuth in the static scheme, it does not support mobility and needs a signal analyzer to obtain the average baseband power of RF signals. Considering other advanced properties of BCAuth, we think BCAuth has great potential for practical application.

VII. CONCLUSION

This paper proposed BCAuth, a multi-stage BD authentication and attack tracing scheme for offering enhanced authentication security in BC systems with resource-limited static and mobile BDs. It mainly exploits physical layer features, including RSS, AoA and position information for BD authentication by holistically considering the specific characteristics of BDs and BC. Besides, we designed an attack tracing method to determine the number of attackers and trace their positions and trajectories. We analyzed the security of BCAuth and validated its feasibility and performance through numerical simulations with various experimental settings. The results show that BCAuth can achieve desirable performance in the tested scenarios and is capable of enhancing high security without introducing significant overhead. In the future, we will continue our research by prototyping BCAuth with real BD hardware to further evaluate its performance and improve our design towards practical applications.

REFERENCES

[1] W. Liu, K. Huang, X. Zhou, and S. Durrani, "Next generation backscatter communication: Systems, techniques, and applications," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, p. 69, Dec. 2019, doi: 10.1186/s13638-019-1391-7.

[2] P. Wang, N. Wang, M. Dabaghchian, K. Zeng, and Z. Yan, "Optimal resource allocation for secure multi-user wireless powered backscatter communication with artificial noise," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2019, doi: 10.1109/INFOCOM.2019.8737501.

[3] V. Talla, J. Smith, and S. Gollakota, "Advances and open problems in backscatter networking," 2020, *arXiv:2011.03242*.

[4] Y. Xu, B. Gu, and D. Li, "Robust energy-efficient optimization for secure wireless-powered backscatter communications with a non-linear EH model," *IEEE Commun. Lett.*, vol. 25, no. 10, pp. 3209–3213, Oct. 2021.

[5] Y. Xu, B. Gu, R. Q. Hu, D. Li, and H. Zhang, "Joint computation offloading and radio resource allocation in MEC-based wireless-powered backscatter communication networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 6200–6205, Jun. 2021.

[6] Y. Xu, Z. Qin, G. Gui, H. Gacanin, H. Sari, and F. Adachi, "Energy efficiency maximization in NOMA enabled backscatter communications with QoS guarantee," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 353–357, Feb. 2020.

[7] D. Ma, G. Lan, M. Hassan, W. Hu, and S. K. Das, "Sensing, computing, and communications for energy harvesting IoTs: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 2, pp. 1222–1250, Dec. 2019.

[8] Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 2, pp. 668–695, Feb. 2021.

[9] P. Wang, L. Jiao, K. Zeng, and Z. Yan, "Physical layer key generation between backscatter devices over ambient RF signals," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, May 2021, pp. 1–10, doi: 10.1109/INFOCOM42981.2021.9488885.

[10] A. Mitrokotsa, M. R. Rieback, and A. S. Tanenbaum, "Classifying RFID attacks and defenses," *Inf. Syst. Frontiers*, vol. 12, no. 5, pp. 491–505, Nov. 2010.

[11] X. Li *et al.*, "Hardware impaired ambient backscatter NOMA systems: Reliability and security," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2723–2736, Apr. 2021.

[12] M. G. Samaila, M. Neto, D. A. B. Fernandes, M. M. Freire, and P. R. M. Inácio, "Challenges of securing Internet of Things devices: A survey," *Secur. Privacy*, vol. 1, no. 2, p. e20, Mar. 2018.

[13] G. Tsudik, "YA-TRAP: Yet another trivial RFID authentication protocol," in *Proc. 4th Annu. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PERCOMW06)*, 2006, p. 643.

[14] G. Wang, C. Qian, H. Cai, J. Han, H. Ding, and J. Zhao, "Towards replay-resilient RFID authentication," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, vol. 2, Oct. 2018, pp. 385–399.

[15] K. Zeng, K. Govindan, D. Wu, and P. Mohapatra, "Identity-based attack detection in mobile wireless networks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1880–1888.

[16] J. Zhang, S. Rajendran, Z. Sun, R. Woods, and L. Hanzo, "Physical layer security for the Internet of Things: Authentication and key generation," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 92–98, Oct. 2019.

[17] D. Zanetti, B. Danev, and S. Apkun, "Physical-layer identification of UHF RFID tags," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw. (MobiCom)*, 2010, pp. 353–364.

[18] N. Xie, Q. Zhang, J. Chen, and H. Tan, "Privacy-preserving physical-layer authentication for non-orthogonal multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1371–1385, Apr. 2022.

[19] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks [security and privacy in emerging wireless networks]," *IEEE Wireless Commun.*, vol. 17, no. 5, pp. 56–62, Oct. 2010.

[20] Q. Wang, "A novel physical layer assisted authentication scheme for mobile wireless sensor networks," *Sensors*, vol. 17, no. 2, p. 289, Feb. 2017.

[21] X. Qiu, X. Sun, and M. Hayes, "Enhanced security authentication based on convolutional-LSTM networks," *Sensors*, vol. 21, no. 16, p. 5379, Aug. 2021.

[22] J. Han *et al.*, "GenePrint: Generic and accurate physical-layer identification for UHF RFID tags," *IEEE/ACM Trans. Netw.*, vol. 24, no. 2, pp. 846–858, Apr. 2015.

[23] A. Mehmood, W. Aman, M. M. U. Rahman, M. A. Imran, and Q. H. Abbasi, "Preventing identity attacks in RFID backscatter communication systems: A physical-layer approach," in *Proc. Int. Conf. U.K.-China Emerg. Technol. (UCET)*, Aug. 2020, pp. 1–5.

[24] G. Wang *et al.*, "Towards replay-resilient RFID authentication," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw.*, Oct. 2018, pp. 385–399.

[25] B. Danev, T. S. Heydt-Benjamin, and S. Capkun, "Physical-layer identification of RFID devices," in *Proc. USENIX Secur. Symp.*, 2009, pp. 199–214.

[26] K. Zeng, K. Govindan, and P. Mohapatra, "Non-cryptographic authentication and identification in wireless networks [security and privacy in emerging wireless networks]," *IEEE Wireless Commun.*, vol. 17, no. 5, pp. 56–62, Oct. 2010.

[27] Z. Luo, W. Wang, Q. Huang, T. Jiang, and Q. Zhang, "Securing IoT devices by exploiting backscatter propagation signatures," *IEEE Trans. Mobile Comput.*, early access, Jun. 2, 2021, doi: 10.1109/TMC.2021.3084754.

[28] Y. Li, S. Liu, Z. Yan, and R. H. Deng, "Secure 5G positioning with truth discovery, attack detection and tracing," *IEEE Internet Things J.*, early access, Jun. 14, 2021, doi: 10.1109/JIOT.2021.3088852.

[29] H. Ostaffe, "RF-based wireless charging and energy harvesting enables new applications and improves product design," Powercast Corp., Pittsburgh, PA, USA, Tech. Rep., 2017.

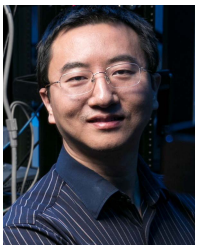
[30] H. A. Song, B. Hooi, M. Jereminov, A. Pandey, L. Pileggi, and C. Faloutsos, "Powercast: Mining and forecasting power grid sequences," in *Proc. Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Cham, Switzerland: Springer, 2017, pp. 606–621.

[31] A. Juels, "RFID security and privacy: A research survey," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 2, pp. 381–394, Feb. 2006.

[32] P. Kitsos and Y. Zhang, *RFID Security*, vol. 233. Cham, Switzerland: Springer, 2008.
[33] E. Global, "EPC radio-frequency identity protocols class-1 generation-2 UHF RFID protocol for communications at 860 MHz–960 MHz," *Version*, vol. 1, p. 23, Dec. 2008.
[34] Y. Komori, K. Sakai, and S. Fukumoto, "Fast and secure tag authentication in large-scale RFID systems using skip graphs," *Comput. Commun.*, vol. 116, pp. 77–89, Jan. 2018.
[35] G. Tsudik, "YA-TRAP: Yet another trivial RFID authentication protocol," in *Proc. 4th Annu. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PERCOMW)*, 2006, p. 4.
[36] H.-Y. Chien, "SASI: A new ultralightweight RFID authentication protocol providing strong authentication and strong integrity," *IEEE Trans. Dependable Secure Comput.*, vol. 4, no. 4, pp. 337–340, Oct./Dec. 2007.
[37] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, "Practical user authentication leveraging channel state information (CSI)," in *Proc. 9th ACM Symp. Inf., Comput. Commun. Secur.*, Jun. 2014, pp. 389–400.
[38] J. Liu and X. Wang, "Physical layer authentication enhancement using two-dimensional channel quantization," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 4171–4182, Jun. 2016.
[39] P. L. Yu, J. S. Baras, and B. M. Sadler, "Physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 38–51, Mar. 2008.
[40] Y. Huang, W. Wang, T. Jiang, and Q. Zhang, "Detecting colluding sybil attackers in robotic networks using backscatters," *IEEE/ACM Trans. Netw.*, vol. 29, no. 2, pp. 793–804, Apr. 2021.
[41] K. Fan, S. Sun, Z. Yan, Q. Pan, H. Li, and Y. Yang, "A blockchain-based clock synchronization scheme in IoT," *Future Gener. Comput. Syst.*, vol. 101, pp. 524–533, Dec. 2019.
[42] Z. Yan, L. Peng, W. Feng, and L. T. Yang, "Social-chain: Decentralized trust evaluation based on blockchain in pervasive social networking," *ACM Trans. Internet Technol.*, vol. 21, no. 1, pp. 1–28, Feb. 2021.
[43] Z. Li, Y. Liu, K. G. Shin, J. Liu, and Z. Yan, "Interference steering to manage interference in IoT," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10458–10471, Dec. 2019.
[44] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.
[45] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, 2020.
[46] K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, "Enhanced fingerprinting and trajectory prediction for IoT localization in smart buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1294–1307, Jul. 2016.
[47] J.-Y. Lee, C.-H. Yoon, H. Park, and J. So, "Analysis of location estimation algorithms for WIFI fingerprint-based indoor localization," in *Proc. 2nd Int. Conf. Softw. Technol.*, vol. 19, 2013, pp. 89–92.
[48] B. F. Baccelli, B. Blaszczyszyn, and P. Muhlethaler, "An Aloha protocol for multihop mobile wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 421–436, Feb. 2006.
[49] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, vol. 344. Hoboken, NJ, USA: Wiley, 2009.
[50] D. Salvatore and D. Reagle, "Schaum's outline of theory and problems of statistics and econometrics," in *Statistical Method Econometric Model* (Schaum's Outline Series), 2nd ed. New York, NY, USA: McGraw-Hill, 2002, p. 328.
[51] J. Yang, Y. Chen, W. Trappe, and J. Cheng, "Detection and localization of multiple spoofing attackers in wireless networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 1, pp. 44–58, Jan. 2012.
[52] R. A. Johnson, I. Miller, and J. E. Freund, *Probability and Statistics for Engineers*. London, U.K.: Pearson, 2000.
[53] K. Calder, *Statistical Inference*. New York, NY, USA: Holt, 1953.
[54] J. R. Smith, *Wirelessly Powered Sensor Networks and Computational RFID*. Cham, Switzerland: Springer, 2013.
[55] S. T. Ali, V. Sivaraman, and D. Ostry, "Eliminating reconciliation cost in secret key generation for body-worn health monitoring devices," *IEEE Trans. Mobile Comput.*, vol. 13, no. 12, pp. 2763–2776, Dec. 2013.
[56] J. Tang *et al.*, "Identity-based attack detection and classification utilizing reciprocal RSS variations in mobile wireless networks," *IEEE Trans. Mobile Comput.*, vol. 21, no. 5, pp. 1657–1671, May 2022.
[57] Y. Chen, J. Yang, W. Trappe, and R. P. Martin, "Detecting and localizing identity-based attacks in wireless and sensor networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2418–2434, Mar. 2010.
[58] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. 1, pp. 53–65, 1987.
[59] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *Statist. Methodol.*, vol. 63, no. 2, pp. 411–423, 2004.
[60] C. Bettstetter, H. Hartenstein, and X. Pérez-Costa, "Stochastic properties of the random waypoint mobility model," *Wireless Netw.*, vol. 10, pp. 555–567, Sep. 2004.
[61] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications With MATLAB*. Hoboken, NJ, USA: Wiley, 2010.
[62] R. Baeza-Yates *et al.*, *Modern Information Retrieval*, vol. 463. New York, NY, USA: ACM Press, 1999.

Pu Wang (Student Member, IEEE) received the Ph.D. degree in cyberspace security from Xidian University in 2021. His research interests are in backscatter communication, wireless information and power transfer, physical layer security, and information security in the Internet of Things.

Zheng Yan (Senior Member, IEEE) received the Doctor of Science degree in electrical engineering technology from the Helsinki University of Technology, Helsinki, Finland, in 2007. She is currently a Professor with the School of Cyber Engineering, Xidian University, Xi'an, China, and also a Visiting Professor and an Academy Fellow at Aalto University, Helsinki. Her research interests are in trust, security, privacy, and data analytics. Her recent achieved awards include the 2021 N^2Women: Stars in Computer Networking and Communications, the Nokia Distinguished Inventor Award, the Aalto ELEC Impact Award, the Best Journal Paper Award issued by the IEEE Communication Society Technical Committee on Big Data, and the Outstanding Associate Editor of 2017 and 2018 for IEEE ACCESS. She has served as the General Chair or the Program Chair for numerous international conferences, including IEEE TrustCom 2015 and IFIP Networking 2021. She is the Founding Steering Committee Co-Chair of IEEE Blockchain Conference. She is an Area Editor or an Associate Editor of IEEE INTERNET OF THINGS JOURNAL, *Information Fusion*, *Information Sciences*, *IEEE Network* Magazine, IEEE ACCESS, and *Journal of Network and Computer Applications*.



Kai Zeng (Member, IEEE) received the Ph.D. degree in electrical and computer engineering from Worcester Polytechnic Institute (WPI) in 2008. He was a Post-Doctoral Scholar with the Department of Computer Science, University of California at Davis (UCD), from 2008 to 2011. He has worked with the Department of Computer and Information Science, University of Michigan–Dearborn, as an Assistant Professor, from 2011 to 2014. He is currently an Associate Professor with the Department of Electrical and Computer Engineering, the Department of Computer Science, and the Center for Secure Information Systems, George Mason University. His current research interests include cyber-physical system security and privacy, physical layer security, network forensics, and cognitive radio networks. He was a recipient of the U.S. National Science Foundation Faculty Early Career Development (CAREER) Award in 2012. He has won the Excellence in Postdoctoral Research Award at UCD in 2011 and the Sigma Xi Outstanding Ph.D. Dissertation Award at WPI in 2008. He is an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.