Abstract—Resorting to the exploitation of physical attributes, physical-layer authentication (PLA) is a promising technology to supplement and enhance current cryptography-based security mechanisms in wireless communications. In the fifth-generation (5G) communications, many disruptive technologies spring up, such as millimeter-wave communication (mmWave), massive multiple-input and multiple-output (MIMO) and non-orthogonal multiple-access (NOMA). PLA schemes in 5G networks are facing challenges while exposed to opportunities at the same time. This article seeks to identify the critical technology gaps as well as the feasible enablers in terms of the unique characteristics of 5G networks. The investigation consists of four hierarchical parts. In the first part, existing PLA schemes are reviewed, and the corresponding challenges are discussed in the next part. In the third part, we investigate potential enablers based on the unique characteristics of 5G networks and provide three corresponding case studies. Furthermore, open problems and research directions on PLA for 5G and beyond are discussed in the last part, including waveform design, feature learning, mobile users, and terahertz communications.

I. INTRODUCTION

Recent years have witnessed a spurt of progress in the fifth-generation (5G) communications. With the adoption of emerging 5G wireless communication technologies, such as millimeter wave (mmWave) communication, massive multiple-input and multiple-output (MIMO) and non-orthogonal access (NOMA), 5G networks are expected to provide reliable wireless communication with ultra-low delay and extremely high throughput. Meanwhile, the capability to ensure security, trust, identity, and privacy is of the utmost importance for 5G communications. In this regard, as a supplement and enhancement strategy, physical-layer authentication (PLA) is emerging as an effective approach to provide a high-security and low-complexity security solution utilizing the unique properties at the physical-layer in 5G networks [1], [2].

Generally, existing PLA schemes can be categorized into two types: radio-frequency (RF)/hardware-based PLA and location/channel-based PLA. RF/hardware-based PLA explores the imperfect features extracted from hardware or modulation to form an RF fingerprint that can identify devices [3], and location/channel-based PLA can be used to counter identity spoofing attacks based on location/channel features [4], [5]. However, these existing PLA schemes have corresponding disadvantages. The extracted features in the RF/hardware-based PLA usually require a feature extracting device, e.g., a signal analyzer, which implies a high deployment cost or overhead. Location/channel-based PLA methods are vulnerable to co-located spoofing attacks, where the attacker is very close to the victim. Furthermore, the emerging radio security threats in 5G networks call for new effective countermeasures. For instance, pilot contamination attacks could corrupt the channel estimation using the same pilot signal as the legitimate users in massive MIMO and NOMA [6], [7].

For these issues, the unique characteristics of 5G wireless communications may suggest new solutions. Take the beam pattern in mmWave massive MIMO communications as an example. Due to the imperfection of electronic circuits and the different packaging and placement of the antenna, the antenna arrays in different transmitters are not exactly identical and form a unique beam for each individual transmitter. This uniqueness of beam patterns can be used for location-based PLA methods against co-located spoofing attacks. Another example lies in the sector-level-sweep (SLS) signal-to-noise-ratio (SNR) traces obtained in the 5G communication protocol (e.g., IEEE 802.11ad protocol). These SLS SNR traces are sensitive to different devices and can be employed to achieve a new RF-based PLA scheme free from signal analyzers. Furthermore, resorting to the high directivity of mmWave communications, a new PLA scheme based on the sparsity of the channel virtual representation can be exploited to detect pilot contamination attacks [7].

Based on these observations, this article aims to investigate these challenges and opportunities of PLA schemes in 5G communications. In particular, existing PLA schemes are reviewed and the corresponding challenges are discussed. For these challenges, we introduce the potential solutions based on the unique characteristics of 5G communication from three aspects: 5G antenna characteristics, 5G communication protocols, and 5G signal processing technologies. The corresponding case studies are provided, involving beam patterns, SLS SNR traces, and channel virtual representations. Furthermore, open research problems and future research topics on the PLA for 5G and beyond are discussed. We hope this article can help to stimulate further research in this area.
In the remainder of this article, we will review existing PLA schemes in Section II and discuss the challenges in Section III. In Section IV, we introduce the potential benefits from 5G communications and provide three corresponding case studies. Section V discusses open topics of PLA in 5G and beyond. Conclusions are given in Section VI.

II. EXISTING PLA SCHEMES

In general, PLA approaches can be broadly classified into two categories: RF/hardware-based and location/channel-based PLA.

1) RF/hardware-based PLA: Based on the fact that different wireless transceivers emit RF signals with distinctive features/patterns in analog and modulation domains, the goal of the RF/hardware-based PLA is to extract these distinctive features/patterns to recognize the corresponding devices. In the analog domain, unique RF signal emission patterns can be observed during the transition of a transmitter [8]. In the modulation domain, the proposed features for PLA include I/Q offset, phase and magnitude errors, and power spectral density (PSD) [3]. In general, RF/hardware-based PLA methods need a high-end signal analyzer (e.g., oscilloscope and spectrum analyzer) or software-defined radio device to extract these features. Then, white-list based strategies are used to achieve device recognition, where legitimate users need to set up a database of the feature space of legitimate devices. Here, the prior information of legitimate devices is required and multiple-class classification machine-learning algorithms are the commonly employed detection strategies. This PLA technique offers another means for authenticating a device analogous to the fingerprinting method for authenticating a human. It can be combined with cryptographic mechanisms to enhance authentication security strength, e.g., enabling multi-factor authentication.

2) Channel/location-based PLA: Different from RF/hardware-based PLA, channel/location-based PLA does not need extra feature extracting equipment since the exploited physical-layer features are readily extractable from the off-the-shelf devices. Generally, channel/location-based PLA schemes depend on the fact that devices at different locations present different channel/location-based feature profiles. The popular features employed in these PLA schemes include received signal strength (RSS), channel frequency response (CFR), and channel impulse response (CIR) [4], [5]. For these channel/location-based PLA methods, legitimate users are free from collecting the prior information of legitimate devices, and the detection algorithms based on similarity measurement between location or channel features are commonly used. This PLA strategy can enhance the physical-layer protection of off-the-shelf devices against masquerade attacks, in which multiple devices use the same identifier (e.g., spoofing attacks) or a single device claims multiple identifiers (e.g., Sybil attacks).

Different from these existing PLA schemes, in this work, we focus on PLA methods adapted in 5G wireless networks and aim to exploit the unique characteristics of 5G communications to achieve desirable PLA performance. To this end, three salient features of 5G communications will be exploited, involving the uniqueness of beam patterns, the distinguishability of SLS SNR traces, and the sparsity of channel virtual representation. These unique features can be utilized to substantially boost PLA performance in 5G networks.

III. PROBLEM STATEMENT

To achieve a desirable PLA scheme for 5G networks, we need to tackle the following two main issues:

• Minimize extra hardware deployment overhead. RF/hardware-based PLA schemes have high requirements for communication environment stability and fingerprinting extraction. In practice, the difference of the selected RF feature of different devices is usually small and could be corrupted by both noise and interference. As a result, the extracted features in the RF/hardware-based schemes usually require a high-end feature extracting device to improve the accuracy of estimating features. This implies a high cost or overhead for the communication systems.

• Counter co-located attackers. Current channel/location-based PLA schemes work only when the environment is relatively stable and the attacker is at a different location from the legitimate user. They are hard to handle the co-located attacks, where the attacker is very close to the victim. However, under 5G communication scenarios, co-located attacks are one of the most common attack strategies. In mmWave massive MIMO 5G communications, beamforming is a commonly employed physical layer technique, which can form a space region that improves communication performance. In this case, attackers prefer to approach legitimate users to obtain the channel gains. As a result, co-located attacks are more likely to appear in 5G networks, and that will challenge existing channel/location-based PLA schemes.

IV. POTENTIAL OPPORTUNITIES AND CASE STUDIES

To tackle these challenges above, in this section, we will discuss potential benefits for PLA from 5G communications and provide the corresponding case studies.

A. Potential opportunities

We discuss the potential opportunities from three specific cases: 5G antenna characteristics, 5G communication protocols and 5G signal processing technologies.

• 5G antenna characteristics. Thanks to 5G key wireless communication technologies, such as massive MIMO and mmWave, new physical layer features emerge. The beam pattern of antennas based on massive MIMO and mmWave is a case in point. Due to the imperfection of electronic circuits, even for the same type of devices
from the same manufacturer, the antenna arrays are hard to be exactly identical. Meanwhile, the packaging and placement of the antenna inside a device also can affect the radiation characteristics of the beam pattern. As a result, the beam pattern of an individual device in 5G networks is unique, and it could be exploited by channel/location-based PLA to curb co-located spoofing attacks. A corresponding case study will be given in Section IV-B1.

- **5G communication protocols.** In order to adapt to the characteristics of 5G communications, 5G communication protocols have many characteristics that are different from traditional communication protocols. Take the IEEE 802.11ad protocol for 60GHz mmWave WLAN as an example. In this protocol, transceivers have to perform a sector/beam level sweep to find a beam-pattern pairing with the best channel gain. The SNR traces in the SLS show a significant distinguishability between different devices. More importantly, these SLS SNR traces are stable and easy to obtain. That is, this feature suggests a new RF/hardware-based PLA method without high-end feature extracting devices. A corresponding case study will be provided in Section IV-B2.

- **5G signal processing technologies.** With the rise of 5G communications, new signal processing technologies are burgeoning. For example, a 5G signal processing technique is introduced to combine the antenna space and beamspace through spatial Fourier transform, which is called channel virtual beamspace representation [9]. This virtual representation is served as an effective signal processing tool to reduce the constraints on the hardware design in the mmWave 5G communication system [9]. Meanwhile, thanks to the excellent spatial discrimination, this channel virtual representation is sensitive to the different transmitters and the number of transmitters. This unique feature may gain an edge over traditional channel features in channel/location-based PLA schemes for protecting 5G networks. The case study based on the channel virtual representation against pilot contamination attacks in 5G NOMA networks is introduced in Section IV-B3.

### B. Case studies

To illustrate these opportunities, three corresponding case studies are provided, where the employed 5G communication technologies include beam pattern tracking, sector/beam sweeping, and channel virtual representation.

1) **Beam pattern-based PLA:** Identity spoofing attacks pose one of the most serious threats to wireless networks, where the attacker can pretend to be a legitimate user using a faked identity, such as the media access control (MAC) address and IP address. Channel-based PLA is a promising technology to counter this threat, however, most existing channel-based PLA methods cannot tackle the co-located spoofing attacks. To address this issue, the uniqueness of the beam pattern in mmWave massive MIMO 5G communications may suggest a cure.

In mmWave massive MIMO 5G communications, beamforming techniques can significantly improve communication performance. For an individual device, the antenna array is responsible for providing various beam patterns to support the beamforming for different directions. One of the interesting features of the beam pattern in mmWave MIMO is that the beam patterns’ shape is relatively stable for an individual device when the antenna array is fixed. That is, beam patterns do not change drastically over time especially the transmitter’s patterns. An example is given in Fig. 1, which shows two different beam patterns’ SNR at different measurement time in the same environment. The azimuth angle is between $[-90^\circ, +90^\circ]$ and the corresponding SNR is measured every $1.8^\circ$. The time interval between measurements is much longer than the channel coherence time. We can see that different

![Fig. 1. Beam pattern SNR at different azimuth angles (the azimuth angle $\in [-90^\circ, +90^\circ]$ and SNR measured every $1.8^\circ$). (a) is the pattern 1 at the time $T_1$; (b) is the pattern 1 at the time $T_2$; (c) is the pattern 2 at the time $T_3$; (d) is the pattern 2 at the time $T_4$, $T_2 - T_1 \geq 60\text{min}$. (The experimental device is Talon AD7200 router with 36 transmitting and 4 receiving beam patterns.)](image1)

![Fig. 2. Beam pattern based PLA.](image2)
patterns show different SNR curves, while the same patterns present high correlations even over a long measuring time.

Based on this observation, a beam pattern-based PLA scheme can be proposed, as shown in Fig. 2. With beam pattern tracking technologies, the variation of the channel gain can be used to estimate the perturbation of the antenna array [10]. In this case, when we provide a set of perturbation (e.g., $P_1, P_2, ..., P_6$) in the antenna array at Alice (Transmitter), Alice’s antenna array is changed. Since the variation of the channel gain is proportional to the variation of the array gain, when Bob (Receiver) extracts the corresponding channel gains, the changes of the channel gain at Bob can reflect the changes of the antenna array gain at Alice [10]. Thus, the perturbation of the channel gains at Bob will exhibit the characteristics of the beam patterns at Alice. Since the beam pattern of Alice is unique, based on these channel gains under the different angle perturbations, Bob can establish a channel-based PLA to distinguish Alice from other devices even there are co-located attackers.

Remarks: (1) In contrast to existing channel-based PLA schemes, this beam pattern-based PLA method relying on channel gain measurement can effectively counter co-located spoofing attacks; (2) Since the channel gain estimation is a common communication process in massive MIMO, this beam pattern-based PLA does not require complex feature extraction processes or a high-end signal analyzer.

2) Sector/beam sweeping-based scheme: Physical layer fingerprinting is a viable technique to provide authentication services identifying wireless devices. However, for most existing RF fingerprinting schemes, extracting RF/hardware-based features usually requires a high-end signal analyzer or software-defined radio device, which is hard to be supported by off-the-shelf devices. To tackle this challenge, the beam sweeping process in 5G communication protocols offers an opportunity.

According to the IEEE 802.11ad standard for the 60GHz band, both transmitter and receiver have to perform a sector/beam level sweep, i.e., SLS, to find an antenna pattern pair with the best channel gain to initiate a communication link. During this procedure, SNR values corresponding to each Tx-Rx beam pattern pair will be obtained. Therefore, it does not need to change or add extra communication overhead over the current IEEE 802.11ad protocol to obtain the SNR traces of SLS. Fig. 3 shows an example of this phenomenon, where the curves represent SLS SNR traces which are sequences of SNR values corresponding to transmitter beam pattern index. With the same receiving device, the SLS SNR traces of three different devices show different patterns even the transmitters are fixed at the same location. Meanwhile, the curves of the SLS SNR traces from the same transmitting devices show a similar pattern although the absolute values of SNRs may be different.

To evaluate the identification performance of the SLS SNR traces, we conducted a preliminary real-world experiment, where off-the-shelf 802.11ad devices (Talon AD7200 routers) were selected. For the Talon AD7200, there were 36 default transmitting sectors (beam patterns) and 4 semi-omnidirectional receiving sectors. In one SLS, the transmitter sequentially changed the transmitting sectors. For each transmitting sector, 25 frames were transmitted and the corresponding SNR traces were logged. One router was set as an access point (AP), and the other was set as a client. AP collected the SLS SNR traces to identify the clients. In the experiment, we used three different clients at the same location, marked as Device-1, Device-2, and Device-3.

Moreover, to achieve identification, the one-class classification algorithm based on the support vector machine (SVM) is employed. One-class classification machine learning algorithms can achieve the classification task only resorting to positive training samples. This can greatly simplify the requirement of the training samples for the identification of SLS SNR traces. Fig. 4 shows identification results for the
authentication performance. We can see that all of the curves are very close to the upper left corner. That is, the experiment results show a great identification performance, where the detection rates \( P_D \) can reach around 99% even when the false alarm rate \( P_{FA} \) is only 0.01.

**Remarks:** (1) This SLS SNR trace-based PLA is completely compatible with the IEEE 802.11ad protocol without introducing any extra communication overhead. No high-end signal analyzer is needed to obtain the SNR values, which can be directly read out from the off-the-shelf 802.11ad devices; (2) SLS SNR traces have been used to counter spoofing identity attacks in [11] and achieve RF-based fingerprinting in [12].

3) **Based on channel virtual representation:** Pilot contamination attacks are becoming a serious physical layer threat in 5G massive MIMO and NOMA communications, where the attacker can inject the same pilot as the legitimate user’s to control the channel estimation results [7], [13]. To thwart this attacker, the emerging signal processing technique in 5G communication, e.g., the channel virtual representation, may imply a desirable PLA solution.

In 5G communication systems, to alleviate the constraints on the hardware partly due to high frequency and bandwidth communication channels, the channel virtual representation serving as an effective signal processing technology is exploited in the design of mmWave massive MIMO communication systems [9]. Channel virtual representations are to characterize the mmWave channel by fixed virtual receive and transmit direction, where it consists of path gain, virtual angle-of-arrival (AoA) and virtual angle-of-departure (AoD). When increasing the number of antennas, fewer paths of mmWave would contribute to each spatial bin of the virtual AoA and virtual AoD, as shown in Fig. 5. As a result, the sparsity of the channel paths in the channel virtual representation can reflect the number of the transmitters. Under the same communication environment, the sparsity of paths with one transmitter as shown in Fig. 5(a) is significantly higher than that of paths with two transmitters as shown in Fig. 5(b).

Based on this observation, channel virtual representations can be used to enhance channel-based PLA against pilot contamination attacks in 5G NOMA communications [7]. In NOMA communications, pilot contamination attackers could send the same pilot signals as legitimate users to join the NOMA superimposed signals to affect the channel estimation at the receiver. It challenges the traditional pilot contamination attack detection schemes that will raise alarms when more than one transmitter gets detected [13]. If these traditional detection schemes are directly applied in NOMA communications, excessive false alarms will be raised [7]. Based on channel virtual representations, channel-based PLA can provide an effective pilot contamination attack detection even if the attacker is very close to the victim (i.e., co-located attacks).

A preliminary evaluation is illustrated in Fig. 6, where the detection performance of the PLA is presented by the receiver that operates characteristic (ROC) curve, where \( P_D \) denotes the detection rate and \( P_{FA} \) is the false alarm rate. We can see that this PLA scheme has a notable detection performance. For instance, when the false alarm rate is \( P_{FA} = 0.03 \), all of the detection rates can exceed 98%. It implies that the performance of the channel-based PLA schemes can get a significant boost with the help of the channel virtual representations.

**Remarks:** (1) This channel-based PLA based on channel virtual representations can effectively counter the emerging pilot contamination attacks in 5G NOMA communications; (2) Resorting to the 5G signal processing technologies, this PLA scheme does not need extra feature extracting processes and can defend against co-located attacks.

**V. Future trends and open topics**

This section discusses research trends and potential future research topics of PLA in 5G and beyond, involving waveform design, feature learning, mobile users, and terahertz communications.
A. Waveform design for PLA

Waveform design has been used to improve target identification and classification [14]. Therefore, it is possible to design a new dialect waveform of an existing standard, that will increase the probability of correctly identifying the transmitters while preserving communication rates. To achieve this task, waveform synthesis has to search the design space based on the information from test data. Thus, waveform synthesis is a challenging stochastic optimization problem, which involves both the objective function value (i.e., detection rate) and the feasibility of a waveform (i.e., communication rate). For example, a standard OFDM waveform can be generated by setting the corresponding components to reach the communication requirement in an OFDM system. In this process, some methods and algorithms are required to deal with the possible negative factors, such as I/Q imbalance and peak-to-average power ratio (PAPR). Therefore, to enhance the transmitter identification, we can generate variations in the OFDM waveform. These variations can be handled in signal processing at the receiver but may enhance the device identification performance compared with the standard waveform.

B. Feature learning for PLA

In general, for a wireless receiver, the received signal is a mixture of the signal content, device impairment, and other noise like that from the environment and receiver. Thus, it is challenging to find the features that are truly discriminative with good model generalization. To address this problem, machine learning models may suggest a solution. There are two basic ideas: (1) Utilize machine learning algorithms to find a combined feature. This combined feature may be an integration of several existing features with different weights, and machine learning algorithms are exploited to search the best composite mode; (2) Generate a new feature based on deep learning models. The feature extraction and recognition in PLA can be seen as a black-box operation, just like the deep learning process in deep neural networks. The generated feature can significantly enhance PLA schemes, but it is hard to explain the corresponding specific physical meaning. Taking advantage of the burgeoning machine learning technique to accelerate PLA is an interesting and meaningful topic.

C. PLA for mobile users

Mobile and dynamic communications are prevalent in 5G wireless networks, such as the unmanned aerial vehicle (UAV) and Internet of Vehicle (IoV) communications. Mobile users will have to frequently switch between different base stations or access points, which results in frequent authentication handover. This case becomes even more challenging in heterogeneous networks (HetNets). For this issue, one feasible solution lies in developing physical layer features that are not sensitive to the mobile environment. As we introduced earlier in Section V-B, with the help of the booming machine learning techniques, we can combine existing features or generate new features to find the features which can be applicable to mobile scenarios. PLA schemes that can handle the high mobility in 5G networks are highly expected.

D. PLA in terahertz communication

Terahertz communications have significant potential for 5G networks and beyond, and have attracted great interest from academia and industry. Possessing THz frequency bands, this technique can support terabit-per-second (Tb/s) wireless indoor communications. However, although its highly directional and narrow beams, terahertz communications could be corrupted by an eavesdropper who can intercept signals in line-of-sight transmissions [15]. Few works have yet been reported on the security design in Terahertz communications. A new area of research on PLA for terahertz communications is starting to form.

VI. Conclusion

In this paper, we reviewed existing PLA schemes and pointed out the challenges and opportunities for PLA in 5G networks. The potential solutions based on the unique characteristics of 5G communications were introduced and the corresponding case studies were provided to illustrate these benefits. We also discussed future topics and research trends related to PLA in 5G and beyond.

ACKNOWLEDGMENT

This work was supported in part by the Commonwealth Cyber Initiative (CCI) and its Northern Virginia (NOVA) Node, an investment in the advancement of cyber R&D, innovation and workforce development (For more information about CCI, visit cyberinitiative.org).

REFERENCES


