Light-Weight Physical Layer Enhanced Security Schemes for 5G Wireless Networks

Jie Tang, Hong Wen, Kai Zeng, Run-fa Liao, Fei Pan, and Lin Hu

Abstract

Due to the broadcast nature of wireless radio propagation channels, 5G wireless networks face serious security threats. Security mechanisms that leverage physical layer characteristics have been considered as potential complements to enhance 5G wireless network security. This article proposes physical layer security enhancements to defend against security attacks from the three most significant aspects: wireless secure communication, physical layer assisted authentication and secret key distribution. By integrating physical layer security with novel techniques and application scenarios in 5G communication, we elaborate several promising strategies, including massive MIMO beamforming with security code communication, self-adaptive mobile and cooperative secure communication with dynamic channel prediction, physical layer assisted authentication, malicious node detection and key generation of massive MIMO millimeter wave system. Research opportunities and future challenges are further discussed.

Introduction

Fifth generation (5G) networks [1] promise to meet the continuously increasing demands of wireless mobile applications. They intend to provide an ultra-high data rate with fast access, which is 10 times faster than 4G networks. Furthermore, a 5G network is a heterogeneous network that consists of various types of sub-networks and supports various applications, such as self-driving vehicles, smart cities, military and government applications, and so on. Thus in a 5G network, various types of sub-networks with various transmit power levels, coverage areas, and access protocols are deployed to form a multi-tier hierarchical network architecture. Meanwhile, a massive amount of resource-constrained IoT (Internet-of-Things) devices and sensors will be connected by 5G wireless networks [2]. Traditional security protocols based on conventional cryptographic techniques tend to introduce heavy computation or key management overhead, which are not optimal for securing communication among resource-constrained IoT devices.

Recently, there has been increasing interest in enhancing security for 5G wireless networks with various physical layer security strategies [3–11]. By intelligently leveraging the properties of the wireless physical medium, physical layer security has been identified as a promising strategy that can effectively provide security enhancement for 5G networks. Different from traditional cryptographic approaches relying on computational complexity, physical layer security relies on unique physical layer properties such as wireless channel state information (CSI), receive signal strength (RSS), hardware fingerprints and so on, which can provide lightweight security strategies with high scalability for 5G networks. By integrating novel techniques and application scenarios in 5G [1] with physical layer security, we propose a few PHY-layer security complements to defend against security attacks on 5G wireless networks. In this article, we mainly focus on three significant aspects: wireless secure communication, physical layer assisted authentications, and secret key distribution. We point out several promising directions and discuss research opportunities and future challenges, including massive MIMO beamforming with security code communication, self-adaptive mobile and cooperative secure communication with dynamic channel prediction, physical layer assisted authentication, malicious node detection, and key generation of massive MIMO millimeter wave system.

The rest of this article is organized as follows. In the following section, 5G security requirements and physical layer key technologies are presented. Following that, we introduce a massive MIMO beamforming technique with security code framework for 5G wireless communication, which can directly provide confidential information transmission without any pre-shared key. The self-adaptive mobile and cooperative secure communication with dynamic channel prediction is then discussed. We further elaborate cross layer design by combining physical layer security with traditional cryptographic techniques, which can provide stronger secrecy. Then we present lightweight physical layer assisted authentication and malicious node detection schemes. We highlight the opportunities and challenges of physical layer key generation in massive MIMO millimeter wave systems. The final section concludes the article.

5G Security Requirements and Physical Layer Key Technology

Security Requirements of 5G Systems

We list the key security requirements of 5G wireless network systems and the corresponding challenges below.

Jie Tang, Hong Wen, Run-fa Liao, and Fei Pan are with the University of Electronic Science and Technology of China (UESTC); Jie Tang and Kai Zeng are with George Mason University; Fei Pan is with Sichuan Agricultural University; Lin Hu is with Chongqing University of Posts and Telecommunications of China.

Digital Object Identifier:
10.1109/MNET.001.1700412

Flexible Wireless Secure Communication: Conventional cryptographic mechanisms provide communication security based on pre-shared keys to encrypt/decrypt the confidential data. However, given the large number of 5G users, devices, and their huge amount of data exchanged among resource-constrained devices, the traditional security approach makes cryptographic key distribution and management highly challenging. For 5G networks, securing confidential data transmissions with flexible security levels can reduce overhead and improve overall network performance. To satisfy the security requirements, networks must adopt flexible, compatible and extensible security strategies to provide diverse security levels for various types of confidential data. Physical layer and cross-layer security approaches [1, 2] can provide secure communications with low computation overhead, less bandwidth occupation and low power consumption, which are the key features to be considered for 5G wireless secure communication.

User Authentication and Trust Management in Heterogeneous Networks: In 5G based large scale IoT (Internet of Things) [6], a large number of sensor nodes have limited energy resources in different locations and operate unattended by humans. The low computational capabilities and short battery life of IoT devices make them unable to run complex cryptographic authentication and encryption techniques. Moreover, the authentication of mobile nodes in 5G networks faces significant challenges. The huge number of high mobility nodes moving across small cells of different tiers will result in frequent handovers and authentication, which incurs high communication overhead. Thus, lightweight authentication protocols based on various physical layer schemes [5, 10] are in dire need to achieve lightweight user authentication without causing significant overheads. Physical layer authentication can also be used to detect clone and Sybil attacks [8]. As it is known, various attacks, for example, man-in-the-middle attacks, session hijacking, denial of service (DoS), and data modification, can be launched based on clone and Sybil attacks. In 3G and 4G wireless networks, traditional cryptographic techniques are ineffective and fall short from detecting those malicious attacks. By combining lower/physical layer schemes [5, 10] into detection protocols, the clone and Sybil attacks can be detected with relatively lower complexity.

Key Distribution and Management: Conventional cryptography-based security mechanisms need complex key management to distribute, update, and revoke the keys. However, key management is difficult to implement in high-density networks, for example, IoT and D2D (Device to Device) networks [2, 6] where a large number of nodes join and leave the networks frequently. Also, the computation overhead of setting up shared secret keys using the Diffie-Hellman (D-H) protocol is high, which is undesirable for many 5G resource-constrained devices, such as embedded sensors, wearable devices, and so on. Furthermore, with the ever-increasing computing power of attackers, the encryption mechanism has to increase the key length in order to maintain a certain level of security strength, which in turn aggravates the computation overhead. To cope with these challenges, physical layer properties can be utilized to facilitate the distribution of cryptographic keys [4] with lower cost and higher speed in 5G networks.

The mmWave antenna has very small size (half millimeter wavelength) and is used in massive MIMO. The unique propagation characteristics of mmWave are quite different from those at lower frequency, which can be exploited to greatly enhance physical layer security.

Physical Layer Key Technologies for 5G

5G wireless networks integrate new potential transmission technologies [1–3], such as massive multiple-input multiple-output (MIMO), millimeter wave communication (mmWave), non-orthogonal multiple access (NOMA), and full-duplex technology. Some new propagation characteristics and physical properties can be potentially exploited to enhance 5G physical layer security, which bring in great opportunities along with challenges [3]. We mainly discuss massive MIMO and mmWave techniques below, which are two key revolutionary enabling technologies for 5G wireless networks.

Massive MIMO: Massive MIMO [3] is regarded as a revolutionary technology for 5G wireless networks. The benefits of massive MIMO techniques are realized by using very large antenna arrays (typically hundreds) at the transmitter and/or the receiver, which can provide high power and spectrum efficiencies. Massive MIMO also provides rich spatial freedom and channel resources to develop novel physical layer approaches to overcome the security threats of 5G networks. Recently, how to leverage the advantages of massive MIMO for physical layer security has become an important research topic [3]. However, many challenges still need to be resolved in massive MIMO physical layer security designs, such as channel reciprocity, pilot contamination and power allocation.

mmWave: mmWave communication can use a huge range of spectrum, from 30 GHz to 300 GHz, which can alleviate the burden on the nearly fully occupied spectral band of current wireless networks. The mmWave antenna has very small size (half millimeter wavelength) and is used in massive MIMO. The unique propagation characteristics of mmWave are quite different from those at lower frequency, which can be exploited to greatly enhance physical layer security. For example, the mmWave beamforming has a much narrower beam with higher directional property, which can accurately orientate toward legitimate users [12]. Oppositely, users outside of the beam can hardly receive the signal. mmWave is a completely new and promising research frontier with great potential for future physical layer security.

Physical Layer Secure Communication Enhancement for 5G

Massive MIMO Beamforming with Security Code Communication

Figure 1 illustrates the proposed massive MIMO beamforming with security code communication framework. Alice can be viewed as a BS equipped

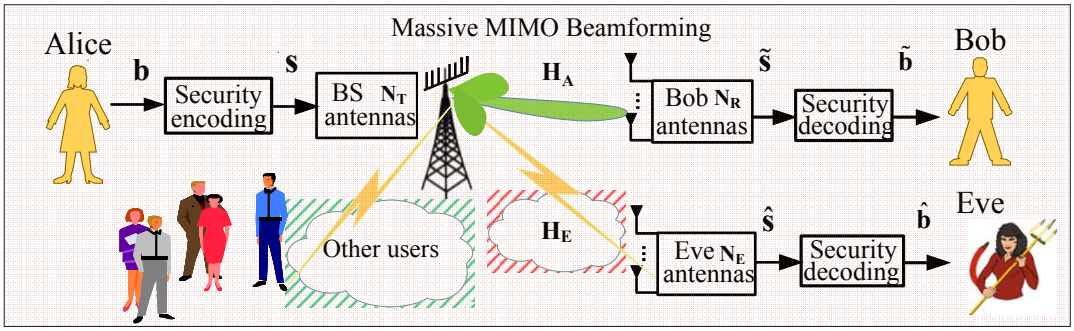

FIGURE 1. Massive MIMO beamforming with security code for secure communication.

SNR	−1 dB	0 dB	1 dB
BER after MB	$\Pr(\hat{\mathbf{s}}) = 0.0734$ $\Pr(\tilde{\mathbf{s}}) = 1.194\ 10^{-4}$	$\Pr(\hat{\mathbf{s}}) = 0.0509$ $\Pr(\tilde{\mathbf{s}}) = 1.958 \times 10^{-5}$	$\Pr(\hat{\mathbf{s}}) = 0.048$ $\Pr(\tilde{\mathbf{s}}) = 2.763 \times 10^{-6}$
Security codes	(255, 56)	(511, 9)	(2047, 22)
BER after SC decoding	$P_{Eve} = 0.4603$ $P_{Bob} < 7.9 \times 10^{-3}$	$P_{Eve} = 0.4795$ $P_{Bob} < 4.8 \times 10^{-3}$	$P_{Eve} = 0.4753$ $P_{Bob} < 2.8 \times 10^{-3}$

TABLE 1. The performance with $N_T = 8$, $N_R = 2$, and $N_E = 2$.

with a massive MIMO antenna array attempting to utilize beamforming (use full or part of the antennas in the array) to send confidential data to Bob. Bob is a legitimate user in the cell who subscribes to the special service offered by Alice. Other users in the cell have not paid for this service. Alice only wants Bob to receive the information, but other users should not get any useful information from the signal. Without loss of generality, assume an eavesdropper Eve who attempts to intercept the signal from Alice to Bob and decode it to compromise the secure information.

The unique channel superiorities in Alice and Bob can be created naturally by beamforming [3], which has been widely investigated for 5G massive MIMO as an essential transmit technique. For the high directional characteristic of the beams directed toward Bob [3], other users in the cell (including Eve) only receive extremely weak signal compared to Bob. The framework achieves secure communication through two steps, namely, the first step is to build a superiority channel by massive MIMO beamforming, while the second step aims to ensure secure communication by enabling the eavesdropper to have about 0.5 bit error rate (BER) with an implementable security code. Security code is also called wiretap channel code [7, 8], which is a specific coding scheme based on the wiretap channel model. Under the degraded wiretap channel, the transmitter can satisfy both secure and reliable communication by security code and it does not need any pre-shared information at the transmitter and receiver.

In Fig. 1, Alice first encodes message bits $\mathbf{b} = (b_1, b_2, ..., b_m)$ by suitable secure coding as $\mathbf{s} = (s_1, s_2, ..., s_n)$. After practical modulation, for example, BPSK, the massive MIMO beamforming process is performed. The number of antennas for Alice, Bob and Eve are denoted by N_T, N_R, and N_E, respectively. Assume that we have a rich scattering environment with Rayleigh fading, and Alice chooses transmit beamformer $\mathbf{f}$ corresponding to the largest diversity gain at the direction of Alice to Bob's channel matrix $\mathbf{H}_A$. For any other receivers (including Eve) with a separation from Bob by at least half a wavelength, the channel will almost suffer independent channel fading with channel gain $\mathbf{H}_E$. From the MIMO diversity, other users (including Eve) will benefit nothing from the massive multiple transmit antennas at Alice. Assume that Bob adopts an MRC combiner to get his optimal receiving SNR. As for Eve, she cannot adopt an MRC combiner for the unknown beamformer $\mathbf{f}$. After the communication between the legitimate partners, the sequence $\tilde{\mathbf{s}} = (\tilde{s}_1, \tilde{s}_2, ..., \tilde{s}_n)$ received by Bob is the noisy version of sequence $\mathbf{s}$. Meanwhile, the eavesdropper Eve can also observe the noisy sequence $\hat{\mathbf{s}} = (\hat{s}_1, \hat{s}_2, ..., \hat{s}_n)$. We always have:

$$\sum_{i=1}^{n} \Pr(\tilde{s}_i \neq s_i) < \sum_{i=1}^{n} \Pr(\hat{s}_i \neq s_i) \tag{1}$$

Eve's channel is a degraded version of Bob's channel. According to Wyner [8], secure communication can be realized.

Both Bob and Eve attempt to perform secure decoding, respectively. Bob and Eve can decode $\tilde{\mathbf{b}} = (\tilde{b}_1, \tilde{b}_2, ..., \tilde{b}_m)$ and $\hat{\mathbf{b}} = (\hat{b}_1, \hat{b}_2, ..., \hat{b}_m)$ from their respectively received signal $\tilde{\mathbf{s}}$ and $\hat{\mathbf{s}}$. The model's final secrecy performance can be measured by conditions [8] which ensure that the legitimate parties can communicate reliably while the eavesdroppers are unable to receive any useful information. In a binary transmission system, the reliability and security conditions can be characterized by:

$$P_{Bob} = \frac{1}{m} \sum_{i=1}^{m} \Pr(\tilde{b}_i \neq b_i) \to 0, \quad m \to +\infty \tag{2a}$$

$$P_{Eve} = \frac{1}{m} \sum_{i=1}^{m} \Pr(\hat{b}_i \neq b_i) \to 0.5. \quad m \to +\infty \tag{2b}$$

The simulation results are shown in Table 1, where $\Pr(\hat{\mathbf{s}})$ and $\Pr(\tilde{\mathbf{s}})$ denote the BER of Bob and Eve after massive MIMO beamforming (MB), respectively. In each SNR (−1 dB, 0 dB and 1 dB), P_{Bob} and P_{Eve} denote the BER of Bob and Eve after secure decoding, respectively. In the experiment, a family of (n, m) linear security codes [7] constructed by BCH code and Hamming code is employed in our security system, where n is the block length, and m is the number of secret information bits. The simulation results verify that the effectiveness of the proposed model approaches Eq. 2.

The proposed model forms a well-integrated security solution that efficiently secures confidential data, such as paid video and high-definition live broadcast. The network's communication and computation overhead can be fully reduced. However, the performance of the proposed model operating in actual channel environments deserves further investigation, because the interference, channel estimation errors, and pilot contamination can all pose as performance limiting factors. Moreover, most works on current MIMO physical-layer security [3–8] conclude that the secrecy rate can be achieved by secure code schemes with infinite block length [8]. In practice, long secure codes [7] are needed to approach the secrecy rate. However, for 5G low latency requirements, the achievable secrecy performance with short length code [7] can be very low and a more effective design should be investigated. The widely discussed artificial noise (AN) [2, 9, 10] can be utilized in Alice's transmit signal to decrease Eve's receive SNR. However, it may interfere other nearby users, which should be further addressed.

Cooperative Secure Communication Strategies in Mobile Networks

The widely discussed physical layer cooperative secure communication [9] offers huge benefits for 5G ultra-high density and multi-tier network security, especially for mobile terminals and IoT nodes. Consider a downlink communication scenario as shown in Fig. 2. Alice is the macro base station intending to send secure information to Bob, in the presence of a possible eavesdropper Eve in the same macro cell. Assume in the pico cell, the access point (AP) Charlie has plenty of resources to act as a cooperator that assists Alice to enhance secure communication from Alice to Bob, while providing service to its own intended users in this cell. In 5G ultra dense multi-tier networks, the nearby BS, APs in the femto cell and even D2D [2] devices can act as cooperators. Proper cooperative communication strategies with lower complexity and power consumption [9] should be investigated well under 5G complex network structures. Moreover, because nodes are deployed in various locations, which may suffer from diverse channel propagation impairments, the physical layer secrecy performance [9–14] should be evaluated under more practical wireless propagation environments, for example, by comprehensively taking into consideration the impact of large scale path loss, shadow fading and small scale channel fading of different users and cooperators.

Furthermore, most existing works only study the physical layer secrecy performance under static scenarios. However, in highly mobile 5G networks with dynamic nodes, for example, connected cars, cellular users walking on the street or riding by a bus, the physical layer secrecy performance under mobility is not well investigated. Most physical layer strategies are based on TDD, which needs CSI estimation (or CSI feedback from the terminal). For high mobility 5G networks, such as vehicular networks, the channel of the mobile nodes may change dramatically during the estimate/feedback delay, which brings serious challenges for secrecy performance analysis and protocol design. Thus, cooperative physical layer secrecy performance, relay strategies and power allocation should be carefully investigated for practical dynamic wireless networks.

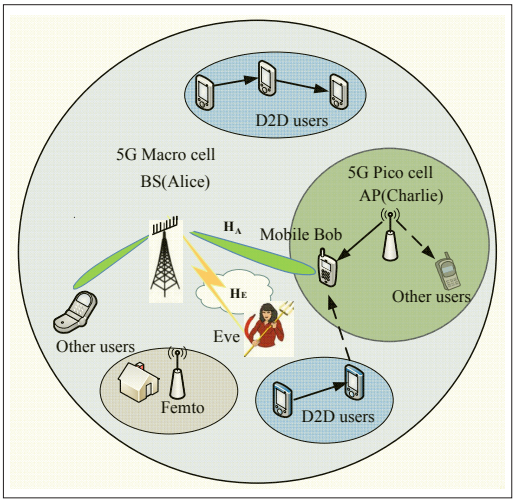

FIGURE 2. Cooperative secure communication in 5G mobile and multi-tier wireless network.

Self-Adaptive Mobile User Secrecy with Dynamic Channel Prediction

Physical layer secure communication can provide enhanced and lightweight security schemes at the expense of transmission rate degradation. Therefore, it is worth investigating when and where physical layer security is suitable in 5G systems. By combining channel prediction with QoS (Quality of Service), a self-adaptive decision can be made.

We address physical layer secrecy adaptation as follows. As a sender, Alice adapts channel prediction to a rapidly changing channel. Some intelligent learning algorithms [10, 15] such as artificial neural network (ANN) can be used to predict channel coefficients well within several coherence time periods. For example, mobile user Bob is assumed to move around in the cell at time t_0 and feeds enough CSI and related information (e.g., speed, locations) back to Alice. Then Alice can track and predict Bob's channel by an ANN within a time interval Δt. Utilizing the previously estimated channels as the training samples, the k-th channel can be tracked by using the previous k – 1estimated channel gains, as $\boldsymbol{H}_b(t_k) = f(\boldsymbol{H}_b(t_1, t_2, ..., t_{k-1}))$, where f means the multi-layer BP neural network. Similarly, the $(k + 1)$th channel is predicted by the k channels before as $\boldsymbol{H}_b(t_{k+1}) = f(\boldsymbol{H}_b(t_2, t_3, ..., t_k))$. Thus for any time slot $t_i > t_0$, Alice can predict Bob's future channel $\boldsymbol{H}_b(t_{i+1})$, $\boldsymbol{H}_b(t_{i+1})$, ..., $\boldsymbol{H}_b(t_{i+\Delta t})$. Figure 3 compares the tracked and real channel at t_{30} by training the channels at timeslots t_1, t_2, ..., t_{29}. Based on the predicted channel, Alice can execute various kinds of security strategies, for example, to estimate Bob's secrecy rate $R_{S,i+\Delta t}$ at the time slot $t_{i+\Delta t}$ and adopt on-off QoS strategies to transmit signal only when $R_{S,i+\Delta t} \geq \tau$ (τ is the desired secrecy rate threshold), or adopt some adaptive code and modulation schemes. How to combine physical layer security with artificial intelligence and machine learning [10, 15] is highly interesting. However, such research is still in its infancy and it is worth further investigating.

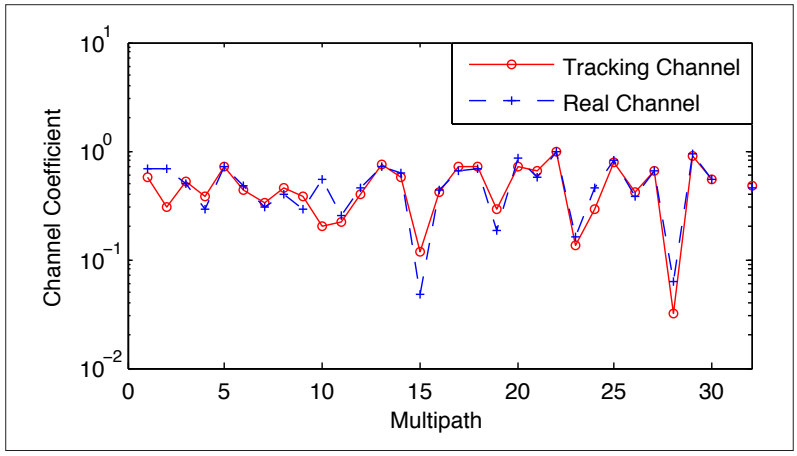

FIGURE 3. The ANN channel tracking.

Cross Layer Design for Secure Communication Enhancement

In real world communications, current physical layer security techniques purely based on wireless channel property may not be able to guarantee the security with probability 1. Hence, current physical layer security can only ensure data secrecy from the information theory perspective, but not guarantee that every data block is secure. This weakness makes it unsuitable to protect information of stronger secrecy requirement, such as financial data. In practice, a powerful eavesdropper can utilize more antennas to get higher receive SNR from legitimate partners and the secrecy capacity will degrade. Therefore, cross-layer design by combining physical-layer security with traditional high layer cryptographic techniques is highly desirable. The work in [13] combines DFRFT (Discrete Fractional Fourier Transform) with stream cipher in a DFRFT modulation system. The work in [14] presents a cross-layer design by randomly replacing OFDM training symbols and inserting dummy data with a pre-shared sequence, which can prevent eavesdroppers from synchronizing and estimating the channel. Under the cross-layer secure system, the potential attackers have to attack the cipher key under noisy signal, which is much more difficult than that of finding the cipher key under error-free signal in the traditional cipher system. For a proper cross-layer design, the protocol complexity should not increase too much while a stronger security enhancement is achieved. Therefore, more efficient and lower complexity cross-layer secure communication design is a promising research direction.

Lightweight Authentication

Cross Layer Authentication

Due to high computation complexity, 4G networks adopt EAP/AKA (Extensible Authentication Protocol Method /Authentication and Key Agreement) authentication protocol. In 5G networks, cross-layer access authentication combining with various lower/physical layer schemes [5, 8, 10] can both enhance the security of user authentication and supplement data packet integrity protection. As shown in Fig. 4a, for user authentication, Alice and Bob implement upper layer authentication first based on the EAP/AKA protocol while the physical authentication information is checked, which can defend against a man-in-the-middle attack. If EAP/AKA authentications succeed, they continue to implement physical layer authentication for data packet integrity protection. The receiver can compare a measured channel response with a prior channel response to distinguish if the data frames are coming from the same authenticated sensor nodes. For example, node Bob needs to authenticate the signal coming from a previously authorized node Alice. Assume the attackers in the cell are in different locations from Alice. Alice transmits data frames to Bob and Bob consciously estimates channel H_k at k-th frame and compares H_k with its previous channel response H_{k-1} (the interval between H_k and H_{k-1} should be designed within the channel coherence time). If their correlation coefficients are higher than a specific threshold, Bob believes the data frames are sent from the same transmitter Alice. Otherwise, Bob can conclude that some of the data frames are sent by attackers and discard those frames.

Malicious Node Detection

Physical layer properties can also assist in detecting malicious nodes, such as clone and Sybil attacks [8]. As shown in Fig. 4b, the normal sensor nodes are organized into local clusters, with one node acting as the local cluster head. The clone attacker can capture a legitimate node and clone its security elements to masquerade it in the network at different locations. The clone nodes have the same ID, key and other related information as the captured node, which causes the traditional up-layer authentication to fail. Then those clone nodes can conduct many types of attacks, such as a man-in-the-middle attack, session hijacking, DoS attack, and data modification. However, the node in different position has different channel responses, by integrating the node's physical channel response into the protocols, the clone nodes can be detected with relatively low complexity. For example, the head node of the cluster can periodically request all nodes in this cluster for authentication and record the channel responses and IDs feeding back from nodes. If the head node finds two different feedback channel responses coming from the same ID, it can find and revoke the cloned nodes.

Similarly, physical layer authentication can also be adopted to detect Sybil attacks. As shown in Fig. 4b, in Sybil attacks, the malicious nodes can fabricate a fake node with different IDs in the network. The goal of detecting Sybil attacks is to validate that each node's identity is the only identity presented by the corresponding physical node. On the contrary, these Sybil nodes generated from the same malicious node are located in the same position. The head node of the cluster can periodically request all nodes in its cluster for authentication and record the channel responses and IDs fed back from those nodes. If two or more nodes have different IDs but have very "close" channel information, a conclusion can be drawn that these nodes are located at the same position and the Sybil attack is detected.

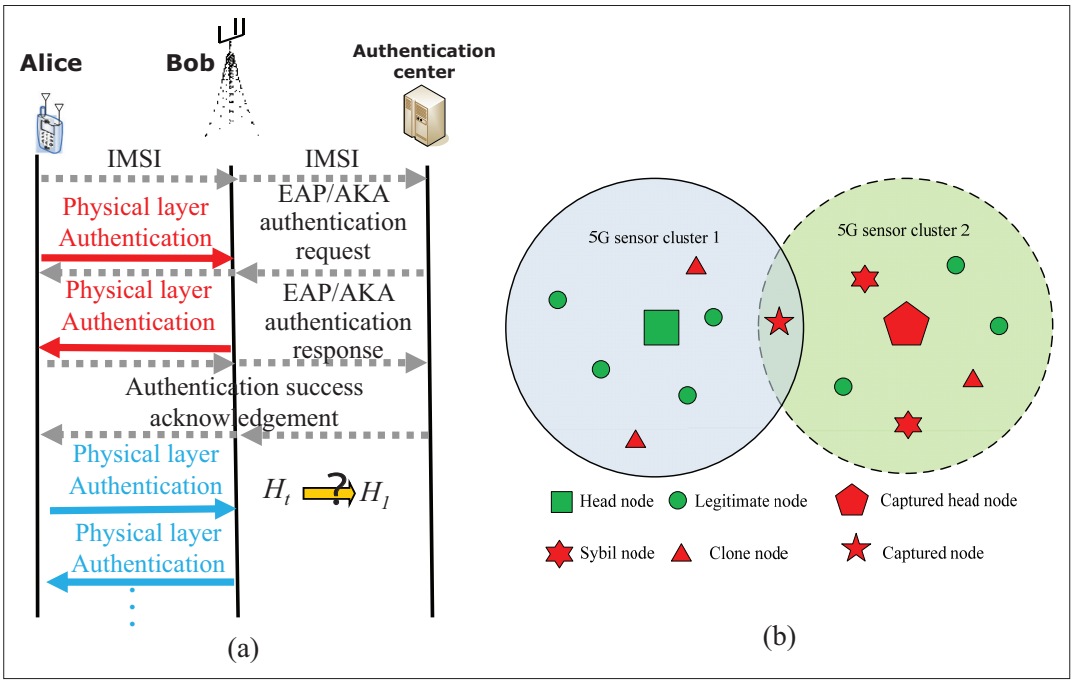

FIGURE 4. The physical layer authentication for 5G: a) the cross-layer authentication; b) malicious node detection.

However, when utilizing the schemes in different propagation environments, the specific protocols and key parameters, for example, decision thresholds, should be designed carefully and the corresponding challenges should be well investigated. For example, if the cluster-head is captured, for example, in cluster 2, the cloned nodes in this cluster cannot be detected because the captured head node will not declare this attack. Also, the captured node in cluster 1 can deploy cloned nodes in cluster 2. In addition, how to design the protocol that can detect a Sybil and clone attack simultaneously is also an interesting topic.

Some intelligent learning algorithms, such as generative adversarial networks (GAN) [15], can be designed to detect malicious nodes. The discriminator function of GAN can be trained to learn the characteristics and distribution of the input data. Meanwhile, the generator function can be designed to generate false data to confuse the discriminator. Therefore, we can train the discriminator function by inputting legitimate nodes' channel information. Meanwhile, when there exist malicious nodes, a well-designed generator function is needed to generate false data to confuse the discriminator. By designing the learning process carefully, the discriminator function can learn to distinguish channel information of legitimate nodes from that of a malicious node intelligently. How to combine the detection of malicious nodes with artificial intelligence and machine learning [10, 15] is highly interesting. However, such research is still in its infancy and it is worth further studying.

Physical Layer Key Generation for 5G mmWave Massive MIMO

For some specific circumstances, physical layer key generation mechanisms do not require expensive computation but can offer key generation at high rates. Therefore, they are expected to work well for certain 5G network scenarios, such as D2D and IoT networks [2, 6]. A typical physical layer secret key generation technique includes the following processes, such as mutual channel probing, reconciliation, and privacy amplifications [4]. For a 5G massive MIMO system with a large number of antennas, the key generation rate is expected to be significantly increased with the increase of the number of antennas. Thus, to utilize physical layer key generation with 5G new techniques, such as massive MIMO and mmWave communications systems, is attractive. Figure 5a depicts the massive MIMO mmWave key generation scenarios. Assume BS Alice with multi-antennas N_t attempting to share a key with mmWave device Bob. Since millimeter antenna size is very small, it is possible that terminal Bob is integrated with a large number of antennas N_r. For traditional physical layer key generation in typical TDD systems, based on the channel reciprocity, Alice and Bob perform mutual channel probing within a very short time interval to estimate channel matrix $\mathbf{H}_A$ and $\mathbf{H}_B$, respectively. However, physical layer key generation in mmWave massive MIMO communications faces many challenges, although for a massive MIMO system with a large number of antennas, the key generation rate is expected to be significantly increased with the increase of the number of antennas. However, as dimensions of the channel matrix grow large, channel estimation becomes particularly challenging at the mutual channel probing stages [4].

The virtual channel representation [12] characterizes the mmWave massive MIMO channel which can be written as $\mathbf{H} = \mathbf{U_r}\tilde{\mathbf{H}}\mathbf{U_t^H}$, where $\mathbf{U_r}$ and $\mathbf{U_t}$ are dimension of $N_r \times N_r$ and $N_t \times N_t$ unitary discrete Fourier transform (DFT) matrices, respectively. The matrix $\tilde{\mathbf{H}}$ denotes

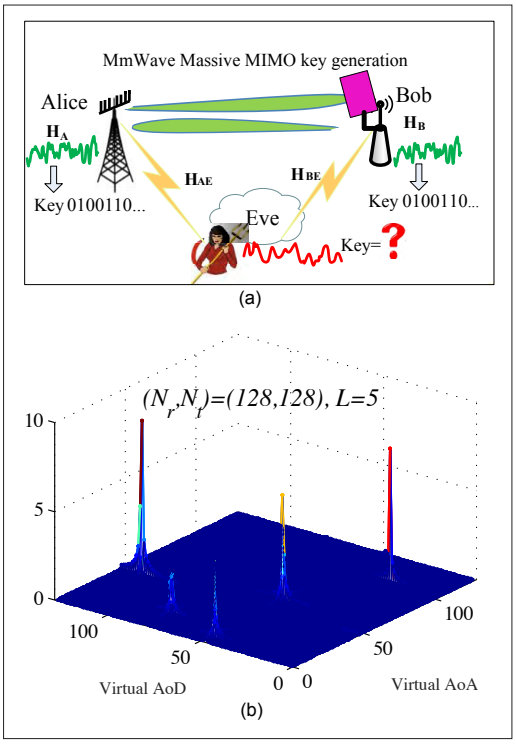

FIGURE 5. Physical layer key generation for 5G mmWave massive MIMO: a) physical layer key generation based on CSI; b) MIMO mmWave virtual channels under $N_t = N_r = 128$ and $L = 5$.

the virtual channel matrix, whose entries capture the gains of the corresponding paths. For a mmWave massive MIMO virtual channel, as the numbers of antennas N_t and N_r increase, the virtual channel becomes more and more sparse. It can be seen that it contains L non-zero peaks corresponding to L sparse paths. Figure 5b shows a mmWave virtual channel with $N_t = N_r = 128$ and $L = 5$. For uniform scattering propagation environments, the positions of L impulses in $\tilde{\mathbf{H}}$ are verified to be uniform, thus this unique property can be considered as a potential random source to generate the secret keys. However, the virtual channel's correlation property should be investigated in order to evaluate its performance when a possible attacker Eve is located very close to Alice or Bob.

As for real world applications, there are more challenges we should consider. First, the mmWave channel coherence time will be significantly small (inversely proportional to the carrier frequency), which introduces challenges when ensuring reciprocity of the mutual channel probing process. The 5G full-duplex technology can be integrated with millimeter-wave communications to achieve high channel reciprocity. Second, the security strength of the proposed key generation scheme under particular attack should be investigated thoroughly, because an advanced attacker with full-duplex capability can jam and eavesdrop at the same time. Also, the research on 5G multi-user key generation is largely open. When considering key distribution for multi-user in real network applications, the end-to-end security key strength, energy consumption, and delay on each link may be different, and thus the flexible control protocol design and security performance evaluation are worth investigating.

Conclusions

In this article, we have discussed physical layer security enhancement mechanisms to fight against security threats in 5G wireless networks. By integrating physical layer security with 5G novel physical layer and up-layer techniques, we have developed a novel framework for massive MIMO beamforming with security code communication. We analyze the validities and benefits of the proposed framework and discuss self-adaptive cooperative security schemes and dynamic channel prediction for 5G mobile security. The potential of physical layer assisted authentication and malicious node detection for 5G networks are also demonstrated. Furthermore, physical layer key generation for 5G massive MIMO millimeter wave systems is discussed. We expect that this article can shed light on the physical layer security design for 5G wireless networks.

Acknowledgment

The work is partially supported by the National Natural Science Foundation of China (no. 61572114) and U.S. NSF grant CNS-1619073

References

[1] N. Yang *et al.*, "Safeguarding 5G Wireless Communication Networks Using Physical Layer Security," *IEEE Commun. Mag.*, vol. 53, no. 4, Apr. 2015, pp. 20–27.
[2] J. Liu *et al.*, "Device-to-Device Communication in LTE-Advanced Networks: A Survey," *IEEE Commun. Surveys. Tuts*, vol. 17, no. 4, Fourth Quarter 2015, pp. 1923–40.
[3]D. Kapetanovic, G. Zheng, and F. Rusek, "Physical Layer Security for Massive MIMO: An Overview on Passive Eavesdropping and Active Attacks," *IEEE Commun. Mag.*, vol. 53, no. 6, June 2015, pp. 21–27.
[4] K. Zeng, "Physical Layer Key Generation in Wireless Networks: Challenges and Opportunities," *IEEE Commun Mag.*, vol. 53, no. 6, June 2015, pp. 33–39.
[5] X. Wang, P. Hao and L. Hanzo, "Physical-Layer Authentication for Wireless Security Enhancement: Current Challenges and Future Developments," *IEEE Commun. Mag.*, vol. 54, no. 6, June 2016, pp. 152–58.
[6] S. Verma *et al.*, "A Survey on Network Methodologies for Real-Time Analytics of Massive IoT Data and Open Research Issues," *IEEE Commun. Surveys. Tuts.*, vol. 19, no. 3, Third Quarter 2017, pp. 1457–77.
[7] H. Wen, P.-H. Ho, and B. Wu, "Achieving Secure Communications over Wiretap Channels via Security Codes from Resilient Functions," *IEEE Wireless. Commun. Lett.*, vol. 3, no. 3, June. 2014, pp. 273–76.
[8] H. Wen, Physical Layer *Approaches for Securing Wireless Communication Systems*, New York, NY, USA: Springer-Verlag, 2013.
[9] L. Hu *et al.*, "Cooperative Jamming Aided Secrecy Enhancement in Wireless Networks with Passive Eavesdroppers," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, Mar. 2018, pp. 2108–17.
[10] L. Xiao *et al.*, "Game Theoretic Study on Channel-Based Authentication in MIMO Systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, Aug. 2017, pp. 7474–84.
[11] J. Tang *et al.*, "Associating MIMO Beamforming with Security Code to Achieve Unconditional Security," *IET Commun.*, vol. 10, no. 12, Apr. 2016, pp 1522–31.
[12] Q. Duan *et al.*, "AoD and AoA Tracking with Directional Sounding Beam Design for Millimeter Wave MIMO Systems," *Proc. 2015 IEEE 26th Annual Int'l. Symposium on Personal Indoor and Mobile Radio Commun. (PIMRC)*, Hong Kong, 2015, pp. 2271–76.
[13] H. Wen *et al.*, "A Cross-Layer Secure Communication Model Based on Discrete Fractional Fourier Transform (DFRFT)," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 1, Mar. 2015, pp. 119–26.
[14] J. Zhang *et al.*, "Design of an OFDM Physical Layer Encryption Scheme," *IEEE Trans. Veh. Technol*, vol. 66, no. 3, Mar. 2017, pp. 2114–27.
[15] C. Jiang *et al.*, "Machine Learning Paradigms for Next-Generation Wireless Networks," *IEEE Wireless Commun.*, vol. 24, no. 2, April 2017, pp. 98–105.

Biographies

Jie Tang (cs.tan@163.com) received his Ph.D. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China. He is currently a postdoctor at George Mason University, Fairfax, VA, USA. His main interests lie in wireless communications and physical layer security.

Hong Wen (sunlike@uestc.edu.cn) received her Ph.D. degree from the Communication and Computer Engineering Department at Southwest Jiaotong University, Chengdu, P. R. China, in 2004. She then worked as an associate professor at the National Key Laboratory of Science and Technology on Communications at the University of Electronic Science and Technology of China (UESTC), P. R. China. From January 2008 to August 2009, she was a visiting scholar and postdoctoral fellow in the ECE Department at the University of Waterloo. Her current main interests are in wireless communication systems security.

Kai Zeng (kzeng2@gmu.edu) is an associate professor in the Department of Electrical and Computer Engineering at George Mason University, U.S.A. He received his Ph.D. degree in electrical and computer engineering from Worcester Polytechnic Institute (WPI) in 2008. He was a postdoctoral scholar in the Department of Computer Science at the University of California, Davis (UCD) from 2008 to 2011. His current research interests are in cyber-physical system security and privacy, physical layer security, network forensics, and cognitive radio networks.

Run-Fa Liao (runfa.liao@std.uestc.edu.cn) is working toward his Ph.D. degree in communication and information system at the National Key Laboratory of Science and Technology on Communications at the University of Electronic Science and Technology of China. His current main interests are wireless communication security and intelligent algorithms.

Fei Pan (panfeivivi@hotmail.com) received her bachelor degree in communication engineering from the Northwest University, Xi'an, China, in 2011, and her Ph.D. degree in communication and information systems from the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2019. She was a visiting Ph.D. student at KTH-Royal Institute of Technology, Stockholm, and a Ph.D. intern at ABB Corporate Research Center, Vasteras, Sweden, from 2017 to 2018. She is currently working as a lecturer at the College of Information Engineering, Sichuan Agricultural University, Ya'an, China. Her research interests include wireless communications, physical layer security, and high performance wireless communications in industrial automation.

Lin Hu (lin.hu@ieee.org) is a lecturer in the School of Communication and Information Engineering at Chongqing University of Posts and Telecommunications. He received a Ph.D. degree in communication and information systems from the National Key Laboratory of Science and Technology on Communications, the University of Electronic Science and Technology of China in 2017. His research interests include cooperative communications and physical layer security.