Spoofing Attack Detection in 5G Network

Monika Singh[1, 2,*] and Navin Kumar[3]

[1]Amrita School of Engineering, Amrita Vishwa Vidyapeetham (University), Bengaluru, Karnataka, India
[2] Department of Electronics and Communication Engineering, CMR Institute of Technology, Bengaluru, Karnataka, India
[3] Department of Electronics and Communication Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham (University), Bengaluru, Karnataka, India
Email: monika.singh@cmrit.ac.in (M.S.); navin_kum3@yahoo.com (N.K.)
*Corresponding author

***Abstract*—Spoofing Attack (SA) is a challenging issue in mobile wireless communication especially with huge traffic in 5G and beyond where attacker inserts counterfeit data with false identification to intercept a valid transmission. Detection and corrective action become very important in these cases. A potential method to prevent identity spoofing threats is channel-based Physical-Layer (PL) privacy. It is of interest to a broad spectrum of people and organizations engaged in network administration and computer security. In this work, channel-based SA identification method is proposed to prevent serious consequences. The Physical Layer (PL) properties are utilized in order to detect SA. As a unique channel feature, the Prime Elements of a Digital Channel Representation (PE-DCR) are identified. In this work, a detection method is developed which is built on PE-DCR to detect SA in stable and varying radio surroundings. The challenge of SA detection is changed into a 1st-Class Categorization (1-CC) issue for the changing radio setting where the channel covariance is fluctuating. An active detection system based on Bidirectional Long Short-Term Memory (BiLSTM) Neural Networks (NN) as Back Propagation Forward Scheme (BPFS) is proposed to effectively manage this issue. Results from simulations validate the viability of the proposed detection methods. The proposed method achieved detection accuracy of 80%.**

***Keywords*—spoofing attacks identification, bidirectional long short-term memory, 5G**

I. Introduction

5G and advanced mobile communication systems are being considered to form a fully connected network [1]. That is, very high dense network allowing large amount of traffic, heterogeneous networks, different kinds and priority of traffic with high degree of operational flexibility, scalability and so on. Additionally, because of large supported applications and services, this network is expected to provide high level of security, authenticity within limited time-bound and coverage area [2]. In fact, security and authenticity has always been a challenge in all networks. But currently, this challenge has become increasingly critical for societal interest. Conventionally, security has been executed in the intellectual levels of communications systems, above the levels of the mechanical data transmission [3]. Encryption is the primary technique for the data confidentiality in most of the current situation and networks. In highly dense networks, data encryption becomes difficult and computationally complex. Therefore, Physical Layer Security (PLS) has become very important for current and future networks. PLS provides an additional layer of safety on leading edge of cryptosystem and encryption [1]. PLS systems make use of the difference between the main and monitoring networks as well as the inherent irrationality and reciprocity of wireless media [1]. It also makes use of a difference in transmission strength received by legal and hostile networks, whereas the security feature of cryptographic relies on the (limited) processing capability of the adversary. According to communication theory, the quantity of data gathered by malware may be decreased if the transmission signals are created to increase the similarity measurement between the transmitters and the receiver. As a result, sensitive information included in transmissions cannot be intruded over due to PLS protections.

Spoofing Attack (SA) is a kind of PLS threat as it tries to forge identity of the senders. The attacker uses false identities like an Internet Protocol (IP) address to appear to be a valid user. The attacker may then proceed illegally from this, to carry out more sophisticated threats, such as Denial-of-Service (DoS) attacks and Man in the Middle (MITM) attacks [4]. Therefore, detection of spoofing attack is important. SA detection uses the built-in characteristics of communication networks to identify various emitters in various places. Detection of identity of SA is done by exploiting wireless PL characteristics, such as Received Signal Strengths (RSS) [5], channel impulse responses [6], channel frequency responses [7], etc. Several specific physical layered security measures deal with fixed instances, and often need a collection of trained data identification [4]. Typically, wireless carriers use a communication security mechanism to verify the integrity of consecutive images on the premise that the images' intervals fall inside the program's calibration period [8]. In addition, the current monitoring systems for PL, SA only identify the hackers each individually. Rapid and effective identification is challenging to perform whenever a

Manuscript received November 8, 2023; revised January 21, 2024; accepted February 6, 2024; published July 16, 2024.

doi:10.12720/jcm.19.7.308-316

massive proportion of Legitimate Users (LUs) must be considered, as in the case in denser communication networks and Internet of Things (IoT) devices. The SA could be determined if the monitoring system finds many transmissions with a certain identifier but originating from separate communication systems. Upon this, SA warning can be triggered and further countermeasures including resuming communication and changing the key might be used by authorized users. These SA detection techniques are already in use, however, still improvement is needed in future networks. Given the distinctive qualities of 5G communications, it is challenging to sustain good detectability using the traditional channel attributes included in current identification techniques [8]. In addition, the Signal-to-Noise Ratio (SNR) of the eavesdropper may be comparable to or even superior to that of the legal channel owing to the inherent unpredictability of the medium; particularly when the eavesdropper is nearer the source than the true recipient. Wyner's theories thus may not work in such circumstances [9]. In the broadcast channel and the Gaussian channel, studies of the achievable theoretical level against eavesdropping were conducted in part as a result of Wyner's work. These methods have motivated a significant amount of past study initiatives for various fading channels from an information-theoretic perspective [10]. In particular, we examine the fading channel models that have successfully captured mm-Wave situations in 5G.

Failure to verify the communication's source is a typical error in the detection of spoofing attacks. In other words, failing to verify if an incoming message is coming from a reliable source or whether it has been faked. Neglecting to observe user behavior in order to spot abnormalities is another error. Attacks that use spoofing sometimes entail strange activity, including signing in from an odd place or at an odd hour. Utilizing single-factor authentication only: While it may effectively stop spoofing attacks, multi-factor authentication is often not employed in many systems. Spoofers may easily acquire access by stealing or guessing passwords in the absence of multi-factor authentication. Many algorithms and methods, including as signature-based detection, Machine Learning (ML), Bayesian networks, deep learning, and biometric authentication, are used to identify spoofing attacks.

This work proposes a new SA identification approach for mmWave 5G networks. The prime elements of digital channel virtual representation PE-DCR are recognized as a distinctive channel property. To recognize SA in both steady and dynamic radio environments, two detection techniques are considered based on PE-DCR. For the stable radio environment, Neyman-Pearson (NP) testing-driven SA detection is proposed, where the channel link is steady based on the l2-norm of PE-DCR. To enable channel-based SA detection in 5G communications, PE-DCR is introduced which is inspired by the newly developed signal-processing technology in mmWave communication namely, channel virtual representation. PE-DCR is more responsive to the transmitter's position. The main contributions in this work are as follows:

- Proposed a new channel characteristic called channel virtual representation to prevent SA in mmWave 5G communications.
- Developed NP testing-based SA detection based on the 12-norm of PE-DVR in a static radio environment.
- Proposed an active detection system based on "Bidirectional Long Short-Term Memory (BiLSTM) Neural Networks (NN)" as Back Propagation Forward Scheme (BPFS) to effectively manage Cross-Correlation Spectral Magnitude Learning (CCSML) issue. In fact, CCSML is related to SA detection for the dynamic radio environment.

The rest of the paper is structured as follows. State of art, the most recent and relevant literature is briefly included in Section II. The system framework and proposed framework/methods is described in Section III. Simulation findings are presented and discussed in Section IV. The study's conclusion is presented in Section V.

II. Relevant Literature

Recently, PL authentication/security in wireless communication has attracted significant research interest. It provides information-theoretic security by exploiting the randomness of PL characteristic of wireless channel. PLS analysis across a variety of 5G supporting technologies has been carried out. It includes massive MIMO, millimeter wave communications, network information, non-orthogonal multiple access, and full-duplex. Unlike encrypted communication approaches which assume an investigator lacks the computational power to solve complicated mathematical tasks in a specific timeframe [11, 12].

An in-depth investigation at 5G networks, with an emphasis on how to use machine intelligence to fix the most pressing issues. Many issues, including subpar beam-forming and slow synchronization with large time spans [13], must be considered via signaling methods before the resulting data transmission rates can be determined. A complete analysis of the state of HetNets cybersecurity, including the numerous underlying techniques and 5G developments are reviewed. Most PLS strategies may be used in harsh environments and with restricted embedded platforms [14]. There is a lack of an extensible, general, and theoretical methodology for classifying many PLS methods presently in use to prevent remote active surveillance. This is developed into a broader idea and strong alternative that can supplement or even replace cryptographic techniques, which present its own set of challenges and difficulties [15]. Focus has been given to hostile ML oriented attacks on transmission schemes with the central server using a Deep Neural Network (DNN) to serve a large number of User Equipment's (UEs) and dividing the transmission capacity across many symmetrical sub bands.

As the fundamental optimizing technique is a challenging topic that can't be effectively handled by methodological tools [16], a data-driven deep learning solution is usually necessary. Making available ubiquitous, security, and almost real-time communication is the primary challenge to creating a smart integrated society

[17]. It includes a summary of the state of research in Intelligent Reflecting Surface (IRS)-aided wireless technology with an emphasis on feasible outcomes to real-world technical challenges. Using a large number of active reflectors, data may be reflected in real time [18].

Therefore, it is important to prevent the SA in mmWave 5G environment and to apply the NN tool for detection and optimize performance analysis.

III. Proposed Methodology

In spoofing assaults, the adversary inserts fictitious messages into a conversation that is otherwise genuine while using a false identity. This false message will contaminate the original messages and by using false identity they can steal the information. Channel-based PL detection is one potential solution towards this PL threat. In this work, it is suggested to use the special properties of virtualized channels to effectively implement channel-based identification in mmWave and Single Input Single Output (SISO) 5G communications. For the purposes of this discussion, let's consider a 5G Wireless Network (WN) consisting of a Base Station (BS), a Legitimate Client (LC), potentially Spoofing Hackers (SHs), and users as shown in Fig. 1. Once a message has already been obtained, the BS might analyze its network parameters. It could be possible to deduce the frequency of the connected transmitters (Tx) from the planes or preambles of this message. SISO and mmWave technology would be installed in transmitters and Receivers (Rx). Standardized 5G communication methods between Base Stations (BS) and Logical Nodes (LNs) may benefit from beamforming technology. Moreover, all LCs are permanently installed in the network at predetermined nodes, while SHs are placed at random but cannot be at the same nodes as LCs. The 5G massive Multiple Input Multiple Output (MIMO) deployment is shown in Fig. 2 but Single Input Single Output (SISO) is considered as initial work before implementing MIMO. It aids in boosting transfer speeds, expanding network reach, and strengthening wireless connections' dependability.

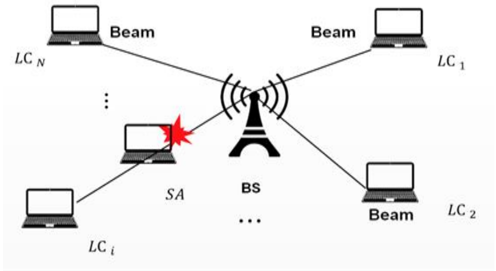

Fig. 1. 5G Massive MIMO application setting [19].

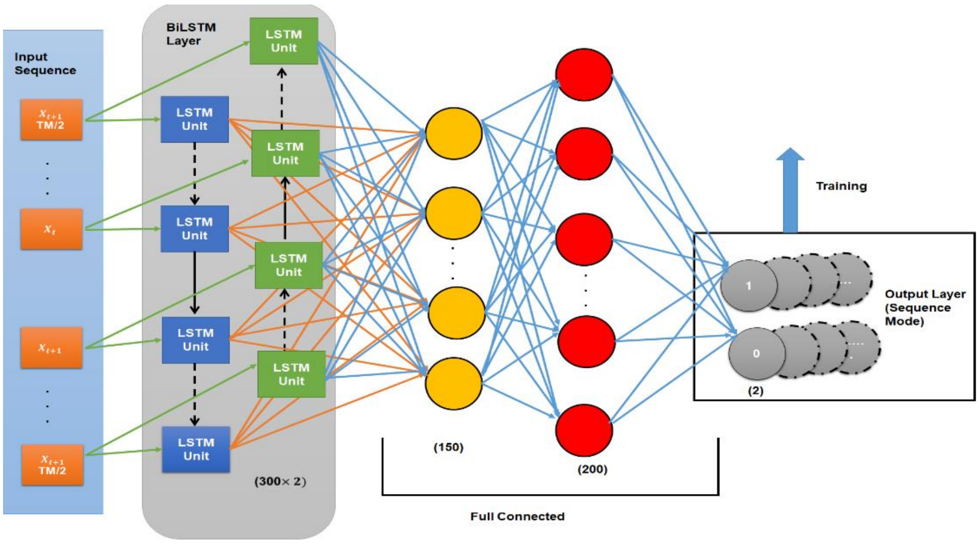

Fig. 2. Framework of the Bi-LSTM modal.

Furthermore, let's assume that SH is a skilled subscriber-spoofing hacker. The SH has the ability to change a variety of data, including the IP/MAC addresses of the sending and receiving nodes, the sequence number, the frames validation, and more. For identifying the key to be compromised, it would need persistent monitoring of the channel in between BS and the permitted users. Even if the victim can never be replaced, the SH can find relief from the identical mmWave SISO. Any time during the transmission session, the intruder may launch the attack under a fake identity and send bogus packets during that time.

Identification framework: Hypothesis analysis is often used in channel-based spoofing identification studies to determine whether the identity data package was indeed sent by the intended transmitter. Let's suppose that a package containing the channel data H is sent by a transmitter, with the identity data i(H) announcing the real emitter. The channel-based spoofing attack identification may be developed using Eq. (1) because H can be identified.

$$H_0 : \zeta(\mathrm{H}) = \epsilon \tag{1}$$

$$H_1 : \zeta(\mathrm{H}) \neq \epsilon \tag{2}$$

where the neutral hypothesis H_0 indicates that ϵ is the actual transmitter of this packet. The transmitter ϵ did not send the package, according to the contrary hypothesis H_1.

The uniqueness of channel modes is the foundation for H identification. The propagation idea states that when the transmitter's location moves by the degree of a frequencies, the channel decorrelation will occur rapidly. Theoretically, relevant channel conditions will change greatly as long as the spacing between the transmitters exceeds the frequency, which would be 10mm for devices at 50GHz. As a result, the transmitter of such acquired data packet is believed to remain unchanged if the receiver may retain the channel state for the most latest transmissions and the channel information of the data obtained and the channel recordings are similar. Nevertheless, the spoofing attack may be identified since the acquired packets might come from several broadcasters. The spoofing attack warning is produced when the spoofing attack identification is effective, enabling authorized clients to take extra safeguards like continuing communication and modifying the code. It is crucial to note that even the spoofing assailant's channel is included in the channel record, the identification technique still works as the acquired packets and the channel account will alter dramatically when a real user sends packets. The obtained signal may be identified by BS using a machine learning based identification system and the ML technique might execute in the application layer process while the objective indicators and training dataset (TD) arrive from the PL. Moreover, both a steady radio field and a mobile dynamic radio setup are taken into consideration. The channel connection between the packets that are acquired and those that are recorded is constant in a stable radio situation, but it will change frequently in a dynamic radio settings.

A. *Channel Virtual Representation*

It is possible to build a channel virtual description for mmWave communications by using a geometrical distribution channels. Considering the fragmented multipath layout in mmWave, as shown in Eq. (3), a geometrical channel notion with dispersion caused by ray tracing could be utilized to explain the network.

$$H = \sqrt{\frac{N_{Tx}N_{Rx}}{\sigma}} \textstyle\sum_{i=1}^{D} \omega a \; \alpha_{Rx}(\varphi_{Rx}, S)\alpha^{*}_{Tx}(\varphi_{Tx}, S) \quad (3)$$

where N_{Tx} and N_{Rx} indicate the antenna numbers of the *tx* and *rx*, correspondingly. The mean direction is denoted by σ. ωa is the appropriate attenuation factor for a complicated Gaussian distribution with a 0 averages, and D stands for the number of dispersion. The actual "angel of departure and angle of arrival" angles on the broadcast and receiving sides are indicated by the symbol φ_{Tx}, S and φ_{Rx}, S. The antenna matrix replies are represented by vectors α_{Rx} and α^{*}_{Tx}. The mmWave SISO channel is represented by defined simulated reception and broadcast orientations using channel digital modeling. The virtual recognition correlates to the network description about evenly spaced temporal angles if the antenna arrangement is a D_m directional uniform linear matrix. Discrete Fourier transform (DFT) matrix obtained as:

$$M = \frac{1}{\sqrt{D_m}} [b(\theta_0), \ldots, a(\theta_{D_m - 1})]^T \quad (4)$$

The digital channel description illustrated in Eq. (5) is predicated on this DFT unified matrix.

$$G = U_r G_V U_t^{*} = \sum_{x=1}^{N_r} \sum_{y=1}^{N_t} G_V(x, y) a_r(\theta_{r,q}) a_i^{*}(\theta_{t,p}) \; (5)$$

where, U_r and U_t are unified DFT matrix that may represent the constant digital obtain angle and constant digital transmission angle that evenly sampled the unit angular space, respectively. The item G_V in the digital channel matrix $G_V(x,)$ captures the gains of the respective pathways. Virtual bins are denoted by the letters $\theta_{r,q}$, and $\theta_{t,p}$. Eq. (6) may be used to express the connection between the actual channel model and the route virtual description based on Eqs. (3)–(5).

$$G_V(x, y) = \sum_{l=1}^{L} \propto_L f(N_r, \varphi_{r,l} - \frac{q}{N_r}) f \times$$

$$\left(N_t, \varphi_{t,l} - \frac{P}{N_T}\right) \quad (6)$$

$$g(\beta, \gamma) = \frac{1}{\beta} \textstyle\sum_{l=0}^{\beta - 1} e^{-j2\pi\gamma l} \quad (7)$$

where the $g(\beta, \gamma)$function is denoted in Eq. (7).

The Eq. (6) conclude that datasets of a flattened form of scattering at digital orientations make up the digital image $G_V(x, y)$. This means that the PC-location CVRs on the digital angle matrices will display the characteristics of the angles of all primary scatter. A SISO channel is investigated, operating at 60GHz with 1×1 antenna and the number of main scatters is 7. Typical channel characteristics illustrate a chaotic environment, which makes it hard to depict the mmWave channel's direction and sparseness. Instead, the channel virtual portrayal might be used to denote scatter characteristics, as the PE-DVR can characterize scatters of different orientations using corresponding sets of virtual bins. All scattering can be detected once the antenna separation is large enough (N_r, N_T). These characteristics might be more useful than conventional channel qualities for identifying mmWave SISO channels.

B. *Identification Premised on Ml for a Dynamic Radio Field*

In order for ML-based approaches to accomplish classification, a classifier must be developed using both optimistic and pessimistic Training dataset. Two issues need to be addressed before using PE-DCR for SA detection in a dynamic radio environment. The most effective means of collecting information about unsuccessful training scenarios are discussed. (i) Valid clients have trouble obtaining samples of these hackers due to the difficulty of doing so in a radio setting where the network connection variables are dynamically changing; (ii) how to effectively upgrade the categorization modal to accommodate the varying radio setting where the network connection variables are dynamically changing. This work proposes a novel digital Bi-LSTM architecture and recasting the problem of identifying spoofing attacks as a 1-CC subject as a means to tackle this challenge. Structure of the Bi-LSTM modal is shown in Fig. 2.

Five levels, including an input layer, a BiLSTM layer, and three completely connected layers, are proposed. Because of the restricted storage capacity of a PL, the input to the BiLSTM layer is a digital data series of length L. In addition, for each set of intricate information, a real two-dimensional matrix is constructed. This results in a 2xTM size for the data sent into the BiLSTM layer. To construct a system for series-to-series regeneration, similar framework to that of a sequential analysis is used, but the process of producing the BiLSTM layer to a serial one is altered. The system's output in series mode is a $2 \times L$ array, however selecting a different mode has no effect on the architecture. The multiplexer of the system is the procedure of the sequenced phase to collect the intended output at each sample interval.

One-class categorization identification: 1-CC identifies the intended messages by using an ML technique designed specifically for positive training dataset. Hence, the ML classifier can only use one kind of TD. For the spoofing attack identification training sets, only typical cases were allowed.

Let $A_{tf} = [a_i, b_i]$ denotes the training factors, $a_i \in \mathbb{R}^n$ $where\ i = 1, \dots, M$ and S is the amount of the instructing sets. The intended label is the single property of the related phase, i.e. $y = [y_i]\ [2, \dots.,2]_{l \times M}$. As a result, the 1-CC seeks to build an appropriate machine learning system using the distance measure $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. The one-class model may provide the appropriate prediction values for a group of testing information $D_{tf.a} = [\hat{a}_1, \dots.., \hat{a}_K]$.

Let $\hat{X}_k = 0$ represent the normal situation, and $\hat{X}_k = 1$ represent the spoofing attack scenario. Eq. (8) represents the illustration of the spoofing attack identification issue.

$$p(\hat{y}_k) = \begin{cases} \hat{X}_k = 0 & for\ normal \\ \hat{X}_k = 1 & for\ SA \end{cases} \tag{8}$$

The Bi-LSTM based architecture will then be introduced to perform this ML-based spoofing attack identification.

C. Proposed Spoofing Identification Using Bi-Lstm Framework

Three processes make up the Bi-LSTM architecture which has been displayed: data processing, training, and digital updating. An illustration of this system is shown in Fig. 3. During data preparation, the original data are normalized and assessed. The generator and classifier are then built using these inputs in the training phase.

Adjustments to the communication environment may be made to the classifier at any time throughout the identification process. The details of each step is discussed as follows:

Data processing: Both measurement and normalisation are often used in data processing.

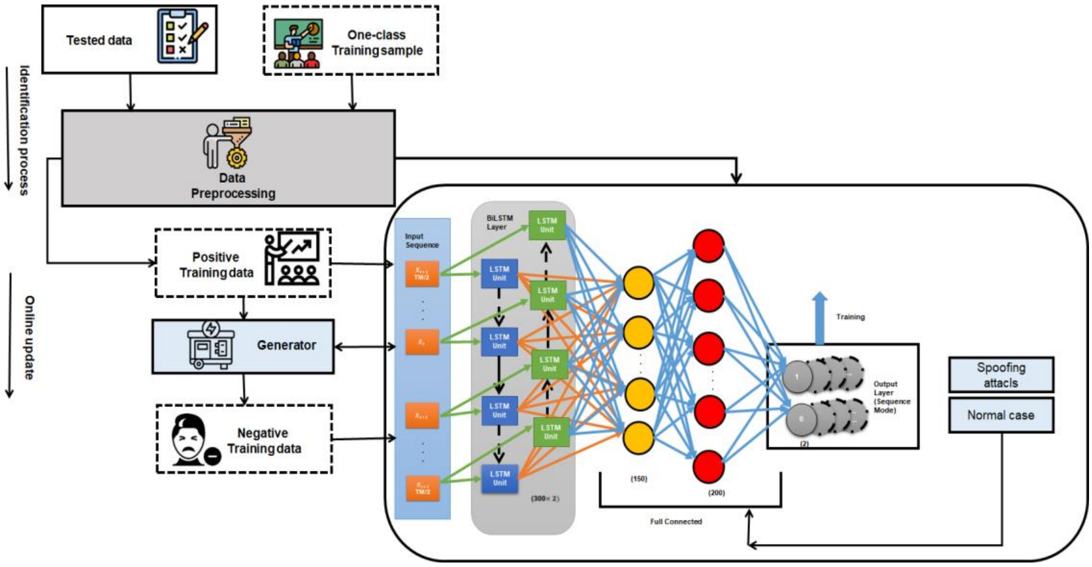

Fig. 3. Proposed flow diagram.

Normalization: The goal of normalization is to translate the relevant data from a specific data collection onto the virtual network channel array. To locate PE-DVR, filtering is used with a threshold of t. The array of a channel may be transformed into a vector and a particular data region.

Measurement: Two different criteria are utilized to quantify the real division between the receiving channel and the channel record. The initial unit of measurement is the Euclidean distance (ED), which is given in Eq. (9).

$$I^{(ED)}(C, C_\Delta) = ||h, h_\Delta||^2 \tag{9}$$

where, C stand for the acquired channel, C_Δ is for the channel history, and $||.||^2$ is for the Frobenius standard. The Pearson correlation coefficient (PCC) is a second metric in Eq. (10).

$$I^{(PCC)}(C, C_\Delta) = \frac{\sum_{i=1}^{N_{rl}}(C_i - \hat{C})(C_{\Delta i}\overline{C_\Delta})}{\sqrt{\sum_{i=1}^{N_{rl}}(C_i - C)^2}\sqrt{\sum_{i=1}^{N_{rl}}(C_{\Delta i} - \overline{C_\Delta})^2}} \tag{10}$$

where, C_i and $C_{\Delta i}$ denote a channel vector component and C_Δ and $\overline{C_\Delta}$ represent the mean readings.

Training procedure: The training process includes both the preparation of the generator and the guidance of the classifier. Using the target information (i.e. positive training dataset (PTD)), the classifier distinguishes the SA condition from the average case, while the producer generates negative training data (NTD). The PTD's characteristics are used as a basis for teaching the generator. If the recording data is to be used in an investigation of a spoofing attack, the PCC between the recording data and the NTD must be less than if the data were to be used in a trial of a spoofing attempt. The ED in a normal case also has to be larger than it would be in a spoofing attack. This finding may inspire the development of a system for the production of NTD based on the characteristics of PTD. Let *N_d* stand for the negative dataset shown in Eq. (11), and let *R* stand for an asymmetric region from which the NTD may be selected at random.

$$N_d = [N_{d,i}^{(PCC)}, N_{d,i}^{(ED)}]_{2\times nd} \in \mathrm{R} \tag{11}$$

Therefore, Eq. (12) may be used to depict the creation of the negative dataset.

$$R\,\text{subject to} \quad \begin{array}{c} 0 \le I_{d,i}^{(PCC)} < I_C^{(PCC)} \\ X \ge I_{d,i}^{(ED)} > I_C^{(PCC)} \\ ||N_d - P_D||^2 \ \le R_s \\ R_s = D_s + \delta \end{array} \tag{12}$$

The surface radius of the PTD is denoted by R_s, while the center of the collection of the positive data is denoted by P_D. The D_s value denotes the difference in size between the positive and NTD. X is just intended as an empirical number and represents the potential upper limit of the produced NTD. δ may be improved during the training of the discriminator since it is a soft margin slack parameter. A solitary concealed layer of LSTM NN serves as the foundation for the classifier. To produce input weights *X*, output weights *Y*, and bias *B*, which are shown in Eq. (13), is the goal of training a NN with one hidden layer for categorization.

$$\min_{X,B,Y} P_{loss} = \sum_{i=1}^{n}(\sum_{i=1}^{K} Y_{ig}(X_i . a_i + B_i) - b_j \tag{13}$$

where *K* and *N* stand for the number of neurons in the concealed layer and the number of TD, correspondingly, and stand for the training samples. *B* is also the designation for the practice sample.

Eq. (14) may be used to determine the output matrices of the concealed layer if the activity factor of the NN is g(H).

$$\ominus = \begin{bmatrix} g(X_1, b_1, a_1) \ldots . g(X_k, b_k, a_1) \\ \vdots \qquad\qquad \vdots \\ g(X_1, b_1, a_n) \ldots . g(X_k, b_k, a_n) \end{bmatrix} \tag{14}$$

Hence, the task of solving a linear system may be reduced to the training of a single hidden layer NN.

$$\ominus\, Z = b \ \ominus\ominus_1\ominus_1 = b \tag{15}$$

where $b = [b_1, b_2, \ldots . b_n]$ is the label of the TD.

Eq. (16) may be used to compute the output levels of the NNs if the matrix ⊖ is established through activation and TD.

$$z = \ominus^+ B \qquad \ominus\, Z = \ominus^+, \ominus^+ B \tag{16}$$

where, $\ominus^+$ is the moores-Penrose generation inverse of ⊖. As a result, Eq. (17) may be used to express the training efficiency ϕ if the forecast attribute for the TD is $\widehat{B}$.

$$\phi = \frac{\|b - \hat{B}\| l_0}{n} \tag{17}$$

where $\|.\|$ indicates l_0 norm that detects the number of non-zero values.

Moreover, Eq. (18) defines a trade-off between ϕ and R_s that must be made to improve the classifier and generator. Simple linear computing may be used to tackle this efficiency challenge.

$$\begin{array}{l} \min 1 - \Phi + R_s \\ \ \Psi \end{array} \tag{18}$$

$$subject\ to \begin{cases} 1 - \Phi \ge 0.99; \\ R_s = R_C + \Psi\ ; \\ \Psi \in [0, \infty); . \end{cases}$$

Identification procedure: While the detection process progresses; the proposed system might periodically update the classifier to take into account new information about the surrounding environment. To illustrate, suppose that *N* is the total number of TD, Qx is the set of PTDs that have just been confirmed, $Y\ ((0))$ is the set of output levels for the previous hidden layer, and *Y* ((1)) is the set for the current hidden layer. As shown in Eq. (19), the original output matrix is $\odot\ ((0))$, and a new one, $\odot\ ((1))$, may be generated using the new data.

$$\odot_{(1)} = \begin{bmatrix} g(X_1, b_1, \tilde{a}_1) & \cdots & g(X_L, b_K, \tilde{a}_1) \\ \vdots & \ddots & \vdots \\ g(X_1, b_1, \tilde{a}_{\tilde{N}}) & \cdots & g(X_L, b_K, \tilde{a}_{\tilde{N}}) \end{bmatrix} \tag{19}$$

Eq. (20) may be used to resolve the novel output weights X_1, where $\gamma = \odot_{(0)}^T \odot_{(0)} + \odot_{(1)}^T \odot_{(1)}$.

$$Y_{(1)} = Y_{(0)} + \gamma^{-1} \odot_{(1)}^T (B_{(1)} - \odot_{(1)} Y_{(0)} \tag{20}$$

Algorithm 1 presents the suggested Bi-LSTM method's pseudo-code.

Algorithm 1: Proposed Framework

```
Require Training sample $G_V$.
Iterate for every cycle
Data preparation
    i.  Normalization;
    ii. Determine the PTD's using Eq. (15) and (16);
Training procedure:
    i.   Acquire the NTD by Eq. (17);
    ii.  Train classifier based on Eq. (18);
    iii. Optimization based on Eq. (19);
Identification process:
Depending on the classifier, determine the forecast value
```

```
        if y= 0
                Accept this message x̃
        else
                Raise alarm.
        end if
    Adjust the classifier by eq. (20)
End Repeat
```

To simulate, MATLAB tool is used on a 64-bit Windows desktop with an i7-7700 processor, 16 GB of storage, and the "Monte Carlo experiment" data set. In this investigation, it is shown show that the practical angle of departure and angle of arrival (AoD/AoA), i.e., ∅_t, l, and ∅_r, l, can be uniformly generated arbitrarily without compromising generality (0, 2π). It is assumed that each scattering creates a new channel in the propagation matrix, and that L is the maximum number of scatters. The attacker's channel in a spoofing assault is also distinct from the legitimate channel. In a typical setting, the channel correlation may be represented by Eq. (21) of Jakes' model.

$$H_d(k+1) = xH_d(k) + \omega(k) \tag{21}$$

where $H_d(k+1)$ and $H_d(k)$ represent the channel data collected from two subsequent data from similar tx. A is the network connection variable, and $\omega(k)$ denotes an independence of H_d zero-mean complicated Gaussian procedure $\omega(k)$. Eq. (22) defines the variance of ω(k).

$$\sigma_\omega^2 = (1 - x^2)\sigma_A^2 \tag{22}$$

Practically stated, the phrase $J_0(2\pi vT/\varepsilon)$, where ε is frequency, v is the node's velocity of motion, and J_0 stands for the Bessel functionality of the initial type and 0th type, may be used to denote the channel correlation coefficient x. Consider a range of sustainable and communication factors, SNRs, the number of antennas, and the number of TD, to assess the efficacy of the suggested ML-based method. Hence, the 1-CC identification method is suggested. In addition, Eq. (23) efficiency is taken into account and used the detection accuracy P_{DA} as a quality criterion.

$$P_{DA} = 1 - (P_{MD} + P_{FA}) \tag{23}$$

where P_{FA} stands for the false alarm rate and $P_{MD} = 1 - P_D$ stands for the miss detection rate.

The probability of detecting a signal accurately or event in a certain system or channel is referred to as probability detection accuracy (P_{DA}). On the other hand, the channel correlation parameter measures the correlation or connection between various parts or sections of a communication channel.

IV. Result and Discussion

In this section, simulation results are described to verify the channel-dependent SA identification methods. In signal processing and communication engineering, the probability of detection accuracy P_{DA} vs SNR graph is often used to assess how well a detection system or a communication system performs in the presence of noise. The trade-off between the chance of detection and the probability of false alarms at various levels of SNR is examined by plotting the P_{DA} against SNR.

As the communication channel is not constant, the network connection variables are kept constant throughout training and testing of the proposed system in static scenario. Fig. 4 shows the detection accuracy P_{DA} results for various values of the channel correlation parameter. The findings demonstrate that the Bi-LSTM frameworks suggested model i.e. BPFS offers detection accuracy almost in the range of 60-80% with respect to the channel correlation value. For instance, it is observed that when channel correlation value is 0.7, then detection accuracy is about 80%. This indicates spoofing attack can be detected up to 80%. As channel correlation increases, detection accuracy will improve. Detection accuracy will be slightly lower if correlation parameter is lower. Here bidirectional approach is discovered which has not been considered earlier by researcher to the best of our knowledge. Thus, it can be concluded that it provides reliable accuracy for the variety of channel correlations.

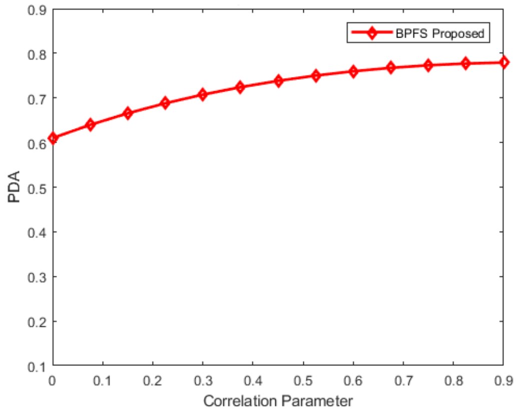

Fig. 4. Detection accuracy vs channel correlation in static scenario.

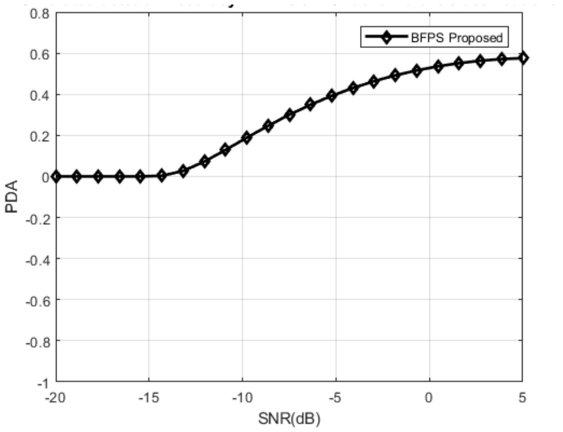

Fig. 5. PDA vs SNR performance in static scenario.

Next, for the same static scenario, detection accuracy at different SNR ratios is seen in Fig. 5. In a communication system, the SNR is a crucial metric that measures how powerful the intended signal is in comparison to the ambient noise. In general, there is a significant correlation between SNR and the detection accuracy. It is observed from the simulation results that detection accuracy becomes better as the SNR rises. This is due to the fact that a greater SNR makes it simpler to separate the signal from the noise since the signal power is stronger in comparison to the noise power. The proposed detection system with optimum SNR depicts that detection effectiveness is improved. For example when the SNR is 5 then the detection accuracy of the proposed BPFS framework reaches 60%. At the lower SNR of 20 dB to -15 dB detection accuracy is observed as 0% and it is gradually increases from −13 dB.

Next, in dynamic scenario, the performance of SNR and detection accuracy is observed as shown in Fig. 6. A higher SNR results in more effective spoofing detection. The suggested approach has a high accuracy at higher SNR. The accuracy decreases at lower value of SNR.

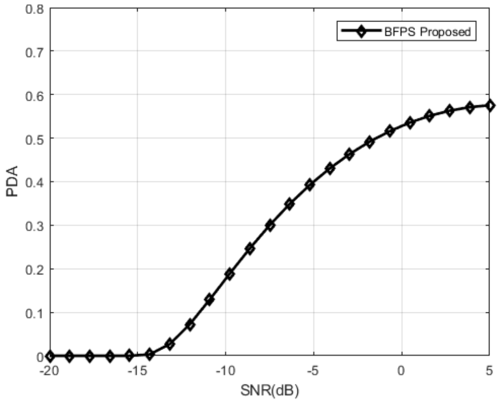

Fig. 6. PDA vs SNR performance in dynamic scenario.

Furthermore, in dynamic scenarios, the networking environment might cause the system's TD and identification input to come from different places. Fig. 7 shows the detection accuracy P_{DA} results for various values of the channel correlation parameter. The findings demonstrate that the Bi-LSTM framework's suggested model i.e. BPFS offers detection accuracy almost in the range of 68-80% with respect to the channel correlation value. For instance, it is observed that when channel correlation value is 0.9 then detection accuracy is about 80%. This indicates spoofing attack can be detected up to 80%. Compare to static scenario, in case of dynamic scenario, detection accuracy performance is improved at lower channel correlation. The proposed work is limited to SISO.

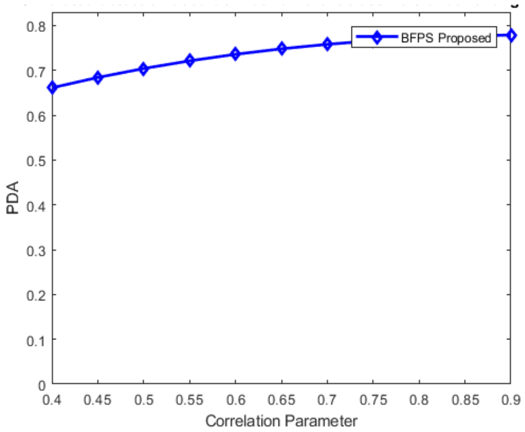

Fig. 7. Detection accuracy vs channel correlation in dynamic scenario.

V. Conclusion

The Bi-LSTM architecture is proposed for mmWave 5G network to prevent PL-SA. The proposed method manages ML-based SA detection based on prime elements of a PE-DCR. Simulation results show that the detection rate of the proposed method is much superior to that of the standard methods. The detection accuracy of a traditional system is only around 60% in a static radio scenario and for the proposed approach is 80%. In terms of training effectiveness and detection accuracy, the proposed Bi-LSTM framework outperforms the most well-known 1-CC classifiers, and the detection accuracy in a variable radio setting exceed 99%. The proposed work is limited to SISO and hence it can be extended further in future to implement the MIMO and massive MIMO scenario. The system can be also analysed using NN methods and performances can be compared.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Monika Singh conducted the research work. Navin Kumar supervised the work as a supervisor and analyzed the data; all authors had approved the final version.

References

[1] Y. Gao *et al.*, "Physical layer security in 5G based large scale Social networks: Opportunities and challenges," *IEEE Access*, vol. 6, pp. 26350–26357, 2018.

[2] S. Wang, X. Xu, K. Huang, X. Ji, Y. Chen, and L. Jin, "Artificial noise aided hybrid analog-digital beamforming for secure transmission in MIMO millimeter wave relay systems," *IEEE Access*, vol. 7, pp. 28597–28606, 2019.

[3] K. N. Le and T. A. Tsiftsis, "Wireless security employing opportunistic relays and an adaptive encoder under outdated CSI and dual-correlated nakagami-m fading," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2405–2419, March 2019.

[4] H. Boche and C. Deppe, "Secure identification under passive eavesdroppers and active jamming attacks," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 472–485, Feb. 2019.

[5] D. P. Moya Osorio, H. Alves, and E. E. B. Olivo, "On the secrecy performance and power allocation in relaying networks with untrusted relay in the partial secrecy regime," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2268–2281, 2020.

[6] S. W. Lai, R. Zhao, S. P. Tang, J. J. Xia, F. S. Zhou, and L. S. Fan, "Intelligent secure mobile edge computing for beyond 5G wireless networks," *Physical Communication*, vol. 45, p. 101283, 2021.

[7] T. Thangappan and B. Therese, "Overview of fronthaul technologies and the DBA algorithms in XGPON-based FH technology in CRAN architecture in 5G network," *Lecture Notes in Electrical Engineering*, vol 792, pp. 271–280, 2022.

[8] V. Sanchez *et.al*., "On the statistics of the ratio of nonconstrained arbitrary α-μ random variables: A general framework and applications," *Transactions on Emerging Telecommunications Technologies*, vol.31, p. 3832, 2020.

[9] R. M. Borges *et. al*., "Integrating optical and wireless techniques towards novel fronthaul and access architectures in a 5G NR framework," *Journal of Applied Sciences*, vol. 11, p. 5048, 2021.

[10] T. Sharma, A. Chehri, and P. Fortier, "Review of optical and wireless backhaul networks and emerging trends of next-generation 5G and 6G technologies," *Transactions on Emerging Telecommunications Technologies*, vol. 32, p. 4155, 2021.

[11] N. Wang, L. Jiao, P. Wang, M. Dabaghchian, and K. Zeng, "Efficient identity spoofing attack detection for IoT in mm-Wave and massive MIMO 5G communication," in *Proc. 2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1–6.

[12] J. D. V. Sánchez *et al*., "Survey on physical layer security for 5G wireless networks," *Ann. Telecommun*, vol. 76, pp. 155–174, 2021.

[13] J. Tanveer, A. Haider, R. Ali, and A. Kim, "Machine learning for physical layer in 5G and beyond wireless networks: A survey," *Electronics*, vol. 11, no. 1, p. 121, Dec. 2021.

[14] F. Irram, M. Ali, M. Naeem, and S. Mumtaz, "Physical layer security for beyond 5G/6G networks: Emerging technologies and future directions," *Journal of Network and Computer Applications*, vol. 206, p. 103431, 2022.

[15] J. M. Hamamreh, H. M. Furqan, and H. Arslan, "Classifications and applications of physical layer security techniques for confidentiality: A comprehensive survey," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1773–1828, 2019.

[16] B. Kim, Y. Shi, Y. E. Sagduyu, T. Erpek, and S. Ulukus, "Adversarial attacks against deep learning-based power control in wireless communications," in *Proc. 2021 IEEE Globecom Workshops (GC Wkshps)*, Madrid, Spain, 2021, pp. 1–6.

[17] M. Vaezi *et al*., "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 2, pp. 1117–1174, Secondquarter 2022.

[18] B. Zheng, C. You, W. Mei, and R. Zhang, "A Survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 2, pp. 1035–1071, Secondquarter 2022.

[19] W. Li, N. Wang, L. Jiao, and K. Zeng, "Physical layer spoofing attack detection in MmWave massive MIMO 5G networks," *IEEE Access*, vol. 9, pp. 60419–60432, 2021.