


Deep-Learning-Based Physical-Layer Lightweight Authentication in Frequency-Division Duplex Channel

Yuta Matsuzaki, Shun Kojima, *Member, IEEE*, and Shinya Sugiura , *Senior Member, IEEE*

Abstract—This letter proposes a lightweight authentication scheme based on secret key generation for frequency-division duplexing. Firstly, a base station predicts downlink channel state information (CSI) from uplink CSI with the aid of deep learning. Then, a secret key is shared between the BS and a mobile user by quantizing the downlink CSI. Since this key generation method uses physical-layer features, the costs of the calculation complexity, the key distribution, and the management, which are typically imposed by the conventional upper-layer key generation, are significantly reduced. Furthermore, the generated key is utilized to carry out low-latency and low-complexity authentication, which is suitable for Internet of things applications.

Index Terms—Authentication, deep learning, frequency-division duplex, grant-free access, secret key generation.

I. INTRODUCTION

THE communication improves toward beyond 5G (B5G) standard [1] is expected to support a diverse number of Internet of things (IoT) devices and is also necessary for boosting fundamental wireless performance. Furthermore, maintaining IoT security [2] against denial of service attacks, resource consumption, masquerade attacks, replay attacks, information disclosure, and message modification is also vital in B5G [3] [4], [5], [6]. In order to encrypt information, a secret key is commonly shared between two legitimate users by public-key algorithms, such as RSA [7] or elliptic curve discrete logarithm problem (DLP) [8] ones. However, public-key algorithms cannot achieve quantum resistance and typically impose high encoding/decoding complexity, especially when employed for encryption with a long public key. Hence, key exchange by a public-key algorithm may not be suitable for future IoT devices in terms of security and energy efficiency.

By contrast, as a part of physical-layer security, the concept of secret key generation (SKG) has been extensively investigated [9], where the costs of key sharing/managing, the overhead, the latency, and the encoding/decoding complexity are reduced [10]. More specifically, a secret key is generated by the quantization of channel state information (CSI) shared between two legitimate users [11]. In most previous SKG

studies [12], the use of a reciprocal time-division duplex (TDD) channel is assumed to allow two legitimate users to share the channel information and acquire the same secret key from quantized CSI. To relax the reciprocal channel constraint, SKG in a non-reciprocity frequency-division duplex (FDD) channel was developed [13], [14], where the correlation between the uplink and downlink channels is exploited. As an additional benefit, SKG in FDD allows a reduction of latency for the SKG in comparison to its TDD counterpart.

The use of secret keys for authentication has been considered with the aid of a public key infrastructure (PKI) [15], which uses a pair of a secret key and a public key in the public-key algorithm, whose security performance depends on the difficulty of the mathematics problems, such as factoring problems or discrete logarithm problems, or elliptic curve DLP. To avoid the complexity inherent to a PKI, an authentication method using a physical-layer feature was proposed with the aid of the confirmation of characteristic distortion [16], where a secret key in TDD is used to assign time slots to each user. In this scheme, the complexity and the latency are significantly low. However, to the best of our knowledge, physical-layer authentication in a non-reciprocal FDD channel, where the uplink and downlink channels are not the same, has not been developed, nor has a detailed analysis of the proposed method been performed.

Against this background, the novel contributions of this letter are as follows. We propose a lightweight authentication scheme based on SKG in a non-reciprocal FDD channel. More specifically, the downlink channel is estimated from the uplink pilot symbols with the aid of DL at a base station (BS), while a user directly estimates the downlink channel from the pilot symbols transmitted from the BS. This allows us to reduce the feedback information typically needed for conventional SKG in the FDD channel, hence reducing the risk of information leakage to an eavesdropper, as well as reducing the delay. Furthermore, a secret key is generated and shared between the BS and each user by quantizing the associated downlink channel coefficients. In the presence of fading, an attacker cannot access the associated downlink channel, and hence the generated key is quantum resistant while benefiting from significant reductions in latency and computational cost, as well as power consumption in comparison to the conventional public key cryptography [17], [18]. Moreover, lightweight authentication is invoked for IoT communication in the scenario of low-latency FDD grant-free access. The BS authenticates each legitimate user when a user transmits data in the time slots allocated by the physical-layer secret key.

Manuscript received 18 April 2023; revised 19 May 2023; accepted 12 June 2023. Date of publication 14 June 2023; date of current version 12 August 2023. This study was supported in part by the Japan Science and Technology Agency (JST) Fusion Oriented Research for disruptive Science and Technology (FOREST) (Grant Number JPMJFR2127), and in part by the National Institute of Information and Communications Technology (NICT), Japan. The associate editor coordinating the review of this letter and approving it for publication was C. Kundu. (*Corresponding author: Shinya Sugiura.*)

The authors are with the Institute of Industrial Science, The University of Tokyo, Tokyo 153-8505, Japan (e-mail: sugiura@iis.u-tokyo.ac.jp).
Digital Object Identifier 10.1109/LCOMM.2023.3286043

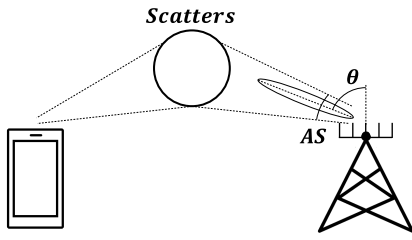


Fig. 1. Multipath channel model between the BS, equipped with a uniform linear array, and the single-antenna user. P scatters positioned from the BS to the direction of $\theta \in [\theta_p - \Delta\theta/2, \theta_p + \Delta\theta/2]$ ($1 \leq p \leq P$).

II. SYSTEM MODEL

A. FDD Non-Reciprocity Channel Model

Fig. 1 illustrates a multipath channel model considered in this letter in which each path is the same in uplink and downlink, similar to [13]. The BS is equipped with M antenna elements in the form of a uniform linear array (ULA), and each user is equipped with a single antenna element. In this letter, only a single-user scenario is considered for the sake of simplicity, but this model is readily applicable to a multi-antenna multi-user scenario.

The channel is assumed to consist of P paths, and the direction of each path from the BS to the p th scatterer obeys the uniform random distribution over $[\theta_p - \Delta\theta/2, \theta_p + \Delta\theta/2]$, where $\Delta\theta$ is an angular spread (AS). The channel vector at the carrier frequency of f between the BS and the user is represented by

$$\mathbf{h}(f) = \sum_{p=1}^P \alpha_p e^{-j2\pi f \tau_p + j\phi_p} \mathbf{a}(\theta_p) \in \mathbb{C}^M, \quad (1)$$

where α_p , ϕ_p , τ_p , and θ_p are the attenuation, phase shift, delay, and direction of arrival (DOA) for the p th path, respectively. Moreover, $\mathbf{a}(\theta_p)$ is the array manifold vector, defined as

$$\mathbf{a}(\theta_p) = \left[1, e^{-j\chi \sin \theta_p}, \dots, e^{-j\chi(M-1) \sin \theta_p} \right]^T \in \mathbb{C}^M, \quad (2)$$

where we have $\chi = 2\pi df/c$, while d is the antenna spacing of the ULA at the BS and c is the speed of light. Note that α_p is a function of the length l_p of the p th path, which is represented by $\alpha_p = (\lambda/4\pi l_p)^2$. The phase ϕ_p depends on the scatter materials and angles of the incident wave. The delay τ_p is calculated based on the distance traveled by the signal along the p th path. In this letter, we consider the carrier frequencies f_U and f_D for uplink and downlink in an FDD channel, respectively. Hence, it can be seen from (1) that the uplink and downlink channel vectors $\mathbf{h}(f_U)$ and $\mathbf{h}(f_D)$ are non-reciprocal. This implies that simple quantization of each channel vector, which is typically considered in a TDD scenario, does not generate identical secret keys in an FDD channel.

B. Deep-Learning-Based Downlink Channel Estimation at BS

In this section, we introduce DL-aided channel estimation of the downlink channels from the uplink channels for the FDD system at the BS [12], [13]. More specifically, our neural network model to be trained, the received pilot signals, the loss

function used for our DL, and the optimization algorithm are given. The channel vector estimated by our neural network is given by

$$\hat{\mathbf{h}}(f_D) = \mathbf{F}^{(L)} \in \mathbb{C}^M, \quad (3)$$

where

$$\mathbf{F}^{(l)} = \begin{cases} \mathbf{f}^{(l-1)}(\mathbf{F}^{(l-1)}) & \text{for } 2 \leq l \leq L \\ \bar{\mathbf{h}}(f_U) & \text{for } l = 1 \end{cases} \quad (4)$$

Also, L is the artificial neural network (ANN) number of layers, and $\bar{\mathbf{h}}(f_U)$ represents the uplink channel vector estimated by the traditional (non-DL) algorithm at the BS. Moreover, $\mathbf{f}^{(l)}$ is a nonlinear transformation function, such as the rectified linear unit (Relu) function [19], in the l th layer, which is written as

$$\mathbf{f}^{(l)}(z) = \begin{cases} \mathbf{g}(\mathbf{W}^{(l)}z + \mathbf{b}^{(l)}) & \text{for } 1 \leq l < L-1 \\ \mathbf{W}^{(l)}z + \mathbf{b}^{(l)} & \text{for } l = L-1 \end{cases} \quad (5)$$

where $\mathbf{W}^{(l)} \in \mathbb{C}^{M \times M}$ and $\mathbf{b}^{(l)} \in \mathbb{C}^M$ are the parameters of a neural network, which are trained according to our algorithm presented below. Furthermore, \mathbf{g} is the activation function, which is given by

$$\mathbf{g}(z) = \max\{\Re[z], \mathbf{0}\} + j \max\{\Im[z], \mathbf{0}\}, \quad (6)$$

where $\Re[\cdot]$ and $\Im[\cdot]$ are the real and imaginary parts of a vector, respectively.

In the proposed downlink CSI estimation, we have two stages, namely, the training stage and the deployment stage. In the training stage, during the channel coherence time, the BS transmits a pilot symbol from each antenna element to the user while the user also transmits a pilot symbol to the BS, which is repeated T times. The associated received signals at the BS and the user are modeled, respectively, by

$$\mathbf{y}_{\text{BS}}^{(t)} = \mathbf{h}^{(t)}(f_U)x + \mathbf{n}_{\text{BS}}^{(t)}, \quad \text{for } t = 0, \dots, T-1 \quad (7)$$

$$\mathbf{y}_{\text{u}}^{(t)} = \mathbf{h}^{(t)}(f_D)x + \mathbf{n}_{\text{u}}^{(t)}, \quad \text{for } t = 0, \dots, T-1, \quad (8)$$

where x comprises a pilot symbol, while $\mathbf{n}_{\text{BS}}^{(t)}$ and $\mathbf{n}_{\text{u}}^{(t)}$ are both additive white Gaussian noise that obeys the complex-valued Gaussian distribution $\mathcal{CN}(0, \sigma^2)$. Also, σ^2 is the noise variance, and the transmit power of a pilot symbol is represented by $P_{\text{tx}} = E[|x|^2]$. Furthermore, $E[\cdot]$ represents the expectation operation.

The downlink channel vector $\bar{\mathbf{h}}(f_D)$ estimated at the user, based on a traditional algorithm, such as zero-forcing (ZF) or minimum mean-square error (MMSE), is fed back to the BS. Then, the neural network is trained based on $\bar{\mathbf{h}}(f_D)$ and $\bar{\mathbf{h}}(f_U)$, which are the response variable and the explanatory variable, respectively. More specifically, the neural network is trained to minimize the loss function

$$\text{Loss}(\Omega) = \frac{1}{TM} \sum_{t=0}^{T-1} \left\| \hat{\mathbf{h}}^{(t)}(f_D) - \bar{\mathbf{h}}^{(t)}(f_D) \right\|_2^2, \quad (9)$$

where $\Omega = \{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}_{l=1}^{L-1}$, and $\hat{\mathbf{h}}^{(t)}(f_D)$ is the downlink channel vector, estimated by DL at the BS. Also, $\|\cdot\|_2$ denotes the l_2 norm. In our scheme, the parameters of the neural

network, $\Omega = \{\mathbf{W}^{(l)}, \mathbf{b}^{(l)}\}_{l=1}^{L-1}$, are optimized by minimizing the loss function $\text{Loss}(\Omega)$ with the aid of the adaptive moment estimation (ADAM) algorithm [20]. In the deployment stage, the BS and the user send pilot symbols to each other, and the BS estimates the uplink channel vector $\mathbf{h}(f_U)$ based on the MMSE algorithm as $\bar{\mathbf{h}}(f_U)$. Then, $\bar{\mathbf{h}}(f_U)$ is input into DL to output $\hat{\mathbf{h}}(f_D)$, where the parameters optimized at the training phase are used.

C. SKG From Estimated Downlink Channel Vector

While the BS estimates $\mathbf{h}(f_D)$ according to the DL-based scheme of Section II-B, the user estimates $\mathbf{h}(f_D)$ from the pilot symbols transmitted from the BS. Then, a secret key is generated by quantizing the estimated downlink channel vector at the BS and the user. In this letter, only the phase information for the estimated downlink channel vector is used for the sake of simplicity. In order to attain quantization of an n -bit secret key per channel, 2^n -level phase demodulation is carried out. Therefore, an $(M \times n)$ -bit secret key is generated for each channel vector $\mathbf{h}(f_D) \in \mathbb{C}^M$.

D. Allocation of Active Time Slots

The BS authenticates the legitimate user by L_1 specific active time slots within a frame, which are allocated based on the shared secret key. More specifically, there is $L_2 C_{L_1}$ combination to specify L_1 out of $L_2 (\geq L_1)$ time slots. Hence, a shared secret key with the length of $n = \lceil \log_2 L_2 C_{L_1} \rceil$ bits are used for the authentication between the BS and the user.

Compared with the conventional public-key-based cryptographic methods, the proposed scheme does not require the rounds of the system setup, key generation, distribution, refreshment, or revocation, as well as the presence of a third-party certificate authority. Different from the physical-layer authentication schemes using distortion characters [21], the proposed scheme provides robust continuous authentication without any continuous parameter update. Furthermore, the proposed scheme enables continuous authentication by simply checking the active time slots without generating or verifying the secret keys periodically, hence achieving lightweight authentication. In the proposed framework, the user sends data to the BS in a grant-free manner. Under the presence of uncorrelated fading, spoofers cannot send data in the same time slots as those of the legitimate user, which are activated by the secret key in our scheme¹.

Fig. 2 illustrates a successful case of authentication based on active time-slot allocation. A spoofer's attack may be successful only when the spoofer instantaneously specifies all the active time slots per frame.

III. PERFORMANCE RESULTS

In this section, we provide our performance results to characterize the proposed scheme. The channel model between the BS and the user considered in our simulations is illustrated

¹Note that the previous SKG schemes of [13] and [14] are readily applicable to our physical-layer authentication framework, while the such application has not been provided in the literature. The detailed investigations are beyond the scope of this letter and are left for future studies.

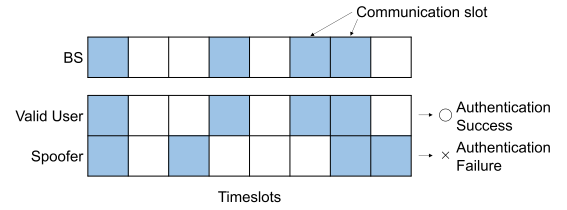


Fig. 2. Active time-slot allocation based on a generated secret key.

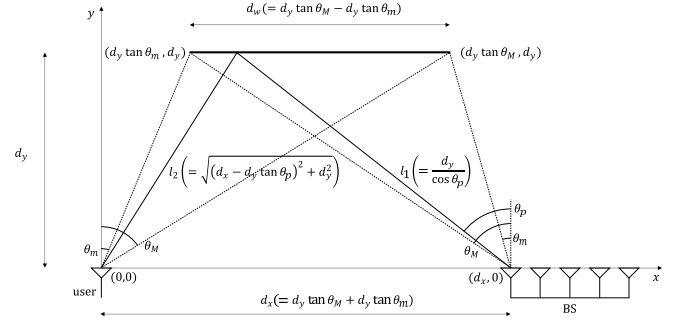


Fig. 3. Channel model employed in our simulations.

TABLE I
BASIC PARAMETERS USED IN THE SIMULATIONS

| | |
|--|-----------------------------------|
| Uplink carrier frequency f_U | 2.4 GHz |
| Downlink carrier frequency f_D | $[f_U, f_U + 80 \text{ MHz}]$ |
| Number of antenna elements at BS M | 16 |
| Antenna spacing at BS d | 75 mm |
| Channel | frequency-flat Rayleigh fading |
| Number of paths P | 20 |
| Attenuation of p th path | $\alpha_p = (\lambda/4\pi l_p)^2$ |
| Phase rotation of p th path | $\phi_p \in [0, 1\pi]$ |
| Distance d_y | 200 m |
| Angles (θ_m, θ_M) | $(42.5^\circ, 47.5^\circ)$ |
| Average SNR $P_{\text{tx}} E[\ \mathbf{h}(f_D)\ ^2] / M\sigma^2$ | 25 dB |
| Number of symbols per frame L_2 | 16 |

in Fig. 3. For simplicity, scatterers are positioned in a line along the x -axis, which reflects a wave with no amplitude attenuation or phase rotation. In our simulations, P scatterers are uniform-randomly selected from the line per channel generation of $\mathbf{h}(f_U)$ and $\mathbf{h}(f_D)$. The y -axis distance from the scatterers to the BS, as well as that from the scatterers to the user, is d_y , as shown in Fig. 3. Note that θ_m and θ_M are the maximum and minimum angles of $\theta_p \in [\theta_m, \theta_M]$. The distance between the user's antenna element and the closest BS antenna element is given by $d_x = d_y \tan \theta_M + d_y \tan \theta_m$. The length of the scatterers d_w is given by the angular spread $\Delta\theta = \theta_M - \theta_m$ and the distance d_y as follows: $d_w = d_y \tan \theta_M - d_y \tan \theta_m$. The edges of the scatterers are given by $(x, y) = (\tan \theta_m, y_d)$ and $(\tan \theta_M, y_d)$. Moreover, similar to most previous studies, we assume the absence of a BS-to-user direct link due to blockage by an obstacle, and hence the generated channel tends to obey Rayleigh fading.

As also listed in Table I, the basic system parameters are set as $M = 16$, $d = 75$ mm, $f_U = 2.4$ GHz, $\Delta f = f_D - f_U = 0, 20, 40, 60, 80$ MHz, $P = 20$, $\tau_p = l_p/c$ s, $\phi_p \in [0, 2\pi]$, $\theta_m = 42.5^\circ$, $\theta_M = 47.5^\circ$, $P_{\text{tx}} E[\|\mathbf{h}(f_D)\|^2] / M\sigma^2 = 25$ dB, and $d_y = 200$ m, which are used in our simulations unless otherwise noted.

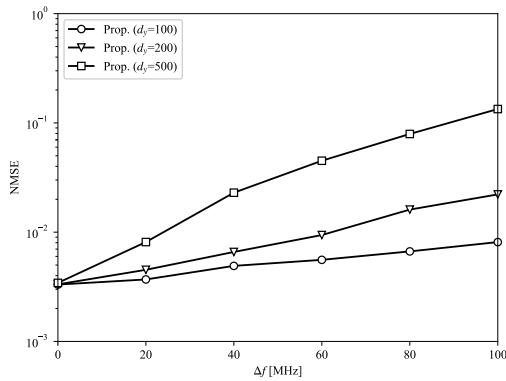


Fig. 4. NMSE of an estimated downlink channel vector at the BS.

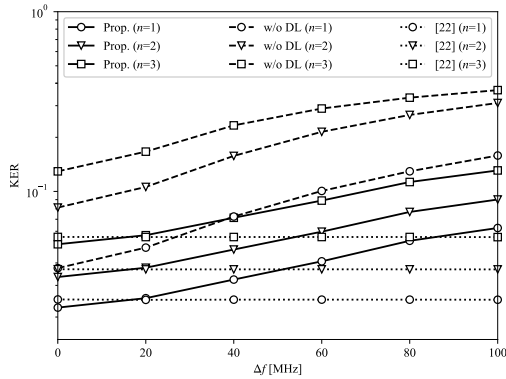


Fig. 5. KER between the secret key generated at the BS and that of the user, where we considered $n = 1, 2$, and 3 . The energy-based SKG and the PASKey scheme were employed as the benchmarks.

A. Performance of Deep-Learning-Based Channel Estimation

In our DL-based estimation of the downlink channel vector at the BS, each input and output layer has 32 nodes, where we train the network to predict $[\Re[\mathbf{h}(f_D)]^T \Im[\mathbf{h}(f_D)]^T]$ from $[\Re[\mathbf{h}(f_U)]^T \Im[\mathbf{h}(f_U)]^T]$. The nodes in the hidden layer are set as (64, 128, 64), and the nodes of the five layers are represented by (32, 64, 128, 64, 32). We employed the ADAM algorithm [20] with a learning rate of 0.001 for optimization and 100 epochs. The training data were collected with $T = 512$ in (7) and (8).

Fig. 4 shows the normalized mean-square error (NMSE) of the estimated downlink channel vector at BS, which is defined as follows:

$$\text{NMSE} = E \left[\frac{\|\hat{\mathbf{h}}(f_D) - \mathbf{h}(f_D)\|_2^2}{\|\mathbf{h}(f_D)\|_2^2} \right]. \quad (10)$$

Here, the distance d_y was set as 100 m, 200 m, and 500 m. Observe in Fig. 4 that even for a high Δf , accurate channel estimation was achievable, especially for a low distance d_y .

B. SKG Performance

Fig. 5 shows the key-error ratio (KER) between the generated secret key at the BS and that of the user. We considered the two benchmark schemes, i.e., the energy-based SKG scheme without DL and the pilot assistant secret key generation (PASKey) scheme [22] that relies on amplified feedback.

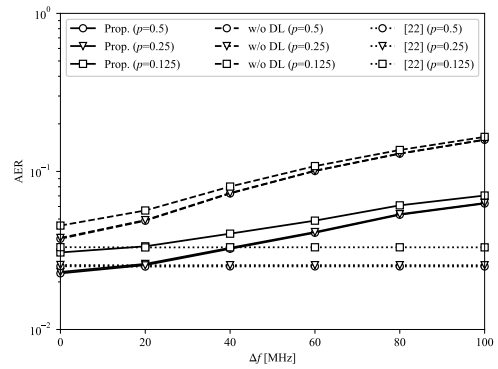


Fig. 6. AER comparisons between the proposed scheme, the energy-based SKG scheme, and the PASKey scheme, where active time-slot ratio was given by $p = 0.125, 0.25$, and 0.5 .

More specifically, in the energy-based SKG scheme, the amplitudes of the downlink channel coefficients are regarded as the amplitudes of the uplink ones, which are estimated from the pilot signals transmitted from the user. Then, the estimated amplitude is quantized to generate a secret key. Furthermore, in the PASKey scheme, additional amplified pilot feedback from the user allows us to estimate the downlink channel coefficient in a stable manner while suffering from the doubled latency, the noise amplification of the feedback signal, as well as information leakage to an eavesdropper, unlike the proposed scheme.² As shown in Fig. 5, upon decreasing the generated key length n per channel coefficient and the frequency difference Δf , the KER improved. The proposed scheme outperformed the energy-based benchmark scheme without DL, where the performance advantage increased with the increase of Δf . The KER of the idealistic PASKey scheme remains unchanged regardless of Δf while suffering from information leakage to the eavesdropper, as well as the increased SKG latency. Note that while information reconciliation with channel coding and privacy amplification with a hash function is typically implemented to improve the reliability of SKG [11], we considered only a channel-uncoded scenario for simplicity.

C. Authentication Performance

Fig. 6 shows the authentication-error ratio (AER), where an authentication error is counted when all the time slots randomly generated by a spoofer match those activated in the proposed scheme and when the generated secret key at the BS does not agree with that of the user. We assumed that the spoofer knows the ratio of the activated time slots over the time slots per frame $p = L_1/L_2$, where the ratio was set as $p = 0.125, 0.25$, and 0.5 . The quantization level was given by $n = 1$. Observe in Fig. 6 that the proposed scheme exhibited benefits similar to those shown in Fig. 5 while maintaining

²To elaborate a little further, in the PASKey scheme, each of the BS and the user transmits pilot-related overhead twice higher than in the conventional SKG and the proposed scheme. Furthermore, even for the FDD scenario, the BS and the user have to transmit the pilot in a different time slot in the PASKey scheme. Hence, the latency imposed by the SKG of the PASKey scheme is approximately four times higher than that of the proposed scheme. This may result in increased performance degradation over the practical time-varying channel.

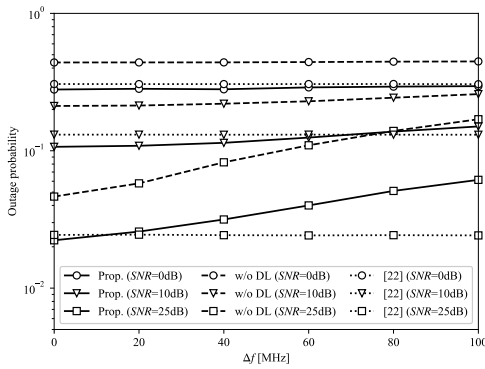


Fig. 7. Outage probability comparisons between the proposed scheme, the energy-based SKG scheme, and the PASKey scheme, where the system parameters of $(n, p) = (1, 0.5)$ were employed while considering the receive SNRs of 0 dB, 10 dB, and 25 dB.

the lower overhead and latency in comparison to the PASKey scheme.

D. Communication Performance

Fig. 7 shows the outage probability of the proposed scheme, which is affected by either authentication or data detection. More specifically, an unsuccessful event for data detection is induced when at least one symbol in each frame is mis-detected at the receiver due to the effects of fading, AWGNs, and channel estimation errors. Also, the definition of authentication error is the same as that used in Section III-C. We considered 16 symbols in each frame while we set $p = 0.5$ and $n = 1$ while the average SNR was given by 0 dB, 10 dB, and 25 dB. The modulation scheme was quadrature phase-shift keying, and the ZF algorithm was used for demodulation. The other parameters are the same as those used in Fig. 6. As shown in Fig. 7, the outage probability improved upon decreasing Δf while outperforming the energy-based SKG scheme in each scenario. More specifically, the proposed scheme's performance benefits increased upon decreasing the receive SNR.

IV. CONCLUSION

In this letter, we proposed DL-based physical-layer channel estimation and lightweight authentication in a non-reciprocal FDD channel. In our scheme, the downlink channel is estimated from the uplink pilot symbols based on DL at the BS, hence reducing the feedback information, the delay, and the information leakage to an eavesdropper, in comparison to the conventional SKG assuming the reciprocal channel. Each legitimate user is authenticated when a user transmits data in the time slots allocated by the physical-layer secret key. Our performance results demonstrated that our authentication functioned while achieving lower latency and error rates in an FDD fading channel than the conventional energy-based benchmark scheme.

REFERENCES

- [1] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.
- [2] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 11891–11915, Aug. 2021.
- [3] A. Hameed and A. Alomary, "Security issues in IoT: A survey," in *Proc. Int. Conf. Innov. Intell. Informat., Comput., Technol. (3ICT)*, Sep. 2019, pp. 1–5.
- [4] Y.-S. Shiu, S. Chang, H.-C. Wu, S. Huang, and H.-H. Chen, "Physical layer security in wireless networks: A tutorial," *IEEE Wireless Commun.*, vol. 18, no. 2, pp. 66–74, Apr. 2011.
- [5] R. Nakai and S. Sugiura, "Physical layer security in buffer-state-based max-ratio relay selection exploiting broadcasting with cooperative beamforming and jamming," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 2, pp. 431–444, Feb. 2019.
- [6] S. Sugiura, "Secrecy performance of eigendecomposition-based FTN signaling and NOFDM in quasi-static fading channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5872–5882, Sep. 2021.
- [7] A. J. Menezes, T. Okamoto, and S. A. Vanstone, "Reducing elliptic curve logarithms to logarithms in a finite field," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1639–1646, Sep. 1993.
- [8] Y. Wang, B. Ramamurthy, and X. Zou, "The performance of elliptic curve based group Diffie–Hellman protocols for secure group communication over ad hoc networks," in *Proc. IEEE Int. Conf. Commun.*, vol. 5, Jun. 2006, pp. 2243–2248.
- [9] N. Xie, Z. Li, and H. Tan, "A survey of physical-layer authentication in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 1, pp. 282–310, 1st Quart., 2021.
- [10] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. Wong, and X. Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 679–695, Apr. 2018.
- [11] H. Fang, X. Wang, N. Zhao, and N. Al-Dahir, "Lightweight continuous authentication via intelligently arranged pseudo-random access in 5G-and-beyond," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4011–4023, Jun. 2021.
- [12] C. Huang, G. C. Alexandropoulos, A. Zappone, C. Yuen, and M. Debbah, "Deep learning for UL/DL channel calibration in generic massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.
- [13] Y. Yang, F. Gao, G. Y. Li, and M. Jian, "Deep learning-based downlink channel prediction for FDD massive MIMO system," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 1994–1998, Nov. 2019.
- [14] X. Zhang, G. Li, J. Zhang, A. Hu, Z. Hou, and B. Xiao, "Deep-learning-based physical-layer secret key generation for FDD systems," *IEEE Internet Things J.*, vol. 9, no. 8, pp. 6081–6094, Apr. 2022.
- [15] R. Perlman, "An overview of PKI trust models," *IEEE Netw.*, vol. 13, no. 6, pp. 38–43, 1999.
- [16] P. L. Yu, J. S. Baras, and B. M. Sadler, "Physical-layer authentication," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 38–51, Mar. 2008.
- [17] C. T. Zenger, M. Pietersz, J. Zimmer, J.-F. Posielek, T. Lenze, and C. Paar, "Authenticated key establishment for low-resource devices exploiting correlated random channels," *Comput. Netw.*, vol. 109, pp. 105–123, Nov. 2016.
- [18] J. Zhang, A. Marshall, and L. Hanzo, "Channel-envelope differencing eliminates secret key correlation: LoRa-based key generation in low power wide area networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 12462–12466, Dec. 2018.
- [19] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. 14th Int. Conf. Artif. Intell. Statist.*, 2011, pp. 315–323.
- [20] C. Trabelsi et al., "Deep complex networks," in *Proc. Int. Conf. Learn. Represent.*, Vancouver, BC, Canada, 2018, pp. 1–19.
- [21] Y. Liu, H. Chen, and L. Wang, "Physical layer security for next generation wireless networks: Theories, technologies, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 347–376, 1st Quart., 2017.
- [22] D. Qin and Z. Ding, "Exploiting multi-antenna non-reciprocal channels for shared secret key generation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 12, pp. 2693–2705, Dec. 2016.