Distributed Physical Layer Authentication: Overview and Opportunities

Tianhui Zhang, Yan Huo, *Senior Member, IEEE,* Liran Ma, *Member, IEEE,* and Ela Guo

***Abstract*—Distributed physical layer authentication (DPLA) is a novel authentication framework, which not only exploits the collaborative computing of multiple devices to enhance overall efficiency, but also alleviates the degradation of processing performance caused by resource-constrained terminals. It is considered a promising architecture for solving access security issues in future communications. Considering DPLA's potential, in this article, we review existing DPLA schemes to provide a comprehensive summary of the strategies and technical approaches adopted during each implementation stage. Our simulation results show that the voting-assisted DPLA scheme has better authentication performance than the centralized PLA. In addition, we also present some open research issues on DPLA, addressing new opportunities ahead and potential research directions.**

I. Introduction

AUTHENTICATION plays a crucial role in securing wireless communications systems and applications as it is the basis for most types of access control. Typically, an authentication scheme depends on a type of authenticator (e.g., secrets), which is a means used to confirm a claimed identity. For example, a memorized secret (e.g., password) is a common authenticator. Another popular authenticator for networked systems is a secret key managed by a cryptographic system (e.g. asymmetric cryptography). As new wireless communications systems, such as the Internet of Things (IoT), emerge with technology advancements (e.g., 5G), these common authenticators may no longer be effective due to expensive management or computational overhead [1].

Information-theoretic security based authenticators, therefore, become more attractive as they are more suitable for resource-constrained devices and delay-sensitive applications in IoT. This type of authenticator, called a physical layer authenticator (PLA), exploits natural features of wireless links such as channel reciprocity, randomness, space-time uniqueness for authentication. PLA authenticates by comparing extracted physical layer features of a user's signal to a previously defined threshold of features unique to that user. The new signal must be consistent with the features to be accepted. In this way, PLA fulfills the integration of communications and authentication by taking advantage of the signal itself, effectively reducing additional network resource consumption.

Broadly speaking, there are two categories of PLA: centralized and decentralized. Centralized PLA (CPLA) relies on a single device to accomplish the entire authentication process including channel information collection, processing and identification. Thus, CPLA schemes are faced with the risk of single-point failure at cluster-head [2], which is fairly common in highly dynamic and unreliable communication scenarios. Decentralized PLA (DPLA), on the other hand, relies on several collaborative nodes, which evaluates authenticity as follows: i) Share the tasks of channel sensing, data processing or message identifying from the central node to a group of candidate edge nodes; ii) Select non-malignant, reliable and self-confident candidate nodes acting as real collaborators in a micro-alliance, according to their respective channel quality, identification ability or historical reputations, etc; iii) Formulate a unified authentication decision in this micro-alliance with the help of collaborative communications, model parameter interaction and decision-level fusion among multiple trusted edge devices. DPLA's authentication process has several advantages over CPLA's process [3]. In specific, multi-directional channel sensing, regarded as distributed acquisition, provides DPLA a more robust physical layer dataset with spatial diversity that attackers can hardly imitate. Multi-node aided data processing and identifying, regarded as distributed training, brings DPLA better real-time performance and lowers computing burden on individual device [4]. Multiparty involved consensus formation, regarded as decision fusion, offers DPLA a more comprehensive authentication result with higher accuracy. Through personalized task assignment of collaborators and decision fusion at a coordinator, DPLA can make authentication decisions more accurate, robust and efficient.

To the best of our knowledge, this is the first work that gives a broad overview of current DPLA technology and literature. We first review the theoretical foundations of PLA, and summarize the limitations of CPLA to derive the motivations for DPLA studies. Then, we explain the DPLA frameworks in detail, expounding strategies, algorithms, and techniques used in each stage of setting up a DPLA system. We also present some preliminary comparison results to verify the effectiveness of the mentioned weighted voting-assisted DPLA scheme compared to a classical CPLA scheme. And finally, we discuss several open research issues by considering new perspectives and conclude the article.

II. Overview of PLA

In this section, we first review the theoretical basis and most commonly used mechanisms of PLA. Then, we summarize the limitations faced by CPLA and conclude with the motivations for DPLA research.

Manuscript received XX XX, 2021; revised XX XX, 2021. This work was partly supported by the Fundamental Research Funds for the Central Universities (Grants #2019JBZ001) and the Science and Technology Innovation Project of Xiongan (Grant #2022XACX0400).

Corresponding author: Yan Huo

A. PLA Basics

Considering a typical PLA system as shown in Fig. 1, Bob receives several messages, including legitimate messages from Alice and fabricated messages from Eve, which impersonate Alice's signals. Once a new message arrives, Bob authenticates its origin by analyzing the latest extracted physical layer features. Based on certain testing-decision mechanisms described in the next subsection, only messages with similar physical features to legitimate users will be received, while others will be denied. Since Eve's signals have different physical layer features from Alice's, his/her message is rejected, maintaining the integrity of the system.

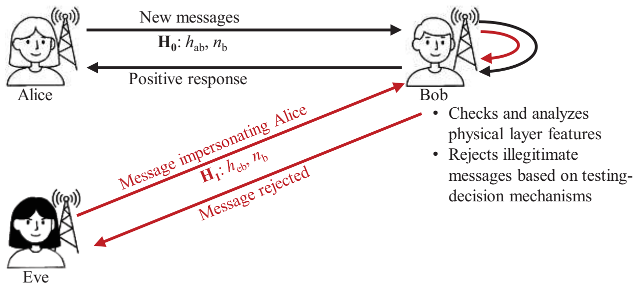
New messages
$\mathbf{H}_0$: h_{ab}, n_b
Positive response
Alice
Bob
• Checks and analyzes physical layer features
• Rejects illegitimate messages based on testing-decision mechanisms
Message impersonating Alice
$\mathbf{H}_1$: h_{eb}, n_b
Message rejected
Eve

Fig. 1: A typical PLA framework.

In this way, PLA accomplishes light-weight authentication by using the signal itself. This is due to the fact that physical layer features are spatially uncorrelated between different geographic locations. Therefore, as long as the distance between the legitimate user and the attacker is greater than a half wavelength, PLA can effectively differentiate the signals [5].

B. Authentication-Decision Mechanisms

Two types decision mechanisms are exploited in PLA. One method, named as the threshold-based mechanism, follows the Neyman Pearson lemma to systematically construct a binary hypothesis test, i.e., $\mathrm{T}(x_n) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \delta_n^*$, where $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the case of messages from a legitimate user and a malicious node, respectively. To obtain an optimal decision boundary δ_n^* in the method, it is guaranteed to minimize the Type-2 error (i.e., the probability of missed detection) for a given Type-1 error (i.e., the probability of false alarm). Meanwhile, the test statistic $\mathrm{T}(x_n)$ is formulated according to channel difference measurements such as the likelihood ratio test.

The other method is designed based on the machine learning (ML) technology. It models PLA as a classify issue to find out an optimal segmentation plane for identity signatures by training the parameters of the neural network, thus achieving threshold-free. Different ML methods are applied in PLA with different emphases. For example, deep neural network (DNN) has stronger fitting and classification capabilities than other algorithms, which can assist PLA to make decision more accurate, while support vector machine (SVM) can provide better authentication performance for PLA on limited offline datasets. Unsupervised ML methods, such as K-means clustering, k-nearest neighbor (k-NN) algorithm, etc. have also been applied to aggregate multiple sampled instantaneous physical layer features into clusters, which can effectively combat estimation errors or unreliable judgment caused by dynamic fluctuation of physical channels.

C. Limitations of CPLA

Although CPLAs have been able to guarantee the legitimacy of access in most cases, all of them rely on a single device to accomplish the entire authentication process from channel acquisition, data processing to final decision-making. However, the emergence of new attacks, the limitation of network resources and the highly dynamic nature of wireless channels all lead to the fact that single-point PLA may not be sufficient to secure the system. The limitations are summarized as follows. Firstly, intelligent PLA-aware attackers can evade detection through power manipulation and spatial position optimization [6], resulting in a significant decline in CPLA's detection performance. Secondly, in light of the demand for IoT massive connections, the workload of the central processor has increased dramatically, leading to higher processing latency and difficulty in training the complex neural network independently [4]. Thirdly, imperfect estimation of physical layer features due to dynamic interference or user mobility in fast changing environments will cause single-point failure of the centralized system [1]. As a result, the robustness of authentication is seriously reduced. Moreover, we can expect that the security of certification solely relying on one device is often more limited than that of cooperative authentication. These limitations have motivated researchers to look for a better authenticator with higher safety and credibility.

D. Motivations of DPLA

DPLA brings new opportunities to overcome the shortcomings of CPLA. The idea was first introduced in [7], where the DPLA scheme was investigated for a simple scenario containing a single legitimate transmitter and a spoofing attacker. Multiple supervised nodes (edge devices) located at different spatial position are involved to complete the authentication collaboratively. As such, the edge computing assisted DPLA not only improves the authentication performance without increasing the computational burden on each device, but also allows tasks to be tailored to each node's computing power and properties, which complements the diversity and optimization that newer communication scenarios require. Moreover, the distributed architecture provides natural scalability, better anti-attack capability and higher robustness, since the dynamic joining, leaving or failure of a single point will not bring significant impact on the entire authentication system. And as confirmed in [6], DPLA effectively mitigates the degradation of detection performance when subjected to PLA-aware attacks, thereby overcoming the vulnerability of CPLA.

E. Respective Application Scenarios

Both CPLA and DPLA are employed as aided schemes for scenarios where traditional upper-layer protocols are hard to be deployed or limited in efficiency. However, the choice between them is mainly dependent on the network topology. The authentication mode of CPLA is single-to-many, so only a central processor with powerful computing capabilities can guarantee the reliability of identification. CPLA is more applicable for authentication of small-scale user groups that are relatively geographically concentrated. And such schemes often require less extensibility and environmental adaptability. DPLA completes the authentication in a many-to-many manner, and the security risk of the system is spread out over each collaborative peers, greatly reducing the possibility of authentication failures. Also, the parallel processing capabilities and more efficient data sharing brought by its branch/mesh topology enable DPLA to better cope with large-scale concurrent user access. Scenarios such as industrial wireless edge networks (IWENs), wireless sensor networks (WSNs) and so on. The common point of these networks is that the available edge peers they contain are numerous, and the number of packets they need to handle are tremendous. Furthermore, DPLA can be competent for scenarios that require greater flexibility and scalability in certification. For example, in a distributed topology, peer nodes i, j, k have already authenticated User A, while others still have no knowledge of him. The certification results of i, j, k could be passed to other neighbors if the interaction among peers is close and based on trust. That way, authentication may become transitive. And this potential ability will make DPLA more competitive than CPLA in such self-organizing networks.

III. Discussion of A DPLA Framework

Current DPLA frameworks can be categorized into three types, each with a different level of distribution. The first one can be called as a semi-distributed PLA framework. The authentication decision is still made directly by the central node. The framework employs collaborators only as perception points, and the raw physical layer features are uploaded for combination. The second one is referred to as decision-level fusion based DPLA. This framework allows collaborators to evaluate each incoming packet preliminary. The information forwarded to the central node is the local outputs/opinions gleaned from collaborative peers. Acting as a fusion center (FC), the central node coordinates inconsistent opinions on authentication of the same packet for final decision-making. The third one is a fully-distributed PLA, with no coordinated node at all. This FC-free DPLA framework is more stable as damage to any node has minimal impact on system capability. Also, it can greatly enhance the network flexibility and scalability [8]. With these advantages, it is believed as an attractive model for collaborative authentication. As of now, the only study in [9] that introduced this mode is a secure physical layer voting scheme where each participant serves as a decentralized leader, and independently tallies the votes generated by other members to compute the final outcome. The interaction continues until all participants end up with the same voting estimate.

Next, we expound algorithms and techniques in each implementation stage of DPLA, including distributed acquisition, distributed training, collaborator selection, and final decision formation, which are illustrated in Fig.2.

A. Distributed Acquisition

Distributed acquisition refers to a set of radio-heads equipped with multiple antennas capturing channel state information (CSI) from different geographic locations. The pioneering work of [5] proposed a logistic regression-based DPLA strategy by taking advantage of several randomly deployed landmarks to enhance the spatial resolution and to improve spoofing detection rate. In view of this work, a few early studies have been carried out focusing on the performance enhancements brought by distributed acquisition. Yet, these studies did not consider the fact that randomization of acquisition locations would reduce the similarity of physical layer feature, thereby increasing divergence among collaborators or even leading to the failure to reach a consensus. Recently, Wang et al. put forward a horizontal federated learning (FL) aided DPLA scheme by scheduling multiple trusted edge devices to jointly accomplish channel sensing and authentication in [10]. Different from previous studies, the collaborators are no longer randomly scattered, but closely surround the central coordinator. This ensures more feature overlap and ultimately facilitates cooperative learning based on sample associations.

In essential, the utilization of multi-directional perception would enhance the robustness of authentication to perturbations, since sharing of observations among collaborators will compensate for uncertainties or imperfect measurements on a single isolated node. Also, it greatly increases the difficulty for an attacker to successfully replicate channel information of legal users. Thus, DPLA realizes improved security compared to CPLA.

B. Distributed Training

Distributed training aims to relieve the excessive computational load of a centralized control system. Its studies focus on sample segmentation and information interaction/parameter exchange among collaborative peers. Sample segmentation is to find an optimal strategy for appropriately splitting large amounts of datasets into sub-samples. A modified grouping method based on downsampling proposed in [4] has been proved to display better sub-classification results while reducing the data volume. Efficient information exchange is for immediate updates, accelerated model convergence, and to complement each other. Xiao et al. constructed a DPLA scheme in [11] applying distributed Frank-Wolfe (d-FW) algorithm to solve the coefficient estimation problem of the complex ML-based model. The main contribution lies in the introduction of a data sharing mechanism between collaborators. It reduces communication and computational costs of the distributed architecture, thereby enabling online authentication. In [12], the authors presented an autonomous collaborative PLA framework compatible with FL technology. This approach is more privacy-preserving that allows multiple independent collaborators with individual physical layer observations to establish a shared authentication model, without leaking their underlying data among peers. In this way, malicious behaviors of stealing and abusing sensitive physical data will be completely eradicated.

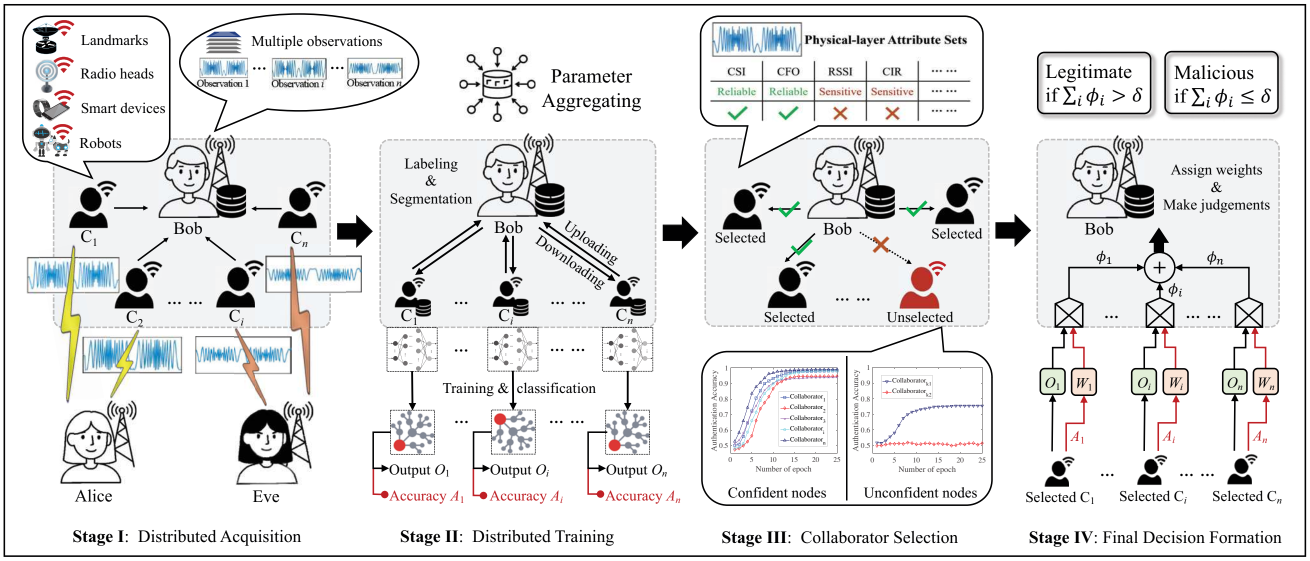

Fig. 2: A DPLA framework.

In brief, distributed training can handle large amounts of training data that is formidable for a single resource-limited device in a parallel manner. Moreover, it achieves substantial enhancements on identification accuracy. Note that the extra communication overhead due to frequent information exchange may cause spectrum shortage and high latency. Hence, how to improve efficiency through the optimal design of data sharing policy calling for further exploration.

C. Collaborator Selection

Collaborator selection is to filter non-malicious devices, with adequate channel quality, strong computing power, and good reputation, as candidates for use. Specifically, the filtering mechanism includes unreliable features elimination and unconfident collaborators cancellation.

For one thing, some physical layer features are sensitive to changeable topology, especially in high-speed mobile scenarios. These features are considered unreliable since authentication based on variable features may lead to unexpected errors. In [2], Wang et al. constructed a situation-aware DPLA customization algorithm in a UAV scenario. In the algorithm, they utilized Gini impurity as an evaluation indicator to measure the reliability of different physical layer features. Accordingly, only the UAVs with reliable feature observations are eligible to participate in following authentication. For another, some malicious or hostile nodes may interfere with authentication progress by injecting misleading information. An effective solution is to introduce authority constraints by computing trust degree and node credibility. In [13], the authors presented a novel DPLA strategy with cluster-head safeguarding mechanism. In the strategy, each UAV is assigned a trust value weighted by detection error level and sensitivity level to represent its own identification capability. The node will be graded as an unconfident/untrustworthy collaborator and removed from the network when its trust does not reach the desired threshold.

In short, a collaborator selection mechanism can customize the suitable feature combination on each trusted node, increase the overall system stability, and further minimize the computation overhead of DPLA.

D. Final Decision Formation

Final decision provides the certification result. For the second and third DPLA frameworks, a voting fusion mechanism can be introduced as an effective means to integrate the outputs of different collaborators for more robust decision results. The goal of it is to maximize the contribution of high-quality local outputs while minimizing the negative impact of erroneous local outputs on the final decision. Existing voting fusion mechanisms include unweighted voting and weighted voting.

For the unweighted voting mechanism, the most straightforward idea is the all-accept/-reject method, which denotes that a message is either accepted if all collaborators outputs are legitimate or rejected if all are illegitimate. Another method is plurality voting. It accepts or rejects messages based on popular votes. However, plurality voting ignores gaps of authentication capacity among collaborators. Even those with less identifying capability have the same voting power as more engaged collaborators. To address this challenge, the authors in [14] predefined appropriate weights for each voter, named as the weighted voting mechanism, which encompasses a variety of forms. For example, simple weighted voting (SWV) requires outputs of all voters to be weighted according to their estimated authentication accuracy. Weighted majority voting (WMV) maximizes the overall accuracy of decision-making by assigning weights in the form of Logit. Re-scale weighted voting removes the ineligible nodes and scales weight values of qualified nodes proportionally. Best-worst weighted voting (BWWV) defines the authorities of the best and worst classifier as 1 and 0, and linearly grades the weights of others.

In a word, decision-level fusion can obtain a more comprehensive certification result by coordinating multi-party opinions, while reducing the additional communication overhead caused by frequent parameter exchange.

IV. Case Studies

This section validates the effectiveness of the distributed framework by comparing the authentication performance of the single-point CPLA and the voting-assisted DPLA.

We consider a DPLA system composed of seven edge collaborative devices ($C_1 \sim C_7$) that placed around the central receiver (Bob). The legitimate user (Alice) intends to communicate with Bob, while the coexisting spoofer (Eve) who impersonates Alice's identity also attempts to muddle through. To be more realistic, we assume that the collaborators are not completely credible/confident. For example, C_4 is suffered from severe extra noise, and C_5 is a hidden internal attacker that seeks to disrupt the decision formation by arbitrarily flipping its local outputs. Based on this, a collaborator selection (CS) strategy is adopted to cancel poorly performing or malicious collaborators, thereby retaining those that yield positive contributions. Specifically, each collaborator is assigned a trust value weighted by false alarm rate and missed detection rate. If the trust is lower than a predefined threshold, the node will be considered invalid and discarded from the micro-alliance. The outputs of selected collaborators are eligible for fusion. Both plurality voting and weighted majority voting are applied for performance analysis.

The CSI is used as the identifying signature for PLA. We build a typical Rayleigh fading channel to simulate indoor communications, which may contain a large open space with a few scattered obstacles. The channel coefficient is affected by both the large-scale and small-scale propagation effects, which are modeled as the log-normal shadow fading and the multi-path flat Rayleigh fading, respectively. Then, the static physical datasets are generated on a single-input-multiple-output (SIMO) orthogonal-frequency-division-multiplexing (OFDM) transmission system, where least squares (LS) algorithm is applied for channel measurements. Detailed communication parameter settings follow the IEEE 802.11a WLAN standard mentioned in [10]. Besides, collaborators do not assume the responsibility for making final decisions, so the system has no strict requirements on their precision. The ML-based network structure at the collaborative devices can be relatively simplified to ensure that their computational burden is not overwhelming. Therefore, the back propagation (BP) neural network, which has faster learning speed, less computation and higher parallelism, is employed as the sub-classifier. It consists of 3 hidden layers with 20, 20 and 10 neurons, sequentially.

Fig.3 shows the authentication performance of the seven local collaborators and their corresponding geographical locations. To be specific, the final accuracy rates they achieved are 98.17%, 94.71%, 92.58%, 75.46%, 51.45%, 98.09% and 99.14%. Note that the gaps in identification capacities are due to the fact that the location-aware sensitivity of each collaborator are quite different. The location-aware sensitivity can be defined as the distance difference between Alice to a collaborator and Eve to it. The greater it is, the higher the resolution of channel characteristics will be, therefore, more accurate identification could be realized. Besides, the performance of C_4 and C_5 is significantly inferior, resulting in poor trust values assigned by Bob. That way, unconfident collaborators with unreliable channel observations and internal attackers can be detected effectively.

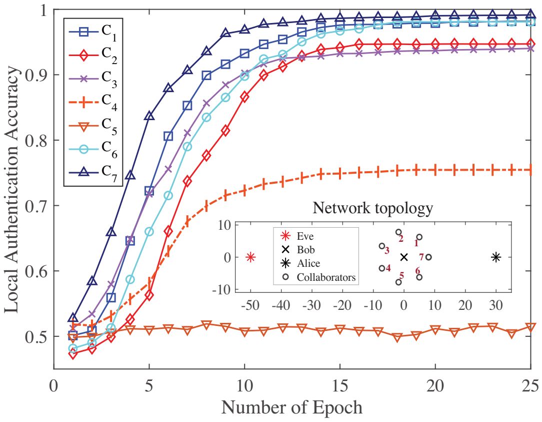

Fig. 3: Local authentication performance of each collaborator.

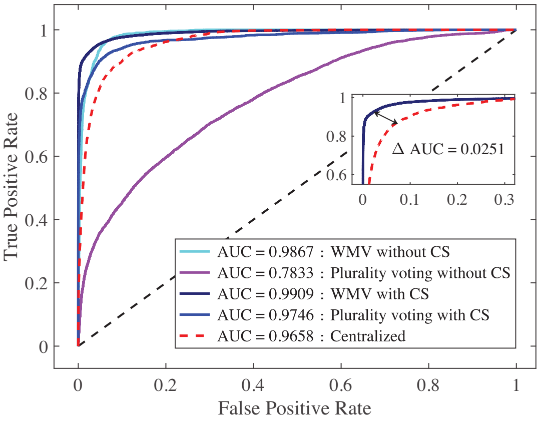

Fig. 4: ROC curve of different authenticators.

After generating local outputs and experiencing the CS, the final decision is made among selected edge devices by voting. To evaluate the effect of authenticators, the receiver operating characteristic (ROC) curve with the area under the curve (AUC) is introduced as a measure. The ROC plots the true positive rate (TPR) against the false positive rate (FPR) to illustrate the performance of a binary classier with different thresholds, which graphically reflects the correlation between sensitivity and specificity.

Fig.4 demonstrates the comparison results between CPLA and DPLA. It can be found that the AUC difference of the WMV-assisted DPLA and the CPLA is 0.0251, which verifies the superiority of distributed cooperative authentication. Moreover, the performance gain introduced by collaborator selection is discussed. We can observe that the WMV-assisted DPLA maintains excellent performance in terms of AUC even without CS, while that of plurality voting decreases dramatically to a worse level than the CPLA. The reason is that the WMV mechanism can weaken the authority of poorly performing collaborators by adaptively lowering their weights,

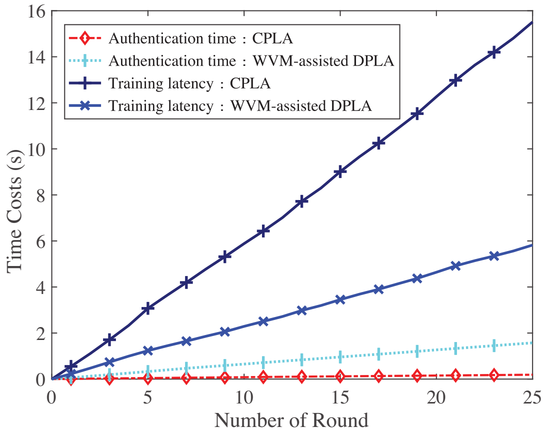

Fig. 5: Time cost of the CPLA and the voting-assisted DPLA at different stages.

thereby mitigating the negative impact of internal attackers on decision-making. This provides DPLA with performance gains similar to collaborator selection. By contrast, the plurality voting mechanism requires CS more to ensure its reliability.

Fig.5 depicts the time cost of executing DPLA and CPLA. Delays among collaborators in each stage of distributed acquisition, training and decision/authentication are considered. For channel acquisition, the time consumption for them is almost the same. This is because the pilot signal is sent as a broadcast, collaborators can capture it simultaneously with Bob. Note that the distributed acquisition inevitably increases the overhead of signature extraction in exchange for enhanced spatial resolution of physical layer features. This concern may be alleviated by striking a balance between performance and cost. For instance, we can set an appropriate budget on the number of participating collaborators, or distribute rewards (e.g. communication resources) to collaborators to compensate their costs. For each iteration, 1200 CSI samples are fed for model updates, and 200 packets are randomly selected from the validation set for identification. As can be seen, the training time of CPLA takes about three times longer than DPLA, which proves the efficiency of distributed training. The latency is significantly reduced because the complexity of neural network and the data volume processed on the cooperative devices are much lower than that of CPLA, while the parallel processing of CPUs applied on edge devices also plays a great role. In addition, it is noted that the single-point CPLA accomplishes authentication faster in milliseconds, while the WMV-assisted DPLA performs slightly poorly. The reason is that the DPLA scheme takes more time to evaluate and assign weights to the local outputs of the sub-classifiers for attaining a unified decision, whereas the CPLA scheme can make judgements directly via an established authenticator.

The results demonstrate that CPLA can achieve faster authentication when the model completes offline training. However, the shortcoming is once the environment changes, the retraining at the central authenticator may be quite time-consuming. It is believed that edge intelligence enabled D-PLAs will be more efficient for relearning, since each node is not heavily burdened. And the little time spent on decision fusion is acceptable in exchange for a higher detection rate of external spoofers.

V. Open Research Issues

In this section, we discuss a few open issues in DPLA with consideration of novel techniques and innovative concepts.

A. *Fully-distributed PLA based on Gossip protocols*

Introducing consensus-based gossip learning (GL) to the authentication procedure is one of the possible methods to achieve fully-distributed PLA. GL, originated from FL, is one of state-of-the-art decentralized machine learning protocols [8]. Different from traditional FL, GL requires no aggregation server or any central component. It addresses the challenges of single point of failure, poor scalability, and weak connectivity in FL, thus enabling higher fault-tolerance and robustness for DPLA systems. Specifically, collaborators in GL collect physical layer features for local training. The collaborators then transmit their model updates to one-hop neighbors and aggregate the parameters received from these neighbors until the network converges to an average consensus state. That way, each collaborator acts as both server and client simultaneously, enabling peer-to-peer communication without infrastructure. In a GL network, any collaborator can make the final decision on authentication, preventing over-dependence on any point. However, this structure may involve extensive model parameters/data exchange among peers, resulting in reduced communication efficiency of the GL protocol. To address this concern, model sparsification technologies and adaptive peer selection mechanisms should be introduced. By removing redundant information and constructing the communication topology adaptively, consensus formation will be accelerated.

B. *Node reputation management against internal attacks*

While DPLA can achieve better security against outer attacks, the utilization of collaborative devices may introduce new security challenges. The faulty or malicious nodes (e.g., vote tampering attacks and Byzantine attacks) hidden in them may inject falsified reports, which will adversely affect the reliability of final decisions, or even lead to a failure to reach consensus. To guard against such internal attacks, node reputation-rating can be introduced into the DPLA design. This will have a positive impact on the confidence of the interactive information. Reputation measures the reliability or credibility of an entity in view of its past behaviors and current performance, thereby eliminating fluctuations caused by a single observation. Numerous reputation models studied in mobile Ad-hoc networks can be naturally migrated to DPLA. For example, the authors in [15] build a Beta reputation system and evaluates the node reputation through clustering-based and distance-based decision rules. The evaluation result is then used as a reference for weighted value of model aggregation in the subsequent P2P communications. Moreover, node reputation should be dynamically updated according to the performance of different time slots, combined with certain accountability mechanisms. Collaborators with low reputation scores need to be punished, while those with good reputation need to be rewarded.

C. Trade-off between security performance and cost

Previous studies on DPLA have lacked a comprehensive analysis of system-level costs, such as total latency, communication overhead, and so on. Although the enhanced detection performance is an attractive benefit of DPLA, it requires additional edge devices for channel acquisition and computing, bringing about extra costs (e.g. bandwidth, energy) to incentivize or pay collaborators. As pointed out in [3], the balance between performance and costs is strongly rely on the deployment scenario, processing architecture, resource allocation and the number of participants involved. To find which deployment framework is a more preferable option, we should compare different DPLA frameworks with varying degrees of distribution, ranging from completely centralized to fully-distributed processing. In addition, jointly optimization issues can be formulated by considering both the resource management and the collaborator selection, such that better a trade-off between authentication performance and consumption can be realized.

D. Deployment challenges for DPLA in practical networks

For the semi-distributed PLA framework, many critical factors need to be considered in the actual implementation, including the optimal design of acquisition locations and the heterogeneity of collaborative devices. By exploring the channel-to-location mapping relations and quantifying the quality of channel observations at different devices, the performance of the DPLA framework can be further improved. For the decision-level fusion based DPLA framework, good coordination of sensing and computing among collaborators is of paramount importance. The difficulty lies in the joint design of communication resource and computation performance, i.e. integrated sensing and computation (ISAC), to schedule collaborators more flexibly, improve resource utilization, and achieve more accurate certification. For the fully-distributed PLA framework, how to organically combine a reputation calculation/delivery mechanism with the interactive authentication process of the GL-based framework is a tricky issue in deployment. Moreover, in real communication networks with frequent data interactions, once malicious nodes occupy a certain proportion or there is a concealed Byzantine data cooperative attack, the results of GL may deviate from the theoretical eventual consistency. In this case, additional security measures other than reputation need to be introduced.

VI. Conclusion

The advantages of DPLA lie in the high robustness brought by distributed acquisition and the efficiency and flexibility brought by distributed training. In this article, we present a DPLA framework, survey various technologies used in each stage of DPLA, and compare authentication performance and time efficiency between CPLA and DPLAs. Finally, we discussed a few open research issues on addressing new perspective and opportunities of feasible DPLA designs.

References

[1] H. Fang, X. Wang, and S. Tomasin, “Machine Learning for Intelligent Authentication in 5G and Beyond Wireless Networks,” *IEEE Wireless Communications*, vol. 26, no. 5, pp. 55–61, 2019.

[2] H. Wang, H. Fang, and X. Wang, “Edge Intelligence Enabled Soft Decentralized Authentication in UAV Swarm,” in *2021 IEEE/CIC International Conference on Communications in China (ICCC)*, 2021, pp. 86–91.

[3] H. Forssell, R. Thobaben, and J. Gross, “Delay Performance of Distributed Physical Layer Authentication Under Sybil Attacks,” in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–7.

[4] F. Xie, Z. Pang, H. Wen, W. Lei, and X. Xu, “Weighted Voting in Physical Layer Authentication for Industrial Wireless Edge Networks,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2796–2806, 2022.

[5] L. Xiao, X. Wan, and Z. Han, “PHY-Layer Authentication With Multiple Landmarks With Reduced Overhead,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1676–1687, 2018.

[6] H. Forssell and R. Thobaben, “Worst-Case Detection Performance for Distributed SIMO Physical Layer Authentication,” *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 485–499, 2022.

[7] A. Mahmood, W. Aman, M. O. Iqbal, M. M. U. Rahman, and Q. H. Abbasi, “Channel Impulse Response-Based Distributed Physical Layer Authentication,” in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, 2017, pp. 1–5.

[8] I. Hegedüs, G. Danner, and M. Jelasity, “Decentralized learning works: An empirical comparison of gossip learning and federated learning,” *J. Parallel Distributed Comput.*, vol. 148, pp. 109–124, 2021.

[9] N. Ghose, B. Hu, Y. Zhang, and L. Lazos, “Secure Physical Layer Voting,” *IEEE Transactions on Mobile Computing*, vol. 17, no. 3, pp. 688–702, 2018.

[10] S. Wang, N. Li, S. Xia, X. Tao, and H. Lu, “Collaborative Physical Layer Authentication in Internet of Things Based on Federated Learning,” in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021, pp. 714–719.

[11] X. Wan, L. Xiao, Q. Li, and Z. Han, “FHY-layer Authentication with Multiple Landmarks with Reduced Communication Overhead,” in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[12] H. Fang, X. Wang, Z. Xiao, and L. Hanzo, “Autonomous Collaborative Authentication with Privacy Preservation in 6G: From Homogeneity to Heterogeneity,” *IEEE Network*, vol. 36, no. 6, pp. 28–36, 2022.

[13] H. Wang, H. Fang, and X. Wang, “Safeguarding Cluster Heads in UAV Swarm Using Edge Intelligence: Linear Discriminant Analysis-Based Cross-Layer Authentication,” *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1298–1309, 2021.

[14] F. Moreno-Seco, J. M. Iñesta, P. J. P. de León, and L. Micó, “Comparison of Classifier Fusion Methods for Classification in Pattern Recognition Tasks,” in *Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR)*, 2006, pp. 705–713.

[15] N. S. Fernando, J. M. Acken, and R. B. Bass, “Developing a Distributed Trust Model for Distributed Energy Resources,” in *2021 IEEE Conference on Technologies for Sustainability (SusTech)*, 2021, pp. 1–6.

Biographies

Tianhui Zhang (21120171@bjtu.edu.cn) received the B.E. degree in Communication and Information Engineering from Beijing Jiaotong University, Beijing, China, in 2021. She is currently a master student in Beijing Jiaotong University. Her research interests include physical layer authentication, distributed machine learning and collaboration techniques.

Yan Huo (yhuo@bjtu.edu.cn) is a Professor with the School of Electronics and Information Engineering, Beijing Jiaotong University. His current research focuses on wireless communications, security and privacy, and signal processing. He is a senior member of the IEEE.

Liran Ma (l.ma@tcu.edu) is a Professor in the Department of Computer Science at Texas Christian University. His current research focuses on cybersecurity in information and intelligent systems, and instruments for effective education on security and artificial intelligence to computer science and STEM majors.

Ela Guo (epguo2004@gmail.com) is a high school student in Dallas. Current interests include data analytics and machine learning applications.