# GALAMC: Guaranteed Authentication Level at Minimized Complexity Relying on Intelligent Collaboration

Huanchi Wang, *Member, IEEE,* Xianbin Wang, *Fellow, IEEE,* He Fang, *Member, IEEE,* and Lajos Hanzo, *Life Fellow, IEEE*

***Abstract*—Conventional centralized authentication techniques based on both digital cryptography and physical-layer attributes are prone to single-point failure due to either compromised digital security keys or an abrupt change in the physical communication environment. Although these particular challenges could be mitigated by the joint use of decentralized authentication and physical-layer attributes, such schemes often exhibit unpredictable performance. Simultaneously, the necessary involvement of multiple parties and the imperfect observation of the physical communication environment can also significantly increase the latency and computational complexity. As a remedy, a decentralized authentication scheme is proposed in this paper to achieve *Guaranteed Authentication Level at Minimized Complexity* (GALAMC) based on the intelligent use of distributed collaboration and available distributive physical-layer attributes. Specifically, we aim for minimizing the complexity of the proposed collaborative authentication process by harnessing the minimum number of collaborative nodes and the selected authentication attributes at each node across the different environments while guaranteeing the required authentication level. The related physical-layer authentication scheme is implemented at each collaborative node where different physical-layer attributes can be selected based on their usefulness which is time-varying. The simulation results demonstrate that our scheme maintains the target level of authentication and it is more immune to sudden environmental changes than the conventional centralized physical-layer authentication scheme. It can also be observed that our proposed scheme can adaptively select the minimum number of collaborative nodes for adaptively minimizing the computational cost.**

***Index Terms*—Collaborative authentication, Decentralized mechanism, Physical-layer attributes, Unmanned Aerial Vehicles (UAVs).**

## I. INTRODUCTION

WITH the proliferation of smart devices and the Internet-of-Things (IoT) enabled applications, secure communications over the wireless infrastructure have become the critical foundation of the increasingly connected society and industries [1]. As an important aspect of communication security provision, authentication is essential for ascertaining the true identity of each communication device and for distinguishing the potential attackers from the legitimate users [2]. This task has become more challenging in wireless networks due to their security weaknesses, such as their open broadcast nature of the wireless communication environment, especially in the face of the rapidly growing computational power of the attackers [3], [4].

Conventional wireless security provision has been achieved by using a centralized server, such as an authentication server, or any trusted third party to provide the required credentials [5]. These centralized servers have numerous benefits, such as predictable overhead, high interoperability and compatibility with different platforms [6]. However, they also face many new challenges. To be more specific, the centralized authentication process can be very inefficient due to the growing complexity and the dynamic nature of the IoT network. Since all devices have to contact a certain entity in the centralized authentication, this trusted entity has to be always available and authentic. This assumption is sometimes unrealistic in many emerging applications, such as high-velocity vehicular networks, Unmanned Aerial Vehicles (UAV) networks, and so on. Furthermore, the centralized security schemes are always susceptible to the single-point failure, where attacks can be launched against the security server to disrupt the operation of the network and the related devices [7]. Ultimately, the conventional centralized authentication schemes is best-effort based due to the inherent weaknesses of conventional digital credential-based security provision, when the digital security keys and passwords are compromised.

To overcome these challenges, decentralized authentication becomes extremely important for improved security provision, where multiple collaborative nodes are engaged jointly for the authentication of the same device. It could become even more robust by addressing the weakness of the commonly used digital security schemes by embedding the physical-layer authentication, where the devices-specific channel and hardware characteristics are observed at the collaborative entities. However, the decentralized authentication schemes usually result in significantly increased computational cost, processing delay,

H. Wang and X. Wang are with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B9, Canada. (E-mails: {hwang589, xianbin.wang}@uwo.ca).

H. Fang is with the School of Electronic and Information Engineering, Soochow University, Soochow 215301, China, and was with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 5B9, Canada (e-mail: fanghe@suda.edu.cn).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K (E-mail: lh@ecs,soton.ac.uk).

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Program under Grant RGPIN2018-06254 and in part by the Canada Research Chair Program.

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/X01228X/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

and communication overhead due to the extra resources and collaboration involved. Additionally, the decentralized authentication resources or physical-layer inputs arriving from some collaborators may in fact become excessive or redundant for satisfying certain security requirements. Hence, it is critical to improve the trade-off between the authentication requirement, authentication performance and the computational complexity under the decentralized environment.

### A. State of the art

Achieving reliable authentication is more challenging in mobile ad hoc networks, e.g., vehicular networks and flying UAV networks. Due to the unpredictable channel conditions and dynamic network topology, establishing and maintaining a reliable connection between a device to be authenticated and the central security server is not always guaranteed in these applications. To elaborate on the related security challenges and the proposed intelligent collaborative authentication techniques, a flying UAV network is considered as an application scenario in this paper.

UAVs play increasingly more important roles in emerging applications. In supporting surveillance and disaster relief, the UAVs can form multiple collaborative groups, also known as UAV swarms, to provide a self-managed flying ad hoc network (FANET) for rapid deployment in different applications [8]–[10]. These UAV swarms usually have a star-topology in which the cluster head (CH) communicates directly with each member UAV within the swarm or with the CHs from other swarms if needed without the help of intermediate nodes. By breaking the UAVs into multiple swarms, the flexibility of each swarm can be increased and the latency within each swarm can be reduced. The provision of security guarantees for UAV swarms can be extremely challenging given their low cost, flexible maneuvering and harsh or even hostile operating environment. Explicitly, they exhibit rapidly evolving network topology change, intermittent connection with the ground station and might be readily discovered by their adversaries [11]–[14]. As shown in Fig. 1, the CH of the FANET can access the ground station and use it as a centralized authentication server. However, the high-power long-distance wireless transmission and the increased latency make the cloud-based authentication schemes less attractive in FANETs. Hence, the on-site resources, which are the resources within the UAV swarm (i.e., using the CH as the central authentication server), should be utilized to support prompt and reliable security provision.

To protect the system from malicious attacks, one of the conventional on-site authentication techniques is the classic cryptography-based centralized authentication scheme [15]. To be more specific, these schemes usually utilize either a symmetric-key such as the Advanced Encryption Standard (AES) or an asymmetric-key such as the Rivest-Shamir-Adleman (RSA) solution relying on a key management center to encrypt the transmission [16]. In the UAV swarm, these digital key-based authentication techniques use the CH as the authentication server and aim for ascertaining the identity of each UAV by verifying the security key. However, these

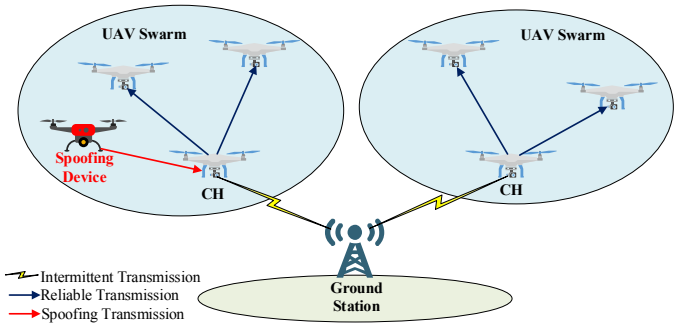


Fig. 1. Adversarial environment in a flying UAV network. The connection between the ground station and the cluster head (CH) can become intermittent which makes it difficult to provide a constant cloud-based centralized security provision. Hence, on-site authentication resources (i.e., CH and member UAVs) should be utilized to perform the authentication.

security provision techniques rely on the storage of secret keys and might become compromised in the face of the adversaries' rapidly increased computational power [17]. Once the digital key is compromised by a brute force attack, it is almost impossible for the central node to verify the true identity of the authentication requester, which catastrophically degrades the security performance [18].

Considering the weakness of the digital security provision, sophisticated physical-layer authentication techniques have been developed to extract the unique hardware and channel characteristics of each device for authentication [19]–[27]. They also bring many other advantages, such as low computational complexity and latency, since the observation of the physical-layer attributes usually does not require additional signalling or protocol changes [28]–[30]. However, the performance of physical-layer authentication schemes cannot be guaranteed due to the imperfect estimation and owing to the time-varying nature of the specific physical attributes used for authentication and noisy communication environment, such as the UAV network. The limited dynamic range of the specific attributes could also be insufficient to provide a guaranteed authentication result, when the number of devices to be authenticated increases [31]. Ultimately, observing and analyzing multiple attributes and devices at the same instance may create a bottleneck, which may further lead to single-point authentication failure and reduced application traffic [32].

To mitigate the probability of single-point failure and increase the overall reliability, decentralized authentication techniques have been developed, where a group of collaborative nodes are utilized for arriving at a final authentication decision. A popular decentralized authentication solution is the blockchain-based technique, where duplicated transactional databases are distributed over multiple nodes within a peer-to-peer network [33]–[35]. These nodes form a chain of interlocked blocks and each block contains the cryptographic hash of the previous block. Naturally, the longer the blockchain is, the safer the system becomes. Hence, it is extremely hard for the attacker to forge or delete the information, since the attacker has to overwrite or remove the history on all nodes before the next block record arrives [36]. However, due to the limited energy and storage space of UAVs, it is extremely

difficult to harness the blockchain-based method due to its excessive computational and communication overhead.

Ultimately, the physical-layer security (PLS) techniques have also been integrated into decentralized security schemes for mitigating the risks associated with digital security credentials and processes [37]–[39]. The estimations of physical attributes at distributed collaborative nodes are capable of mitigating the uncertainty caused by the time-varying environment and noisy estimations by a single observer. To optimize the performance of each collaborative node, a different number of PLS attributes can be selected at each node based on the specific hardware capability. The difficulty for the attacker to impersonate the legitimate devices is thus dramatically increased, since it is extremely hard to predict and impersonate different observed physical-layer attributes at different locations and at the same time. Moreover, the distributed authentication techniques do not require a static network topology for the authentication process. This can significantly improve both the reliability and the robustness of the authentication scheme, especially in a hostile environment, where the link between the UAVs and the CH is intermittent. However, upon involving more devices in the authentication process, the computational complexity and overall network latency will be dramatically increased which raises many challenges in resource-constraint devices.

In a nutshell, a major challenge of the conventional centralized digital-key based authentication or physical-layer authentication techniques is that they are best-effort based and usually require extra authentication resources for improving the overall performance [16]–[27]. On the other hand, some state-of-the-art decentralized authentication schemes are capable of improving the authentication security robustness even in the complex use cases with the aid of rich authentication resources. However, the excessive computational and communication overheads are unrealistic for resource-constrained networks [33]–[39]. More importantly, it is extremely challenging to utilize a static authentication scheme in hostile time-varying environments relying on complex distributed collaboration, especially in dynamically evolving UAV networks.

### *B. Contributions*

To overcome the above challenges, we propose the novel concept of Security-as-a-Service (SaS) for decentralized collaborative physical-layer authentication. To be more specific, in contrast to the best-effort based centralized authentication schemes, a guaranteed level of authentication performance is achieved by involving a minimal amount of 'just-sufficient' authentication resources, namely a limited number of collaborative nodes and their physical-layer attributes. The authentication performance is not maximized when using 'just-sufficient' collaborative nodes and physical-layer attributes, which is in contrast to the decentralized authentication schemes relying on all possible authentication resources. Instead, we aim to guarantee a specific target authentication requirement at a minimal computational cost. A fluid authentication topology can be customized for different time-varying environments so that the most reliable, robust and efficient model can be selected. Hence, an optimal equilibrium can be achieved between the authentication requirement, the required resources and the security performance. The novelty of this treatise is boldly and explicitly contrasted to the state-of-the-art in Table I. A list of key notations used in this manuscript are summarized in Table II. Moreover, a list of abbreviations and acronyms used throughout the paper is also given in Table III.

The contributions of this paper are summarized as follows:

- A novel concept of Guaranteed Authentication Level at Minimized Complexity (GALAMC) is proposed for decentralized collaborative physical-layer authentication. As a major benefit, the computational complexity of the distributed collaboration can be intelligently minimized by harnessing 'just-sufficient' collaborators and authentication attributes.
- To select the collaborative nodes and the corresponding authentication attributes, a Gini-impurity-based attribute evaluation algorithm is proposed for assessing the reliability of each time-varying physical-layer attribute at each collaborative node. A collaborative node evaluation algorithm is also developed for assessing the authentication benefits of each collaborative node based on their relative locations and their past authentication contributions.
- An intelligent authentication customization algorithm is proposed for integrating the above two factors and provide SaS. By activating this algorithm at each authentication instance, a customized authentication model will be generated for selecting the best combination of collaborative nodes. Reliable authentication decisions may then be generated at these selected collaborative nodes and fused into the final authentication decision.

This paper is organized as follows. Section II introduces our system model and problem formulation. Section III outlines the proposed Gini-impurity-based attributes evaluation algorithm, the collaborative node evaluation algorithm, intelligent twin-component authentication customization algorithm and our authentication decision fusion algorithm. The analysis and performance evaluation are presented in Section IV and V. Finally, the last section (Section VI) concludes this paper.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

To elaborate on the proposed authentication technique, the decentralized authentication process of a flying UAV network is considered. As shown in Fig. 2, a flying UAV swarm consists of $M$ legitimate UAVs including a cluster head (CH). The member UAVs are within the communication range of the CH. Hence, a star topology is formulated, where the CH and the member UAVs communicate directly without further routing via intermediate nodes. The relative velocity of the UAVs is in the typical range between 0 km/h and 60 km/h [40]. The line-of-sight (LoS) propagation and the channel fading conditions depend on the environment. For example, the urban area usually features non-line-of-sight (NLoS) channel condition with a Rayleigh fading distribution. Spoofing devices coexist in this network, which aim to actively impersonate legitimate UAVs. We assume that the spoofing UAVs are sufficiently far

TABLE I
OVERVIEW OF THE EXISTING STATE-OF-THE-ART TECHNIQUES

| **References ⇒ / keywords ⇓** | **Proposed Approach** | [37] | [15] | [19]–[21], [23]–[26] | [33], [35] | [38], [39] |
|---|---|---|---|---|---|---|
| **Physical-layer-based** | ✓ | ✓ | | ✓ | | ✓ |
| **Digital-based** | | | ✓ | | ✓ | |
| **Centralized Topology** | | | ✓ | ✓ | | |
| **Decentralized Topology** | ✓ | ✓ | | | ✓ | ✓ |
| **Best-effort-based** | | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Performance-guaranteed** | ✓ | | | | | |
| **Customized authentication model** | ✓ | | | | | |

TABLE II
LIST OF KEY NOTATIONS

| Notation | Explanation |
|---|---|
| $M$ | Number of UAVs (including CH) in the swarm |
| $\boldsymbol{H_m^I}$ | Physical-layer estimation observed by the UAV $m$ in phase $I$ |
| $N$ | Number of observed physical-layer attributes |
| $()^{\mathrm{T}}$ | Transposition of a vector |
| $J$ | Number of selected physical-layer attributes |
| $K$ | Number of selected collaborative UAV nodes |
| $\phi_k$ | Soft authentication decision generated by the selected UAV $k$ |
| $\phi_0$ | Authentication requester is legitimate |
| $\phi_1$ | Authentication requester is illegitimate |
| $\nu$ | Authentication decision threshold |
| $\mathcal{E}$ | Authentication error rate |
| $\mathcal{E}_D$ | Operator defined security requirement in terms of error rate |
| $\mathcal{E}_A$ | Actual authentication performance in terms of error rate |
| $G_n$ | Gini-impurity of the $n$-th attribute |
| $\mathcal{U}_{pq}$ | Usability index of a collaborative node |
| $P_k(\Phi_0)$ | Probability of legitimacy |
| $P_k(\Phi_1)$ | Probability of illegitimacy |
| $\mathcal{K}$ | K-factor |

TABLE III
LIST OF ACRONYMS AND ABBREVIATIONS

| Acronym | Explanation |
|---|---|
| **CFO** | Carrier Frequency Offset |
| **CH** | Cluster Head |
| **FA** | False Alarm |
| **FANET** | Flying Ad Hoc Network |
| **GALAMC** | Guaranteed Authentication Level at Minimized Complexity |
| **IoT** | Internet-of-Things |
| **I/Q** | In-phase/Quadrature |
| **LOF** | Local Outlier Factor |
| **LOS** | Line-of-Sight |
| **MD** | Missdetection |
| **NLOS** | Non-Line-of-Sight |
| **OF** | Objective Function |
| **PLS** | Physical-Layer Security |
| **RSSI** | Received Signal Strength Indicator |
| **SaS** | Security-as-a-Service |
| **UAV** | Unmanned Aerial Vehicles |

away from the legitimate UAVs, so that it is hard for the spoofing UAVs to predict the exact physical-layer attributes (e.g. channel conditions) of the legitimate UAVs. Due to the potential connectivity outages or long distances from the ground server, on-site authentication within the UAV swarm is always preferred for avoiding the related delay. Again, authentication coordinated by the CH of the UAV swarm is preferred due to the limited security related information and computational resources at every single UAV. In a nutshell, our main objective is to authenticate the devices at a guaranteed security by harnessing the minimum amount of authentication resources within the flying UAV swarm. The CH selects multiple collaborative nodes for generating edge authentication decisions based on the physical-layer estimations and fuses these decisions into a final authentication verdict. The full process of the decentralized authentication contains three phases:

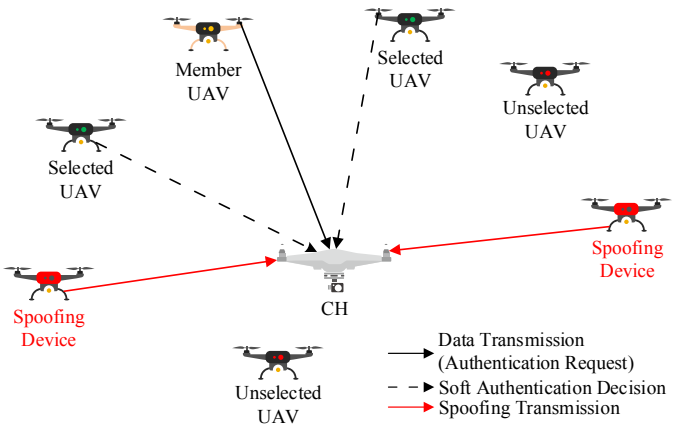


Fig. 2. System model of a flying UAV network. The legitimate communications in a UAV network suffer from attacks initiated by the spoofing devices. The CH coordinates the collaborative nodes and fuses authentication decisions to guarantee the SaS with minimum effort.

*Phase I*: At the time $t_1$, one or more messages have been transmitted to the CH and the CH aims to appropriately select the collaborative nodes available for authentication based on the observations of the message. Due to the interference or noise imposed by the environment, some collaborative nodes including the CH may observe a noisy physical-layer estimate $\boldsymbol{H^I}$. The estimate at UAV $m$ is given by:

$$\boldsymbol{H_m{}^I} = (H_{m1}^I, H_{m2}^I, , H_{mN}^I)^{\mathrm{T}}, \tag{1}$$

where $N$ is the number of physical-layer attributes observed and $()^{\mathrm{T}}$ represents the transposition of a vector. The physical-layer attributes may include the carrier frequency offset (CFO), in-phase/quadrature (I/Q) imbalance, received signal strength indicator (RSSI) and so on. These physical-layer attributes are then concatenated into an array with time stamps and stored at each node for future analysis. Since some authentication attributes may be less accurate, an attribute's reliability evaluation may be implemented at each available collaborative node as a result. The number of reliable attributes being selected for improving the performance at minimal computational complexity at each node is denoted as $J$. By evaluating the reliability of the attributes at each node, the unavailable node(s) having no reliable estimates will be temporarily removed from the authentication process. Then, based on their relative location and past contributions, the CH selects the least number of collaborative nodes from the set of available and reliable collaborators, which is denoted by $K$.

*Phase II*: At the time $t_2$, each selected collaborative node generates an edge authentication decision and reports back to the CH. For example, at the selected UAV $k$, a soft edge authentication $\phi_k$ is generated, where we have $\phi_k = [0, 1]$. The collaborative node evaluates how likely it is that the authentication requester is legitimate based on the physical-layer observations. For example, $\phi_k = 0.5$ represents that the UAV $m$ estimates a 50% probability for the authentication requester to be legitimate.

*Phase III*: At the time $t_3$, the CH generates the final authentication decision based on the $K$ received edge authentication decisions as follows:

$$\begin{cases} \Phi_0, & \frac{1}{K}\sum_{k=1}^{K}\phi_k > \nu; \\ \Phi_1, & \frac{1}{K}\sum_{k=1}^{K}\phi_k \leq \nu, \end{cases} \tag{2}$$

where $K$ is between 1 and $M-1$. Furthermore, $\Phi_0$ represents that the transmitter is legitimate, while $\Phi_1$ indicates that the signal is transmitted from a spoofing device; $\phi_k$ represents the soft authentication decision gleaned from the $k$-th receiver node. Moreover, $\nu$ is the authentication decision threshold at the CH in the range of $[0, 1]$, which can be dynamically configured by the operators for different scenarios.

To evaluate the performance of a collaborative decentralized authentication, the False Alarm rate and the Missdetection rate are considered, which can be formulated as:

1) False Alarm (FA) rate: The probability that a legitimate UAV is rejected as a suspected spoofing device, which is formulated as:

$$P_{\mathrm{FA}} = \Pr(\frac{1}{K}\sum_{k=1}^{K}\phi_k \leq \nu|\Phi_0). \tag{3}$$

where $\Pr()$ represents the probability of an event.

2) Missdetection (MD) rate: The probability that a spoofing device is approved as a legitimate UAV. It can be defined as:

$$P_{\mathrm{MD}} = \Pr(\frac{1}{K}\sum_{k=1}^{K}\phi_k > \nu|\Phi_1). \tag{4}$$

To define the security requirement and evaluate the performance of the actual authentication, the false alarm rate and the missdetection rate can be combined into the authentication error rate ($\mathcal{E}$) as:

$$\mathcal{E} = w_1 P_{\mathrm{FA}} + w_2 P_{\mathrm{MD}}, \tag{5}$$

where $w_1$ and $w_2$ are the weights of the false alarm and the missdetection, since each may have a different impact on the system in different scenarios. The weights directly represent the security requirements and should be defined by the operator based on the specific application. To formulate the security requirement ($\mathcal{E}_D$) in terms of authentication's error rate, the operator has to also define the target false alarm rate and missdetection rate. Similarly, the actual authentication performance can also be formulated by measuring the actual false alarm rate and missdetection rate as $\mathcal{E}_A$. To guarantee the target SaS, the actual security performance should always be 'just-above' the security requirement ($\mathcal{E}_D > \mathcal{E}_A$), so that the security guarantee is met without utilizing an excessive amount of authentication resources, such as the number of selected collaborative nodes ($K$) and the number of attributes selected at each node ($J$). Hence, the SaS is defined as:

$$\min_{J,K,\nu} \mathcal{E}_D - \mathcal{E}_A, \tag{6}$$

which is also the objective function (OF) of our problem formulation. Therefore, it is critical to select the most reliable collaborative nodes for computing the soft edge authentication decisions so that the SaS can be guaranteed with minimum effort.

## III. GUARANTEED SAS AT MINIMIZED COMPLEXITY BY INTELLIGENT COLLABORATIVE AUTHENTICATION

To solve the problem of (6) by forming a customized authentication model, it is critical to utilize the necessary and minimal authentication resource to achieve the 'just-sufficient' authentication performance. Therefore, quantifying the reliability and selecting the most appropriate physical-layer attributes at each collaborative node becomes the first challenging design dilemma. This can help us to remove the unnecessary attributes, while eliminating the nodes that failed to collect any reliable physical-layer estimates for various reasons. Then, the CH aims for selecting the most reliable combination of collaborative nodes to further reduce the amount of authentication resources being utilized. To be more specific, a collaborative node evaluation algorithm is developed for characterizing and for ranking the available nodes by considering both their relative locations and past authentication contributions. An intelligent twin-component authentication customization algorithm is also proposed for selecting the collaborative nodes based on the previous algorithms. Once the reliable nodes and the reliable attributes at each node are selected, a soft authentication decision is generated at each selected collaborative node and transmitted to the CH. Ultimately, the CH intelligently fuses the soft authentication decisions into the final authentication decision and updates the authentication record for future use. The flow of the above procedures is shown in Fig. 3.

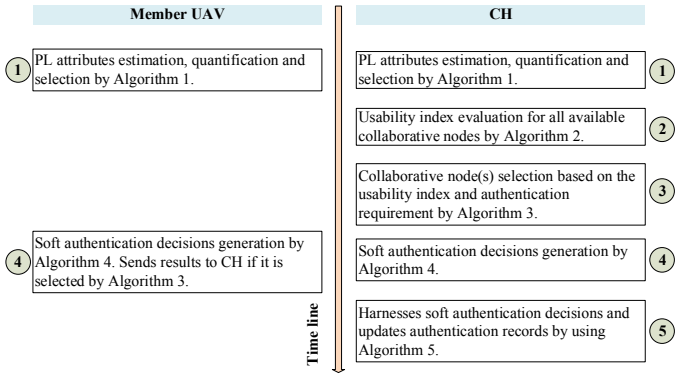


Fig. 3. Authentication processing flow for member UAVs and CH of the proposed algorithms

### A. Gini-impurity-based Attributes Evaluation Algorithm

To guarantee the target SaS at minimum effort, we first have to verify, whether the physical-layer attributes are reliable at each available collaborative node so that we can eliminate any excessive authentication resources. The conception of a reliable attribute evaluation algorithm that can competently quantify the contribution of the specific attributes to the SaS performance becomes the next challenge. To achieve this goal, the past observations have to be stored and exploited for continuously monitoring the behaviour of each attribute at the collaborative nodes. If no estimate is observed or none of the observed attributes are reliable, the associated collaborative node should be removed from the authentication process. This eliminates the ambiguity caused by unreliable decisions and also reduces the overall computational cost.

To achieve an attribute reliability evaluation criterion that can be generalized, we compute the Gini-impurity for all available physical-layer attributes. Explicitly, the Gini-impurity represents the probability of an attribute misleading the authentication decision [41]–[45]. To be more specific, a high Gini-impurity means that the attribute has a high probability of misleading the authentication decision; hence, the attribute is unreliable. The Gini-impurity of the $n$-th attribute can be formulated as:

$$G_n = \sum_{c=1}^{C} f_{nc}(1 - f_{nc}) = \sum_{c=1}^{C} f_{nc} - \sum_{c=1}^{C} f_{nc}^2 = 1 - \sum_{c=1}^{C} f_{nc}^2, \tag{7}$$

where we have $C = 2$ since the authentication decision can only be either correct or false. Furthermore, $f_{n1}$ and $f_{n2}$ are the frequency of the decision being correct by using only the $n$-th attribute, and that of being false, respectively. To get the value of $f_{nc}$, the previous authentication decisions using only the $n$-th attribute are computed and compared to the final authentication decisions fedback from the CH. For example, if the authentication decision using the attribute is different from the final authentication decision, it would be deemed to be a false decision. Then, the total number of correct decisions and false decisions are divided by the total number of available decisions as $f_{nc}$. It should be noted that when the total number of decisions is too high, the accuracy of $f_{nc}$ may become less sensitive to real-time environmental changes. The collaborative node should consider removing the early observations in order to remain sensitive and to reduce the data storage required. Ultimately, to evaluate whether an attribute is reliable and hence it is worth selecting, a Gini-impurity threshold ($\tau$) can be introduced and selected by the operator based on the specific scenario. Explicitly, the attributes are deemed reliable if the Gini-impurity is lower than a specific operator-defined threshold, i.e., $G_n < \tau$. If no observation is made or none of the attributes is reliable at a collaborative node, the node will not be considered by the CH at this instant and hence will be assigned a usability index of 0. Meanwhile, this excluded node will keep collecting the physical-layer estimates and rejoin the authentication process, when it meets the minimum performance requirement. Hence, to find the optimum number of selected attributes ($J$) in (6), an optimized $\tau$ value has to be found based on the specific security requirement by the operator. The relationship between the $\tau$ selection has been analyzed in Section IV. The proposed Gini-impurity-based attribute evaluation procedure is shown in Algorithm 1.

**Algorithm 1** Gini-impurity-based attributes evaluation algorithm

Given the previous observations of the other devices from each collaborative node and the authentication decision feedback of those devices from the CH. All the physical-layer attributes are considered in the UAV network.

1: Gini impurity ($G_n$) value of each physical-layer attribute is obtained via (7);
2: **if** $G_n < \tau, n = 1, 2, ..., N$ **then**
3: $n$-th attribute will be deemed as non-informative and dropped;
4: **else**
5: $n$-th attribute will be selected;
6: **end if**
7: each collaborative node calculates the number of attributes being selected ($J$);
8: **if** $J \neq 0$ **then**
9: self-reports to the CH as an available and reliable collaborative node;
10: updates the local authentication scheme to utilize $J$ selected attributes for the next authentication instance;
11: **else**
12: temporarily eliminates from the authentication process.
13: **end if**

### B. Collaborative Node Evaluation Algorithm

After selecting the attributes at each collaborative node, the next step is to optimize the number of collaborative nodes ($K$) for making the final authentication decision. Since the objective of (6) is to guarantee the SaS rather than maximize the security performance, some of the collaborative nodes can be eliminated for reducing the computational cost. Hence, it is critical to quantify a figure of merit for the remaining collaborative nodes termed as the usability index at the authentication

instance. The less important collaborative nodes can then be eliminated for reducing the computational cost.

To evaluate the usability of the remaining collaborators, the factors that cause the accuracy fluctuation of the physical-layer observations have to be studied. The analog physical-layer attributes estimates tend to be more environment-dependent than the upper layer attributes. For example, a longer distance between the authentication requester and the collaborative node results in lower received signal strength, while the noise level is almost constant. Therefore, the noise may significantly increase the measurement deviation and increase the uncertainty for distant requesters. On the other hand, different locations also result in different channel fading statistics. The measured physical-layer estimates may become significantly different from the previous estimates and hence may result in a flawed decision. Hence, from the challenges listed above, the relative location between the collaborative node ($p$) and the authentication requester ($q$) is chosen as one of the evaluation attributes for quantifying the usability of the collaborative node. The longitude ($X$), latitude ($Y$) and altitude ($Z$) are utilized to define the location of each node. By analyzing the relative location as an attribute for quantifying the usability index, collaborator nodes that failed to observe reliable physical-layer estimates will be temporarily eliminated from the authentication as an outlier.

To detect an outlier that leads to an unreliable physical-layer estimate based on the relative location, the Local Outlier Factor (LOF) can be utilized for defining the local neighborhood of the data point [46]. It can reveal how isolated a data point is with respect to its neighborhood based on a single parameter $\mathcal{N}$, which is the number of nearest neighbors used in defining the local neighborhood. The distance between the data point ($\alpha$) and the $\mathcal{N}$-th neighbor is defined as $kdist(\alpha)$. The judgment of the outlier is based on the density between each data point and its neighbor points [47]. If the density of reliable estimate is lower than normal, it is more likely to be identified as an outlier, since it has a lower probability to make a reliable edge authentication decision [48]. Then, the reachability distance, which is an intermediate parameter, can be expressed as:

$$rdist(\alpha,\beta) = \max\{dist(\alpha,\beta), kdist(\beta)\}, \tag{8}$$

where $\alpha$ is the current data point and $\beta$ is the target point. Since there are 3 attributes, namely longitude, latitude and altitude, in each data point, the $dist(\alpha,\beta)$ can be expressed as:

$$dist(\alpha,\beta) = \sqrt{(X_\alpha - X_\beta)^2 + (Y_\alpha - Y_\beta)^2 + (Z_\alpha - Z_\beta)^2}. \tag{9}$$

Then, the local reachability density $lrd(\alpha)$ of the data point $\alpha$, which quantifies the average reachability distance of $\mathcal{N}$ neighbors can be expressed as:

$$lrd(\alpha) = \frac{|R(\alpha)|}{\sum_{\beta\in R(\alpha)} rdist(\alpha,\beta)}, \tag{10}$$

where $|R(\alpha)|$ denotes the size of $R(\alpha)$, which can be written as:

$$R(\alpha) = \{\beta | dist(\alpha,\beta) < kdist(\alpha)\}. \tag{11}$$

Finally, the LOF can be calculated as:

$$lof(\alpha) = \frac{\sum_{\beta\in R(\alpha)} \frac{lrd(\beta)}{lrd(\alpha)}}{|R(\alpha)|}. \tag{12}$$

If the LOF is near or smaller than 1, it is more likely to be a normal data point. By contrast, if the LOF is higher than 1, it is more likely to be an outlier. To be more specific, the relative location will be converted to a binary flag of either 1 or 0, where 1 indicates that the collaborative node is capable of generating a reliable edge authentication decision at this location as a normal data point and 0 means that the collaborative node fails to generate a reliable edge authentication, as it is an outlier. This binary decision can be formulated as:

$$D_{pq} = \begin{cases} 0, & lof(\alpha_p) > \mathcal{L}; \\ 1, & lof(\alpha_p) \leq \mathcal{L}, \end{cases} \tag{13}$$

where $D_{pq}$ is the binary index that judges whether the collaborative node ($p$) can make a reliable physical-layer estimate at its relative location with respect to the authentication requester ($q$). Furthermore, $\mathcal{L}$ is the LOF threshold selected by the operator and the data points used to calculate the LOF are previous authentication contributions collected at the CH.

On the other hand, there exists a scenario in which some of the soft authentication decisions are received from attackers. It is critical to monitor the behaviour of each collaborative node and eliminate any suspicious collaborative nodes. To achieve this goal, the authentication contribution of each collaborative node has to be considered. If a collaborative node has a high probability of giving a flawed authentication decision, the usability index should be adjusted to reflect the node's unreliable behaviour. Therefore, the authentication reliability rate ($R_{pq}$) of a collaborative node ($p$) with respect to the authentication requester ($q$) can be calculated by using the $U$ most recent authentication decisions as:

$$R_{pq} = \frac{\sum_{u=1}^{U}(1 - w_1 \text{Pr}(\phi_{pq} \leq \nu | \Phi_0) - w_2 \text{Pr}(\phi_{pq} > \nu | \Phi_1))}{U}, \tag{14}$$

where $\nu$ is the authentication threshold at the CH used to evaluate whether the contribution of the collaborative node is positive or negative. Furthermore, $w_1$ and $w_2$ are the weights used in (5) to reflect the different importance levels of the misdetection and false alarm scenarios. Then, to formulate the usability index of the collaborative node ($p$) with respect to the authentication requester ($q$), the distance estimate and the reliability rate can be combined as:

$$\mathcal{U}_{pq} = R_{pq} D_{pq}, \tag{15}$$

where we have $\mathcal{U}_{pq} = [0, 1]$. Therefore, if the collaborative node is deemed unreliable due to its relative location, the usability index will be set to 0, since $D_{pq} = 0$. Collaborative nodes that are eliminated in Algorithm 1 will automatically have a usability index of 0. Then, the calculated usability index will be passed to the next step for ultimately selecting the collaborative node and the combination of the attributes. The proposed procedure is shown at a glance in Algorithm 2.

**Algorithm 2** Collaborative Node Evaluation Algorithm

Given the collaborative node $p$, authentication requester $q$, and $U$ previous contributions,

1: acquire the relative location of the collaborative node ($p$) with respect to the authentication requester ($q$).
2: compute the binary distance index ($D_{pq}$) via (13);
3: calculate the reliability index ($R_{pq}$) via (14);
4: fuse the calculated $D_{pq}$ and $R_{pq}$ as the usability index ($\mathcal{U}_{pq}$) via (15);

### C. Intelligent Twin-component Authentication Customization Algorithm

After calculating the usability index of each collaborative node by using Algorithm 2, the set of usability indices can be formulated as $\boldsymbol{\mathcal{U}} = (\mathcal{U}_{1q}, \mathcal{U}_{2q}, ..., \mathcal{U}_{Mq})^{\mathrm{T}}$ in descending order where $\mathcal{U}_{1q}$ has the highest usability index and $\mathcal{U}_{Mq}$ has the lowest usability index. If multiple collaborative nodes have the same usability index, the node that has a higher proportion of correct authentication decisions beyond the most recent $U$ authentication decisions will have a higher rank. For example, if a pair of collaborative nodes has a usability index of 1, the node with more correct authentication decisions will be ranked as $\mathcal{U}_{1q}$ and the other one will be ranked as $\mathcal{U}_{2q}$. The usability index of all $M$ devices is included in this set, where the unreliable collaborative nodes flagged in Algorithm 1 and Algorithm 2 have a usability index of 0. To select the collaborative nodes based on situational-awareness, it is critical to understand the authentication performance requirement to be met by the application. To be more specific, a military application usually has a lower tolerance for wrong authentication decisions than a civilian application due to the more severe outcomes caused by the fault. Therefore, the optimization goal of (6) can be met by relying on the minimum number $K$ of collaborative nodes.

To guarantee the SaS as given in the problem formulation (6), the usability index obtained from Algorithm 2 can be utilized, since it can be converted to the missdetection rate and the false alarm rate, which are the pair of attributes that characterize the error rate. Therefore, the goal of the collaborative node selection may be reformulated as:

$$\min_{K}[\mathcal{E}_D - (1-\mathcal{U}_{1q})(1-\mathcal{U}_{2q})...(1-\mathcal{U}_{Kq})], \tag{16}$$

where $K = 1, 2, ..., M$ and $[\mathcal{E}_D - (1-\mathcal{U}_{1q})(1-\mathcal{U}_{2q})...(1-\mathcal{U}_{Kq})] \geq 0$. A stricter security requirement generally indicates that more collaborative nodes have to be utilized to fuse the final authentication decision. To elaborate on the authentication models, if multiple UAVs have a usability index of 1, including the CH, only the CH will be harnessed for performing the authentication to minimize the computational overhead. On the other hand, if multiple UAVs have the same usability index and the algorithm decides that it does not need all of them, the collaborative node having a higher rank will be selected. The proposed intelligent twin-component authentication customization procedure is given in Algorithm 3.

**Algorithm 3** Intelligent Authentication Customization Algorithm

Given the usability index of each device from Algorithm 2 and the operator-defined security requirement ($\mathcal{E}_D$).

1: rank the usability index into a set of usability in descending order as $\boldsymbol{\mathcal{U}} = (\mathcal{U}_{1q}, \mathcal{U}_{2q}, ..., \mathcal{U}_{Mq})^{\mathrm{T}}$;
2: **if** there exist multiple collaborative nodes that have the same usability index **then**
3: a collaborative node with more correct authentication decisions in the past will have a higher rank;
4: **end if**
5: select the top $K$ collaborative nodes from the set of usability index that meet the requirement of $\min_K(\mathcal{E}_D - (1-\mathcal{U}_{1q})(1-\mathcal{U}_{2q})...(1-\mathcal{U}_{Kq}))$;

### D. Authentication Decision Fusion Algorithm

After customizing the authentication model, all $K$ selected collaborative nodes have to submit their soft authentication decisions to the CH, where the final authentication decision is fused. To generate the soft authentication decision ($\phi_k$) at node $k$, it is critical to evaluate how likely the authentication requester is legitimate based on the $J$ selected physical-layer attributes. Hence, the probability of legitimacy generated by node $k$ is used as the soft authentication decision and is expressed as $P_k(\Phi_0)$ and the binary hypothesis test of (2) can be reformulated as:

$$\begin{cases} \Phi_0, & \frac{1}{K}\sum_{k=1}^{K} P_k(\Phi_0) > \nu; \\ \Phi_1, & \frac{1}{K}\sum_{k=1}^{K} P_k(\Phi_0) \leq \nu. \end{cases} \tag{17}$$

Ideally, an optimized regression model is used to map the physical-layer attributes to the probability of legitimacy. However, it is a challenge to fit such a model to the probability within [0,1], since the boundary of the regression model is usually $(-\infty, \infty)$ [49]. Hence, to simplify the regression model, we utilize the natural logarithm of the odd, also known as the logit, so that the domain of [0,1] is relaxed to $(-\infty, \infty)$ [50]. The logit ($L$) is the natural log of the ratio between the probability of being legitimate and the probability of being illegitimate. Since the identity of the authentication requester UAV is also binary, the logit can be formulated as:

$$L = ln(\frac{P_k(\Phi_0)}{P_k(\Phi_1)}) = ln(\frac{P_k(\Phi_0)}{1-P_k(\Phi_0)}), \tag{18}$$

where $\Phi_0$ represents that the device is legitimate and $\Phi_1$ means that the device is illegitimate. $P_k(\Phi_0)$ indicates the probability for the authentication requester to be legitimate and $P_k(\Phi_1)$ represents the probability to be illegitimate and $P_k(\Phi_0) = 1 - P_k(\Phi_1)$. To find $P_k(\Phi_0)$, different linear regression models can be utilized to map the $J$ selected physical-layer attributes to the logit. However, given the resource-constrained nature of the UAVs, complex regression model such as a neural network may cause excessive long latency. Hence, the linear regression model is implemented in the UAV swarm, and then (18) can be rewritten as:

$$ln(\frac{P_k(\Phi_0)}{1-P_k(\Phi_0)}) = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}, \tag{19}$$

where $\boldsymbol{B}$ is the vector of the regression coefficient computed by the linear regression model and $\boldsymbol{X}$ is the vector of the physical-layer attributes. Therefore, by combining (18) and (19), $P_k(\Phi_0)$ can be expressed as:

$$\begin{aligned} P_k(\Phi_0) &= e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}[1 - P_k(\Phi_0)] \\ &= e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})} - e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}P_k(\Phi_0) \\ &= \frac{e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}}{1 + e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}}. \end{aligned} \tag{20}$$

Hence, the authentication decision fusion algorithm can be broken down into two algorithms, namely the soft authentication decision algorithm implemented at each collaborative node as Algorithm 4 and the final decision fusion algorithm harnessed by the CH as Algorithm 5. Both algorithms are given in the detailed flow chart of the proposed intelligent collaborative authentication scheme in Fig. 4. Fig. 4 also outlined the flow of our proposed scheme in detail as a complement to Fig. 3. The optimization task of this algorithm is to find the best $\nu$ value that fulfills (6). After the final authentication decision is made, the CH will transmit the judgment back to each collaborative node for future analysis and then updates the authentication record, which constitutes as the *contribution* used in Algorithm 2.

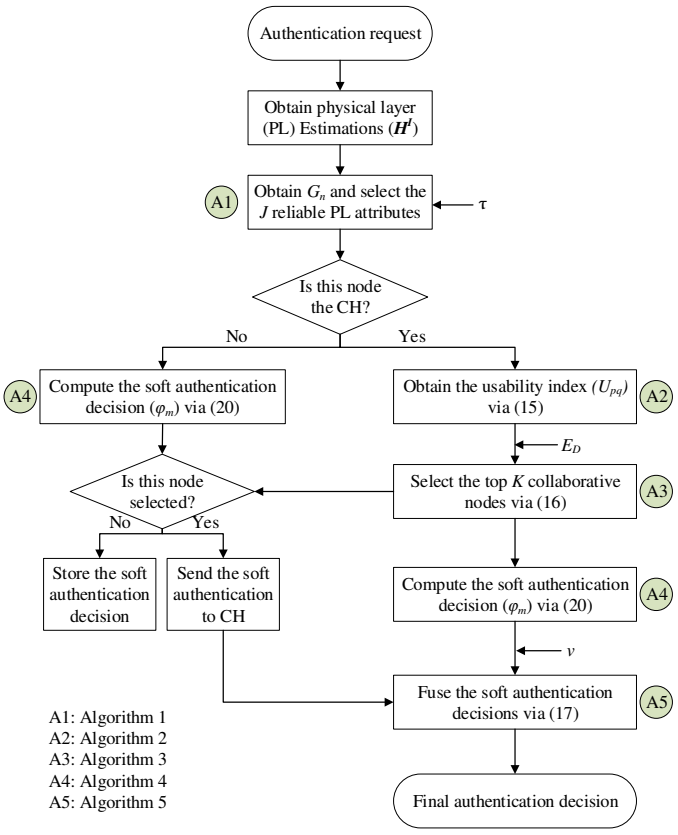


Fig. 4. Flow chart of the proposed intelligent collaborative authentication scheme

## IV. Analysis

In this section, we first analyze the effects of Gini-impurity threshold selection on the Gini-impurity-based attribute evaluation algorithm's performance. Then, the impact of the authentication threshold $\nu$ will also be discussed.

### A. Analysis of the Gini-impurity Threshold ($\tau$) Selection

In order to minimize the authentication resource utilization at each node, the unreliable physical-layer attribute have to be removed by using the Gini-impurity-based attributes evaluation algorithm. Since the Gini-impurity threshold is defined by the operator, it is critical to analyze its impact on the authentication performance. To achieve this goal, (20) which represents the soft authentication is firstly combined with the FA rate of (3) and the MD rate of (4). Then, the error rate of (5) can be reformulated as:

$$\begin{aligned} \mathcal{E} &= w_1 P_{\mathrm{FA}} + w_2 P_{\mathrm{MD}} \\ &= w_1 \Pr\Big(\frac{1}{K}\sum_{k=1}^{K}\phi_k \le \nu|\Phi_0\Big) + w_2 \Pr\Big(\frac{1}{K}\sum_{k=1}^{K}\phi_k > \nu|\Phi_1\Big) \\ &= w_1 \Pr\Big(\frac{1}{K}\sum_{k=1}^{K}\frac{e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}}{1+e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}} \le \nu|\Phi_0\Big) \\ &\quad + w_2 \Pr\Big(\frac{1}{K}\sum_{k=1}^{K}\frac{e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}}{1+e^{(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X})}} > \nu|\Phi_1\Big), \end{aligned} \tag{21}$$

where $\boldsymbol{X}$ is the vector of the selected physical-layer attributes. Only the physical-layer attribute having a Gini-impurity ($G_n$) smaller than the Gini-impurity threshold $\tau$ can be included into this vector. We assume that there exist two Gini-impurity thresholds where $0.5 \ge \tau_1 > \tau_2 > 0$. Since $\tau_1 > \tau_2$, the Gini-impurity of the physical layer attribute vector $\boldsymbol{X}_{\tau 1}$ is higher than or equal to the Gini-impurity of $\boldsymbol{X}_{\tau 2}$. To be more specific, $\boldsymbol{B}_{\tau 2}$ will fit the regression model better compared to $\boldsymbol{X}_{\tau 1}$ due to the associated information gain, which is the difference between the Gini-impurity of $\boldsymbol{X}_{\tau 1}$ and $\boldsymbol{X}_{\tau 2}$. Hence, we can get $\Pr(\frac{1}{K}\sum_{k=1}^{K}\frac{e^{(\boldsymbol{B}_{\tau 1}^{\mathrm{T}}\boldsymbol{X}_{\tau 1})}}{1+e^{(\boldsymbol{B}_{\tau 1}^{\mathrm{T}}\boldsymbol{X}_{\tau 1})}} \le \nu|\Phi_0) \ge \Pr(\frac{1}{K}\sum_{k=1}^{K}\frac{e^{(\boldsymbol{B}_{\tau 2}^{\mathrm{T}}\boldsymbol{X}_{\tau 2})}}{1+e^{(\boldsymbol{B}_{\tau 2}^{\mathrm{T}}\boldsymbol{X}_{\tau 2})}} \le \nu|\Phi_0)$ and $\Pr(\frac{1}{K}\sum_{k=1}^{K}\frac{e^{(\boldsymbol{B}_{\tau 1}^{\mathrm{T}}\boldsymbol{X}_{\tau 1})}}{1+e^{(\boldsymbol{B}_{\tau 1}^{\mathrm{T}}\boldsymbol{X}_{\tau 1})}} > \nu|\Phi_1) \ge \Pr(\frac{1}{K}\sum_{k=1}^{K}\frac{e^{(\boldsymbol{B}_{\tau 1}^{\mathrm{T}}\boldsymbol{X}_{\tau 1})}}{1+e^{(\boldsymbol{B}_{\tau 1}^{\mathrm{T}}\boldsymbol{X}_{\tau 1})}} > \nu|\Phi_1)$ which are equivalent to $\mathcal{E}_{\tau 1} \ge \mathcal{E}_{\tau 2}$. In an extreme case, where $\tau_1 = 0.5$, all physical layer attributes are chosen and Algorithm 1 can be deemed as excluded. It can be safely concluded that the security performance of using Algorithm 1 ($0.5 > \tau_2 > 0$) will be better than excluding Algorithm 1 ($\tau_1 = 0.5$). However, it should be noted that if $\tau$ is too small, no attribute will be selected to construct the regression model. Hence, the $\tau$ value selection should be carefully adjusted based on the specific application use case. The simulation result of the analysis is given in Section V, Fig. 6 and 7.

### B. Analysis of the Authentication Threshold ($\nu$) Selection

As the final authentication decision judgment, it is critical to select an optimal authentication threshold to minimize the error rate. To analyze the impact of the threshold selection, we assume that the operator defines an authentication threshold $\nu$,

where $1 \geq \nu \geq 0$. When $\nu = 1$, all devices will be deemed to be spoofing devices. In the extreme case, the false alarm rate and the missdetection rate will become $\Pr(\frac{1}{K}\sum_{k=1}^{K}\phi_k \leq \nu|\Phi_0) = 1$ and $\Pr(\frac{1}{K}\sum_{k=1}^{K}\phi_k > \nu|\Phi_1) = 0$ so that (21) can then be simplified as $\mathcal{E} = w_1$. Similarly, when $\nu = 0$, all devices will be authenticated as legitimate and the error rate can be simplified as $\mathcal{E} = w_2$, respectively. Hence, in extreme cases, the error rate can be reformulated as:

$$\mathcal{E} = \begin{cases} w_1, & \nu = 1; \\ w_2, & \nu = 0. \end{cases} \tag{22}$$

In order to place $\nu$ in the range that is biased neither to the missdetection nor the false alarm, it is important to select the value of $\nu$ within the optimal range. However, it is extremely hard for the human operator to find the range before the application starts due to the nature of each unique environment. Therefore, it is critical to improve the optimal range of $\nu$ at each node to compensate for the potential error. As discussed previously in Section IV-A, the regression model will fit better due to its information gain by applying the Gini-impurity-based attribute evaluation algorithm. Given more accurate soft authentication decisions at each collaborative node, the optimal range of $\nu$ can be relaxed. The simulation results gleaned from this analysis can be found in Section V, Fig. 6.

## V. PERFORMANCE EVALUATION

In this section, the performance of the proposed intelligent collaborative authentication scheme providing a guaranteed SaS is studied using MATLAB based simulation. To represent different UAV swarm sizes, the simulations consist of three UAV swarms that contain 2, 4 and 6 member UAVs as well as a CH. 30 spoofing devices coexist in the UAV swarm which aim for actively impersonating each legitimate member UAV at each time instance for malicious purposes. Each UAV has a random motion path and a random starting position that is less than 30 meters from each other. The relative travelling speed between two UAVs is in the typical range of 0 km/h and 60 km/h [40]. The analysis represents the last of the 5 simulations that have been initialized differently and all simulations have similar results. A dynamic environment having 600 observations is constructed in which both urban and rural areas are considered along with a transitional period. The physical-layer attributes simulated include the RSSI, CFO and IQI.

A 3D motion trajectory is considered for the UAVs. An altitude-dependent Rician model is considered in the line-of-sight (LOS) conditions in the rural area, since the flight altitude varies from 150 to 300m [51]–[53]. The probability density function of the Rician distribution can be expressed as:

$$P_{\text{Rician}}(s) = \frac{s}{\sigma^2}\exp\left(\frac{-s^2 - A^2}{2\sigma^2}\right) I_0\left(\frac{As}{\sigma^2}\right), s \geq 0, \tag{23}$$

where $s$ is the amplitude of the received signal, $A$ is the peak amplitude of the LOS component, $I_0(\cdot)$ is the modified Bessel function of the first kind with order zero, $\sigma$ is the root-mean-square of the received signal and the K-factor is defined as $\mathcal{K} = \frac{A^2}{2\sigma^2}$. Then, in the urban area, where buildings may exist, a non-line-of-sight (NLOS) condition with a flight altitude between 15 and 30 m [54] is considered. In the NLOS condition, since there may be no dominant path, the Rician fading reduces to a Rayleigh fading which can be formulated as:

$$P_{\text{Rayleigh}}(s) = \frac{s}{\sigma^2}\exp\left(\frac{-s^2}{2\sigma^2}\right). \tag{24}$$

To model the LoS and NLoS condition in MATLAB, we use the embedded wlanTGaxChannel which conveys the signal through the 802.11ax channel. The MATLAB predefined Model-A with $\mathcal{K} = 0$ is used to simulate the Rayleigh fading with 1 propagation path under the LoS condition. By contrast, the MATLAB predefined model-F associated with $\mathcal{K} = 6$ is selected to simulate the Rician fading having 6 propagation paths to represent the NLOS condition in a complex environment.

To simulate a traveling UAV swarm, we construct a scenario, where the UAV swarm travels from a complex terrain (i.e., urban area) to a simple terrain (i.e., suburban area). In the complex terrain, the transmissions within the UAV swarm are fully in NLoS conditions. Then, when the UAV swarm is close to the simple terrain, a transitional period emerges, where some of the transmissions take place in the LoS condition, while the majority of the transmissions are still in NLoS condition. To simulate this period, a random token was used at each collaborative UAV to toggle between model-A and model-F with a 25% probability being LOS condition at each UAV. Hence, each collaborative UAV may opt for a different authentication decision due to the different observation conditions. Lastly, a sudden environmental change takes place to represent the UAV swarm arriving at the simple terrain. In this stage, all transmissions among the UAVs suddenly move into a full LoS condition to test the robustness of our proposed scheme. To evaluate the error rate of (5), we set $w_1 = w_2 = 0.5$.

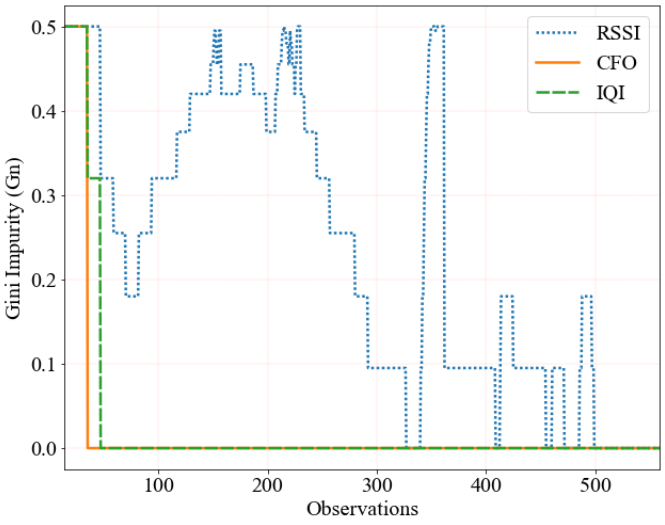


Fig. 5. Gini-impurity measurements of the physical-layer attributes across different environments at a collaborative node

To select a collaborative node associated with reliable attributes to guarantee the SaS, the Gini-impurity-based attribute evaluation algorithm is proposed first to evaluate the reliability of each physical-layer attribute at each collaborative node. To examine whether the Gini-impurity measurement can reflect the different characteristics of each attribute, we selected a random member UAV within the network and plotted the relationship between the Gini-impurity and the time-varying environment as shown, in Fig. 5. It can be observed that the Gini impurity of each physical-layer attribute fluctuates with respect to the environmental change. In this example, the Gini-impurity of the RSSI fluctuates dramatically due to the high mobility of the UAVs since RSSI is distance sensitive. Therefore, it can be concluded that each attribute has a different level of reliability and robustness in a time-varying environment.

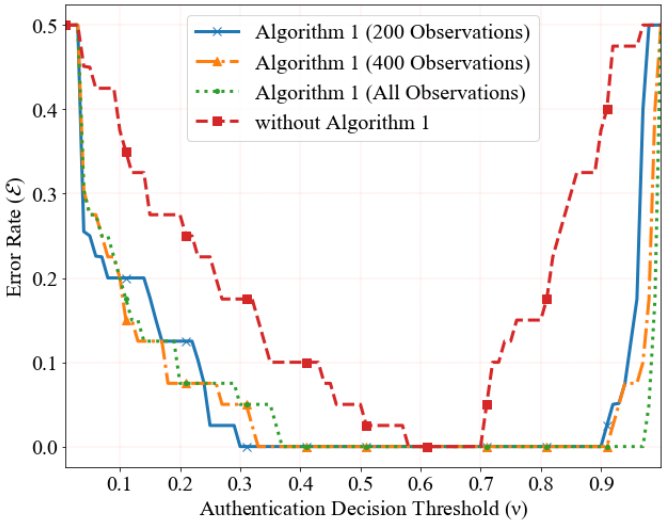


Fig. 6. Error rate comparison results with Algorithm 1 at different parameters and without Algorithm 1 with all attributes.

Furthermore, to examine whether it is necessary to eliminate the less reliable attributes by utilizing the Gini-impurity-based attribute evaluation algorithm, we selected a random member UAV within the network again. As discussed in the analysis of Section IV, we set $\tau = 0.5$ to utilize all physical-layer attributes for error rate comparison purposes. The error rate of (5) is computed with respect to the number of observations, as shown in Fig. 6. This demonstrates that our proposed algorithm has a better error rate performance than utilizing all physical-layer attributes across different $\nu$. It can also be concluded that our proposed scheme significantly increases the optimal authentication decision threshold interval, as discussed in Section IV-B, which makes it easier to select the authentication threshold ($\nu$). On the other hand, we have also tuned Algorithm 1 by utilizing a different number of observations to obtain $f_{nc}$. To be more specific, we have included 3 scenarios where the latest 200 observations, the latest 400 observations and all observations are used. It can be observed that the security performance between the three cases is very similar, because it is predominantly determined by the number of the observations which is 600 in total.

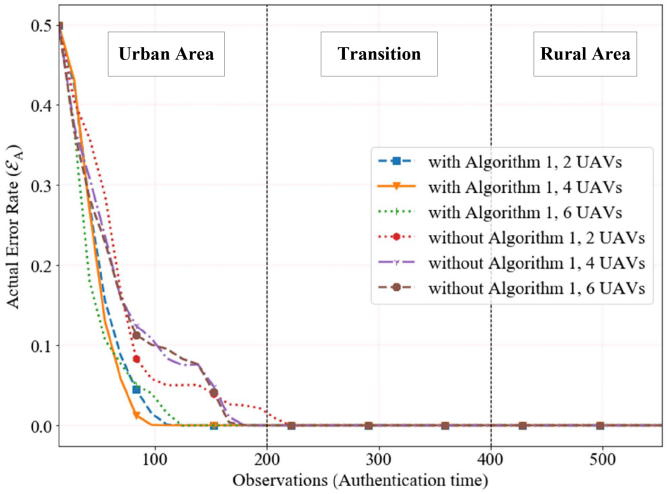


Fig. 7. Actual error rate comparison results with and without using Algorithm 1 in the UAV swarm

In Fig. 7, we compare the error rate of using Algorithm 1 to the scenario of employing all physical-layer attributes in different UAV swarm sizes. Similar to the previous figure, we set $\tau = 0.5$ to simulate the case using all available physical-layer attributes. It can be observed that our proposed scheme achieves a better or at least similar performance to that using all physical-layer attributes, which also verifies our analytical results. Although all techniques achieve an optimized authentication performance, when the number of observations increases, the Gini-impurity-based attribute evaluation algorithm has a better performance even when the number of observations is small. The associated reduced number of observations are extremely beneficial for resource-constrained applications, where it is more difficult to find the collaborative nodes for observation purposes.

Since the computational complexity is proportional to the number of selected collaborative nodes, one of the objectives in (6) is to minimize the number of selected collaborative nodes ($K$). To examine our scheme's ability to select the minimum number of collaborative nodes, we considered 3 different security requirements evaluated in terms of the operator-defined error rate as $\mathcal{E}_D = 0.1$, $\mathcal{E}_D = 0.01$ and $\mathcal{E}_D = 0.001$. To guarantee the SaS, the actual error rate ($\mathcal{E}_A$) should be lower than the operator-defined error rate ($\mathcal{E}_D$) despite utilizing a minimal number of authentication resources. The different selections of $\mathcal{E}_D$ reflect the unique security requirements in different scenarios ranging from civilian to military applications.

As shown in Fig. 8, when $\mathcal{E}_D = 0.1$, only 1 collaborative node is selected across different environments to guarantee the SaS. This demonstrates that only modest authentication resources are required to achieve a limited authentication requirement. When $\mathcal{E}_D = 0.01$, it can be observed that more collaborative nodes are selected at the beginning of the authentications compared to the result of Fig. 8. This demonstrates that a training stage is required at each node to approach the optimal performance. It can also be observed that during the environmental change, our proposed scheme

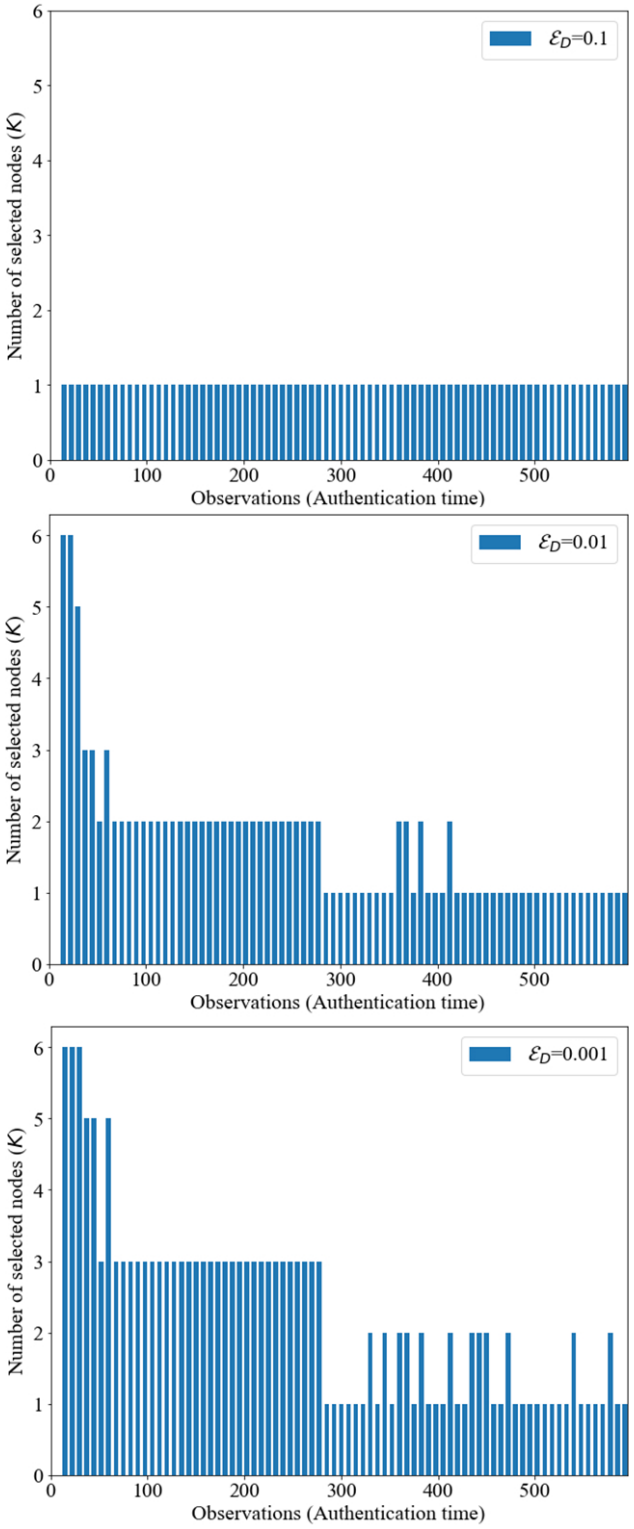


Fig. 8. Subplots of the Security requirements ($\mathcal{E}_D$) and the number of selected collaborative node(s)

can dynamically select a different number of collaborative UAVs to meet the authentication requirement. Then, to study the computational cost under an extremely strict authentication performance requirement, a subplot is also included to characterize the collaborative node selection at $\mathcal{E}_D = 0.001$. Similar to the case of $\mathcal{E}_D = 0.01$, our proposed scheme selected more authentication nodes during the environmental change to guarantee the authentication requirement. We can also observe that more collaborative nodes are selected across different environments. However, this demanding authentication performance can still be achieved without utilizing all possible authentication resources. When the collaborative nodes are well-trained, our proposed scheme succeeds in reducing the number of selected nodes to as low as one. Note that although only a single collaborative node is being selected, it is not equivalent to the centralized authentication schemes, since the single collaborative node selected can be different from one time instant to another. Our proposed scheme can intelligently select the most reliable combination of collaborative node(s) to achieve guaranteed performance with minimum effort in a distributed system.

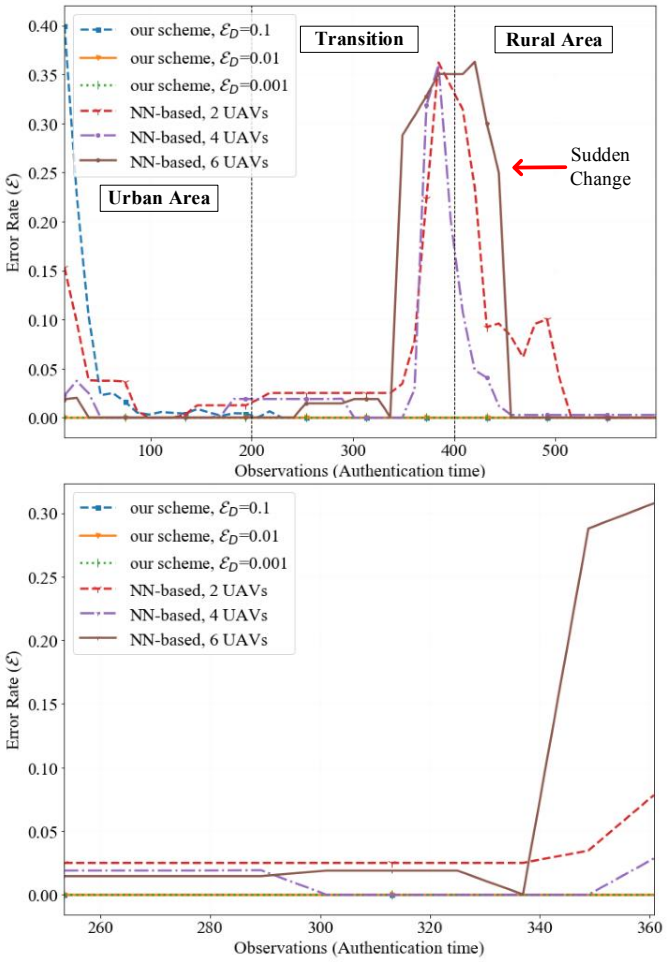


Fig. 9. Performance comparison between our proposed scheme and the NN-based centralized authentication scheme

Ultimately, to demonstrate that our proposed scheme performs better in satisfying the target authentication requirement,

Fig. 9 is given by considering the same security requirement for $\mathcal{E}_D = 0.1$, $\mathcal{E}_D = 0.01$ and $\mathcal{E}_D = 0.001$, as used in the previous step. It can be observed from the zoomed-in plot that the actual security performance ($\mathcal{E}_A$) can indeed satisfy the defined requirement, which successfully demonstrates that the SaS can be guaranteed with fewer authentication resources. For the authentication performance comparison, a nearest-neighbor-based (NN-based) centralized authentication scheme proposed in [55] is selected for testing under three different UAV swarm sizes of 2, 4 and 6 UAVs. It can be observed that the NN-based centralized authentication scheme has a similar trend in all three UAV sizes. Furthermore, compared to Fig. 7, the training stage of our proposed scheme is shorter, since the decisions gleaned from the reliable collaborative nodes are fused together. Moreover, to demonstrate the robustness of our proposed scheme, a sudden environmental change is simulated within the transitional stage, as labeled on the plot. It can be observed that the performance of our proposed scheme is significantly more robust against the sudden environmental change imposed.

## VI. CONCLUSION

An intelligent collaborative authentication scheme was proposed for employment in diverse environments. The novel SaS concept was conceived to achieve a specifically defined level of authentication with minimal authentication resources as an attractive design alternative to best-effort-based techniques. To achieve our ambitious design objective, the reliability of the physical-layer attributes had to be considered at each collaborative node. Hence, a Gini-impurity-based attribute evaluation algorithm was first developed to evaluate how likely the attributes would contribute to a reliable authentication decision. If none of the attributes were considered sufficiently reliable based on the operator-defined threshold ($\tau$) or no observation was made at the time instance, the collaborative node would be temporarily removed from the authentication process in conjunction with a usability index of 0. By applying this procedure, only the reliable physical-layer attributes of each node would be selected at each instance based on situational-awareness. Then, the collaborative node evaluation algorithm was proposed for quantifying the usability index of each collaborative node associated with a reliable physical-layer attribute by considering both the relative distance and the past contributions. The result is further utilized by the intelligent authentication customization algorithm to select the most suitable combination of collaborative nodes. This procedure guarantees to satisfy the specific authentication requirement at a minimal computational cost by comparing it to the decentralized authentication which utilizes all nodes. Finally, the proposed scheme was critically appraised against other state-of-the-art centralized authentication schemes to demonstrate its superior authentication performance versus computational complexity.

## REFERENCES


[1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[2] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 94–104, 2016.

[3] J. Zhang, G. Li, A. Marshall, A. Hu, and L. Hanzo, "A new frontier for IoT security emerging from three decades of key generation relying on wireless channels," *IEEE Access*, vol. 8, pp. 138406–138446, 2020.

[4] H. Fang, X. Wang, N. Zhao, and N. Al-Dhahir, "Lightweight continuous authentication via intelligently arranged pseudo-random access in 5G-and-Beyond," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 4011–4023, 2021.

[5] K. Xue, X. Luo, H. Tian, J. Hong, D. S. L. Wei, and J. Li, "A blockchain based user subscription data management and access control scheme in mobile communication networks," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2021.

[6] N. Hong, "A security framework for the Internet of Things based on public key infrastructure," in *Advanced Materials Research*, vol. 671, pp. 3223–3226, Trans Tech Publ, 2013.

[7] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2604–2616, 2021.

[8] S. Rosati, K. Kruelecki, G. Heitz, D. Floreano, and B. Rimoldi, "Dynamic routing for flying ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1690–1700, 2016.

[9] H. Bastami, M. Letafati, M. Moradikia, A. Abdelhadi, H. Behroozi, and L. Hanzo, "On the physical layer security of the cooperative rate-splitting-aided downlink in UAV networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 5018–5033, 2021.

[10] W. Alnumay, U. Ghosh, and P. Chatterjee, "A trust-based predictive model for mobile ad hoc network in Internet of Things," *Sensors*, vol. 19, no. 6, p. 1467, 2019.

[11] S. Bhandari, X. Wang, and R. Lee, "Mobility and location-aware stable clustering scheme for UAV networks," *IEEE Access*, vol. 8, pp. 106364–106372, 2020.

[12] U. S. D. of Defense, *Unmanned Systems Roadmap: 2007-2032*. AD-a475 002, Department of Defense, 2007.

[13] I. U. Khan, I. M. Qureshi, M. A. Aziz, T. A. Cheema, and S. B. H. Shah, "Smart IoT control-based nature inspired energy efficient routing protocol for flying ad hoc network (FANET)," *IEEE Access*, vol. 8, pp. 56371–56378, 2020.

[14] S. Hu, W. Ni, X. Wang, A. Jamalipour, and D. Ta, "Joint optimization of trajectory, propulsion, and thrust powers for covert UAV-on-UAV video tracking and surveillance," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1959–1972, 2021.

[15] H.-M. Sun, Y.-H. Chen, and Y.-H. Lin, "oPass: A user authentication protocol resistant to password stealing and password reuse attacks," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 651–663, 2012.

[16] H. Zhao, Y. Zhang, X. Huang, Y. Xiang, and C. Su, "A physical-layer key generation approach based on received signal strength in smart homes," *IEEE Internet of Things Journal*, vol. 9, no. 7, pp. 4917–4927, 2022.

[17] N. Aldaghri and H. Mahdavifar, "Physical layer secret key generation in static environments," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2692–2705, 2020.

[18] J. Zhang, S. Rajendran, Z. Sun, R. Woods, and L. Hanzo, "Physical layer security for the Internet of Things: Authentication and key generation," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 92–98, 2019.

[19] 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*, pp. 1–1, 2021.

[20] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2564–2573, 2012.

[21] F. Zhu, B. Xiao, J. Liu, and L.-j. Chen, "Efficient physical-layer unknown tag identification in large-scale RFID systems," *IEEE Transactions on Communications*, vol. 65, no. 1, pp. 283–295, 2017.

[22] X. Wang, P. Hao, and L. Hanzo, "Physical-layer authentication for wireless security enhancement: Current challenges and future developments," *IEEE Communications Magazine*, vol. 54, no. 6, pp. 152–158, 2016.

[23] W. Hou, X. Wang, J.-Y. Chouinard, and A. Refaey, "Physical layer authentication for mobile systems with time-varying carrier frequency offsets," *IEEE Transactions on Communications*, vol. 62, no. 5, pp. 1658–1667, 2014.

[24] 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.

[25] J. Liu and X. Wang, "Physical layer authentication enhancement using two-dimensional channel quantization," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4171–4182, 2016.

[26] H. Fang, X. Wang, and L. Hanzo, "Adaptive trust management for soft authentication and progressive authorization relying on physical layer attributes," *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2607–2620, 2020.

[27] 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.

[28] Y. Liu, H.-H. Chen, and L. Wang, "Physical layer security for next generation wireless networks: Theories, technologies, and challenges," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 347–376, 2017.

[29] Z. Gu, H. Chen, P. Xu, Y. Li, and B. Vucetic, "Physical layer authentication for non-coherent massive SIMO-enabled industrial IoT communications," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3722–3733, 2020.

[30] Y. Chen, P.-H. Ho, H. Wen, S. Y. Chang, and S. Real, "On physical-layer authentication via online transfer learning," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[31] H. Fang, X. Wang, and L. Hanzo, "Learning-aided physical layer authentication as an intelligent process," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2260–2273, 2019.

[32] S. Xia, X. Tao, N. Li, S. Wang, T. Sui, H. Wu, J. Xu, and Z. Han, "Multiple correlated attributes based physical layer authentication in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1673–1687, 2021.

[33] Z. Cui, F. Xue, S. Zhang, X. Cai, Y. Cao, W. Zhang, and J. Chen, "A hybrid blockchain-based identity authentication scheme for Multi-WSN," *IEEE Transactions on Services Computing*, vol. 13, no. 2, pp. 241–251, 2020.

[34] N. B. Truong, K. Sun, G. M. Lee, and Y. Guo, "Gdpr-compliant personal data management: A blockchain-based solution," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1746–1761, 2020.

[35] M. Zhaofeng, M. Jialin, W. Jihui, and S. Zhiguang, "Blockchain-based decentralized authentication modeling scheme in edge and IoT environment," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2116–2123, 2021.

[36] M. B. Mollah, J. Zhao, D. Niyato, Y. L. Guan, C. Yuen, S. Sun, K.-Y. Lam, and L. H. Koh, "Blockchain for the internet of vehicles towards intelligent transportation systems: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4157–4185, 2021.

[37] X. Duan and X. Wang, "Fast authentication in 5G HetNet through SDN enabled weighted secure-context-information transfer," in *2016 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2016.

[38] 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) (IEEE ICCC 2021)*, (Xiamen, China), July 2021.

[39] H. Forssell and R. Thobaben, "Worst-case detection performance for distributed SIMO physical layer authentication," *IEEE Transactions on Communications*, pp. 1–1, 2021.

[40] H.-L. Song and Y.-C. Ko, "Beam alignment for high-speed uav via angle prediction and adaptive beam coverage," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10185–10192, 2021.

[41] Y. Yuan, L. Wu, and X. Zhang, "Gini-impurity index analysis," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3154–3169, 2021.

[42] L. Jiang, B. Zhang, Q. Ni, X. Sun, and P. Dong, "Prediction of snp sequences via gini impurity based gradient boosting method," *IEEE Access*, vol. 7, pp. 12647–12657, 2019.

[43] K. Filus, P. Boryszko, J. Domaska, M. Siavvas, and E. Gelenbe, "Efficient feature selection for static analysis vulnerability prediction," *Sensors*, vol. 21, no. 4, 2021.

[44] J. Tan, S. Jing, L. Guo, and B. Xiao, "Ddos detection method based on gini impurity and random forest in sdn environment," in *2021 International Conference on Security, Pattern Analysis, and CyberneticsSPAC)*, pp. 601–606, 2021.

[45] H.-Y. Lin and Z.-Y. Wu, "Development of automatic gear shifting for bicycle riding based on physiological information and environment sensing," *IEEE Sensors Journal*, vol. 21, no. 21, pp. 24591–24600, 2021.

[46] Z. Yuan, H. Chen, T. Li, X. Zhang, and B. Sang, "Multigranulation relative entropy-based mixed attribute outlier detection in neighborhood systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–13, 2021.

[47] Y. Peng, Y. Yang, Y. Xu, Y. Xue, R. Song, J. Kang, and H. Zhao, "Electricity theft detection in AMI based on clustering and local outlier factor," *IEEE Access*, vol. 9, pp. 107250–107259, 2021.

[48] W. Wang and P. Lu, "An efficient switching median filter based on local outlier factor," *IEEE Signal Processing Letters*, vol. 18, no. 10, pp. 551–554, 2011.

[49] J. H. Aldrich and F. D. Nelson, *Linear probability, logit, and probit models*. No. 45, Sage, 1984.

[50] D. A. Hensher and W. H. Greene, "The mixed logit model: the state of practice," *Transportation*, vol. 30, no. 2, pp. 133–176, 2003.

[51] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on uav communications for 5g and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, 2019.

[52] H. Lei, D. Wang, K.-H. Park, I. S. Ansari, J. Jiang, G. Pan, and M.-S. Alouini, "Safeguarding uav iot communication systems against randomly located eavesdroppers," *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1230–1244, 2020.

[53] M. Al-Jarrah, A. Al-Dweik, E. Alsusa, Y. Iraqi, and M.-S. Alouini, "On the performance of irs-assisted multi-layer uav communications with imperfect phase compensation," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8551–8568, 2021.

[54] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[55] L. Senigagliesi, M. Baldi, and E. Gambi, "Comparison of statistical and machine learning techniques for physical layer authentication," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1506–1521, 2021.


## VII. Biography

**Huanchi Wang** (Member, IEEE) received his B.E.Sc and M.E.Sc degree in electrical engineering from Western University, Canada, in 2019 and 2021, respectively. His research interests include the machine learning, intelligent authentication and distributed security provisioning.

**Xianbin Wang** (Fellow, IEEE) received his Ph.D. degree in electrical and computer engineering from the National University of Singapore in 2001.

He is currently a Professor and a Tier-1 Canada Research Chair with Western University, Canada. Prior to joining Western University, he was with the Communications Research Centre Canada as a Research Scientist/Senior Research Scientist from 2002 to 2007. From 2001 to 2002, he was a System Designer at STMicroelectronics. His current research interests include 5G/6G technologies, Internet of Things, communications security, machine learning, and intelligent communications. He has over 500 highly cited journals and conference papers, in addition to 30 granted and pending patents and several standard contributions.

Dr. Wang is a Fellow of the Canadian Academy of Engineering and a Fellow of the Engineering Institute of Canada. He has received many prestigious awards and recognitions, including the IEEE Canada R. A. Fessenden Award, Canada Research Chair, Engineering Research Excellence Award at Western University, Canadian Federal Government Public Service Award, Ontario Early Researcher Award, and six IEEE best paper awards. He was involved in many IEEE conferences, including GLOBECOM, ICC, VTC, PIMRC, WCNC, CCECE, and CWIT, in different roles, such as General Chair, Symposium Chair, Tutorial Instructor, Track Chair, Session Chair, TPC Co-Chair, and Keynote Speaker. He serves/has served as the Editor-in-Chief, Associate Editor-in-Chief, and editor/associate editor for over ten journals. He was the Chair of the IEEE ComSoc Signal Processing and Computing for Communications (SPCC) Technical Committee and is currently serving as the Central Area Chair for IEEE Canada.

**He Fang** (Member, IEEE) is a full professor with the School of Electronic and Information Engineering, Soochow University, China. She received her Ph.D. degree in Electrical and Computer Engineering from Western University, Canada, in 2020. Her research interests include intelligent security provision, trust management, machine learning, distributed optimization and collaboration techniques. She currently serves as a Guest Editor and Topical Advisory Panel Member for several journals, and was involved in many IEEE conferences including IEEE GLOBECOM, VTC, and ICCC, in different roles such as Session Chair and TPC member. She also served as the Vice-Chair of Communication/Broadcasting Chapter, IEEE London Section, Canada, from Sep. 2019 to August 2021.

**Lajos Hanzo** (http://www-mobile.ecs.soton.ac.uk, https://en.wikipedia.org/wiki/Lajos_Hanzo) (FIEEE'04) received Honorary Doctorates from the Technical University of Budapest and Edinburgh University. He is a Foreign Member of the Hungarian Science-Academy, Fellow of the Royal Academy of Engineering (FREng), of the IET, of EURASIP and holds the IEEE Eric Sumner Technical Field Award.