Trust-Aware Authentication and Authorization for IoT: A Federated Machine Learning Approach

Kazi Istiaque Ahmed, Mohammad Tahir, *Senior Member, IEEE,* Sian Lun Lau, *Senior Member, IEEE,* Mohamed Hadi Habaebi, *Senior Member, IEEE*, Abdul Ahad, and Amna Mughees

***Abstract*—The need for strong authentication and authorization (AA) security measures is growing with the proliferation of the Internet of Things (IoT). This paper presents an advanced trust-aware authentication and authorization system for IoT environments. Using real-world data collected from Zigbee Zolertia Z1 devices, a Federated Machine Learning model was developed that utilizes Physical Layer properties such as Received Signal Strength Indicator (RSSI), Link Quality Indicator (LQI), device Internal Temperature, device Battery Level, and device MAC address. The proposed solution for AA IoT utilizes a trust calculation algorithm based on Federated Learning (FL), which is suitable for IoT environments and enables data privacy and scalability. Incorporating device-specific information, such as internal temperature and battery level, helps a more nuanced evaluation of the device's status, improving the precision of trust calculations. The proposed architecture performs particularly well for unauthorized intrusion attempts modelled using spoofing, replay and Sybil attacks. Specifically, the proposed methodology can detect malicious AA activities classified as Writing + Reading attempts with 100% accuracy, demonstrating its effectiveness in protecting IoT devices from attacks. Furthermore, the model achieves 99.18% accuracy in reading access permissions and 99.99% accuracy in identifying Write + Read + Execute permissions, highlighting its reliability in implementing access control restrictions for improving security in IoT environments. This research helps improve IoT security by addressing crucial challenges in the ever-expanding world of networked devices.**

***Index Terms*—Internet of Things (IoT), Machine learning (ML), Artificial Neural Networks (ANN), Federated Learning (FL), Authentication, Authorization, Access control, Security, Networking, Trust, and Trust Management.**

I. Introduction

THE proliferation of Internet of Things (IoT) devices has transformed how we engage with the digital world, penetrating various facets of our daily lives, ranging from smart homes to industrial automation. Currently, there are nearly 23 billion devices connected to the Internet. If the forecast is to be believed, the number will rise to 51 billion by 2025 [1]–[3] and will continue to increase rapidly. Providing adequate security procedures becomes increasingly essential as the IoT ecosystem evolves, particularly in contexts where crucial data and sensitive information are exchanged. Authentication and Authorization, two critical components of IoT security, prevent undesirable activities by malicious actors on protected resources.

A. Security Challenges

The information sharing between nodes must be appropriately protected and recognized in applications facilitated by the IoT network. Authentication and Authorization (AA) is the first primary stage. AA is intended to restrict actions and operations on IoT environments that authenticated users can do [4]. AA aims to prevent safety breaches through the enforcement of user restrictions, resulting in the exposure of essential resources to unwanted opponents. Several competing solutions and methodologies are available for AA, which makes it difficult to have a standardized method that can be applied to IoT networks. For example, OAuth [5] provides safe delegated third-party access and is one example of a standard and widely used authorization scheme. OAuth restricts access to resources on behalf of the resource owner through a centralized system.

Furthermore, ensuring the trustworthiness of IoT nodes in the given field is an essential task during message exchange. Security is a major concern when it comes to the IoT, and building trust is essential to ensure secure operations on IoT devices. This includes secure booting, key management, data protection, secure communication sessions, safe patching of hardware and software, and monitoring and auditing of data exchange between devices. Besides, sharing data among IoT nodes directly affects IoT protection and is highly based upon the trustworthiness of the data and the standard of service rendered.

In this regard, this paper presents a novel trust-aware authentication and Authorization for IoT networks, emphasizing single-node and multi-node scenarios. Our methodology introduces a novel perspective by incorporating physical layer properties such as RSSI, Link Quality Indicator (LQI), device internal temperature, device battery level, and the device's

Corresponding Authors: Mohammad Tahir

Kazi Istiaque Ahmed, Sian Lun Lau, and Amna Mughees are with the Department of Computing and Information Systems, Sunway University, Petaling Jaya 47500, Selangor, Malaysia; kazi.i@imail.sunway.edu.my; sianlunl@sunway.edu.my; 19099621@imail.sunway.edu.my;

Mohammad Tahir is with Department of Computing, University of Turku, FI-20014, Turun Yliopisto, Finland; tahir.mohammad@utu.fi; He is also Adjunct Professor at Chitkara University Institute of Engineering and Technology;

Mohamed Hadi Habaebi, is with IoT & Wireless Communication Protocols Lab, Department of Electrical and Computer Engineering, International Islamic University Malaysia, Gombak 53100, Selangor, Malaysia; habaebi@iium.edu.my

Abdul Ahad is with the School of Software, Northwestern Polytechnical University, Xian, Shaanxi, 710072, P.R. China; He is also Adjunct Scientific Research at the Department of Electronics and Communication Engineering, Istanbul Technical University (ITU), 34467, Istanbul, Turkey; ahad9388@gmail.com

Media Access Control (MAC) address as critical parameters for trust computation based on real-world data collected from Zigbee Zolertia Z1 devices.

Furthermore, Unlike traditional systems that rely primarily on cryptographic methods (such as RSA [6], Secure Hash [7], etc.) or conventional authentication protocols [8], our architecture takes advantage of Federated Machine Learning. Machine learning (ML) based physical layer authentication (PLA) has gained significant interest in recent years due to its ability to achieve threshold-free authentication and higher performance [9]. Incorporating physical layer attributes in our trust computation model recognizes IoT devices' intrinsic qualities that go beyond software-based considerations. This novel combination of trust computing and ML enables the system to dynamically adapt to changing threat environments and operational situations.

B. Motivation

The motivation for pursuing Trust-based Authentication and Authorization for IoT stems from the escalating deployment of interconnected devices in diverse applications, ranging from smart homes to industrial automation. As the IoT landscape expands, ensuring the security and trustworthiness of data exchanged between IoT devices becomes imperative. Existing solutions often fall short in addressing the unique challenges posed by resource-constrained devices, necessitating the development of a tailored approach. Collecting new data is critical to understand this ecosystem's specific security requirements and dynamics comprehensively. Utilizing existing data is limited in its applicability, as it may not capture the intricacies of IoT devices and their communication patterns. Furthermore, the need for Federated Learning (FL) arises due to IoT networks' distributed and decentralized nature. Traditional centralized models struggle with scalability and may compromise user privacy. FL enables the training of models across multiple devices, preserving data privacy and ensuring a more resilient and adaptable solution to the evolving security challenges in IoT [10]–[12]. Moreover, in federated learning, the weights are only shared instead of the actual data to save bandwidth, energy, and computation and provide a secure form of exchanging data, making it suitable for IoT devices [13].

C. Contribution

The article proposes a novel method for authentication and authorization using trust management based on the physical layer characteristics. In particular, the main contributions are as follows:

1) A trust-aware authentication and authorization system for IoT environments, utilizing federated learning that incorporates physical layer properties such as RSSI, LQI, device internal temperature, battery level, and device MAC address for enhanced security.
2) Addressing data privacy and heterogeneity by employing federated learning, our framework ensures data privacy and supports the decentralized training of models, effectively handling the diversity of IoT devices and operational conditions.
3) Our proposed methodology is grounded in practical application, demonstrated through the use of data collected from Zigbee Zolertia Z1 devices. This not only validates the model under real-world conditions but also ensures its relevance and applicability to IoT networks.
4) Comprehensive evaluation of the proposed methodology to accurately detect unauthorized access attempts with high precision underscoring the effectiveness of the proposed solution in protecting IoT devices from unauthorized access.

D. Organization of article

The rest of the paper is organized as follows: Section-II comprehensively analyzes related work on IoT security, identifying gaps and emphasizing the importance of advanced trust-aware techniques. Section-III presents the data collection procedure for training ML models, describing real-world data collections from Zigbee Zolertia Z1 devices. The section also introduces the federated machine learning model for trust management for AA. Section-IV describes the experimental setup and findings, which demonstrate the effectiveness of our strategy in both single-node and multi-node IoT contexts. Finally, Section-V summarizes significant findings, discusses ramifications, and suggests avenues for further research in the dynamic landscape of IoT security.

II. Related Works

Recent research has focused on improving security in IoT systems by developing different methodologies and schemes for device-based authentication and authorization. The following subsection briefly discusses various methods and advances in AA schemes for IoT. Table-I summarizes existing ML-based AA techniques, the gaps in existing techniques of IoT authentication and authorization schemes and how the proposed work fills the gaps in the existing literature.

A. General AA's

A technique presented in [14] utilizes mutual authentication to establish a secure communication channel between clients and servers. This is achieved by employing the lightweight Application Layer Protocol CoAP and the Advanced Encryption Standard (AES) for further security. The suggested approach also deals with the resilience, real-time registration of IoT nodes, and performance metrics, such as average response time and duration of the handshake process. However, privacy preservation and the placement of pre-shared keys during the provisioning phase are not explicitly considered. Work in [15] proposes a streamlined two-factor authentication system that employs digital signatures and device capacity. This approach seeks to enhance security at both the application and physical levels by implementing a secure TLS channel and addressing the vulnerability to Man-in-the-Middle (MITM) attacks.

Various studies investigate authentication techniques specifically designed for individual IoT applications. An example is

the study by Hamidi et al. [16], which explicitly examines the authentication of the Internet of Medical Things (IMT) and utilizes biometric techniques to ensure a safe connection. The work in [17] introduces a novel authentication technique called Bubbles-of-Trust, which uses blockchain technology to establish virtual zones for grouping devices. However, using a public blockchain presents difficulties, such as the duration required for transaction validation and the lack of efficiency in applications that require real-time processing. In addition, the authors of [18] introduce a two-factor authentication method for IoT devices connected to cloud computing. This method focuses explicitly on password registration, verification, and updating. These studies provide essential insights for improving authentication procedures in various IoT applications. However, it is important to note that each solution has limitations when applied to resource-constrained devices.

B. Trust Based AA

Authors in [19] propose a computational strategy that establishes trust in the social Internet of Things (SIoT) by utilizing reputations, suggestions, and knowledge. Although it improves efficiency and productivity in specific contexts, such as for service distributors, its reliability is constrained. Authors of [20] present a reliable framework for discovering services that consider quality of service (QoS) and use ontological models for matching services, propagating trust, and making decisions for clients. A hierarchical blockchain-based method for safeguarding identification in the context of mobility support is presented in [21]. However, their methodology fails to address essential trust factors such as credibility and accuracy. Abderrahim et al. [22] propose a clustering framework to manage trust in IoT that utilises the Kalman filter to forecast confidence values. However, the proposed work faces challenges in dealing with diverse IoT environments. A trust analysis and forecasting model focusing on data is presented in [23], but the ability to handle large amounts of data is restricted. Researchers of [24] provide a trust-based decision-making framework for the health system in the IoT, adjusting to declining trust levels and optimizing decision precision. However, it considers an IoT setting exclusively comprising sensor nodes. Chen et al. [25] provide a trust management technique that utilizes several approaches, such as maximum ratio combining (MRC) and selection combining (SC). However, the scalability of this strategy is not assured due to the limited range of nodes in the testing phase. In [26], a proposed sensing technique prioritizes trustworthiness and is guided by policies. This strategy relies on descriptive information and aberrant IoT details. However, a drawback of this approach is that it may mistakenly see an outdated policy as malicious, resulting in possible problems.

The study of Caminha et al. [27] propose a Smart Middleware Architecture that automatically identifies IoT devices. This architecture also includes calculating semantic attributes to estimate the trust value of these devices. The objective is to establish reliable communication between IoT nodes. The paper addresses the security difficulties and presents a Trust Aggregation Authentication Protocol that utilizes Machine Learning (TAAPML). The protocol, outlined in [28], utilizes Support Vector Machines (SVM) to determine trust thresholds dynamically. It surpasses the Trust Management Model (TMM) in terms of both efficiency and effectiveness across several measures.

C. ML-Based AA

The work in [38] examines the issue of safeguarding IoT devices at the physical layer, where conventional cybersecurity methods are insufficient. The authors suggest employing PHY-layer authentication methods that leverage the spatial correlation of wireless channel attributes, including RSSI, RSS, CIR, CSI, and MAC addresses. The authors utilize hypothesis testing to compare the PHY-layer attributes of the received data with the transmitted data, where the authentication performance is contingent upon threshold values. To address the difficulties associated with selecting appropriate thresholds, reinforcement learning (RL) is used, which treats authentication as a Markov decision process.

An unsupervised learning authentication method utilizing IGMM is presented in [39]. The study highlights the significance of using lightweight authorization mechanisms for IoT nodes with limited resources. The authors investigate using RL techniques, namely Q-learning, to optimize network security against assaults at the physical or media access control (MAC) layer, such as jamming and spoofing [40].

The authors of [29] present a deep learning classifier that utilizes LSTM networks to identify low-power IoT devices even when adversaries are present. The classifier is shown to be resilient to signal disruptions. The study by Hammad et al. in [41] introduces a novel way for real-time authentication on edge computing platforms utilizing convolutional neural networks (CNNs) and electrocardiogram (ECG) data. The study demonstrates that this approach outperforms conventional strategies in terms of performance. Li et al. in [42] present a solution to the difficulties in handling and distributing keys in extensive IoT networks, which

achieves a high accuracy rate of 94.7%. In the realm of training, machine-learning-based physical layer authentication strategies for IoT, the study conducted by the authors of [32] introduces the utilization of data augmentation through deep neural networks. The results demonstrate enhanced robustness and authentication rates. Similarly, in [30], the authors investigate authentication in IoT using CNN and Random Projections with an accuracy of 98.08%. However, it is observed that it is vulnerable to noise when the signal-to-noise ratio is low. The study in [37] investigates using deep reinforcement learning and Channel State Information (CSI) for identity authentication. The research demonstrates that this approach can achieve efficient and dependable authentication without intricate cryptographic techniques. In addition, authors in [31] suggest a centralized Adaptive Neural Network (ANN) for intelligent authentication at the physical layer. This solution surpasses current approaches in communication scenarios that involve temporal fluctuations.

The method in [43] requires minimal computational resources, making it a cost-effective solution for large-scale deployments in mobile wireless networks with the k-Nearest

TABLE I
SUMMARY OF EXISTING LITERATURE

Techniques	Year	Input Parameters			A	A*	Perf. Metrics		Attacks		Charac		Limitation
		RF	CSI	Others			Accuracy	MCR	Sp	Sy	C	D	
LSTM [29]	2018	✓			✓		97.75	2.25			✓		No detailed discussion on the selected Machine Learning Technique. Moreover, FAR, FRR, False/True Positive, and False/True Negative have not been analyzed
CNN [30]	2019	✓			✓		98.08	1.92			✓		Other than Accuracy and MCR, no other statistical measures are considered or been analyzed.
Adaptive NN [31]	2020	✓			✓		95.89	4.11	✓		✓		Other than Accuracy and MCR, no other statistical measures are considered or been analyzed. Moreover, the numerical results of authentication are not discussed.
DNN [32]	2020	✓			✓		93.70	6.30	✓		✓		FAR, FRR, False/True Positive, False/True Negative have not analyzed (can lead to diversified insights). Moreover, the accuracy decreased when the number of requests increased, and the Sybil attack was considered but not discussed.
SVM [33]	2021		✓		✓		90.00	10.00		✓	✓		FPR and TPR discussed without numerical results. Moreover, other performance measures matrices are not considered.
FL/SL [34]	2021			✓	✓		92.16	7.84			✓		No discussion on the selected ML/DL techniques with optimization Function. Moreover, no discussion on the mobile active authentication attacks.
F-DNN [12]	2021	✓			✓		94.16	5.84				✓	Only SGD optimization function which has limited the system to low heterogeneity problem. Moreover, no results have been shown when there is more than one device.
ANN [35]	2022			✓	✓		75.00	25.00				✓	Not Autonomic and Other benchmarking Metrices are not discussed. However, there was no discussion of the attacks.
CART [36]	2022			✓	✓		95.43	4.57	✓		✓		Other statistical measures are not considered or been analyzed. ID spoofing was considered, but there was no discussion on how to tackle node spoofing.
DRL [37]	2023		✓		✓		96.10	3.90			✓		Firstly the model is complex. On the other hand, other statistical measures have not been considered or analyzed.
Proposed work	2024		✓	✓	✓	✓	99.79	0.21	✓	✓	✓	✓	Limited to Stationary Nodes

A=Authentication, A*=Authorization, Perf.=Performance, Sp=Spoofing, Sy=Sybil, Charac=Characteristic, C=Centralized, D=Distributed

Neighbors (k-NN) algorithm. It also improves the robustness of phy-layer authentication (PLA) by accurately identifying legitimate nodes based on unique channel characteristics.

The authors of [11] explore privacy-conscious access control in healthcare systems enabled by the IoT. Federated deep learning is used to improve accuracy and establish trustworthy connections. The federated deep learning approach attains an overall accuracy rate of 90%, with a false acceptance rate (FAR) of 15% and a Misclassification Rate (MCR) of 13%, highlighting the importance of dynamic and context-aware access management. Yazdinejad et al. [12] investigate the issue of authentication by employing federated learning with DNN+SGD for RF signal and achieving an Accuracy of 90.7%.

Based on the existing literature and summary provided in Table-I that in the existing literature, various schemes have been proposed in the literature for IoT AA. The most considered input parameters are MAC, RSSI, RF fingerprint, CSI, and CIR and are considered either centralized or distributed implementations.

In addition, the issue of resource constraints for IoT devices is significant for developing any Phy-Layer security schemes, which has posed significant challenges in implementing effective security solutions. Moreover, implementing either centralized or distributed AA techniques is not enough to tackle the security and privacy of IoT nodes.

Therefore, in this work, we propose a security solution to implement AA at the physical layer using the phy-layer characteristics utilizing a dataset collected from IoT nodes (Zolertia Z1 nodes). Furthermore, the proposed solution incorporates hybrid implementation through the use of Federated Learning to address the issue of resource limitations on the nodes.

Therefore, our proposed scheme will fill the gap using not only MAC but also the RSSI, LQI, Battery Level of the device, and Device Internal temperature from the real dataset with a federated machine learning model. The proposed solution can be implemented in any communication stack (e.g. TCP/IP) without affecting the other layer, thereby adding another layer of security at the physical layer, which usually does not have any security mechanism implemented. This type of security is

also referred to as physical layer security and is an active area of research for future networks such as 6G.

III. System Model

The system model considered is depicted in Figure-1. In the proposed model, we incorporate FL for authentication and authorization of the IoT nodes. This is an alternative methodology for training machine learning models on a distributed approach, encompassing several IoT or edge devices. Unlike traditional ML, this approach eliminates the need to centralize data, safeguarding data privacy. Using FL for AA involves using distributed controllers in the Local nodes, such as gateway/edge, to gather data from constrained devices and run local ML tasks. Whereas, the centralized controller in the cloud contains the global model visibility of the network topology. Solely relying on centralized access management may result in a single point of failure and challenges related to scalability, particularly as the network expands. However, distributed controllers that contain a copy of the global model provide resilience and scalability by assigning control to specific nodes, although they may lack overall visibility and coordination. The FL framework achieves a balance between centralized supervision and decentralized decision-making by utilizing both centralized and distributed controllers. This hybrid approach utilizes the advantages of each model: centralized access control enforces global policies and coordinates actions, and distributed access control allows local independence and resilience to faults. This guarantees effective administration while improving the system's resilience and scalability, rendering it highly suitable for dynamic IoT contexts.

The proposed system model consists of various IoT devices connected to the gateway. The role of these IoT devices is to observe and transmit the information to centralized and distributed controllers. The Gateway Nodes(GWNs) participate in communication and share model weights with the global TM node for aggregation of weight in FL. The global TM node will centrally establish, store, and manage trust among all IoT and Gateway Nodes (GWNs). The distributed model, embedded in the GWNs of the IoT network layer, uses the Feed-Forward Neural Network of ANN to estimate and aggregate the local trust levels based on the trust values of the forwarding devices in their respective neighbourhood over time for AA utilizing various physical layer features.

Utilizing the RSSI and LQI as wireless channel fingerprints for node identification is a significant AA method at the physical layer. The distinctive attributes of the wireless channel, including its temporal and spatial properties, enable the mapping of diverse sites with varying spatiotemporal environmental factors. Furthermore, additional parameters, like the node temperature and the battery level, can enhance the ability to make more precise decisions for the AA process. During communication, the instantaneous RSSI, LQI, and other relevant metrics obtained while accessing the Edge/Gateway can differentiate between legitimate and illegitimate nodes at varying distances. Other features can be retrieved from the instantaneous RSSI and LQI to represent the Spatiotemporal environment accurately, which enables the identification of transmitting nodes located at various places.

A. Dataset Collection

The following steps were carried out during dataset generation. A detailed discussion on the hardware setup and data collection procedure to create the real dataset is presented in [44].

- **Data Collection Setup:** Zigbee Zolertia Z1 nodes with 802.15.4 low-power, short-range radios are used in a series of systematic experiments to collect data on physical layer properties relevant to AA in indoor environments.
- **Parameters Captured:** The dataset includes measurements such as Received Signal Strength Indicator (RSSI), Link Quality Indicator (LQI), device internal temperature, battery level, internal acceleration (across X, Y, and Z axes), Channel Check Rate, Radio Channel, and Transmission power (Tx in dBm). Altogether there were six main features and two derived features, as shown in Table-II.
- **Stationary Nodes:** During experiments, nodes were stationary and strategically placed to collect data from various ranges and antenna orientations to reflect realistic usage scenarios.
- **Scenarios and Antenna Orientations:** Three end nodes were fixed at distances of $1m$, $2m$, and $3m$ from a gateway node, with data collected over 24-hour intervals to assess the impact of day and night temperatures on key parameters like RSSI and LQI. Antenna orientations of $90, 180$, and 0 degrees were explored to evaluate their impact on the dataset.
- **Data Instances:** A total of $347,200$ instances were collected, with detailed breakdowns based on distance from the gateway node and antenna orientation, providing a rich source for evaluating AA mechanisms in IoT environments.

B. Threat Model

The wireless channel openness renders the Edge/Gateway vulnerable to spoofing attacks or impersonation attacks when IoT nodes connect to the Edge/Gateway. Spoofing attacks involve malicious nodes capturing the identity of legitimate nodes and subsequently monitoring or tampering with the transmitted data. On the other hand, Sybil nodes can assume multiple false identities of legitimate nodes, leading to the initiation of rapid access requests to the Edge/Gateway. Consequently, this excessive demand on Edge/Gateway resources causes network congestion, leading to a Denial of service (DoS) attack. The attack model considered in this work is presented in Figure-2. The work considers diverse physical layer features instead of using only MAC addresses to identify and counteract attacks in wireless networks. This model primarily emphasizes the following features:

1) **Received Signal Strength Indicator (RSSI):** RSSI quantifies the power level a radio receiver receives upon detecting a radio frequency carrier signal. The transmission

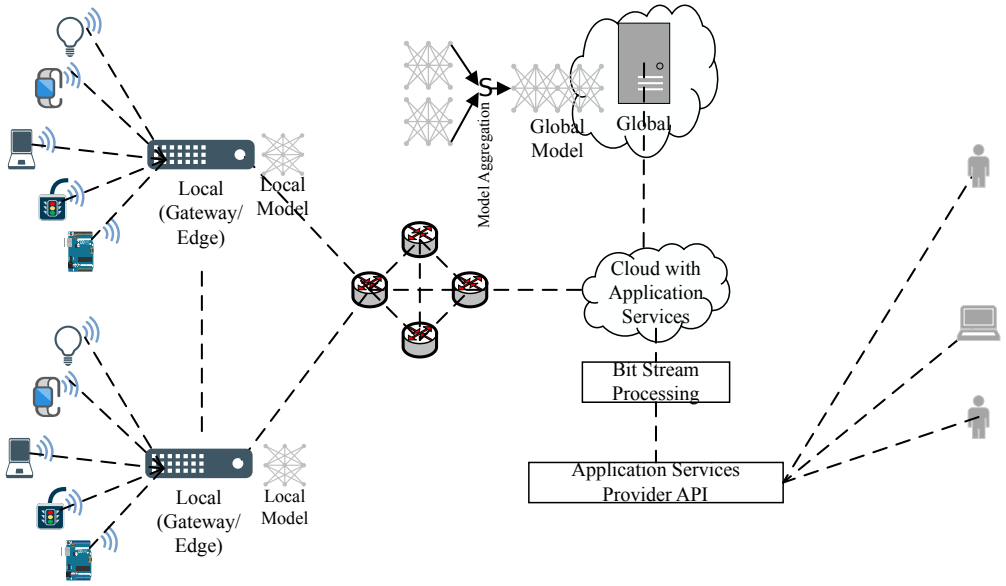

Fig. 1. Trust-Aware AA System model Architecture.

TABLE II
DATASET FEATURES [44]

Parameters	Description
Node MAC	Node Media Access Control (MAC) Address
Node RSSI (dBM)	Received Signal Strength Indicator (RSSI) in dBM
Node LQI	Link Quality Indicator (LQI), which indicates the Link Quality
Node Temperature (Celsius)	Node Temperature in (Celsius)
Node Battery level (mV)	Battery Level of the node in millivolts (mV)
Node Antenna Orientation	Antenna Orientation of the sender in degrees (90/0/180)
Node Temperature difference (derived)	Temperature difference between past two readings (Celsius)
Node battery difference (derived)	Battery level difference between past two readings (mV)

power, the distance between the transmitter and receiver, and environmental conditions like obstructions and interference exert impact. In spoofing detection, RSSI data may be used to establish each device's fingerprint profile. Authentic devices display consistent RSSI patterns, but faked devices may reveal irregularities attributable to their distinct physical locations or transmission characteristics.

2) **Link Quality Indicator (LQI):** LQI quantifies the communication link quality between two devices. It is frequently utilized alongside RSSI to offer a more thorough perspective of the communication channel. LQI aids in the detection of spoofing attempts by recognizing inconsistencies between anticipated link quality and historical data. A significant decline in LQI without a concomitant alteration in RSSI may suggest a spoofing effort.
3) **Battery Status:** The battery level of a node may indicate its operating state and assist in identifying spoofing attacks. Authentic nodes provide consistent battery consumption patterns, however counterfeit nodes may display irregular battery usage due to the extra energy needed to sustain the fraudulent identity. Monitoring battery levels can assist in detecting nodes that exhibit abnormal behaviour, hence facilitating the identification of spoofing attempts.
4) **Internal Temperature of Nodes:** The internal temperature of a node may serve as a physical layer metric for detecting spoofing. Authentic nodes have steady temperature profiles, however counterfeit nodes may display irregular temperature variations due to the extra processing necessary to sustain the falsified identity. Monitoring the interior temperature enables the identification of nodes under stress or functioning abnormally, perhaps signalling a spoofing assault.

By including these physical layer parameters—RSSI, LQI, battery level, and nodes' internal temperature—the PHY-layer threat model establishes a comprehensive framework for identifying and countering considered attacks in IoT networks. This method utilizes the distinct attributes of each parameter to develop a robust security framework that is more resistant to spoofing tactics.

C. Trust-aware AA and Attack Detection

Algorithm-1 presents a trust establishment approach for IoT devices. This method involves monitoring the RSSI, LQI, Temperature, Battery Level (mV), Radio Channel, and antenna Orientation of the Gateways and their connected nodes to detect abnormal resource consumption patterns. Specifically, it identifies devices that exhibit unusual consumption behaviours and deviate from the expected minimal regular consumption, indicating potential failures or malicious IoT

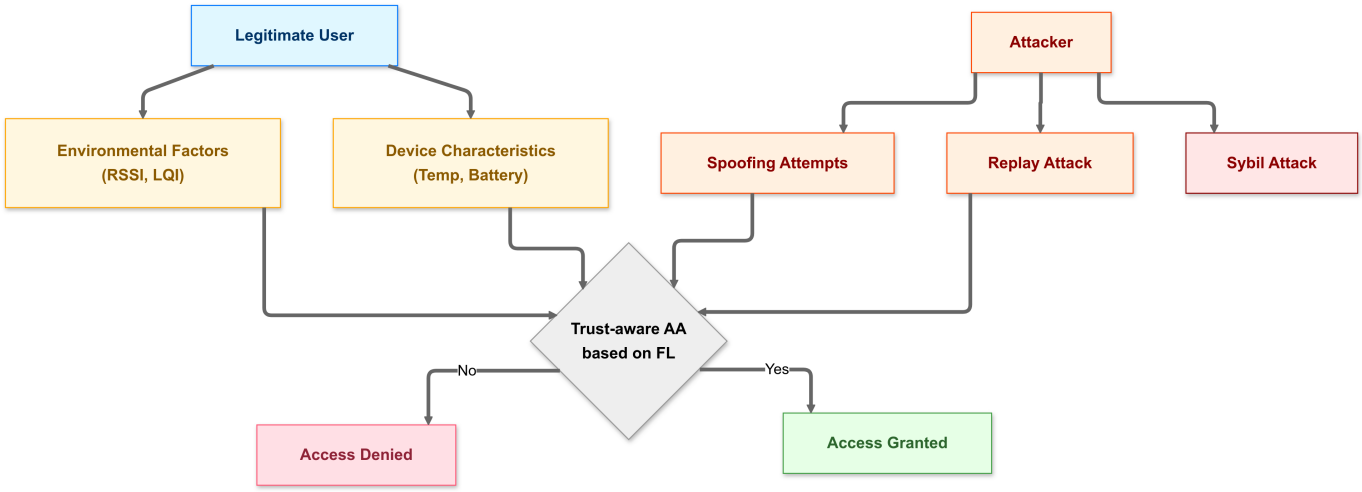

Fig. 2. Threat model for Trust-Aware AA.

devices. Identifying devices that allocate insufficient resources to fulfil the FL tasks and those that demonstrate excessive resource consumption, which may suggest potential malicious behaviour, is crucial. As an illustration, specific malicious devices may prioritize certain actions to produce specific misclassifications. These devices are anticipated to allocate more resources than conventional devices that solely focus on optimizing the underlying federated task.

Each edge server is responsible for monitoring IoT devices in its designated coverage area. Therefore, Algorithm-1 is executed on every edge server. The method being proposed leverages the modified Z-score statistical technique [45]. The Modified Z-score is a standardized metric used to assess the magnitude of outliers, precisely the extent to which a given score deviates from the average score by evaluating the relationship between a specific score and a designated typical score. This approach demonstrates a higher level of resilience to extreme values in comparison to certain alternative statistical methods, such as the traditional Z-Score, Tukey method, and others. This is achieved by leveraging the median $\bar{x}$ with μ instead of the only mean μ. This method estimates the disparity between a specific score and the median value within the presented algorithm. This estimation is achieved by utilizing the median absolute deviation (MAD) of a given metric, such as RSSI, LQI, Temperature, Battery Level (mV), Radio Channel, Or Antenna Orientation exhibited by a device (referred to as "j") over a defined time window $[t-\delta, t]$. This specific line of code can be found in Algorithm-1, the line where $MAD_j^z(t)$ is computed.

The modified Z-score $\alpha_j^z(i,t)$ is determined by dividing the discrepancy between the consumption $TestNodeMetrics_i$ of device j in relation to the resource metric z at the time i within the time interval $[t-\delta, t]$ and the median consumption of that device in terms of the same metric within the time interval $[t-\delta, t]$, by the median absolute deviation of the metric z (as stated in the variables $MAD_j^z(t)$ of Algorithm-1).

Outliers are designated as such when the $\alpha_j^z(i,t)$ exceeds the threshold φ, which has been established as 3.5 according to the argument presented in reference [45]. As mentioned earlier, the limit quantifies each IoT device's maximum and minimum habitual utilization patterns within a specific time frame. Therefore, any subsequent consumption that surpasses or falls below this threshold is considered atypical. The algorithm subsequently examines potential future usage of the IoT to identify any instances of usage that exceed or fall below the calculated abnormal threshold (Algorithm-1 `outlier_detection (OD) 1,2,3` section. If such a scenario arises, this particular observation is recorded in a tabular format that documents any atypical consumption of the IoT device, as indicated in Algorithm-1, specifically on lines where stated, removing the first row from $BaseMetrices$ and Adding $TestNodeMetrics_j$ at the end of $BaseMetrices$. The mean atypical consumption for each metric is subsequently calculated. The algorithm computes the trust value of each IoT device by dividing the sum of the proportional abnormal consumption across all metrics by the number of metrics that the device has either overused or underused, indicated as individual trust of the input parameters $[T1, T2, T3, T4]$ in line-5,7,9, and 10 of Algorithm-1. If there is no excessive or insufficient utilization of any metric, the initial level of trust in the IoT system's reliability would be assigned a value of 1. This value signifies complete trust in the respective device, and based on the trust levels Γ_j 0.00, 0.70, 0.85, and 1.00, the device will get authorized to perform the authorized task as shown in the Table-III where E = Execute, W=Write, R=Read, and D=Denied. On the Other hand, the trust levels are considered as either $\Gamma_j = 0.00$ or 1.00 when OD observes $\alpha_j^z(i,t) > \varphi$ or $\alpha_j^z(i,t) \leq \varphi$ for the attack detection scenario illustrated in Figure-4.

TABLE III
TRUST LEVELS AND AUTHORIZATIONS

Trust Levels	AA
$\Gamma = 1.00$	E + W + R
$\Gamma = 0.85$	W+R
$\Gamma = 0.70$	R
$\Gamma = 0.00$	D

Algorithm 1: Trust Establishment Model

Input : j: an IoT, monitored by Edge Server, M = MAC,$RSSI$,LQI,$TEMP$,$BATT$, $PREV\_RSSI$,$PREV\_LQI$,$PREV\_TEMP$, $PREV\_BATT$,$PREV\_Access\_Level$: Set of IoT's metrics for analyzing by Edge, δ: the size of the time window after which the algorithm is to be repeated X_{Max}: Device max Battery Level

Variables: R: A table recording the registered MAC address of IoT devices, $M_j^z(t)$: a table recording $x_j^z(i)(i = t-\delta, t-\delta+1, ..., t)$, the RSSI, LQI and Temperature and Battery Level difference of $z \in M$ j during the time interval $[t-\delta, t]$

$MAD_j^z(t)$: the median absolute deviation of $M_j^z(t)$ with mean μ, i.e., $MAD_j^z(t) = median|x_j^z(i) - x_j^z(t)|$ for all $t-\delta \leq i \leq t$

Output : Γ_j: trust value of j.

Initialize $BaseMetrices = M_j^z \in \delta$; $TestNodeMetrics_j$ to 0; **foreach** *Rows* $z \in M$ **do**
  if $M_j^{MAC} \in R$ **then**
    $TestNodeMetrics_j = BaseMetrices - FirstRow \cdot M_z^j$;
    OD1 of $TestNodeMetrics_j$ using $MAD_j^z(t)$ of $\alpha_j^z(i,t) \leq (\varphi)$;
    OD2 of $TestNodeMetrics_j$ using $MAD_j^z(t)$ of $\alpha_j^z(i,t) \leq (\varphi - 1)$;
    OD3 of $TestNodeMetrics_j$ using $MAD_j^z(t)$ of $\alpha_j^z(i,t) \leq (\varphi - 2)$; where $\alpha_j^z(i,t)$: the Z-score of $x_j^z(i) \in$M$_j^z(t)$;
    $\alpha_j^z(i,t) = \frac{\varrho(x_j^z(i) - \bar{x}_j^z(t))}{MAD_j^z(t)}$;
    and OD observe $\alpha_j^z(i,t) \leq \varphi$ where $\varphi = 3.5$;
    if $all(OD1 == 0) == 1$ **then**
      $T1, T2, T3, T4 = 0.70$;
      Remove the first Row of $BaseMetrices$;
      Add $TestNodeMetrics_j$ at the end of $BaseMetrices$;
    else if $all(OD2 == 0) == 1$ **then**
      $T1, T2, T3, T4 = 0.85$;
      Remove the first Row of $BaseMetrices$;
      Add $TestNodeMetrics_j$ at the end of $BaseMetrices$;
    else if $all(OD3 == 0) == 1$ **then**
      $T1, T2, T3, T4 = 1.0$;
      Remove the first Row of $BaseMetrices$;
      Add $TestNodeMetrics_j$ at the end of $BaseMetrices$;
    else
      $T1, T2, T3, T4 = 0$;
    end
    Γ_j = $[T1, T2, T3, T4]$;
  end
  return Γ_j
end

TABLE IV
AA CLASSES, TRUST LEVELS, AND AUTHORIZATIONS

Classes	Trust Levels	Authorizations
Class-4	$\Gamma = 1.00$	E + W + R
Class-3	$\Gamma = 0.85$	W+R
Class-2	$\Gamma = 0.70$	R
Class-1	$\Gamma = 0.00$	D

D. Trust Aware AA with Federated Learning Model

The previous section discussed various features that contribute to building trust, including physical layer attributes such as RSSI, LQI, device temperature, battery level, and other Zolertia Z1 motes metrics. The trust levels acquired from the trust establishment algorithms for the AA model are illustrated in Figure-3, with values of 0.00, 0.75, 0.85, and 1.00. The threshold chosen for each level is arbitrary and can be tuned depending on the level of security required. For example, for high-security IoT applications, the trust level required can raised to up to 0.95. This essentially means that only highly trusted nodes can perform read, write, and execute operations.

Accuracy, precision, sensitivity, and selectivity are used to assess the performance of the three ANN approaches of Levenberg-Marquardt (LM), Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG) during the training phase.

A federated learning model is built by combining the three aforementioned models. The purpose of this combination is to assess the effectiveness of the trained model and determine if it meets the desired performance standards. This approach helps to ensure that the final model provides accurate and reliable outcomes that can be used to make informed decisions. Upon completion of the training process, a federated model undergoes a testing phase to assess its efficacy. If the model achieves a satisfactory level of performance, it is then stored in the centralized server and distributed to the gateway nodes. Alternatively, if it does not meet the required performance in terms of defined metrics, the model is subjected to further training rounds.

Each client-side model of FL gathers information and determines the best weights to be sent to the server-side algorithm. Three methods were used simultaneously because there are four groups to choose from in the dataset (Class-1, Class-2, Class-3, and Class-4) based on the trust level 0.00, 0.70, 0.85, and 1.00, respectively, as mentioned in the Table-IV. The FL was used to make the AA prediction once the weights from the training model were delivered to the server.

This research considers the AA using a three-layer feed-forward neural network for all nodes comprising input, hidden, and output layers. The reason for a lower number of layers is due to the fact that there are not too many features involved (only eight features) and, therefore, there is a lower requirement for generalization. If the required feature increases significantly, as in the case of image classification, the number of layers required increases significantly.

The elements in each layer are represented by integers *t, f,* and *k*, while the input characteristics are written as $[s1, s2, s3, ...s_n]$ while $b1$ and $b2$ represent the bias introduced

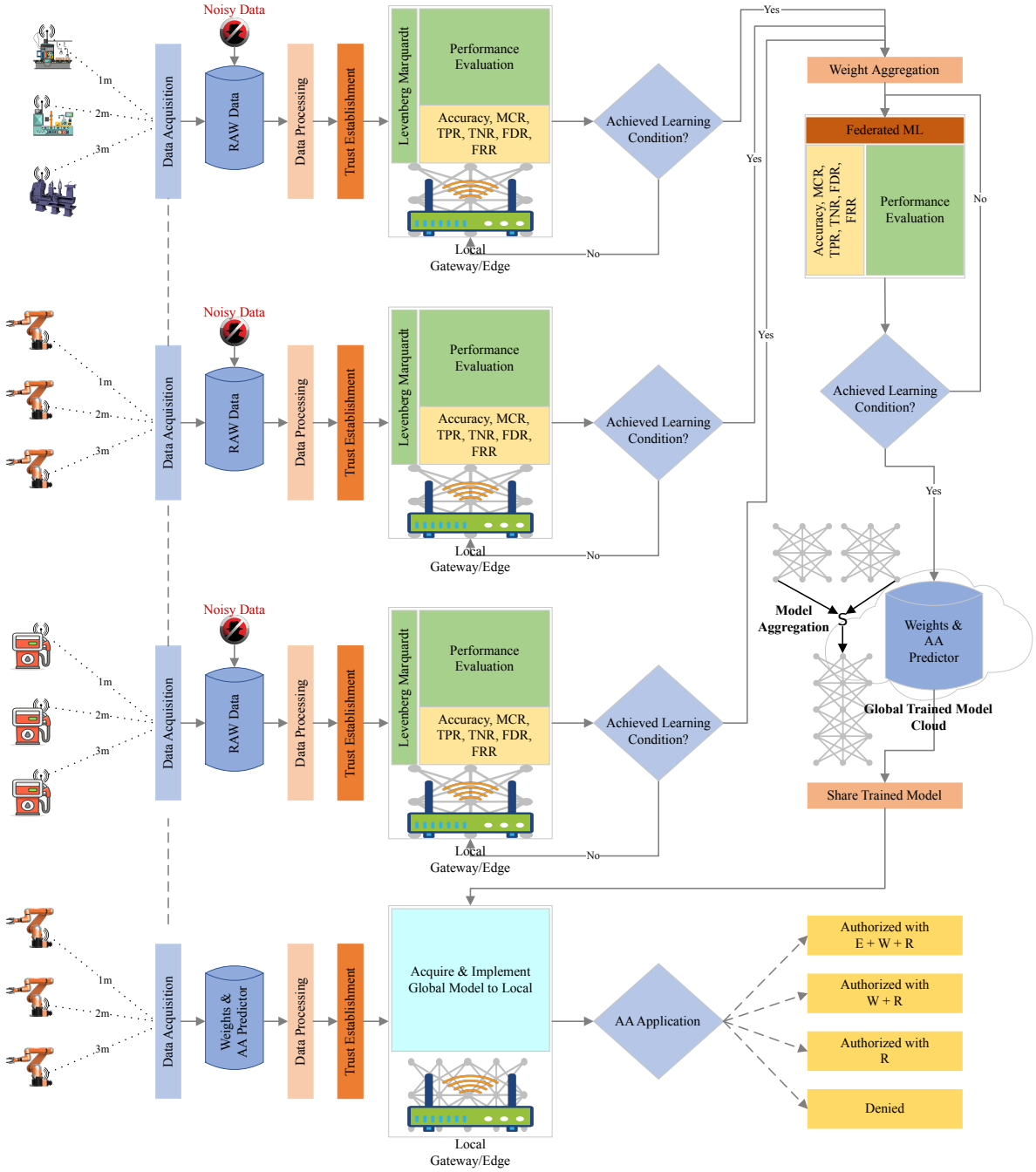

Fig. 3. Trust Aware AA with FL.

at each layer, respectively. The weights connecting the input and hidden layers are displayed as $q_{f,n}$, while those relating the hidden and output layers are shown as $p_{f,n}$. The input, hidden, and output layer dimensions n, k, and g represent each layer's total number of elements, respectively. Equation-(1) [40] may be used to determine the output of each neuron in the f_{th} hidden layer, where w_f^{cli} is the output of the neuron serving client i (cl_i).

$$w_f^{cli} = \frac{1}{1 + e^{-(b_1 + \sum_{t=1}^{k}(p_{t,f}^{cli} \times S_t))}}; \ where \ 1 \leq f \leq k \quad (1)$$

Similarly, in Equation-(2) [41], x_{cli}^n refers to the nth neuron's output at the corresponding output layer.

$$x_n^{cli} = \frac{1}{1 + e^{-(b_2 + \sum_{f=1}^{k}(q_{f,n}^{cli} \times w_f^{cli}))}}; \ where \ 1 \leq n \leq g \quad (2)$$

$$F^{cli} = \frac{1}{2} \sum_{n} \left(\beta_n^{cli} - x_n^{cli}\right)^2 \quad (3)$$

where F^{cli} denotes the i^{th} client error, β_{cli}^n for the n^{th} client's expected output, and x_{cli}^n is the n^{th} client's expected and projected outputs in equation-(3) [42].

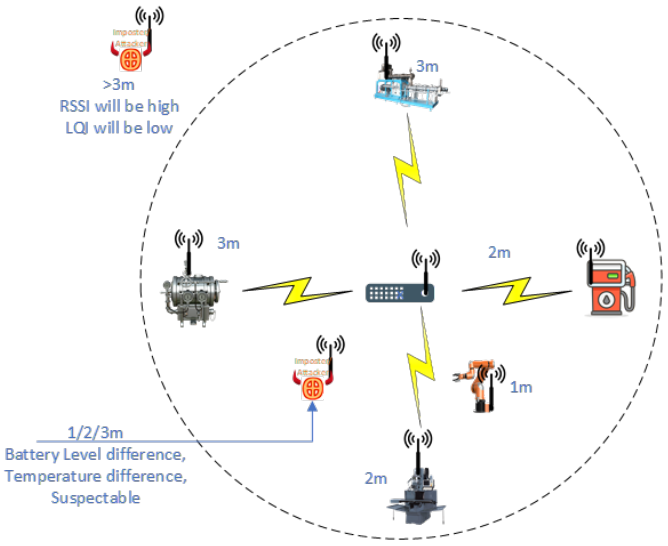

Fig. 4. Attack scenario of proposed Trust-Aware AA.

As mentioned in Equation-(4) [40,41], the output layer's dynamic weight is

$$\Delta P \propto -\frac{\partial F^{cli}}{\partial P^{cli}} \tag{4}$$

$$\Delta p_{f,n}^{cli} \propto -\frac{\partial F^{cli}}{\partial p_{t,f}^{cli}} \tag{5}$$

$$\Delta Q \propto -\frac{\partial F^{cli}}{\partial Q^{cli}} \tag{6}$$

$$\Delta q_{f,n}^{cli} \propto -\frac{\partial F^{cli}}{\partial q_{f,n}^{cli}} \tag{7}$$

Using the chain rule technique, we can rewrite the above equation as

$$\Delta q_{f,n}^{cli} = -\zeta \frac{\partial F^{cli}}{\partial x_n^{cli}} \times \frac{\partial x_n^{cli}}{\partial q_{f,n}^{cli}} \tag{8}$$

in which ζ is a constant. It is proven in Equation-(7) that by exchanging the numbers in Equation-(6), we may get the weight value after the change.

$$\Delta q_{f,n}^{cli} = \zeta \left(\beta_n^{cli} - x_n^{cli}\right) \times x_n^{cli} \left(1 - x_n^{cli}\right) \times w_f^{cli} \tag{9}$$

$$\Delta q_{f,n}^{cli} = \zeta \lambda_n^{cli} w_f^{cli} \tag{10}$$

where

$$\Delta \lambda_{f,n}^{cli} = \left(\beta_n^{cli} - x_n^{cli}\right) \times x_n^{cli} \left(1 - x_n^{cli}\right) \tag{11}$$

The chain rule updates the weights connecting the input and hidden layers.

$$\Delta p_{t,f}^{cli} \propto -\left[\sum_n \frac{\partial F^{cli}}{\partial x_n^{cli}} \times \frac{\partial x_n^{cli}}{\partial w_f^{cli}}\right] \times \frac{\partial w_f^{cli}}{\partial p_{t,f}^{cli}} \tag{12}$$

$$\Delta p_{t,f}^{cli} = -\zeta \left[\sum_n \frac{\partial F^{cli}}{\partial x_n^{cli}} \times \frac{\partial x_n^{cli}}{\partial w_f^{cli}}\right] \times \frac{\partial w_f^{cli}}{\partial p_{t,f}^{cli}} \tag{13}$$

$$\Delta p_{t,f}^{cli} = \zeta \sum_n \left(\beta_n^{cli} - x_n^{cli}\right) \times x_n^{cli} \left(1 - x_n^{cli}\right) \times \left(q_{f,n}^{cli}\right) \times w_f^{cli} \left(1 - w_f^{cli}\right) \times s_t \tag{14}$$

$$\Delta p_{t,f}^{cli} = \zeta \left[\sum_n \lambda_n^{cli} \times q_{f,n}^{cli}\right] \times w_f^{cli} \left(1 - w_f^{cli}\right) \times s_t \tag{15}$$

After some simplification, the equation may be written as:

$$\Delta p_{t,f}^{cli} = \zeta \alpha_f^{cli} \times s_t \tag{16}$$

where

$$\alpha_f^{cli} = \left[\sum_n \lambda_n^{cli} \times q_{f,n}^{cli}\right] \times w_f^{cli} \left(1 - w_f^{cli}\right) \tag{17}$$

$$q_{f,n}^{cli}(d+1) = q_{f,n}^{cli}(d) + \gamma \Delta q_{f,n}^{cli} \tag{18}$$

Equation-(18) [41,42] is used to fine-tune the weights between the output and hidden layers. Equation-(19) [40,42] is used to adjust the weights connecting the input and hidden layers.

$$p_{t,f}^{cli}(d+1) = p_{t,f}^{cli}(d) + \gamma \Delta p_{t,f}^{cli} \tag{19}$$

E. Training proposed FL Model

This section describes the steps involved in training the FL model for trust-aware AA. The training steps are as follows:

*1) **Local Machine Learning Algorithm**:* The pseudo-code for the training locally ML method using a feed-forward neural network at the edge/gateway is shown in Algorithm-(2). At this stage, we train the local edge/gateway node after establishing trust in the acquired and pre-processed data. This algorithm will return the optimum weights to the FL server as described in the next section-III-E2.

*2) **Weights Transfer**:* A federated server receives these weights and uses them accordingly. It is possible to encrypt these weights before sending them, making the system more secure. The work does not include encrypted weights, but optional encryption can be implemented if needed.

*3) **Federated Machine Learning Server**:* Each client sends the federated server its optimal weight $(P_{InpHid}^{cli}, Q_{HidOut}^{cli})$. In our scenario, we use one of the following ANN training methods for each client: Choose from (1) Levenberg-Marquardt (LM), (2) Bayesian Regularization (BR), or (3) Scaled Conjugate Gradient (SCG). Equations-(20-22) provide optimal weights for the LM, BR, and SCG algorithms.

$$P_{InpHid}^{cl_1}(LM) = \left[\begin{pmatrix} p_{11}^1 & p_{12}^1 & \cdots & p_{1c_n}^1 \\ p_{21}^1 & p_{22}^1 & \cdots & p_{2c_n}^1 \\ \vdots & \vdots & \vdots & \vdots \\ p_{r_m1}^1 & p_{r_m2}^1 & \cdots & p_{r_{mcn}}^1 \end{pmatrix}\right]_{d1 \times d2} \tag{20}$$

$$P_{InpHid}^{cl_2}(BR) = \left[\begin{pmatrix} p_{11}^2 & p_{12}^2 & \cdots & p_{1c_n}^2 \\ p_{21}^2 & p_{22}^2 & \cdots & p_{2c_n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ p_{r_m1}^2 & p_{r_m2}^2 & \cdots & p_{r_{mcn}}^2 \end{pmatrix}\right]_{d3 \times d4} \tag{21}$$

Algorithm 2: Proposed ML Algorithm for training locally at the Edge/Gateway node

Initialize Local Data splitting to small groups B_s;
Initialize both layers i.e., input and hidden layer weights $(P^{cl_i}_{InpHid}, Q^{cl_i}_{InpHid}), F^{cl_i} = 0$ and number of epochs d=0;
foreach *small groups* B_s **do**
  Stage-1: Apply the FeedForward Network to
  a. Calculating w_f^{cli} from Equation-(1);
  b. Calculating estimated output (x_n^{cli}) from Equation-(2);
  Stage-2: Calculate the Error values (F^{cli}) using Equation-(3);
  Stage-3: Weights updating phase
  a. Calculating $\Delta q_{f,n}^{cli}$ from Equation-(9);
  b. Calculating $\Delta p_{f,t}^{cli}$ from Equation-(16);
  c. Updating the weights between hidden and output layers $q_{f,n}^{cli}(d+1)$ from Equation-(18);
  d. Updating the weights between input and hidden layers $p_{f,t}^{cli}(d+1)$ from Equation-(19);
  if *achieved learning condition* **then**
    go to stage 4;
  else
    repeat;
  end
  Stage-4: Return optimum weights $(P^{cli}_{InpHid}, Q^{cli}_{HidOut})$ to the Federated Learning Server;
end

$$P^{cl_3}_{InpHid}(SCG) = \begin{bmatrix} \begin{pmatrix} p^3_{11} & p^3_{12} & \cdots & p^3_{1c_n} \\ p^3_{21} & p^3_{22} & \cdots & p^3_{2c_n} \\ \vdots & \vdots & \vdots & \vdots \\ p^3_{r_m1} & p^3_{r_m2} & \cdots & p^3_{r_mcn} \end{pmatrix} \end{bmatrix}_{d5\times d6} \tag{22}$$

Equation-(23) can be used to express the best weights for the federated server's input layer to the hidden layer, where $P^n_{InpHid}(FS)$ is the sum of the weights of all locally trained clients.

$$P^n_{InpHid}(FS) = P^{cl_1}_{InpHid}(LM) + P^{cl_2}_{InpHid}(BR) + P^{cl_3}_{InpHid}(SCG) \tag{23}$$

With the same number of rows and columns, the matrices $P_{InpHid-LM}$, $P_{InpHid-BR}$, and $P_{InpHid-SCG}$ can be combined. Using equation-(24), we can derive the centralized server, also called the global model.

$$P_{InpHid-FS} = \mu f()(P^{cl_1}_{InpHid}(LM), P^{cl_2}_{InpHid}(BR), P^{cl_3}_{InpHid}(SCG)) \tag{24}$$

$P_{InpHid-FS}$ represents the optimal input-to-hidden federated weights in equation-(24). Depending on the performance, different scaling factors are assigned to the locally trained nodes.

*4) **Optimal hidden output Layers Weights**:* Similar to the optimal weights from the input layer to the hidden layer, equations can establish the optimal weights from the hidden layer to the output layer-(25 - 27).

$$Q^{cl_1}_{HidOut}(LM) = \begin{bmatrix} \begin{pmatrix} q^1_{11} & q^1_{12} & \cdots & q^1_{1c_n} \\ q^1_{21} & q^1_{22} & \cdots & q^1_{2c_n} \\ \vdots & \vdots & \vdots & \vdots \\ q^1_{r_m1} & q^1_{r_m2} & \cdots & q^1_{r_mcn} \end{pmatrix} \end{bmatrix}_{d7\times d8} \tag{25}$$

$$Q^{cl_2}_{HidOut}(BR) = \begin{bmatrix} \begin{pmatrix} q^2_{11} & q^2_{12} & \cdots & q^2_{1c_n} \\ q^2_{21} & q^2_{22} & \cdots & q^2_{2c_n} \\ \vdots & \vdots & \vdots & \vdots \\ q^2_{r_m1} & q^2_{r_m2} & \cdots & q^2_{r_mcn} \end{pmatrix} \end{bmatrix}_{d9\times d10} \tag{26}$$

$$Q^{cl_3}_{HidOut}(SCG) = \begin{bmatrix} \begin{pmatrix} q^3_{11} & q^3_{12} & \cdots & q^3_{1c_n} \\ q^3_{21} & q^3_{22} & \cdots & q^3_{2c_n} \\ \vdots & \vdots & \vdots & \vdots \\ q^3_{r_m1} & q^3_{r_m2} & \cdots & q^3_{r_mcn} \end{pmatrix} \end{bmatrix}_{d11\times d12} \tag{27}$$

Equation-(27) can be used to determine federated weights, although the same dimensional inconsistency problem can be encountered. We will follow the same process for individuals to standardize the size of client nodes' weight arrays.

$$Q^n_{HidOut}(FS) = Q^{cl_1}_{HidOut}(LM) + Q^{cl_2}_{HidOut}(BR) + Q^{cl_3}_{HidOut}(SCG) \tag{28}$$

$$Q_{HidOut-FS} = \mu f()(Q^{cl_1}_{HidOut}(LM), Q^{cl_2}_{HidOut}(BR), Q^{cl_3}_{HidOut}(SCG)) \tag{29}$$

In particular, $Q_{HidOut-FS}$ in Equation-(29) refers to the federated optimal weights from the hidden layer to the output layer. Various scaling factors are applied to the locally trained edge/gateway client, each of which is determined by the performance of the edge node.

*5) **Federated Learning Algorithm**:* The pseudo-code for the proposed FL algorithm is shown in Algorithm 3 for trust-aware AA. This algorithm is run on the centralized server and updates itself to share the updated model with the existing or new edge/gateway node.

*6) **Gateway/Edge Node**:* The global model weights are transmitted to a local network or edge devices to enable them to detect storage activity using the global model. Subsequently, the proposed trust-aware algorithm imports the stored data to the cloud to predict the TM AA in the validation phase. The proposed FL model classifies AA into four categories based on AA Classes: Class-1 represents Access Denied, Class-2 represents Read Only, Class-3 represents Write + Read Only access levels, and Class-4 represents Execute + Write + Read Only access levels. The proposed trust model enables the determination of the optimal access level for the IoT devices.

TABLE V
SIMULATION AND TRAINING PARAMETERS

Parameters	Values
Area of deployed Sensor Area	1-3 meters
Total Dataset Instances	347,200 (100%)
Randomized Training Instances	243,040 (70%)
Randomized Testing Instances	104,160 (30%)
Number of Hidden Layers Neurons	32
Learning Rate, α	0.01
Momentum	0.8
Maximum Number of Epochs	1000
Activation Function	Sigmoid
Maximum Number of validation Check Fail	6

IV. RESULTS AND DISCUSSIONS

To evaluate the effectiveness of the proposed method, we examined three distinct setups in which sensor nodes attempt to connect with the Edge/Gateway at varying distances. According to the findings, the LM, BR, and SCG algorithms perform better than other machine learning (ML) algorithms with respect to mean squared error (MSE) over epochs and regression analysis. According to the Trust-Aware AA with FL model depicted in Figure 3, we analyzed nodes situated at distances of $1m$, $2m$, and $3m$. These nodes sought authentication and authorization from the Edge/Gateway node to carry out a designated task. In the setup as shown in Figure-3, Setup 1 (Local Edge-1) employed the LM, Setup 2 (Local Edge-2) utilized the BR, and Setup 3 (Local Edge-3) used the SCG. The MSE over Epoch(s) of all the Local ML models is depicted in Figure-5 with the simulation parameters mentioned in Table-V. The optimal weights obtained from LM, BR, and SCG are combined in the Federated Learning server as discussed in the methodology section-III. On the other hand, the trained Federated Learning model will distribute the trained network to the edge devices from the Federated server as per algorithm-3 for AA and to detect attacks.

A comprehensive examination was conducted on a sample size of $104,160$ instances, wherein the sensor nodes attempted to connect with the Edge/Gateway at a distance of $1m, 2m$, and $3m$ for AA. Similarly, for the same sample size, it is considered that the spoofer nodes located at $2m$ and $3m$ distances or different antenna orientations try to spoof, i.e., by changing actual MAC as 1m distance nodes MAC to take control of the Edge/Gateway and vice versa.

The data was randomized every time, and the test was conducted using MATLAB. The proposed trust-aware AA model's outcome is illustrated in figure-6 in the form of a confusion matrix, and the effectiveness is assessed by applying various statistical measures to individual classes. The overall performance of the model is determined by the statistical measures, including accuracy, misclassification rate (MCR), Recall, precision, specificity, negative predictive value (NPV), false positive rate (FPR), false rejection rate (FRR), false discovery rate (FDR), $F_{0.5}$ score, and F_1 score in the figure-7, figure-8 and figure-9 and Table-VI provide the summary of the performances of the AA and Attack detection using the following equations to evaluate the effectiveness of the proposed model.

Algorithm 3: Proposed FL Learning Algorithm for trust-aware AA

Initialize weights $(P_{InpHid-FS}, Q_{HidOut-FS})$
foreach *cycle* **do**
  foreach *client* **do**
    $\left[P^{cl_i}_{InpHid}, Q^{cl_i}_{HidOut}\right] = Client\left(d, P^{cl_i}_{InpHid}, Q^{cl_i}_{HidOut}\right)$
  end
  At this Stage: Calculating $Q_{HidOut-FS}$ from Equation-(29)
  At this Stage: Calculating $P_{InpHid-FS}$ from Equation-(24)
  At this Stage: for prediction of unknown data samples
  foreach *number of Samples, I* **do**
    Calculate $w^{FS}_f = \frac{1}{1+e^{-(b_1+\sum_{t=1}^{k}(p^{cli}_{f,t}\times S_t))}};\ where\ 1 \le f \le k;$
    Calculate $x^{FS}_n = \frac{1}{1+e^{-(b_2+\sum_{f=1}^{k}(q^{cli}_{f,n}\times w^{cli}_f))}};\ where\ 1 \le n \le g;$
    Calculate error $F^{FS} = \frac{1}{2} \times \sum_{n=1}^{g}\left(\beta^{FS}_n - x^{FS}_n\right)^2$
  end
end

$$Accuracy = \frac{(TN+TP)}{\{(FN+FP)+(TN+TP)\}} \tag{30}$$

$$MCR = \frac{(FN+FP)}{\{(FN+FP)+(TN+TP)\}} \tag{31}$$

$$TPR = \frac{(TP)}{\{(FN+TP)\}} \tag{32}$$

$$Specificity = \frac{(TN)}{\{(FP+TN)\}} \tag{33}$$

$$Precision = \frac{(TP)}{(TP+FP)} \tag{34}$$

$$NPV = \frac{(TN)}{(FN+TN)} \tag{35}$$

$$FDR = \frac{(FP)}{(FP+TP)} \tag{36}$$

$$FAR = \frac{(FP)}{(FP+TN)} \tag{37}$$

$$FRR = \frac{(FN)}{(FN+TP)} \tag{38}$$

$$F_1\ Score = 2 \times Precision \times \frac{Recall}{Precision\ +\ Recall} \tag{39}$$

$$F_{0.5}\ Score = 1.25\ \times\ Precision \times \frac{Recall}{0.25\ \times\ Precision\ +\ Recall} \tag{40}$$

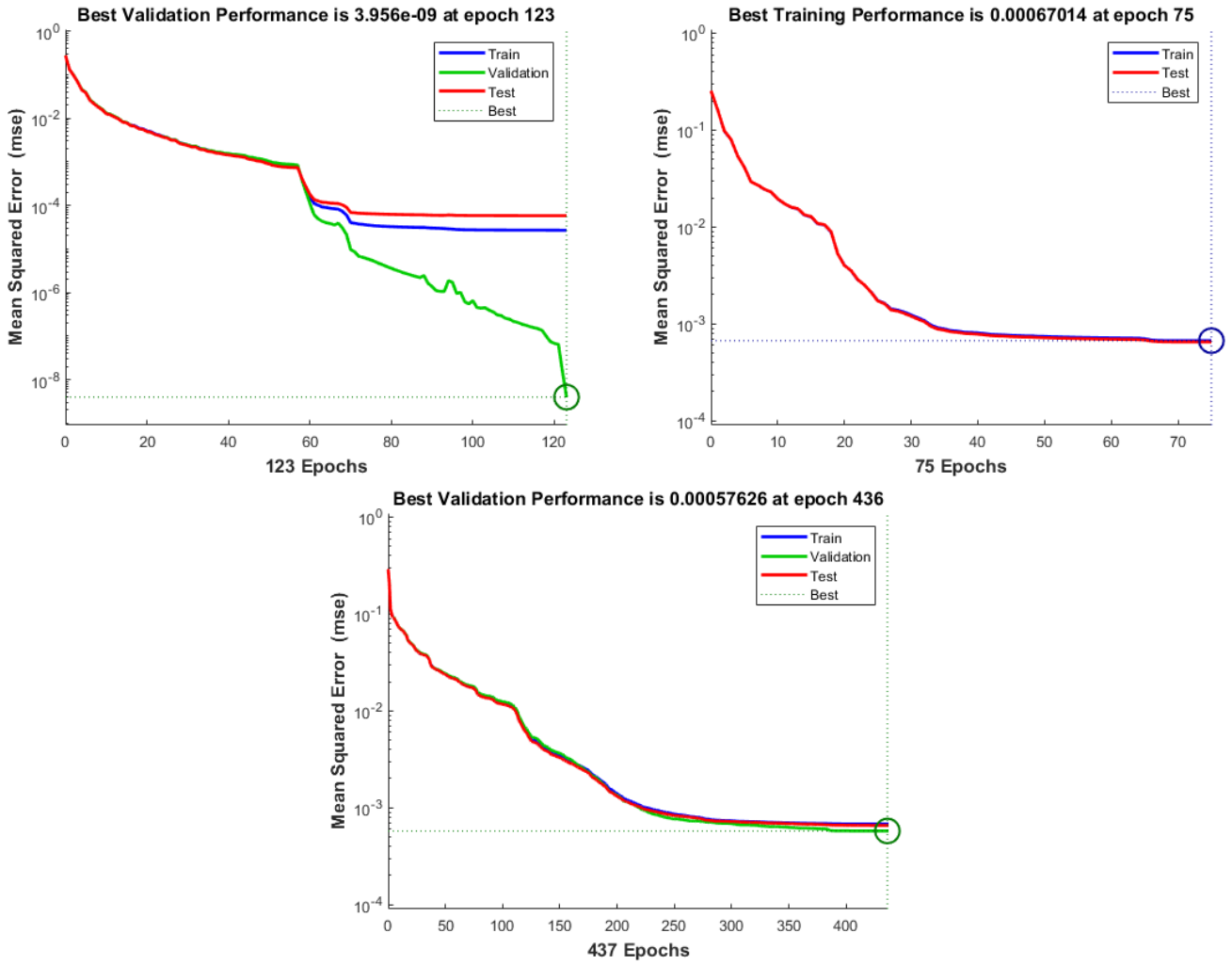

Fig. 5. Local Model performance - MSE vs Epochs for LM, BR, SCG (from Left to Right)

$$MCC = \frac{\{(TN \times TP) - (FN \times FP)\}}{\sqrt{\{(FN + TN) \times (FN + TP)\}}} \times \frac{1}{\sqrt{\{(FP + TN) \times (FP + TP)\}}} \quad (41)$$

Note that in the equations above, True Positives (TP) are the number of positive instances correctly classified as positive, True Negatives (TN) are the number of negative instances correctly classified as negative, False Positives (FP) are the number of negative instances incorrectly classified as positive, and False Negatives (FN) are the number of positive instances incorrectly classified as negative.

The prediction outcomes for the classification tasks are depicted in Figure-6. The figure summarizes the correct and wrong predictions, presenting the count values for each class. The accuracy in identifying Class-1 and Class-4 is close to 100%. However, one Class-2 sample was wrongly classified as Class-1, two as Class-3, and two Class-3 samples were improperly classified as Class-1 and one as Class-2. As observed, the number of incorrect predictions decreased in the proposed trust-aware AA. Furthermore, the overall performance is summarized in Table-VI.

TABLE VI
AA AND ATTACK DETECTION STATISTICAL PERFORMANCE EVALUATION

Metrics	AA - Overall (%)	Attack - Overall (%)
Accuracy	99.7925	99.99602
MCR	0.207476068	0.003982
TPR	99.7925	99.99602
Specificity	99.925	100
Precision	99.745	100
NPV	99.92127187	99.99873
F1-Score	99.77876751	99.99801
F0.5-Score	99.77140937	99.9992
MCC	99.70236443	99.99738
FAR	0.075545711	0
FDR	0.233123411	0
FRR	0.207476068	0.003982

We compared the accuracy and misclassification rate (MCR) of our IoT trust-aware AA for IoT with various ML techniques as depicted in Figure-10, Our proposed scheme results in better accuracy, achieving a rate of 99.99602%. However, the overall MCR is nominal, standing at only 0.003982%. The

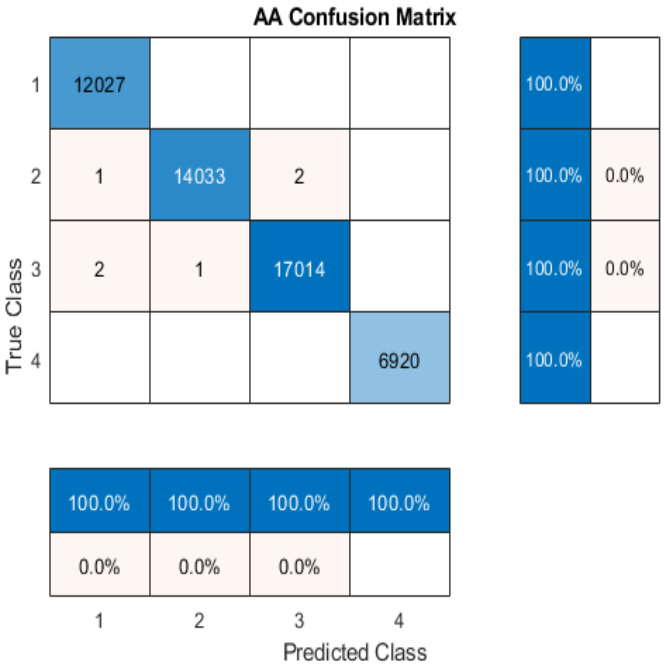

Fig. 6. FL Confusion Matrix for trust-aware AA.

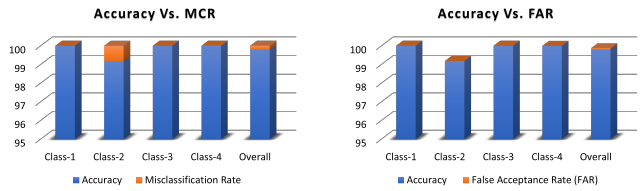

Fig. 7. Trust-Aware AA Performance Evaluation. Accuracy Vs. Misclassification Rate, Accuracy Vs. False Acceptance Rate(from Left to Right)

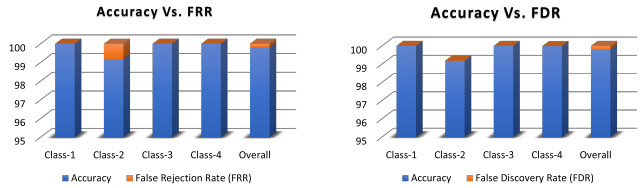

Fig. 8. Trust-Aware AA Performance Evaluation. Accuracy Vs. False Rejection Rate, Accuracy Vs. False Discovery Rate (from Left to Right)

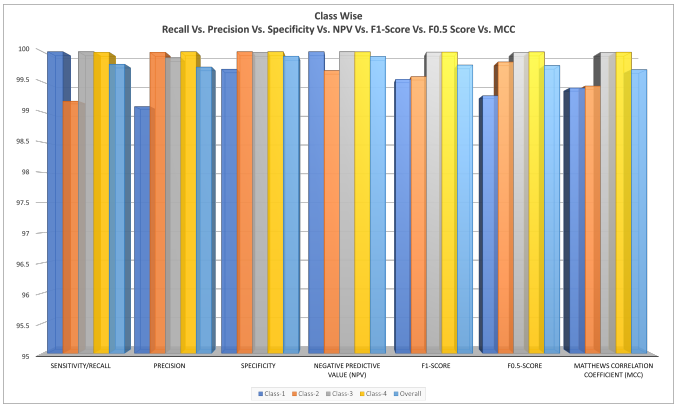

Fig. 9. Trust-Aware AA Performance Evaluation for Other matrices.

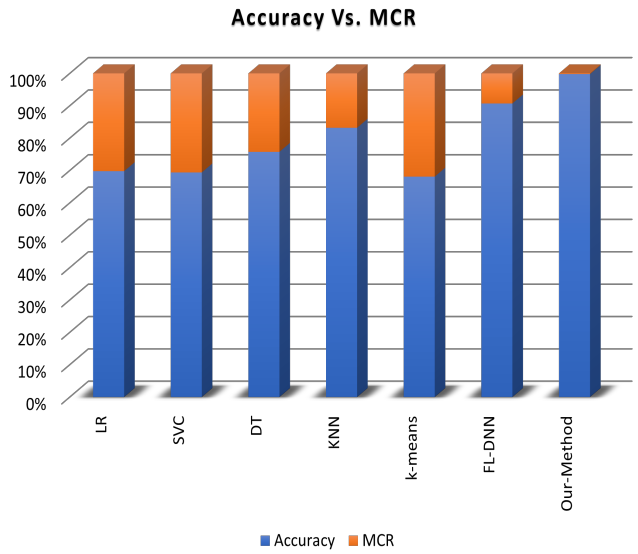

Fig. 10. Accuracy Vs. MCR

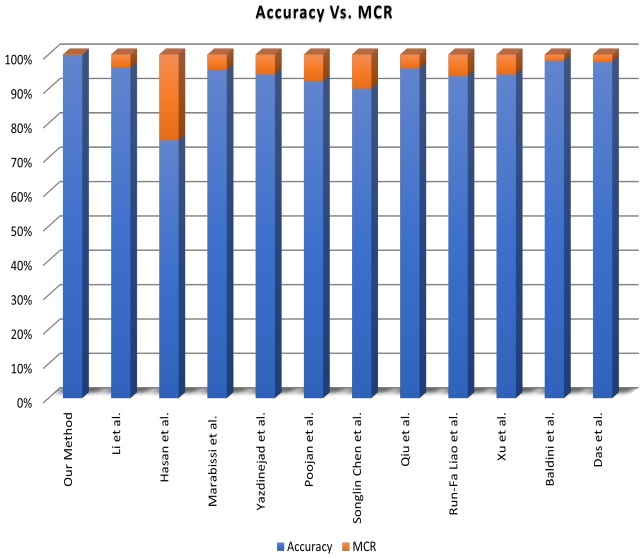

Fig. 11. Accuracy Vs. MCR

model accuracy exhibits impressive performance, attaining almost perfect classification outcomes while minimizing false positives and negatives. As anticipated, our model performs better than competing machine learning models. Moreover, it can be observed from Table-VI that the overall performance to detect and mitigate spoofing attacks is quite high based on the performance metrics that we considered in our work.

Furthermore, we benchmarked the proposed method's accuracy and misclassification rate (MCR), as depicted in Figure-11, with the existing literature. In the conducted benchmarking study, an assessment was made of the effectiveness of our strategy in comparison to various recent methodologies within the respective field. The proposed method demonstrated a remarkable level of accuracy, reaching 99.79%, which surpasses the closest rival, Li et al. [37], by a substantial mar-

gin of 3.69% (96.10%) with our trust establishment model. Furthermore, regarding the MCR, our approach exhibited outstanding outcomes, exhibiting a minimal MCR of 0.21%, thus emphasizing its resilience. In the accuracy evaluation, our findings demonstrated a commendable precision rate of 100%, which was in line with the results reported by Yazdinejad et al. (95.8%) [12]. In addition, our approach showed a significantly lower False Rejection Rate (FRR) of 0.003982% compared to the study conducted by Hasan et al. (5.03%) [35]. This outcome highlights our proposed method's overall improved performance within the benchmarking context.

A. Discussion

At the core of security within the IoT landscape lies the imperative for robust authentication and authorization processes. These processes are essential in managing access to nodes and their corresponding data. However, the heterogeneity in communication topologies and protocols across IoT nodes, coupled with the absence of a standardized approach, underscores the need for trust management. The proposed architecture comprising Federated ML enhances the performance of IoT AA, showing improved performance with our trust management algorithm compared to alternative machine learning models with an Accuracy of 99.7925% in AA. Additionally, the suggested model can regularly update and enhance its performance through continuous learning when new nodes are added to the network. Moreover, the proposed architecture has exhibited resilience in potential vulnerabilities, with over 99.8% accuracy in identifying security breaches such as spoofing and impersonation.

The physical layer security solutions are viable for future communication networks, including IoT networks. These solutions can be incorporated into the communication stack with standard physical layer characteristics with significant changes to existing protocol stacks and adding an additional layer of security.

B. Limitation and Future Research Directions

Our proposed work is limited to indoor stationary nodes and edge devices that are fixed in location and antenna orientation, as described in the dataset section focusing only on the physical layer parameters. The movement of the nodes is not considered in this experiment. However, it is an important research direction in which mobile IoT nodes need to be considered for more practical applications. Moreover, our work does not consider the network-level performance parameters, i.e., packet dropping and throughput. Since the network-level characteristics have been widely considered in the literature, the focus of this work is to utilize the physical layer characteristics. However, it will be interesting to explore the implementation of the concept of security at the physical layer along with network security to secure the entire communication stack.

V. Conclusion

The primary objective of this study was to tackle the machine learning-specific challenges associated with the central training process, privacy preservation, accuracy, and heterogeneity concerns in the authentication and authorization of IoT systems. To achieve this objective, we presented a federated machine learning-based trust-aware scheme for the IoT Authentication and Authorization (AA) model. Trust management and federated machine learning are crucial components in enhancing the reliability and security of nodes in a network. The proposed Trust-aware AA scheme evaluates the Node Trust Level to identify malicious nodes and grant appropriate access to the nodes. The physical layer parameters, such as Media Access Control (MAC), Received Signal Strength Indicator (RSSI), Link Quality Indicator (LQI), Node Battery Level, and Node's internal Temperature, play a vital role in evaluating node reliability and credibility. Therefore, the proposed scheme is founded on the physical layer characteristics of nodes within IoT networks. The presented scheme leverages decentralized data provided by Zigbee Zolertia Z1 motes while incorporating our trust establishment algorithm and a privacy-preserving advantage through the federated ML approach to train data locally. The proposed architecture of trust-aware AA can enhance the performance of IoT security, exhibiting superior performance with our trust establishment algorithm compared to alternative machine learning models. The proposed method achieves an accuracy rate of 99.7925% in AA, accurately classifying and detecting unauthorized and authorized attempts.

References

[1] S. M. Jayadeva, A. Al Ayub Ahmed, R. Malik, A. A. Shaikh, M. N. E. Siddique, and M. Naved, *Roles of Cloud Computing and Internet of Things in Marketing Management: A Critical Review and Future Trends*. Springer Nature Singapore, 2023, vol. 290, no. January.

[2] A. P. Monali Suthar and D. S. Khara, "Survey of Secure IoT using Machine Learning Approach," *Interantional Journal of Scientific Research in Engineering and Management*, vol. 07, no. 04, 2023.

[3] T. Alam, "A Reliable Communication Framework and Its Use in Internet of Things (IoT)," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol. © 2018 IJSRCSEIT*, vol. 5, no. 10, pp. 450–456, 2018.

[4] G. D. Putra, V. Dedeoglu, S. S. Kanhere, and R. Jurdak, "Trust Management in Decentralized IoT Access Control System," *arXiv Prepr.*, p. arXiv:1912.10247, 2019. [Online]. Available: http://arxiv.org/abs/1912.10247

[5] D. Hardt, "The OAuth 2.0 Authorization Framework," 2012. [Online]. Available: https://tools.ietf.org/id/draft-ietf-oauth-v2-31.html

[6] A. K. Singh and D. Saxena, "A Cryptography and Machine Learning Based Authentication for Secure Data-Sharing in Federated Cloud Services Environment," *Journal of Applied Security Research*, vol. 17, no. 3, pp. 385–412, 2022. [Online]. Available: https://doi.org/10.1080/19361610.2020.1870404

[7] K. Mabodi, M. Yusefi, S. Zandiyan, L. Irankhah, and R. Fotohi, "Multi-level trust-based intelligence schema for securing of internet of things (IoT) against security threats using cryptographic authentication," *Journal of Supercomputing*, vol. 76, no. 9, pp. 7081–7106, 2020. [Online]. Available: https://doi.org/10.1007/s11227-019-03137-5

[8] R. Meng, X. Xu, H. Sun, H. Zhao, B. Wang, S. Han, and P. Zhang, "Multiuser Physical-Layer Authentication Based on Latent Perturbed Neural Networks for Industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 10, no. 1, pp. 637–652, 2023.

[9] R. F. Liao, H. Wen, J. Wu, F. Pan, A. Xu, Y. Jiang, F. Xie, and M. Cao, "Deep-Learning-Based Physical Layer Authentication for Industrial Wireless Sensor Networks," *Sensors (Basel, Switzerland)*, vol. 19, no. 11, pp. 1–17, 2019.

[10] S. Wang, N. Li, S. Xia, X. Tao, and H. Lu, "Collaborative Physical Layer Authentication in Internet of Things Based on Federated Learning," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, vol. 2021-Septe, pp. 714–719, 2021.

[11] H. Lin, K. Kaur, X. Wang, G. Kaddoum, J. Hu, and M. M. Hassan, "Privacy-Aware Access Control in IoT-enabled Healthcare: A Federated Deep Learning Approach," *IEEE Internet of Things Journal*, vol. 4662, no. c, pp. 1–10, 2021.

[12] A. Yazdinejad, R. M. Parizi, A. Dehghantanha, and H. Karimipour, "Federated learning for drone authentication," *Ad Hoc Networks*, vol. 120, no. December 2020, p. 102574, 2021. [Online]. Available: https://doi.org/10.1016/j.adhoc.2021.102574

[13] H. Xie, Z. Qin, X. Tao, and Z. Han, "Towards Intelligent Communications: Large Model Empowered Semantic Communications," *arXiv*, pp. 1–7, 2024. [Online]. Available: http://arxiv.org/abs/2402.13073

[14] M. A. Jan, F. Khan, M. Alam, and M. Usman, "A payload-based mutual authentication scheme for Internet of Things," *Futur. Gener. Comput. Syst.*, vol. 92, pp. 1028–1039, 2019. [Online]. Available: http://dx.doi.org/10.1016/j.future.2017.08.035

[15] Z. A. Alizai, N. F. Tareen, and I. Jadoon, "Improved IoT Device Authentication Scheme Using Device Capability and Digital Signatures," *ICAEM 2018 - 2018 Int. Conf. Appl. Eng. Math. Proc.*, pp. 115–119, 2018.

[16] H. Hamidi, "An approach to develop the smart health using Internet of Things and authentication based on biometric technology," *Futur. Gener. Comput. Syst.*, vol. 91, pp. 434–449, 2019. [Online]. Available: https://doi.org/10.1016/j.future.2018.09.024

[17] M. T. Hammi, B. Hammi, P. Bellot, and A. Serhrouchni, "Bubbles of Trust: A decentralized blockchain-based authentication system for IoT," *Comput. Secur.*, vol. 78, pp. 126–142, 2018. [Online]. Available: https://doi.org/10.1016/j.cose.2018.06.004

[18] L. Zhou, X. Li, K. H. Yeh, C. Su, and W. Chiu, "Lightweight IoT-based authentication scheme in cloud computing circumstance," *Futur. Gener. Comput. Syst.*, vol. 91, pp. 244–251, 2019. [Online]. Available: https://doi.org/10.1016/j.future.2018.08.038

[19] U. Jayasinghe, N. B. Truong, and G. M. Lee, "RpR : A Trust Computation Model for Social Internet of Things ," in *2016 Intl IEEE Conf. Ubiquitous Intell. Comput. Adv. Trust. Comput. Scalable Comput. Commun. Cloud Big Data Comput. Internet People, Smart World Congr.*, 2016, pp. 930–937.

[20] J. Li, Y. Bai, N. Zaman, and V. C. Leung, "A Decentralized Trustworthy Context and QoS-Aware Service Discovery Framework for the Internet of Things," *IEEE Access*, vol. 5, pp. 19 154–19 166, 2017.

[21] R. T. Frahat, M. M. Monowar, and S. M. Buhari, "Secure and Scalable Trust Management Model for IoT P2P Network," *2nd Int. Conf. Comput. Appl. Inf. Secur. ICCAIS 2019*, pp. 1–6, 2019.

[22] O. Ben Abderrahim, M. H. Elhdhili, and L. Saidane, "TMCoI-SIOT: A trust management system based on communities of interest for the social Internet of Things," *2017 13th Int. Wirel. Commun. Mob. Comput. Conf. IWCMC 2017*, pp. 747–752, 2017.

[23] U. Jayasinghe, A. Otebolaku, T. W. Um, and G. M. Lee, "Data centric trust evaluation and prediction framework for IOT," *Proc. 2017 ITU Kaleidosc. Acad. Conf. Challenges a Data-Driven Soc. ITU K 2017*, vol. 2018-Janua, no. March, pp. 1–7, 2017.

[24] H. Al-Hamadi and I. R. Chen, "Trust-Based Decision Making for Health IoT Systems," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1408–1419, 2017.

[25] J. I. Z. Chen, "Embedding the MRC and SC Schemes into Trust Management Algorithm Applied to IoT Security Protection," *Wirel. Pers. Commun.*, vol. 99, no. 1, pp. 461–477, 2018. [Online]. Available: https://doi.org/10.1007/s11277-017-5120-4

[26] W. Li, H. Song, and F. Zeng, "Policy-Based Secure and Trustworthy Sensing for Internet of Things in Smart Cities," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 716–723, 2018.

[27] J. Caminha, A. Perkusich, and M. Perkusich, "A Smart Trust Management Method to Detect On-Off Attacks in the Internet of Things," *Secur. Commun. Networks*, vol. 2018, 2018.

[28] S. Chinnaswamy and A. K, "Trust aggregation authentication protocol using machine learning for IoT wireless sensor networks," *Computers and Electrical Engineering*, vol. 91, no. December 2020, p. 107130, 2021. [Online]. Available: https://doi.org/10.1016/j.compeleceng.2021.107130

[29] R. Das, A. Gadre, S. Zhang, S. Kumar, and J. M. Moura, "A Deep Learning Approach to IoT Authentication," in *IEEE International Conference on Communications*, vol. 2018-May, 2018.

[30] G. Baldini, R. Giuliani, and F. Dimc, "Physical layer authentication of Internet of Things wireless devices using convolutional neural networks and recurrence plots," *Internet Technology Letters*, vol. 2, no. 2, p. e81, 2019.

[31] X. Qiu, J. Dai, and M. Hayes, "A Learning Approach for Physical Layer Authentication Using Adaptive Neural Network," *IEEE Access*, vol. 8, pp. 26 139–26 149, 2020.

[32] R. F. Liao, H. Wen, S. Chen, F. Xie, F. Pan, J. Tang, and H. Song, "Multiuser Physical Layer Authentication in Internet of Things with Data Augmentation," *IEEE Internet of Things Journal*, vol. 7, no. 3, pp. 2077–2088, 2020.

[33] S. Chen, Z. Pang, H. Wen, K. Yu, T. Zhang, and Y. Lu, "Automated Labeling and Learning for Physical Layer Authentication against Clone Node and Sybil Attacks in Industrial Wireless Edge Networks," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2041–2051, 2021.

[34] P. Oza and V. M. Patel, "Federated Learning-based Active Authentication on Mobile Devices," *Computer Vision and Pattern Recognition (cs.CV)*, vol. arXiv:2104, 2021.

[35] S. S. Ul Hasan, A. Ghani, I. U. Din, A. Almogren, and A. Altameem, "IoT devices authentication using artificial neural network," *Computers, Materials and Continua*, vol. 70, no. 2, pp. 3701–3716, 2022.

[36] D. Marabissi, L. Mucchi, and A. Stomaci, "IoT Nodes Authentication and ID Spoofing Detection Based on Joint Use of Physical Layer Security and Machine Learning," *Future Internet*, vol. 14, no. 2, pp. 1–21, 2022.

[37] Y. Li, Y. Wang, X. Liu, P. Zuo, H. Li, and H. Jiang, "Deep-Reinforcement-Learning-Based Wireless IoT Device Identification Using Channel State Information," *Symmetry*, vol. 15, no. 7, pp. 1–21, 2023.

[38] L. Xiao, Y. Li, G. Han, G. Liu, and W. Zhuang, "PHY-Layer Spoofing Detection with Reinforcement Learning in Wireless Networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 10 037–10 047, 2016.

[39] L. Xiao, Q. Yan, W. Lou, G. Chen, and Y. T. Hou, "Proximity-based security techniques for mobile users in wireless networks," *IEEE Trans. Inf. Forensics Secur.*, vol. 8, no. 12, pp. 2089–2100, 2013.

[40] L. Xiao, C. Xie, T. Chen, H. Dai, and H. V. Poor, "Mobile offloading game against smart attacks," *Proc. - IEEE INFOCOM*, vol. 2016-Septe, pp. 403–408, 2016.

[41] M. Hammad, A. M. Iliyasu, I. A. Elgendy, and A. A. El-Latif, "End-to-End Data Authentication Deep Learning Model for Securing IoT Configurations," *Human-centric Computing and Information Sciences*, vol. 12, no. 04, 2022.

[42] X. Li, K. Huang, S. Wang, and X. Xu, "A physical layer authentication mechanism for IoT devices," *China Communications*, vol. 19, no. 5, pp. 129–140, 2022.

[43] D. Marabissi, A. Stomaci, and L. Mucchi, "Adaptive Security in Mobile Wireless Networks: Machine Learning-Enhanced Continuous Physical Layer Authentication for Dynamic Environments," in *2024 IEEE International Workshop on Metrology for Industry 4.0 and IoT, MetroInd4.0 and IoT 2024 - Proceedings*. IEEE, 2024, pp. 310–315.

[44] K. I. Ahmed, M. Tahir, S. L. Lau, M. H. Habaebi, A. Ahad, and I. M. Pires, "Dataset for authentication and authorization using physical layer properties in indoor environment," *Data in Brief*, vol. 55, p. 110589, 2024. [Online]. Available: https://doi.org/10.1016/j.dib.2024.110589

[45] T. Crosby, B. Iglewicz, and D. C. Hoaglin, *How to Detect and Handle Outliers*. ASQC, 1994, vol. 36, no. 3.