Journal of Computer Science

Research Article

A Novel Approach Based on Federated Learning for the Identification of Man in the Middle Attacks in IOT Networks Using Blockchain

Sarumathi S, Juliet Johny, Minal Khandare, Lijimol K, Keerthi V and Krishnameena P

Department of Computer Science and Engineering, HKBK College of Engineering, Bengaluru, Karnataka, India

Article history
Received: 08-04-2025
Revised: 22-06-2025
Accepted: 25-09-2025

Corresponding Author:
Sarumathi S
Department of Computer Science and Engineering, HKBK College of Engineering, Bengaluru, Karnataka, India
Email: sarumathisakadevan@gmail.com

Abstract: Federated Learning (FL) enables cooperative model training across dispersed edge devices while protecting data privacy and providing localized insights without the need for centralized data aggregation. In the Internet of Things (IoT), federated learning enables cooperative model training among dispersed edge devices while protecting data privacy and providing localized insights without the need for centralized data aggregation. Nevertheless, Federated Learning's local model sharing approach makes it susceptible to Man-in-the-Middle (MITM) attacks. In FL, attackers have the ability to manipulate local models. As a result, a global model produced from the altered local models may be inaccurate. In this paper, we suggest a blockchain-based FL architecture to prevent intermediaries from readily altering the FL parameters throughout the transmission process. All of the clients' parameters are combined by a cloud server, which is acting as the federated parameter server. By integrating blockchain technology into the overall architecture, we link all cloud and edge servers. This paper uses a PoC algorithm based on the SHA-256 hashing function in order to validate the data. The results and comparison analysis show that the suggested framework has a low false-positive rate and a high accuracy in detecting MITM attacks in their early stages, with a detection rate ranging from 98 to 100%.

Keywords: Blockchain Technology, Intrusion Detection, MITM, Federated Learning, Proof of Accuracy, Introduction

Introduction

Machine Learning (ML) has impacted several industries and significantly changed people's lives. Machine learning models could be trained by utilizing the massive amounts of data produced every day by multiple end users, leading to improved services and an improvement in people's quality of life. Due to its expansion, the Internet of Things (IoT) is now a crucial part of many intelligent applications. It encompasses manufacturing, transportation, vital system infrastructure, healthcare, agriculture, etc. IoT devices use machine learning algorithms to collect vast amounts of data and function independently as part of an intelligent system. An important factor in training a machine learning algorithm system is the vast amount of IoT data (Rathee et al., 2021) Significant security dangers, such as network infiltration, data pollution, malware threats, and vulnerability exploitation, are introduced to IoT systems by the growth of diverse devices and network connections. Thus, it becomes crucial to put protective measures in place to protect edge servers, reduce the leakage of private data, and maintain production safety.

Federated Learning (FL)

Traditional machine learning techniques have difficulties with communication overhead and privacy issues since they necessitate moving data to a centralized server. With the introduction of Federated Learning, a local model is trained on the device itself using data that is acquired locally, distributing the model training process over multiple devices. The various trained model parameters are shared with a centralized device for global model training after several devices have completed the local training. For the following model training run, the global model is then transmitted back to every device involved in the FL process. But Communication efficiency declines and security issues arise as the number

of devices involved in the FL process rises. When a malicious node mimics the identity of an authenticated node, it renders all other nodes that communicate with it vulnerable to attack, and it will eventually reduce the global model's accuracy (Song et al., 2021) So FL is unreliable due to single points of failure, trust issues, as well as potential vulnerabilities to malicious attacks. It is necessary to have a system that functions openly and independently of a centralized aggregator, guaranteeing that calculations are carried out in a reliable way to prevent malicious updates.

Blockchain Technology

In blockchain technology, a block is composed of a header and a body. The header includes the time stamp, transaction information, and the hash of the previous block. Every transaction on the blockchain is digitally signed, and the hash is saved so that it may be retrieved later. This makes it possible to record the history of every transaction. Blockchain uses consensus, which is an algorithm that defines agreement that has to be agreed upon by all the nodes in the system. In a decentralized network, a consensus algorithm is an agreement that is used collectively to reach a conclusion when necessary. The popular consensus algorithms are PoW, PoS, and PoA. Following successful consensus validation of the transaction, the blocks are connected to form a chain.

Blockchain-Enabled Federated Learning

The FL system's unreliability can be eliminated by using blockchain nodes in place of the central aggregator to confirm model updates prior to aggregating the global model. Heterogeneous devices can securely authenticate shared electronic records without depending on a central, reliable third party through the use of blockchain technology. The data is stored in the blockchain as blocks.

Blockchain technology improves the security of the exchange of training data between FL clients and FL servers (Li et al., 2022). Data is stored using blockchain technology in a unique manner that makes it difficult for it to be altered, compromised, or misused (Fig. 1).

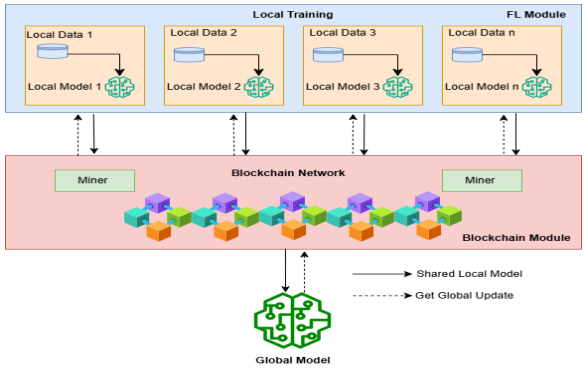

Fig. 1: Architecture of Blockchain-Enabled FL

In typical FL systems, a global model is kept on a single central server. Blockchain-based FL ensures that model updates are stored on multiple nodes instead of just one. Integrating blockchain technology into FL to improve reliability, efficiency, and security. No one node in a blockchain network has complete control over the network. Every node keeps a copy of the complete blockchain and takes part in the consensus process to decide when to create and add new blocks. Any attempt to change data in a block would require simultaneously changing all copies throughout the entire blockchain due to its distributed structure, making it impossible to modify data. Furthermore, the decentralization aspect increases the blockchain system's stability and resilience by removing the possibility of single points of failure.

FL clients upload their local updates to the blockchain-connected miners, where miners can be servers, personal computers, or cloud-based nodes. Every FL participant and data holder has a direct line of communication with the miner, guaranteeing continuous connection. The local model updates from participating FL devices or participants must be sent to the miners. Additionally, a block is uploaded to the blockchain network after aggregation is completed using the consensus procedure. The blockchain network receives validated new blocks. Until it achieves the necessary learning rate, the FL model process keeps operating. The global model can then be downloaded to be utilized by FL clients or other participants upon request. Lastly, miners can download the global model, and FL participants can obtain the model from them.

Intrusion Detection

In order to guarantee the security and privacy of data transfer between local devices and the FL Server, intrusion detection systems are essential. A hardware or software system that monitors a network for unauthorized activity or policy violations is called an Intrusion Detection System (IDS). Due to their success in achieving high classification accuracy, machine learning and deep learning with intrusion detection systems have become increasingly popular. However, the requirement to store and communicate data to a centralized server may compromise privacy and security. Federated Learning (FL), on the other hand, is a decentralized learning method that protects privacy. Instead of transferring data, FL trains models locally and sends the parameters to a centralized server. The enormous computational cost and privacy issues of keeping massive volumes of data on a single server along the cloud-edge continuum have made the use of centralized machine learning techniques for IDS impractical. One appealing development in tackling the problem is the combination of blockchain technology and federated learning. So, a new blockchain-based federated learning framework is proposed for intrusion detection. Blockchain-driven FL maintains a record of

every block that has been linked into a chain.

Man-in-the-Middle (MITM) Attacks

Ensuring the availability, confidentiality, and integrity of data is crucial in today's digital environment. One popular kind of cybersecurity assault that gives attackers the ability to listen to the conversation of two targets is called a Man-In-The-Middle Attack (MITM). The concept of "man-in-the-middle" refers to the fact that the attack occurs between two hosts that are actually interacting, enabling the attacker to "listen" to a discussion that they shouldn't normally be able to. When sending emails, texts, or video calls, none of the participants realize that an attacker has entered the conversation and is taking their information.

Figure 2 illustrates a typical Man-in-the-Middle (MITM) attack scenario in a Federated Learning (FL) environment, where an adversary intercepts the communication between clients and the global FL model. This attack poses significant threats, as the attacker can tamper with the exchanged model parameters, inject malicious gradients, or extract sensitive information during transmission. However, this vulnerability is effectively addressed in the proposed blockchain-enabled FL framework. By incorporating blockchain technology, the framework ensures secure, transparent, and tamper-proof communication between clients and the server. Smart contracts are used to authenticate clients and validate updates before aggregation, while the decentralized ledger maintains an immutable record of transactions, thereby eliminating the need for trust in a centralized authority. Additionally, cryptographic techniques such as end-to-end encryption and digital signatures safeguard the integrity and confidentiality of the data being transmitted.

Network-wide interactions between FL clients and servers reveal exploitable weaknesses to man-in-the-middle attackers. By impersonating the client, an attacker can send purposefully altered model parameters to the server (Park et al., 2024). An attacker can spread injected data to nodes across the network after spreading and aggregating parameters.

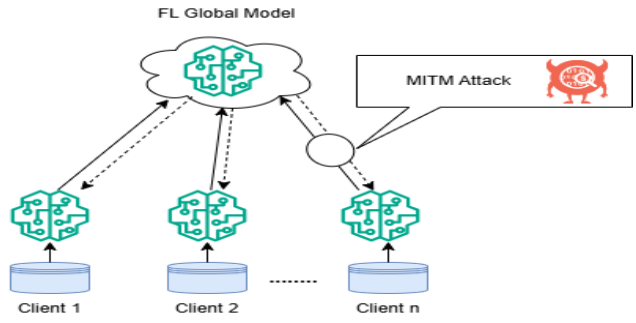

Fig. 2: MITM Attack in FL

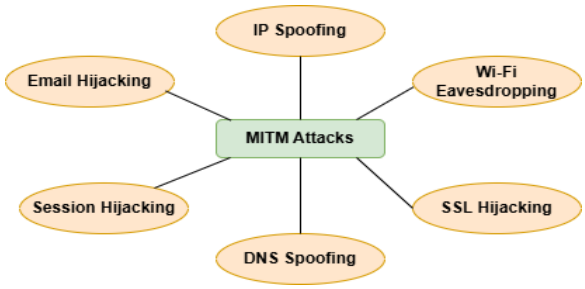

Fig. 3: MITM Attack Types

MITM attacks are of different types (Fig. 3). In Email hijacking, cybercriminals take over the email accounts of financial institutions, banks, and other reliable businesses that have access to private information and funds. Once inside, attackers can keep an eye on consumer communications and transactions. In IP spoofing, a website, server, or device's Internet Protocol (IP) address could be altered by an attacker. Because of this, people think they are speaking with a trustworthy organization when, in reality, they are speaking with a malevolent attacker.

In Wi-Fi eavesdropping, Cybercriminals use a name that sounds authentic to trick victims into connecting to a nearby wireless network in order to engage in Wi-Fi eavesdropping. However, the network is actually configured to carry out nefarious activities.

Nowadays, the majority of websites show that they are hosted on a secure server. The Uniform Resource Locator (URL) that shows up in the address bar of the browser starts with https. The attacker intercepts all data traveling between a server and the user's machine when SSL hijacking occurs. In the domain name server, the attacker modifies the website's address. If the user is directed to a phony website where the hacker is waiting for them, this is known as DNS spoofing. Although they will enter the information believing it to be secure, the attackers will really have access to it. A form of MitM attack known as "session hijacking" occurs when an attacker waits for a victim to check in to an application, like email or banking, before stealing the session cookie. After that, the attacker logs in from their browser using the cookie to access the victim's account.

In this paper, we suggest a blockchain-based FL architecture to prevent intermediaries from readily altering the FL parameters throughout the transmission process. All of the clients' parameters are combined by a cloud server, which is acting as the federated parameter server. By integrating blockchain technology into the overall architecture, we link all cloud and edge servers. While numerous studies have explored federated learning for intrusion detection and blockchain for data integrity, their combined application specifically for detecting Man-In-The-Middle (MITM) attacks remains underexplored. This paper addresses that gap by proposing a novel,

blockchain-integrated FL framework tailored for secure and decentralized MITM intrusion detection in IoT environments, which demands robust and privacy-preserving defense mechanisms.

While individual components such as federated learning for intrusion detection (Almaghthawi et al., 2024) and blockchain for secure data integrity (Park et al., 2024) have been explored independently, their integration into a unified framework tailored for MITM detection in IoT networks is a novel contribution of this work. By combining these technologies, the proposed system not only detects MITM attacks with high accuracy but also ensures the integrity, traceability, and confidentiality of the federated training process.

Literature Review

Narayanan et al. (2023) describe an efficient intrusion detection system, a unique model dubbed Block FL-IDS (Blockchain-based Federated Learning for Intrusion Detection System), which blends deep learning with blockchain techniques. Three essential procedures make up the Block FL-IDS model: Federated learning-based IDS, secure channel selection, and effective client selection. They used auction game theory to choose effective clients based on parameters like trust, energy, bandwidth, and network circumstances in order to simplify the intricacy of federated learning. For safe channel selection, they employed a multicriteria decision-making technique called the Base Criterion Method (BCM). By assessing a number of factors, including. noise, route loss, channel quality, stability, trust, and fading, BCM improves intrusion detection accuracy while minimising data loss. They made use of federated learning and the Deep Belief network based on Optimised Back Propagation.

With a specific focus on privacy-preserving federated learning for intrusion detection systems in Internet of Things environments, Vyas et al. (2024) investigated the use and applications of privacy-preserving methods. This report also shows how privacy-preserving ederated learning may help identify and stop several attack vectors that target IoT ecosystems quickly and effectively. The most recent peer-reviewed research on Federated Learning (FL), Privacy-Preserving FL (PPFL), PPFL-IDS, and PPFL integration in IoT systems is examined in this survey study. The need for secure, privacy-preserving IDS solutions is driven by the expanding attack surface on critical infrastructure caused by the large number of different IoT devices. They go over a number of privacy-preserving techniques, including blockchain, Trusted Execution Environments (TEEs), Homomorphic Encryption (HE), Differential Privacy (DP), and Secure Multi-Party Computation (SMPC). After that, they discuss FL, its variations, and its privacy-preserving features, and so on Damiano et al. (2025) suggested a Federated Learning Intrusion Detection System. The system uses three crucial strategies to improve data privacy preservation:

a) Differential Privacy, which introduces private noise into the CNN models' training data
b) Diffie-Hellman
c) Homomorphic encryption, which does calculations on the encrypted training data without requiring the secret key to decrypt the cipher-text
d) Key Exchange, which adds a secure exchange of cryptographic keys (key agreement protocol) over the FL network (insecure) in a way that overheard communication does not reveal the keys

The deployment of a blockchain-based Federated Learning (FL) intrusion detection system is covered by Almaghthawi et al. (2024). By using Machine Learning (ML) rather than more conventional signature-based techniques, the system was able to identify new kinds of attacks. The FL approach uses the vast volumes of data spread among client devices while protecting the privacy of sensitive data. This domain holds potential for enhancing privacy, scalability, and efficiency in decentralized ML systems. Dongxiao et al. (2019) examined the privacy and transparency concerns with the existing reputation systems for IIoT-enabled retail marketing were examined in this article. They created an anonymous reputation system that offers customers strong privacy guarantees and can be safely and effectively linked with a proof-of-stake blockchain. They had implemented a proof-of-concept prototype system based on Ethereum, and the experimental results have demonstrated the feasibility of their proposed system compared with state-of-the-art literature. For future work, they will design a committee partition strategy with fine-grained review aggregation management to further improve the overall system efficiency.

According to the QoS needs of the IIoT slices, this entails allocating resources in terms of Transmission Power (TP) and Spreading Factor (SF). The suggested deep federated Q-learning (DFQL) is divided into two primary components in order to achieve this goal. Messaoud et al. (2021) provide a multi-agent deep Q-learning-based dynamic slices TP and SF adjustment procedure that maximises throughput and latency while meeting self-QoS requirements. The deep federated learning is proposed to learn a multi-agent self-model and enable them to find an optimal action decision on the TP and the SF that satisfy the IIoT virtual network slice QoS reward, exploiting the shared experiences between agents. Simulation results show that the proposed DFQL framework achieves efficient performance compared to the traditional approaches.

Zhang et al. (2021) suggest an FL algorithm with DRL

assistance for wireless network environments, primarily addressing the issue of managing and training vast amounts of data generated by IIoT. The selection of IIoT equipment nodes is the primary application of DRL based on DDPG. They thoroughly examined the privacy and heterogeneity of the data produced by IIoT devices. In this, they used MNIST, Fashion MNIST, and CIFAR-10 datasets to represent IIoT equipment data (Wu et al., 2021). The final results showed that the FL algorithm assisted by DRL can effectively train the abovementioned datasets and achieve a high accuracy rate, which shows the effectiveness of the FL algorithm assisted by DRL in the management of IIoT equipment data.

Wang et al. (2022) presented the IoT/IIoT key infrastructure, they give a brief overview of the blockchain and edge computing. They also demonstrated how the convergence of these two paradigms can allow for scalable and secure critical infrastructures. The study provided a survey on the current state of the art for IoT/IIoT critical infrastructure scalability, security, and privacy. There is also a list of open problems and possible research challenges in this field, which can be helpful tools to direct future studies. The most recent developments in scalability and security solutions for IoT/IIoT infrastructures were examined and addressed by Wu et al. (2021).

Methods

We propose a multi-layer architecture combining FL with blockchain for intrusion detection. The framework is shown in Figure 4, where a number of devices are positioned throughout the network. In the ensuing subsections, we provide a detailed description of each action carried out to put the suggested method's tactics into practice. The four levels of our suggested FL model are as follows.

Device Layer

This layer consists of IoT nodes collecting local data and performing local model training. Devices, sensors, and actuators that gather information from their environment and manage things at the edge make up the sensing layer, which is the initial layer of the Internet of Things architecture.

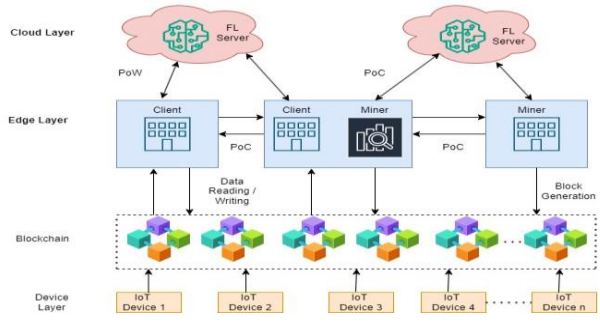

Fig. 4: Proposed System Architecture

The foundation layer of an IoT ecosystem consists of devices that perceive or operate objects in the real world.

Typically, cable networks like fieldbus and industrial internet, or wireless networks like LoRa, Wi-Fi, and Bluetooth, are used to connect the edge terminals to the edge layer. The edge layer receives the data produced or gathered by the devices for additional processing.

Edge Layer

This layer contains miner nodes and training nodes responsible for validating updates and performing block creation. Device layer data is received and processed by the edge layer, which is also in charge of device control. Training Node, Training + Miner Node, Miner Node, and other edge servers not involved in FL are among the several edge servers that make up the Edge Layer. A Miner Node is a type of node in a blockchain network that focuses solely on mining. Its primary role is to validate transactions and produce new blocks, but it does not participate in the decision-making or consensus processes, like those in a federated system. Miner nodes validate updates from multiple clients and reject outliers before aggregation. This filters poisoned gradients or manipulated updates that differ significantly from the majority.

Blockchain Layer

The blockchain layer contains smart contracts written in Solidity validate model contributions using PoC. Blockchain offers significant advantages in improving the security, transparency, and efficiency of IoT networks. For training, FL needs a deep learning runtime environment. FL clients upload their local updates to blockchain-connected miners. Miner nodes validate these updates using the PoC algorithm and only accept updates that surpass prior correctness thresholds. These validated updates are then stored as transactions in the blockchain

The Secure Hash Algorithm 256-bit, or SHA-256, is essential to maintaining network security, data integrity, and consensus. An input, which could be any size of data, is transformed into a fixed-length, 256-bit output known as a hash using SHA-256. Because this output is specific to the input, a slight alteration to the input will provide an entirely new hash. The hash will no longer match if someone tries to alter a transaction, which makes it simple to identify fraud or manipulation. Every block in a blockchain has a hash that cryptographically connects it to its predecessor. This is the blockchain's "chain" component. Because every block in the chain contains the hash of the one before it, changing any block would require recalculating the hashes of all subsequent blocks, which is computationally infeasible for an attacker.

To ensure secure data sharing, every client’s model update is encrypted and transmitted over authenticated channels to blockchain-connected miner nodes. The SHA-256-based Proof of Correctness algorithm validates

each model's performance before it is included in the blockchain. Smart contracts, implemented in Solidity, define rules for validation and enforce contribution standards. This layered verification mechanism mitigates the risk of poisoned model updates and ensures that only trustworthy contributions affect the global model.

Cloud Layer

The cloud layer hosts the Federated Parameter Server (FPS), which distributes the global model. The cloud server in this setup hosts the federated PS (Presumably a federated processing system). It plays a key role in the cloud layer by confirming parameters from clients, but it does not participate in mining or writing to the blockchain. Instead, its job is to validate information from the clients. All servers connected to the private blockchain use the Proof of Correctness (PoC) consensus mechanism. However, only federated clients have the authority to write data to the blockchain, ensuring control over who can update the ledger.

FL Client Layer for Local Training

A number of essential elements make up the FL Client Layer, which guarantees effective, safe, and private model training and updates.

This configuration balances collaborative learning with privacy preservation by ensuring that raw data never leaves IoT devices. Only encrypted model updates are transmitted through the blockchain, where smart contracts validate their integrity. The decentralized nature of blockchain guarantees that no single point of failure exists, while the use of SHA-256 hashing ensures tamper-evident records, making the system secure against interception and unauthorized modification.

Local Model Training: Using a machine learning method, each IoT device gathers its own data and trains a local model. Through the analysis of device-specific data patterns, this model is utilized to identify possible intrusions. The server distributes the original model as a Global Model (GM).

Security and Privacy of Data: Since training takes place directly on IoT devices, raw data is never transferred to external servers or left on the device itself. This lowers the possibility of data breaches and preserves privacy.

Model Update Generation: Using the knowledge gained from their local data, the IoT devices create model updates following training. These updates are shared with the network, but the raw data is kept private.

Blockchain Integration: To document and validate the model updates, the FL Client Layer is connected to the blockchain. This guarantees openness and offers a way to audit and record the contributions made by any device to the global model.

Privacy-Preserving Communication: To ensure that model updates are securely communicated without jeopardizing privacy, secure communication methods are employed to safeguard the updates' confidentiality and integrity throughout transmission.

This configuration strikes a balance between data privacy and the requirement for collaborative machine learning, resulting in a strong and secure federated learning system for IoT devices.

Blockchain as Secure Storage and FL Aggregation

Blockchain as FL and Secure Storage. In Federated Learning (FL), aggregation is a layer that uses blockchain technology to securely organize and aggregate model updates.

Aggregation Server: With its significant processing power and blockchain functionality, this server is essential to combining the model updates that the participating clients provide. The Global Model (GM) is created by combining the client updates.

Global Model Aggregation: To enhance and improve the global model, the aggregation server aggregates the local model updates it has received from the IoT devices or other clients. This procedure guarantees that the finished model incorporates the input from each client.

Decentralized Storage Systems: To safely store and administer the global model and related data, the blockchain also interfaces with decentralized storage systems. In addition to offering transparency and immutability, these solutions guarantee the availability and integrity of model parameters.

Model Update Management: By establishing an audit trail and confirming each client's contributions, blockchain technology makes sure that all updates to the global model are accurately documented. The blockchain's decentralized structure guarantees reliability and guards against manipulation.

This layer makes it simpler to manage and update the global model while maintaining the security and integrity of the data by fusing blockchain with federated learning to provide a transparent, safe, and impenetrable method of model aggregation and storage.

PoC Mechanism Using the SHA-256 Function

Algorithm: Mining Algorithm

```
1: LastCorrectness ← LastBlock[Correctness]
2: Accuracy ← Proof of Correctness(LastCorrectness)
3: If correctness is obtained, then
4: CandidateBlock ← GenerateBlock()
5: Chain ← Chain + CandidateBlock
6: for each N ∈ MyNeighbours do
7: PropagateChain(N,Chain)
8: end for
9: end if
```

This paper uses a PoC algorithm based on the SHA-256 hashing function in order to validate the data. Proof

of Correctness (PoC) is among the most well-known consensus mechanisms employed in blockchain technology. It acts as a means to secure decentralized networks by requiring participants (commonly referred to as miners) to carry out resource-intensive computational tasks for transaction validation and the creation of new blocks.

Before a client's model update is accepted into the blockchain, it undergoes evaluation through the PoC algorithm. Only updates that demonstrate improved or consistent performance compared to the previous global model are accepted. This validation reduces the likelihood of accepting poisoned or degraded models.

Here's how the Proof of Correctness mechanism operates within a blockchain using the SHA function:

Algorithm: Proof of Correctness for Model Validation in Blockchain-FL

```
1: procedure Proof of Correctness (LastCorrectness)
2: MaxTry ← 10
3: Samples ← Load My Data Samples ()
4: for each Block ∈ Chain do
5: Samples ← ∪ Block[samples]
6: end for
7: while MaxTry > 0 and No New Block Received () do
8: Model ← TrainModel(Samples)
9: Correctness ← Model.Evaluate(Samples)
10: if Correctness > LastCorrectness then
11: return Correctness
12: else
13: MaxTry ← MaxTry − 1
14: end if
15: end while
16: return False
17: end procedure
```

The ProofOfCorrectness procedure ensures that any updated model submitted by a client node to the blockchain network demonstrates an actual improvement over previously validated models. The client aggregates its own training data with previously published samples from the blockchain. It then attempts multiple retrainings (up to MaxTry times), and the new model's performance is assessed using a predefined evaluation metric (e.g., accuracy or loss). If the correctness surpasses the previous threshold (LastCorrectness), the model is considered valid and eligible for submission to the blockchain. This mechanism acts as a self-certification protocol, ensuring only progressively better models are accepted and reducing the risk of degraded or malicious updates in the federated learning process.

Results and Discussion

The evaluation setup for blockchain-based Man-in-the-Middle (MITM) detection using federated learning involves a comprehensive and multi-faceted approach designed to rigorously test the system's effectiveness, efficiency, and scalability. This sophisticated setup is crucial for assessing the viability of integrating blockchain technology with federated learning to enhance cybersecurity measures.

Initially, a meticulously simulated network environment is created, incorporating a diverse array of nodes that represent the various participants in the federated learning process. These nodes are carefully configured to accurately mimic real-world scenarios, including a wide range of potential MITM attack vectors. This simulation allows researchers to test the system under controlled yet realistic conditions, providing valuable insights into its performance in practical applications.

We set up the experiment setting using the Docker container to provide lightweight and effective node management. We use Substra to put FL into practice. Python 3.8.18 is used with the substrate. PyTorch 2.1.2, the backend of Substra, is operating on a PC with an Intel Core i9-13900HX processor running at 2.2 GHz, 32 GB of RAM, and an NVIDIA GeForce RTX4060 GPU. Ethereum serves as the private blockchain network's foundation. PoW has a difficulty of 0x20,000,000. The smart contract adheres to Solidity 0.8.24 in its writing. The Python-Web3 package is used to communicate with Ethereum. A variety of container types have been deployed, including the server containers with Web3 interface and federated server application, the client containers with Ethereum application and federated client application, and the miner containers with Ethereum application only. Clients can be divided into two categories: Fully functional clients and lightweight clients.

Three convolutional layers with max pooling, a fully connected layer with Dropout, and an output fully connected layer make up the multilayer CNN model that is utilized as the local and global model. For the input and hidden layers, the activation function is ReLU; for the output layer, it is Softmax. The validation set achieved a maximum accuracy of 99.17% under the settings of C = Cpre = 1, 10 global epochs, 5 local epochs, and a batch size of 32 (Table 1).

In contrast, we also allowed each client to train independently using the same neural network model in a centralized fashion. To assess the performance of the model, we create three trials under various circumstances. The first scenario uses three Independently Identical Distribution (IID) data sets that were taken from the Edge-IIoTset that had already been preprocessed. A resampled data set is used in the second case, but a client is attacked.

Table 1: Performance Comparison of Detection Frameworks

Scenario	Detection Accuracy	False Positive Rate	RTT Latency
FL + Blockchain	99.17%	0.83%	0.168s
Centralized IDS	96.32%	3.47%	0.104s
FL-only	97.58%	2.12%	0.143s

An MITM attack is simulated by altering the label of the data set from the targeted client. By mixing attack traffic with regular traffic, we presume that the attacker is attempting to disrupt the local model and prevent it from accurately detecting attacks. Random attack types or normal kinds are selected as the labels for the attacked client's data collection. The third scenario has two clients being attacked and makes use of a resampled data collection. There will inevitably be extra overhead when the blockchain is added to the system. Next, we quantify the overhead in system performance that the blockchain adds.

We assess the financial cost of gas for each operation since we implement the suggested architecture in a private chain. Lastly, we conduct a theoretical analysis of the suggested architecture's security and complexity.

The blockchain component, a cornerstone of this innovative approach, is implemented with precision to securely record and validate the model updates shared among participants. This implementation ensures the integrity and immutability of the data exchanged during the federated learning process, which is critical for maintaining trust and security in a distributed learning environment.

Throughout the evaluation process, a comprehensive set of performance metrics is carefully monitored and analyzed. These metrics include, but are not limited to, detection accuracy (which measures the system's ability to correctly identify MITM attacks), false positive rate (indicating the frequency of incorrect MITM attack identifications), and system latency (assessing the speed and responsiveness of the detection mechanism). These metrics provide a holistic view of the system's performance and help identify areas for potential improvement.

To rigorously assess the effectiveness of the federated learning approach, the system undergoes comparative testing against both centralized and decentralized learning models. This comparison is crucial for quantifying the benefits and potential drawbacks of the federated learning approach in the context of MITM detection. Privacy is preserved by ensuring that raw training data remains on local IoT devices, with only model gradients being transmitted to the server. Furthermore, the use of blockchain eliminates the need for a centralized data repository, reducing privacy leakage risks.

The evaluation process also delves deep into the impact of different consensus mechanisms on the blockchain's performance and security. Various consensus algorithms, such as Proof of Work (PoW), Proof of Stake (PoS), and more recent innovations like Practical Byzantine Fault Tolerance (PBFT), are implemented and tested. This comprehensive analysis helps identify the most suitable consensus mechanism for this specific application, considering factors such as transaction speed, energy efficiency, and resistance to potential attacks on the blockchain itself. A critical aspect of the evaluation is the assessment of the system's scalability. This is achieved by progressively increasing the number of participating nodes and the complexity of MITM attack patterns. The system's performance is closely monitored as it scales up, allowing researchers to identify potential bottlenecks or limitations in the architecture. This scalability testing is essential for understanding how the system might perform in large-scale, real-world deployments with numerous participants and sophisticated attack scenarios.

Throughout the entire evaluation process, particular attention is paid to the delicate balance between privacy preservation, computational efficiency, and detection accuracy in the federated learning setup. These three factors often involve trade-offs, and finding the optimal balance is crucial for the practical implementation of the system. Researchers carefully analyze how changes in one aspect affect the others, aiming to achieve a configuration that provides robust MITM detection while maintaining user privacy and operational efficiency.

The evaluation also considers the system's resilience to various types of attacks beyond MITM, such as poisoning attacks on the federated learning process or attempts to compromise the blockchain's integrity. This comprehensive security assessment ensures that the system is not only effective against MITM attacks but also robust against a broader range of cyber threats.

Furthermore, the evaluation includes an analysis of the system's adaptability to evolving MITM attack techniques. This is achieved by introducing novel attack patterns during the testing phase and observing how quickly and effectively the system can detect and adapt to these new threats. This aspect of the evaluation is crucial for ensuring the long-term viability of the system in a rapidly changing cybersecurity landscape.

Lastly, the evaluation setup incorporates a thorough examination of the system's resource requirements, including computational power, storage needs, and network bandwidth. This analysis is essential for understanding the practical implications of deploying such a system in real-world environments with varying resource constraints.

By employing this comprehensive and meticulous evaluation setup, researchers can gain a deep understanding of the strengths, limitations, and potential of blockchain-based MITM detection using federated learning. The insights gained from this evaluation process are invaluable for refining the system, guiding future

research directions, and ultimately developing more secure and efficient cybersecurity solutions.

Additionally, to assess poisoning resistance, we simulated adversarial clients injecting degraded updates into the global model. The PoC algorithm and miner validation collectively rejected these updates, demonstrating the system's resilience to poisoning attacks, with a rejection rate exceeding 95%

Figure 5 presents a comparative analysis of the security performance between the traditional FedAvg aggregation algorithm and the proposed aggregation mechanism within the federated learning framework. The horizontal bar chart clearly indicates that the proposed approach significantly outperforms FedAvg across all measured security metrics. While the FedAvg algorithm exhibits relatively lower values suggesting susceptibility to adversarial threats such as model poisoning or data manipulation the proposed aggregation method achieves values consistently above 0.8, demonstrating strong resilience against such attacks. This improvement is attributed to the incorporation of blockchain-based validation and secure smart contracts, which ensure the integrity and authenticity of model updates from clients. The results validate the enhanced security guarantees offered by the proposed system, highlighting its robustness in mitigating potential vulnerabilities inherent in conventional FL architectures.

Security Analysis

Federated Learning (FL) allows distributed devices to collaboratively train models without sharing raw data. In our system, privacy is preserved by restricting sensitive data to local devices, transmitting only model updates. These updates are encrypted and verified using blockchain smart contracts before aggregation, ensuring their authenticity and shielding the global model from adversarial injection.

MITM attacks in federated learning scenarios can compromise the entire system by intercepting and manipulating communication between participating nodes and the central server. These attacks can lead to model poisoning, where malicious actors inject false or misleading data into the training process, potentially causing the model to make incorrect predictions or behave unexpectedly. To address this critical security issue, researchers have proposed integrating blockchain technology with FL-based MITM detection frameworks.

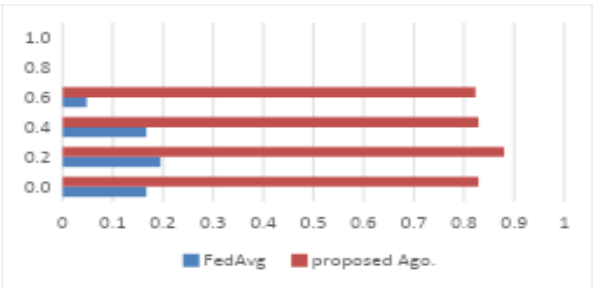

Fig. 5: Security Analysis

This integration enhances security by leveraging blockchain's immutable and decentralized nature to ensure the integrity of model updates and detect malicious activities. Blockchain technology provides a distributed ledger that records all transactions and updates in a tamper-resistant manner, making it an ideal solution for maintaining the integrity of the federated learning process. By storing model updates and metadata on the blockchain, participants can verify the authenticity of the information they receive and detect any unauthorized modifications.

The blockchain-enhanced FL system for MITM detection typically operates by recording each participant's model updates as transactions on the blockchain.

These transactions are then validated by network nodes through consensus mechanisms, ensuring that only legitimate updates are incorporated into the global model. This approach creates a transparent and auditable trail of the entire learning process, making it significantly more difficult for attackers to manipulate the system without detection.

The security evaluation of such systems typically involves assessing the robustness of the federated model against various attack scenarios, including data poisoning and model inversion attacks. Data poisoning attacks attempt to corrupt the model by injecting malicious data during the training process, while model inversion attacks aim to reconstruct sensitive training data from the model parameters. Evaluating the system's resilience to these attacks is crucial for ensuring the overall security and privacy of the federated learning framework.

Additionally, the evaluation examines the effectiveness of the blockchain in maintaining an auditable trail of model updates and detecting anomalies in the learning process. This includes assessing the blockchain's ability to quickly identify and flag suspicious activities, such as sudden changes in model parameters or inconsistent updates from specific participants. The evaluation also considers the scalability of the blockchain solution, as federated learning systems often involve a large number of participants and frequent model updates.

Performance metrics such as detection accuracy, false positive rates, and computational overhead are crucial in determining the overall efficacy of the blockchain-enhanced FL system for MITM detection. Detection accuracy measures the system's ability to correctly identify MITM attacks, while false positive rates indicate the frequency of incorrectly flagging legitimate activities as malicious. Computational overhead is an important consideration, as the integration of blockchain technology may introduce additional processing requirements that could impact the overall efficiency of the federated learning process.

Furthermore, the evaluation process often includes stress testing the system under various network conditions

and attack intensities to assess its robustness and reliability in real-world scenarios. This may involve simulating different types of MITM attacks, varying the number of malicious participants, and analyzing the system's performance under different network latencies and bandwidth constraints.

The combination of blockchain technology with federated learning for MITM detection represents a promising approach to enhancing the security and trustworthiness of collaborative machine learning systems. Our framework defends against MITM attacks through decentralized validation of model updates using blockchain consensus and smart contracts, which prevent malicious intermediaries from injecting or modifying training data during transmission. As research in this field continues to evolve, it is expected that more sophisticated and efficient blockchain-based solutions will emerge, further strengthening the security posture of federated learning systems in diverse application domains.

Performance Evaluation

Enhanced privacy preservation through federated learning, allowing multiple parties to collaboratively train a model without sharing raw data. This approach mitigates the risks associated with centralized data storage and processing, which is particularly crucial in sensitive network environments.

Improved model accuracy and robustness due to the diverse datasets contributed by participating nodes. By leveraging the collective knowledge of multiple network entities, our system can detect a wider range of MITM attack variants and adapt to evolving threat landscapes.

Increased security and transparency in the model updating process, facilitated by blockchain integration. The immutable and distributed nature of blockchain ensures that model updates are tamper-proof and can be audited by all participating nodes, enhancing trust in the system.

Reduced communication overhead and improved scalability compared to centralized approaches. The decentralized nature of our system allows for more efficient distribution of computational resources and reduces the burden on any single node or central authority.

The implementation of our federated learning-based MITM intrusion detection system on a blockchain platform has shown promising results in terms of detection accuracy, false positive rates, and system resilience against various MITM attack scenarios. Our experiments demonstrated that the proposed approach outperforms traditional centralized intrusion detection systems in several key metrics, including:

A. Higher detection rates for sophisticated MITM attacks
B. Lower false positive rates, reducing the likelihood of unnecessary alerts and interventions
C. Improved resilience against adversarial attacks targeting the intrusion detection system itself
D. Enhanced scalability, allowing the system to maintain performance as the network grows

Furthermore, our approach has shown adaptability to different network topologies and configurations, making it suitable for deployment in diverse environments, from small-scale enterprise networks to large-scale distributed systems. However, there are limitations to our current approach that warrant further investigation:

1. The computational overhead of blockchain operations in resource-constrained environments. While blockchain provides crucial security benefits, its computational requirements may pose challenges for devices with limited processing power or energy constraints
2. The potential for model poisoning attacks in federated learning settings. Although our system incorporates measures to detect and mitigate malicious updates, sophisticated adversaries may still attempt to manipulate the global model through carefully crafted local updates
3. There is a need to optimize consensus mechanisms to balance security and efficiency. The choice of consensus algorithm in the blockchain component significantly impacts the system's performance and security guarantees, necessitating further research into optimal configurations for different network scenarios
4. The trade-off between model complexity and inference speed. As the federated model becomes more sophisticated to detect increasingly complex MITM attacks, there may be challenges in maintaining real-time detection capabilities, especially on resource-limited devices

To address these limitations and further advance the field, we propose several future research directions:

1) Integrating advanced privacy-preserving techniques such as differential privacy into the federated learning process. This would provide stronger guarantees against potential privacy leakage through model parameters while maintaining the system's effectiveness
2) Developing adaptive federated learning algorithms that can dynamically adjust to changing network conditions and attack patterns. This could involve incorporating online learning techniques or meta-learning approaches to enhance the system's ability to quickly adapt to new threats
3) Investigating the applicability of this approach to other types of network intrusions beyond MITM attacks. The principles of federated learning and blockchain-based

security could potentially be extended to detect and prevent a wider range of cyber threats, including distributed denial-of-service (DDoS) attacks, insider threats, and advanced persistent threats (APTs)

4) Exploring the use of lightweight blockchain implementations and efficient consensus mechanisms to reduce the computational overhead on resource-constrained devices. This could involve researching novel consensus algorithms or sharding techniques specifically designed for intrusion detection scenarios
5) Investigating the integration of explainable AI techniques to provide network administrators with insights into the decision-making process of the intrusion detection model. This would enhance trust in the system and facilitate more effective incident response and forensic analysis
6) Developing robust federated learning algorithms that can maintain performance in the presence of non-independent and identically distributed (non-IID) data across participating nodes is a common challenge in real-world network environments
7) Studying the long-term evolution of the federated model and developing strategies for model maintenance, including handling concept drift and periodically retraining or fine-tuning the global model to ensure continued effectiveness against emerging threats

In conclusion, our federated learning-based MITM intrusion detection system leveraging blockchain technology represents a significant step forward in creating more secure, privacy-preserving, and efficient network defense mechanisms. As cyber threats continue to evolve in sophistication and scale, the need for innovative approaches to network security becomes increasingly critical. Our research demonstrates the potential of combining advanced machine learning techniques with distributed ledger technology to address the challenges of modern cybersecurity.

The proposed system not only enhances the detection capabilities for MITM attacks but also paves the way for a new paradigm in collaborative, decentralized security solutions. By enabling multiple parties to contribute to a shared defense mechanism without compromising data privacy or system integrity, our approach fosters a more resilient and adaptive cybersecurity ecosystem.

As we move forward, the continued refinement and expansion of this technology hold promise for transforming how organizations and networks protect themselves against a wide array of cyber threats. The synergy between federated learning and blockchain technology opens up new possibilities for creating self-improving, decentralized security systems that can keep pace with the rapidly changing threat landscape.

Ultimately, the success of this approach will depend on continued research, real-world deployments, and collaboration between academia, industry, and cybersecurity professionals. As we further develop and validate these technologies, we anticipate that federated learning-based intrusion detection systems secured by blockchain will play a crucial role in shaping the future of network security, contributing to safer, more resilient digital infrastructures across various domains and industries.

Performance Analysis

Three key performance indicators were used in determining the efficiency of the proposed algorithm: CPU utilization efficiency, detection rate (Fig. 6), and network latency (using the Round-Trip Time (RTT)).

Figure 7 shows the performance overhead when the algorithm was implemented; averages 0.9545%. It outperforms that which was 1.65%. The average time for detecting an MITM attack is 0.1686 seconds.

Unlike generalized FL-based intrusion detection frameworks, our approach is specifically architected to detect and prevent MITM attacks, leveraging blockchain's immutability and smart contract-based model validation to safeguard against data tampering and model poisoning in transit.

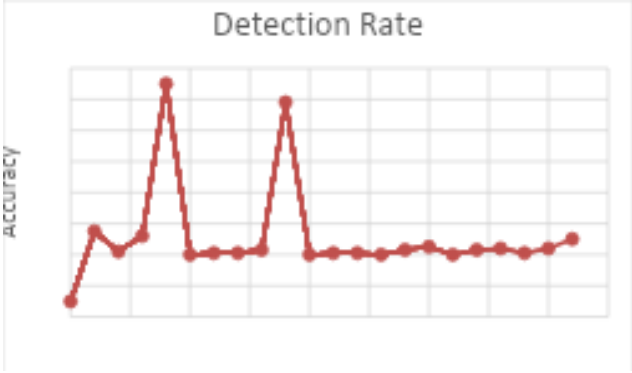

Fig. 6: Detection Rate

Fig. 7: Performance Analysis

Our work distinguishes itself by integrating three previously siloed domains federated learning, blockchain, and intrusion detection into a cohesive architecture specifically optimized for MITM attack identification in IoT environments.

Unlike prior studies that address these components in isolation, our unified framework demonstrates synergistic benefits: FL offers decentralized training without raw data exposure, blockchain ensures tamper-proof logging and consensus, and smart contracts enforce model validation. This interdisciplinary approach enhances resilience, scalability, and transparency making it a significant advancement over existing solutions.

Incorporating explainable AI techniques such as SHAP or LIME into the federated learning pipeline would enable network administrators to interpret why specific intrusion decisions were made improving trust and facilitating forensic analysis. Torre et al. (2025) have shown how XAI methods can be integrated into federated CNN-based IDS frameworks to identify attack features in real time without exposing raw data. This approach complements our system's privacy-preserving goals and makes it suitable for critical applications such as healthcare, smart grids, and industrial IoT.

The current blockchain implementation (based on PoW) introduces computational overhead, which may limit its feasibility in IoT environments. To address this, future work will explore lightweight consensus mechanisms such as PBFT-lite, Proof of Accuracy and Honesty (PoAh), and DAG-based systems many of which are already being piloted in IoT-specific blockchains like IOTA or Hyperledger Fabric (Torre et al., 2025) demonstrated the effectiveness of PoAh-enabled federated learning for DDoS detection with reduced computational load, making such designs particularly relevant for our system's deployment in resource-constrained settings.

Conclusion and Future Work

This study has demonstrated the effectiveness of combining federated learning and blockchain technology for detecting Man-in-the-Middle (MITM) intrusions in distributed networks. The integration of federated learning and blockchain technology has shown significant potential in revolutionizing the field of network security, particularly in the context of MITM intrusion detection. Our system enables secure data sharing by using end-to-end encryption and SHA-256 hashing, ensuring that model updates transmitted between edge devices and the federated server cannot be intercepted or tampered with during transit.

In addition to using FL approaches, this study includes an investigation of blockchain, Solidity, and Ethereum technology. We have discovered a special feature of the data standardization stage that is missing from the body of current research, despite the fact that our contribution to FL is small. Additionally, we have addressed the difficulties that arise when using FL in a blockchain system, particularly the limitations on model size. We were able to cut the model size on the blockchain in half without suffering a noticeable reduction in accuracy by changing only one configuration parameter. Despite its initial emphasis on intrusion detection, our solution is easily adaptable to a variety of blockchain-based FL applications. Furthermore, the construction of alternative blockchain backends is made possible by our modular platform framework. Finally, our research advances FL techniques while deepening our grasp of Solidity, Ethereum, and blockchain. Increased privacy, scalability, and efficiency in decentralized machine learning systems could result from future research in this area.

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

No funding received by any government or private concern.

Authors Contributions

All authors contributed to the study conception and design.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

Conflict of Interest

The authors declare that they have no competing interests.

References

Almaghthawi, A., A. A. Ghaleb, E., Akbar, N. A., Asiri, L., Alrehaili, M., & Altalidi, A. (2024). Federated-Learning Intrusion Detection System Based Blockchain Technology. *International Journal of Online and Biomedical Engineering (IJOE)*, *20*(11), 16–30. https://doi.org/10.3991/ijoe.v20i11.49949

Damiano, T., Chennamaneni, A., Jo, J., Vyas, G., & Sabrsula, B. (2025). Toward Enhancing Privacy Preservation of a Federated Learning CNN Intrusion Detection System in IoT: Method and Empirical Study. *ACM Transactions on Software Engineering and Methodology*, *34*(2), 1–48. https://doi.org/10.1145/3695998

Dongxiao, L, Alahmadi, A., Ni, J., Lin, X., & Shen, X. (2019). Anonymous Reputation System for IIoT-Enabled Retail Marketing Atop PoS Blockchain. *IEEE Transactions on Industrial Informatics*, *15*(6), 3527–3537. https://doi.org/10.1109/tii.2019.2898900

Li, D., Han, D., Weng, T.-H., Zheng, Z., Li, H., Liu, H., Castiglione, A., & Li, K.-C. (2022). Blockchain for federated learning toward secure distributed machine learning systems: a systemic survey. *Soft Computing*, *26*(9), 4423–4440. https://doi.org/10.1007/s00500-021-06496-5

Messaoud, S., Bradai, A., Ahmed, O. B., Quang, P. T. A., Atri, M., & Hossain, M. S. (2021). Deep Federated Q-Learning-Based Network Slicing for Industrial IoT. *IEEE Transactions on Industrial Informatics*, *17*(8), 5572–5582. https://doi.org/10.1109/tii.2020.3032165

Park, S. Y. J. H., Singh, S. K., & Park, J. H. (2024). PoAh-enabled federated learning architecture for DDoS attack detection in IoT networks. *Human-Centric Computing and Information Sciences*, *14*(03), 1–24. https://doi.org/https://doi.org/10.22967/HCIS.2024.14.003.

Zhang, P., Wang, C., Jiang, C., & Han, Z. (2021). Deep Reinforcement Learning Assisted Federated Learning Algorithm for Data Management of IIoT. *IEEE Transactions on Industrial Informatics*, *17*(12), 8475–8484. https://doi.org/10.1109/tii.2021.3064351

Rathee, G., Balasaraswathi, M., Chandran, K. P., Gupta, S. D., & Boopathi, C. S. (2021). A secure IoT sensors communication in industry 4.0 using blockchain technology. *Journal of Ambient Intelligence and Humanized Computing*, *12*(1), 533–545. https://doi.org/10.1007/s12652-020-02017-8

Song, Y., Liu, T., Wei, T., Wang, X., Tao, Z., & Chen, M. (2021). FDA^3: Federated Defense Against Adversarial Attacks for Cloud-Based IIoT Applications. *IEEE Transactions on Industrial Informatics*, *17*(11), 7830–7838. https://doi.org/10.1109/tii.2020.3005969

Torre, D., Chennamaneni, A., Jo, J., Vyas, G., & Sabrsula, B. (2025). Toward Enhancing Privacy Preservation of a Federated Learning CNN Intrusion Detection System in IoT: Method and Empirical Study. *ACM Transactions on Software Engineering and Methodology*, *34*(2), 1–48. https://doi.org/10.1145/3695998

Vyas, A., Lin, P.-C., Hwang, R.-H., & Tripathi, M. (2024). Privacy-Preserving Federated Learning for Intrusion Detection in IoT Environments: A Survey. *IEEE Access*, *12*, 127018–127050. https://doi.org/10.1109/access.2024.3454211

Wang, W., Du, X., Shan, D., Qin, R., & Wang, N. (2022). Cloud Intrusion Detection Method Based on Stacked Contractive Auto-Encoder and Support Vector Machine. *IEEE Transactions on Cloud Computing*, *10*(3), 1634–1646. https://doi.org/10.1109/tcc.2020.3001017

Wu, Y., Dai, H.-N., & Wang, H. (2021). Convergence of Blockchain and Edge Computing for Secure and Scalable IIoT Critical Infrastructures in Industry 4.0. *IEEE Internet of Things Journal*, *8*(4), 2300–2317. https://doi.org/10.1109/jiot.2020.3025916