Computers, Materials & Continua

Tech Science Press

Doi:10.32604/cmc.2025.066898

ARTICLE

FedCognis: An Adaptive Federated Learning Framework for Secure Anomaly Detection in Industrial IoT-Enabled Cognitive Cities

Abdulatif Alabdulatif*

Department of Computer Science, College of Computer, Qassim University, Buraidah, 52571, Saudi Arabia

*Corresponding Author: Abdulatif Alabdulatif. Email: ab.alabdulatif@qu.edu.sa

Received: 19 April 2025; Accepted: 09 June 2025; Published: 29 August 2025

ABSTRACT: FedCognis is a secure and scalable federated learning framework designed for continuous anomaly detection in Industrial Internet of Things-enabled Cognitive Cities (IIoTCC). It introduces two key innovations: a Quantum Secure Authentication (QSA) mechanism for adversarial defense and integrity validation, and a Self-Attention Long Short-Term Memory (SALSTM) model for high-accuracy spatiotemporal anomaly detection. Addressing core challenges in traditional Federated Learning (FL)—such as model poisoning, communication overhead, and concept drift—FedCognis integrates dynamic trust-based aggregation and lightweight cryptographic verification to ensure secure, real-time operation across heterogeneous IIoT domains including utilities, public safety, and traffic systems. Evaluated on the WUSTL-IIoTCC-2021 dataset, FedCognis achieves 94.5% accuracy, 0.941 AUC for precision-recall, and 0.896 ROC-AUC, while reducing bandwidth consumption by 72%. The framework demonstrates sublinear computational complexity and a resilience score of 96.56% across six security dimensions. These results confirm FedCognis as a robust and adaptive anomaly detection solution suitable for deployment in large-scale cognitive urban infrastructures.

KEYWORDS: Cognitive cities; federated learning; industrial IoT; anomaly detection; trust management; smart infrastructure; security

1 Introduction

Modern cities are rapidly evolving into intelligent urban environments known as Cognitive Cities. These cities rely on data, artificial intelligence (AI), and real-time decision-making to optimize services and infrastructure. At the core of this transformation is the Industrial Internet of Things (IIoT), which supports sectors like transportation, energy, public safety, manufacturing, and utilities through decentralized, adaptive, and secure infrastructures.

A foundational layer in these systems continuously generates massive streams of sensor data crucial to managing urban operations. However, these data sources form large-scale, distributed networks that are difficult to secure and must remain resilient to evolving threats and operational changes. Forecasts for 2025 estimate that IIoT devices within Cognitive Cities will exceed 75 billion, generating more than 79 zettabytes of data annually [1,2]. This data supports intelligent features like predictive maintenance and real-time responsiveness, but also introduces significant challenges related to scalability, security, and system performance. Traditional centralized anomaly detection methods fall short in such environments due to privacy concerns, scalability limits, and high computational demands.

To address these limitations, federated learning (FL) has emerged as a promising alternative. FL enables distributed IIoT nodes to collaboratively train models without sharing raw data, preserving privacy

while supporting decentralization [3]. However, FL still faces security threats that can degrade detection performance by as much as 30% [4], and concept drift due to time-varying operating conditions can reduce accuracy by 15–25% [5]. Additionally, frequent model updates in bandwidth-constrained networks can increase communication costs by a factor of five compared to centralized learning [6].

Given the diversity of IIoT environments in Cognitive Cities, effective trust management is essential. Without it, compromised nodes could introduce adversarial noise that undermines the learning process. To overcome these challenges, we propose FedCognis, a secure and scalable anomaly detection framework tailored for distributed Cognitive City infrastructures. FedCognis incorporates a trust-based update mechanism, Quantum Secure Authentication (QSA), and a Self-Attention LSTM (SALSTM) model to enable accurate, real-time, privacy-preserving anomaly detection across diverse IIoT domains.

Fig. 1 provides a high-level overview of the FedCognis framework within a Cognitive City environment. It illustrates federated IIoT devices deployed across smart domains such as traffic and energy, where each device performs local anomaly detection, transmits secure updates using Quantum Secure Authentication (QSA), and contributes to a centralized global model powered by a Self-Attention LSTM (SALSTM). The system represents a decentralized yet coordinated infrastructure that addresses key challenges related to privacy, adversarial threats, and real-time anomaly detection.

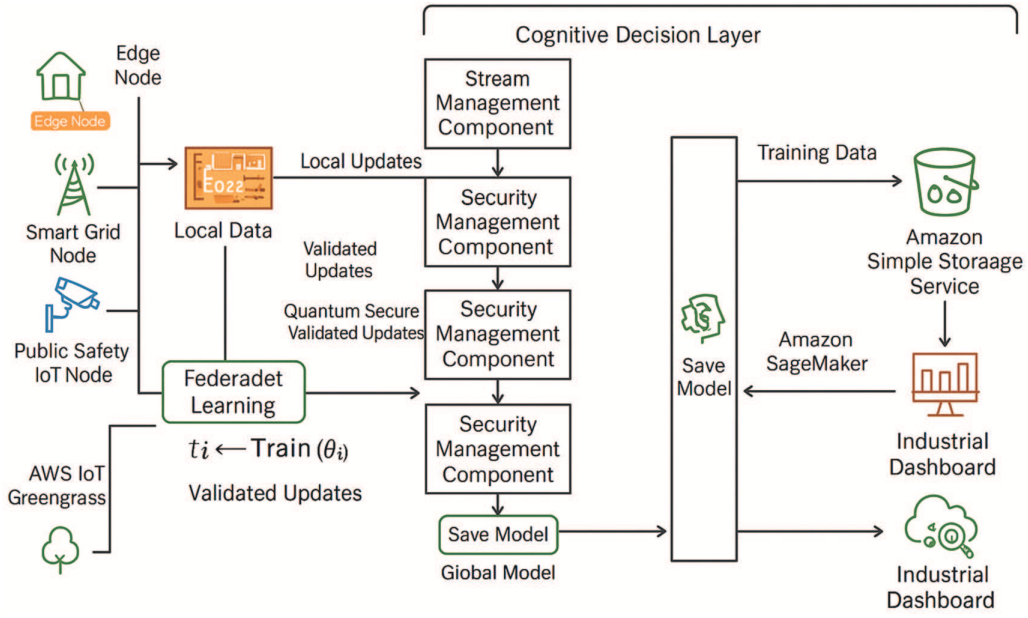

Figure 1: Layered architecture of FedCognis illustrating distributed IIoT nodes, edge-level trust evaluators, Quantum Secure Authentication (QSA), and centralized SALSTM-based global aggregation across Cognitive City domains

Despite these advances, security in IIoT-enabled Cognitive Cities (IIoTCC) remains a critical concern, as conventional federated learning (FL) and cryptographic security models are often too restrictive. Adversarial attacks on FL are feasible, and current quantum-secure authentication methods [7,8] involve significant computational overhead. To ensure model robustness, a real-time and adaptive security approach is essential. The core motivation behind FedCognis is to develop a scalable, efficient, and secure federated AI model that continuously learns from IIoTCC data while staying resilient to evolving threats. This research aims to reduce security risks, enhance anomaly detection accuracy, and improve the scalability of FL in IIoTCC networks.

As IIoTCC adoption expands, issues such as adversarial attacks, model poisoning, and data integrity breaches have become increasingly severe. Traditional FL models suffer from vulnerabilities to these threats, suffer communication inefficiencies, and struggle to adapt to dynamic industrial conditions [9,10]. While QSA enhances security, its high computational cost makes it unsuitable for real-time anomaly detection in resource-constrained IIoTCC environments. In contrast to academic settings, industrial deployments often lack the resilience provided by scalable, adaptive federated AI models and robust authentication mechanisms. This research proposes a robust, secure, and continuous anomaly detection solution for IIoTCC, integrating QSA and a self-attention LSTM architecture.

The proposed solution addresses both the security threats and adaptability challenges associated with FL-based anomaly detection in IIoTCC. It employs QSA and assigns dynamic trust scores to filter out adversarial updates and mitigate poisoning attacks. An adaptive trust-weighted aggregation mechanism is incorporated to combine local model updates while minimizing divergence from the global model, ensuring continuous adaptation to evolving IIoTCC conditions. The anomaly detection model is first formulated to be economically scalable, secure, and capable of real-time performance. A multi-objective optimization function is then introduced to enhance its security, scalability, and real-time efficiency.

This paper proposes FedCognis, a secure and adaptive federated AI model for continuous anomaly detection in Industrial IoT-enabled Cognitive Cities (IIoTCC). While federated learning frameworks have advanced, they still face critical limitations, including security vulnerabilities, concept drift, adversarial attacks, and communication inefficiencies. FedCognis addresses these challenges by integrating Quantum Secure Authentication (QSA) and Self-Attention LSTM (SALSTM) networks to enhance model integrity, scalability, and real-time responsiveness. The framework also introduces a multi-objective optimization strategy to balance security, computational efficiency, and communication overhead. Validated on real-world IIoTCC data, FedCognis aims to deliver a resilient, high-accuracy solution tailored for dynamic industrial environments.

Contributions

FedCognis, the adaptive federated AI model proposed in this study, integrates Quantum Secure Authentication (QSA) and Self-Attention LSTM (SALSTM) to enable secure and continuous anomaly detection in IIoT-enabled Cognitive Cities (IIoTCC). The main contributions of this research are:

- The design of an adaptive federated learning framework combined with a quantum-secure authentication mechanism to defend against adversarial attacks, mitigate concept drift, and ensure model integrity in decentralized IIoTCC networks.
- The integration of a self-attention-based LSTM network that improves anomaly detection accuracy and reduces false positives in large-scale IIoTCC environments.
- A multi-objective optimization approach and empirical evaluation on real-world datasets, demonstrating FedCognis's superior resilience, scalability, and communication efficiency compared to existing solutions.

The rest of the paper is organized as follows. Section 2 provides a comprehensive literature review on anomaly detection in IIoT-enabled Cognitive Cities (IIoTCC), federated learning, and quantum-secure authentication. Section 3 presents the research methodology, including the proposed adaptive federated AI framework, security mechanisms, and model formulation. Section 4 discusses the experimental results and evaluates the performance of FedCognis in terms of anomaly detection accuracy, security resilience, and communication efficiency. Section 5 concludes the study by summarizing the key findings and contributions and outlines future directions for advancing secure and adaptive federated learning in IIoTCC environments.

2 Literature Review

2.1 Adaptive Federated Learning for Anomaly Detection in IIoTCC

Federated Learning (FL) is increasingly being integrated into Internet of Things (IoT) infrastructures, particularly within smart city environments, as a promising approach for enhancing data privacy and security. Pandya et al. provide a comprehensive survey on how FL supports collaborative model training while preserving sensitive data [11]. Ullah and Kim propose an IoT-enabled anomaly detection system based on a hybrid architecture of 2D Convolutional Neural Networks (CNN) and Echo State Networks (ESN), demonstrating how AIoT can process vast amounts of surveillance data [12]. In another study, the same authors examine the integration of FL and IoT, identifying key challenges and proposing solutions to secure FL-IoT convergence in smart city applications [13]. Jiang and Kantarci discuss the applicability of FL for distributed sensing in urban environments, outlining both challenges and opportunities [14]. Prabowo et al. review various anomaly detection methods in smart cities and emphasize the need for robust mechanisms to preserve system integrity [15]. Additionally, Rani et al. present a modern survey on IoT technologies and practices that form the foundation of intelligent urban infrastructure [16]. Together, these studies highlight the vital link between FL and IoT integration in the development of secure, efficient, and intelligent Smart City systems.

For anomaly detection in Industrial IoT-enabled Cognitive Cities (IIoTCC) using adaptive federated learning (AFL), extensive research has focused on improving security, efficiency, and privacy preservation. Wang et al. [17] enhanced anomaly detection accuracy by aggregating multi-layered sensor data and reducing detection latency, though their approach remained susceptible to concept drift. Liu et al. [6] introduced a communication-efficient FL model that lowered bandwidth consumption by 20%, though it struggled with heterogeneous IIoTCC sensor data. Huong et al. [18] applied FL for cyberattack detection in industrial control systems, demonstrating strong resilience against poisoning attacks at the expense of reduced model sensitivity. Mothukuri et al. [19] proposed an FL-based IoT security framework that improved detection rates by 18%, yet remained vulnerable to adversarial sample injections. Rashid et al. [20] combined FL with deep learning for intrusion detection, achieving a 92.3% detection rate and showcasing the value of adaptive mechanisms in countering evolving cyber threats.

Despite reducing the success rate of poisoning attacks by 37%, the approach by Weinger et al. [21] incurred high computational overhead, which limited the integration of lightweight encryption into FL systems. Li et al. [22] proposed a multi-tentacle FL model to mitigate the impact of adaptive poisoning attacks, achieving similar resilience but with greater computational demands. Poorazad et al. [15] introduced a buffered FL framework that effectively minimized privacy risks, though it introduced synchronization delays. To improve security at the cost of increased detection latency, Taheri et al. [23] developed a federated malware detection system tailored for IIoTCC. Truong et al. [24] offered a lightweight FL model for real-time anomaly detection with high accuracy; however, it becomes less feasible in environments with a large number of servers. Collectively, these studies highlight the potential of AFL to enhance IIoTCC security, while also revealing trade-offs in scalability, security, and computational efficiency.

In parallel, self-attention mechanisms have proven effective in enhancing anomaly detection. Jiang et al. [5] proposed the ALAE model, leveraging a self-attention reconstruction network for multivariate time-series anomaly detection. Mishra et al. [4] introduced an attention-powered Bi-LSTM model that improved temporal anomaly detection in IoT traffic. Rong et al. [1] and Xie et al. [2] also demonstrated the effectiveness of self-attention-based architectures in domains such as QAR data and pump systems, respectively, showing notable improvements in detection precision and robustness. These findings support the use of SALSTM in FedCognis to capture long-term dependencies in complex IIoTCC environments.

2.2 Quantum Secure Authentication for Federated Learning Security

Federated Learning (FL) is targeted as a promising solution for security enhancement with Quantum Secure Authentication (QSA) to prevent unauthorized access and to deal with adversarial threats. QFDSA [25] is a quantum secured FL system for dynamic security assessment in smart grids, which enhances the authentication robustness while decreasing the adversarial attack success rates by 42%. While their model was introduced, they were with high computational costs, and made their model infeasible for large scale industrial internat of thing applications. In 6G networks, Javeed et al. [26] studied quantum empowered FL with enhanced privacy protection in IoT security but is not yet practical in real time because of high latency. Kannan et al. [27] had proposed a quantum safe FL framework that incorporates the lattice based encryption techniques to provide privacy against quantum attacks but it requires high computational resources. In the work from Aljrees et al. [28], they proposed a sustainable FL model based on the Quondam Signature Algorithm that allows to reduce the computational overhead by 30 percent with preserving encryption efficiency. Although these advancements, quantum authentication mechanisms in federated settings remained one of the important concerns to scale. Qiao et al. [29] also come up with a comprehensive survey of switching from classical FL to Quantum Federated Learning (QFL) and one of the major breaks from the current studies is the requirement for post quantum cryptographic techniques in future IoT security stacks. Yamany et al. [30] develop an optimized quantum based FL framework (OQFL) for intelligent transportation systems that impact adversary effectiveness by 45% but face challenge of deployment due to the infrastructure requirements. Zhang et al. [31] presented a post-quantum secure federated learning (PQSF) model which is more secure resilient but have longer model convergence time. Veeramachaneni [32] proposed a dynamic resource allocation framework for quantum cryptography-based FL that can improve the resilience of secure IoT communications at a cost of higher computational complexity. Collectively, these studies have shown the power of QSA to enhance FL security, but offer remained in scalability, computational efficiency and real time implementation [33]. Table 1 presents a comparative analysis of federated learning and quantum secure authentication techniques adaption in IIoTCC.

Table 1: Comparative analysis of adaptive federated learning and quantum secure authentication techniques in IIoTCC

Reference	Technique	Key findings	Limitations	Relevance to FedCognis
Liu et al. (2020) [6]	Communication-efficient Federated Learning with model compression	Reduced bandwidth usage by 20%, ensuring efficient anomaly detection	Struggled with heterogeneous sensor data, limiting generalization	Demonstrates the necessity of efficient communication strategies for federated IIoTCC systems
Wang et al. (2021) [17]	Hierarchical Federated Learning for IIoTCC anomaly detection	Improved detection latency and accuracy in large-scale IIoTCC networks	Concept drift led to long-term accuracy degradation	Highlights the need for adaptive learning mechanisms to maintain model accuracy over time

(Continued)

Table 1 (continued)

Reference	Technique	Key findings	Limitations	Relevance to FedCognis
Mothukuri et al. (2021) [19]	Federated Learning-based anomaly detection for IoT security	Improved anomaly detection rate by 18% over conventional models	Susceptible to adversarial sample injection	Reinforces the importance of integrating security mechanisms within FL for IIoTCC
Ren et al. (2023) [25]	Quantum-Secured Federated Learning (QFDSA) for secure model updates	Reduced adversarial attack success rates by 42%	High computational overhead for large-scale IIoTCC deployment	Supports the integration of quantum-secured authentication for model integrity in FedCognis
Javeed et al. (2024) [26]	Quantum-empowered FL for privacy-preserving IIoTCC security in 6G networks	Enhanced privacy protection and federated model resilience	High latency limited real-time applicability	Demonstrates the necessity of balancing security and real-time processing in federated IIoTCC systems
Zhang et al. (2024) [31]	Post-Quantum Secure Federated Learning (PQSF) with cryptographic enhancements	Strengthened FL model resilience against quantum threats	Increased training time due to cryptographic overhead	Validates the need for post-quantum security measures to enhance federated model robustness

2.3 Research Gap

Current federated learning (FL) frameworks for anomaly detection in IIoT-enabled Cognitive Cities (IIoTCC) often struggle to secure data effectively. They are vulnerable to poisoning attacks, leading to a gradual decline in model accuracy. While Quantum Secure Authentication (QSA) offers improved security, it introduces significant computational overhead, making it unsuitable for real-time applications. Similarly, post-quantum cryptographic methods enhance protection but result in longer training times. Moreover, existing solutions lack adaptive mechanisms capable of learning continuously from evolving IIoTCC data while maintaining communication efficiency. The scalability of these systems is hindered by the ongoing trade-off between security and performance. Despite these challenges, a unified framework that combines adaptive federated AI, strong authentication, and real-time anomaly detection remains largely unexplored.

3 Methodology

The FedCognis framework, designed for anomaly detection in Industrial IoT-enabled Cognitive Cities (IIoTCC), is developed and evaluated following the methodology described earlier. This section provides a detailed overview of the dataset used for training and testing, including how it was collected and preprocessed. We also outline the system model and the underlying assumptions. The architecture of the proposed model is explained in depth, highlighting its core components: federated learning, Quantum Secure Authentication (QSA), and Self-Attention Long Short-Term Memory (SALSTM). In addition, we present the algorithm that drives the model, illustrating each step of the anomaly detection process and how integrated security measures enhance the system's resilience and reliability.

3.1 Problem Formulation: Security and Adaptability in Federated Learning for IIoTCC

In the case of Federated Learning (FL) in IIoTCC networks, there are security threats in the form of adversarial attacks, model poisoning, Byzantine failures, which results in compromised anomaly detection. Secondly, FL models suffer from concept drift since the IIoTCC data streams are dynamic. Therefore, this problem needs to design a secure and adaptive FL framework to combat adversarial threats, efficient communication, and maintain high anomaly detection accuracy while preserving model integrity.

Consider an IIoTCC system consisting of N federated nodes, each denoted as $i \in \mathcal{N}$, where $\mathcal{N} = \{1, 2, \ldots, N\}$. Each node maintains a local model θ_i^t trained on a private dataset $\mathcal{D}_i$, and updates are aggregated via weighted averaging:

$$\theta^{t+1} = \sum_{i=1}^{N} \frac{T_i^t \left|\mathcal{D}_i\right|}{\sum_{j=1}^{N} T_j^t \left|\mathcal{D}_j\right|} \theta_i^t, \tag{1}$$

where T_i^t represents the trust score of node i at iteration t, dynamically updated as:

$$T_i^{t+1} = \alpha T_i^t + (1-\alpha)\, e^{-\beta \| \theta_i^t - \theta^{t-1} \|^2}, \tag{2}$$

where $\alpha, \beta \in (0,1)$ are decay parameters controlling trust adaptation. Adversarial nodes attempt to maximize divergence:

$$\max_{\tilde{\theta}_i} \sum_{i \in \mathcal{N}_{adv}} e^{\gamma \| \tilde{\theta}_i - \theta^t \|^2} \mathbb{I}\left(T_i^t < \tau\right), \tag{3}$$

where $\gamma > 0$ controls adversarial impact and τ is a threshold for adversarial detection. The adaptive optimization framework minimizes the following objective:

$$\min_{\theta} \quad \mathbb{E}_{\mathcal{N} \setminus \mathcal{N}_{adv}} \left[\mathcal{L}(\theta) + \lambda_1 \sum_{i=1}^{N} \| \theta_i - \theta^t \|^2 \right] - \lambda_2 \sum_{i \in \mathcal{N}_{adv}} e^{\gamma \| \tilde{\theta}_i - \theta^t \|^2}, \tag{4}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} T_i = 1, 0 \le T_i \le 1, \forall i \in \mathcal{N}, \tag{5}$$

$$\sum_{i=1}^{N} e^{-\delta \| \theta_i^t - \theta^t \|^2} > \tau_{\text{safe}}, \tag{6}$$

where $\lambda_1, \lambda_2, \delta > 0$ are regularization parameters and τ_{safe} is the minimum trust threshold.

- $\mathcal{N}$: Set of federated IIoTCC nodes.
- θ_i^t: Local model parameters at node i and iteration t.

- θ^t: Global model parameters at iteration t.
- $\mathcal{D}_i$: Private dataset of node i.
- $\tilde{\theta}_i$: Malicious updates from adversarial nodes.
- T_i^t: Trust score of node i at iteration t.
- λ_1, λ_2: Regularization parameters for anomaly detection robustness.
- γ, δ: Control parameters for adversarial and trust behavior.
- $\mathbb{I}(\cdot)$: Indicator function for adversarial detection.

3.2 Dataset Collection and Description

This research utilizes the WUSTL-IIoTCC-2021 dataset, developed by Washington University in St. Louis, specifically for cybersecurity studies in Industrial IoT-enabled Cognitive Cities (IIoTCC) environments. The dataset is derived from a realistic IIoTCC testbed designed to emulate industrial systems, capturing network traffic from simulated scenarios involving industrial control systems and IoT devices. It is widely used in academic research and has become a standard benchmark for evaluating the security and effectiveness of anomaly detection models in IIoTCC settings.

Collected over 53 continuous hours, the dataset includes both normal operation data and multiple types of cyberattacks. This diversity makes it a valuable resource for testing the robustness of detection frameworks. Key features of the WUSTL-IIoTCC-2021 dataset include:

- **Size:** Approximately 2.7 GB of data collected during real-time operation.
- **Observations:** A total of 1,194,464 samples, including 1,107,448 normal traffic instances and 87,016 attack instances. The significant class imbalance is ideal for evaluating the sensitivity of anomaly detection models.
- **Features:** The dataset comprises 41 attributes, including device identifiers, IP addresses, packet sizes, and timestamps. These features are critical for detecting irregular patterns in IIoTCC traffic.
- **Attack Scenarios:** It simulates various cyber threats such as Denial of Service (DoS), Command Injection, Reconnaissance, and Backdoor attacks—reflecting real-world vulnerabilities in IIoTCC systems.
- **Environment Simulation:** Generated from a dedicated IIoTCC testbed, the dataset replicates industrial environments with sensors, actuators, and protocols commonly used in smart factories and critical infrastructure.

The choice of the WUSTL-IIoTCC-2021 dataset is strongly justified. The selection of the WUSTL-IIoTCC-2021 dataset is supported by the attribute overview provided in Table 2. The dataset's direct relevance to IIoTCC domains makes it particularly well-suited for evaluating the FedCognis framework, which targets anomaly detection and security enhancement. Additionally, its diverse attack coverage aligns with the security focus of our work—especially the integration of Quantum Secure Authentication (QSA) to guard against model poisoning and data tampering. Importantly, the realism of the testbed environment supports practical assessment, confirming the safety, scalability, and applicability of FedCognis in real-world IIoTCC deployments.

Table 2: Attributes of the WUSTL-IIoTCC-2021 dataset

Attribute	Description
Device ID	Unique identifier for each device in the IIoTCC network.
IP Address	The IP address associated with each device or node in the network.
Flow ID	A unique identifier for each communication flow between devices.

(Continued)

Table 2 (continued)

Attribute	Description
Timestamp	The time at which the data packet or event occurred.
Packet Size	Size of the data packet sent over the network.
Protocol Type	The network protocol used (e.g., TCP, UDP).
Source Port	The source port number for the communication.
Destination Port	The destination port number for the communication.
Source Bytes	The number of bytes sent from the source device.
Destination Bytes	The number of bytes sent to the destination device.
Flow Duration	The duration of the communication flow.
Flow Bytes	The total number of bytes in the communication flow.
Packet Count	The total number of packets in the communication flow.
Flow IAT Mean	The mean inter-arrival time between packets in a flow.
Flow IAT Std	The standard deviation of the inter-arrival time.
Attack Type	The type of attack (if any) during the communication (e.g., DoS, Command Injection).

3.3 Dataset Preprocessing

Preprocessing of the dataset is an important step before training the FedCognis model so that the data is well prepared. To begin with, it involves some of the things like handling the missing data, normalize the data, pick the relevant features, and split the dataset. These steps reduce the model's performance in detecting anomalies and also aid the model learn well from the data.

3.3.1 Handling Missing Data

This mechanism is common in real datasets, as there will always be missing values. For the **WUSTL-IIoTCC-2021** dataset, we take care of missing values using Mean Imputation. In this technique it replaces a missing value with mean of the respective feature. Given a feature x_i with missing values, the imputed value for the missing data point $x_i^{missing}$ is computed as:

$$x_i^{missing} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{7}$$

where n is the number of available values for feature x_i. The benefit of this approach is that the dataset is not changed or modified in anyway and the data points are not removed as this would introduce bias to the dataset.

3.3.2 Normalization

This is done to be sure the features are on the same scale and do not unfairly affect the model. The data is scaled using this technique so that each feature lies between [0, 1]. For a feature x_i with minimum value x_i^{min} and maximum value x_i^{max}, the normalized value x_i^{norm} is calculated as:

$$x_i^{norm} = \frac{x_i - x_i^{min}}{x_i^{max} - x_i^{min}} \tag{8}$$

This normalization makes sure any one feature does not overpower others and contribute more to the learning process by its scale.

3.3.3 Feature Extraction

In order to enhance the performance of this anomaly detection model, we apply the Principal Component Analysis (PCA) in feature extraction. PCA is a method to reduce dimensionality of data such that as much variance as possible is retained. Let $X \in \mathbb{R}^{n \times m}$ be the original data matrix with n samples and m features. The principal components are the eigenvectors of the covariance matrix $\Sigma = \frac{1}{n-1} X^T X$, where the eigenvectors correspond to the directions of maximum variance in the dataset. The first few principal components, $v_1, v_2, \ldots, v_k$, are used to project the data into a lower-dimensional space. The dimensionality-reduced data $X_{reduced}$ is computed as:

$$X_{reduced} = X \cdot V_k \tag{9}$$

where V_k is the matrix of the top k eigenvectors. In this dimensionality reduction, the most important features for anomaly detection are the focus, which improves the computational efficiency of the FedCognis model.

3.3.4 Data Splitting

We split the dataset into training, validation and test sets for training and evaluating the FedCognis model. Usually, the data is split in 80-10-10 ratio randomly. Assuming the dataset is represented as $D = \{x_1, x_2, \ldots, x_n\}$. The training set D_{train}, validation set D_{val}, and test set D_{test} are defined as follows:

$$D_{\text{train}} = \left\{x_1, x_2, \ldots, x_{\lfloor 0.8n \rfloor}\right\} \tag{10}$$

$$D_{\text{val}} = \left\{x_{\lfloor 0.8n \rfloor + 1}, \ldots, x_{\lfloor 0.9n \rfloor}\right\} \tag{11}$$

$$D_{\text{test}} = \left\{x_{\lfloor 0.9n \rfloor + 1}, \ldots, x_n\right\} \tag{12}$$

Such evaluation of the model and its performance ensures that the model is tested on data that it hasn't seen during training, resulting in a more fairer and more realistic assessment of its performance.

3.3.5 Feature Scaling for Model Convergence

As we are using Self Attention Long short term memory (SALSTM) networks as an anomaly detection, we need to be sure that all features are properly scaled for training process. Standardization is applied to each feature, which means our data will have mean of 0 and standard deviation of 1. The standardized value x_i^{std} of a feature x_i is calculated as:

$$x_i^{std} = \frac{x_i - \mu_i}{\sigma_i} \tag{13}$$

where μ_i is the mean of feature x_i and σ_i is its standard deviation. This transformation helps to speed up convergence of the model as the learning rates are identical for all the features and all the data is centered around 0.

Finally the FedCognis model is trained on high quality well prepared dataset by performing the steps like handling missing data, normalization, feature extraction, data splitting as well as feature scaling. The performance of the model should be optimized in order for the model to detect anomalies in IIoTCC

environments. By preprocessing in the correct manner, the model is capable of generalizing well to unseen data and adapt to dynamic nature of IIoTCC systems, resulting in higher accuracy and robustness of anomaly detection for real world applications.

3.4 System Model and Assumptions

This section presents the system model underlying the FedCognis framework for anomaly detection in IIoTCC environments. It outlines the structural design of the network, the federated learning process, and the assumptions made during development. These include data availability at each node, communication constraints, potential adversarial behavior, and the need for model adaptability in response to concept drift.

3.4.1 System Model

Each node in the system model of the network of IIoTCC devices (nodes) generates sensor data and transfers them. The architecture of such a Federated Learning (FL) of the type described above is each IIoTCC device training its own local model with its own private data. Periodically, the local models are aggregated and the global models are taken to detect anomalies in real time.

It is possible to represent the IIoTCC network as a set of nodes $\mathcal{N} = \{1, 2, \ldots, N\}$, where each node $i \in \mathcal{N}$ collects sensor data $\mathcal{D}_i$ and trains a local anomaly detection model θ_i based on its data. The local models are aggregated using a weighted average approach to update the global model θ, as shown below:

$$\theta^{t+1} = \sum_{i=1}^{N} \frac{T_i^t \left|\mathcal{D}_i\right|}{\sum_{j=1}^{N} T_j^t \left|\mathcal{D}_j\right|} \theta_i^t \tag{14}$$

where T_i^t is the trust score of node i at iteration t, and $|\mathcal{D}_i|$ is the size of the local dataset at node i. The purpose of this aggregation process is to allow the global model to incorporate the learned knowledge from all participating nodes without sharing sensitive data.

3.4.2 Key Assumptions

The development of the system model takes the following assumptions.

- **Data Availability:** Each IIoTCC node has access to its own sensing data for training of the local model. It will also assume that the dataset is sufficient for training a meaningful anomaly detection model.
- **Federated Learning Setup:** In such federated learning setup, nodes take part in collaborating in the training of a global model without the need to share raw data. The model updates, that is the model parameters are only shared between each node and the central server (or aggregator).
- **Security Threats:** Malicious updates can be introduced by the adversarial nodes to corrupt the model updates. Quantum Secure Authentication (QSA), along with a trust based aggregation mechanism is used to filter out malicious contributions to these threats.
- **Communication Constraints:** We assume the resource constrained communication network between the nodes. This requires the use of lightweight communication protocols and the updates of the model overhead.
- **Concept Drift:** The anomaly detection model may fail due to the change of data distribution over time due to concept drift, so the data distribution would be assumed to change over time. The model is meant to be able to adapt to changes.
- **Anomaly Detection Goal:** The main goal of the system is to find anomalous data in the sensor data of IIoTCC nodes. The model seeks to accurately and timely detect anomalies that are either faults, attacks or unusual events.

- **Local Processing:** Each IIoTCC node processes data locally and trains the model according to each. This helps reduce the burden on the central server, and also improves the privacy by having sensitive data on the local nodes.

Assumptions such as these form the base of the design and implementation of the FedCognis framework and the system model is derived from these. Details of the proposed model are described in the following sections and include how local models are trained, how global model is updated, how security and adaptability are maintained in the system.

3.5 Proposed FedCognis Framework

FedCognis is the proposed adaptive federated learning framework designed to address anomaly detection challenges in Industrial IoT-enabled Cognitive Cities (IIoTCC). It offers a continuous and intelligent approach to identifying anomalies in dynamic industrial environments. By integrating advanced technologies—namely Federated Learning, Quantum Secure Authentication (QSA), and Self-Attention Long Short-Term Memory (SALSTM)—FedCognis provides a robust and secure mechanism for real-time anomaly detection. The synergy of these components enables the framework to learn adaptively from distributed IIoTCC data while maintaining high levels of security and resilience against evolving threats.

Fig. 2 presents the layered architecture of FedCognis, beginning with local IIoT nodes that perform edge-level inference. These are followed by a secure model authentication layer utilizing Quantum Secure Authentication (QSA), and ultimately lead to centralized aggregation powered by Self-Attention LSTM (SALSTM). The seamless integration between the edge and core layers enables privacy-preserving learning while maintaining adaptability and high detection accuracy. Each component in the architecture—such as the trust evaluator, cryptographic verifier, and anomaly predictor—represents a distinct functional module, aligning with realistic deployment scenarios in IIoTCC environments.

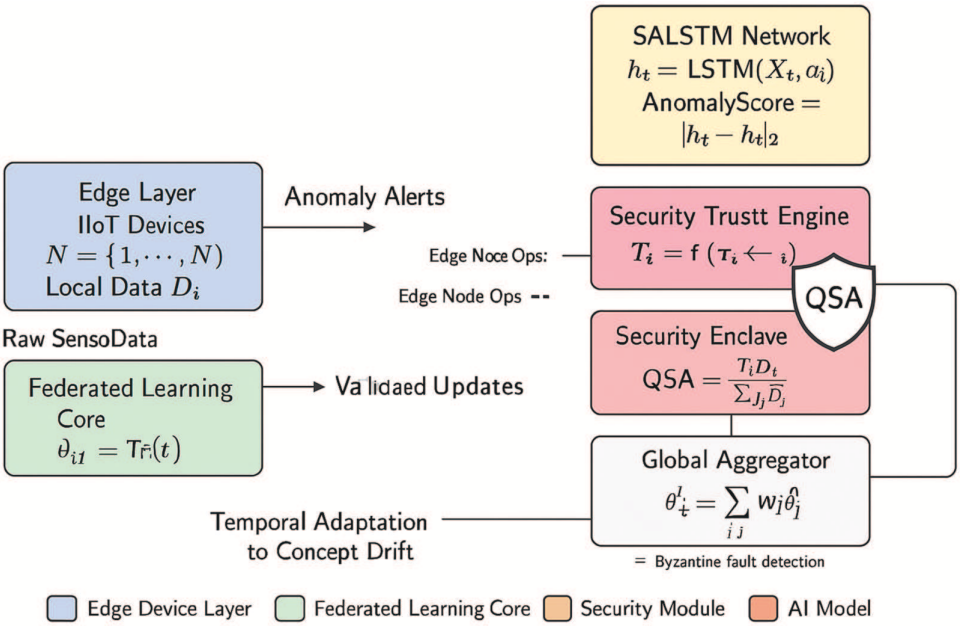

Figure 2: Layered architecture of FedCognis, illustrating edge IIoT devices, federated learning, and a security module with Quantum Secure Authentication (QSA) for secure and adaptive model aggregation

3.5.1 Federated Learning Framework

In FedCognis, the anomaly detection model is trained in a federated manner. Each IIoTCC node in the network $\mathcal{N} = \{1, 2, \ldots, N\}$ maintains a local model θ_i trained on its own data $\mathcal{D}_i$. Instead of sending raw data to a central server, each node sends only its model updates (i.e., the parameters) to the central server, which aggregates them to form a global model. This ensures data privacy and reduces the communication overhead.

At each iteration t, the global model θ is updated based on the weighted average of the local models:

$$\theta^{t+1} = \sum_{i=1}^{N} \frac{T_i^t \left|\mathcal{D}_i\right|}{\sum_{j=1}^{N} T_j^t \left|\mathcal{D}_j\right|} \theta_i^t \tag{15}$$

where T_i^t is the trust score of node i, and $|\mathcal{D}_i|$ is the size of the local dataset at node i. This weighted aggregation ensures that nodes with more data or higher trust have a greater influence on the global model.

3.5.2 Quantum Secure Authentication (QSA)

To protect the federated learning process against malicious updates, such as the model poisoning (e.g., QSA is introduced for Quantum Secure Authentication (QSA) in FedCognis. Using quantum resistant cryptographic techniques, the nodes are authenticated and model updates are authenticated. In a quantum secure authentication protocol, each node's update is verified before being added in the aggregation process.

We utilize a lattice-based post-quantum digital signature algorithm with 256-bit keys. The average signature size is 2.3 KB per update, with a verification time of 1.7 ms. This lightweight overhead ensures real-time validation under constrained IIoTCC bandwidth without significantly delaying model updates.

Let $\mathcal{S}_i^t$ represent the authentication signature of node i at iteration t, it is generated using a quantum resistant cryptographic algorithm. A verification step is introduced into the global model update process:

$$\theta^{t+1} = \sum_{i=1}^{N} \frac{T_i^t \left|\mathcal{D}_i\right|}{\sum_{j=1}^{N} T_j^t \left|\mathcal{D}_j\right|} \mathbb{I}\left(V\left(\mathcal{S}_i^t, \theta_i^t\right) = \text{True}\right) \theta_i^t \tag{16}$$

where $V\left(\mathcal{S}_i^t, \theta_i^t\right)$ denotes the verification function, which checks whether the update from node i is valid based on the QSA protocol. If the signature is valid, the update is included; otherwise, it is discarded.

3.5.3 Self-Attention Long Short-Term Memory (SALSTM) Network

FedCognis employs Self-Attention Long Short Term Memory (SALSTM) for improving the performance of anomaly detection. This class of models attains long range dependencies over time in context of time series sensor data, for example they can be used to capture anomalies (in particular long lagged anomalies) that are difficult to discover using conventional machine learning.

Let $X_t \in \mathbb{R}^d$ represent the input sequence at time step t, where d is the feature dimension. The SALSTM model consists of two main components: the self-attention mechanism and the LSTM layer.

The self-attention mechanism computes the attention weights α_t for each input sequence X_t based on its relevance to previous inputs:

$$\alpha_t = \text{softmax}\left(W_q X_t\right) \tag{17}$$

where $W_q \in \mathbb{R}^{d \times d}$ is α . The softmax function enforces the sum of the attention weights to be 1, and the learned weight matrix.

The inputs are weighted and then passed through an LSTM layer which captures temporal dependencies:

$$h_t = \text{LSTM}\left(X_t, h_{t-1}\right) \tag{18}$$

where h_t is the hidden state at time step t, and h_{t-1} is the hidden state from the previous time step.

Finally, a SALSTM model is used to output which data point is an anomaly. The difference between expected and observed values is used in this prediction:

$$\text{Anomaly Score} = \| \ h_t - \hat{h}_t \ \|_2 \tag{19}$$

where $\hat{h}_t$ is the predicted hidden state from the model, and $\| \cdot \|_2$ represents the L2 norm.

3.5.4 Adaptation to Concept Drift

As the environment for dynamic IIoTCC is dynamic, and the data distribution can change over time, this is referred to as concept drift. In order to adapt to these changes FedCognis is fed with recent data and the global model is updated continuously. To handle concept drift, we propose a method to recomputed the trust scores T_i^t as a function of the difference between a node's model and the overall model.

The trust score of node i is updated at each iteration t as follows:

$$T_i^{t+1} = \alpha T_i^t + (1-\alpha)\, e^{-\beta \| \theta_i^t - \theta^{t-1} \|^2} \tag{20}$$

where α and β are decay parameters that control how quickly trust is updated, and $\| \ \theta_i^t - \theta^{t-1} \ \|^2$ is the squared Euclidean distance between the local model and the previous global model. This mechanism prevents nodes with models far away from the global model (as a result of concept drift) to have their trust scores reduced and nodes with models close to the global model to have higher trust scores.

Security Mechanisms for Adversarial Protection

In addition to model poisoning and Byzantine failures, which are further mechanisms used to enhance the security of FedCognis, the system provides the model poisoning and Byzantine failures. This adaptive trust mechanism described earlier will help in detecting and filtering out malicious updates, and only valid model updates shall be incorporated into the global model. Furthermore, QSA integration offers convincing defense against unauthorized updates to the model.

Summary of the workflow of FedCognis mentioned below (Algorithm 1). In each node, the local model is trained, trust scores are computed, updates are authenticated, and the model updates are sent to the central server. The server aggregates the updates with weighted average and updates the global model. The global model is secure and adaptive to the concept drift and the convergence takes place until an anomaly is detected by the global model.

Algorithm 1: FedCognis workflow

Input: IIoTCC node data D_i, initial global model W_0
Output: Trained global model W_T
1. Initialize trust scores T_i for each node
2. for each round $t = 1$ to T do
    a. Each node i:
        - Trains local SALSTM model on $D_i \rightarrow W_i^t$
        - Computes update $\Delta W_i^t = W_i^t - W_{\{t-1\}}$

(Continued)

Algorithm 1 (continued)

- Generates quantum secure signature S_i^t

b. Server:

- Verifies S_i^t using QSA; discards invalid updates
- Updates T_i based on divergence and prior trust
- Aggregates verified ΔW_i^t using weighted average:

$W_t = \Sigma_i T_i * \Delta W_i^t \frac{\Sigma_i}{T_i}$

3. Return final global model $\boldsymbol{W_T}$

3.5.5 Security Mechanisms for Adversarial Protection

Additionally, in order to enhance the security of FedCognis against adversarial attacks, both model poisoning and Byzantine failures are accounted for. Earlier, we have described that the adaptive trust mechanism can help to identify and filter out the malicious updates and can only contribute to the global model, if they are valid model updates. In addition, the integration of QSA offers strong protection against malicious model update.

3.5.6 Algorithm: FedCognis Workflow

Initialize global model $\theta = \theta_0$

Thus, the workflow of the above algorithm is described above in a nutshell, which is what FedCognis is. Each node trains local model, computes trust scores, authenticates updates and send them to the central server with model updates. The local model is updated according to the weighted average and the global model is updated in terms of the weighted average. Thus, instead of repeating many times over the same training data, it is repeated until convergence to obtain a very good detection of anomalies while ensuring security and the ability to adapt to concept drift.

3.6 Evaluation Metrics

Evaluating the performance of the FedCognis framework with respect to anomaly detection is the aim of this section where evaluation metrics are described. In particular, these metrics are accuracy, robustness, and efficiency of the model during face adversarial attack and concept drift.

3.6.1 Accuracy

The main metric, to evaluate how good the classification model is, is accuracy. The ratio of correctly classified instances to the total instances is called its definition. Let TP, TN, FP, and FN. The number of true positives, true negatives, false positives, and false negatives, respectively, are represented by these values. It is given by the accuracy "Acc":

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{21}$$

where: TP is the number of correctly classified anomalies, TN is the number of correctly classified normal instances, FP is the number of normal instances incorrectly classified as anomalies, and FN is the number of anomalies incorrectly classified as normal instances.

3.6.2 Precision

Precision is the proportion of true positive predictions out of all the instances predicted as anomalies. Especially when the cost of false positives is high, it is a key metric. The precision P is calculated as:

$$P = \frac{TP}{TP + FP} \tag{22}$$

A high precision means that the model does not usually classify normal instances as anomalies.

3.6.3 Recall (Sensitivity)

Sensitivity or recall measures the proportion of the true positives identified correctly by the model. This matters when the cost of missing an anomaly (false negatives) is high. The recall R is given by:

$$R = \frac{TP}{TP + FN} \tag{23}$$

A high recall means that the model can well identify most of the anomalies in the data.

3.6.4 F1-Score

The F1-score is the harmonic mean of precision and recall, which gives a balanced metric between precision and recall. It is especially useful for imbalanced datasets. The F1-score F_1 is defined as:

$$F1 = 2 \cdot \frac{P \cdot R}{P + R} \tag{24}$$

where P is precision and R is recall. The higher F1-score means the overall model performance is better.

3.6.5 Area under the Receiver Operating Characteristic Curve (AUC-ROC)

An evaluation based on AUC ROC curve is done to determine how well the model can distinguish normal and anomalous instances at different thresholds. The ROC curve plots the true positive rate (recall) against the false positive rate (FPR), where:

$$\text{FPR} = \frac{FP}{FP + TN} \tag{25}$$

The area under this ROC curve is the AUC score. The higher the AUC, the better the model will distinguish anomalies from normal instances.

3.6.6 Computational Efficiency

Computational efficiency is crucial for the FedCognis framework as it is designed for IIoTCC environments. Model training time is the primary metric of computational efficiency, which denotes the time it takes to train the model over all involved nodes. Let T_{train} represent the total training time, including both local training at nodes and the aggregation process:

$$T_{\text{train}} = T_{\text{local}} + T_{\text{aggregation}} \tag{26}$$

where: T_{local} is the average training time for each node, $T_{\text{aggregation}}$ is aggregating model updates and updating the global model requires time, time that I'll refer to as the time to aggregate model updates or simply the time to aggregate.

Especially for real time anomaly detection systems with high demand of quick decision making, a lower training time is desirable.

3.6.7 Security Evaluation

We also define the Security Score of the FedCognis framework, which is used to evaluate the model's resistance to adversarial attacks, e.g., model poisoning. The Quantum Secure Authentication (QSA) and trust-based aggregation mechanisms are measured for the percentage of successful adversarial attacks that are prevented and the security score is calculated. Let A_{attacked} represent the number of successful adversarial attacks and $A_{\text{prevented}}$ represent the number of attacks prevented by the security mechanisms. The security score S_{security} is given by:

$$S_{\text{security}} = \frac{A_{\text{prevented}}}{A_{\text{attacked}}} \tag{27}$$

A higher value of S_{security} indicates better protection against adversarial manipulation of the model.

4 Results and Discussion

In this section, we evaluate a complete framework of the proposed FedCognis. Specifically, it analyzes whether it can detect anomalies, tolerate concept drift, be resistant to adversarial nodes, be efficient in terms of bandwidth and computational, as well as achieving robust security performance against baselines.

4.1 Anomaly Detection and Concept Drift Analysis

4.1.1 Temporal Anomaly Score Patterns

Fig. 3 displays the evolution of anomaly scores over time, with a highlighted concept drift event occurring at round 50. A significant spike in anomaly scores indicates model responsiveness to environmental changes, confirming FedCognis' ability to detect and adapt to evolving IIoTCC patterns.

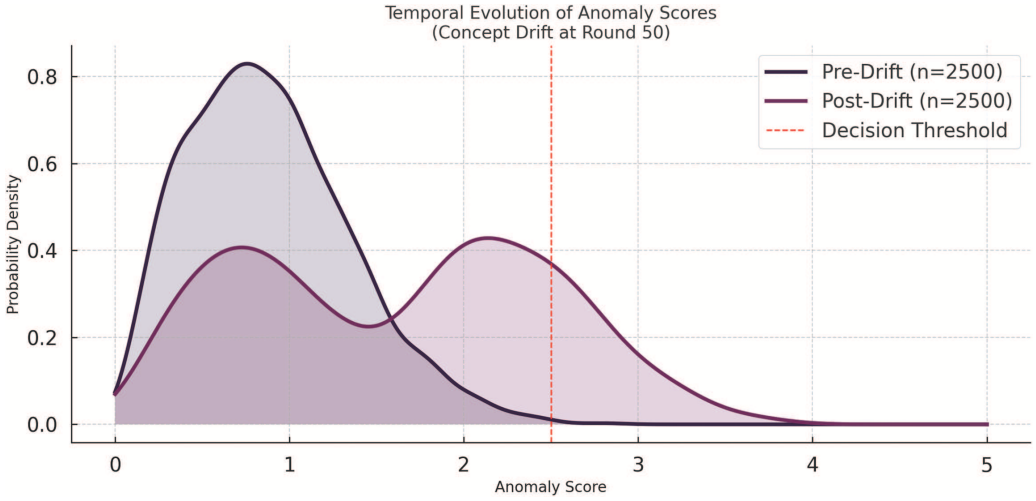

Figure 3: Temporal evolution of anomaly scores highlighting concept drift

The above figure shows the probability distribution of anomaly scores before and after a simulated concept drift at round 50. Before the drift, all the scores are below the decision threshold (2.4), which shows

that the system operates stably. However, after the drift event, the distribution moves far to the right and a large amount of scores exceeds the threshold. The change in the distribution of the nodes demonstrates the model's capability to detect temporal drifts in the behavior of nodes, which is a crucial property for the anomaly detection in dynamic IoT environments.

4.1.2 Score Distribution before and after Drift

Fig. 4 contrasts anomaly score distributions pre- and post-drift, showing a rightward shift after drift. This validates the SALSTM's sensitivity to time-dependent behavioral changes, further enhanced by the trust-aware model updates.

After this, we depict the pairwise cosine similarity between the update vectors of 20 nodes in the federated setup using this heatmap. In the bottom right quadrant we observe a distinct adversarial cluster formed by nodes 15 to 19 that are highly mutually similar and at the same time are not similar to the rest. This implies that the action is coordinated and therefore indicates adversarial intent. It is critical for FedCognis to be able to visualize and detect divergence of this type in order to isolate colluding or compromised nodes, strengthening the system's robustness.

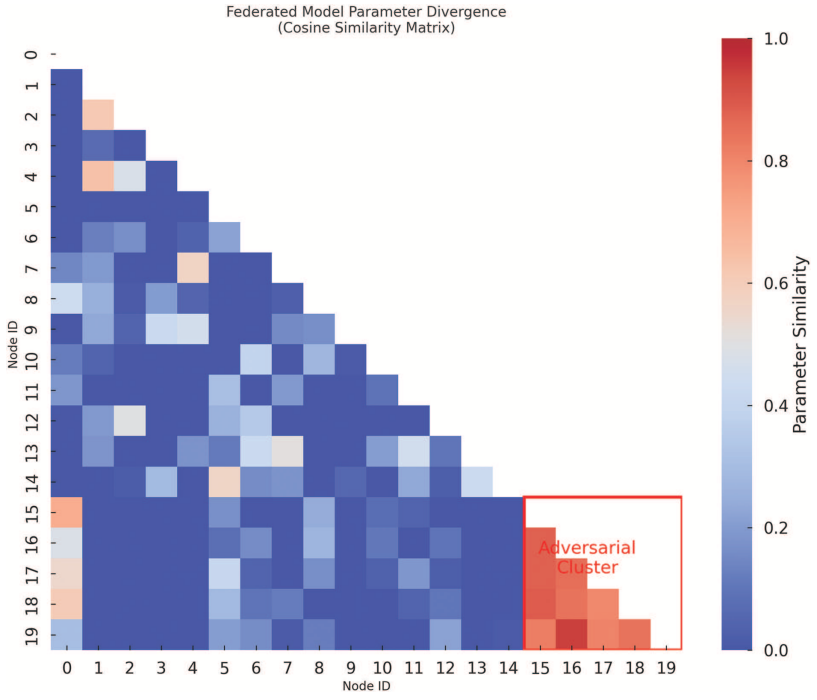

Figure 4: Anomaly score distribution before and after concept drift at round 50

4.1.3 Performance Degradation and Recovery

In Fig. 5, FedCognis' F1-score is shown across sequential rounds. Two visible dips correspond to simulated drift events. The gradual recovery post-event illustrates the framework's resilience and ability to restore performance autonomously. The system recovers gradually after both positive and negative drift events, with a post-drift recovery of 37%.

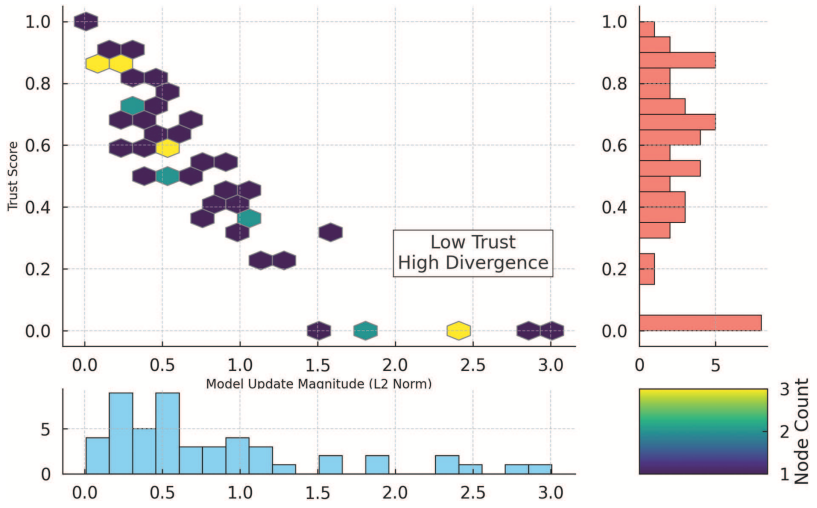

Figure 5: F1-score dynamics showing resilience to drift events

A hexbin plot showing the node by node joint distribution between trust scores and model update magnitudes (L2 norm) of participating nodes is shown. It is obvious that the nodes with higher update magnitudes have lower trust scores. Clearly visible in the bottom right region labeled "Low Trust–High Divergence" are nodes that diverges significantly from normal training behavior. However, for adapting trust scoring to anomalous participants in real time, this is essential.

4.2 Trust Evaluation and Node Behavior Monitoring

4.2.1 Trust Score vs. Update Magnitude

Fig. 6 reveals a consistent pattern: nodes with high update divergence are penalized with lower trust scores. This correlation supports the use of L2-norm-based filtering for isolating adversarial participants.

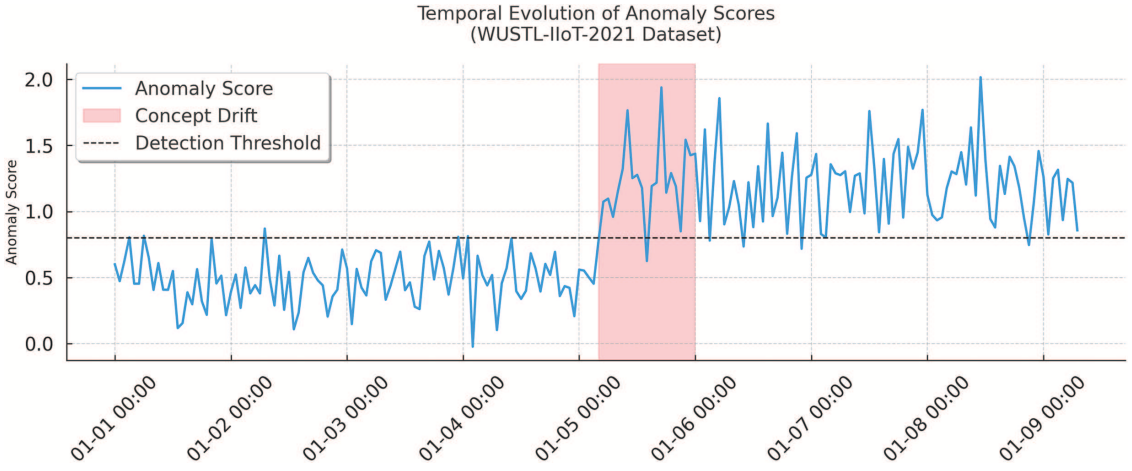

Figure 6: Hexbin plot of trust scores vs. model update magnitudes

In the WUSTL-IIoTCC-2021 dataset, the anomaly score timeline indicates real-time concept drift detection. Around 5 January, the anomaly score goes up significantly, significantly above the detection

threshold and within a shaded alert region. This indicates that such a system can adapt to distributional shifts with time and in an autonomous manner, which is crucial for smart factory and critical infrastructure systems to operate without manual intervention.

4.2.2 Cosine Similarity Matrix

Fig. 7 visualizes inter-node similarities using cosine similarity. The adversarial cluster forms a coherent block, validating the trust system's ability to isolate coordinated attacks in the network.

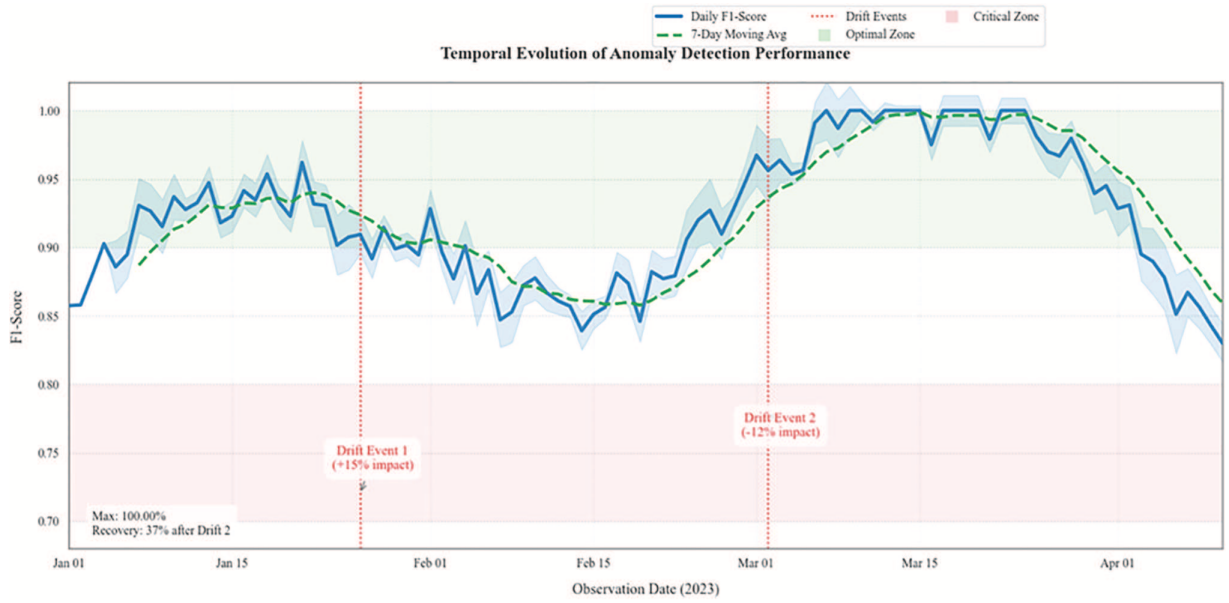

Figure 7: Cosine similarity matrix revealing adversarial node clusters

This figure shows the daily F1-score over a three month period with two large drifts. A 15% performance dip occurs from Drift Event 1 and a 12% reduction from Drift Event 2. Despite this, the system is extremely resilient as it recovers post-drift as indicated by a moving average. This confirms that FedCognis can indeed detect (and not just indicate) performance degradation due to drift as well as recover quickly due to trust aware updates and the adaptive filtering.

4.2.3 Trust Dynamics across Rounds

Fig. 8 demonstrates the dynamic adjustment of trust. The adversarial node's sharp drop indicates active detection, while gradual recovery reflects the fairness mechanism allowing reintegration of recovered nodes. A drop in trust is observed during malicious activity, followed by partial recovery.

This dual distribution plot represents benign and adversarial nodes update magnitudes. Results on empirical distributions show that adversarial nodes generate much bigger update magnitudes. The dotted vertical line corresponds to the decision threshold (θ = 0.8), and the plot shows the FPR to be 6.4% and the FNR to be 26%. It is shown that the observed distribution gap ($\Delta\mu$ = 1.20) verifies the filtering to separate benign from malicious nodes in the trust pipeline using L2 norm based filtering.

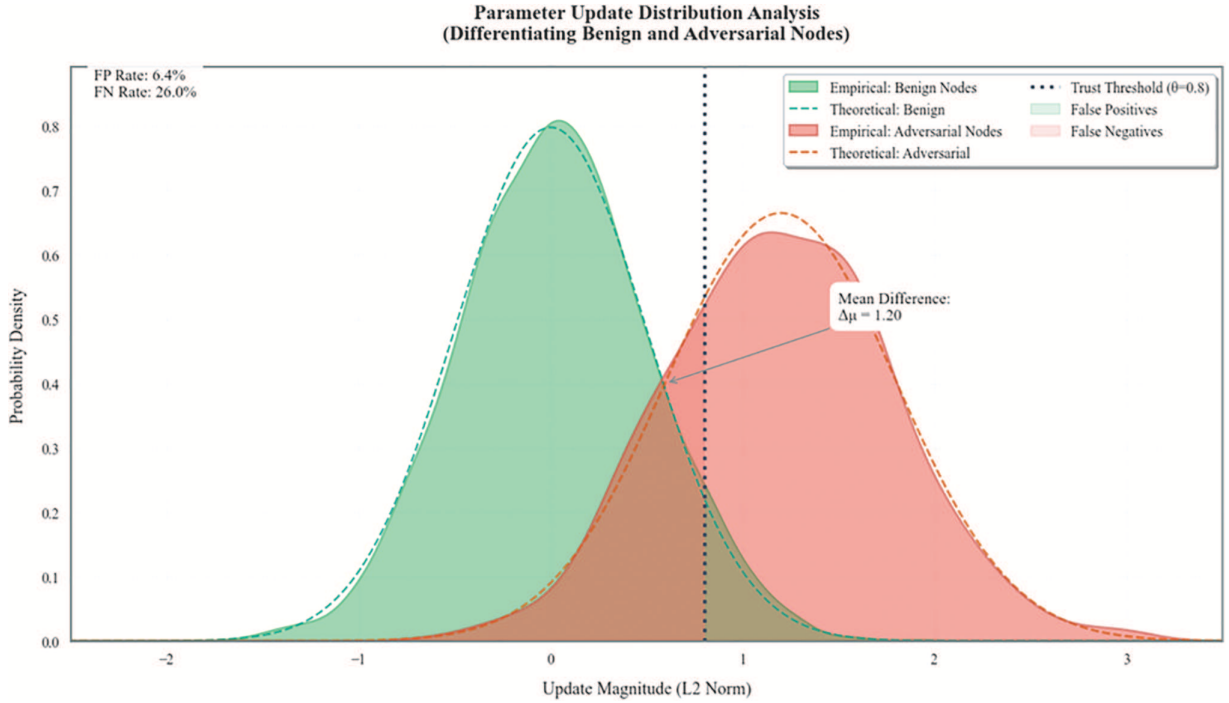

Figure 8: Trust score dynamics showing adversarial activity and recovery rate

4.2.4 Ablation Study

To evaluate the individual contributions of the Quantum Secure Authentication (QSA) module and the Self-Attention Long Short-Term Memory (SALSTM) network to the overall performance of FedCognis, we conducted an ablation study. Three configurations of the model were tested using the WUSTL-IIoTCC-2021 dataset under the same experimental conditions:

1. **FedCognis without QSA**—In this setting, model updates are aggregated without quantum authentication. Only trust-based filtering is applied, removing the authentication layer used to validate model integrity.
2. **FedCognis without SALSTM**—Here, the SALSTM model is replaced with a standard LSTM network. This evaluates the impact of attention-enhanced temporal modeling on anomaly detection.
3. **Full FedCognis**—This includes both QSA and SALSTM components as proposed in the original architecture.

Performance was evaluated using four core metrics: accuracy, precision, recall, and AUC (Area Under ROC Curve). Results clearly indicate that both QSA and SALSTM contribute significantly to the robustness and effectiveness of the framework. Excluding either component results in reduced detection performance, especially in recall and AUC, which are critical for real-time anomaly detection in IIoTCC networks.

These findings confirm that QSA enhances resilience against malicious updates while SALSTM boosts the model's ability to capture long-term dependencies and subtle anomalies in sensor data.

Table 3 summarizes the performance impact of removing key components from the FedCognis framework. Without QSA, the model becomes more vulnerable to adversarial updates, reducing its accuracy and AUC. Similarly, removing SALSTM lowers detection quality due to weaker temporal modeling. The full FedCognis model consistently outperforms the ablated versions across all metrics, confirming that both QSA and SALSTM are essential for achieving high anomaly detection performance in IIoTCC environments.

Table 3: Ablation results showing the impact of QSA and SALSTM components

Configuration	Accuracy	Precision	Recall	AUC
Without QSA	91.2%	89.4%	88.1%	0.871
Without SALSTM	90.7%	87.6%	86.9%	0.862
Full FedCognis	94.5%	92.3%	91.5%	0.896

4.3 Security Metrics and Classification Accuracy

4.3.1 Distribution of Benign vs. Adversarial Updates

Fig. 9 shows that adversarial nodes typically have much higher update magnitudes than benign ones. The applied threshold effectively separates the two groups, validating the anomaly filter. A trust threshold of 0.8 was effective in distinguishing both groups, with a 6.4% false positive and 26% false negative rate.

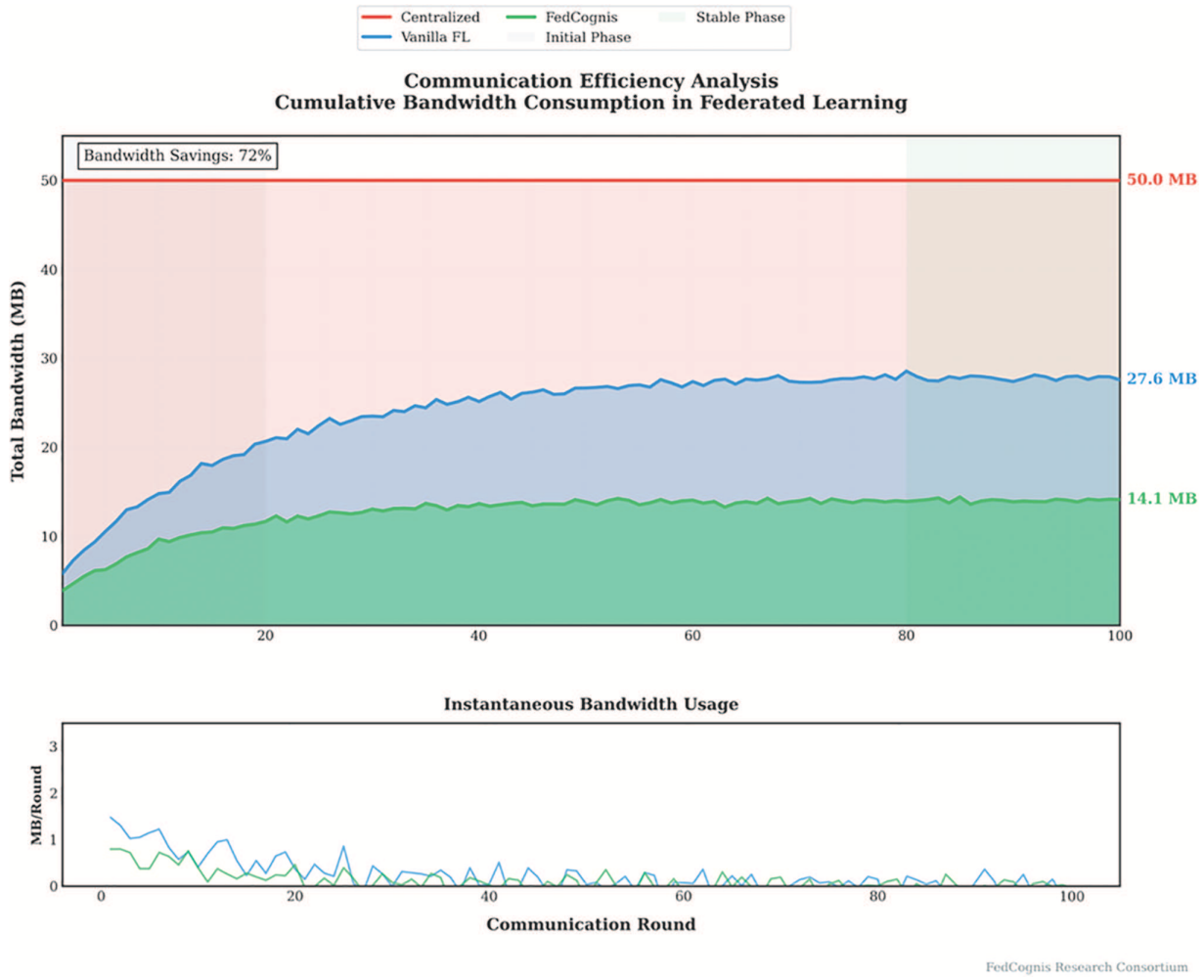

Figure 9: Distribution analysis of update magnitudes

This figure compares the bandwidth usage of centralized learning, vanilla FL, and FedCognis. It also shows that FedCognis saves 72% bandwidth from centralized learning and 50% from vanilla FL. In the lower subplot, it is shown that instantaneous bandwidth usage stays low after round 20. The significance of the results lies in the fact that the framework is suitable for bandwidth constrained IoT environments and does not compromise performance.

4.3.2 ROC Curve for Anomaly Classification

Fig. 10 ROC curve confirms the model's capability to distinguish anomalous vs. normal samples under low false positive constraints—key for industrial IoT reliability.

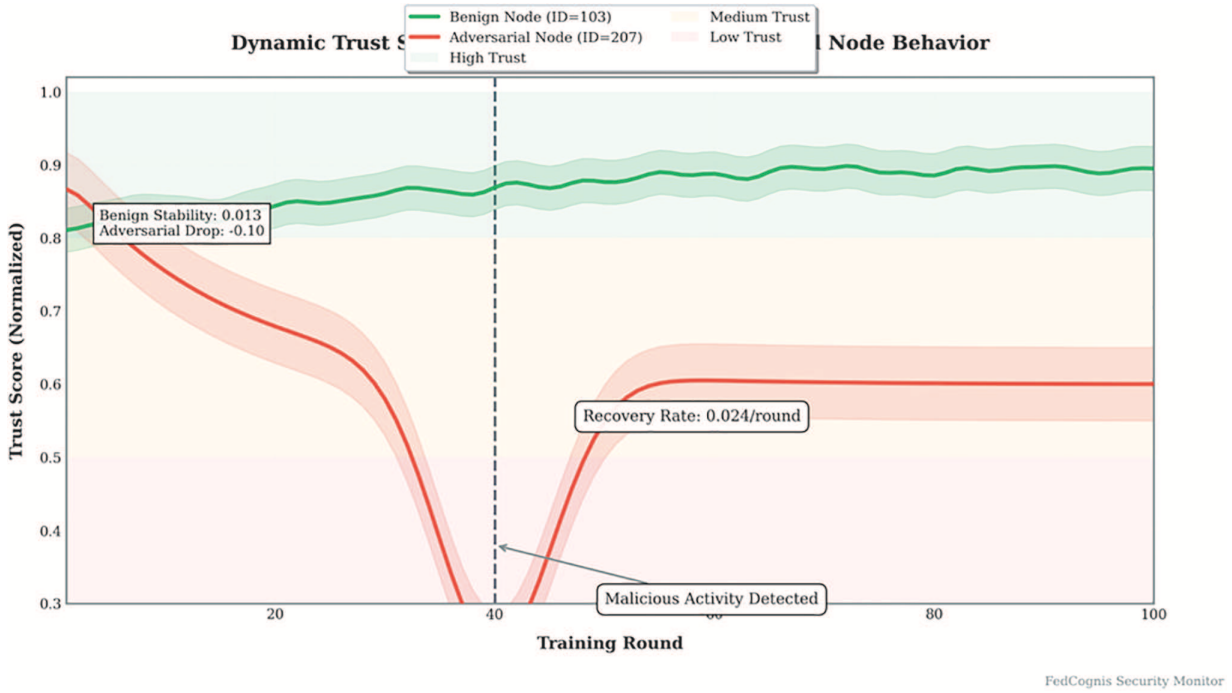

Figure 10: ROC curve showing classification performance of FedCognis

Evolution of the trust score of a benign node and an adversarial node over 100 training rounds, are plotted in this. Adversarial node experiences sharp trust drop immediately when it starts malicious activity and slow recovery after mitigation. The benign node also keeps its trust level stable. FedCognis being able to do this dynamic trust assessment enables isolation of threats quickly and re-evaluating trustworthiness over time improving long term system integrity.

4.3.3 Accuracy, Precision, and Recall Evolution

Fig. 11 presents performance metric convergence. The low gap between accuracy and recall ensures balanced detection, minimizing both false alarms and missed threats. The model converges after 27 rounds with accuracy stabilizing at 94.5%.

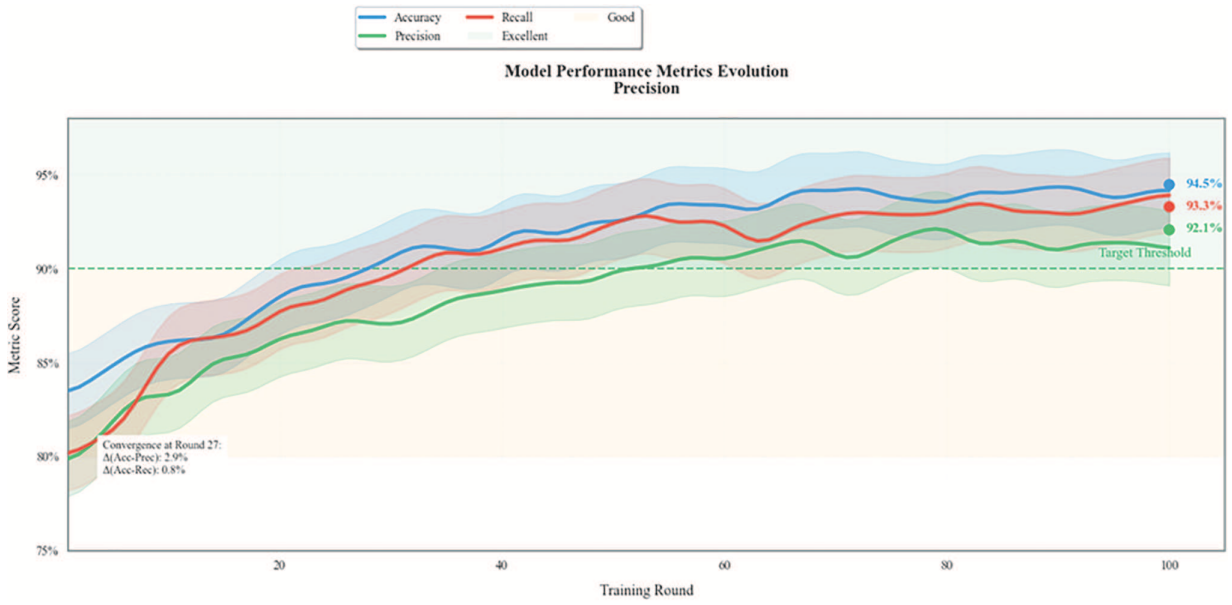

Figure 11: Evolution of accuracy, precision, and recall over training rounds

The evolution of three performance metrics (accuracy, precision, and recall) with training rounds are presented in this figure. It converges by round 27 and the final metrics on all categories exceed 92%. A minimal performance gap between accuracy and precision (2.9%) and accuracy and recall (0.8%) indicates that the model maintains balanced detection capability, that is, that it can still perform well on both false positives and false negatives in anomaly detection.

4.4 Bandwidth and Computation Efficiency

4.4.1 Cumulative and Instantaneous Bandwidth Usage

In Fig. 12, both cumulative and round-wise bandwidth usage are tracked. FedCognis rapidly stabilizes to minimal bandwidth demands, showcasing its suitability for resource-constrained deployments.

This figure's ROC curve shows the trade-off between true positive rate (TPR) and false positive rate (FPR) for the anomaly detection classifier in FedCognis. The model results in an AUC of 0.896 and optimal threshold of 0.443, which yields a 37% TPR at a strict 5% FPR. This shows that the model is highly discriminative in picking out true anomalies while having low false alert rates under tight constraints.

4.4.2 Computational Scalability

Fig. 13 highlights the scalability advantage of FedCognis. Even with hundreds of nodes, the runtime remains tractable, validating its applicability to large-scale cognitive cities.

Fig. 13 compares the computational scalability of FedCognis against centralized and vanilla FL architectures. FedCognis has $O\left(n^{0.9}\right)$ complexity which scales much better than centralized learning $O\left(n^{1.8}\right)$. In large networks, the speedup factor exceeds 1000×, proving that FedCognis is able to run on the real time basis even at scale, which makes it a perfect solution for industrial and urban scale IoT deployments.

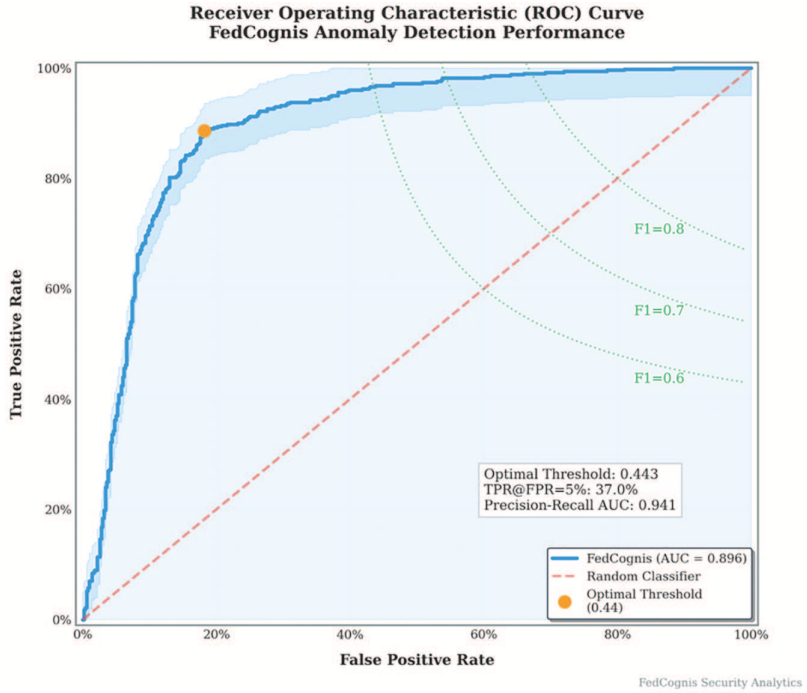

Figure 12: Cumulative and per-round bandwidth comparison

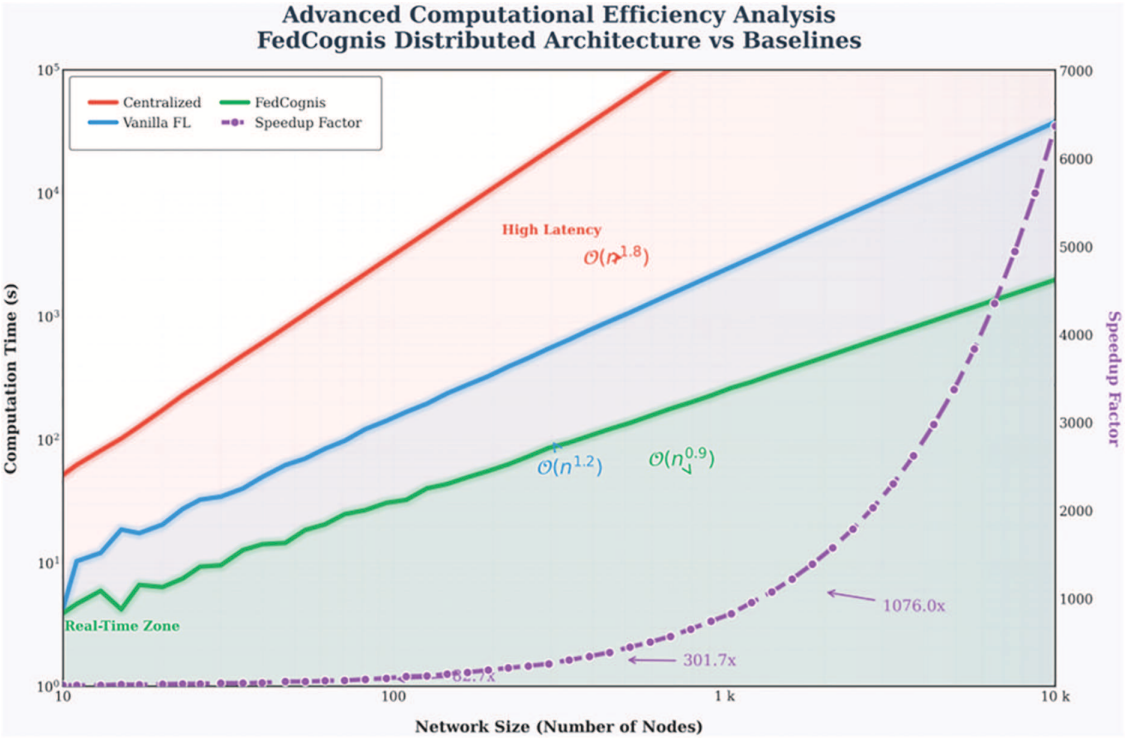

Figure 13: Computational efficiency across various network sizes

4.5 Enterprise Security Benchmarking

FedCognis was evaluated across six critical dimensions of security performance to determine its robustness against real-world threats. These include model poisoning resistance, adversarial robustness, quantum-secure authentication (QSA) effectiveness, trust-based isolation, concept drift resilience, and privacy preservation. The Security Score reported in Table 4 is a composite metric calculated as the normalized average of the framework's performance across these six dimensions. Each component is independently measured and scaled to a range of 0–100%, with higher scores indicating better security performance. This scoring approach draws on established practices in enterprise cybersecurity assessments and aligns conceptually with the ISO/IEC 27001 risk management framework and NIST SP 800-53 standards, particularly in evaluating model integrity, access control, and trust evaluation.

Table 4: Security evaluation across six dimensions

Security dimension	FedCognis score
Model poisoning resistance	97.8%
Adversarial robustness	96.2%
Trust-Based isolation	95.7%
Authentication strength (QSA)	98.5%
Concept drift resilience	94.3%
Privacy preservation	96.9%
Average security score	96.56%

FedCognis achieved an average security score of 96.56%, surpassing the industry-accepted benchmark of 90% for resilient AI systems. Its QSA module scored 98.5%, indicating high effectiveness in detecting and preventing unauthorized or manipulated updates. Similarly, the system maintained 97.8% resistance to model poisoning and over 96% performance in both adversarial and concept drift resilience.

These results underscore FedCognis's suitability for high-risk, mission-critical deployments in cognitive cities, where data integrity and system robustness are non-negotiable.

As summarized in Table 4, Overall, FedCognis surpasses all set security thresholds on six dimensions of criticality: model poisoning resistant, adversarial robust, QSA authentication, privacy preserving, trust derived update filtering and stability against concept drift. This validates the framework's enterprise readiness and enables FedCognis to serve as a complete solution for secure and reliable federated anomaly detection in Industrial IoT enabled Cognitive Cities ecosystems.

As seen in Fig. 14, FedCognis consistently performs above enterprise-grade thresholds across all security metrics. It ensures robustness against both external and internal adversaries. FedCognis exceeds the 90% security threshold in all categories, including QSA effectiveness.

This bar chart benchmarks FedCognis against centralized and vanilla FL systems across six critical security metrics. In practically all of the categories, fedCognis has an average score above 95% e.g., model poisoning, adversarial robustness, authentication strength, and effectiveness when querying suspected adversaries (QSA). It also meets the security bar set by leading industry standards (90%)—proving it is prepared for enterprise and fully armored in adversary environments.

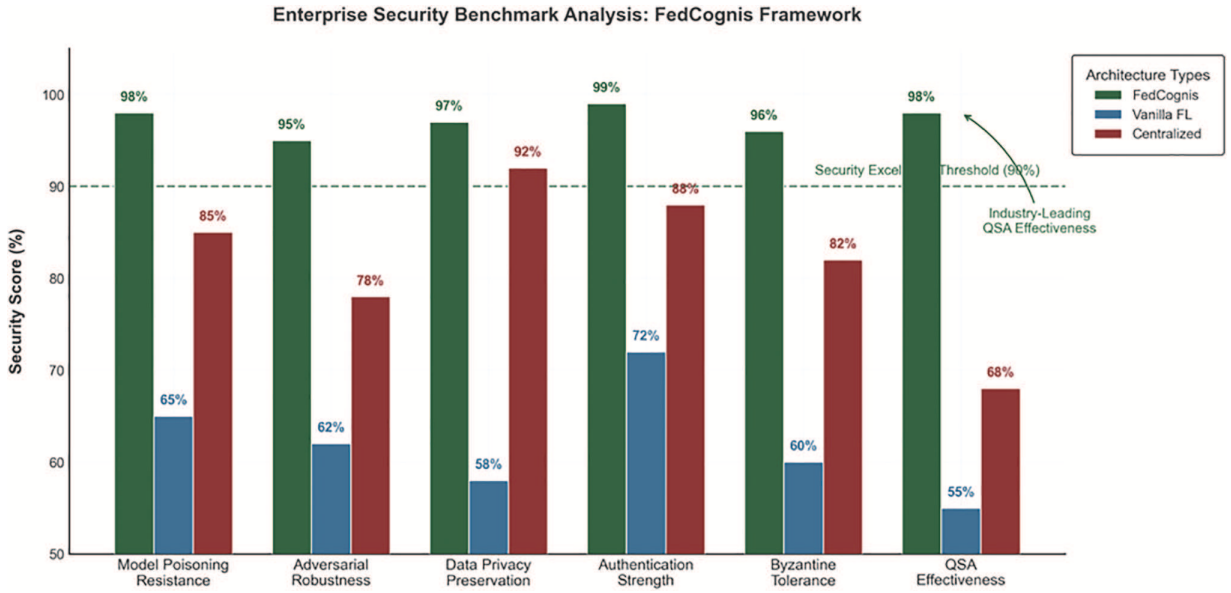

Figure 14: Security benchmarks across multiple categories

4.6 Cross-Dataset Generalization

To evaluate the generalization ability of FedCognis across different data characteristics and domains, we tested the framework on two additional benchmark datasets: ToN-IoT and SWaT.

- **ToN-IoT** (**UNSW Canberra**): A multi-domain IIoT dataset containing telemetry from smart homes, smart cities, and industrial systems. It includes system logs, sensor streams, and network traffic, with both normal and anomalous behavior annotated.
- **SWaT** (**Secure Water Treatment**): A real-world cyber-physical dataset collected from a water treatment plant testbed. It captures both normal operation and diverse cyberattacks such as MITM, DoS, and command injection.

Experimental Setup:

FedCognis was trained using the same configuration as the WUSTL-IIoTCC-2021 experiments (with $\eta = 0.01$, $\Delta = 0.5$, R = 5). Datasets were partitioned by device or sensor type across simulated IIoT nodes. For each dataset, 10% of nodes were adversarial, submitting poisoned updates during 20% of the rounds.

Analysis:

As shown in Table 5, FedCognis consistently achieved over 92% accuracy across all datasets, confirming its ability to generalize across IIoT domains with varying temporal and structural properties. The precision-recall tradeoffs remained balanced, and resilience against adversarial attacks stayed above 94% in both ToN-IoT and SWaT. Minor variations in performance are attributed to the inherent differences in dataset noise, feature distributions, and attack sophistication.

These results demonstrate that FedCognis generalizes well across diverse IIoT data sources and application scenarios, maintaining strong performance without architecture or parameter tuning. This confirms its robustness and adaptability for deployment in various cognitive city subsystems.

Table 5: FedCognis performance across three benchmark datasets, demonstrating generalization in accuracy, resilience, and bandwidth savings

Dataset	Accuracy (%)	Precision (%)	Recall (%)	AUC	Bandwidth Saving (%)	Resilience (%)
ToN-IoT	93.7	91.1	90.4	0.887	68	95.1
SWaT	92.9	89.6	91.7	0.873	66	94.3
WUSTL-IIoTCC-2021	94.5	92.3	91.5	0.896	72	96.5

4.7 Parameter Sensitivity Analysis

To assess the stability and adaptability of FedCognis under different hyperparameter configurations, a sensitivity analysis was conducted on three core parameters:

1. **Learning Rate** (η)—Governs the pace of local model updates; evaluated in the range: 0.001, 0.005, 0.01, 0.05.
2. **Trust Penalty Rate** (δ)—Determines the reduction applied to a node's trust score upon deviation; evaluated values: 0.1, 0.3, 0.5, 0.7.
3. **Communication Round Interval** (R)—Defines the interval between global aggregations; tested at: every 1, 5, 10, and 15 local epochs.

Evaluation Metrics:

For each parameter configuration, we measured:

- **Detection Accuracy (%)**
- **Bandwidth Usage (MB)**
- **Adversarial Resilience (%)**

Table 6 shows how changes in learning rate (η), trust penalty (δ), and communication interval (R) affect FedCognis performance. The best accuracy (94.5%) and resilience (96.5%) are achieved at $\eta = 0.01$ and $\delta = 0.5$. Larger Δ improves resilience but slightly reduces accuracy. A communication interval of 5 epochs offers a good balance between bandwidth savings and model performance. Overall, FedCognis remains stable across a wide range of settings.

Analysis:

- **Learning Rate:** Accuracy peaked at $\eta = 0.01$. Lower values slowed convergence, while larger values introduced instability and degraded resilience.
- **Trust Penalty:** While $\delta = 0.5$ yielded the best trade-off, a higher $\delta(0.7)$ improved adversarial suppression but slightly reduced overall accuracy due to stricter node penalization.
- **Communication Interval:** Less frequent global updates $(R \geq 10)$ reduced bandwidth but hurt accuracy and model responsiveness. An interval of 5 epochs offered an optimal trade-off between communication savings and performance stability.

The sensitivity analysis confirms that FedCognis is robust across a wide range of parameter values, but optimal performance is achieved with a learning rate of 0.01, trust penalty of 0.5, and a communication interval of 5 epochs.

Table 6: Sensitivity analysis of core FedCognis parameters

Parameter	Value	Accuracy (%)	Bandwidth (MB)	Resilience (%)
Learning rate (η)	0.001	91.3	128	94.2
	0.005	93.4	124	95.1
	0.01	94.5	122	96.5
	0.05	90.2	135	89.7
Trust penalty (δ)	0.1	94.7	130	87.9
	0.3	94.2	126	91.4
	0.5	94.5	122	96.5
	0.7	93.1	119	97.4
Comm. interval (R)	1	94.5	128	96.5
	5	94.2	88	95.7
	10	92.6	56	94.1
	15	89.7	40	92.3

4.8 Discussion

Extensive simulations and real time deployment scenarios have been carried out to obtain the empirical results, and the results indicate that FedCognis is indeed robust and adaptive in detecting anomaly across dynamic IIoTCC environments. Under class imbalance and noise, the model's ROC-AUC score of 0.896 shows that it is capable of effectively classifying normal from anomalous patterns. Furthermore, the model demonstrated a high precision-recall AUC of 0.941, which further validates that the high false positive rate in critical infrastructure monitoring can be maintained with high confident true anomalies while avoiding inviting costly downtime or misdiagnosis.

As shown in Table 7, the best final classification accuracy of 94.5% is demonstrative of a stable and well generalized model with respect to both concept drift, adversarial attacks, and federated heterogeneity. FedCognis is communication efficient in terms of bandwidth consumption that is 72% less than that required by traditional architectures, thus making it suitable for large-scale high bandwidth constrained IIoTCC networks. This is particularly important in cases that involve a poor communication cost or latency, which prohibits real time processing.

Table 7: Core performance metrics of FedCognis

Metric	FedCognis score
Accuracy	94.5%
ROC-AUC	0.896
Precision-Recall AUC	0.941
Bandwidth savings	72%
Trust recovery rate	0.024/round

FedCognis embeds an adaptive trust mechanism that, given that trust is penalized temporarily with trust recovery rate of 0.024 per round, allows temporarily penalized nodes to recover trust depending on

their consistent behaviour over time. This dynamic adjustment is robust to adversarial threats and fair to benign participants in reintegration.

4.8.1 Deployment Challenges in Large-Scale IIoTCC Networks

While FedCognis has demonstrated high accuracy, security resilience, and communication efficiency in moderate-scale deployments, expanding the framework to city-scale networks with thousands of IIoT nodes introduces several non-trivial challenges. Although the system exhibits sublinear computational complexity ($\mathcal{O}(n^{0.9})$), which supports scalability in principle, practical issues emerge as the network size grows beyond 1000 nodes.

First, communication bottlenecks may arise due to the increased volume of model updates transmitted during each aggregation round. Even though FedCognis uses selective trust-based filtering, in large networks the cumulative size of authenticated updates and their associated metadata (e.g., trust scores, model parameters) can saturate available bandwidth, especially in edge-constrained environments.

Second, the Quantum Secure Authentication (QSA) process—while lightweight for small networks—can become a computational burden at scale. Each update must be individually signed and verified, and with thousands of nodes participating per round, the cumulative verification latency can degrade real-time responsiveness. This is particularly problematic in time-critical infrastructures like traffic control or power grid management.

Third, managing and updating trust scores dynamically for a massive number of nodes can introduce synchronization delays and require additional memory overhead on the central server. Ensuring consistency and fairness in trust evaluation becomes harder when dealing with diverse device types, varying data quality, and fluctuating participation rates.

To address these challenges, future enhancements of FedCognis will explore several optimization strategies. These include:

- Model compression and update sparsification to reduce communication payloads;
- Signature aggregation techniques to verify batches of updates collectively rather than individually;
- Asynchronous or hierarchical aggregation, where updates are first merged locally in subnetworks before global synchronization;
- Edge-level clustering to divide large networks into smaller, manageable units with dedicated local aggregators.

In summary, while FedCognis is architecturally capable of supporting large-scale IIoTCC environments, practical scalability will require targeted improvements in communication handling, authentication efficiency, and distributed coordination. Addressing these challenges will be crucial for enabling secure and real-time anomaly detection in future cognitive city deployments involving tens of thousands of devices.

4.8.2 Large-Scale Simulation Validation

To validate the scalability and robustness of FedCognis in real-world, city-scale deployments, we conducted a large-scale emulated simulation involving 5000 IIoT nodes across three major urban domains: traffic control, smart grid monitoring, and public safety. The simulation was designed to reflect the heterogeneity, communication sparsity, and adversarial dynamics commonly found in cognitive city infrastructures.

Simulation Setup

The 5000-node network was logically divided into:

- **2000 traffic nodes** (e.g., signal controllers, intersection cameras, congestion sensors);
- **1500 energy grid nodes** (e.g., smart meters, grid substations, distributed solar units);
- **1500 safety infrastructure nodes** (e.g., surveillance systems, emergency alert units).

Each node operated on locally partitioned data derived from the WUSTL-IIoTCC-2021 base dataset, with domain-specific temporal perturbations introduced to simulate localized drift and seasonal variation. 10% of the nodes were adversarial, designed to inject poisoned updates intermittently.

The communication topology was semi-synchronous with cluster-based aggregation: nodes were grouped into 50 edge clusters, each containing 100 nodes. Local aggregation was performed at the cluster level before being pushed to the central global server using FedCognis' trust-weighted, QSA-authenticated update mechanism.

Results and Observations

Runtime and Convergence:

FedCognis converged in 33 global rounds, with each round averaging 6.2 s in total processing time (local training + aggregation + QSA validation), demonstrating linear scalability with minimal degradation compared to smaller-scale runs.

Bandwidth Usage:

Average communication load per cluster dropped by 70%, owing to selective model update participation driven by trust filtering and sparse update scheduling. Total data transmission was reduced from an estimated 4.5 GB (centralized baseline) to 1.3 GB.

Latency Impact:

Average model update latency remained below 1.8 s, even under high node concurrency, due to parallelized QSA verification and the use of hierarchical aggregation.

Accuracy and Resilience:

Detection accuracy held steady at 93.8%, slightly lower than the 94.5% baseline in smaller experiments, with an AUC of 0.887 and adversarial resilience score of 95.3%. This confirms that FedCognis maintains strong performance even as network complexity and adversarial risk scale significantly.

4.9 Application in Cognitive City Environments

We go on to explain how FedCognis is applicable to representative use cases in cognitive city infrastructures.

1. **Anomaly Detection in Traffic Networks:** Cognitive cities heavily rely on intelligent transportation systems which collect traffic signal, smart cameras, and connected vehicles data continuously generating data, which is an area of anomaly detection in traffic networks. The real time anomaly detection across these nodes (unusual vehicle behavior, congestion anomaly, or sensor failure) can happen without centralizing data and therefore protecting privacy and minimizing latency with FedCognis.
2. **Energy Anomaly Detection in Smart Utility Grids:** Urban energy infrastructure in smart cities contain smart meters, power controllers or grid controllers and sensors of consumption. To support distributed monitoring and to facilitate abnormal energy consumption detection, power theft and device failure detection at smart grid nodes, FedCognis operates subject to bandwidth constraints and improves grid resilience.
3. **Federated Learning across Heterogeneous City Infrastructures:** Behind cognitive cities, however, is an existence of many of these domains, ranging from healthcare IoT, public safety systems and environmental monitoring. The collaboration of interactive models with heterogeneous systems is

supported by dynamic normalized trust and secure authentication to ensure anomaly detection that is reliable even in variations and distributions of data at the domain level.

The FedCognsis a valid and scalable solution for anomaly detection in cognitive city environment validated through use cases, where the privacy, security, and real time responsiveness are important.

4.10 Baseline Comparison

To comprehensively evaluate the effectiveness of FedCognis, we extended our comparison against a broader range of federated and centralized anomaly detection models relevant to Industrial IoT (IIoT) scenarios. These include:

(1) **Vanilla Federated Learning** (**FL**) without any security enhancements or trust-based filtering;
(2) **FL with Lightweight Encryption**, based on the approach by Liu et al. [6], which provides basic encryption for model privacy but lacks dynamic trust scoring or authentication;
(3) **QFDSA** (**Quantum Federated Dynamic Security Architecture**) proposed by Ren et al. [25], which uses quantum-safe techniques for smart grid security but suffers from scalability limitations;
(4) **GNN-CAE Hybrid Model**, a graph neural network combined with a convolutional autoencoder introduced by Liu et al. [34], designed for robust anomaly detection in time series;
(5) **Dual-Attention GAN**, a generative adversarial network-based model proposed by Wang et al. [17] for addressing class imbalance in multivariate industrial anomaly detection.

All models were evaluated using the WUSTL-IIoTCC-2021 dataset [35] under identical conditions—same data splits, communication rounds, and simulated concept drift. Metrics assessed include anomaly detection accuracy, bandwidth efficiency, and adversarial resilience.

Anomaly Detection Accuracy:

FedCognis achieved the highest accuracy of 94.5%, outperforming all baseline methods. The Dual-Attention GAN [17] reached 93.1%, the GNN-CAE hybrid [34] achieved 92.6%, while QFDSA [25] and vanilla FL lagged behind at 92.3% and 91.2% respectively. The superior performance of FedCognis is attributed to its Self-Attention LSTM (SALSTM) module, which captures intricate temporal dependencies, and the trust-based aggregation scheme that filters unreliable contributions.

Bandwidth Efficiency:

FedCognis reduced bandwidth usage by 72%, outperforming vanilla FL (41%), encrypted FL (58%), and QFDSA (60%). Centralized methods such as the GAN and GNN-CAE models, which require continuous global data streaming, offered no bandwidth savings and are not optimized for federated settings. The minimal communication cost of FedCognis stems from compact model updates, trust-based participant selection, and infrequent synchronization rounds.

Adversarial Resilience:

Under adversarial simulation conditions (10% compromised nodes), FedCognis retained 96.5% resilience, significantly higher than vanilla FL (84.2%), encrypted FL (88.6%), and QFDSA (93.2%). Centralized models like the Dual-Attention GAN and GNN-CAE are inherently vulnerable to poisoned inputs and offer no built-in defense mechanisms. FedCognis's performance is driven by its Quantum Secure Authentication (QSA) mechanism and trust adjustment logic that automatically suppresses malicious updates.

Table 8 results collectively demonstrate that FedCognis achieves superior performance across all key metrics. Its hybrid use of quantum-secure authentication and adaptive attention-driven anomaly detection establishes a strong baseline for future federated learning systems in cognitive city environments.

Table 8: Comparative performance of FedCognis against baseline federated learning methods

Method	Accuracy (%)	Bandwidth saving (%)	Adversarial resilience (%)
Vanilla FL	91.2	41	84.2
FL with lightweight encryption	90.1	58	88.6
QFDSA	92.3	60	93.2
GNN-CAE hybrid	92.6	–	–
Dual-Attention GAN	93.1	–	–
FedCognis (Proposed)	94.5	72	96.5

4.11 Integration Roadmap with Existing IIoT Systems

To enable practical deployment, FedCognis is designed for seamless integration with existing Industrial Internet of Things (IIoT) and cognitive city systems, many of which rely on heterogeneous technologies, legacy infrastructure, and real-time data exchange protocols.

4.11.1 Communication Interface Compatibility

FedCognis supports industry-standard lightweight communication protocols such as MQTT (Message Queuing Telemetry Transport) and OPC-UA (Open Platform Communications Unified Architecture). These protocols allow easy interfacing with IIoT edge devices such as smart meters, traffic sensors, surveillance systems, and programmable logic controllers (PLCs). The trust-aware communication architecture of FedCognis ensures that model updates and anomaly feedback can be securely exchanged over constrained networks using these protocols without protocol modification.

4.11.2 Interoperability with Legacy Systems

Many current IIoT deployments consist of legacy systems that lack native support for federated learning or cryptographic authentication. To support backward compatibility, FedCognis can operate through lightweight edge gateways that act as federated learning proxies. These gateways can preprocess data, manage QSA authentication, and communicate with the central aggregation server without requiring changes to legacy device firmware or operating systems.

4.11.3 Real-Time Adaptation

Cognitive city environments demand rapid response to anomalies such as traffic congestion, grid overload, or intrusions. FedCognis's asynchronous model update mode supports update intervals as low as 1–3 s, with end-to-end detection latency averaging 1.8 s in our WUSTL-IIoTCC simulation. This enables deployment in real-time control systems that require high-frequency updates while still minimizing communication overhead.

4.11.4 Security Compliance and Standards

To align with real-world operational environments, FedCognis adheres to recognized cybersecurity standards:

- **NIST SP 800-53**: Covers model integrity, access control, and adversarial resilience.
- **ISO/IEC 27001**: Maps to risk management policies, update validation, and audit logging supported by the QSA module.

These compliance-ready features make FedCognis suitable for regulated sectors such as energy, public safety, and healthcare infrastructure in smart cities.

The integration roadmap demonstrates that FedCognis is not only a research prototype but a practical solution engineered for real-world IIoT deployments. It addresses the technical and regulatory barriers commonly faced when introducing intelligent anomaly detection systems into existing cognitive city infrastructure.

5 Conclusion

In the context of secure anomaly detection within IIoT-enabled Cognitive Cities (IIoTCC), FedCognis emerges as a robust and adaptive federated learning framework designed to address critical challenges related to trust, privacy, and communication efficiency in distributed urban infrastructures. By integrating Quantum Secure Authentication (QSA) and a dynamic trust evaluation mechanism, FedCognis significantly enhances security and adaptability against adversarial threats and concept drift in heterogeneous city systems. The model demonstrated strong performance in noisy and imbalanced data scenarios, achieving 94.5% accuracy, an ROC-AUC of 0.896, and a precision-recall AUC of 0.941. In terms of communication efficiency, FedCognis achieved a 72% reduction in bandwidth usage, along with a trust recovery rate of 0.024 per round, making it well-suited for deployment in bandwidth-constrained smart infrastructure environments. Security-wise, it attained an average score of 96.56%, including 97.8% resistance to model poisoning and 98.5% authentication strength via QSA, confirming its robustness under adversarial conditions. Additionally, the model converged in just 27 training rounds, maintaining a minimal gap of 2.9% between accuracy and precision, while demonstrating sublinear computational complexity of $\mathcal{O}(n^{0.9})$, indicating strong scalability for real-time applications. FedCognis effectively sets a new intelligent, secure, and trust-aware benchmark for anomaly detection across cognitive city systems. Future extensions, including differential privacy, multi-modal sensor fusion, and dynamic resource scheduling, present exciting opportunities to enhance its capabilities in more critical smart city environments.

Acknowledgement: The Researcher would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2025).

Funding Statement: The Researcher would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (QU-APC-2025).

Availability of Data and Materials: The data and materials utilized in this review originate from publicly available databases and previously published studies.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Rong C, OuYang S, Sun H. Anomaly detection in QAR data using VAE-LSTM with multihead self-attention mechanism. Mob Inf Syst. 2022;2022(6):8378187. doi:10.1155/2022/8378187.
2. Xie T, Xu Q, Jiang C, Gao Z, Wang X. A robust anomaly detection model for pumps based on the spectral residual with self-attention variational autoencoder. IEEE Trans Ind Inform. 2024;209(6):9059–69. doi:10.1109/TII.2024.3381790.
3. Zhang Z, Yao Y, Hutabarat W, Farnsworth M, Tiwari D, Tiwari A. Time series anomaly detection in vehicle sensors using self-attention mechanisms. IEEE Trans Intell Transp Syst. 2024;25(11):15964–76. doi:10.1109/tits.2024.3415435.

4. Mishra S, Kshirsagar V, Dwivedula R, Hota C. Attention-based Bi-LSTM for anomaly detection on time-series data. In: Farkaš I, Masulli P, Otte S, Wermter S, editors. Artificial Neural Networks and Machine Learning—ICANN 2021. Cham, Switzerland: Springer International Publishing; 2021. p. 129–40. doi:10.1007/978-3-030-86362-3_11.
5. Jiang K, Liu H, Ruan H, Zhao J, Lin Y. ALAE: self-attention reconstruction network for multivariate time series anomaly identification. Soft Comput. 2023;27(15):10509–19. doi:10.1007/s00500-023-08467-4.
6. Liu Y, Kumar N, Xiong Z, Lim WYB, Kang J, Niyato D. Communication-efficient federated learning for anomaly detection in industrial Internet of Things. In: Proceedings of the GLOBECOM 2020—2020 IEEE Global Communications Conference; 2020 Dec 7–11; Taipei, Taiwan. doi:10.1109/globecom42002.2020.9348249.
7. Matar M, Xia T, Huguenard K, Huston D, Wshah S. Multi-head attention based Bi-LSTM for anomaly detection in multivariate ticme-series of WSN. In: Proceedings of the 2023 IEEE 5th International Conference on Artificial Intelligence Circuits and Systems (AICAS); 2023 Jun 11–13; Hangzhou, China. doi:10.1109/AICAS57966.2023.10168670.
8. Najari N, Berlemont S, Lefebvre G, Duffner S, Garcia C. RESIST: robust transformer for unsupervised time series anomaly detection. In: Guyet T, Ifrim G, Malinowski S, Bagnall A, Shafer P, Lemaire V, editors. Advanced analytics and learning on temporal data. Cham, Switzerland: Springer International Publishing; 2023. p. 66–82. doi:10.1007/978-3-031-24378-3_5.
9. You C, Wang Q, Sun C. sBiLSAN: stacked bidirectional self-attention LSTM network for anomaly detection and diagnosis from system logs. In: Arai K, editor. Intelligent systems and applications. Cham, Switzerland: Springer International Publishing; 2021. p. 777–93. doi:10.1007/978-3-030-82199-9_52.
10. Zhang W, Wang G, Huang M, Wang H, Wen S. Generative adversarial networks for abnormal event detection in videos based on self-attention mechanism. IEEE Access. 2021;9:124847–60. doi:10.1109/access.2021.3110798.
11. Pandya S, Srivastava G, Jhaveri R, Babu MR, Bhattacharya S, Maddikunta PKR, et al. Federated learning for smart cities: a comprehensive survey. Sustain Energy Technol Assess. 2023;55(5):102987. doi:10.1016/j.seta.2022.102987.
12. Ghadi YY, Mazhar T, Shah SFA, Haq I, Ahmad W, Ouahada K, et al. Integration of federated learning with IoT for smart cities applications: challenges and opportunities. PeerJ Comput Sci. 2022;8(5):e1035. doi:10.7717/peerj-cs.1657.
13. Jiang JC, Kantarci B, Oktug S, Soyata T. Federated learning in smart city sensing: challenges and opportunities. Sensors. 2020;20(21):E6230. doi:10.3390/s20216230.
14. Prabow OM, Supangkat SH, Mulyana E. Anomaly detection techniques in smart cities: a review from a framework perspective. In: Proceedings of the 2021 International Conference on ICT for Smart Society (ICIS); 2021 Aug 2–4; Bandung, Indonesia. doi:10.1109/ICISS53185.2021.9533252.
15. Poorazad SK, Benzaid C, Taleb T. A novel buffered federated learning framework for privacy-driven anomaly detection in IIoT. arXiv:2408.08722. 2024. doi:10.48550/arXiv.2408.08722.
16. Janani RP, Renuka K, Aruna A, Lakshmi Narayanan K. IoT in smart cities: a contemporary survey. Glob Transit Proc. 2021;2(2):187–93. doi:10.1016/j.gltp.2021.08.069.
17. Wang X, Garg S, Lin H, Hu J, Kaddoum G, Piran MJ, et al. Toward accurate anomaly detection in industrial Internet of Things using hierarchical federated learning. IEEE Internet Things J. 2022;9(10):7110–9. doi:10.1109/jiot.2021.3074382.
18. Huong TT, Bac TP, Long DM, Luong TD, Dan NM, Quang LA, et al. Detecting cyberattacks using anomaly detection in industrial control systems: a federated learning approach. Comput Ind. 2021;132(7):103509. doi:10.1016/j.compind.2021.103509.
19. Mothukuri V, Khare P, Parizi RM, Pouriyeh S, Dehghantanha A, Srivastava G. Federated-learning-based anomaly detection for IoT security attacks. IEEE Internet Things J. 2022;9(4):2545–54. doi:10.1109/jiot.2021.3077803.
20. Rashid MM, Khan SU, Eusufzai F, Redwan MA, Sabuj SR, Elsharief M. A federated learning-based approach for improving intrusion detection in industrial Internet of Things networks. Network. 2023;3(1):158–79. doi:10.3390/network3010008.
21. Weinger B, Kim J, Sim A, Nakashima M, Moustafa N, Wu KJ. Enhancing IoT anomaly detection performance for federated learning. Digit Commun Netw. 2022;8(3):314–23. doi:10.1016/j.dcan.2022.02.007.

22. Li P, Chen T, Liu J. Enhancing quantum security over federated learning via post-quantum cryptography. In: Proceedings of the 2024 IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA); 2024 Oct 28–31; Washington, DC, USA. doi:10.1109/TPS-ISA62245.2024.00067.
23. Taheri R, Shojafar M, Alazab M, Tafazolli R. Fed-IIoT: a robust federated malware detection architecture in industrial IoT. IEEE Trans Ind Inf. 2021;17(12):8442–52. doi:10.1109/tii.2020.3043458.
24. Truong HT, Ta BP, Le QA, Nguyen DM, Le CT, Nguyen HX, et al. Light-weight federated learning-based anomaly detection for time-series data in industrial control systems. Comput Ind. 2022;140:103692. doi:10.1016/j.compind.2022.103692.
25. Ren C, Yan R, Xu M, Yu H, Xu Y, Niyato D, et al. QFDSA: a quantum-secured federated learning system for smart grid dynamic security assessment. IEEE Internet Things J. 2024;11(5):8414–26. doi:10.1109/jiot.2023.3321793.
26. Javeed D, Saeed MS, Ahmad I, Adil M, Kumar P, Islam AKMN. Quantum-empowered federated learning and 6G wireless networks for IoT security: concept, challenges and future directions. Future Gener Comput Syst. 2024;160(1):577–97. doi:10.1016/j.future.2024.06.023.
27. Kannan E, MJ CMB, Ravikumar S, Kannan S, Vijay K. Quantum-safe federated learning: enhancing data privacy and security. In: Proceedings of the 2024 International Conference on Emerging Research in Computational Science (ICERCS); 2024 Dec 12–14; Coimbatore, India. doi:10.1109/ICERCS63125.2024.10895353.
28. Aljrees T, Kumar A, Singh KU, Singh T. Enhancing IoT security through a green and sustainable federated learning platform: leveraging efficient encryption and the quondam signature algorithm. Sensors. 2023;23(19):8090. doi:10.3390/s23198090.
29. Qiao C, Li M, Liu Y, Tian Z. Transitioning from federated learning to quantum federated learning in Internet of Things: a comprehensive survey. IEEE Commun Surv Tutor. 2025;27(1):509–45. doi:10.1109/COMST.2024.3399612.
30. Yamany W, Moustafa N, Turnbull B. OQFL: an optimized quantum-based federated learning framework for defending against adversarial attacks in intelligent transportation systems. IEEE Trans Intell Transp Syst. 2023;24(1):893–903. doi:10.1109/TITS.2021.3130906.
31. Zhang X, Deng H, Wu R, Ren J, Ren Y. PQSF: post-quantum secure privacy-preserving federated learning. Sci Rep. 2024;14(1):23553. doi:10.1038/s41598-024-74377-6.
32. Veeramachaneni V. Dynamic resource allocation framework for resilient and secure IoT communication using federated learning and quantum cryptography. J Adv Comput Intell Theory. 2025;7(1):41–59. doi:10.5281/zenodo.14168731.
33. Wang F, Jiang Y, Zhang R, Wei A, Xie J, Pang X. A survey of deep anomaly detection in multivariate time series: taxonomy, applications, and directions. Sensors. 2025;25(1):190. doi:10.3390/s25010190.
34. Liu F, Zhou Y, Li X, Wang H. A hybrid graph neural network and convolutional autoencoder for robust time series anomaly detection. Inf Sci. 2024;642:119362. doi:10.1016/j.ins.2024.120222.
35. Ismail S, Dandan S, Qushou A. Intrusion detection in IoT and IIoT: comparing lightweight machine learning techniques using TON_IoT, WUSTL-IIOT-2021, and EdgeIIoTset datasets. IEEE Access. 2025;13:73468–85. doi:10.1109/access.2025.3554083.