# Privacy Integrity-Aware Blockchain Communication in Federated Edge Learning Platform

Original Scientific Paper

**Chitresha Jain***
Pandit Deendayal Energy University,
Computer Science & engineering Department
Raysan, Gandhinagar, India
Chitresha.research@gmail.com

**Payal Chaudhari**
Pandit Deendayal Energy University,
Computer Science & engineering Department
Raysan, Gandhinagar, India
Payal.Chaudhari@sot.pdpu.ac.in

*Corresponding author

***Abstract*** *– The blockchain architecture offers transparent security mechanisms in a decentralized manner; due to this, it has attained increasing growth in a federated edge-server learning environment. In federated learning, the data model is executed in multiple edge servers in a collaborative manner, increasing users' privacy and data-integrity breach because of single point failure attack in the main computational server. Blockchain employing a rewarding mechanism in federated edge-learning platform aids the model to overcome single-point aggregation failure. However, the current method failed to identify selfish and baized workers; further, reaching global consensus model to assure privacy-integrity in blockchain-enabled federated edge-server is difficult. This paper presents privacy-integrity-aware blockchain communication (PIABC) in federated edge-server learning platform. The PIABC model is very effective in comparison with existing blockchain-privacy preserving schemes for identifying the correctly aggregated packets and eliminating malicious packets within the federated edge-server learning platform.*

***Keywords****: Blockchain, Consensus, Federated Learning, Integrity, Secure aggregation, Privacy*

*Received: May 8, 2025; Received in revised form: August 28, 2025; Accepted: August 29, 2025*

## 1. INTRODUCTION

In the past few years, there has been a notable increase in the popularity of Federated-Learning (FL) [1]. Artificial-Intelligence (AI) like Machine-Learning (ML) and Deep Learning (DL) approaches, can be trained immediately on devices used by users as well as at edge of network using FL, which eliminates the need to centralize unprocessed information [2]. As a result, data breaches are less likely to occur, and users' privacy is protected whenever confidential data is stored on their devices. Additionally, when employees collaborate, they can access a wealth of information, which enhances efficiency and makes FL models more adaptable and effective. However, despite these advantages, FL also presents several challenges and limitations [3]. The primary features of FL make it vulnerable to novel attacks, which include (i) system-heterogeneity; (ii) the necessity of a reliable centralized entity to coordinate analysis of locally-trained approaches; (iii) vulnerable to inference attacks and information counterfeiting; (iv) absence of a reward approach for involved nodes; (v) communication-security; along with (vi) regulating issues [3, 4]. Moreover, scholars have started looking into methods that facilitate the utilization of blockchain since the FL method's present implementation lacks the necessary capabilities to deal with such issues [3, 5]. Both the public and private sectors are interested in FL because of its endless possibilities, which arise because of decentralized framework. FL depends on the likelihood of carrying out transactions that are legitimate and verifiable without requiring the participation of an unauthorized third-party, while also assuring the tracking and storage of information securely. Therefore, the integration of blockchain along with FL enhances the existing framework, guaranteeing the protection of private information, reliability, and framework safety in decentralized collaborating-learning applications [6].

Moreover, while the FL and Edge-Computing (EC) environment provides data-privacy and data-security preserving frameworks, it still faces challenges and threats, which are mentioned below [7, 8]. ***Data security attacks***: as the FL is executed in internet-of-things and multi-edge server in a collaborative manner there is a higher chance of data being attacked thus impacting data integrity and various security vulnerability by giving access to unauthorized person. Hence, one of the most important things to think about when designing FL security approaches is how to make a trusted framework in a place that is unreliable. ***Data Privacy-Preserving Problems***: Since the task is executed across different service nodes on the edge, there is a higher chance of privacy leaks. Thus, researchers studying privacy-preserving FL approaches across EC face new challenges due to the ever-changing nature of attack types. ***Computation and Communication Overhead***: The swift proliferation of Smart-IoT (SIoT) services has resulted in a significant increase in volume of information at edge-nodes, resulting in increased computation and communication cost [9]. Thus, exploration of new FL approaches in EC environments is constrained by the inadequate computing efficiency, restricted transmission bandwidth, real-time networking, and high standards of service demands of edge-devices. ***Diverse attack***: the FL is prone to different attack like Free-Rider and Poisoning Attacks, Sybil, and inference attack [10]. Thus, there is a need for a more enhanced model that deals with different kinds of security attacks. ***Single-Point Failures***: FL exhibits vulnerability towards single-point failures due to its reliance on a central server for the transmission of model variables required for updating the model. FL frameworks, when integrated with blockchain technologies, enhance local decentralization and provide an efficient approach. To keep information stored securely, blockchain nodes work together. Their combined abilities allow them to check all stored models and information for malicious activity on any node [9]. However, the current blockchain model poses certain challenges in reputation design in detecting poisoning and backdoor attack behavior [10].

To address the above issue, several studies have been published in the literature [11-18] that utilize particular reward processes. The fundamental concept of the current reward or incentive-based mechanisms is that individuals provide modified information by introducing noise to maintain integrity and privacy, while fusion-centers compensate for the compromise of users' integrity and privacy [11, 12]. Nonetheless, this brings forth two additional challenges:

(i)determining an appropriate level of noise to maintain the necessary privacy and (ii) ensuring effective information trustworthiness while ultimately reducing the effects of compromised information. It is essential to initially measure the threshold to preserve integrity and privacy, followed by concurrent improvement of aggregation accuracy while ensuring that users are provided with suitable thresholds for integrity and privacy preservation [12]. However, creating a reward system that guarantees integrity is essential. The compensation of users who provide altered confidential information about their respective preserved privacy-integrity threshold established by the federated coordinator is essential [16]. This indicates that the compensation coming from aggregation/fusion-centers to users is connected with the user's trust, which is derived from the dependability of their submitted information through an effective verification procedure [17, 18]. This work introduces an effective learning-driven approach for privacy-integrity aware blockchain communication (PIABC) for a federated edge-learning platform, aimed at addressing the aforementioned challenges. The proposed model offers the best possible trade-off among fusion accuracy and integrity-privacy preservation level, resulting in optimum integrity-privacy preservation data fusion. Then, it validates the dependability of the information and finally updates worker and user trust (reputations) and calculates the fused weights accordingly. **The contributions of work are as follows**:

- The paper introduces an innovative blockchain-based communication model designed to ensure both data privacy and integrity in federated edge learning.
- A novel trust mechanism is developed, utilizing blockchain for secure user authentication while preserving user privacy.
- A global consensus model is designed to rigorously enforce privacy and integrity standards within the federated edge learning framework.
- The model demonstrates higher throughput, improved detection rates, and fewer misclassifications of attacks compared to existing methods.

The paper is organized as follows: Section 2 discusses various current methods designed to provide blockchain-enabled federated learning to mitigate different attacks. The section highlighted the benefits and limitations of the current security mechanism for a federated edge learning platform. Section 3 aims at designing a novel approach to address both privacy and integrity issues, adopting a trust and consensus model. Section 4 validates the result of the proposed approach over existing methodologies. Finally, the significance of research in terms of performance parameters is discussed alongside the future direction to enhance the security model.

## 2. LITERATURE SURVEY

This section reviews security schemes for federated learning (FL) ensuring user privacy and data integrity, emphasizing their contributions and limitations. H. Liu *et al.* [11] proposed a blockchain-based trust model using a trust-computing sandbox and state-channel blockchain with smart contracts to handle malicious activity and task scheduling via Deep Reinforcement Learning (DRL). Simulations with the OPENAI GYM framework showed

improvements in task completion, cost, and SLA, but lacked security evaluation. W. E. Mbonu *et al.* [12] introduced a blockchain-enabled secure aggregation method to protect central servers, improve scalability, and reduce single points of failure, while using fault-tolerant servers for stragglers and callbacks to cut training time and storage. Evaluations on MNIST confirmed better accuracy, lower communication cost, and scalability, but the model did not fully address attacks.M. A. Mohammed *et al.* [13] presented an approach for healthcare, where they utilized blockchain-based FL for scheduling and offloading data to central servers, presenting an energy-efficient model called Energy-Efficient Distributed FL Offloading-Scheduling (ED- FOS). The main aim was to reduce energy, training time, and provide better Quality-of-Service. Simulations showed that EDFOS minimized energy consumption by 39%, training duration by 29% and resource consumption by 36%. This EDFOS provided better outcomes for training FL, yet failed to provide any security for users. M. Zirui *et al.* [14] presented a blockchain-based privacy-preserving approach for the healthcare sector, i.e., to help users overcome depression because of COVID-19. This work utilized a consensus blockchain-based privacy-preserving approach for providing security, privacy, trust, and interoperability. For simulation, a blockchain environment was created and evaluations were conducted in terms of cost, latency, and trust, where the best outcomes in comparison with existing approaches were achieved, yet failed to provide any outcomes on providing security. H. Javed *et al.* [15] presented a security model for monitoring systems used in smart healthcare. Their main aim was to handle insider malicious attacks. Hence, this work focused on providing security in the cloud for presenting a model called Cloud-Access Security-Broker (CASB), which collected all actions (logs) performed by users and provided security using blockchain. Evaluations were conducted by simulating an environment where patients' data was collected, and whenever a user retrieved the data, the user id was evaluated. Results were evaluated in terms of data storage duration and overall blockchain performance. The proposed approach provided integrity, scalability, privacy, accessibility, and transparency. S. T. Ahmed *et al.* [8] presented an approach for smart healthcare that utilized blockchain-based FL. Their main aim was to provide privacy by using a global-based aggregation approach, which indexed data in a central server and synchronized the knowledge server. Evaluations were conducted by considering various IoT devices, where the evaluated IoT behavior and classification delay were considered. Each IoT device was labeled and delay was evaluated within the FL environment, and findings show that their approach reduced labeling time and delay.

I. U. Din *et al.* [16] presented a trust-based approach called Context-Aware-Cognitive Memory-Trust Management-System (CACMTM) for smart transportation. The trust interactions among vehicles were constructed using game theory, where different trust modules were built. The main focus of this work was to utilize past trust established with vehicles (IoT nodes) to understand the behavior of vehicles, hence improving the trust approach. Also, considered a historical trust module to reduce attacks. After the trust module, it utilized blockchain for providing better security, accountability, and transparency. The CACMTM architecture was built in the following manner: first, trust was evaluated between vehicles, then trust was decided (i.e., to establish a connection or not), then update trust modules (all different trust modules), and finally utilize the trust in blockchain to provide security and prevent attacks. Simulations were conducted using the popular simulator OMNet++, where different attacks were evaluated. Evaluations were conducted in terms of accuracy, computation overhead, attack detection, and time, where CACMTM achieved the best results when compared with other approaches. Also, they evaluated results for different trust thresholds and for different numbers of IoT nodes, where the best result was obtained. Q. Xie *et al.* [17] presented an approach for the Internet of Vehicles (IoV) to provide security and identify attacks. This work utilized a cryptography key encryption and decryption approach along with Physical-Unclonable-Functions (PUFs) to identify an attack. Evaluations were conducted in terms of cost and how they can handle different attacks. Findings show that the approach can identify inside and outside attacks efficiently. Z. Ma *et al.* [18] presented an approach for IoV that was reliant on Road-Side-Units (RSUs), for which they presented a blockchain-based security-distributed authentication approach. This approach first collected data, preprocessed and stored data at the edge to decrease delay in communication and response time, which was done using a trusted authority. Further, smart contracts were used for authentication along with an enhanced Practical-Byzantine Fault-Tolerant Consensus approach for providing authentication for the blockchain ledger. Further, provided security using a Real-or-Random approach. Performance was measured in terms of execution and communication cost. However, the current method failed to identify selfish and biased workers; further, reaching a global consensus model to assure privacy-integrity in blockchain-enabled federated edge-server is difficult. In the next section, addressing the research core issues, the following methodology is presented.

## 3. PROPOSED METHODOLOGY

In this study, we propose a novel, blockchain-enabled, trust-based privacy-preserving authentication model to enhance security and privacy in federated learning (FL) environments. By leveraging blockchain technology, the system ensures secure and verifiable interactions between users and nodes. The model integrates dynamic trust evaluation mechanisms that continuously assess and update trust levels of participating nodes, based on recent interaction data and connection metrics. Blockchain's immutable ledger ensures transparency and accountability in this process, enabling real-time adaptation to node behavior and effectively isolating malicious entities.

Simultaneously, the methodology includes a privacy and integrity-aware consensus-based aggregation scheme to safeguard data privacy and integrity during model training. Blockchain technology underpins the aggregation process, providing a transparent and secure environment for model fusion. By incorporating secure computation techniques and differential privacy, the system ensures that the aggregation models remain robust against adversarial attacks and privacy breaches. The fusion of blockchain with federated learning provides a secure, trustworthy, and privacy-preserving framework, boosting the overall resilience and reliability of the system.

### 3.1. ARCHITECTURE

The architecture of the Privacy-Integrity-Aware Blockchain Communication (PIABC) framework, applied within a federated edge-server learning platform, is illustrated in Figure 1. This architecture leverages blockchain technology to establish a secure federated learning environment. The system consists of user nodes and computing nodes interconnected through the blockchain. Initially, users interact with the system via a blockchain-based trust privacy-preserving authentication model. Blockchain ensures that every user's identity and actions are securely validated before any model training begins, preventing malicious actors from accessing the system. Once the authentication is completed, the FL server assigns model training tasks to the computing nodes, which use blockchain to log the task assignments and track the training progress, ensuring transparency and trust in the training process.

Upon completion of model training, the nodes aggregate the trained models and send the fused results back to the users through a privacy and integrity-aware consensus-based fusion approach. The fusion process is secured by blockchain's immutable ledger, ensuring that the aggregation is transparent and that no unauthorized modifications can occur during the fusion process. Blockchain technology also guarantees that the model updates and aggregated results are auditable, further enhancing trust in the system.

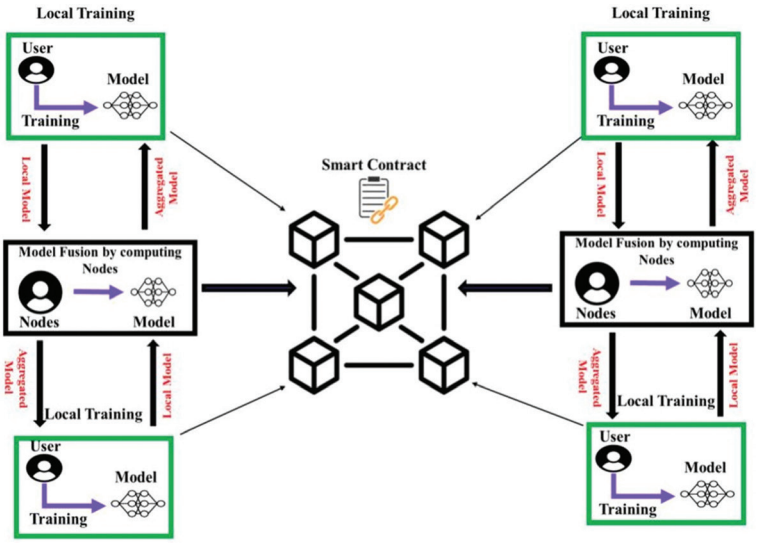


**Fig. 1.** Architecture of proposed Privacy-Integrity-Aware Blockchain Communication (PI- ABC) in federated edge-server learning platform

The PIABC approach, supported by blockchain's decentralized and secure infrastructure, ensures privacy and data integrity for each user without requiring any data sharing between users. To prevent potential security threats, such as inference or pollution attacks, a firewall is established between users. This firewall enforces a strict no-data-sharing policy while still allowing nodes to act as intermediaries for any necessary data transmission between users, thus ensuring enhanced security. In this setup, the nodes are responsible for executing model training and aggregation, while the FL server assigns the model training tasks. Blockchain provides a transparent and immutable record of task assignments, node activities, and model updates, ensuring the security and accountability of the entire system. The pseudocode of the same is presented in Algorithm 1.

### 3.2. BLOCKCHAIN-BASED TRUST PRIVACY PRESERVING AUTHENTICATION MODEL

The blockchain-based trust privacy-preserving authentication model ensures secure and verifiable interactions between users and computing nodes by utilizing dynamic trust evaluations combined with blockchain's immutable features. This model is designed to continuously assess and adjust trust levels between users and computing nodes based on a variety of metrics. These metrics include recent and past interactions, connection failures, and indirect trust derived from other nodes within the network, which are explained in further detail in the following sections. To achieve dynamic trust allocation, the model assigns trust weights to nodes based on their performance and reliability, allowing the system to identify and isolate malicious nodes effectively.

This approach guarantees that only trustworthy nodes are engaged in the federated learning process, maintaining the integrity of the system. Malicious or unreliable nodes are penalized, while trustworthy nodes are rewarded, incentivizing positive behavior and ensuring a secure and privacy-preserving environment for all users.

The trust levels are continually updated and validated through an exponential-average updating process, which ensures that the system remains resilient to manipulation and changes in node behavior. This ongoing validation of trust protects the system from fraudulent activities while fostering secure and reliable connections between users and nodes.

#### 3.2.1. Direct and Indirect Trust Establishment

This work presents a validation and trustworthiness approach that authenticates the user and establishes secure interactions with computing nodes (workers). Initially, in this model, a trust level is calculated, which involves evaluating the trust a user has in a computing node (worker). It is important to note that each computing node stores relevant data, including user ID, time, data type, data size, and trust level, for the entire established connection between the user and the computing node.

To optimize storage overhead and efficiently manage the data collected by the computing node, this work allocates specific weights for storing the data. An exponential-average updating process is used to directly store this data into the Interplanetary File System (IPFS), ensuring scalability and decentralization.

Let $Sec_o^u(x, y)$ ,where $x$ denotes user, $y$ denotes computing node, o denotes total interaction time for given data-type, and u denotes time-period. The parameter $Sec_o^u$ is used for evaluating trust-level between user and computing node. In this work, the direct trust is established between computing node $y$ and user $x$ and also between computing node $y$ to computing node $p$ and computing node $p$ to user $x$.

The trust calculation process starts with an initial security metric. This initial value is typically set to a baseline value of $L_o^u = 0.5$ for any new interaction. This baseline represents neutral trust, implying that there is no prior information or bias about the trustworthiness of the computing node or user. Essentially, it establishes a starting point where neither trust nor distrust is assumed. The value of $L_o^u$ is then updated dynamically based on the ongoing interactions between the user and the computing node, with more recent interactions having a greater influence on the trust level.

Direct trust is established in this work in three phases: between the user $x$ and computing node $y$, between computing node $y$ and computing node $p$, and between computing node $p$ and user $x$. This connection is represented by the direct trust metric $L_o^u(x, y)$, which is mathematically expressed in Eq. (1).

$$\mathbb{L}_o^u(x, y) = Sec_o^u(x, y) \quad (1)$$

**Algorithm 1** PIABC - Privacy-Integrity-Aware Blockchain Communication Framework

**Require:**

- User nodes $X = \{x_1, x_2, ..., x_n\}$ (perform local model training)
- Computing nodes $Y = \{y_1, y_2, ..., y_m\}$ (perform model aggregation)
- Federated Learning server $FL$
- Initial model $M$
- Blockchain ledger $B$
- Time period $u$, interaction duration $o$

**Ensure:** Aggregated global model $M_{agg}$ delivered to authenticated users

**Step 1: User Authentication**
**for** each user $x_i \in X$ **do**
  Verify user $x_i$ via blockchain-based authentication
  Log authentication event in blockchain ledger $B$
**end for**
**Step 2: Model Assignment and Local Training**
**for** each authenticated user $x_i$ **do**
  Distribute initial model $M$ to user $x_i$
  User $x_i$ performs local training to generate model $M_i$
  Record training status and progress in blockchain $B$
**end for**
**Step 3: Computing Node Selection**
**for** each computing node $y_j \in C$ **do**
  Evaluate node $y_j$ for aggregation eligibility
  **if** node $y_j$ meets selection criteria **then**
    Assign aggregation task to node $y_j$
    Log task assignment in blockchain $B$
  **else**
    Mark node as untrustcd and log rejection in $B$
  **end if**
**end for**
**Step 4: Model Aggregation**
**for** each selected computing node $y_j$ **do**
  Collect local models $\{M_1, M_2, M_n\}$ from user nodes
  Perform privacy-aware consensus-based fusion to generate $M_{agg}$
  Record aggregation process and final model hash in blockchain B
**end for**
**Step 5: Secure Model Distribution**
**for** each authenticated user $x_i$ **do**
  Deliver final aggregated model $M_{agg}$ to user $x_i$
  Log delivery event in blockchain ledger $B$
**end for**
  **return** $M_{agg}$

From Eq. (1), if computing node $y$ provides better execution, then user $x$ establishes connection having best trust-level. Similar happens with the computing node $y$ if it gives better execution, then user $x$ establishes connection with best trust-level. This helps user $x$ to achieve direct trust establishment.

In this work, the indirect trust is established between the computing node $y$ and user $x$ and also between computing node $y$ to computing node $p$ and computing node $p$ to user $x$ by considering past established connection. To gain knowledge about past established connections, the computing node $y$ connects with computing nodes p to collect the trust-level previously established by user $x$. Finally, computing node $y$ fuses (aggregates) trust-level established from computing nodes $p$ using Eq. (2).

$$\mathbb{G}_o^u(x,y) = \begin{cases} \frac{\sum_{p \in Z-\{x\}} \mathbb{F}_o^u(x,p) * \mathbb{L}_o^u(x,y)}{\sum_{p \in Z-\{x\}} \mathbb{F}_o^u(x,p)}, if\ |Z-\{x\}| > 0 \\ 0, if\ |Z-\{x\}| = 0 \end{cases} \quad (2)$$

In Eq. (2), $G_o^u(x, y)$ denotes fused trust-levels collected from computing nodes $p$ and $Z=S(y)$ denotes computing node $p$ which had established connection with computing node $y$.

This work utilizes weight-based approach, where weights are allocated dynamically for computing nodes. For allocation, higher weights are allocated for highly-trusted computing nodes, whereas lower weights are allocated for less-trusted computing nodes. Consider $F_o^u(x, y)$ which is used to denote the evaluation of validation-based security trustworthiness for a computing node $y$. The $F_o^u$ (x,y) is mathematically evaluated using Eq. (3).

$$\mathbb{F}_o^u(x,y) = \begin{cases} 1 - \frac{\log\left(Sec_o^u(x,y)\right)}{\log\theta}, & if\ \mathbb{R}_o^u(x,y) > \theta, \\ 0, & else \end{cases} \quad (3)$$

In Eq. (3), log$\theta$ denotes similar least-tolerable variable and $R_o^u(x, y)$ denotes relationship I between user $x$ and other computing nodes $y$ (where $p=y$ for one established connection between worker and computing node).

### 3.2.2. Evaluation of Latest and Past Established Trust

In this work, the latest trust is established between computing node $y$ and user $x$ and also between computing node $y$ to computing node $p$ and computing node $p$ to user $x$ is evaluated by considering both direct and indirect-trust. During evaluation of latest established trust, direct established trust is given higher trust-level, because the computing node $y$ or computing node $p$ interacts more with the user $x$. Hence, the latest established trust is mathematically evaluated using Eq. (4).

$$\mathbb{C}_o^u(x,y) = \delta * \mathbb{L}_o^u(x,y) + (1-\delta) * \mathbb{G}_o^u(x,y) \quad (4)$$

In Eq. (4), $C_o^u(x, y)$ denotes latest established trust metric and $\delta$ denotes trust-level weight assigned to direct established trust. $\delta$ is weighted function that can be optimized dynamically according to users $x$ interaction $T^u(x, y)$ on respective worker nodes $y$ considering time $u$; however, in this work average interaction $\vec{T}^u(x, y)$ time is considered to dynamically optimize the $\delta$ weighted value.

In Eq. (4), as u increases, the latest establish trust becomes old, which can be termed as past established trust and is denoted as $Q_o^u(x, y)$. The evaluation of $Q_o^u(x, y)$ is done similar to trust-level evaluation, i.e., using EAUP. Hence $Q_o^u(x, y)$ can be mathematically evaluated using Eq. (5).

$$\mathbb{Q}_o^u(x,y) = \frac{\varphi * \mathbb{L}_{o-1}^u(x,y) + \mathbb{C}_{o-1}^u(x,y)}{2} \quad (5)$$

In Eq. (5), φ(0≤φ≤1) is the incentive parameter and whenever L_0^0 (x,y), the whole evaluation changes to 0. By utilizing past established trust, the malicious computing nodes y connecting with computing nodes p or user x cannot change their process, i.e., computing node p cannot connect with y or x when a y has already established a connection with x. Hence, this metric allows a computing node y or computing node p to establish a connection with user x in a cooperative way, thereby reducing attacks. Also, the trust-level for the latest established trust changes to past established trust only when a computing node y or computing node p has made more connections with user x, hence increasing privacy.

### 3.2.3. Evaluation of Upcoming Trust Establishment

In this work, the upcoming trust is established between computing node $y$ and user $x$ and also between computing node $y$ to computing node $p$ and computing node $p$ to user $x$ is evaluated by considering both latest and past established trust. Consider $Future_o^u(x, y)$ as upcoming trust which will be established from computing node $y$ to user $x$ can be mathematically represented using Eq. (6).

$$Future_o^u(x,y) = \begin{cases} 0, if\ neither\ \mathbb{Q}\ or\ \mathbb{C}\ is\ available \\ \alpha\mathbb{C}_o^u(x,y) + (1-\alpha)\mathbb{L}_o^u(x,y)\ if\ either\ \mathbb{Q}\ or\ \mathbb{C}\ is\ available \end{cases} \quad (6)$$

In Eq. (6), α is a dynamic variable and in this work $\alpha$=0. The $\alpha$ can be dynamically changed using a deviating variable $\omega$ according to application/task requirement. Also, by making $\omega$ dynamic, a computing node can change its past established trust to latest established trust. Moreover, it is important that $\omega$ should not be set very less as malicious computing nodes can use this parameter for changing their behavior, i.e., they may change from malicious to non-malicious computing node, hence leading to attack to user $x$.

### 3.2.4. Security Metric for classification of Malicious Computing Node

For identifying and classifying malicious computing node, this work considers a security metric denoted as $F_o^u(x, y)$, which is evaluated by considering upcoming trust establishment and unfair and changing trust metric, which is mathematically represented using Eq. (7).

$$\mathcal{F}_o^u(x,y) = \mathbb{Q}_o^u(x,y) * Future_o^u(x,y) \quad (7)$$

Using Eq. (7), the computing nodes having higher upcoming trust-level, will result in less unfair and changing trust-level. Hence, the malicious computing nodes will have lesser trust-level, thereby reducing the attack on user $x$. Also, to having higher trust-levels during upcoming trust establishment, it is necessary that it should not change its process. Hence, using Eq. (7), user $x$ can select the computing node $y$ having higher trust-level, thereby reducing attack and increasing security.

### 3.3. PRIVACY AND INTEGRITY-AWARE CONSENSUS-BASED FUSION APPROACH

This section provides an efficient approach for fusing (aggregating) data to provide security, privacy, confidentiality and integrity. The aggregation process takes place after ensuring trust-level security and authenticating data privacy. Moreover, for providing integrity for authenticated data, a consensus-based fusion approach is presented, where the data which is unsecured is discarded

#### 3.3.1. Privacy-Integrity Consensus Approach

This work provides integrity by using a consensus-based fusion approach to prevent diverse attack which includes pollution and inference attack for federated-learning environment. Consider a blockchain-based federated-learning environment, where there exists $x$ users. By using graph-theory, this work considers users $x$ as a graph $H$={$E$, $V$}, where $H$ is a graph consisting of edges $E$ (set of connections or interactions between users) and vertices $V$ (set of users). Further, consider ($j$, $k$)∈$E$ (which implies a directed interaction from vertex $j$ to vertex $k$, meaning that user j sends data or model updates to user $k$), only if users are interconnected with each other. Consider the starting state of users as $y_j$ (0), having different time-session $l$. As the user $j$ will have contact or might establish a connection with other users.

The drawback of current consensus-based security approaches is that users in federated-learning environment try to get knowledge of other users starting states, i.e., $y$(0), hence impacting other user's privacy. The proposed consensus model during integrity assurance makes sure it preserves privacy requirements. Thus, the proposed consensus-based fusion approach utilizes $y(l)$ to converge to $\bar{y}$, for preserving user's privacy. The process of convergence is done in repetitive four steps so that it converges to $\bar{y}$. The steps are given below:

**1. Assumption of Random Data**: Consider that each user communicates random data $w_j$ ($l$) for time-session $l$, where variance=1 and mean=0. Also consider that $w_j$ ($l$) for various users is represented as $\{w_j(l)\}_{j=1,\ldots,\ o,\ l=0,\ 1,\ldots}$ which is uniformly distributed. For each user generating data and communication, a noise can be induced for every $y_j$ ($l$) which can be represented using $x_j$ (l). The noise $x_j$ (l) can be denoted using Eq. (8).

$$x_j(l) = w_j(o) \tag{8}$$

**2. Evaluation of Noise**: The Eq. (8) is only correct whenever $l$=0, else $x_j$ ($l$) is evaluated using Eq. (9).

$$x_j(l) = \beta^l w_j(l) - \beta^{l-1} w_j(l-1) \tag{9}$$

In Eq. (9), $\beta$ is constant variable for every user and changes from 0 to 1. The novel state for new user can be obtained using Eq. (10).

$$y_j''(l+1) = b_{jj} y_j''(l) + x_j(l) \tag{10}$$

**3. Interaction with Adjacent Users**: Further, the users interacting with nearby (adjacent) users and their state mean is evaluated using Eq. (11).

$$y_j(l+1) = b_{jj} y_j''(l) + \sum_{k \in \mathcal{O}(j)} b_{jk} y_j''(l) \tag{11}$$

**4. Updating States:** Revise $l$+1 states and return back to Step 1. This process is repeated until $y(l)$ converges to $\bar{y}$. Further, Eq. (10) and Eq. (11) are converted to matrix format as presented in Eq. (12).

$$y(l+1) = By''(l) = B\big(y(l) + x(l)\big) \tag{12}$$

By utilizing the 4 Steps, the $y(l)$ converges to $\bar{y}$ having accurate average state values. Also, during convergence it is important to assure that noise too reduces. Moreover, to achieve best convergence result, the asymptotic-sum has to be 0. The presented consensus-based security approach provides better outcome for both general and gaussian noise by utilizing highest-probability evaluation method. Also, the consensus-based security approach does not require any communication-channel, hence, utilizes less energy and resources and in federated-learning environment. The presented consensus-based fusion approach also provides privacy and integrity as it need less values for constructing consensus as users can optimize noise $x(l)$ independently, rather than depending on any variable. Moreover, it is important to know that each user or multiple users can choose $x(l)$ independently with considering the exponential-decaying co-variance matrix, but has to ensure that $x(l)$ does not affect consensus variables to achieve correct average consensus.

The data fused using Eq. (12) has better trust-level security privacy and integrity, but verifying data is important, hence, it is important to identify if there is any malicious packet present in network or not. For this, consider $J_0$ where data is fused for achieving better security to prevent attacks and $J_1$ denotes non-efficient data where attacks could happen. Hence from this, the attack can be identified by considering non-malicious packets interaction. A non-malicious packet will show no changes in state, whereas malicious packet will show a change. The state of normal packet can be denoted as $R_h$=$R$ ($J_1 \mid J_0$) and malicious packet can be represented as $R_m$=$R$ ($J_0 \mid J_1$). Hence a static-test can be developed for this evaluation, where initially the data variance is evaluated using Eq. (13).

$$N = \| \, \mathrm{y}_j(l) - \hat{y}_j(l) \, \|^2 \tag{13}$$

Eq. (13), provides difference among original data and original + noise data. For identifying malicious packet and non-malicious packets, the variable $N \lessgtr_{J_1}^{J_0} (\vartheta)$ is used. If there exists malicious packet, then the users can be changed to malicious, else they remain normal. Also, it has to be noted that the malicious packets consume more resources as they try to attack other nodes/packets; this helps in identifying the probability of attack within the federated learning environment. The proposed use of both trust-based authentication combined with a consensus-based privacy-integrity assurance model is effective in authenticating the user and eliminating false packets within the federated edge-server learning environment; the process improves overall throughput with higher detection accuracy and less misclassification, as proved in the following section.

#### 3.3.2. Theoretical Analysis and Security Guarantees

The previous section outlined the equations for trust computation and consensus-based privacy–integrity fusion. In this section, we present theoretical analyses and bounds demonstrating the model's behavior across different scenarios.

***Convergence Analysis of Trust Evaluation***: We provide a mathematical discussion showing that the exponential-average updating process used for both trust estimation and consensus fusion converges to a stable value over time, assuming bounded variance in interaction behavior. A formal convergence guarantee has been added by analyzing the recursive structure of Equations (4), (5), and (6) using principles from Markov decision processes and stochastic averaging.

***Security Proof of Malicious Node Isolation***: We analytically derive that malicious nodes characterized by fluctuating or degrading trust values are penalized in successive rounds due to diminishing trust weights (as shown in Equation (7)). This ensures that their influence on the federated learning process is minimized, and they are progressively isolated.

***Consensus Robustness under Noise and Attack***: We added a theoretical explanation of how the consensus protocol resists data pollution and inference attacks. By modeling user interaction as a graph and the injected noise as a bounded Gaussian variable, we show that the mean state converges with high probability to the correct average (Equation 12), supported by an asymptotic variance reduction analysis.

***Formal Attack Detection Bound***: For the malicious data detection metric (Equation 13), we now provide a statistical hypothesis testing model based on the Neyman-Pearson Lemma to distinguish between hypotheses $J_0$ (benign) and $J_1$ (malicious). This includes specifying decision thresholds $\vartheta$ and false positive/negative rates.

## 4. RESULTS AND DISCUSSION

To validate the practical applicability of the proposed PIABC security model, we simulate its deployment in a federated IoT environment under varying adversarial conditions. The objective is to assess the model's ability to detect and isolate malicious participants effectively while maintaining low misclassification rates and high throughput. PIABC's performance is benchmarked against an existing blockchain-aided privacy-preserving framework (BPPF) [18], using both synthetic and real-world datasets. Simulation scenarios are designed to reflect dynamic attack intensities, diverse threat types, and realistic network conditions, as detailed in the following section.

### 4.1. SIMULATION SCENARIO FOR REALISTIC DEPLOYMENT

Experiments were conducted to evaluate the PIABC model against the blockchain-aided privacy-preserving framework (BPPF) [18] using the CIC Federated Learning Dataset [19]. Performance was measured in terms of detection rate, misclassification rate, and throughput under attack intensities ranging from 10% to 40%. Both models were implemented in the C#-based SENSORIA simulator [20] with blockchain support via IoTSim-Osmosis [21].

To further evaluate the robustness of the PIABC model, simulations were performed on the UNSW-NB15 and CIC-IoT2023 datasets [20], which contain diverse multi-stage cyberattacks. Sybil and collusion attacks are mitigated through cross-node trust validation and anomaly detection, while backdoor threats are countered by tracking temporal trust variations and applying consensus-based integrity verification.

Before training, both datasets were preprocessed using a standard pipeline: removal of duplicate and incomplete records, one-hot encoding for categorical variables, and Min-Max normalization for numerical attributes. Highly correlated and low-variance features were filtered out, and final feature subsets were chosen based on statistical relevance and domain knowledge—25 features from UNSW-NB15 and 28 from CIC-IoT2023. The processed data was then split into training (70%), validation (15%), and testing (15%) sets using stratified sampling to maintain class balance. A fixed random seed ensured reproducibility, and 5-fold cross-validation was employed for hyperparameter optimization and anomaly threshold selection.

The simulation results demonstrate that PIABC consistently achieves higher detection accuracy and lower misclassification rates across varied attack scenarios, confirming its effectiveness and scalability in real-world federated IoT deployments.

### 4.2. DETECTION RATE

This section studies the detection rate performance of identifying the attack using both PIABC and BPPF under

the same simulation configuration. A higher value of detection rate indicates superior performance. The detection rate performance of both models is graphically shown in Figure 2. The result shows the proposed PIABC model has a higher detection rate in identifying the attack in comparison with BPPF, considering varied attack percentages. The enhancement achieved is due to the adoption of an effective trust model implemented in Eq. (7).

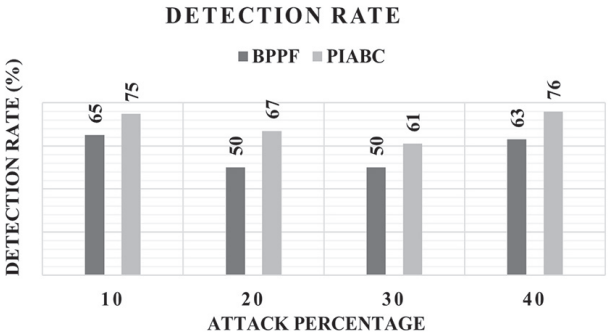


**Fig. 2.** Detection rate vs varied attack percentage

### 4.3. MISCLASSIFICATION RATE

This section studies the misclassification rate performance of wrongly identifying the attack using both PIABC and BPPF under the same simulation configuration. A lower value of the misclassification rate indicates superior performance. The misclassification rate performance of both models is graphically shown in Figure 3. The result shows the proposed PIABC model has a higher misclassification rate in wrongly identifying the attack in comparison with BPPF, considering varied attack percentages. The enhancement achieved is due to the adoption of an effective consensus model designed in Eq. (12).

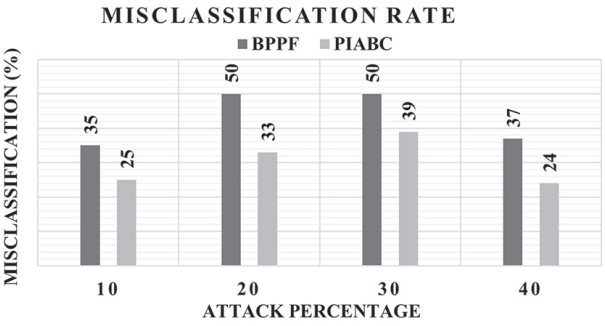


**Fig. 3.** Misclassification rate vs varied attack percentage

### 4.4. NORMALIZED THROUGHPUT

This section studies the normalized throughput performance by varying attack rate using both PIABC and BPPF under the same simulation configuration. The throughput is measured in terms of bits transmitted per second; however, in this research article normalized throughput is considered for validation.

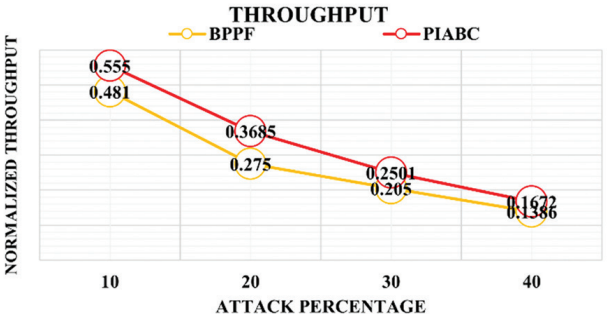


**Fig. 4.** Normalized throughput vs varied attack percentage

A higher value of normalized throughput indicates superior performance. The detection rate performance of both the models is graphically shown in Fig. 4. The result shows the proposed PIABC model has higher normalized throughput in comparison with BPPF considering varied attack percentage. The normalized throughput enhancement achieved is due to adoption of effective trust model implemented in Eq. (7) and consensus model designed in Eq. (12).

### 4.5. DISCUSSION

The PIABC attains convergence, by providing a formal proof showing that the exponential-average updating process converges under bounded interaction variance. Regarding security, we use statistical hypothesis testing to demonstrate resilience against Sybil and backdoor attacks based on trust deviation detection. For computational complexity, we analytically derive that the trust evaluation and consensus fusion algorithms operate with polynomial time complexity $O(n \cdot t)$, where n is the number of nodes and t is interaction time, ensuring scalability. The PIABC system employs a Proof-of-Authority (PoA) consensus protocol, integrated via the IoTSim-Osmosis framework, due to its lightweight nature and suitability for resource-constrained IoT environments. PoA allows for faster block confirmations and lower energy consumption compared to Proof-of-Work, making it ideal for real-time intrusion detection in federated learning-based systems. Blockchain-induced delays, including block generation and smart contract execution times, are simulated using event-driven modeling within SENSORIA, incorporating realistic network latency based on exponential distribution patterns. The storage model is designed for efficiency, with only hash-verified metadata, such as access logs and model update references, stored on-chain, while bulk data remains securely stored off-chain in cloud repositories. This hybrid approach ensures transparency, data integrity, and scalability. The measured computational overhead of the blockchain integration remains under 5%, while communication overhead is limited to 7–8%, primarily due to compact transaction sizes (~256 bytes). These overheads are directly correlated with key performance metrics. Despite the additional cost and processing load, the proposed PIABC model achieves

improved throughput, higher detection rates, and reduced misclassification rates when compared with the baseline BPPF framework. The reduction in false positives and increased detection accuracy justifies the minimal added overhead, while delay remains within acceptable thresholds.

## 5. CONCLUSION

This work shows that in federated learning, the data model is executed in multiple edge-server in a collaborative manner; as a result, it increases users' privacy and data breach because of a single point failure attack in the main computational server. Blockchain employing rewarding mechanism in a federated edge-learning platform aids the model to overcome single-point aggregation failure. However, the current method failed to identify selfish and biased workers; further, reaching global consensus model to assure privacy-integrity in blockchain-enabled federated edge-server is difficult. This work introduced privacy-integrity-aware blockchain communication (PIABC) in federated edge-server learning platform. An experiment is conducted to study the performance considering a varied attack size. The results show the proposed model can resist different attacks using the CIC-IOT federated edge attack dataset. The average percentage reduction in detection rate by PIABC over BPPF is approximately 18.46%. The average percentage improvement in misclassification rate by PIABC over BPPF is approximately 26.13%. The average percentage improvement in throughput by PIABC over BPPF is approximately 44.53%.

The PIABC model is very effective in comparison with the existing blockchain-privacy preserving scheme for identifying the correctly aggregated packets and eliminating malicious packets within the federated edge-server learning platform. Future work would consider developing more effective detection strategies to detect more complex attacks employing different benchmarks and also further optimizing the model. In the current work, we conducted detailed simulations to evaluate throughput, detection rate, misclassification rate, and computational complexity to measure the overhead of the proposed model. These evaluations demonstrate the model's effectiveness and low processing cost in simulated environments. However, aspects such as energy consumption, real-time latency, and deployment on actual edge hardware have not been covered in this study. Therefore, the future research will aim to validate the model's effectiveness and resource efficiency through deployment on actual edge computing platforms.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the Gujarat Council on Science and Technology (GUJCOST), India, for the research project, under project number GUJCOST/STI/2021-22/3867. This support has been instrumental in the successful completion of this research work.

## 6. REFERENCES:

[1] H. Li, L. Ge, L. Tian, "Survey: federated learning data security and privacy-preserving in edge-Internet of Things", Artificial Intelligence Review, Vol. 57, No. 5, 2024, p. 130.

[2] T. Alam, R. Gupta, A Ullah, S. Qamar, "Blockchain-Enabled Federated Reinforcement Learning (B-FRL) model for privacy preservation service in IoT systems", Wireless Personal Communications, Vol. 136, No. 4, 2024, pp. 2545-2571.

[3] Y. Jia, L. Xiong, Y. Fan, W. Liang, N. Xiong, F. Xiao, "Blockchain-based privacy- preserving multi-tasks federated learning framework", Connection Science, Vol. 36, No. 1, 2024, p. 2299103.

[4] J. Shen, S. Zhou, F. Xiao, "Research on Data Quality Governance for Federated Cooperation Scenarios", Electronics, Vol. 13, No. 18, 2024, p. 3606.

[5] K. M. Sameera, S. Nicolazzo, M. Arazzi, A. Nocera, R. R. KA, P. Vinod, M. Conti," Privacy-preserving in Blockchain-based Federated Learning systems", Computer Communications, Vol. 222, 2024, pp. 38-67.

[6] C. Dhasaratha, M. K. Hasan, S. Islam, S. Khapre, S. Abdullah, T. M. Ghazal, A. I. Alzahrani, N. Alalwan, N. Vo, M. Akhtaruzzaman, "Data privacy model using blockchain reinforcement federated learning approach for scalable internet of medical things", CAAI Transactions on Intelligence Technology, 2024. (in press)

[7] Z. Jovanovic, Z. Hou, K. Biswas, V. Muthukkumarasamy, "Robust integration of blockchain and explainable federated learning for automated credit scoring", Computer Networks, Vol. 243, 2024, p. 110303.

[8] S. T. Ahmed, T. R. Mahesh, E. Srividhya, V. Vinoth Kumar, S. B. Khan, A. Albuali, A. Almusharraf, "Towards blockchain based federated learning in categorizing healthcare monitoring devices on artificial intelligence of medical things investigative framework", BMC Medical Imaging Vol. 24, No. 1, 2024, p. 105.

[9] W. Moulahi, I. Jdey, T. Moulahi, M. Alawida, A. Alabdulatif," A blockchain-based federated learning mechanism for privacy preservation of healthcare

IoT data", Computers in Biology and Medicine, Vol. 167, 2023, p. 107630.

[10] L. Wang, C. Guan, "Improving security in the internet of vehicles: A blockchain-based data sharing scheme", Electronics, Vol. 13, No. 4, 2024, p. 714.

[11] H. Liu, H. Zhou, H. Chen, Y. Yan, J. Huang, A. Xiong, S. Yang, J. Chen, S. Guo, "A federated learning multi-task scheduling mechanism based on trusted computing sandbox", Sensors, Vol. 23, No. 4, 2023, p. 2093.

[12] W. E. Mbonu, C. Maple, G. Epiphaniou, "An end-process blockchain-based secure aggregation mechanism using Federated Machine Learning", Electronics, Vol. 12, No. 21, 2023, p. 4543.

[13] M. A. Mohammed, A. Lakhan, K. H. Abdulkareem, D. A. Zebari, J. Nedoma, R. Martinek, S. Kadry, B. Garcia-Zapirain, "Energy-efficient distributed federated learning offloading and scheduling healthcare system in blockchain based networks", Internet of Things, Vol. 22, 2023, p. 100815.

[14] M. Zirui, G. Bin, "A Privacy-Preserved and User Self-Governance Blockchain-Based Framework to Combat COVID-19 Depression in social media", IEEE Access, Vol. 11, 2023, pp. 35255-35280.

[15] H. Javed, Z. Abaid, S. Akbar, K. Ullah, A. Ahmad, A. Saeed, H. Ali, Y. Y. Ghadi, T. J. Alahmadi, H. K. Alkahtani, A. Raza, "Blockchain-based logging to defeat malicious insiders: The case of remote health monitoring systems", IEEE Access, Vol. 12, 2023, pp. 12062-12079.

[16] I. U. Din, K. A. Awan, A. Almogren, "Secure and privacy-preserving trust management system for trustworthy communications in intelligent transportation systems", IEEE Access, Vol. 11, pp. 65407-65417.

[17] Q. Xie, Z. Sun, Q. Xie, Z. Ding," A cross-trusted authority authentication protocol for Internet of Vehicles based on blockchain", IEEE Access, Vol. 11, 2023, pp. 97840-97851.

[18] Z. Ma, J. Jiang, H. Wei, B. Wang, W. Luo, H. Luo, D. Liu, "A blockchain-based secure distributed authentication scheme for internet of vehicles", IEEE Access, Vol. 12, 2024, p. 81471-81482.

[19] Datasets — Research — Canadian Institute for Cybersecurity — UNB, https://www.unb.ca/cic/datasets/index.html (accessed: 2024)

[20] N. Ababneh, J. N. Al-Karaki, "On the lifetime analytics of IoT network", Proceedings of the International Conference on Communication and Signal Processing, Chennai, India, 28-30 July 2020, pp. 1086-1090.

[21] A. Albshri, A. Alzubaidi, M. Alharby, B. Awaji, K. Mitra, E. Solaiman, "A conceptual architecture for simulating blockchain-based IoT ecosystems", Journal of Cloud Computing, Vol. 12, No. 1, 2023, p. 103.