

Crowdsensing with Federated Trust Management: Adaptive Defense Against Malicious Contributions

Jiantao Xu¹, Chen Zhang², Liu Jin³, and Chunhua Su[✉]

¹University of Aizu, Aizuwakamatsu 965-8580, Japan
d8252108@u-aizu.ac.jp

²University of Aizu, Aizuwakamatsu 965-8580, Japan
d8252109@u-aizu.ac.jp

³University of Aizu, Aizuwakamatsu 965-8580, Japan
d8242103@u-aizu.ac.jp

[✉]University of Aizu, Aizuwakamatsu 965-8580, Japan
chsu@u-aizu.ac.jp

Abstract. Crowdsensing, as a powerful paradigm, enables the collection of large-scale sensor data from distributed participants. However, malicious contributors providing unreliable or tampered data pose significant challenges to data quality and task reliability. This paper proposes a Federated Trust Management (FTM) framework that integrates a trust assessment mechanism into Federated Learning to mitigate the impact of adversarial contributions. The proposed method dynamically adjusts the credibility of participants and introduces a weighted aggregation strategy in FL, thereby reducing the influence of low-quality or malicious data. Experimental results demonstrate that FTM significantly improves global model accuracy, adversarial impact, and enhances the robustness of crowdsensing applications.

Keywords: Federated Learning; Crowdsensing; Reputation Assessment; Malicious Detection; Adaptive Defense

1 Introduction

With the continuous expansion of mobile and IoT devices, the paradigm of crowdsensing has garnered increasing attention in a wide range of application domains. Crowdsensing leverages large-scale user participation to collect and aggregate data for tasks such as environmental monitoring, smart city services, and intelligent transportation systems [1, 2]. Compared to traditional centralized data-gathering schemes, crowdsensing can significantly reduce operational costs, while offering broader geographic coverage and finer-grained sensing capabilities. However, because any user can upload sensing data to the platform, assuring the quality and trustworthiness of contributed data is a pivotal challenge. In particular, malicious contributors may submit forged or low-quality data, or employ collusion and sybil attacks to undermine the system [3–5].

Existing solutions for data quality in crowdsensing often rely on centralized trust or reputation models. These approaches track users’ historical contributions and penalize those with consistently poor or suspicious behavior [6, 7]. While such mechanisms can enhance data reliability, they have several drawbacks. First, a single point of failure may arise if the central server is compromised or malfunctions. Second, collecting and storing extensive user data brings privacy risks. Third, in dynamic and large-scale networks, adversaries can quickly coordinate or modify their patterns of behavior to circumvent conventional reputation systems.

Federated Learning has recently been proposed as a way to enable large-scale distributed model training while preserving participant privacy [8–10]. In FL, clients train locally on their devices and only upload model updates (e.g., parameters or gradients) to a central server, thereby reducing the need to share raw data. Nevertheless, FL is not immune to security threats: because the server has limited control over local training environments, malicious participants can engage in “data poisoning” or “model poisoning” attacks that degrade global model performance or intentionally steer the model to converge on an erroneous objective [11–13]. Various robust aggregation strategies have been explored to mitigate malicious gradient updates [14, 15], but most of these approaches rely on predefined assumptions about attacker behavior or apply static filters that may not adapt well to adversaries who vary their strategies over time.

To address these gaps, this paper proposes a Federated Trust Management (FTM) framework, which integrates a dynamic, adaptive trust assessment into FL for crowdsensing scenarios. In contrast to traditional reputation-based methods, the key innovation is a real-time trust adjustment mechanism that accounts for participants’ recent contribution quality, anomaly detection results, and prior track records, thus allowing the system to respond more quickly to emerging threats. Moreover, by incorporating trust scores into the global aggregation step, honest participants gain a proportionally higher influence on the federated model, effectively penalizing malicious or low-quality updates. This form of dynamic trust adjustment provides the system with better flexibility and robustness against rapidly changing attack patterns.

The main contributions of this study include:

- 1) Proposal of the FTM Framework: By combining dynamic trust management with FL, we reduce the impact of malicious contributions without exposing participants’ raw data.
- 2) Adaptive Trust-Based Aggregation: We introduce a real-time trust update formula and integrate it into the federated model aggregation, preventing collusive adversaries from manipulating the system.
- 3) Comprehensive Experimental Evaluation: Through simulations on the MNIST dataset under various attack scenarios, we demonstrate that FTM achieves high global model accuracy, even under diverse and evolving attacks.

2 Related Work

In this section, we review the existing research on the trustworthiness of crowd-sensed data, the security and robustness of Federated Learning, and trust management models in distributed systems. These works form the foundation of the proposed Federated Trust Management framework.

2.1 Data Trustworthiness in Crowdsensing

Crowdsensing relies on contributions from distributed participants, making data reliability a critical issue. Various mechanisms have been explored to ensure the trustworthiness of data in crowdsensing systems. For example, reputation-based methods use contributors' reputation scores to filter unreliable data [16–18]. These methods maintain historical records of user behavior and adjust reputation scores over time. However, they have limitations: firstly, they are vulnerable to collusion attacks where malicious users can collectively boost their reputation to manipulate the system; secondly, there is a delay in trust adjustment as reputation scores require sufficient historical data to stabilize, rendering them ineffective against dynamic adversaries; and finally, there are filtering methods based on statistical and machine learning approaches, where some work employs outlier detection and anomaly detection techniques to identify malicious contributions [19–21]. These methods analyze data consistency and deviations from expected patterns. Although effective in some cases, they face challenges such as adaptive attackers—malicious contributors may generate seemingly reasonable yet incorrect data to evade detection—and computational overhead, as running complex machine learning detection algorithms on resource-constrained edge devices is challenging. These limitations highlight the need for a distributed trust management mechanism that can dynamically adapt to evolving threats.

2.2 Security and Robustness of Federated Learning

Federated Learning is a distributed model training technique that preserves privacy by enabling model training without sharing raw data [8–10]. However, when malicious clients participate in training, FL introduces new security risks.

Existing research indicates that malicious participants can manipulate FL models through several means. First, there is data poisoning, where misleading data is injected to bias the global model [11–13]. Next, there is model poisoning, where local updates are manipulated to degrade overall model performance [8, 22, 23]. To mitigate these threats, various robust FL aggregation strategies have been proposed. For example, Xu et al. [14] introduced a hierarchical adaptive sparsification model aggregation (LASA) method that pre-aggregates client updates by sparsifying them and adaptively selects benign layers based on scale and direction to enhance Byzantine fault tolerance, effectively reducing the impact of malicious clients during FL model training. Alsulaimawi [15] employed gradient-based analysis to detect abnormal gradient patterns, enabling

the identification and filtering of anomalous model updates prior to aggregation, significantly enhancing FL security. Although these methods improve the security of FL, they often assume that malicious clients follow specific attack patterns. The approach presented in this paper further enhances FL’s defense capabilities through dynamic trust assessment.

2.3 Trust Management in Distributed Systems

Trust management frameworks have been widely applied in various distributed computing environments, including blockchain-based reputation systems and peer-to-peer networks [24].

Centralized vs. Decentralized Trust Models Trust management can be divided into two categories. In centralized trust models, a single trusted entity assigns and manages trust scores [25]. Although simple to implement, they are prone to single points of failure and attacks. In decentralized trust models, trust scores are computed in a distributed manner, reducing dependency on a central entity. Some works have proposed graph-based trust propagation methods [26] to infer the trust level between users.

Federated Trust Mechanism Recent work has begun exploring the integration of trust mechanisms into FL [27]. However, existing studies primarily focus on static trust scores or fixed weighting schemes, which are clearly insufficient for dynamic adversarial environments. The proposed FTM framework introduces an adaptive trust aggregation mechanism that continuously adjusts participants’ trust levels based on real-time data contributions and model updates, thereby extending previous work.

The existing literature on the trustworthiness of crowdsensing, FL security, and trust management has the following shortcomings: Traditional crowdsensing trust models rely on static reputation systems, which are difficult to cope with continuously evolving malicious strategies; FL security methods mainly focus on robust aggregation but lack adaptive defense against long-term adversarial behaviors; Existing federated trust mechanisms often adopt predefined trust scores instead of dynamically adjusting based on real-time contributions.

We proposed FTM framework directly integrates trust assessment into FL, enabling the system to dynamically adjust participants’ credibility, thereby improving the robustness of crowdsensing applications. Table 1 shows the Symbol and Definition used in this paper.

3 Proposed Method

This section introduces the Federated Trust Management (FTM) framework, which integrates dynamic trust assessment into Federated Learning (FL) to mitigate the impact of malicious contributions in crowdsensing systems. As shown

Table 1. Notations and Definitions

Symbol	Definition
P	Set of participants in crowdsensing
T_i^t	Trust score of participant i at round t
Q_i^t	Contribution quality of participant i at round t
w_i^t	Local model update from participant i at round t
w_t^{global}	Global model at communication round t
α	Forgetting factor for historical trust
β	Penalty factor for trust degradation
D_i^t	Deviation of participant i 's update for global model
λ	Regularization parameter for trust fluctuation
$R(T)$	Regularization term penalizing unstable trust changes

in Fig. 1, the framework primarily comprises the following components: (i) *Local Trust Computation*, where each participant evaluates its own credibility based on the quality of its contributions; (ii) *Federated Trust Aggregation*, where the central server integrates trust values into the global model update; and (iii) *Adaptive Adversarial Defense*, which dynamically adjusts trust scores to counter continuously evolving malicious behavior.

3.1 System Model

This paper considers a crowdsensing system where multiple distributed participants collect and submit sensor data for a global task. The system follows the Federated Learning paradigm, where clients train local models on their respective devices and the central server aggregates client updates without sharing raw data. The system consists of: A set of participants or clients $P = \{p_1, p_2, \dots, p_n\}$, responsible for contributing data and training local models; A trusted central server responsible for collecting local model updates, integrating trust values, and generating the global model; Adversarial participants, where some malicious users attempt to manipulate the system by injecting fake data or tampering with model gradients.

The threat model includes data poisoning attacks, where malicious contributors provide low-quality or deliberately tampered data, and model poisoning attacks, where adversaries manipulate model updates to degrade global performance.

The FTM framework aims to dynamically detect and mitigate the above threats through trust-based adaptive aggregation.

3.2 Local Trust Computation

Each participant maintains a local trust score T_i to reflect its reliability. The trust score is dynamically updated based on the following factors: data consistency (the deviation of submitted data from global consensus), historical contributions

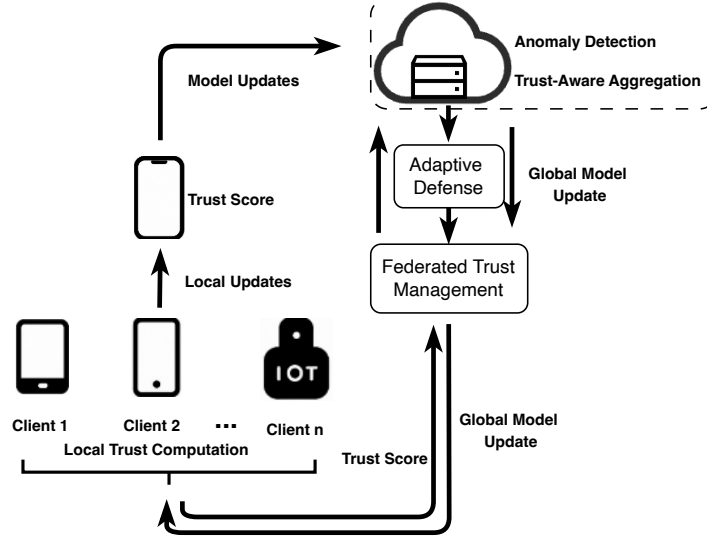


Fig. 1. FTM system architecture

(the evolution of the participant's trust score over time), and anomaly detection (using statistical and machine learning models to detect outlier data).

Trust Score Update Mechanism The trust score T_i is updated using a weighted function:

$$T_i^{t+1} = \alpha \cdot T_i^t + (1 - \alpha) \cdot Q_i^t \quad (1)$$

where Q_i^t represents the quality of the contribution in round t , and α is the forgetting factor that balances historical performance and recent behavior.

3.3 Federated Trust Aggregation

To reduce the impact of low-trust participants, the central server incorporates trust values into the model aggregation process.

Trust-Aware Model Aggregation Unlike standard FedAvg aggregation, the proposed method adopts a trust-weighted aggregation:

$$w_t^{global} = \sum_{i \in P} T_i^t \cdot w_i^t \quad (2)$$

where w_i^t is the local model update of participant p_i . This mechanism ensures that high-trust contributors have a greater impact on the global model.

Trust Score Regularization To prevent the malicious exploitation of trust scores, a regularization term is introduced to penalize participants with excessive fluctuations in their trust scores:

$$R(T) = \lambda \sum_{i \in P} (T_i^t - T_i^{t-1})^2 \quad (3)$$

where λ controls the strength of the penalty.

3.4 Adaptive Adversarial Defense

Dynamic Trust Adjustment To counter malicious behavior, this paper introduces an adaptive trust adjustment mechanism. If a participant's model update deviates significantly from that of the majority, its trust score is penalized:

$$T_i^{t+1} = \max(0, T_i^t - \beta \cdot D_i^t) \quad (4)$$

where D_i^t represents the deviation from the aggregated update, and β is the penalty factor.

Adversarial Robustness Analysis The robustness of FTM against Byzantine attacks is evaluated from two aspects: trust score convergence, ensuring that malicious users are gradually downgraded; and the degradation in model accuracy, measuring the impact of malicious participants on global model performance.

3.5 Algorithm Description

Algorithm 1 outlines the training process of FTM.

Algorithm 1 Federated Trust Management (FTM) Training Process

- 1: **Input:** Participants P , initial trust scores $\{T_i^0\}$, learning rate η
- 2: **for** each communication round $t = 1, 2, \dots, T$ **do**
- 3: **for** each participant $p_i \in P$ (executed in parallel) **do**
- 4: Train local model and compute update w_i^t
- 5: Compute local trust score update T_i^{t+1}
- 6: **end for**
- 7: Compute trust-weighted global aggregation:

$$w_i^{global} = \sum_{i \in P} T_i^t \cdot w_i^t$$

- 8: Update trust scores based on adversarial detection
 - 9: **end for**
 - 10: **Output:** Final global model w_T^{global}
-

The FTM framework presented in this section integrates trust-aware aggregation into Federated Learning to mitigate malicious contributions in crowdsensing

systems. By dynamically adjusting trust scores and penalizing anomalous behavior, the proposed method enhances data reliability and improves the robustness of FL against malicious attacks.

4 Experimental Results

This section evaluates the effectiveness of the proposed Federated Trust Management (FTM) framework in mitigating malicious contributions in crowdsensing applications. First, the experimental setup is introduced, followed by an analysis of key performance indicators such as data reliability, global model accuracy, and adversarial resilience, demonstrating the advantages of the proposed method.

4.1 Experimental Setup

Dataset We simulate a crowdsensing environment where participants contribute sensor data for a classification task. The dataset is based on MNIST. To simulate adversarial behavior, malicious participants are introduced with actions including: adding random noise to submissions (random noise attack); submitting biased data to distort the overall distribution (data poisoning attack); and manipulating model updates by constructing adversarial gradients (model poisoning attack).

Baseline Methods To evaluate the effectiveness of FTM, we compare it with the following baseline methods: Standard Federated Learning (FedAvg): Aggregates model updates without considering trustworthiness. Reputation-Based Filtering (RepFilter): Assigns static reputation scores based on historical contributions. Byzantine Fault-Tolerant Aggregation: Identifies and filters out anomalous updates.

Evaluation Metrics The performance of FTM is assessed using the following metrics: Global Model Accuracy (GMA): The accuracy of the federated model on the test set. Adversarial Influence (AI): The impact of malicious contributions on model convergence.

4.2 Experimental Results and Analysis

Impact on Global Model Accuracy 2 compares the global model accuracy of different methods under varying proportions of malicious participants. As shown in (a), FTM rapidly improves and maintains high model accuracy across different malicious proportions. Under a 25% malicious node scenario, FTM achieves nearly the same accuracy as the attack-free case; even at a 75% malicious proportion, FTM still demonstrates superior convergence speed and stability, with significantly lower accuracy fluctuations than other methods. In contrast, (b)

shows that FedAvg experiences severe accuracy fluctuations under high malicious proportions, struggling to converge to an optimal level in later stages; (c) indicates that RepFilter exhibits increased accuracy variance and slower convergence when facing 50% or more malicious nodes; (d) illustrates that Byzantine Fault-Tolerant Aggregation also encounters significant fluctuations when malicious nodes exceed half of the total participants. Overall, FTM maintains stable and high-level global model accuracy even in strongly adversarial environments.

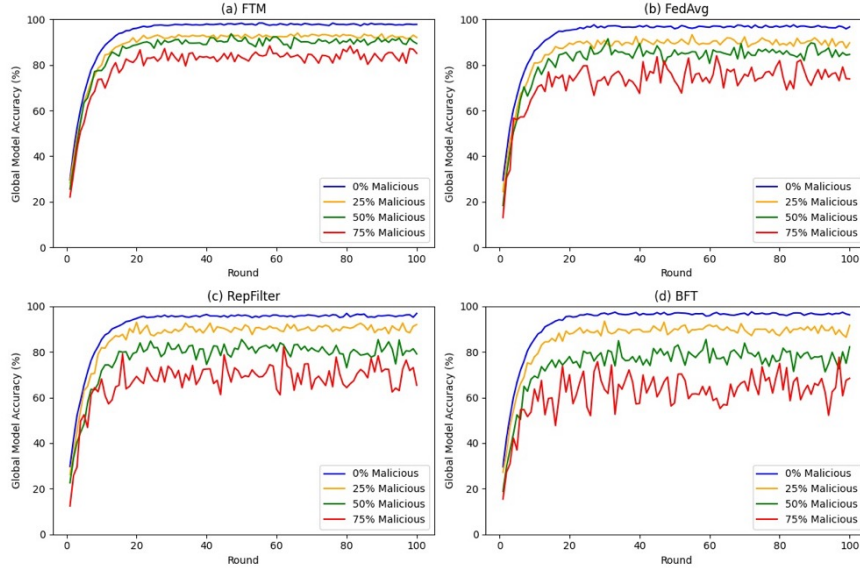


Fig. 2. Model accuracy of different methods.

Analysis of Adversarial Influence Figure 3 presents the Adversarial Influence (AI) metric, revealing the extent of malicious client impact on the global model. As shown in (a), the FTM scheme effectively suppresses attack interference under various malicious node proportions: while initial fluctuations exist, the AI level rapidly decreases and stabilizes with training iterations, maintaining a low impact range even at a 75% malicious proportion. In contrast, (b) shows that AI curves remain high and fluctuate drastically under high malicious proportions; (c) indicates that AI is somewhat alleviated when facing 50% or more malicious nodes but still fails to significantly decrease further; (d) demonstrates that when malicious nodes exceed half, AI remains at a medium-to-high level for a prolonged period. The overall comparison suggests that FTM can suppress adversarial influence faster and more effectively, showcasing strong robustness in resisting malicious interference.

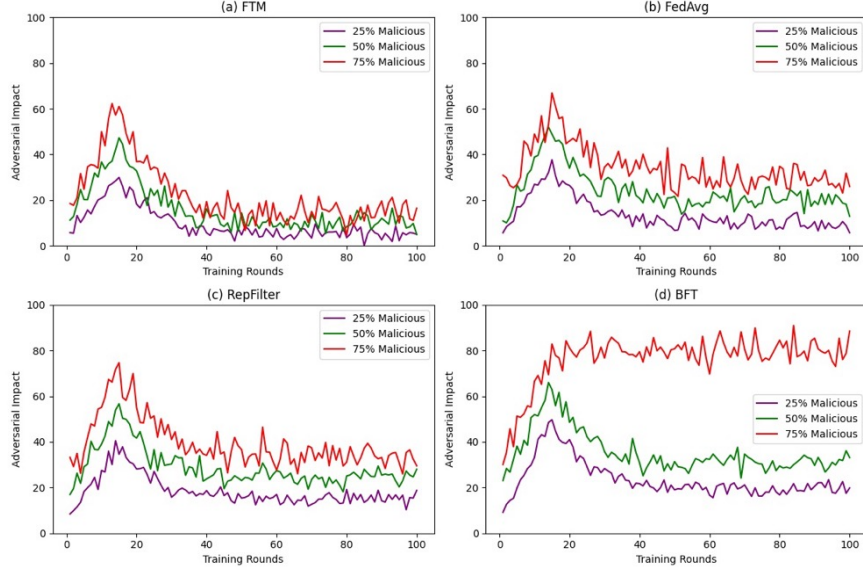


Fig. 3. Impact of malicious participants on model convergence.

5 Discussion

We discuss the broader implications, limitations, and future directions for improvement of the proposed method in this section.

Regarding the effectiveness of dynamic trust adjustment, the results indicate that dynamically adjusting trust scores significantly enhances the robustness of Federated Learning against malicious participants. Unlike traditional static reputation systems, FTM continuously updates trust scores based on recent contributions, enabling the system to better adapt to evolving threats.

In terms of the trade-off between trust sensitivity and model convergence, while a highly sensitive trust mechanism can effectively detect and isolate malicious users, it may also lead to instability in FL model training. Excessive penalization of participants based on limited data samples could prematurely exclude legitimate but noisy contributors. This paper mitigates this issue by employing a gradual trust decay strategy, ensuring that trust adjustments are not overly abrupt.

We evaluated various adversarial behaviors, including random noise attacks, data poisoning attacks, and model poisoning attacks. The results demonstrate that FTM maintains high model accuracy and stable trust score distributions under all attack scenarios. However, there are limitations for more sophisticated adversarial strategies, such as malicious participants gradually adjusting their behavior to evade detection. A notable limitation is the increased computational burden on client devices. The local trust assessment process requires analysis of

historical contributions and execution of lightweight anomaly detection models, which may increase resource consumption, particularly on low-power IoT devices.

Another limitation is vulnerability to other attacks. Although the framework effectively reduces the impact of individual malicious participants, it may still be susceptible to attacks where a malicious entity creates multiple fake identities to manipulate the trust aggregation.

Furthermore, regarding scalability and communication overhead, Federated Learning inherently requires multiple rounds of communication between clients and the central server. Introducing trust-weighted aggregation increases the complexity of model updates. Although our method is comparable to existing FL approaches in terms of communication efficiency, further research is needed to develop low-overhead trust aggregation techniques, such as compressing trust vectors or federated trust pruning strategies.

Regarding future improvement directions, several aspects are outlined. One is the integration of blockchain to achieve trust transparency. To further enhance trust management in crowdsensing, blockchain technology can be incorporated into FTM to ensure immutable and tamper-resistant trust records. By storing trust scores on a distributed ledger, transparency can be improved and collusion attacks mitigated. Another direction is adaptive trust learning based on reinforcement learning; future research can explore trust management based on reinforcement learning, enabling the system to learn optimal trust update strategies through interaction with the environment rather than relying on manually designed rules, thereby allowing trust dynamics to better adapt to continuously evolving adversarial behavior.

6 Conclusion

We proposed a Federated Trust Management (FTM) framework that mitigates malicious contributions to enhance the reliability of crowdsensing applications by integrating dynamic trust assessment into Federated Learning, achieving adaptive defense against adversarial participants. Experimental results show that compared to traditional FL aggregation methods, FTM significantly improves data reliability, model robustness, and adversarial resilience.

We demonstrate that integrating trust mechanisms into Federated Learning is crucial for applications in decentralized and adversarial environments. By introducing real-time trust adjustments, FTM provides a promising solution for enhancing the security and reliability of crowdsensing applications, such as in smart cities—where filtering unreliable data improves the accuracy of environmental and traffic monitoring systems—in healthcare, ensuring high-quality data aggregation in federated medical AI models; and in edge computing scenarios, enhancing the robustness of distributed learning based on IoT networks.

Although FTM performs well in mitigating adversarial contributions, several challenges remain. Scalability optimization is critical as the number of participants increases, requiring efficient trust computation and communication overhead management. Adaptive trust learning should be explored using reinforce-

ment learning (RL) techniques to enable the system to learn optimal trust update strategies in dynamic environments. Privacy-preserving trust management can be investigated using techniques such as secure multi-party computation (MPC) or differential privacy to further enhance privacy protection in trust assessment. Finally, integrating blockchain-based authentication mechanisms can help prevent common attacks in federated trust aggregation.

Acknowledgments

This work was partially supported by JSPS Grant-in-Aid for Scientific Research (C) 23K11103.

References

1. Zhang, T., Zhang, D., Zhang, P., Zhang, J., Tian, S.: A novel method of reliable data transmission for internet of vehicles based on crowd sensing strategy. *International Journal of Communication Systems* **37**(15), e5891 (2024)
2. Davidovic, B., Dejanovic, S., Davidovic, M.: A crowdsensing-based framework for sound and vibration data analysis in smart urban environments. *Smart Cities and Regional Development (SCRD) Journal* **8**(2), 39–46 (2024)
3. Wang, J., Zhao, D., Zhao, G.: Malicious participants and fake task detection incorporating gaussian bias. *ACM Transactions on Internet Technology* **24**(4), 1–19 (2024)
4. Li, J., Gu, B., Gong, S., Su, Z., Guizani, M.: Can we enhance the quality of mobile crowdsensing data without ground truth? *IEEE Transactions on Mobile Computing* (2025)
5. Owoh, N., Riley, J., Ashawa, M., Hosseinzadeh, S., Philip, A., Osamor, J.: An adaptive temporal convolutional network autoencoder for malicious data detection in mobile crowd sensing. *Sensors* **24**(7), 2353 (2024)
6. Truong, N.B., Lee, G.M., Um, T.W., Mackay, M.: Trust evaluation mechanism for user recruitment in mobile crowd-sensing in the internet of things. *IEEE Transactions on Information Forensics and Security* **14**(10), 2705–2719 (2019)
7. Banti, K., Louta, M., Baziana, P.: Data quality in human-centric sensing based next generation iot systems: A comprehensive survey of models, issues and challenges. *IEEE Open Journal of the Communications Society* (2023)
8. Yazdinejad, A., Dehghantanha, A., Karimipour, H., Srivastava, G., Parizi, R.M.: A robust privacy-preserving federated learning model against model poisoning attacks. *IEEE Transactions on Information Forensics and Security* (2024)
9. Liu, Z., Guo, J., Yang, W., Fan, J., Lam, K.Y., Zhao, J.: Privacy-preserving aggregation in federated learning: A survey. *IEEE Transactions on Big Data* (2022)
10. Dodda, S.B., Maruthi, S., Yellu, R.R., Thuniki, P., Reddy, S.R.B.: Federated learning for privacy-preserving collaborative ai: Exploring federated learning techniques for training ai models collaboratively while preserving data privacy. *Australian Journal of Machine Learning Research & Applications* **2**(1), 13–23 (2022)
11. Xie, X., Hu, C., Ren, H., Deng, J.: A survey on vulnerability of federated learning: A learning algorithm perspective. *Neurocomputing* p. 127225 (2024)
12. Xia, F., Cheng, W.: A survey on privacy-preserving federated learning against poisoning attacks. *Cluster Computing* **27**(10), 13,565–13,582 (2024)

13. Kaushal, V., Sharma, S.: Securing the collective intelligence: a comprehensive review of federated learning security attacks and defensive strategies. *Knowledge and Information Systems* pp. 1–39 (2025)
14. Xu, J., Zhang, Z., Hu, R.: Achieving byzantine-resilient federated learning via layer-adaptive sparsified model aggregation. *arXiv preprint arXiv:2409.01435* (2024)
15. Alsulaimawi, Z.: Federated learning with anomaly detection via gradient and reconstruction analysis. *arXiv preprint arXiv:2403.10000* (2024)
16. Deng, J., Wu, Q., Wang, S., Ye, J., Wang, P., Du, M.: A novel joint neural collaborative filtering incorporating rating reliability. *Information Sciences* **665**, 120,406 (2024)
17. Qi, P., Chiaro, D., Guzzo, A., Ianni, M., Fortino, G., Piccialli, F.: Model aggregation techniques in federated learning: A comprehensive survey. *Future Generation Computer Systems* **150**, 272–293 (2024)
18. Olateju, O., Okon, S.U., Igwenagu, U., Salami, A.A., Oladoyinbo, T.O., Olaniyi, O.O.: Combating the challenges of false positives in ai-driven anomaly detection systems and enhancing data security in the cloud. Available at SSRN 4859958 (2024)
19. Alghushairy, O., Alsini, R., Alhassan, Z., Alshdadi, A.A., Banjar, A., Yafaz, A., Ma, X.: An efficient support vector machine algorithm based network outlier detection system. *IEEE Access* (2024)
20. Darabseh, A., Faizan, M.: Outlier detection in wireless sensor networks using machine learning and statistical based approaches. *Revue d’Intelligence Artificielle* **38**(4) (2024)
21. Cerdà-Alabern, L., Iuhasz, G., Gemmi, G.: Anomaly detection for fault detection in wireless community networks using machine learning. *Computer Communications* **202**, 191–203 (2023)
22. Ali, W., Umer, K., Zhou, X., Shao, J.: Hidattack: An effective and undetectable model poisoning attack to federated recommenders. *IEEE Transactions on Knowledge and Data Engineering* (2024)
23. Xu, Z., Jiang, F., Niu, L., Jia, J., Li, B., Poovendran, R.: Ace: A model poisoning attack on contribution evaluation methods in federated learning. *arXiv preprint arXiv:2405.20975* (2024)
24. Zheng, J., Xu, J., Du, H., Niyato, D., Kang, J., Nie, J., Wang, Z.: Trust management of tiny federated learning in internet of unmanned aerial vehicles. *IEEE Internet of Things Journal* (2024)
25. Mlika, F., Karoui, W., Romdhane, L.B.: Blockchain solutions for trustworthy decentralization in social networks. *Computer Networks* p. 110336 (2024)
26. Zhang, M., Xiao, D.: Trustgat: Sparse trust data mining with graph attention for mobile social networks. In: *2023 24th IEEE International Conference on Mobile Data Management (MDM)*, pp. 21–29. IEEE (2023)
27. Vashishth, T.K., Sharma, V., Kumar, B., Sharma, K.K., Chaudhary, S., Panwar, R.: Blockchain for securing federated learning systems: Enhancing privacy and trust. *Model Optimization Methods for Efficient and Edge AI: Federated Learning Architectures, Frameworks and Applications* pp. 299–320 (2025)