Blockchain-Enabled Federated Learning to Enhance Security and Privacy in Internet of Medical Things (IoMT)

Zahra Eskandari[a*], Mohammad Rezaee[b]

Department of Computer Engineering, Quchan University of Technology, Quchan, Iran; z.eskandari@qiet.ac.ir[a], rezaee@qiet.ac.ir[b]

ABSTRACT

Federated learning is a distributed data analysis approach used in many IoT applications, including IoMT, due to its ability to provide acceptable accuracy and privacy. However, a critical issue with Federated learning is the poisoning attack, which has severe consequences on the accuracy of the global model caused by the server's lack of access to raw data. To deal with this problem effectively, a distributed federated learning approach involving blockchain technology is proposed. Using the consensus mechanism based on reputation-based verifier selection, verifiers are selected based on their honest participation in identifying compromised clients. This approach ensures that these clients are correctly identified and their attack is ineffective. The proposed detection mechanism can efficiently resist the data poisoning attack, which significantly improves the accuracy of the global model. Based on evaluation, the accuracy of the global model is compared with and without the proposed detection mechanism that varies with the percentage of poisonous clients and different values for the fraction of poisonous data. In addition to the stable accuracy range of nearly 93%, the accuracy of our proposed detection mechanism is not affected by the increase of α in different values of β.

Keywords— Blockchain, Consensus Algorithm, Federated Learning, Internet of Medical Things, Poisoning Attack.

1. Introduction

Federated learning is a distributed machine learning paradigm that trains an algorithm via multiple independent sessions, each using its dataset [1]. In this approach, nodes are distributed in the network, sense data, and send it to the regional edge node, where the local dataset is formed. However, sharing this information with the server questions the privacy of customers, which is an obstacle in the way of data analysis applications. Federated learning overcomes this challenge by aggregating local models instead of needing raw datasets to extract the global model, preserving confidentiality and providing sufficient accuracy. In IoMT diagnosis applications, where the privacy of patients is of vital importance, hospitals act as edge nodes that collect data from their devices and extract the local model without sending raw data to the server, then local models are sent to the server [2].

A critical issue with Federated learning is the lack of access to raw data; it is possible that devices or edge nodes may intentionally or even due to failure, produce the wrong data/model and submit them to the server, and consequently, the accuracy of the global model is reduced. This attack type is called the poisoning attack [3]. Erroneous data are produced in devices as clients, and the model is extracted with an error at the edge node (data poisoning attack). Additionally, it may be possible for the edge node, after receiving the data from the clients, to produce an incorrect model and send it to the server (model poisoning attack). As a result of aggregating these poisonous models, the global model is associated with errors.

In various attack techniques, the accuracy of the target model is sometimes disturbed or the classification in a certain category is accompanied by problems. In both ways, the attacker aims to prevent the global model from working correctly [4].

There are some approaches to deal with these problems such as clustering the received models in the server [5], checking received local models with validation or generated datasets [9]. In some works, such as [6, 7], aggregation is performed with robust techniques to minimize the effect of poisonous models. in some works, some clients are selected randomly and their datasets are used to audit the local models [8]. This improves the auditing performance, provided that all these auditing nodes are honest. Unfortunately, this is an incorrect assumption in a distributed trustless network where some nodes may be compromised and report incorrect information. Recently, the use of the blockchain [10] in the trustless environments of the Internet of Things has been widely discussed [11, 12, and 13].

In this article, we employ the concept of blockchain and its special features such as distribution and establishment of trust in a trustless environment, to effectively deal with poisoning attacks. To overcome the single point of failure and security risks associated with centralized server-based Federated learning, the authors propose a distributed Federated learning approach that involves blockchain technology. By using the consensus mechanism in blockchain, the authors attempt to defeat the poisoning attack, more specifically we focus here on label flipping attack. To do this, the miners in the consensus

process verify the local models with their datasets and vote for poisonous or non-poisonous behavior of the clients. After verification and voting, poisonous models are detected and eliminated in the aggregation process. To defend against the poisoning attack, malicious clients should be identified using reliable votes. To achieve this, a reputation-based verifier selection mechanism is applied. The main contributions of this paper are summarized as follows:

- Moving from centralized server-based federated learning to distributed blockchain-based federated learning.

- Overcoming poisoning attacks by recognizing poisonous clients using distributed datasets of the verifiers.

- Applying the reputation-based consensus mechanism to deal with compromised verifiers.

In the current work, blockchain-enabled federated learning refers to the use of blockchain to evaluate local models and detect models infected by poisoning attacks through consensus among verifiers, which is an important functionality of blockchain. As explained later, this goal has been achieved accurately, correctly, and sustainably under various relevant attack parameters. Therefore, one of the reasons for the importance of this research and its innovation in covering the gap in related works is the reliance on the inherent features of blockchain to ensure communication security and user privacy by applying anonymity in block production and using hash to prevent block data from being changed. This approach considers user privacy, which is one of the important and reliable factors in the IoMT discussion. In addition to not sharing raw user data and only sharing local models and evaluating them in a blockchain-based environment, blockchain-based features have been relied upon to improve security and privacy in IoMT applications.

Structure of the work: The rest of this paper is organized as follows: A review of initial preliminaries follows in section 2 and Section 3 reviews related works. We present the enhanced reputation-based DPOS consensus scheme to defeat poisoning attacks in Section 4 and finally, performance evaluation of the proposed approach is discussed in Section 5. Section 6 concludes the paper and discuss its future work.

2. Preliminaries

Here, the initial review of the preliminaries related to the paper is studied.

2.1. Blockchain and DPOS consensus algorithm

Blockchain is a distributed and trustless approach to agreement, with distributed processing and shared communication, and without a central authority. Its properties such as decentralization, anonymity, and untrustworthiness make it suitable for use in a distributed and trustless environment such as the Internet of Things (IoT) [11, 12, and 13]. Initially, Proof-of-Work (POW) was used as a consensus mechanism to find agreement between peers. However, this computation-intensive consensus scheme is not appropriate for resource-limited devices in IoT. As a result, Proof-of-Stake (POS) and Delegated Proof-of-Stake (DPOS) have been proposed to perform the consensus process on a group of miners with moderate cost by allowing high-stack

stockholders to vote for nodes to be selected as miners. However, there are problems such as collusion between these high-stack stockholders and compromised miners [6].

To overcome these problems, reputation-based DPOS has been proposed [17]. In this scheme, miners are selected based on how well they behave in voting rounds. Reputation is defined as the rating of an entity's trustworthiness by others based on its past behaviors. High reputation nodes are selected as miners, and they form a mining group. A block manager in a mining group generates the next block and distributes it to the miners, who participate in block verification. After that, the miners vote about each other's honesty or dishonesty, and the reputation parameter of miners is updated based on their behavior in the verification step.

2.2. Poisoning attacks

Malicious participants in a network apply the poisoning attack to generate malicious models in the form of targeted or untargeted attacks [4]. In targeted form, the attacker first selects some samples in the dataset with specific attributes, then assigns the wrong label to these samples, and finally trains the model. In this approach, the accuracy of the global model is maintained high. Backdoor attacks [14] are one prominent class of targeted poisoning attacks. On the other hand, in untargeted poisoning attacks such as label-flipping attacks [15], the wrong labels are assigned to some random samples with the purpose to reduce the overall accuracy of the global model. In addition to this division, poisoning attacks can be further divided into data poisoning, in which malicious participants manipulate local datasets using label flipping and other methods to affect global model accuracy, and model poisoning, in which random local models are generated by applying certain predefined rules [4].

3. Related Works

In this section, we initially review existing work in the fields of federated learning, specifically focusing on blockchain-based federated learning in the healthcare domain. Subsequently, we concentrate on the research goal of detecting poisoning attacks in federated learning and investigate the existing work in this area. By examining these issues and identifying the problems in these solutions, the research aim, which is to address these problems, will be clarified.

3.1. Federated Learning and Blockchain-based Federated Learning in Healthcare

Federated learning [1] is a method for training a model in a decentralized manner, utilizing the data and resources of multiple users. During each round, the server randomly selects a subset of clients, and each client is tasked with training the initial model on their local data, improving the model, and then sharing it with the server. Following the clients' contributions, the server generates an updated model, and the current global model is established at the conclusion of the round. Until now some works try to eliminate central role of the server and proposed distributed federated learning. Some works use blockchain features to do it such as consensus-driven federated learning [23, 27], cross-cluster blockchain-based federated learning (BFL) [28], blockchain-enabled federated learning (FL-Block) scheme [16] using a Proof-of-Work consensus mechanism.

Smart healthcare is an intelligent infrastructure that leverages advanced technologies like IoT, big data, cloud computing, and artificial intelligence to become more efficient [24]. However, these technologies also raise concerns about data security and privacy. As a result, the research community has focused on developing techniques to protect users' medical data from leakage and misuse. BlockFed [25] is one of the promising techniques that enable training machine learning models with high data security and privacy using a blockchain-based federated learning framework for collecting numerous data from various hospitals in a reliable way. In [26], the authors studies data privacy issues and the lack of high-quality training data sets in the internet of health things. They proposed a lightweight hybrid federated learning framework and used blockchain and smart contracts to manage edge training.

3.2. Poisoning attack Detection in Federated Leaning

In federated learning, where the server does not have direct access to participants' data, poisoning attacks are more likely to occur with increased complexity and impact. To prevent these attacks, various defense strategies have been proposed. Some approaches involve clustering techniques [5, 21] or artificial intelligence approaches [20] are conducted before aggregation. This is under the assumption that the majority of users are benign and only malicious participants are retained. Other approaches directly mitigate poisoning attacks during the training process [7, 22]. In [22] Instead of removing outliers from the training data, a trimmed optimization is deployed to make machine learning robust.

The majority of methods designed to detect and mitigate poisoning attacks rely on server-side validation datasets to evaluate the quality of received local models [9]. However, if the server-side validation dataset has a stable distribution, it may be problematic as it cannot detect all types of attacks. An alternative approach [8] is to use the datasets of other participants in the network instead of just the server-side validation dataset. This approach improves auditing performance, provided that all these auditing participants are honest. Unfortunately, this assumption is unrealistic in distributed networks without trust, as some participants may be compromised and report incorrect information. In [16], a blockchain-enabled federated learning scheme called FL-Block was proposed to address these issues. This scheme prevents poisoning attacks by replacing the central authority with a blockchain system that has a non-tempering feature. Poisoning attacks were prevented in this work by replacing the central authority with a novel blockchain system that has a non-tempering feature due to the nature of the blockchain. Because of the hash function utilized in blockchain, applying the poisoning attacks and altering the saved information was impossible. However, in this work, there is no attempt to detect and eliminate poisoning attacks.

The previously mentioned methods are limited in their ability to detect poisoned or compromised participants in untrusted Internet of Things environments. As a result, these approaches are ineffective in addressing this type of attack. Therefore, it is essential to propose a solution that relies on the use of various datasets in a reliable manner and is based on the consensus of supervising participants to identify poisoning attacks.

4. Proposed Approach

Federated learning has been widely used in the healthcare domain and the Internet of Medical Things (IoMT) due to its benefits in terms of remote access to data and user privacy. As shown in figure 1, here, each hospital as edge node, has its own devices and dataset of patients' information, and the local model can be extracted. Compared to standard Federated learning, in this work, to apply federated learning, a mining group is formed based on reputation values, and high reputation nodes are selected as miners. The clients send their local models to the block manager, who analyzes them based on its dataset and creates a block including poisonous or non-poisonous labels for each local model. This block is broadcast to miners, who verify it with their datasets and forward the verification results on the blockchain. The block manager receives all votes regarding the block and miners over a while. If a majority of miners verify the block, poisonous clients are identified, their local models are eliminated, and other non-poisonous models are aggregated and uploaded to the global model blockchain, a public ledger that records the global model after each training session. The reputation opinions of miners are calculated based on votes and included in a public ledger. Algorithm 1 contains a schematic of the steps for the work and the notations of the algorithm are described in Table 1.

As can be seen in the Figure 1, compared to standard federated learning, the addition of verifiers in the process of aggregating local models, makes it possible to identify poisoned models and infected clients. These models are removed from the aggregation process, and a high-accuracy global model is obtained in each round, resulting in the rapid convergence of the global model.

Step 1- System Initialization: Every entity in the network, including devices and edge nodes, to cooperate in the federated learning process should authenticate by a global Trust Authority (TA) as a legitimate entity. After successful authentication, the entity obtains its public and private keys and the corresponding certificates for signing, encryption, and decryption. As dealing with these security algorithms is outside the scope of the article, any standard asymmetric cryptography could be adopted for system initialization.

Step 2- Mining group formation: As stated in line 1 of the algorithm, a mining group is first formed based on the reputation of the nodes in which the nodes with the highest reputation are selected to form the mining group. In each round of training, a miner in the group is selected as the block manager, and others are considered verifiers.

Step 3- Local model uploading: In this step as stated in line 2 of the algorithm, some of the edge nodes from non-poisonous nodes are selected as clients to extract their local model from their datasets. As stated in line 3, these clients extract the local model and then send it to the block manager securely.

Step 4- Block generation: In DPOS schemes, the block manager is responsible for block generation, broadcasting, verification, and management during the consensus process. According to lines 4 to 6 of the algorithm, in a time slot, the block manager receives the local models from the clients, analyzes the models based on its dataset, and considers them as poisonous or non-poisonous clients. Then it applies anonymity and puts local models in a block without revealing
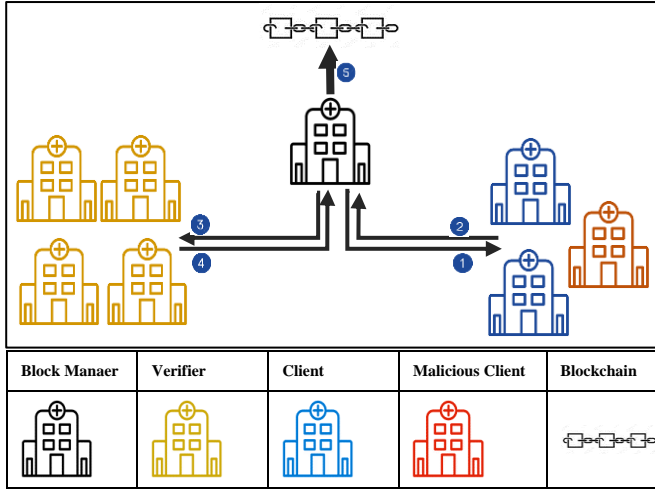
Figure. 1. Enhanced reputation-based DPOS consensus scheme. The procedure consists of: 1) sending global model to clients, 2) sending local models by clients, 3) generating and broadcasting block to the verifiers, 4) sending votes by verifiers, 5) eliminating poisonous and aggregating non-poisonous local models and uploading round's global model to the blockchain public ledger.

the owner's identity. Finally, the block manager generates an unverified block and broadcasts this block to verifiers.

Step 5- Consensus process: according to line 7 of the algorithm, selected verifiers audit the block locally with their dataset and broadcast their audit results with their signatures to each other in a distributed manner. After receiving the audit results, the verifiers compare their results to have a proper understanding of each other and send a reply message consisting of the verifier's audit result about the block and other verifiers to the block manager. The block manager analyzes the received replies; if more than half of the verifiers agree on the block, the detected clients as poisonous are added to the list of poisonous clients to avoid them in the next rounds. After the elimination of the poisonous models, the block manager aggregates the non-poisonous models and the global model is obtained. Finally, this block is formally stored in the global model blockchain. If the block is not verified, the round is repeated with the next block manager.

Step 6- Reputation update: The reputation is updated based on the votes each verifier gives to the others. After each round of the consensus process, the verifiers forward to the block manager their reputation opinions for the others based on their behavior in the round. Then the reputations of the verifiers are updated based on votes and included in a public ledger which can be used for the next consensus round.

It should be mentioned that here we assume the block manager is honest, otherwise the block is not verified by the verifiers so the global model is not updated and the round is repeated. As a result, the manager's reputation is decreased and it is automatically removed from the mining group next round.

5. Performance Evaluation

In this section, we evaluate the performance of the proposed approach. we first describe the dataset and the experiment setup, then evaluate the performance of the proposed approach.

Table 1. Notations

Notation	Definition
$N = \{N_i \mid i = 1 \cdots N\}$	set of clients
$LM = \{LM_i^r \mid i = 1 \cdots N\}$	set of local models at round r
$REP = \{rep_i^r \mid i = 1 \cdots N\}$	set of reputation values at round r
$GM = \{GM^i \mid i = 1 \cdots r - 1\}$	set of global models
r	current round

Algorithm 1: *Proposed* $rep - Dpos$ *scheme* (*in round r*)

Inputs:
$N = \{N_i \mid i = 1 \cdots N\}$
$LM = \{LM_i^r \mid i = 1 \cdots N\}$
$REP = \{rep_i^r \mid i = 1 \cdots N\}$
$GM = \{GM^i \mid i = 1 \cdots r - 1\}$
$P = $ set of poisonous clients
Output: P as an updated set of poisonous clients

1: $Manager.Miners_{group} = MG\_generation(N, REP)$
2: $Clients_{group} = C\_selection(N, P)$
3: $for\ each\ client\ in\ Clients_{group}$:
   $Send\ LM_{client}^r\ to\ Manager$
4: $LM^r = \{LM_{client}^r \mid client \in Clients_{group}\}$
5: $Block = Block\_generation(Manager, LM^r)$
6: $Broadcast\_by\_Manager(Block)$
7: $Votes = Consensus\_algorithm(Miner_{group}, Block)$
8: $if\ majority\_votes(Votes)$:
   $P = P \cup P_{detected}$
   $GM^r = agg(\{LM_{client}^r \mid LM_{client}^r \in LM^r\ and\ client\ !\in P\})$
   $update(GM)$
   $update(REP)$
9: $else$:
   $Repeat\ the\ round$

Experiment setup: Considering the importance of privacy in medical data and the reluctance of owners to share such information, and on the other hand, the need for a massive dataset to have a high-precision diagnostic tool, federated learning is very effective in healthcare applications [2]. Our evaluation is based on the Thyroid Disease Dataset [18], which is a public dataset that has been studied in some studies [19]. The dataset contains 3,163 instances; each instance has 18 features and a label, where an instance without thyroid disease and marked as -1, otherwise labeled as 1. To conduct the experiments, 90% of the data is divided equally among all nodes in the network, following the identical and independent scenario (IID), while the remaining 10% is reserved for model validation. In this study, N=30 nodes are considered in the network, with 10 of them acting as verifiers and each round of training, 10 of the remaining nodes randomly selected as clients. As the evaluation parameter, the effectiveness of the classification models is evaluated based on the accuracy metric, and the goal is to maximize it.

Here, we first conduct a simulation experiment on the impact of the label flipping-based poisoning attack on the federated learning accuracy, then we evaluate the proposed poisoning attack detection mechanism. All experiments are conducted using TensorFlow 2.10.0 on an Ubuntu 16.04 machine having an Intel(R) Core (TM) i7-9750H CPU and 16 GB RAM.

Results: In the experiment, we evaluate the label flipping attack under two parameters, α as the fraction of poisoned data of each participant and β as the fraction of malicious clients. Each poisonous participant clones the fraction α of the local dataset with flipped labels for injecting poisonous data into training, while β is the percentage of clients that are malicious.

We vary the α and β from 0.6 to 0.9 with intervals of 0.1 for the implementation of a label-flipping attack.

To demonstrate the drastic impact of poisoning attacks on the accuracy of global models, we first consider federated learning under label flipping attack without detection and calculate the accuracy of the global model. As shown in Table 2, the increase in the values of α and β parameters severely decrease the accuracy of the global model, indicating the high vulnerability of federated learning to this attack.

To have more clear observation, the accuracy values are demonstrated in Figure 2. It is observed that the larger the value of β, causes the lower value of the accuracy of the global model. Also, in all values of the β parameter, as shown in Fig. 1, by increasing the α value, the accuracy of the global model is decreased.

In continue, to evaluate the performance of the proposed detection mechanism, attack is applied under different values of α and β, and the detection mechanism is used to defeat the attack. As shown in Figure 3, the proposed detection approach can efficiently resist the data poisoning attack, which represents a significant improvement in the accuracy compared with the scenario without the detection of poisonous clients. In addition to the stable accuracy range, of nearly 93%, the accuracy of the proposed detection mechanism isn't affected by the increase of α in different values of β. This indicates that the proposed approach has been able to successfully and effectively overcome the poisoning attack.

Compared to our work, in [19] a random forest is used to detect poisoning attacks in federated learning on the Thyroid Disease Dataset [18]. This work considered 10 nodes in the network and applied the poisoning attack under various α and β parameters in the range of (0,1). The work achieved a detection accuracy of 84.3% to 97.4% across different various α and β values. However, as observed in Table 3, it is important to note that the increase in the accuracy of detecting poisoning attacks is justified by the presence of only 10 nodes in the network and the increase in the nodes' share of the original dataset. Unfortunately, with the increase in the various α and β values, the detection accuracy is decreased to 84.3%. In comparison, our proposed approach achieved an accuracy in the range of 93.18% to 93.99%, indicating more stable

performance. The reason for this is the use of verifiers' datasets to evaluate local models. Therefore, after detecting infected models and removing them from the training process in subsequent rounds, the convergence of the global model is accelerated.

It should be mentioned that all the results are averaged over 50 random scenarios. In each scenario for every α and β pair value, clients and miners are chosen randomly. Then under this random scenario, the impact of the detection of poisonous clients is analysed.

5. Conclusions and future works

In this article, we propose a distributed blockchain-based federated learning approach to effectively deal with poisoning attacks in federated learning applications. The proposed approach involves applying a reputation-based verifier selection technique to have trustful, using reliable votes collected via consensus mechanism, and finally detecting malicious clients. We evaluate the performance of the proposed approach in terms of the accuracy of the global model and the results indicate that the proposed approach has been able to successfully and effectively overcome the poisoning attack. Distributed voting is effective in overcoming poisoning attacks, and the proposed approach works well in the IID scenario, but for non-IID scenarios, it is limited in its ability to detect poisoned or compromised participants. Thus, solutions with a higher level of distribution are needed. This suggests an area for future work.
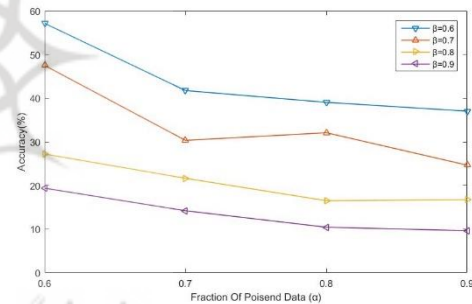

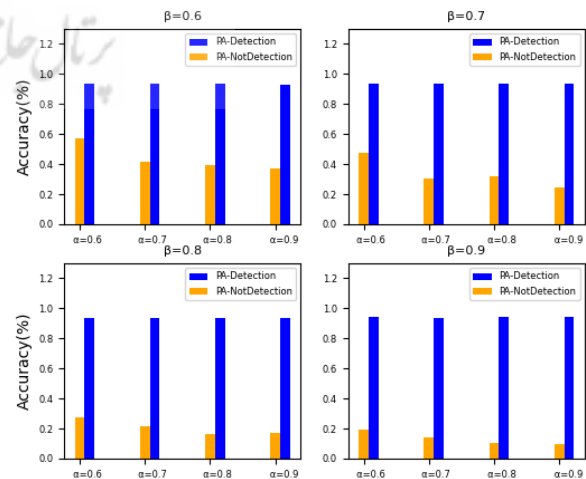Figure. 2. Accuracy of Global model with different values β.

Table 2. Accuracy of Global Model under different values of α and β.

Accuracy		β			
		0.6	0.7	0.8	0.9
α	0.6	57.14%	47.49%	27.19%	19.37%
	0.7	41.74%	30.32%	21.64%	14.20%
	0.8	39.04%	32.08%	16.49%	10.42%
	0.9	37.02%	24.65%	16.73%	9.62%

Table 3: Comparison between proposed approach and [19].

	Num of Nodes	Approach	Accuracy	Ref.
SFPA	10	Random Forest	84.3%-97.4%	[19]
Prop. App.	30	CNN	93%	Here

Figure. 3. Comparison of accuracy of federated learning with/without PA-detection under different values of α and β.

Declarations

References

[1] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges methods and future directions," *IEEE Signal Process. Mag.*, vol. 37, no. 3, pp. 50-60, May 2020. https://doi.org/10.1109/MSP.2020.2975749.

[2] D. C. Nguyen, Q. V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W. J. Hwang, "Federated Learning for Smart Healthcare: A Survey," *ACM Comput. Surv.*, vol. 55, no. 3, pp. 1-37, 2022. https://doi.org/10.1145/3501296

[3] O. Suciu, R. Marginean, Y. Kaya, H. Daume III, and T. Dumitras, "Wen does machine learning fail? Generalized transferability for evasion and poisoning attacks," *Proc. USENIX Conference on Security Symposium (USENIX Security '18)*, 2018 , pp. 1299–1316.

[4] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, and S. Zhao, "Advances and open problems in federated learning," *Foundations and Trends® in Machine Learning*, vol. 14, no.1–2, pp. 1-210, 2019. http://dx.doi.org/10.1561/2200000083

[5] N. Baracaldo, B. Chen, H. Ludwig, and J. A. Safavi, "Mitigating poisoning attacks on machine learning models: A. data provenance based approach," *Proc. ACM 10th ACM Workshop Artif. Intell. Secur.*, 2017, pp. 103–110. https://doi.org/10.1145/3128572.3140450

[6] D. Yin, Y. Chen, R. Kannan, and P. Bartlett, "Byzantine-robust distributed learning: Towards optimal statistical rates," *International Conference on Machine Learning*, 2018, pp. 5650–5659. https://proceedings.mlr.press/v80/yin18a.html.

[7] J. Zhang, C. Ge, F. Hu and B. Chen, "RobustFL: Robust Federated Learning Against Poisoning Attacks in Industrial IoT Systems," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 6388-6397, Sept. 2022, https://doi.org/10.1109/TII.2021.3132954

[8] S. Andreina, G. A. Marson, H. Mollering, and G. Karame, "BaFFLe: Backdoor detection via feedback-based federated learning," *2021 IEEE 41st International Conference on Distributed Computing Systems (ICDCS)*, DC, USA, 2021, pp. 852-863, https://doi.org/10.1109/ICDCS51616.2021.00086.

[9] Y Zh.ao, J. Chen, J. Zhang, D. Wu, M. Blumenstein, and S. Yu, "Detecting and mitigating poisoning attacks in federated learning using generative adversarial networks," *Concurr. Comput. Pract. Exp.* Vol. 34, no. 7, e5906, 2022. https://doi.org/10.1002/cpe.5906

[10] A. A. Monrat, O. Schelén and K. Andersson, "A Survey of Blockchain from the Perspectives of Applications, Challenges, and Opportunities," *IEEE Access*, vol. 7, pp. 117134-117151, 2019, https://doi.org/10.1109/ACCESS.2019.2936094.

[11] H. N. Dai, Z. Zheng and Y. Zhang, "Blockchain for Internet of Things: A Survey," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 8076-8094, Oct. 2019, https://doi.org/10.1109/JIOT.2019.2920987.

[12] X. Wang, X. Zha, W. Ni, R. P. Liu, Y. J. Guo, X. Niu, and K. Zheng, "Survey on blockchain for Internet of Things," *Computer Communications*, vol. 136, pp. 10-29, 2019, https://doi.org/10.1016/j.comcom.2019.01.006.

[13] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasubramanian, "A survey on the adoption of blockchain in IoT: challenges and solutions," *Blockchain: Research and Applications*, Vol. 2, no. 2, 100006, 2021. https://doi.org/10.1016/j.bcra.2021.100006.

[14] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to backdoor federated learning," *Proceedings of the 23th International Conference on Artificial Intelligence and Statistics*. PMLR, Palermo, Sicily, Italy, 2020, pp. 2938-2948. https://proceedings.mlr.press/v108/bagdasaryan20a.html.

[15] V. Tolpegin, S. Truex, M. E. Gursoy, G. Mehmet Emre, and L. Liu, "Data poisoning attacks against federated learning systems," *Computer Security – ESORICS 2020. ESORICS 2020,* Springer, Cham, 2020, pp. 480-50. https://doi.org/10.1007/978-3-030-58951-6_24

[16] Y. Qu, L. Gao, T. H., Luan, Y. Xiang, S. Yu, B. Li, G. Zheng, "Decentralized Privacy Using Blockchain-Enabled Federated Learning in Fog Computing," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5171-5183. https://doi.org/10.1109/JIOT.2020.2977383.

[17] M. T. de Oliveira, L. H. Reis, D. S. Medeiros, R. C. Carrano, S. D. Olabarriaga, and D. M. Mattos, "Blockchain reputation-based consensus: A scalable and resilient mechanism for distributed mistrusting applications," *Computer Networks*, vol. 179, p. 107367, 2020. https://doi.org/10.1016/j.comnet.2020.107367.

[18] http://www.kaggle.com/kumar012/hypothyroid.

[19] Z. Ma, J. Ma, Y. Miao, X. Liu, K. K. R. Choo, R. H. Deng, "Pocket Diagnosis: Secure Federated Learning Against Poisoning Attack in the Cloud," *IEEE Transactions on Services Computing*, vol. 15, no. 6, pp. 3429-3442, 2021. https://doi.org/10.1109/TSC.2021.3090771.

[20] W. Liu, H. Lin, X. Wang, J. Hu, G. Kaddoum, M. J. Piran, and A. Alamri, "D2MIF: A Malicious Model Detection Mechanism for Federated Learning Empowered Artificial Intelligence of Things," *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2141-2151, 2023. https://doi.org/10.1109/JIOT.2021.3081606.

[21] A. N. Bhagoji, S. Chakraborty, P. Mittal, and S. Calo, "Analyzing federated learning through an adversarial lens," *International Conference on Machine Learning,* 2018, pp. 634-643. https://proceedings.mlr.press/v97/bhagoji19a.html.

[22] M. Jagielski, A. Oprea, B. Biggio, C. Liu, C. Nita-Rotaru, and B. Li, "Manipulating machine learning: Poisoning attacks and countermeasures for regression learning," *Proc. IEEE Symposium on Security and Privacy (S&P'18),* IEEE, 2018, pp. 19–35. https://doi.org/10.1109/SP.2018.00057.

[23] L. Barbieri, S. Savazzi, M. Brambilla, and M. Nicoli, "Decentralized federated learning for extended sensing in 6G connected vehicles," *Vehicular Communications*, vol. 33, p. 100396, 2022. https://doi.org/10.1016/j.vehcom.2021.100396.

[24] D. Połap, G. Srivastava, and K.Yu, "Agent architecture of an intelligent medical system based on federated learning and blockchain technology," *Journal of Information Security and Applications (JISA)* vol. 58, p. 102748, 2021. https://doi.org/10.1016/j.jisa.2021.102748

[25] R. Kumar, A. A. Khan, J. Kumar, N. A. Golilarz, S. Zhang, Y. Ting, C. Zheng, W. Wang, "Blockchain-federated-learning and deep learning models for covid-19 detection using CT imaging, *IEEE Sensors Journal,* vol. 21, no. 14, pp. 16301–16314, 2021. https://doi.org/10.1109/JSEN.2021.3076767.

[26] M. A. Rahman, M. S. Hossain, M. S. Islam, N. A. Alrajeh, and G. Muhammad, "Secure and provenance enhanced internet of health things framework: A blockchain managed federated learning approach," *IEEE Access,* vol. 8, pp. 205071–205087, 2020. https://doi.org/10.1109/ACCESS.2020.3037474

[27] S. Savazzi, M. Nicoli and V. Rampa, "Federated Learning with Cooperating Devices: A Consensus Approach for Massive IoT Networks," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4641-4654, May 2020, https://doi.org/10.1109/JIOT.2020.2964162.

[28] H. Jin, X. Dai, J. Xiao, B. Li, H. Li and Y. Zhang, "Cross-Cluster Federated Learning and Blockchain for Internet of Medical Things," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15776-15784, 1 Nov.1, 2021, https://doi.org/10.1109/JIOT.2021.3081578.

Zahra Eskandari received the B.S degree in Computer Engineering from Kharazmi University, Tehran, Iran, in 2006. She received her M.S. and Ph.D. degrees in Computer Engineering from Ferdowsi University of Mashhad, Iran, in 2008 and 2020, respectively. She was with the cybersecurity section at DTU compute, Denmark as a visiting researcher from July 2016 to March 2017. She is a full-time Assistant-Professor in the Department of Computer Engineering at Quchan

University of Technology, Iran. Her research interests include
security in IoT, AIoT and Edge computing.

Mohammad Rezaee received the PhD degree in
Computer Engineering from Ferdowsi University of
Mashhad, Iran, in 2019. Currently, he is an assistant
professor at the Computer Engineering Department,
Quchan University of Technology. His research
interests include Smart Grid Communication, and
Optimization of Communication Networks.