

Federated learning security mechanisms for protecting sensitive data

Asraa A. Abd Al-Ameer^{1,2}, Wesam Sameer Bhaya³

¹Department of Information Networks, University of Babylon, Babel, Iraq

²Department of Mathematics, Al-Zahraa University for Women, Karbala, Iraq

³Department of Information Security, University of Babylon, Babel, Iraq

Article Info

Article history:

Received Sep 10, 2022

Revised Dec 16, 2022

Accepted Feb 1, 2023

Keywords:

Artificial intelligence

Distributed machine learning

Federated learning

Machine learning

Network security

Privacy

Security

ABSTRACT

One of the new trends in the field of artificial intelligence is federated learning (FL), which will have promising roles in many real-world applications due to the work characteristics of its architecture. The learning mechanism for this technique is based on making training in a distributed manner on the local data for each client using decentralized data, then collecting parameters for each local training and uploading it to the server, which in turn will send model updates to all clients to give the final learning result. To provide a broad study on FL from security and privacy aspects, this research paper introduces a general view of FL and its categories, most attacks that can befall it, the safety mechanisms used by existing works in attacks defense, enhancing the safety and privacy of FL whether in the transmission or collecting of data. Then, the usage of FL in network security by many research papers has been presented, and how good results were achieved, and finally a comparison has been made between these papers.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Asraa A. Abd Al-Ameer

Department of Information Networks, University of Babylon

Babel, Iraq

Email: asraaabdalhussien@student.uobabylon.edu.iq

1. INTRODUCTION

In the artificial intelligence field, for building intelligent applications, the elementary and basic requirement is data [1]. Consequently, without data, the training for any model cannot be performed. Big data had been developed, so the amount of it is no further the focus of attention [2]. However, data islands are the form in which the data exists [3]. One of the problems that have devoted a lot of thinking and attention to artificial intelligence is solving data islands. Process and model the data in a centralized manner is the direct solution to data islands [4], [5]. In a traditional machine learning (ML) pipeline, data is collected from different sources and stored in a central location [6]. By such data, the training is done to make a single machine learning model once all data is available at a center. This scheme is called centralized learning as long as the data should be moved from the users' devices to a central device in order to build and train the model [7].

One of the vulnerabilities of decentralized machine learning (DML) is the number of communication desired between the clients and the parameter server since through data transition, the data privacy protection be at its weakest. This means the more communications, the greater the chance of attacks [3]. A technological solution proposed by [8] called federated learning (FL) has newly appeared as a solution to address analogous matters. In FL which is considered a new research direction for artificial intelligence, make the training for its models shifted from the central server to the terminal equipment [2]. FL is a fledgling ML scheme that had been introduced by Google in 2016 to anticipate users text input with many mobile devices whose number reached tens of thousands whilst keeping data on devices [8], [9].

FL process is described as shown in Figure 1. Firstly, a generic global model is downloaded for each device for local training. Secondly, local data of diverse mobile devices will be used to enhance the local model by uploading it in an encryption mode to the cloud. Thirdly, the local models averaged update implemented in the cloud will be conveyed as a revised global model to the device. Finally, the previous procedures repeat until a certain desired performance for the model achieves or the final deadline arrives [10].

This paradigm is mainly presented for two reasons: i) the data may not be sufficiently available to centrally reside on the server-side because of the restrictions on it; and due to the number of devices is boundless, so the number of valuable resources that can be FL use it is so large, and ii) network asynchronous communication becomes usable and provides protection for data privacy by sending sensitive data to the server rather than using local data from the edge server where it prevents the leak of data during its transmission also network asynchronous communication has become an important role [11]. This technique development will solve the conflict between data sharing and data privacy [12]. FL is convenient for application when data are privacy sensitive [13]. Like in mobile devices [14] or industry applications [15] that data are not available to be aggregated with legal concern. The paper is organized as follows. In section 2, introduces the categorization of FL. Section 3 shows security and privacy aspects in FL like the attacks that infect it and defense techniques against these attacks besides security mechanisms used by researchers in data transmission and in aggregation phases of FL. Section 4 shows many research works and their results achieved after applying FL to provide security for the networks. Discussion is in section 5 and the last section, section 6, is the conclusion.

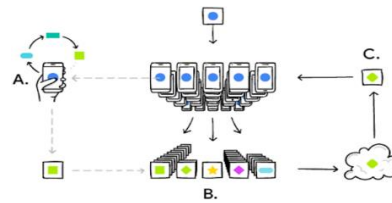


Figure 1. FL process flow [10]

2. CATEGORIZATION OF FEDERATED LEARNING

FL can be divided according to how data is distributed among various parties in feature space and in the sample ID into three categories. These categories are horizontally FL, vertically FL, and federated transfer learning (FTL) as shown in Figure 2 [4].

- Horizontal FL: is called as well sample-based FL. In this case, between the data features across different nodes, there is a certain amount of overlap, while there are quite differences in the sample space of the data as shown in Figure 2(a) [2], [16].
- Vertical FL: is called as well feature-based FL. This case is apropos when there are two data sets that differ in feature space but share the same sample ID space as shown in Figure 2(b) [2], [16].
- FTL: this category is implemented in the scenarios where two data sets vary in samples and in feature space as shown in Figure 2(c) [2], [17]. Limited shared sample sets are used to learn the shared representation between the two feature spaces and after that applied to achieve predictions for samples with only one-side features [4].

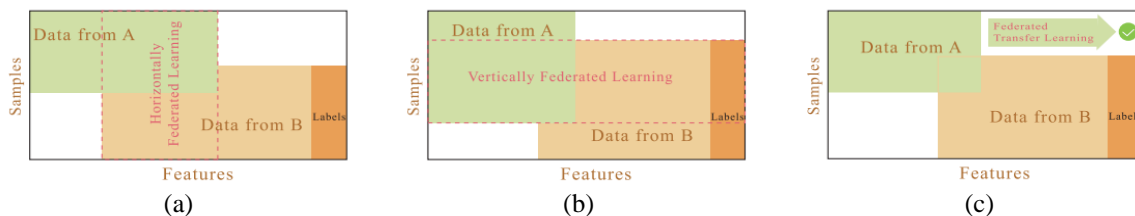


Figure 2. Categorizations of FL (a) horizontal FL, (b) vertical FL, and (c) FTL [2]

3. SECURITY AND PRIVACY IN FEDERATED LEARNING

FL is an enhanced form of distributed ML that offloads the operations that the central server usually performed [18]. Security guarantees and additional privacy are the greatest advantages that FL affords [19].

These advantages make FL useful and have an alluring rule, especially since information theft and data breaches are serious and common threats [16]. This section outlines the last developments in security and data privacy for FL. Current attack models, defense techniques against these attacks, and the various mechanisms used to provide security and privacy have been introduced.

3.1. Attacks on federated learning

FL is a goal for different attacks that intend to manipulate the collaborative learning process [3]. To manipulate and make a change in the global model and to take control of one or more participants, malicious agent benefits from the vulnerabilities [20], [21]. In such a scheme, the attacker targets to access the local data of the various clients, hyper parameters, updated weights in transit, or training procedures [22] in order to launch attacks on the global model and manipulate it. Below some of these attacks have been described.

- a. Poisoning attack: in this type of attack which has a high probability of occurrence in FL, the potential of tampered data weights and adding it to the global ML model is very high and that is for the possibility of any client in FL to access the training data [11], [23]. Poisoning can occur at the time of the training step and can impact either the local model or the training dataset and tamper with the global ML model performance in an indirect manner [16], [24].
- b. Inference attacks: this type of attack is highly like poisoning attacks because the chance of inference attacks is a great from either malicious centralized server or the participants in the FL process. It is considered one of the biggest threats on privacy [11].
- c. Backdoor attacks: is a manner to inject into the existing model a malicious task while maintaining the actual task accuracy [25]. To advance backdoor functionality into the joint model, the malicious participant can use model replacement [22]. Backdoor attacks have the ability to affect ML models by confusing them and forecasting false positives. Moreover, the intensity of backdoor attacks is high because defining attack occurrence takes a long time [24].
- d. Eavesdropping: in the learning process of FL and from clients to the central server, there is an iteration in comprising rounds of communication. Therefore, data can extract by an eavesdropping attacker on a communication channel if a weak channel exists [11].

3.2. Defense techniques for federated learning attacks

Some of defense techniques for FL that have been used to mitigate these types of attacks:

- a. Poisoning attacks in FL can be detected using various anomaly detection techniques. To identify events not match the expected activity or pattern, analytical and statistical methods can be used [7]. Euclidean Distance has been used by [26], to provide a model to detect the aberration in each input parameter of a client is used. Shen *et al.* [27] proposed a defense technique against malicious updates for a client by using clustering on all client updates before the aggregation phase. This method has proven to be beneficial in malicious client updates detection. Autoencoders have been proposed in [28] to prepare anomaly detection defense that helps in identifying malicious local model updates
- b. To mitigate inference attacks one of the model compression techniques can be used like knowledge distillation where knowledge is transferred by a fully trained neural network to a small model gradually on what needs to be done. In training a model the computational costs [29] can be saved through knowledge distillation technique [30]. A federated model distillation has been proposed by [31] in order to provide resilience in using personalized ML models besides using translators to aggregate knowledge that will be shared with each client
- c. One of the problems that appear in FL is when there are large-sized deep neural networks that should be used to train FL environment. A pruning technique has been proposed by [32], to address these issues where it minimizes ML model size to enhance the accuracy and reduce the complexity. Because in this approach it's not needed to share the full-fledged model. A pruning technique helps in identifying communication jams and backdoor attacks more efficiently
- d. Moving target defense is one of the methods used to develop strategies and diverse mechanisms that change continually to boost the complexity for attackers. It's the best type for intrusion protection at the server-level, network-level, and application level. Also, is a defense structure constructed to vague the source of vulnerability from the attackers. Colbaugh and Glass [33] explained the defense of the network-level moving target to prevent eavesdropping-based attacks.

3.3. Security mechanisms in federated learning

3.3.1. Data transmission security mechanisms

One of the most important things is the security of the data. Since the FL technique sends data between the local and the global model, securing mechanisms should be provided. Security mechanisms can be used with cryptographic protocols and algorithms.

- a. Secret sharing schemes: secret sharing schemes [34] were used by [35], to minimize data leakage risk on the server-side and to assure participants' security. Therefore, it's a good choice to provide security and protection for client updates in FL because they can be partition into diverse shares, which helps with the vulnerability concerned with the communication.
- b. Secure multiparty computation: multiple clients updates that were compiled by aggregators in FL contain sensitive information, Therefore, it is necessary to protect them. Multiparty computation schemes very appropriate approach to provide protection for the aggregation process and clients' updates [3], [36].
- c. Homomorphic encryption: homomorphic encryption is an encryption approach that does cipher-text complex mathematical operations without changing the encryption nature. From a security aspect in FL, client updates should not be decrypted by a central server and should be collected by it using only cipher-text. Homomorphic encryption can be provided to meet all these requirements [3]. Paillier federated multi-layer perceptron (PFMLP) has been proposed by [9] which is based on FL and partially homomorphic encryption. The basic idea is all learning parties just transmit the encrypted gradients by homomorphic encryption [37].

3.3.2. Secure aggregation in federated learning

Secure aggregation is a branch of multiparty computation algorithms where a collection of parties hold sensitive information and do not trust each other and in order to calculate an aggregated value, it should collaborate [38]. Any party's information should not reveal by the aggregated value. In order to guarantee a secure transit process, the clients outputs before shared should be encrypted [24].

- a. Federated secure aggregation protocol: a secure aggregation protocol proposed by [39], for FL to preserve the gradients privacy of clients model and assured that the users learn nothing while the server learns the clients inputs sum only. For practical applications [40], advanced a full version of the protocol. The clients raw input is masked through a random number to prevent direct detection to the central server, and each client arises a private-public key pair for each phase of the aggregation process, and all the clients are allowed to couple every other client's public key and its private key and, to generate a private shared key with a hash function [3].
- b. Blockchain FL: blockchain can be used in FL to decentralize the global aggregation process by permissive the blockchain network to exchange the updates of the client' local model while verifying them [41]. Blockchain is suitable and useful to protect the individual local model updates from being disclosed and verify the validity of these updates [16]. To provide secure aggregation in FL, a blockchained FL architecture has been proposed by [42].

4. NETWORK SECURITY USING FEDERATED LEARNING

One of the most important things in computer networks is network security. So its important to improve and ensure networks' safety by trained equipment to avoid the probability of any errors in the machines that organize security [22], [43]. ML has the ability to differentiate between benign data and malicious one in the network. So, it can be used to better analyze preceding cyber attacks and enhance proper security response [44]. FL is one of the learning techniques where each collaborator train a global model cooperatively [45]. Man *et al.* [17] have been proposed federated convolutional neural network (FedACNN), an intelligent intrusion detection mechanism that used CNN deep learning model through the mechanism of FL to complete the task of intrusion detection. Local datasets have been used with computing resources of edge devices for making training for the model and uploading the parameters of the model to a central server for collaborative training. Unlike traditional centralized learning approaches FedACNN does not need to transfer the raw data to a central server and that will reduce data leakage risk and assured model accuracy. By applying FedACNN on the network security layer-knowledge discovery in database (NSL-KDD) dataset, noticed that the detection results have better classification accuracy for attack data can reach 99.76%.

Federated distributed integrated clinical environment (FedDICE) has been proposed in [46]. To implement collaborative learning, mitigation, and detection of ransomware attacks it integrates FL privacy-preserving learning into software defined networks (SDN-oriented security architecture. The results have shown that after applying FedDICE on a clinical environment network traffic dataset it effectively detects ransomware spread detection with a testing accuracy of around 99% in the DICE. FL combined with fog computing in federated learning empowered mitigation architecture (FLEAM) which was proposed by [45], to enhance detection accuracy and minimize mitigation time, enabling defenders to jointly clash botnets, thus consolidating the IoT security. The results have shown that FL improves the accuracy of detection up to approximately 95%.

A two-step learning method called NAFT has been proposed by [47] which based on FL and transfer learning to handle the problem of data scarcity in network anomaly detection. In the first step, a people or

organization, that aims to build a network anomaly detection model Involved in FL in order to drift basic knowledge from other participants. Fine-tune the global detection model after FL experiments had been conducted on the UNSW-NB15 dataset shows that NAFT can accomplish a better anomaly detection performance than other methods when training data is scarce where the accuracy of NAFT reached higher than 90%.

Research by Zhao *et al.* [48], that is based on FL aided long short-term memory framework has proposed an effective independent and identically distributed (IID) method. First, at all user servers the initial long short term memory (LSTM) global model is deployed. Second, every single model is trained for each user, and then its model parameters are uploaded to a central server. Finally, model parameters aggregation is performed to construct a new global model and distribute it to user servers. The proposed method FL-LSTM shows that after applying it to SEA dataset it can detect intrusion and has a higher accuracy detection up to 99.21%.

Blockchain has been supported by [49] to solve issues in fog computing like data privacy. By a comprehensive verification, hybrid identity generation, and off-chain data storage and retrieve, decentralized privacy protection will be enabled by FL-block, while avoiding single-point failure. Besides, poisoning attacks could be defeated from fog servers aspect. After applying FL-block on two datasets the results showed good performances in privacy protection.

In computer networks, threat detection is one of the basic things in cybersecurity that is addressed by [50]. Using community model sharing with a streaming analytic pipeline, they presented an architectural approach. The models train gradually through their streaming scheme, as every log record is processed, thus, adapting to the drift concepts resulting from changing attacks. In addition, the approach of community sharing has been designed to federate learning by combining models without requiring sensitive cyber-log data sharing. Therefore, they provided for the operators of network security the capability to manage the events of cyber threats and the sensitivity of the model through analytic method weighting and community members in the best way suited for their available data and resources. Internal testing for their results indicates the usefulness of their approach.

5. DISCUSSION

This paper reviewed a survey on the FL technique. It explained how different researchers used it to secure the network from many attacks. From the previous explained papers it has been explained that FL enhanced the privacy of data, improved the detection and mitigation of different types of attacks by combining it with other techniques or by using it in different environments as shown in Table 1.

Table 1. FL defense techniques in networking

Reference	Work Description	Methods	Performance measure (%)	Dataset
[17]	Completes the intrusion detection task as an intelligent intrusion detection mechanism	CNN+FL	Accuracy=99.76	NSL-KDD
[46]	Enable collaborative learning, mitigation, and detection of ransomware attacks	FL+SDN-oriented security architecture	Accuracy=99	Clinical environment network traffic dataset
[45]	Boots the accuracy of detection, enabling defenders to jointly combat botnets	FL+fog computing	Accuracy=95	NA
[47]	Provide network anomaly detection that deal with the data scarcity problem	FL+transfer learning	Accuracy=90	UNSW-NB15
[48]	Offered intrusion detection	FL+LSTM	Accuracy=99.21	SEA
[49]	Solve the identified issues in fog computing like data privacy	FL+ blockchain + fog computing + distributed hash table	Accuracy=75	CIFAR-10
[50]	Provided for the operators of network security the capability to manage the events of cyber threats and the sensitivity of the model	FL+streaming architecture	Accuracy=93 NA	Fashion-MINIST Raw HTTP data logs
[51]	Detecting compromised devices in IoT networks	FL	FPR=0 TPR=94.07	Activity, deployment and attack datasets

6. CONCLUSION

FL has been introduced to protect sensitive data on many platforms and had used by many researchers for different applications. In this work, FL technique and its types have been reviewed, the most prominent attacks it may be exposed to, and the defense methods used against these attacks. This paper also

provides a comprehensive study on various solutions that researchers used to offer security mechanisms for FL and finally clarifies multiple methods that researchers have done to provide security and privacy in the networks field. After discussing the results of the researchers' works, it was shown how this technology greatly enhanced the accuracy of detecting attacks and improving the privacy and security of the network.




REFERENCES

- [1] Z. Li, V. Sharma, and S. P. Mohanty, "Preserving data privacy via federated learning: challenges and solutions," *IEEE Consumer Electronics Magazine*, vol. 9, no. 3, pp. 8–16, 2020, doi: 10.1109/MCE.2019.2959108.
- [2] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowledge-Based Systems*, vol. 216, pp. 1–11, 2021, doi: 10.1016/j.knosys.2021.106775.
- [3] S. Shen, T. Zhu, D. Wu, W. Wang, and W. Zhou, "From distributed machine learning to federated learning: in the view of data privacy and security," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 16, pp. 1–19, 2022, doi: 10.1002/cpe.6002.
- [4] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, pp. 1–19, 2019, doi: 10.1145/3298981.
- [5] S. Abdulrahman, H. Tout, H. O. -Slimane, A. Mourad, C. Talhi, and M. Guizani, "A survey on federated learning: the journey from centralized to distributed on-site learning and beyond," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5476–5497, 2021, doi: 10.1109/JIOT.2020.3030072.
- [6] A. A. Al-Ameer, G. A. Hussien, and H. A. Al Ameri, "Lung cancer detection using image processing and deep learning," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 28, no. 2, pp. 987–993, 2022, doi: 10.11591/ijeecs.v28.i2.pp987-993.
- [7] Y. Jin, X. Wei, Y. Liu, and Q. Yang, "A survey towards federated semi-supervised learning," *Computer Science-Machine Learning*, pp. 1–7, 2020.
- [8] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017, vol. 54, p. 10.
- [9] H. Fang and Q. Qian, "Privacy preserving machine learning with homomorphic encryption and federated learning," *Future Internet*, vol. 13, no. 4, pp. 1–20, 2021, doi: 10.3390/fi13040094.
- [10] B. McMahan and D. Ramage, "Federated learning: collaborative machine learning without centralized training data," *Research Scientists*, vol. 2017.
- [11] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021, doi: 10.1016/j.future.2020.10.007.
- [12] A. R. Javed et al., "Integration of blockchain technology and federated learning in vehicular (IoT) networks: a comprehensive survey," *Sensors*, vol. 22, no. 12, pp. 1–24, 2022, doi: 10.3390/s22124394.
- [13] M. S. Jere, T. Farnan, and F. Koushanfar, "A taxonomy of attacks on federated learning," *IEEE Security & Privacy*, vol. 19, no. 2, pp. 20–28, 2021, doi: 10.1109/MSEC.2020.3039941.
- [14] M. Chen, R. Mathews, T. Ouyang, and F. Beaufays, "Federated learning of out-of-vocabulary words," *Computer Science-Computation and Language*, pp. 1–6, 2019.
- [15] M. Chen, O. Semiari, W. Saad, X. Liu, and C. Yin, "Federated echo state learning for minimizing breaks in presence in wireless virtual reality networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 177–191, 2020, doi: 10.1109/TWC.2019.2942929.
- [16] O. A. Wahab, A. Mourad, H. Otkrok, and T. Taleb, "Federated machine learning: survey, multi-level classification, desirable criteria and future directions in communication and networking systems," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1342–1397, 2021, doi: 10.1109/COMST.2021.3058573.
- [17] D. Man, F. Zeng, W. Yang, M. Yu, J. Lv, and Y. Wang, "Intelligent intrusion detection based on federated learning for edge-assisted Internet of Things," *Security and Communication Networks*, vol. 2021, pp. 1–11, 2021, doi: 10.1155/2021/9361348.
- [18] Y. Guo, Z. Zhao, K. He, S. Lai, J. Xia, and L. Fan, "Efficient and flexible management for industrial Internet of Things: a federated learning approach," *Computer Networks*, vol. 192, p. 108122, 2021, doi: 10.1016/j.comnet.2021.108122.
- [19] M. Wazze, H. O. -Slimane, C. Talhi, A. Mourad, and M. Guizani, "Privacy-preserving continuous authentication for mobile and IoT systems using warmup-based federated learning," *IEEE Network*, pp. 1–7, 2022, doi: 10.1109/MNET.121.2200099.
- [20] J. Men et al., "Finding sands in the eyes: vulnerabilities discovery in IoT with EUFuzzer on human machine interface," *IEEE Access*, vol. 7, pp. 103751–103759, 2019, doi: 10.1109/ACCESS.2019.2931061.
- [21] N. R. -Barroso, D. J. -López, M. V. Luzón, F. Herrera, and E. M. -Cámara, "Survey on federated learning threats: concepts, taxonomy on attacks and defences, experimental study and challenges," *Information Fusion*, vol. 90, pp. 148–173, 2023, doi: 10.1016/j.inffus.2022.09.011.
- [22] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to backdoor federated learning," in *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, 2018, vol. 108, pp. 2938–2948.
- [23] V. Shejwalkar, A. Houmansadr, P. Kairouz, and D. Ramage, "Back to the drawing board: a critical evaluation of poisoning attacks on production federated learning," in *2022 IEEE Symposium on Security and Privacy (SP)*, 2022, pp. 1354–1371, doi: 10.1109/SP46214.2022.9833647.
- [24] M. Goldblum et al., "Dataset security for machine learning: data poisoning, backdoor attacks, and defenses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 2, pp. 1563–1580, 2023, doi: 10.1109/TPAMI.2022.3162397.
- [25] C. Xie, M. Chen, P.-Y. Chen, and B. Li, "CRFL: certifiably robust federated learning against backdoor attacks," in *Proceedings of the 38th International Conference on Machine Learning*, 2021, pp. 11372–11382.
- [26] P. Blanchard, E. M. E. Mhamdi, R. Guerraoui, and J. Stainer, "Machine learning with adversaries: byzantine tolerant gradient descent," in *Advances in Neural Information Processing Systems*, 2017, vol. 30, pp. 1–11.
- [27] S. Shen, S. Tople, and P. Saxena, "AUROR: defending against poisoning attacks in collaborative deep learning systems," in *Proceedings of the 32nd Annual Conference on Computer Security Applications*, 2016, pp. 508–519, doi: 10.1145/2991079.2991125.
- [28] M. Fang, X. Cao, J. Jia, and N. Z. Gong, "Local model poisoning attacks to byzantine-robust federated learning," in *Proceedings of the 29th USENIX Security Symposium*, 2020, pp. 1623–1640.
- [29] X. Wu, Y. Zhang, M. Shi, P. Li, R. Li, and N. N. Xiong, "An adaptive federated learning scheme with differential privacy preserving," *Future Generation Computer Systems*, vol. 127, pp. 362–372, 2022, doi: 10.1016/j.future.2021.09.015.
- [30] A. Gajbhiye et al., "Knowledge distillation for quality estimation," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 2021, pp. 5091–5099, doi: 10.18653/v1/2021.findings-acl.452.
- [31] D. Li and J. Wang, "FedMD: heterogenous federated learning via model distillation," *Computer Science-Machine Learning*, pp. 1–8, 2019.
- [32] K. Liu, B. D. -Gavitt, and S. Garg, "Fine-pruning: defending against backdooring attacks on deep neural networks," in *Research*




- in Attacks, Intrusions, and Defenses*, Cham: Springer, 2018, pp. 273–294, doi: 10.1007/978-3-030-00470-5_13.
- [33] R. Colbaugh and K. Glass, "Moving target defense for adaptive adversaries," in *2013 IEEE International Conference on Intelligence and Security Informatics*, 2013, pp. 50–55, doi: 10.1109/ISI.2013.6578785.
- [34] G. R. Blakley, "Safeguarding cryptographic keys," in *Managing Requirements Knowledge, International Workshop on*, 1979, p. 313, doi: 10.1109/AFIPS.1979.98.
- [35] H. Shi, Y. Jiang, H. Yu, Y. Xu, and L. Cui, "MVFLS: multi-participant vertical federated learning based on secret sharing," *The Federate Learning*, pp. 1–9, 2022.
- [36] V. Mugunthan, A. Polychroniadou, D. Byrd, and T. H. Balch, "SMPAI: secure multi-party computation for federated learning," in *Proceedings of the NeurIPS 2019 Workshop on Robust AI in Financial Services*, 2019, pp. 1–9.
- [37] F. Wibawa, F. O. Catak, M. Kuzlu, S. Sarp, and U. Cali, "Homomorphic encryption and federated learning based privacy-preserving CNN training: covid-19 detection use-case," in *EICC 2022: Proceedings of the European Interdisciplinary Cybersecurity Conference*, 2022, pp. 85–90, doi: 10.1145/3528580.3532845.
- [38] J. Guo, Z. Liu, K.-Y. Lam, J. Zhao, Y. Chen, and C. Xing, "Secure weighted aggregation for federated learning," *Computer Science-Cryptography and Security*, pp. 1–18, 2020.
- [39] K. Bonawitz *et al.*, "Practical secure aggregation for federated learning on user-held data," *Computer Science-Cryptography and Security*, pp. 1–5, 2016.
- [40] K. Bonawitz *et al.*, "Practical secure aggregation for privacy-preserving machine learning," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191, doi: 10.1145/3133956.3133982.
- [41] S. R. Pokhrel and J. Choi, "Federated learning with blockchain for autonomous vehicles: analysis and design challenges," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4734–4746, 2020, doi: 10.1109/TCOMM.2020.2990686.
- [42] H. Kim, J. Park, M. Bennis, and S.-L. Kim, "Blockchained on-device federated learning," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1279–1283, 2020, doi: 10.1109/LCOMM.2019.2921755.
- [43] Z. D. A. and M. A. A. Ugli, "Network security issues and effective protection against network attacks," *Bulletin of Science and Practice*, vol. 7, no. 9, pp. 479–485, 2021, doi: 10.33619/2414-2948/70/45.
- [44] S. Li, Y. Cheng, W. Wang, Y. Liu, and T. Chen, "Learning to detect malicious clients for robust federated learning," *Computer Science-Machine Learning*, pp. 1–7, 2020.
- [45] J. Li, L. Lyu, X. Liu, X. Zhang, and X. Lyu, "FLEAM: a federated learning empowered architecture to mitigate DDoS in industrial IoT," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 4059–4068, 2022, doi: 10.1109/TII.2021.3088938.
- [46] C. Thapa, K. K. Karmakar, A. H. Celdran, S. Camtepe, V. Varadharajan, and S. Nepal, "FedDICE: a ransomware spread detection in a distributed integrated clinical environment using federated learning and SDN based mitigation," in *Quality, Reliability, Security and Robustness in Heterogeneous Systems*, Cham: Springer, 2021, pp. 3–24, doi: 10.1007/978-3-030-91424-0_1.
- [47] Y. Zhao, J. Chen, Q. Guo, J. Teng, and D. Wu, "Network anomaly detection using federated learning and transfer learning," in *Security and Privacy in Digital Economy*, Cham: Springer, 2020, pp. 219–231, doi: 10.1007/978-981-15-9129-7_16.
- [48] R. Zhao, Y. Yin, Y. Shi, and Z. Xue, "Intelligent intrusion detection based on federated learning aided long short-term memory," *Physical Communication*, vol. 42, pp. 1–10, 2020, doi: 10.1016/j.phycom.2020.101157.
- [49] Y. Qu *et al.*, "Decentralized privacy using blockchain-enabled federated learning in fog computing," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5171–5183, 2020, doi: 10.1109/JIOT.2020.2977383.
- [50] F. W. Bentrem, M. A. Corsello, and J. J. Palm, "Leveraging sharing communities to achieve federated learning for cybersecurity," *Computer Science-Cryptography and Security*, pp. 1–7, 2021.
- [51] T. D. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A.-R. Sadeghi, "D²IoT: a federated self-learning anomaly detection system for IoT," in *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, 2019, pp. 756–767, doi: 10.1109/ICDCS.2019.00080.

BIOGRAPHIES OF AUTHORS



Asraa A. Abd Al-Ameer    is currently an Assistant Lecturer at Al-Zahraa University for Women, Karbala, Iraq. She received her Bachelor's degree in Information Technology from University of Babylon, Babylon, Iraq, in 2016, and her Master's degree in Information Technology from University of Babylon, Babylon, Iraq, in 2019. She is currently a Ph.D student at the department of information network, college of information technology, University of Babylon, Babylon, Iraq. Her current research is focused on aspects that include SDN, security, network security, privacy, machine learning, and deep learning. She can be contacted at email: asraaabdalhussien@student.uobabylon.edu.iq.



Wesam Sameer Bhaya    Ph.D. in Computer Science. He is a Professor in the Faculty of Information Technology, Information Security department, University of Babylon, Iraq. The current areas of interest are information networking and information security. He can be contacted at email: wesambhaya@uobabylon.edu.iq.