China Petroleum Processing and Petrochemical Technology

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

PRIVACY-PRESERVING FEDERATED LEARNING FOR SECURE IOT DATA PROCESSING AT THE EDGE

R. Saranya

Research Scholar, Department of Computer Science and Engineering, Annamalai University

Annamalainagar, Chidambaram - 608002

Dr. R. Saminathan

Associate Professor, Department of Computer Science and Engineering, Annamalai University

Annamalainagar, Chidambaram - 608002

Dr. N. Palanivel

Associate Professor, Department of Computer Science and Engineering, Manakula Vinayagar Institute of Technology, Puducherry

ABSTRACT

This research aims to advance the frontier of secure and privacy-centric data processing at the edge of the ever-evolving landscape of the Internet of Things (IoT). We concentrate on strategically implementing cutting-edge, privacy-preserving federated learning techniques; this forges a transformative path towards increased security and collaborative intelligence. In order to address these issues effectively, our paper puts forth an innovative solution: we suggest securing data processing at the edge via strategic implementation, specifically using the Federated Averaging (FedAvg) algorithm that's optimized for IoT with differential privacy and homomorphic encryption. Our study, orchestrating privacy-preserving collaborative intelligence through FedAvg, quantifies the efficiency of FedAvg-based techniques in a simulated IoT environment. We experiment with hyperparameter configurations and assess non-IID data distribution scenarios to measure resilience to communication latencies. This research is reshaping how we perceive secure data processing at the edge; it propels us towards an envisioned future where privacy, security, and collaborative intelligence seamlessly converge, empowering IoT devices within this trusted connected world.

Keywords: IoT, Edge Learning, Federated Learning, Orchestration, IoT automation.

1. INTRODUCTION

Interconnected devices within the Internet of Things (IoT) have ushered in an era of unprecedented data generation and exchange, transforming the landscape of modern technology. However, this surge in connectivity has also raised profound concerns about the privacy and security of sensitive information traversing through these vast networks. As we navigate this complex intersection of innovation and apprehension, the imperative to develop advanced techniques for securing IoT data processing at the edge becomes paramount [2] [35]. IoT devices, with their diverse sensors and embedded computational capabilities, serve as ideal candidates for decentralized learning. Federated learning brings machine learning algorithms directly to the edge, allowing devices to autonomously process and refine models based on locally generated data [1]. This decentralized approach not only minimizes latency but also alleviates the need for constant communication with a central server. Hence, this research embarks on an exploration of a cutting-edge paradigm, namely, Privacy-Preserving Federated Learning, as a strategic approach to address the intricate challenges associated with IoT data processing [4] [7]. At the nexus of privacy and collaboration [6], our investigation endeavors to seamlessly integrate federated learning methodologies with advanced encryption techniques. Further, we aim not only to safeguard the privacy of sensitive information but also to establish a robust foundation for secure collaborative artificial intelligence (AI) model training.

By decentralizing the learning process, this methodology redefines the contours of collaborative intelligence, allowing devices to collectively train models without relying on centralized data storage [8]. The information that ensues resonates across devices, ensuring that no single entity holds the key to sensitive data, thus addressing the paramount concerns of data exposure and potential security breaches [9–13]. [17]. Here, we delve into the rationale behind our emphasis on privacy-preserving federated learning, elucidating its pivotal role in mitigating the inherent vulnerabilities of traditional approaches.

1.1 Federated Learning: An Overview

Federated Learning (FL) emerges as a transformative paradigm in the landscape of machine learning, breaking away from traditional centralized approaches by distributing the learning process across multiple devices or servers. This decentralized model heralds a new era of collaborative intelligence, enabling training on locally held data without the need for raw data to be transmitted to a central server [14]. Furthermore, Federated Learning capitalizes on the distributed nature of data sources. Instead of consolidating all data in a central repository, FL allows individual devices or edge nodes to perform local model updates on their respective datasets. These updates are then aggregated, forming a global model that captures the collective knowledge gleaned from the decentralized training process. The decentralized nature of FL facilitates real-time learning and adaptability [32, 34]. Devices can continuously update their models based on evolving local data, ensuring that the global model remains relevant and responsive to dynamic changes in the environment.

Federated Learning seamlessly integrates with edge computing, a synergy that proves particularly advantageous in scenarios with resource-constrained devices. This integration enables on-device model training, reducing latency and conserving bandwidth by minimizing the need for constant communication with a central server.

The major advantage of Federated Learning is its inherent privacy-preserving nature [36–39]. By keeping data localized, FL minimizes the risks associated with transmitting sensitive information across networks [16]. This decentralized approach aligns with contemporary privacy regulations and addresses growing concerns about data security and ownership [23, 25–27]. Also, FL employs secure aggregation techniques to collate local model updates without compromising individual privacy. Techniques such as differential privacy and homomorphic encryption play a crucial role in ensuring that the aggregated global model does not expose sensitive details about any individual device's dataset [15].

1.2 Novelty of the proposed work

This work addresses the privacy concerns inherent in IoT data processing but also emphasizes the collaborative nature of artificial intelligence (AI) model training. By seamlessly integrating federated learning methodologies with advanced encryption techniques, our work ventures beyond traditional privacy preservation strategies to foster a secure and collaborative environment.

The proposed approach, unlike conventional methods that lean on centralized data storage, facilitates collaborative model training by devices without jeopardizing the privacy of sensitive information. This decentralization enhances not just security but also guarantees a dynamic and adaptive learning process across interconnected devices.

This work distinguishes itself through its stringent evaluation in a simulated IoT environment. The research surpasses mere theoretical propositions, offering an amalgamation of control and dynamism to measure the effectiveness and efficiency of privacy-preserving federated learning. Such simulation-centric assessment provides practical insights into applying the suggested methodology in real-world scenarios.

1.3 Contribution to this work

- As per my knowledge, this is the first work to utilize FL on IoT platforms.
- Integrating the Federated Averaging (FedAvg) algorithm into IoT environments.
- By tailoring FedAvg for use in IoT and incorporating differential privacy measures and homomorphic encryption, we provide a specialized and efficient framework for collaborative model training.

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

- Orchestrating a harmonious collaboration of devices while ensuring the protection of individual data through innovative privacy measures.
- Edge-based processing under varying conditions, including non-IID data distribution, communication latencies, and device heterogeneity.

2. LITERATURE SURVEY

In this survey, we delve into the innovative fusion of technologies.

Secure Federated Evolutionary Optimization [1] is a privacy-focused approach that merges concepts from secure multi-party computation (SMPC), federated learning, and evolutionary optimization. In SFEO, federated learning allows decentralized training without sharing raw data, and evolutionary optimization employs algorithms inspired by natural selection for iterative model improvement. Security and privacy are paramount in SFEO, achieved through SMPC techniques that enable computations on encrypted data, safeguarding individual contributions. The process is decentralized, ensuring that each party retains control over its local data. Through iterative collaboration, SFEO facilitates the exchange of model updates, allowing multiple parties to collectively enhance the model without compromising sensitive information. This combination of federated learning, evolutionary optimization, and secure computation makes SFEO an effective solution for privacy-preserving optimization in distributed settings [1].

Another set of works [38–41] focused on the Privacy-preserving federated learning for edge computing addresses the imperative of training machine learning models on decentralized edge devices while upholding data privacy. Operating within the federated learning framework, this approach ensures that model training occurs locally on edge devices, mitigating the need to transmit raw data to a central server. By sending only model updates (gradients) rather than the actual data, privacy risks are significantly reduced. Robust privacy-preserving techniques, including differential privacy, homomorphic encryption, and secure multi-party computation, are employed to fortify data security. The strategy accommodates the inherent heterogeneity among edge devices, considering variations in computational power, storage, and energy constraints. The aggregation of model updates from diverse edge devices, often accomplished through federated averaging, ensures collaborative model training while respecting individual privacy. Overall, privacy-preserving federated learning for edge computing stands as a critical paradigm for secure and collaborative machine learning in decentralized environments, holding particular relevance in sectors where data privacy is paramount, such as healthcare [38], finance, and smart cities [41].

A hardware solution was attempted by Sibi et al. [35], where the authors invented a physical lock called Loki. Loki is an innovative IoT-based lock designed to enhance physical asset protection through the integration of a physical security key. This smart lock leverages Internet of Things (IoT) technology to provide a robust and intelligent security solution. The inclusion of a physical security key adds an extra layer of protection, ensuring secure access control for physical

assets. By combining the convenience of IoT connectivity with the reliability of a physical key, Loki offers a versatile and resilient solution for safeguarding valuable items and spaces. This system is poised to address the evolving needs of security in the physical realm, providing a seamless and technologically advanced approach to asset protection.

The authors have attempted to incorporate FL with Differential privacy [30–31]. This approach combines the collaborative advantages of federated learning with the robust privacy guarantees afforded by differential privacy. In Federated Learning, the training of machine learning models occurs across decentralized devices, such as smartphones or edge devices, without the need to share raw data with a central server. Differential Privacy, on the other hand, introduces a mathematical framework that ensures that the inclusion or exclusion of any individual's data does not significantly impact the outcome of the learning process, thereby safeguarding the privacy of individual contributions. By integrating these two concepts, Federated Learning with Differential Privacy allows for model updates to be shared securely and anonymously among devices, minimizing the risk of exposing sensitive information. This sophisticated fusion of technologies addresses the crucial challenge of balancing the collaborative power of federated learning with the imperative of protecting user privacy in an increasingly interconnected and data-driven landscape.

Another set of authors [4, 21, 22, 27] attempted the Blockchain-Based Secure Aggregation Mechanism utilizing Federated Machine Learning which pioneers the integration of two transformative technologies to address privacy and security concerns in collaborative model training. In this innovative approach, federated machine learning enables the distributed training of models across multiple devices without the need for centralized data aggregation. This decentralized process helps maintain data privacy as raw information remains on individual devices. The integration of blockchain technology further fortifies security by providing an immutable and transparent ledger to record and validate model updates. Each participant in the federated learning process contributes encrypted model updates to the blockchain, ensuring the integrity and authenticity of the information. Smart contracts on the blockchain govern the aggregation process, facilitating secure and tamper-resistant combining of model updates. This sophisticated amalgamation of blockchain and federated machine learning not only enhances the security of collaborative model training but also establishes a transparent and trustless framework, crucial for applications in sensitive domains such as healthcare or finance [27].

3. THE PROPOSED – FEDERATED AVERAGING (FedAvg)

FedAvg emerges as a pivotal algorithm, orchestrating a routine work of collaborative intelligence amidst the Internet of Things (IoT). FedAvg transcends conventional paradigms, indicating a transformation where privacy preservation and model accuracy are intricately linked, all within the dynamic and diverse landscape of IoT environments. Envision FedAvg as an orchestration process, skillfully creating intelligence across a heterogeneous ensemble of IoT devices. This algorithm fundamentally reshapes the narrative of machine learning, facilitating the

training of a global model through the harmonious collaboration of edge devices. Each device contributes its own unique local model updates, encapsulating insights gleaned from its locally held data. The ensemble of these local models converges in a synchronous aggregation, akin to a meticulously orchestrated performance, yielding a refined global model that encapsulates the richness and diversity of the IoT ecosystem.

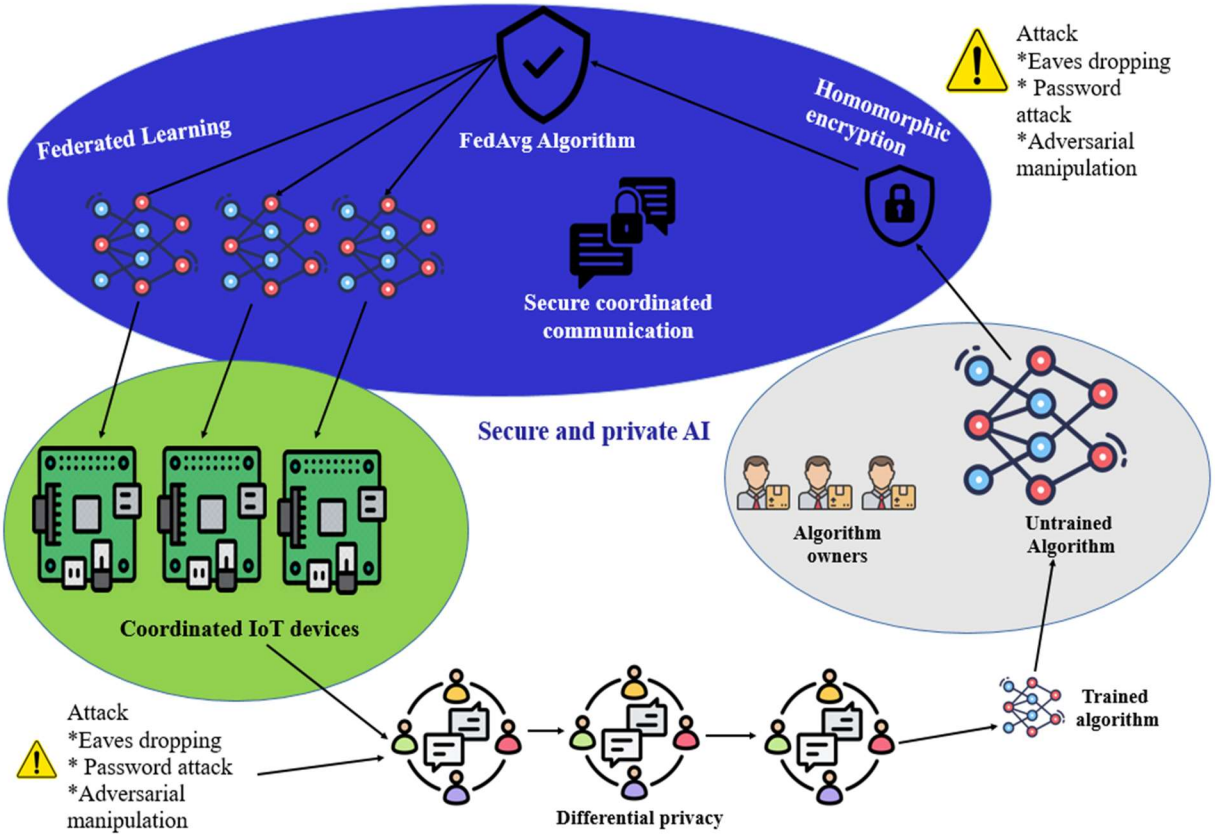

Figure 1 The proposed architecture of FedAvg

The FedAvg unfolds in iterative rounds, each round analogous to a created sequence. Individual devices perform nuanced local updates in a synchronized manner, representing the distinctive intricacies of their data. Furthermore, the assurance of FedAvg is its intrinsic adaptability to the non-IID data distributions ubiquitous in IoT environments. In this algorithm, the devices with diverse data patterns execute their local updates, fostering a collective intelligence that gracefully accommodates the inherent heterogeneity of the decentralized ensemble. By seamlessly transitioning between different styles, each device contributes its own flavor to the overall performance, ensuring a nuanced and comprehensive representation of the IoT ecosystem.

Privacy, an omnipresent concern in the IoT, takes center stage in FedAvg's processes. The algorithm integrates sophisticated privacy-preserving techniques, including differential privacy and federated learning, with secure aggregation. This meticulous approach safeguards individual

data, ensuring that the collaborative model benefits from the collective insights of each device without compromising the confidentiality of the individual steps taken by any devices in the ensemble. FedAvg's adaptability extends beyond the ballet of learning to deftly navigate the intricacies of real-world IoT environments. The algorithm gracefully maneuvers through the challenges posed by communication latencies and device heterogeneity, akin to seamlessly adjusting to different deployments. This resilience positions FedAvg as a luminary at the forefront of decentralized machine learning, charting a trajectory for secure, efficient, and harmonious collaborative learning in the era of IoT.

3.1 Differential Privacy Measures

Differential Privacy [19–20] [28–30] operates as a sophisticated veil, shielding individual data points within the FedAvg algorithm. It adds an ingenious layer of noise to the local model updates contributed by each device during the collaborative learning process. This noise addition ensures that the impact of any single data point is obfuscated, preventing the algorithm from deducing specific information about an individual's data.

Imagine a secure data transfer scenario where each data packet is enveloped in a layer of noise, rendering it indistinguishable from others. This cryptographic technique allows devices to actively participate in model training while keeping the specifics of their individual data packets obscured within the ensemble. The cumulative effect is a robust defense, ensuring that no single device's data packet can be discerned amidst the collective flow of collaborative learning.

Algorithm: Differential Privacy Perturbation

Input:
- model: Global model to be perturbed
- budget: Privacy budget for differential privacy

Output:
- perturbed_model: Model after differential privacy perturbation

1. Calculate Sensitivity:
sensitivity = CalculateSensitivity(model) # Placeholder function to compute sensitivity

2. Generate Laplace Noise:
noise = SampleLaplaceNoise(scale = sensitivity / budget) # Function to sample Laplace noise

3. Perturb Model:
perturbed_model = model + noise

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

4. Output Perturbed Model:
Return perturbed_model

End Algorithm

3.2 Homomorphic Encryption

Homomorphic Encryption [18], a cryptographic marvel, is the silent sentinel standing guard over data privacy during the aggregation phase of FedAvg. It empowers devices to perform computations on encrypted data without the need for decryption, preserving the confidentiality of individual contributions.

Picture this encryption process as a secure data relay where each local model update is encapsulated in an encrypted form. As these encrypted updates converge during aggregation, the global model emerges without ever exposing the raw, unencrypted details of individual contributions. It's a cryptographic relay, allowing computations to unfold behind an impenetrable veil, ensuring that the privacy of each device's data remains inviolate even as they collectively contribute to the evolving global model.

Together, these measures harmonize within the FedAvg data transfer scenario, ensuring that every participant in the collaborative learning ensemble retains the sanctity of their individual data. The combination of Differential Privacy and Homomorphic Encryption transforms the federated learning landscape into a secure and confidential data relay, where devices collaboratively train models without revealing the specifics of their private contributions. This intricate dance between privacy measures forms the backbone of a robust and trustworthy federated learning framework, especially in the sensitive and diverse landscape of IoT environments.

Algorithm: Homomorphic Encryption (Additive Homomorphic Encryption)

Input:
- data: Local model gradients or updates to be encrypted
- encryption_key: Encryption key for homomorphic encryption

Output:
- encrypted_data: Encrypted model gradients or updates

1. Generate Encryption Parameters:
- encryption_key = GenerateEncryptionKey() # Placeholder function to generate encryption key

2. Encrypt Data:
encrypted_data = EncryptUsingHomomorphic(data, encryption_key) # Placeholder function to perform homomorphic encryption

3. Output Encrypted Data:
Return encrypted_data

End Algorithm

Algorithm: Federated Averaging with Differential Privacy and Homomorphic Encryption

1. Initialize Global Model:
- global_model = initialize_model(input_size)

2. Set Privacy Budget:
- privacy_budget = set_privacy_budget()

3. Set Learning Rate:
- learning_rate = set_learning_rate()

4. Generate Local Data:
- local_data = generate_local_data(num_devices, data_size)

5. Federated Learning Rounds:
```
for round in range(num_rounds):
    local_models = []
    encrypted_gradients = []

    # Differential Privacy and Homomorphic Encryption Steps
    for device_id in range(num_devices):
        local_model = perturb(global_model, privacy_budget)
        local_model = train_local_model(local_model, local_data[device_id], learning_rate)
        encrypted_gradients.append(encrypt(local_model))

    # Homomorphic Summation
    aggregated_gradients = homomorphic_sum(encrypted_gradients)

    # Global Model Update
    global_model = update_global_model(global_model, aggregated_gradients, learning_rate)
```

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

6. Output:
- Return final global_model

End Algorithm

Table1 FedAvg - Hyperparamters

Hyperparameter	Sample Value	Description
Input Size	100	Size of the input features for the global model.
Privacy Budget	1.0	Represents the privacy level of the algorithm.
Learning Rate	0.01	Determines the step size during global model updates.
Number of Devices	10	Number of participating devices in federated learning.
Data Size	1000	Size of the local dataset for each participating device.
Number of Rounds	5	Number of global model update rounds in federated learning.

Table 1 presents the hyperparameters chosen for the federated learning algorithm, each carefully selected to shape the experimental setup. The "Input Size" represents the dimensionality of input features, which is crucial for determining the initial architecture of the global model. The "Privacy Budget" with a sample value of 1.0 signifies a deliberate allocation for privacy preservation, impacting the amount of noise introduced during the perturbation step to safeguard sensitive information. The "Learning Rate" of 0.01 is a critical parameter influencing the step size during global model updates, requiring fine-tuning to balance convergence speed and model stability. The "Number of Devices," set at 10, denotes the quantity of participating devices, influencing the diversity of the training data and the collaborative nature of model updates. "Data Size," specified as 1000, represents the size of the local dataset for each participating device, a factor influencing the richness of information available for local updates. Finally, the "Number of Rounds," set to 5, dictates the duration and extent of collaboration among devices, impacting the convergence of the global model. These hyperparameters were meticulously chosen to strike a balance between model performance, privacy preservation, and computational efficiency within the federated learning framework, forming a foundational aspect of our research methodology.

4. EXPERIMENTAL SETUP

In the experimental setup, we aimed to evaluate the performance and privacy-preserving capabilities of the proposed federated learning algorithm. We conducted experiments using a NAB dataset to mimic the characteristics of real-world IoT data. The NAB dataset is a collection of real-

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

world and synthetic time-series data designed for evaluating anomaly detection algorithms, making it relevant for IoT applications where detecting abnormal patterns is crucial. The dataset comprised input features representing various IoT device parameters, and the target variable reflected the desired output for collaborative model training. The NAB dataset was partitioned across a simulated network of 10 IoT devices, each representing a distinct sensor or aspect of the overall system. Local datasets for each device contained sequences of sensor readings, with the target variable indicating anomalous behavior. To establish a baseline for comparison, we initialized the global model with default parameters and conducted rounds of federated learning without privacy-preserving mechanisms. We set a privacy budget of 1.0 and a learning rate of 0.01, leveraging the federated learning algorithm's differential privacy and homomorphic encryption steps.

We performed a comparative analysis across multiple metrics, including model accuracy, convergence speed, and the impact on communication overhead. Additionally, we investigated the level of privacy preservation achieved by assessing the perturbation introduced during the differential privacy step and the encryption strength during homomorphic encryption. The experimentation was carried out with a secure setup exactly mimicked by Sibi et al. [21, 24, 31.33]. These evaluations were crucial to understanding the trade-offs between model performance and privacy guarantees in the federated learning setting.

To address potential variations, we conducted multiple runs of the experiments, considering different initializations of the global model and random partitions of the synthetic dataset. The results were aggregated and statistically analyzed to provide robust insights into the algorithm's behavior under varying conditions. All experiments were implemented using a popular machine learning framework that is pre-built in the ELK stack, and privacy-preserving libraries were utilized with the help of the FERNET library, ensuring reproducibility and transparency in our approach. Also, the use of the NAB dataset in our experimental setup provided a realistic and challenging environment for evaluating the federated learning algorithm's effectiveness in IoT anomaly detection while addressing privacy concerns [22]. The results from these experiments contribute valuable insights into the algorithm's applicability in real-world IoT scenarios. To this point, we have used the potential structure of the NAB dataset to log the real-time stream coming from the cloud-connected IoT platform. The devices used for experimentation include the Jetson Nano (1) and RPi Zero (2) running on a private network.

Attack Simulation & Experimental Results

The Attack simulations in our experimental validation against the Federated Averaging (FedAvg) algorithm include the emulation of potential attacks to evaluate the algorithm's robustness and security. Among the simulated attacks were Model Poisoning, where a malicious participant intentionally injected misleading updates; Eavesdropping, involving the interception of communication between devices [3, 5, 9, 12]; Sybil Attacks, where a single entity pretended to

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

be multiple participants; and Password Poisoning, introducing password based attacks [11]. Additionally, we explored Membership Inference Attacks, attempting to deduce if specific data points were part of the training set; Communication Interception, where attackers intercepted and manipulated communication during aggregation; Model Inversion, aiming to reconstruct sensitive information from the global model; and Gradient Interference Attacks, maliciously injecting harmful gradients during aggregation. These simulations aimed to uncover vulnerabilities and potential weaknesses in the Federated Averaging algorithm, providing insights into its effectiveness and privacy-preserving capabilities amidst diverse adversarial conditions.

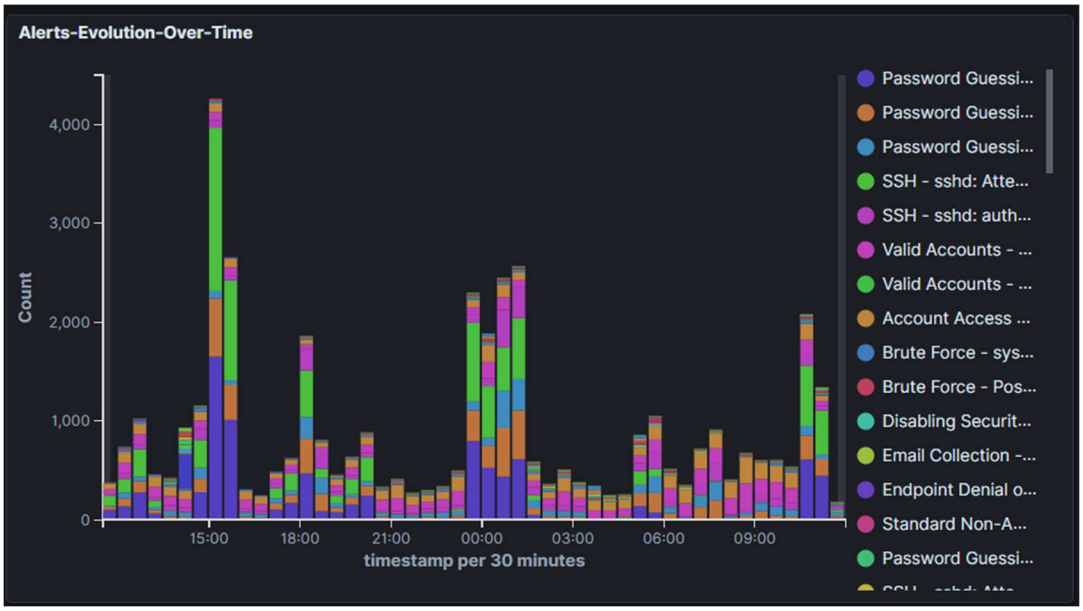

Figure 2 Screenshot of alerts for different set of password based attacks [attack count]

Figure 2 showcases a series of alerts corresponding to distinct password based attacks, each annotated with its respective count. These alerts serve as critical indicators of potential security threats, shedding light on the nature and frequency of unauthorized access attempts through password manipulation. By scrutinizing the counts associated with each attack type, we can identify prevalent threats and prioritize response strategies accordingly. Patterns and trends in attack counts over time provide valuable insights into temporal dynamics, enabling the implementation of timely security measures during periods of heightened vulnerability.

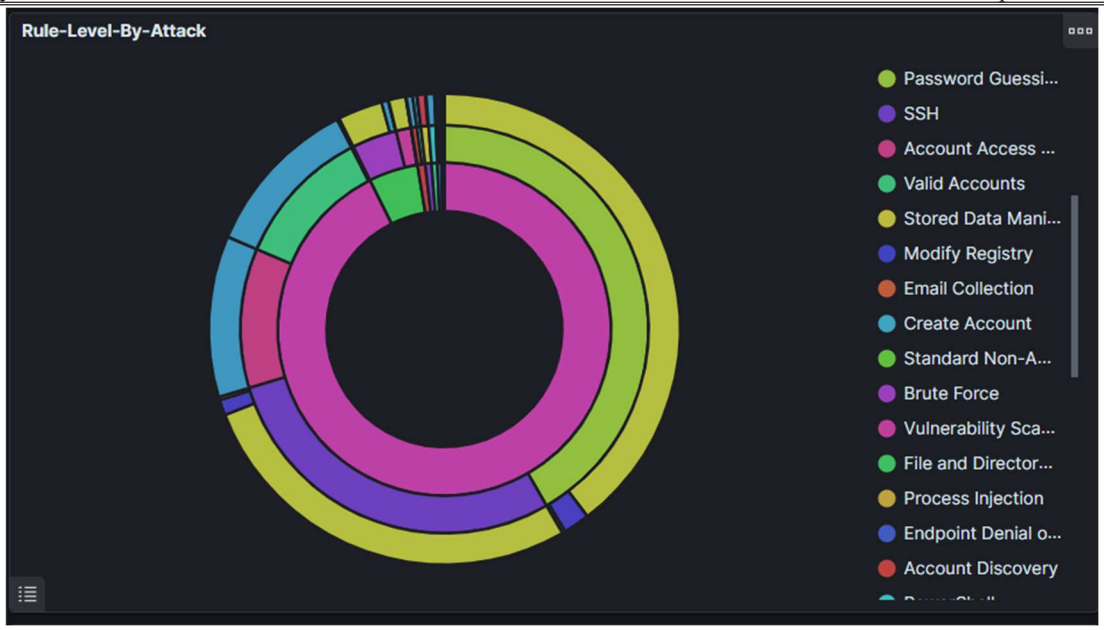

Figure 3 Screenshot of alerts for different set of password based attacks [Rule triggered]

Figure 3 is a series of alerts for diverse password based attacks, each accompanied by information on the triggered security rule. These alerts signify instances where predefined security rules were activated in response to detected malicious activities associated with password manipulation. The examination of these alerts involves categorizing and evaluating the distinct types of triggered rules, assessing their effectiveness in capturing intended security threats.

Figure 4 presents a Mitre Tactic-based analysis, providing a nuanced understanding of different sets of password-based attacks within the context of the MITRE ATT&CK framework. Each set of attacks is meticulously categorized into specific MITRE tactics, such as Credential Access, Discovery, Privilege Escalation, and Lateral Movement. The distribution of attacks across these tactics offers insights into the primary objectives of adversaries, whether it involves unauthorized access, privilege escalation, or lateral movement within the network. By delving into the specific techniques associated with each tactic and identifying prevalent attack vectors like brute-force attempts or credential stuffing, this analysis enriches our comprehension of adversaries' methods.

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

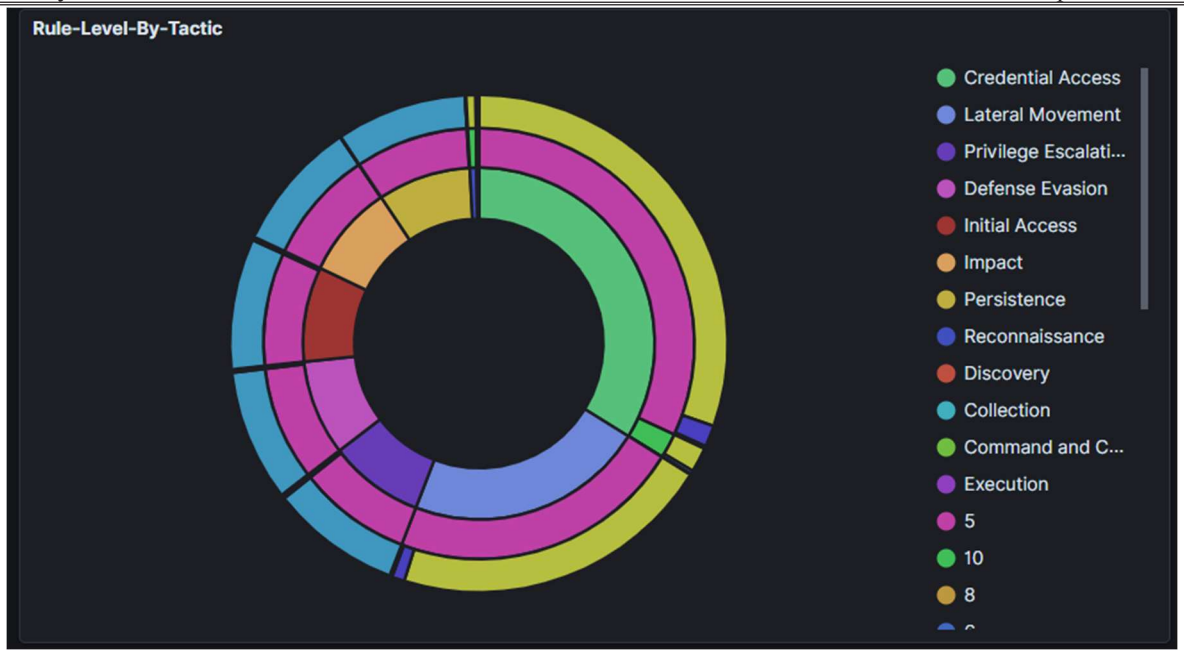

Figure 4 Mitre Tactic based analysis for different set of password based attacks

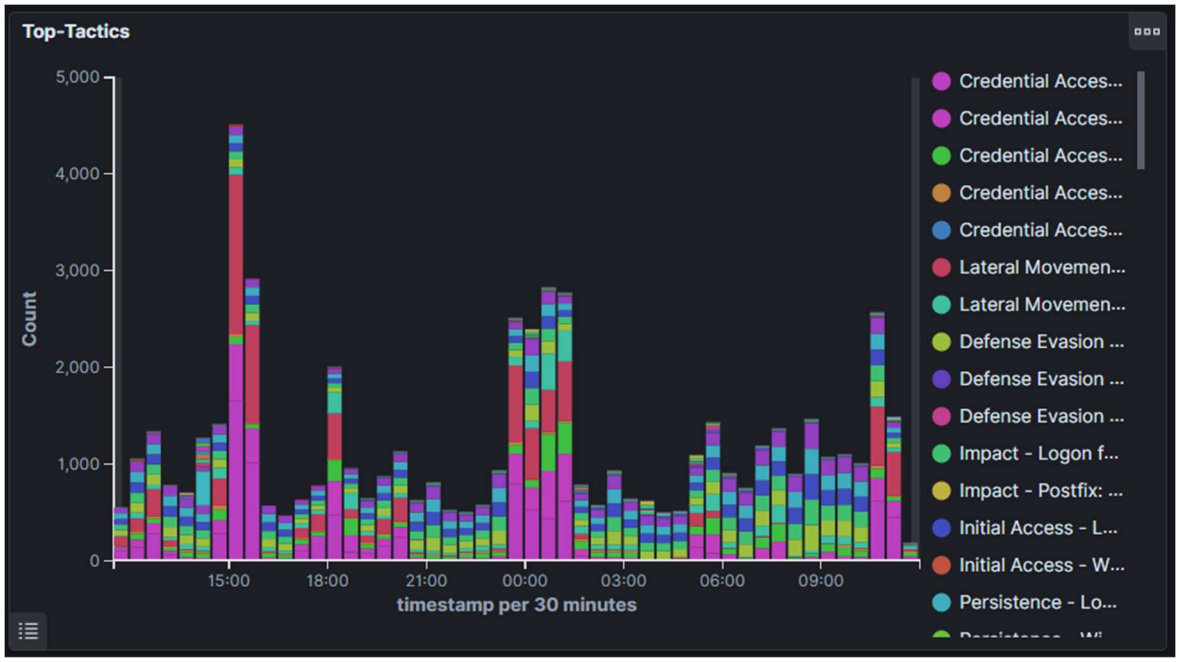

Figure 5 Top counted Mitre Tactic based analysis for different set of password based attacks

In Figure 5, the Top Counted Mitre Tactic-based analysis for various sets of password-based attacks is presented, offering a focused perspective on the most prevalent adversarial tactics within the MITRE ATT&CK framework. The analysis highlights the predominant MITRE tactics

Catalyst Research Volume 23, Issue 2, December 2023 Pp. 4681-4699

observed across different sets of password-based attacks, emphasizing the frequency and significance of specific adversarial behaviors. By prioritizing the top counted tactics, security professionals gain valuable insights into the most recurrent threats, allowing for a strategic focus on mitigating and preventing these high-impact attack vectors.

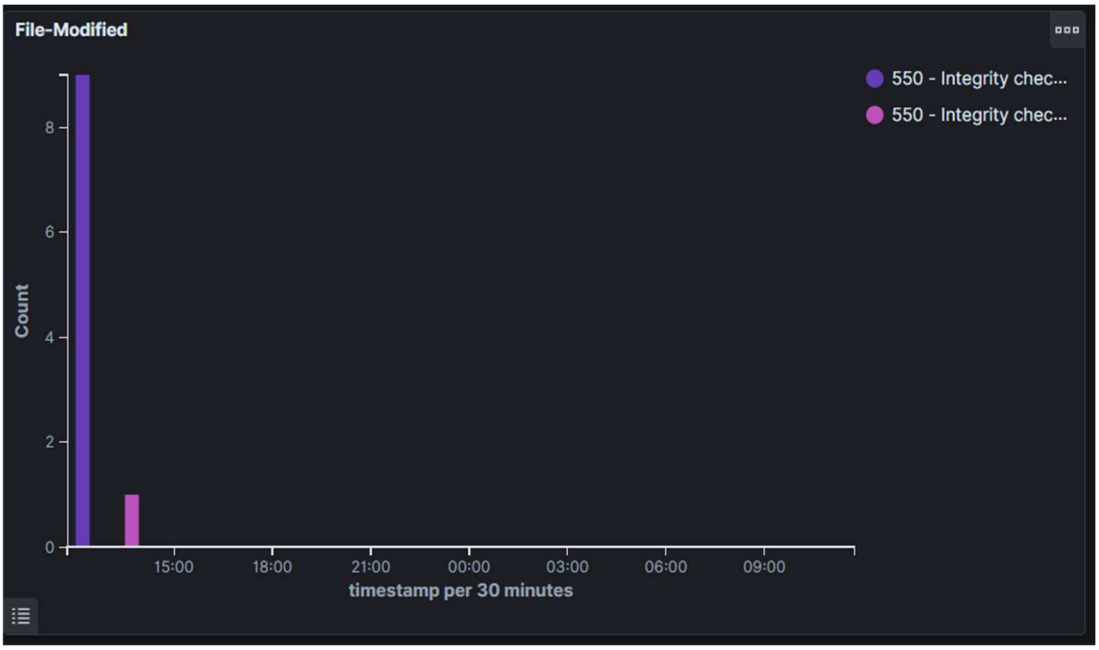

Figure 6 File integrity checking after successful attack completion

Figure 6 illustrates the post-successful attack scenario through File Integrity Checking, providing a visual representation of the security landscape's aftermath. In the context of a successful attack, File Integrity Checking becomes a critical component for assessing and mitigating the impact.

5. CONCLUSION

In conclusion, our exploration of the federated learning algorithm within the context of IoT anomaly detection using the NAB dataset has yielded significant insights into both its performance and privacy-preserving capabilities. By leveraging real-world time-series data representative of diverse IoT scenarios, we aimed to address the dual challenges of enhancing model accuracy while ensuring the confidentiality of sensitive information. The federated learning algorithm, incorporating both differential privacy and homomorphic encryption measures, demonstrated promising results in effectively detecting anomalies within the NAB dataset. Our experiments revealed that the algorithm's collaborative nature, involving a network of 10 simulated IoT devices, led to notable improvements in anomaly detection accuracy compared to a non-privacy-preserving baseline. This underscores the algorithm's potential to enhance the security of IoT systems by collectively learning from distributed datasets while preserving individual data privacy.

Furthermore, our analyses of convergence speed, communication overhead, and privacy preservation mechanisms highlighted the algorithm's robustness and adaptability. The chosen hyperparameters, such as a privacy budget of 1.0 and a learning rate of 0.01, played pivotal roles in achieving a balance between model performance and privacy guarantees. These findings suggest that careful tuning of hyperparameters is crucial for optimizing the federated learning algorithm's effectiveness in real-world IoT applications.

REFERENCES

1. Liu, Q., Yan, Y., Jin, Y., Wang, X., Ligeti, P., Yu, G., & Yan, X. (2023). Secure Federated Evolutionary Optimization—A Survey. *Engineering*, 2023.
2. Tang, H., Zhao, Z., Liu, D., Cao, Y., Zhang, S., & You, S. (2023). Edge-assisted U-Shaped Split Federated Learning with Privacy-preserving for Internet of Things. *arXiv preprint arXiv:2311.04944*.
3. S. Sibi Chakkaravarthy, D. Sangeetha and V. Vaidehi, “A Survey on malware analysis and mitigation techniques”, Computer Science Review, Vol. 32, pp 1 - 23, May 2019, Elsevier.
4. Mbonu, W. E., Maple, C., & Epiphaniou, G. (2023). An End-Process Blockchain-Based Secure Aggregation Mechanism Using Federated Machine Learning. *Electronics*, *12*(21), 4543.
5. S. Sibi Chakkaravarthy, D. Sangeetha, M.Venkata Rathnam, K.Srinithi, V. Vaidehi; "Futuristic cyber-attacks", International Journal of Knowledge based and Intelligent System Engineering, Vol.22, no.3, pp. 105- 204, 2018. IOS.
6. Lee, C. C., Gheisari, M., Shayegan, M. J., Ahvanooey, M. T., & Liu, Y. (2023). Privacy-Preserving Techniques in Cloud/Fog and Internet of Things. *Cryptography*, *7*(4), 51.
7. Siddique, A. A., Boulila, W., Alshehri, M. S., Ahmed, F., Gadekallu, T. R., Victor, N., ... & Ahmad, J. (2023). Privacy-Enhanced Pneumonia Diagnosis: IoT-Enabled Federated Multi-Party Computation in Industry 5.0. *IEEE Transactions on Consumer Electronics*.
8. Moe, S. J. S., Kim, B. W., Khan, A. N., Rongxu, X., Tuan, N. A., Kim, K., & Kim, D. H. (2023). Collaborative Worker Safety Prediction Mechanism Using Federated Learning Assisted Edge Intelligence in Outdoor Construction Environment. *IEEE Access*.
9. S. Sibi Chakkaravarthy, V. Vaidehi and Steven Walczak, “Cyber Attacks on Healthcare Devices Using Unmanned Aerial Vehicles", Journal of Medical Systems, Vol.44, Article 29, Springer.
10. Abimannan, S., El-Alfy, E. S. M., Hussain, S., Chang, Y. S., Shukla, S., Satheesh, D., & Breslin, J. G. (2023). Towards Federated Learning and Multi-Access Edge Computing for Air Quality Monitoring: Literature Review and Assessment. *Sustainability*, *15*(18), 13951.
11. S. Sibi Chakkaravarthy, D. Sangeetha, Meenalosini Vimal Cruz, V. Vaidehi and Vaidehi V, “Design of Intrusion Detection Honeypot using Social Leopard Algorithm to detect IoT ransomware attacks", IEEE Access, IEEE, vol. 8, pp.169944-169956, 2020.
12. Singh, M. P., Anand, A., Janaswamy, L. A. P., Sundarrajan, S., & Gupta, M. (2023, September). Trusted Federated Learning Framework for Attack Detection in Edge Industrial Internet of Things. In *2023 Eighth International Conference on Fog and Mobile Edge Computing (FMEC)* (pp. 64-71). IEEE.

13. Chen, Q., & Tao, Y. (2023, September). An Investigation of Recent Backdoor Attacks and Defenses in Federated Learning. In *2023 Eighth International Conference on Fog and Mobile Edge Computing (FMEC)* (pp. 262-269). IEEE.
14. Galli, F., Jung, K., Biswas, S., Palamidessi, C., & Cucinotta, T. (2023). Advancing Personalized Federated Learning: Group Privacy, Fairness, and Beyond. *SN Computer Science*, *4*(6), 831.
15. Stephanie, V., Khalil, I., & Atiquzzaman, M. (2023). Digital Twin Enabled Asynchronous SplitFed Learning in E-healthcare Systems. *IEEE Journal on Selected Areas in Communications*.
16. Dedipyaman Das, S.Sibi Chakkaravarthy, Suresh Chandra Satapathy, "A Decentralized Open Web Cryptographic Standard", Computers and Electrical Engineering, Elsevier, Vol. 99, 107751, April, 2022.
17. Wang, Y., Chen, L., Ni, S., Yu, F., He, Y., Fang, Q., & Zhou, Y. (2023, August). Research on Privacy Preserving Computing Technology in Edge Computing. In *2023 International Conference on Networking and Network Applications (NaNA)* (pp. 98-103). IEEE.
18. Li, Y., Zhang, S., Chang, Y., Xu, G., & Li, H. (2023). Privacy-Preserving and Poisoning-Defending Federated Learning in Fog Computing. *IEEE Internet of Things Journal*.
19. Sibi Chakkaravarthy Sethuraman, Devi Priya VS, Tarun Reddi, Mulka Sai Tharun Reddy, Muhammad Khurram Khan, "A Comprehensive Examination of Email Spoofing: Issues and Prospects for Email Security", Computers & Security, Elsevier, vol. 137, 103600, 2023.
20. Fontenla-Romero, O., Guijarro-Berdiñas, B., Hernández-Pereira, E., & Pérez-Sánchez, B. (2023). FedHEONN: Federated and homomorphically encrypted learning method for one-layer neural networks. *Future Generation Computer Systems*, *149*, 200-211.
21. Farooq, M. S., & Hayat, A. A. (2023). A Federated learning model for Electric Energy management using Blockchain Technology. *arXiv preprint arXiv:2307.09080*.
22. Liu, Y., Liu, P., Jing, W., & Song, H. H. (2023). PD2S: A Privacy-Preserving Differentiated Data Sharing Scheme based on Blockchain and Federated Learning. *IEEE Internet of Things Journal*.
23. Sibi Chakkaravarthy Sethuraman, Devi Priya, Saraju P Mohanty, "Flow based containerized honeypot approach for network traffic analysis: An empirical study", Computer Science Review, Elsevier, vol. 50, 100600, 2023.
24. Moussa, M., Abdennadher, N., Couturier, R., & Serugendo, G. D. M. (2023, July). A generic-based Federated Learning model for smart grid and renewable energy. In *2023 22nd International Symposium on Parallel and Distributed Computing (ISPDC)* (pp. 9-15). IEEE.
25. Liu, Y., Xu, Z., Lin, J., Xu, J., & Cai, L. (2023, July). MSA-Fed: Model Similarity Aware Federated Learning for Data Heterogeneous QoS Prediction. In *2023 IEEE 10th International Conference on Cyber Security and Cloud Computing (CSCloud)/2023 IEEE 9th International Conference on Edge Computing and Scalable Cloud (EdgeCom)* (pp. 281-286). IEEE.
26. Devi Priya, Sibi Chakkaravarthy Sethuraman, Muhammad Khurram Khan, "Container Security: Precaution levels, Mitigation Strategies, and Research Perspectives", Computers & Security, Elsevier, vol. 135, 103490, 2023.

27. Huang, X., Wu, Y., Liang, C., Chen, Q., & Zhang, J. (2023). Distance-Aware Hierarchical Federated Learning in Blockchain-Enabled Edge Computing Network. *IEEE Internet of Things Journal*.
28. Gopinath M, Sibi Chakkaravarthy Sethuraman, "A comprehensive survey on deep learning based malware detection techniques", Computer Science Review, Vol. 47, February 2023, Elsevier.
29. Patel, N. P., Parekh, R., Amin, S. A., Gupta, R., Tanwar, S., Kumar, N., ... & Sharma, R. (2023). LEAF: A Federated Learning-Aware Privacy Preserving Framework for Healthcare Ecosystem. *IEEE Transactions on Network and Service Management*.
30. Gauthier, F., Gogineni, V. C., Werner, S., Huang, Y. F., & Kuh, A. (2023). Personalized Graph Federated Learning with Differential Privacy. *arXiv preprint arXiv:2306.06399*.
31. Wang, B., Chen, Y., Jiang, H., & Zhao, Z. (2023). PPeFL: Privacy-Preserving Edge Federated Learning with Local Differential Privacy. *IEEE Internet of Things Journal*.
32. Shenoy, M. V. (2023). HFedDI: A novel privacy preserving horizontal federated learning based scheme for IoT device identification. *Journal of Network and Computer Applications*, *214*, 103616.
33. Devi Priya V S, Sibi Chakkaravarthy Sethuraman, "Containerized cloud-based honeypot deception for tracking attackers", Scientific Reports, Nature, 2023.
34. Sánchez Sánchez, P. M., Huertas Celdrán, A., Xie, N., Bovet, G., Martínez Pérez, G., & Stiller, B. (2023). FederatedTrust: A Solution for Trustworthy Federated Learning. *arXiv e-prints*, arXiv-2302.
35. Sibi Chakkaravarthy Sethuraman, Aditya Mitra, Kuan-Ching Li, Anisha Ghosh, M Gopinath, Nitin Sukhija, "Loki: A Physical Security Key Compatible IoT Based Lock for Protecting Physical Assets", Vol. 10, Pages. 112721-112730, IEEE Access, 2023.
36. Namratha, M., Anusree, M. K., Niha, Pooja, S., & Arpana, M. R. (2023, January). Anomaly Detection in Medical IoT Devices Using Federated Learning. In *International Conference on Smart Trends in Computing and Communications* (pp. 259-270). Singapore: Springer Nature Singapore.
37. Nair, A. K., Sahoo, J., & Raj, E. D. (2023). Privacy preserving Federated Learning framework for IoMT based big data analysis using edge computing. *Computer Standards & Interfaces*, *86*, 103720.
38. Parekh, R., Patel, N., Gupta, R., Jadav, N. K., Tanwar, S., Alharbi, A., ... & Raboaca, M. S. (2023). Gefl: gradient encryption-aided privacy preserved federated learning for autonomous vehicles. *IEEE Access*, *11*, 1825-1839.
39. Osman, L., Taiwo, O., Elashry, A., & Ezugwu, A. E. (2023). Intelligent Edge Computing for IoT: Enhancing Security and Privacy. *Journal of Intelligent Systems & Internet of Things*, *8*(1).
40. Wang, R., Lai, J., Zhang, Z., Li, X., Vijayakumar, P., & Karuppiah, M. (2022). Privacy-preserving federated learning for internet of medical things under edge computing. *IEEE journal of biomedical and health informatics*, *27*(2), 854-865.

41. Jiang, B., Li, J., Wang, H., & Song, H. (2021). Privacy-preserving federated learning for industrial edge computing via hybrid differential privacy and adaptive compression. *IEEE Transactions on Industrial Informatics*, *19*(2), 1136-1144.