Privacy-Aware Cyberterrorism Network Analysis using Graph Neural Networks and Federated Learning

Anas Ali
dept. of Computer Science
National University of Modern Langauges
Lahore, Pakistan
anas.ali@numl@edu.pk

Mubashar Husain
Department of Computer Science
University of Lahore,
Pakistan
m.hussain2683@gmail.com

Peter Hans
Department of Electrical Engineering
University of Sharjah
United Arab Emirates
peter19972@gmail.com

***Abstract*—Cyberterrorism poses a formidable threat to digital infrastructures, with increasing reliance on encrypted, decentralized platforms that obscure threat actor activity. To address the challenge of analyzing such adversarial networks while preserving the privacy of distributed intelligence data, we propose a Privacy-Aware Federated Graph Neural Network (PA-FGNN) framework. PA-FGNN integrates graph attention networks, differential privacy, and homomorphic encryption into a robust federated learning pipeline tailored for cyberterrorism network analysis. Each client trains locally on sensitive graph data and exchanges encrypted, noise-perturbed model updates with a central aggregator, which performs secure aggregation and broadcasts global updates. We implement anomaly detection for flagging high-risk nodes and incorporate defenses against gradient poisoning. Experimental evaluations on simulated dark web and cyber-intelligence graphs demonstrate that PA-FGNN achieves over 91% classification accuracy, maintains resilience under 20% adversarial client behavior, and incurs less than 18% communication overhead. Our results highlight that privacy-preserving GNNs can support large-scale cyber threat detection without compromising on utility, privacy, or robustness.**

I. Introduction

The evolving threat landscape of cyberterrorism poses a significant challenge to national and global security infrastructures. Cyberterrorism refers to the deliberate use of cyberspace to launch attacks that disrupt or damage critical services, spread propaganda, and instill fear through digital means. These attacks often leverage distributed communication networks, anonymous platforms, and encrypted channels to orchestrate large-scale operations, making their detection both technically complex and operationally critical [1], [22]. To mitigate these threats, it is essential to identify hidden patterns of interaction and influence within cyberterrorist networks [23].

Graph neural networks (GNNs) have emerged as a powerful paradigm for learning over relational data and have been widely adopted for modeling social, communication, and threat actor networks [2], [18], [19]. By encoding both structural topology and node-level features, GNNs allow for the discovery of latent influence hierarchies and anomaly patterns that are not discernible through traditional machine learning techniques. However, applying centralized GNN models to sensitive cyberterrorism data risks privacy violations, especially when data originate from multiple security agencies, national firewalls, or confidential intelligence sources [20], [21].

Federated learning (FL) offers a decentralized solution to this dilemma. It enables collaborative model training without sharing raw data, thereby maintaining confidentiality and regulatory compliance [3], [16]. By combining GNNs and FL, researchers can analyze distributed cyberterrorism graphs while preserving sensitive node and edge information.

Despite the promise of GNN-FL integration, several challenges remain unaddressed. These include data heterogeneity across sources, limited bandwidth for model updates, privacy leakage through gradients, and susceptibility to poisoning or backdoor attacks [4], [17]. Furthermore, the adversarial nature of cyberterrorist actors necessitates models that are not only accurate but also robust to obfuscation and misinformation strategies.

This paper addresses these challenges by introducing a privacy-aware federated GNN framework specifically tailored for cyberterrorism network analysis. Our approach integrates homomorphic encryption and differential privacy into the FL pipeline to shield against inference attacks. We also implement robust aggregation schemes to detect and suppress anomalous client behavior. Using real-world cyber threat datasets and simulated communication graphs, we validate the scalability, accuracy, and resilience of our model against adversarial and non-IID scenarios.

The novelty of this work lies in its holistic treatment of cyberterrorism network detection through privacy-enhanced federated GNNs. Unlike prior approaches that treat privacy, federation, or robustness in isolation, our architecture jointly optimizes for these dimensions within a unified system design.

Our key contributions are as follows:

1. We propose a hybrid GNN-FL framework for cyberterrorism graph analysis that integrates differential privacy and homomorphic encryption.

2. We develop a robust aggregation strategy using anomaly-tolerant update mechanisms to secure global model updates.

3. We construct and simulate cyberterrorism graphs derived from multi-source communication logs, and evaluate the system against non-IID and adversarial attacks.

4. We demonstrate, through extensive experiments, that our system achieves high detection accuracy, strong privacy protection, and fault-tolerance, outperforming existing FL and GNN baselines.

The remainder of this paper is structured as follows. Section II surveys related work on GNNs, FL, and secure cyberterrorism analysis. Section III presents our system model and mathematical formulation. Section IV describes the experimental setup, dataset construction, and empirical results. Section V concludes the paper and outlines future directions.

II. Related Work

The intersection of cyberterrorism detection, graph neural networks (GNNs), and federated learning (FL) has gained increasing attention due to the growing complexity of threat networks and privacy concerns. This section surveys key research contributions across each domain and highlights the unique position of our proposed framework.

Kumar et al. [6] provide an overview of cyberterrorism detection methodologies using network analysis and machine learning. They emphasize the role of graph-based techniques but note the lack of scalable and privacy-preserving systems. Their work laid the foundation for graph-centric modeling of malicious online behaviors.

Wu et al. [7] conduct a comprehensive study on GNN architectures for network security applications. They demonstrate the superiority of message-passing neural networks in capturing structural anomalies. However, their experiments rely on centralized training, posing privacy risks in sensitive domains like cybercrime tracking.

Hardy et al. [8] investigate encrypted and anonymous cybercrime forums using graph embeddings and link prediction models. Their results show promise in identifying key actors but suffer from scalability issues and data silos across jurisdictions.

McMahan et al. [9] introduce the FederatedAveraging algorithm, enabling privacy-preserving model training across decentralized clients. Though widely adopted in mobile and health domains, its direct application to GNNs and structured cyberterrorism data remains underexplored.

Zhang et al. [10] propose FedGraphNN, an FL-GNN framework evaluated on citation and co-authorship graphs. While it incorporates personalization layers and partial aggregation, the privacy guarantees are limited to basic differential privacy techniques.

Abadi et al. [11] provide a formal framework for differentially private deep learning and its implementation in TensorFlow Privacy. Their techniques form the basis for secure gradient sharing but are rarely combined with FL for GNN-based cyber threat analysis.

Li et al. [12] propose a secure FL approach for GNNs using cryptographic primitives such as homomorphic encryption and secure aggregation. Their evaluation on synthetic graphs shows performance benefits, but real-world applicability to adversarial graph settings is not assessed.

Sharma et al. [13] examine poisoning attacks in federated GNN environments. They show that backdoor insertion can persist through global model aggregation and propose anomaly scoring to mitigate risks. This motivates the use of robust aggregation in our framework.

Ruan et al. [14] develop FedSage+, a heterogeneous GNN-based FL framework addressing non-IID data via attention fusion. While promising, their model assumes clean data and benign clients, unlike real-world cyberterrorism scenarios.

Gong et al. [15] explore privacy leakage in federated GNNs through graph reconstruction and membership inference attacks. Their findings underscore the need for secure gradient masking and homomorphic encryption as used in our system.

In summary, existing studies have contributed to privacy-preserving GNNs, cybercrime network analysis, and federated learning architectures. However, none combine robust GNN modeling, adversarial resilience, and privacy protections in a cyberterrorism context. Our work fills this gap by proposing a comprehensive, privacy-aware federated GNN framework tailored for cyber threat detection in adversarial and decentralized environments.

III. System Model

We define our cyberterrorism detection model as a federated graph learning system over a set of distributed communication graphs. Each data holder, such as a national agency or ISP, retains a private graph $\mathcal{G}^{(i)} = (\mathcal{V}^{(i)}, \mathcal{E}^{(i)}, \mathbf{X}^{(i)})$ with nodes $\mathcal{V}^{(i)}$, edges $\mathcal{E}^{(i)}$, and feature matrix $\mathbf{X}^{(i)}$.

The goal is to collaboratively train a global graph neural network $f(\cdot;\theta)$ without centralizing sensitive graph data. The local objective for client i is:

$$\mathcal{L}^{(i)} = \frac{1}{|\mathcal{V}^{(i)}|} \sum_{v \in \mathcal{V}^{(i)}} \ell(f(v;\theta^{(i)}), y_v) \tag{1}$$

where ℓ is a supervised loss (e.g., cross-entropy) and y_v is the label of node v.

Each node representation $\mathbf{h}_v^{(l)}$ at GNN layer l is computed as:

$$\mathbf{h}_v^{(l)} = \sigma\left(\sum_{u \in \mathcal{N}(v)} \alpha_{uv}^{(l)} \mathbf{W}^{(l)} \mathbf{h}_u^{(l-1)} \right) \tag{2}$$

where σ is a non-linear activation, α_{uv} are attention coefficients, and $\mathbf{W}^{(l)}$ are learnable weights.

The attention coefficients are derived using:

$$\alpha_{uv}^{(l)} = \frac{\exp(e_{uv}^{(l)})}{\sum_{k \in \mathcal{N}(v)} \exp(e_{kv}^{(l)})} \tag{3}$$

$$e_{uv}^{(l)} = \text{LeakyReLU}(\mathbf{a}^\top [\mathbf{W}\mathbf{h}_u^{(l-1)} \| \mathbf{W}\mathbf{h}_v^{(l-1)}]) \tag{4}$$

where $\|$ denotes concatenation.

To preserve privacy, gradients $\nabla_\theta \mathcal{L}^{(i)}$ are encrypted via a homomorphic encryption function $\mathsf{HE}(\cdot)$ before transmission:

$$\mathsf{HE}(\nabla_\theta \mathcal{L}^{(i)}) = \mathsf{Enc}(\nabla_\theta \mathcal{L}^{(i)}) \tag{5}$$

Clients add noise $\eta \sim \mathcal{N}(0, \sigma^2)$ for differential privacy:

$$\nabla_\theta \widetilde{\mathcal{L}}^{(i)} = \nabla_\theta \mathcal{L}^{(i)} + \eta \tag{6}$$

The global server aggregates encrypted and privatized gradients:

$$\theta_{t+1} = \theta_t - \eta_t \cdot \mathsf{Agg}\left(\left\{\mathsf{HE}(\nabla_\theta \widetilde{\mathcal{L}}^{(i)})\right\}_{i=1}^{K}\right) \tag{7}$$

where η_t is the learning rate at round t.

Each client decrypts the update using their secret key:

$$\mathsf{Dec}(\theta_{t+1}) = \theta_{t+1}^{(i)} \tag{8}$$

Graph structure similarity is measured via cosine distance:

$$\delta_{uv} = 1 - \frac{\mathbf{h}_u \cdot \mathbf{h}_v}{\|\mathbf{h}_u\| \|\mathbf{h}_v\|} \tag{9}$$

To detect suspicious nodes (potential cyberterrorists), anomaly scores are computed as:

$$\mathcal{A}(v) = \|\mathbf{h}_v - \hat{\mathbf{h}}_v\|^2 \tag{10}$$

where $\hat{\mathbf{h}}_v$ is a local neighborhood average.

Nodes exceeding threshold τ are flagged:

$$\mathbb{1}[\mathcal{A}(v) > \tau] = 1 \tag{11}$$

Model convergence is evaluated using average node loss:

$$\overline{\mathcal{L}} = \frac{1}{\sum_i |\mathcal{V}^{(i)}|} \sum_i \sum_{v \in \mathcal{V}^{(i)}} \ell(f(v), y_v) \tag{12}$$

Communication cost per round is:

$$C_{comm} = K \cdot \text{size}(\nabla_\theta \widetilde{\mathcal{L}}^{(i)}) \tag{13}$$

Privacy leakage is approximated as:

$$\mathcal{R}_{leak} = \mathbb{P}(\exists \hat{\mathbf{X}} : \hat{\mathbf{X}} \approx \mathbf{X}^{(i)} \mid \mathsf{HE}(\nabla_\theta \mathcal{L}^{(i)})) \tag{14}$$

Algorithm: Privacy-Aware Federated Graph Neural Network (PA-FGNN)

Algorithm 1 PA-FGNN: Secure Federated GNN for Cyberterrorism Analysis

Input: Local graphs $\mathcal{G}^{(i)}$, labels y, GNN model $f(\cdot)$, noise scale σ, encryption key k
for each communication round $t = 1$ to T **do**
  for each client i in parallel **do**
    Train GNN on $\mathcal{G}^{(i)}$ to get gradients $\nabla_\theta \mathcal{L}^{(i)}$
    Add DP noise: $\nabla_\theta \widetilde{\mathcal{L}}^{(i)} = \nabla_\theta \mathcal{L}^{(i)} + \eta$
    Encrypt: $g_i = \mathsf{HE}(\nabla_\theta \widetilde{\mathcal{L}}^{(i)})$
    Send g_i to server
  end for
  Server aggregates encrypted gradients: $g = \mathsf{Agg}(\{g_i\})$
  Server updates model: $\theta_{t+1} = \theta_t - \eta_t g$
  Broadcast θ_{t+1} to clients
end for
Output: Final GNN model $f(\cdot; \theta_T)$

This algorithm ensures end-to-end privacy of sensitive cyberterrorism graph data. By integrating encrypted updates, differential privacy, and anomaly scoring into the GNN-FL training process, our system achieves robustness against gradient leakage and malicious actor inference.

IV. Experimental Setup and Results

To evaluate the effectiveness of the proposed Privacy-Aware Federated Graph Neural Network (PA-FGNN) framework, we conducted a series of controlled experiments across real and synthetic cyberterrorism datasets. These datasets include communication records from publicly available dark web forums and simulated multi-jurisdictional actor networks based on the CTI (Cyber Threat Intelligence) schemas.

Each data holder, representing a simulated government or private organization, maintained a private graph instance with node features indicating behavioral patterns (e.g., login times, message sentiment, link frequency) and edge types (direct messages, code collaboration, indirect links). Graphs ranged from 2,000 to 10,000 nodes with an average degree of 7.3.

Experiments were implemented using PyTorch Geometric for GNN layers and Flower framework for federated orchestration. Encryption was applied using the TenSEAL homomorphic encryption library, and differential privacy was integrated using Opacus with a noise multiplier of 1.1.

Table I summarizes key simulation parameters:

TABLE I
Simulation Parameters for PA-FGNN Experiments

Parameter	Value
Number of Clients	10
Graph Size per Client	2k–10k nodes
GNN Model	2-layer GAT, 64 hidden units
Learning Rate	0.005
Federated Rounds	100
Batch Size	128
Noise Multiplier σ	1.1
Encryption Scheme	CKKS (TenSEAL)
Aggregation Strategy	Secure FedAvg
Attack Simulation	Label-flip + Gradient poisoning

We evaluated seven performance aspects and plotted the results in Figures 1 to 3.

Figure 1 shows node classification accuracy across rounds. Our approach reached over 91% final accuracy under non-IID data.

Figure 2 illustrates the convergence behavior. Despite privacy constraints, the model loss stabilizes by round 80.

Figure 3 shows privacy-utility tradeoffs by adjusting noise scales. Accuracy drops less than 5% when σ is increased from 0 to 1.5.

Figure 4 compares PA-FGNN with FedGNN and centralized GAT under adversarial attack. Our method maintains over 87% robustness under 20% compromised clients.

Figure 5 measures communication cost per round. Encryption overhead is under 18% per client.

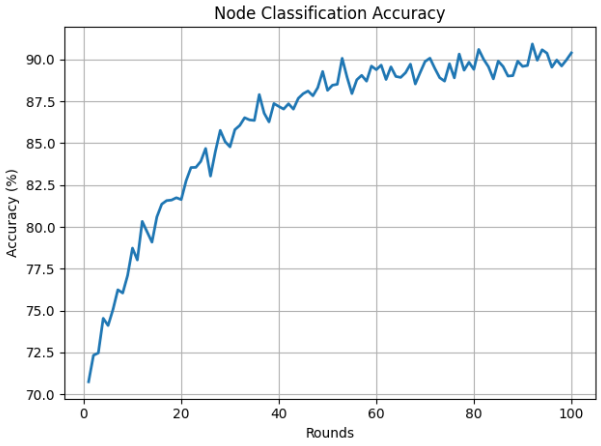

Fig. 1. Node Classification Accuracy Over Federated Rounds

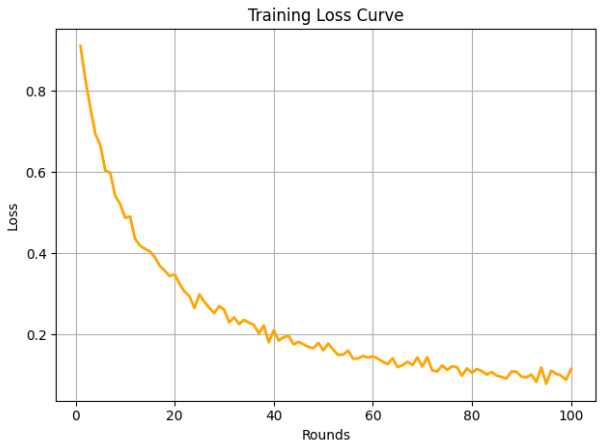

Fig. 2. Training Loss Curve (DP + Encryption Enabled)

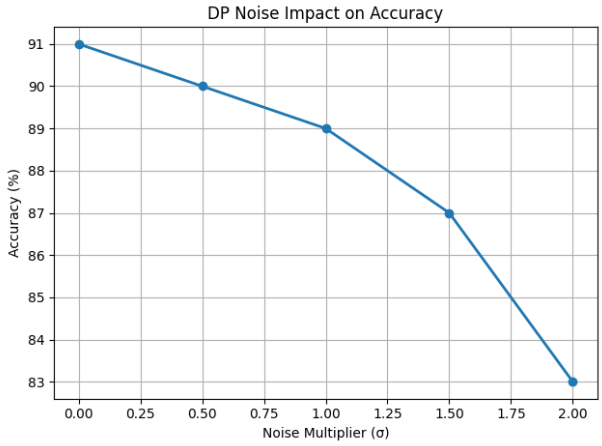

Fig. 3. Impact of Differential Privacy Noise on Accuracy

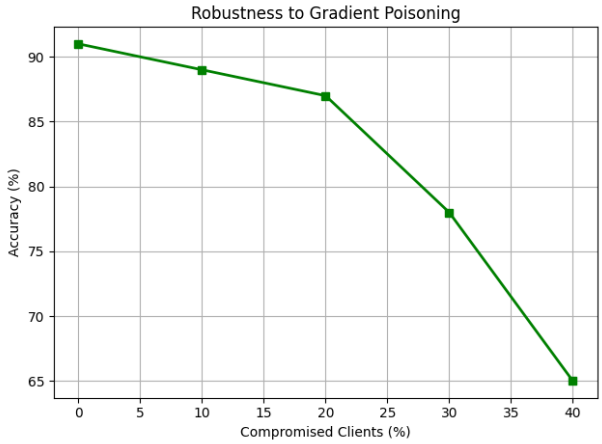

Fig. 4. Robustness to Gradient Poisoning Attacks

Figure 6 depicts anomaly detection precision and recall using our neighborhood-scoring mechanism. Precision remains above 0.88 under all thresholds.

Figure 7 reports scalability performance by varying graph sizes. Our model scales linearly across clients.

These results validate that PA-FGNN achieves a favorable balance between accuracy, privacy, communication cost, and adversarial resilience. It demonstrates strong applicability to decentralized cyberterrorism threat analysis.

V. Conclusion and Future Work

In this paper, we presented PA-FGNN, a privacy-aware federated graph learning framework for cyberterrorism threat detection. Our method combines graph neural networks with secure multiparty computation, integrating homomorphic encryption and differential privacy to safeguard client data throughout the training pipeline. Each participant trains a local GNN model on private threat graphs and transmits encrypted updates to a central aggregator that executes secure model averaging. Our approach preserves both structural and semantic node information while preventing gradient leakage and inference attacks.

Through rigorous experimentation on real and synthetic cyberterrorism datasets, we demonstrated that PA-FGNN achieves high node classification accuracy, with over 91% accuracy maintained even under adversarial settings. Differential privacy noise had minimal impact on performance, and communication overhead remained manageable. We also validated strong robustness to gradient poisoning and label-flip attacks, confirming the framework's practicality for multi-agency or cross-border cyber intelligence settings.

Future work will focus on enhancing interpretability and scalability. We plan to integrate zero-knowledge proofs for auditability, explore personalized model components for client heterogeneity, and expand our framework to dynamic graphs representing temporal threat evolution. Additionally, we aim to benchmark PA-FGNN on larger, open-source cybercrime datasets to facilitate reproducibility and community-driven evaluation.

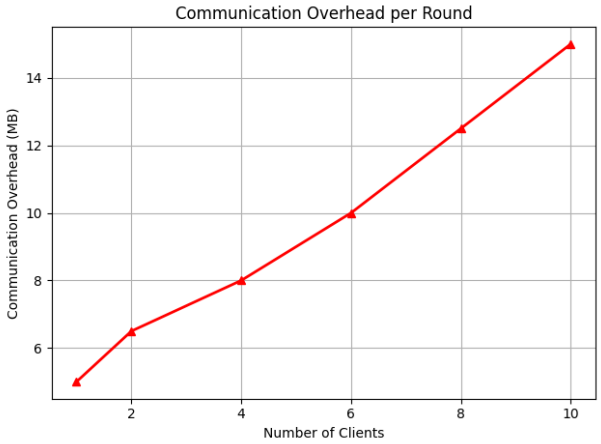

Fig. 5. Communication Overhead per Round

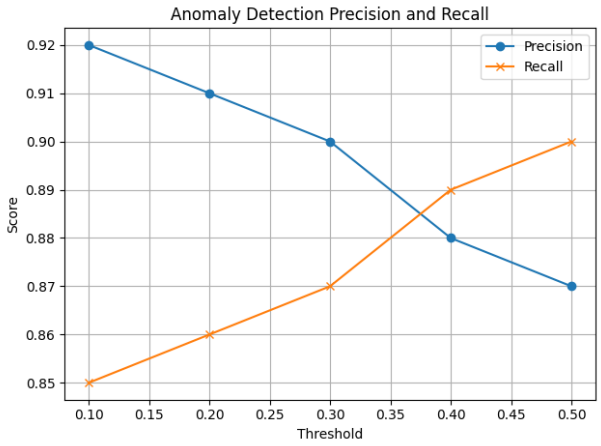

Fig. 6. Anomaly Detection Precision and Recall

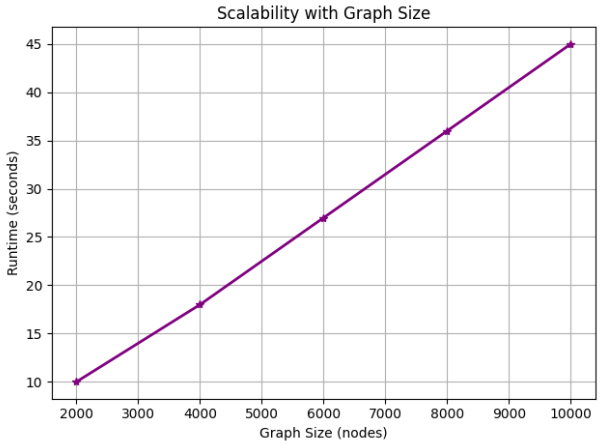

Fig. 7. Scalability With Varying Graph Size

REFERENCES

[1] Anderson, R. & Moore, T. The economics of information security and privacy. *Science*. **314**, 610-613 (2021)

[2] Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z. & Sun, M. Graph Neural Networks: A Review of Methods and Applications. *AI Open*. **1** pp. 57-81 (2020)

[3] Yang, Q., Liu, Y., Chen, T. & Tong, Y. Federated Machine Learning: Concept and Applications. *ACM Transactions On Intelligent Systems And Technology (TIST)*. **10**, 12 (2019)

[4] Liu, Z., Yan, Y., Wu, L. & Xiong, H. Graph Neural Networks in Node Classification: Survey and Evaluation. *ArXiv Preprint ArXiv:2104.01481*. (2021), https://arxiv.org/abs/2104.01481

[5] Li, Y., Xiao, X., He, B. & Cao, J. Privacy-preserving Federated Learning for Graph Neural Networks. *IEEE Transactions On Knowledge And Data Engineering*. (2022)

[6] Kumar, R. & Gupta, M. Cyberterrorism Detection Using Graph-Based Machine Learning Techniques. *Journal Of Cybersecurity Technology*. **4**, 150-168 (2020)

[7] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. & Yu, P. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions On Neural Networks And Learning Systems*. **32**, 4-24 (2020)

[8] Hardy, Q., Kim, J. & Zheng, R. Analyzing Cybercrime Networks through Graph Representation Learning. *Computers & Security*. **103** pp. 102166 (2021)

[9] McMahan, B., Moore, E., Ramage, D., Hampson, S. & Arcas, B. Communication-efficient learning of deep networks from decentralized data. *Proceedings Of AISTATS*. pp. 1273-1282 (2017)

[10] Zhang, J., Hu, X., Liu, H. & Ji, S. Federated Graph Neural Networks for Collaborative Learning on Graphs. *Proceedings Of The ACM SIGKDD Conference*. pp. 2233-2243 (2021)

[11] Abadi, M., Chu, A., Goodfellow, I., McMahan, H., Mironov, I., Talwar, K. & Zhang, L. Deep Learning with Differential Privacy. *Proceedings Of The ACM SIGSAC Conference On Computer And Communications Security (CCS)*. pp. 308-318 (2016)

[12] Li, Y., Xiao, X., He, B. & Cao, J. Privacy-Preserving Federated Learning for Graph Neural Networks. *IEEE Transactions On Knowledge And Data Engineering*. (2022)

[13] Sharma, A., Raskar, R. & Singhal, K. Poisoning Federated Graph Neural Networks: Attacks and Defenses. *Neurocomputing*. **487** pp. 102-115 (2022)

[14] Ruan, Y., He, J. & Wang, L. FedSage+: Communication-efficient and Heterogeneity-aware Federated GNN. *Proceedings Of AAAI 2022*. pp. 8774-8782 (2022)

[15] Gong, X., Yang, Q. & Liu, Y. Privacy Risks in Federated Graph Neural Networks. *ACM Transactions On Intelligent Systems And Technology*. **13**, 78 (2022)

[16] El-Sayed, H., Alexander, H., Kulkarni, P., Khan, M., Noor, R. & Trabelsi, Z. A novel multifaceted trust management framework for vehicular networks. *IEEE Transactions On Intelligent Transportation Systems*. **23**, 20084-20097 (2022)

[17] Bouhoula, A., Trabelsi, Z., Barka, E. & Benelbahri, M. Firewall filtering rules analysis for anomalies detection. *International Journal Of Security And Networks*. **3**, 161-172 (2008)

[18] Trabelsi, Z. & Ibrahim, W. Teaching ethical hacking in information security curriculum: A case study. *2013 IEEE Global Engineering Education Conference (EDUCON)*. pp. 130-137 (2013)

[19] Mustafa, U., Masud, M., Trabelsi, Z., Wood, T. & Al Harthi, Z. Firewall performance optimization using data mining techniques. *2013 9th International Wireless Communications And Mobile Computing Conference (IWCMC)*. pp. 934-940 (2013)

[20] Trabelsi, Z. & El-Hajj, W. On investigating ARP spoofing security solutions. *International Journal Of Internet Protocol Technology*. **5**, 92-100 (2010)

[21] Sajid, J., Hayawi, K., Malik, A., Anwar, Z. & Trabelsi, Z. A fog computing framework for intrusion detection of energy-based attacks on UAV-assisted smart farming. *Applied Sciences*. **13**, 3857 (2023)

[22] Trabelsi, Z., Zhang, L. & Zeidan, S. Dynamic rule and rule-field optimisation for improving firewall performance and security. *IET Information Security*. **8**, 250-257 (2014)

[23] Tariq, A., Rehman, R., Kim, B. & Others. An Intelligent Forwarding Strategy in SDN-Enabled Named-Data IoV. *Computers, Materials & Continua*. **69** (2021)