**RESEARCH PAPER**

# Hierarchical Spatio-Temporal Graph Auto-encoded Federated Contrastive Learning For Attack Detection in VPN-Assisted Wireless IoT Network.

S.Sangeetha M.Sc.,M.Phil.,M.ED [1], Dr. L.Sudha M.Sc.,M.C.A.,M.Phil.,Ph.D[2]

*[1]Research scholar, Department of Computer Science, Navarasam Arts and Science College for Women , Arachalur, Erode-638101 Affiliated to Bharathiar University, Coimbatore, Tamil Nadu, India*
*Email :Sansen1941@gmail.com*
*[2]Research Supervisor,Associate Professor, School Computer Science, VET Institute ofAtrs and Science College , Erode-12*
*Email : sudhal@vetias.ac.in*

**ABSTRACT**
Wireless Internet of Things (IoT) networks utilize Virtual Private Networks (VPN) to secure connections among heterogeneous and devices with limited computational resources. Though, encrypted VPN tunnels are still vulnerable to internal and compromised-node threats which cannot be significantly identified by state-of-the-art systems relying on packet payload analysis and centralized learning. With the aiming at handling these problems, this paper introduces a Hierarchical Spatio-Temporal Graph Auto-Encoder with Federated Contrastive Learning (HST-GA-FCL) for attack identification in VPN-supported wireless IoT environments. The HST-GA-FCL learns communication behavior with the assist of multi-level spatio-temporal graphs that capture device-, cluster-, and network-level relations. An unsupervised spatio-temporal graph auto-encoder in proposed HST-GA-FCL finds normal traffic patterns where a variation in reconstruction and latent representations is designates as attack. Federated learning in proposed HST-GA-FCL facilitates privacy-preserving collaborative training across IoT nodes while contrastive learning enhances representation reliability across assorted devices. The proposed HST-GA-FCL exactly predicts internal adversaries in encrypted VPN traffic with better accuracy. Simulation analysis across various IoT attack situations demonstrates that the proposed HSTGAE-FCL Model attains better performance in terms of detection accuracy, computational complexity, packet delivery ratio and false alarm rate when compared to existing systems.

**Keywords**: Attack, Federated Contrastive Learning, Internet of Things Spatio-Temporal Graph Auto-Encoder and Virtual Private Networks
**How to cite this article:** Sangeetha S, Sudha L.. Hierarchical Spatio-Temporal Graph Auto-encoded Federated Contrastive Learning For Attack Detection in VPN-Assisted Wireless IoT Network...Int J Drug Deliv Technol. 2026;16(3s): 136-144; DOI: 10.25258/ijddt.16.3s.18
**Source of support:** Nil.
**Conflict of interest:** None

## INTRODUCTION

Wireless IoT systems are extensively used in industrial automation, healthcare monitoring, smart cities, and consumer applications where secure data exchange is attained through VPN concepts. While VPNs perform encryption with better confidentiality, but IoT deployments still face considerable security threats owing to limited resource capabilities, varied software environment, and rapidly changing wireless network architecture. These features expand the attack surface and make centralized intrusion monitoring insufficient in strong privacy guarantees, bandwidth, and latency requirements. In the past decade, research in graph-based learning, spatio-temporal representation, self-supervised learning, and federated optimization has presented solution for attack identification. However, few key problems have yet to be resolved.

The state-of-the-art graph-based model are inadequate to static or mildly dynamic network architectures and do not scale well to streaming IoT systems operating over VPN channels. Besides to that, contrastive learning concept depends on relatively homogeneous information sharing's and their reliability is compromised in multi-tenant VPN-secured IoT architectures. As well, federated learning solutions often fail to notice the united modeling of hierarchical spatial configurations ranging from individual devices to clusters and global network performance together with temporal dependence. Accordingly, only limited studies have explored the combined utilize of hierarchical graph representations, spatio-temporal autoencoding, and federated contrastive learning to efficiently predict unseen attacks in VPN network. Motivated by these limitations, HSTGAE-FCL Model is implemented in this research work.

In existing, a self-supervised learning (SSL) concept was intended in [1] with the objective of increasing malicious discovery performance in encrypted network traffic in which traffic metadata and arithmetical flow features are utilized to observe normal actions without considering payload evaluation. The SSL notices threats in encrypted environments while guarding data privacy. But, it impacts scalability across distributed an IoT network which decreases detection accuracy. A Federated Deep Learning (FDL) was planned in [2] by utilizing mutual information regularization to recognize anomalies across IoT devices. Despite privacy advantages, the FDL fail to notices inter-device connections, because it does not integrate graph structure.

**Author for Correspondence: S.Sangeetha*

A contrastive self-supervised approach (CoLA) was introduced in [3] with the goal of finding maclicious nodes in attributed graphs. However CoLA depends on static graph representation limits its usage in dynamic IoT environments with multi-level connectivity. Moreover, a federated framework (FedIoT) was implemented in [4] with the objective of performing on-device anomaly discovery. Though, conventional FedIoT attained better feasibility, but it does not consider multi-scale spatial behaviors in huge IoT environment.

A GraphSAGE was presented in [5] through integrating edge features for identifying network intrusion. Although it enhances detection performance, it still based on labeled traffic information. As well, contrastive self-supervision was performed in [6] with the target of strengthening abnormality detection accuracy while labeled samples are inadequate. Though, diverse wireless IoT data was not considered for stable feature augmentations.

A graph neural network was intended in [7] with the target of increasing threat detection efficiency in IoT traffic patterns. But, it does not deal with decentralized or privacy-limited learning systems. A clustered FL approach was constructed in [8] to diminish inter-device heterogeneity in big IoT environment. Clustering lightens distribution skew however it depends on precise cluster assignments which impacts explicit representation of hierarchical graph structure and time-varying connectivity.

A spatio-temporal graph attention model was introduced in [9] for recognizing video anomalies through utilizing GCN and GAT algorithm. Though it was powerful for dense vision information, but it performed was lacked when sparse data was considered as input. A transformer-enhanced graph auto-encoder was utilized in [10] for finding temporal and spatial communications. However, it limits their operations on resource-constrained IoT hardware.

## LITERATURE STUDY

In the state-of-the-art works, there has been increasing research attention toward highly developed graph-based learning, spatio-temporal representation, and federated intelligence for strong, reliable attack identification. Spatial-temporal graph auto encoders were developed in [11] to significantly discover patterns in video-based datasets. However, sparse IoT network was considered. Moreover, federated contrastive learning combined with transformer models was presented in [12] for device personalization. But, this method was lacked by communication overhead.

A relaxed contrastive loss was focused in [13] to manage heterogeneity in federated conditions. The reduced contrastive constraints enhance association however it based on expensive negative sampling which inappropriate for large-scale IoT applications. A multi-scale contrastive learning was presented in [14] that evaluate patch-level and context-level representations to find out anomalies. However, a data sharing in streaming IoT environments was not focused. As well, gADAM was implemented in [15] to boost GNN robustness via adaptive node augmentations. Though, mixup-based synthetic information can make unnatural graph structures.

A federated ensemble intrusion detection was done in [16] which proving its resilience to label noise. But, integrating ensembles in federated mode considerably raises aggregation complexity and communication cost in massive IoT environment. A review of various federated learning was presented for IoT intrusion discovery and their open issues such as heterogeneity and unreliable client involvement was analzed in [17]. The review also limited work on analyzing hierarchical graph structures in dynamic IoT networks. Privacy-aware anomaly recognition was performed in [18] for smart home application with strict data minimization polices. However, sophisticated coordinated attacks were not concentrated.

A time–frequency contrastive learning was utilized in [19] to increase generalization for temporal signals. An Extreme Gradient Boosting Multifactor Ensemble classifier was presented in [20] to enhance the attack discovery performance in IoT network. Besides to that, a federated contrastive concept was called Fed-CLR for cross-domain representation. Though, localized or rare IoT attacks were not identified. A spatial–temporal GCN was implemented in [21] for abnormality discovery in surveillance videos. But, false alarm rate was higher. For solving the above problem, a novel HSTGAE-FCL Model is introduced in this work.

## PROPOSED HSTGAE-FCL MODEL

The HSTGAE-FCL model is developed to detect anomalies in VPN-enabled wireless IoT networks while ensuring data privacy and robustness. The HSTGAE-FCL model identifies attacks that operate within encrypted tunnels by modeling communication behavior rather than packet payloads. The designed HSTGAE-FCL model integrates four major components i.e. hierarchical graph construction, spatio-temporal graph encoding, anomaly reconstruction with multi-level embedding fusion, and federated contrastive learning for privacy-preserving optimization. The architecture of proposed HSTGAE-FCL model is presented in Figure 1.

**Federated Contrastive Learning**

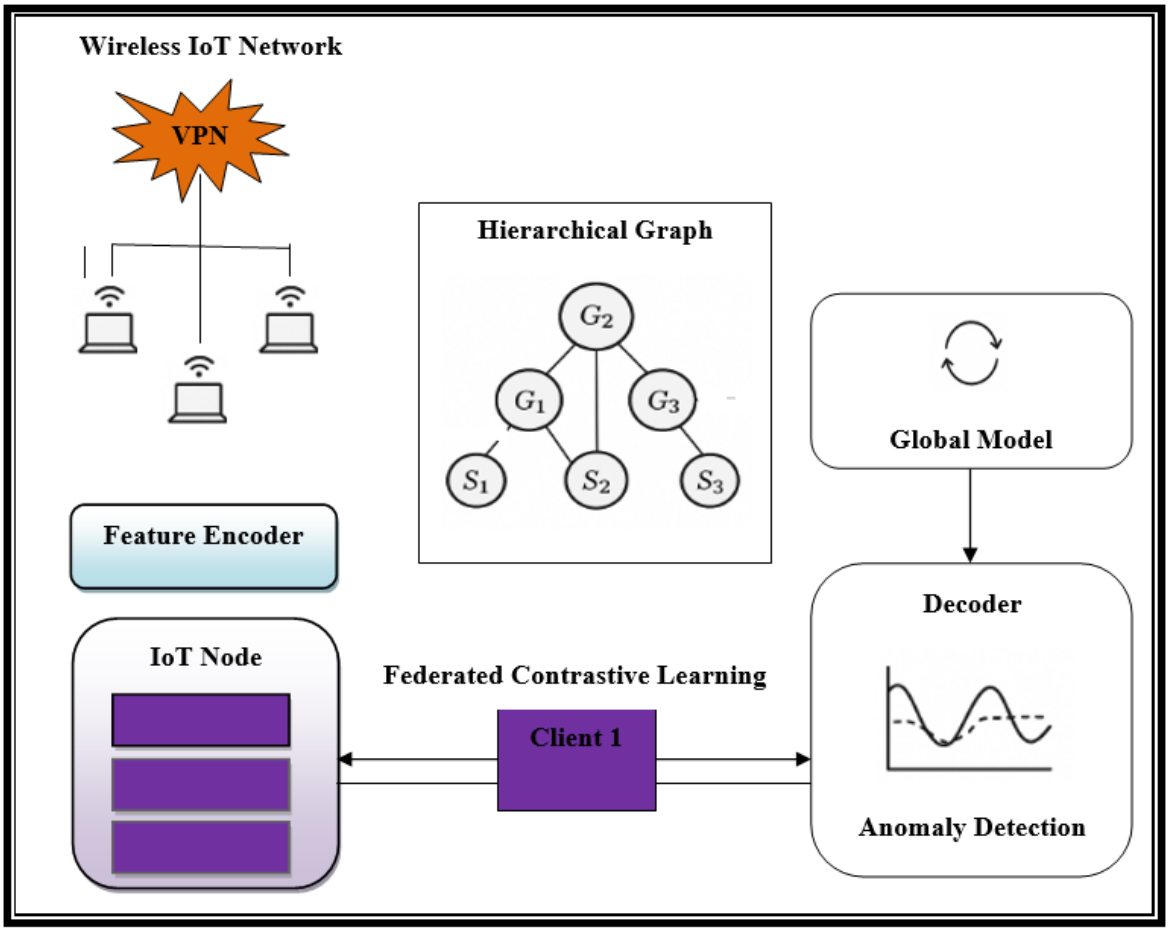


**Figure 1 Architecture of HSTGAE-FCL Model for Attack detection**

Figure 1 illustrates the end-to-end process of HSTGAE-FCL Model for privacy-preserving anomaly discovery. In the network layer, numerous wireless IoT devices communicate through a VPN which ensures encrypted and secure data transmission. Although VPNs guard confidentiality, compromised IoT devices can still formulate malicious traffic i.e. flooding or unauthorized access attempts. Therefore, attack identification is performed inside the VPN tunnel rely on behavioral analysis rather than payload inspection. At each IoT node, raw network traffic metadata are processed in HSTGAE-FCL Model with the help of Feature Encoder. The encoded features are then organized into a hierarchical graph structure. To preserve privacy and scalability, the model is trained using FCL. Each IoT node trains its local hierarchical graph model using contrastive learning objectives, learning normal behavioral patterns without sharing raw data. Only model updates or embeddings are transmitted to the Global Model where they are aggregated to form a robust and generalized representation of normal network behavior across all clients. The decoder and anomaly detection module in HSTGAE-FCL Model reconstructs expected normal behavior from the learned embeddings. Deviations between reconstructed and observed behavior determined using reconstruction error which is considered to estimate an anomaly score. Significant deviations indicate compromised devices, abnormal traffic flows, or insider attacks occurring within the VPN-protected IoT network.

### 3.1 SYSTEM MODEL

Consider a VPN-enabled wireless IoT networks with $N$ devices deployed across multiple clusters and connected through a VPN gateway. Each device generates time-indexed network telemetry such as flow statistics, control messages, and device-level indicators. Since payloads are encrypted, only metadata and behavioral features are available for analysis. The objective is to detect anomalous or malicious behavior by learning the normal spatio-temporal interaction patterns of IoT devices in an unsupervised and privacy-preserving manner without centralized raw data collection. The wireless IoT system model using VPN is shown in Figure 1.

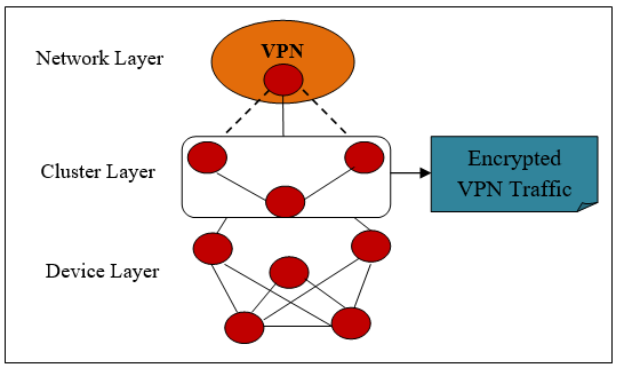


**Figure 1 VPN-Assisted Wireless IoT System Model**

VPN may secure data in wireless IoT network though a compromised device can still flood the network or try unauthorized access. That means VPN encrypts traffic and conceals it from external observers, however if the endpoint device is affected, malicious activities (i.e. network flooding (DDoS), unauthorized access attempts, or insider threats) can occur. Therefore, HSTGAE-FCL model is introduced in this research paper for presenting a robust, privacy-preserved solution for attack discovery in VPN-supported wireless IoT applications.

### 3.2 Hierarchical Multi-Scale Graph Representation

The HSTGAE-FCL model constructs a Hierarchical Multi-Scale Graph Modeling (GMSGM) to capture structural variations in large VPN-enabled wireless IoT networks. This GMSGM design preserves fine-grained local behavior while enabling broader context awareness for detecting complex attacks. The traditional graph models were based on local neighborhood, single-scale representation. To overcome these, GMSGM is implemented in this paper where it captures local, regional, and global interactions in a hierarchical manner and aggregate information across multiple scales to enrich node and graph-level representations. Besides, GMSGM significantly reduce computational complexity by summarizing nodes at higher hierarchical levels. For finding the multi-scale dependencies, the VPN-supported wireless IoT system is considered as a hierarchical graph structure.

**Device-Level-Graph:** The device graph at the time instance '$t$' is represented using

$$g_t^{(d)} = (v^{(d)}, e_t^{(d)}, x_t^{(d)}) \quad (1)$$

In equation (1), $v^{(d)}$ indicates IoT devices and $e_t^{(d)}$ describes wireless or logical communication links whereas $x_t^{(d)} \in \mathrm{R}^{\mathrm{N}\times\mathrm{F}}$ shows node features.

**Cluster-Level-Graph:** Devices are grouped into clusters (i.e. gateways, functional zones) and which is formed using,

$$g_t^{(c)} = (v^{(c)}, e_t^{(c)}, x_t^{(c)}) \quad (2)$$

In equation (2), each cluster node combines statistics from its member devices.

**Network-Level-Graph:** a global graph finds inter-cluster communication through a VPN infrastructure and which is formed using

$$g_t^{(g)} = (v^{(g)}, e_t^{(g)}, x_t^{(g)}) \quad (3)$$

From equation (3), temporal dependencies are integrated with stacking graphs over a sliding window of length $T$ and this gives a Spatio-temporal graph sequence.

### 3.3 Spatio-Temporal Graph Auto-Encoder

To find normal communication activities, an unsupervised spatio-temporal graph auto-encoder is utilized in HSTGAE-FCL model at each hierarchy level.

**Encoder:** The spatial encoder utilize graph convolution to find out relational features using,

$$h_t = \sigma(\tilde{D}^{-1/2}\tilde{a}_t\tilde{D}^{-1/2}x_t w_s) \quad (4)$$

In equation (4), $\tilde{a}_t = a_t + I$ whereas $\tilde{D}$ describes degree matrix and $w_s$ shows learnable weight matrix. Temporal evolution is obtained with the support of gated recurrent units (GRU) mathematically using,

$$z_t = GRU(h_t, z_{t-1}) \quad (5)$$

**Decoder:** The decoder rebuilds both node features and graph structure using,

$$\hat{x_t} = f_d(z_t), \qquad \hat{a_t} = \sigma(z_t z_t^T) \quad (6)$$

**Reconstruction Loss:** The auto-encoder is trained through reducing,

$$\mathcal{L}oss_{re} = \sum_{t=1}^{T}(\|x_t - \hat{x_t}\|_2^2) + \lambda\,\|a_t - \hat{a_t}\|_F^2 \quad (7)$$

In equation (7), high reconstruction errors point out variations from normal characteristics and which are considered as anomalies.

### 3.4 Federated Learning

To achieve privacy, model is trained with the aid of FL concept across IoT nodes or edge gateways. From that, each client $k$ optimizes its local objective using,

$$\mathcal{L}oss_k = \mathcal{L}oss_{re}^k + \alpha\mathcal{L}oss_{con}^k \quad (8)$$

After updating local objective, model parameters are combined at the server using FedAvg which mathematically obtained as,

$$W^{(t+1)} = \sum_{k=1}^{K}\frac{m_k}{m}W_k^{(t)} \quad (9)$$

In equation (9), $m_k$ represents the data size of client $k$.

### 3.5 Federated Contrastive Learning

Because of non-IID data sharing's, FCL is implemented to align representations across varied clients. Given local embeddings $Z_i$ and global prototypes $P$, the contrastive loss is mathematically obtained as,

$$\mathcal{L}oss_{con} = -\log\frac{\exp(sim(Z_i,P^+)/\tau)}{\sum_j \exp(sim(Z_i,P_j)/\tau)} \quad (10)$$

In equation (10), $\tau$ indicate temperature parameter whereas $sim(\cdot)$ describe cosine similarity. This supports reliable latent representations without distributing raw data. In non-IID data sharing's, client/device data is not related and does not follow the same distribution. Each device may have dissimilar patterns, behaviors, feature ranges, attack natures.

### 3.6 Attack Detection using Anomaly Score

For each node at a time , an anomaly score $\omega_t$ is mathematically determined using,

$$\omega_t = \beta\|x_t - \hat{x_t}\|_2 + (1-\beta)\|z_t - \hat{z_t}\|_2 \quad (11)$$

In the above equation (11), $\omega_t > T$ when the anomaly score $\omega_t$ is greater than threshold value ($T$), the proposed model flagged node behavior as anomalous. From that, HSTGAE-FCL model efficiently identifies internal attacks, flooding, impersonation, and lateral movement within encrypted VPN traffic environment with better computational complexity.

The algorithmic step of HSTGAE-FCL model is described as,

**// Hierarchical Spatio-Temporal Graph Auto-Encoder With Federated Contrastive Learning Algorithm**
**Input**: Number of devices '$v_1, v_2, .. v_N$' in VPN enabled IoT environment
**Output:** Achieve better accuracy for attack detection
**Step 1: Begin**
**Step 2:** **For** all the node devices '$v_i$'
**Step 3:** Construct hierarchical graphs from in VPN assisted IoT topology
**Step 4:** Build Device-Level-Graph $g_t^{(d)}$ using (1)
**Step 5:** Create Clsuter-Level-Graph $g_t^{(c)}$ using (2)
**Step 6:** Design Network-Level-Graph $g_t^{(g)}$using (3)
**Step 7:** Encode temporal and device features using (4) and (5)
**Step 8:** Decoder recreates both node features and graph structure using (6)
**Step 9:** Auto-encoder is trained through minimizing reconstruction loss using (7)
**Step 10:** Apply FCL to learn representations across various clients
**Step 11:** Determine contrastive loss $\mathcal{L}oss_{con}$
**Step 12:** Caculate anomaly score $\omega_t$ for each device
**Step 13:** **If** ($\omega_t > T$, **then**
**Step 14:** Device $v_i$ is flagged as anomaly
**Step 15:** **Else**
**Step 16:** Device $v_i$ is flagged as normal
**Step 17:** **End If**
**Step 18:** **End For**
**Step 19:End**

**Algorithm 1 Hierarchical Spatio-Temporal Graph Auto-Encoder with Federated Contrastive Learning**

By using the above algorithmic processes, spatio-temporal modeling in HSTGAE-FCL model captures both device-level and network-level anomalies. Besides, hierarchical graph approach in HSTGAE-FCL model enables scalability to large VPN-assisted IoT networks. As well, proposed HSTGAE-FCL model preserves privacy while improving generalization. Moreover that, proposed HSTGAE-FCL model robust to heterogeneous devices and dynamic VPN enabled IoT network topologies when compared to existing works.

## IV. EXPERIMENTAL SETUP

Proposed HSTGAE-FCL model is implemented in an NS2.34 simulator with the goal of analyzing their attack detection performance. The number 40 to 400 node devices are initially defined in squared area of $A^2$ (1200 m * 1200 m) in VPN supported IoT network environment. Besides, IoT-23 Dataset from **https://www.stratosphereips.org/datasets-iot23** is considered as input dataset during simulation task where it includes of 20 malware captures in IoT devices and 3 captures for benign IoT devices traffic. The simulations parameters are depicted in Table 1.

**Table 1 Simulation Constraints**

| Simulation Parameter | Value |
|---|---|
| VPN Tunnel Type | IPsec / OpenVPN (emulated) |
| Simulation Time | 300 s |
| Mobility Model | Random Waypoint |
| Routing Protocol | DSR |
| Speed of Node | 0 – 20 m/s |
| Traffic Type | CBR / UDP |
| Size of Packet | 512 bytes |
| Data Packet Rate | 25 – 250 packets |
| Transmission Range | 250 m |
| MAC Protocol | IEEE 802.11 |
| Radio Propagation Model | Two-Ray Ground |
| Initial Energy per Node | 1000 Joules |
| Attack Types | Flooding, Unauthorized Access, Impersonation |
| VPN Attack Scenario | Compromised Node Inside VPN |

## PERFORMANCE EVALUATION METRICS

The effectiveness of HSTGAE-FCL model is determined using metrics such as.

Attack Detection Accuracy
Communication Overhead
Packet Delivery Ratio
False Alarm Rate

**1. Attack Detection Accuracy (ADA):** It determines overall exactness of anomaly discovery using,

$$ADA = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

In equation (12), True Positive (TP) defines number of attack devices rightly discovered as attacks and True Negative (TN) describes number of normal devices exactly detected as normal, False Positive (FP) indicates number of normal devices erroneously predicted as attacks and False Negative (FN) refers number of attack devices inaccurately noticed as normal.

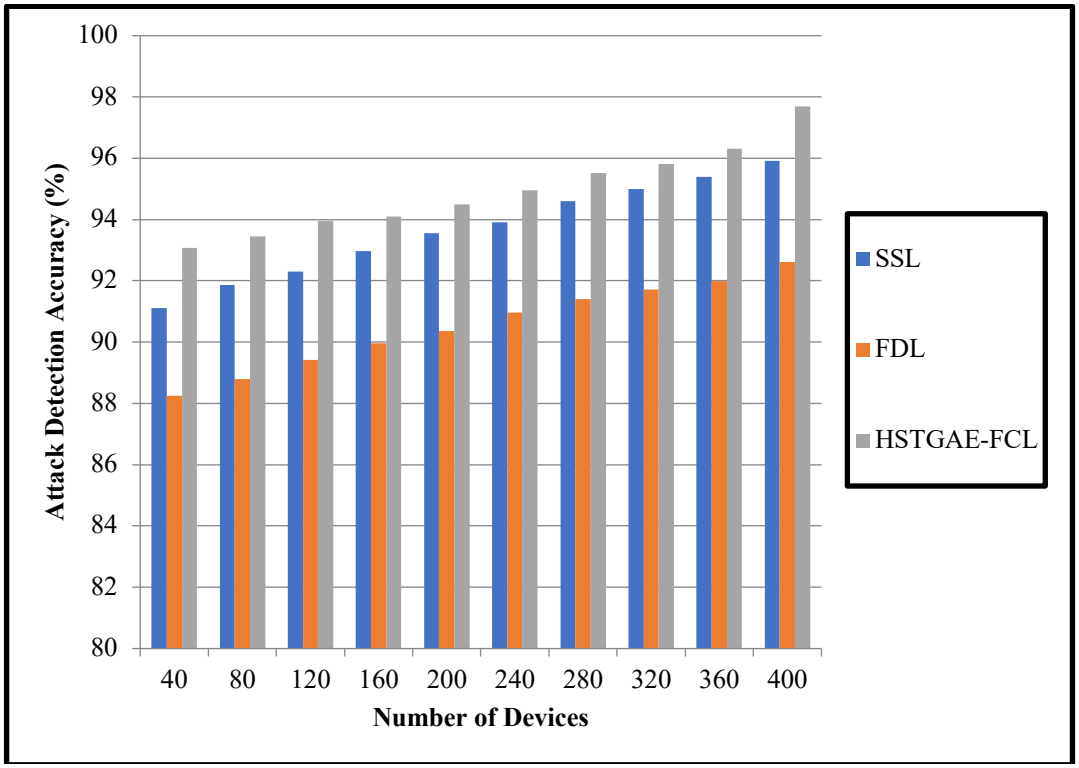


**Figure 3 Graphical Performance of Attack Detection Accuracy**

Figure 3 illustrates the impact of network scale on attack detection accuracy for existing SSL [1], FDL [2] and the proposed HSTGAE-FCL model evaluated across increasing numbers of IoT nodes ranging from 40 to 400. As the number of nodes increases, all three methods exhibit a gradual improvement in detection accuracy. This trend indicates that larger networks provide richer behavioral and traffic patterns, enabling learning models to better characterize normal and anomalous activities. The existing SSL-based method achieves moderate accuracy, benefiting from unsupervised representation learning but lacking collaborative knowledge sharing across distributed devices. The existing FDL approach consistently yields lower accuracy due to limited capability in modeling complex spatio-temporal dependencies and inter-device relationships, particularly in heterogeneous IoT environments. In contrast, the proposed HSTGAE-FCL method consistently outperforms both baseline approaches across all network sizes. As a result, HSTGAE-FCL achieves the highest accuracy, demonstrating strong scalability and robustness in large-scale VPN-assisted wireless IoT networks.

**2. Communication Overhead (CO):** It determines total time utilized to perform privacy preserved data transmission in VPN assisted IoT environment through attack identification which obtained mathematically as,

$$CO = \sum_{i=1}^{N} v_i * time\ [ADN] \quad (13)$$

In equation (16), '$v_i$' denotes a node devices and '$time\ [CSN]$' represented a time employed to precisely detect the node devices as normal or malicious for effective communication. It is observed in terms of milliseconds (ms).

**Table 2 Testing Results of Communication Overhead**

| Number of Devices | Communication Overhead (ms) | | |
|---|---|---|---|
| | SSL | FDL | HSTGAE-FCL |
| 40 | 14.5 | 15.9 | 11.6 |

| | | | |
|---|---|---|---|
| 80 | 16.8 | 17.4 | 13.4 |
| 120 | 18.6 | 19.3 | 15.3 |
| 160 | 20.4 | 21.5 | 17.1 |
| 200 | 22.7 | 23.7 | 19.8 |
| 240 | 24.3 | 25.8 | 21.5 |
| 280 | 26.9 | 27.6 | 23.7 |
| 320 | 28.1 | 29.3 | 25.2 |
| 360 | 30.4 | 31.2 | 27.5 |
| 400 | 32.5 | 33.6 | 29.8 |

Table 2 presents a comparative analysis of communication overhead for conventional SSL [1], FDL [2] and the proposed HSTGAE-FCL model under varying numbers of IoT devices, ranging from 40 to 400. As the number of devices increases, communication overhead rises for all three approaches. The conventional FDL approach consistently exhibits the highest communication overhead across all network sizes due to the transmission of full model parameters during federated aggregation rounds. Similarly, conventional SSL method incurs moderate overhead, as it relies on repeated parameter exchanges and lacks mechanisms to minimize redundant updates. In contrast, the proposed HSTGAE-FCL framework achieves the lowest communication overhead at every network scale. This improvement is primarily attributed to usage of contrastive learning mechanism enables faster convergence, thereby reducing the frequency and duration of communication rounds. The hierarchical structure further minimizes redundant updates by aggregating information at device, cluster, and global levels. Thus, Proposed HSTGAE-FCL reduces communication overhead in VPN-assisted wireless IoT networks.

**3. Packet Delivery Ratio (PDR):** It determines the reliability of data transmission in VPN supported IoT network using,

$$PDR = \frac{Total\ Number\ of\ Packets\ Received}{Total\ Number\ of\ Packets\ sent} \quad (14)$$

Higher PDR indicates that more reliable and secure communication.

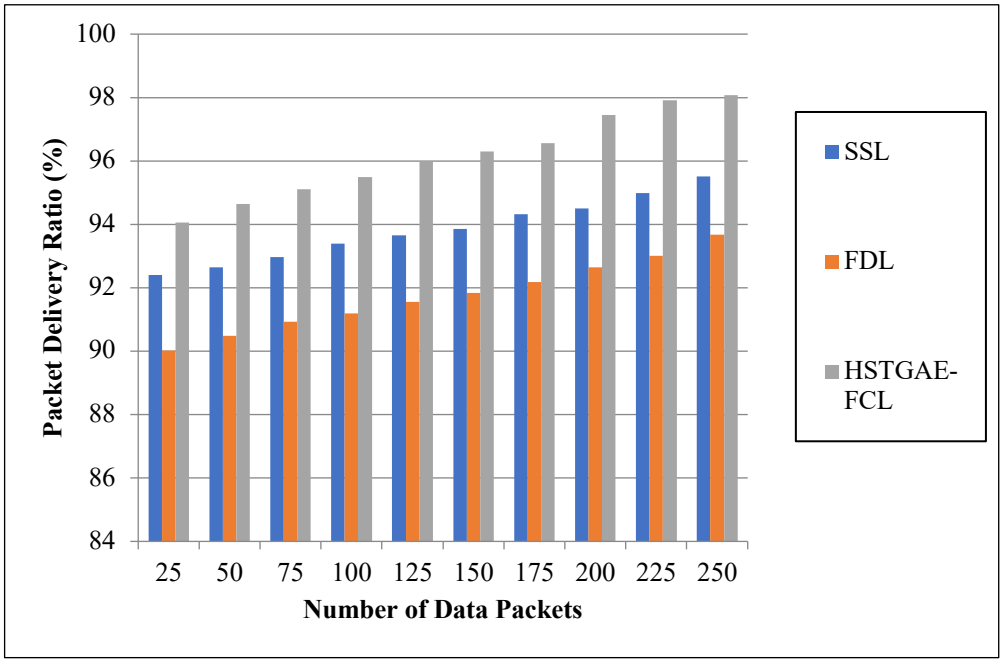


**Figure 4 Graphical Performance of Packet Delivery Ratio**

Figure 4 illustrates the testing results of PDR for state-of-the-art SSL [1], FDL [2] and the proposed HSTGAE-FCL. As the number of data packets increases, all three methods demonstrate a gradual improvement in PDR. This trend indicates that learning-based traffic management and anomaly detection mechanisms become more effective as the system observes richer traffic patterns. The state-of-the-art FDL approach consistently achieves the lowest PDR due to higher communication overhead and frequent model synchronization, which increases network congestion and packet loss, particularly in VPN-assisted environments. The

state-of-the-art SSL method performs better than FDL by avoiding centralized data exchange, yet it lacks collaborative learning across nodes, limiting its ability to adapt to evolving traffic conditions. In contrast, the proposed HSTGAE-FCL model achieves the highest PDR across all packet volumes. This improvement is attributed to its hierarchical spatio-temporal graph modeling, which effectively captures traffic dependencies at device, cluster, and network levels. Moreover, federated contrastive learning reduces unnecessary communication while maintaining accurate anomaly detection, thereby minimizing retransmissions and packet drops. These results confirm that the proposed framework enhances both security and communication reliability in VPN-enabled wireless IoT networks.

**4. False Alarm Rate (FAR):** It determines incorrectly flagged normal traffic using,

$$FAR = \frac{FP}{FP+TN} \quad (15)$$

Lower FAR gives improved system reliability for attack identification.

**Table 3 Testing Results of False Alarm Rate**

| Number of Devices | False Alarm Rate (%) | | |
|---|---|---|---|
| | SSL | FDL | HSTGAE-FCL |
| 40 | 8.89 | 11.76 | 6.92 |
| 80 | 8.14 | 11.21 | 6.54 |
| 120 | 7.7 | 10.58 | 6.05 |
| 160 | 7.02 | 10.04 | 5.9 |
| 200 | 6.45 | 9.65 | 5.51 |
| 240 | 6.08 | 9.03 | 5.05 |
| 280 | 5.4 | 8.6 | 4.49 |
| 320 | 5.01 | 8.29 | 4.2 |
| 360 | 4.6 | 8.02 | 3.68 |
| 400 | 4.09 | 7.38 | 2.3 |

Table 3 presents a comparative evaluation of the FAR for conventional SSL [1], FDL [2] and the proposed HSTGAE-FCL model. As the number of devices increases, the FAR decreases for all three approaches. The conventional FDL approach consistently exhibits the highest FAR across all network sizes. This behavior can be attributed to data heterogeneity across federated clients and the absence of fine-grained spatio-temporal modeling, which leads to misclassification of legitimate traffic patterns. The conventional SSL method performs better than FDL by learning representations directly from encrypted traffic features, but it still lacks coordinated inter-device context, resulting in moderate false positives. In contrast, the proposed HSTGAE-FCL framework achieves the lowest FAR at every network scale. The significant reduction in false alarms is primarily due to its hierarchical spatio-temporal graph representation which captures both local and global normal behavior more accurately. Additionally, federated contrastive learning enhances the separation between normal and anomalous embeddings, reducing ambiguity at decision boundaries. Hence, HSTGAE-FCL reduces the FAR significantly in large-scale VPN-assisted wireless IoT networks.

## CONCLUSION

This paper implemented a novel HSTGAE-FCL for reliable attack identification in VPN-assisted wireless IoT networks. On the contrary to conventional approaches, HSTGAE-FCL rely on payload inspection or centralized data gathering. The implemented HSTGAE-FCL performs behavior-based anomaly discovery in encrypted VPN tunnels and making it well suitable for privacy-preserving IoT environment. By considering IoT devices and their relations using a multi-level hierarchical graph structure, HSTGAE-FCL efficiently predict local, cluster-level, and global network dynamics. The incorporation of spatio-temporal graph autoencoding allows the framework to discover normal communication patterns over time and recognize subtle deviations caused by compromised devices, flooding attacks, or unauthorized access attempts. Moreover, the using of FCL allows distributed IoT nodes to find out robust representations without sharing raw data, ensuring privacy protection while managing data heterogeneity across clients. The simulation evaluation displays that HSTGAE-FCL consistently outperforms state-of-the-art [1] and [2]. The HSTGAE-FCL model attains higher attack recognition accuracy, drastically lower false alarm rates, better packet delivery ratio, and reduced communication overhead as network scale enlarges

## REFERENCE

[1] Sattar, S., Khan, S., Khan, M.I. et al. Anomaly detection

in encrypted network traffic using self-supervised learning. Sci Rep 15, 26585 (2025). https://doi.org/10.1038/s41598-025-08568-0

[2] Xiaofeng Wang, Yonghong Wang, Zahra Javaheri, Laila Almutairi, Navid Moghadamnejad, Osama S. Younes, "Federated deep learning for anomaly detection in the internet of things", Computers and Electrical Engineering, Volume 108, 2023, 108651,ISSN 0045-7906.

[3] Y. Liu, Z. Li, S. Pan, C. Gong, C. Zhou, and G. Karypis, "Anomaly detection on attributed networks via contrastive self-supervised learning (CoLA)," IEEE Transactions on Neural Networks and Learning Systems, 2021

[4] T. Zhang, C. He, T. Ma, L. Gao, M. Ma, and S. Avestimehr, "Federated learning for Internet of Things: A federated learning framework for on-device anomaly data detection (FedIoT)," arXiv: 2106.07976, 2021.

[5] W. W. Lo, "A graph neural network-based intrusion detection system - E-GraphSAGE," in Proc. IEEE/ACM NOMS, 2022.

[6] G.-T. An, J.-M. Park, and K.-S. Lee, "Contrastive learning-based anomaly detection for actual corporate environments," Sensors, 2023.

[7] M. GAO et Al., "Anomaly traffic detection in IoT security using graph neural networks," Computer Networks, 2023.

[8] X. Sáez-de-Cámara, J. J. Astrain, J. Villadangos, and J. A. Sanchez, "Clustered federated learning architecture for network anomaly detection in large-scale heterogeneous IoT networks," arXiv: 2303.15986, 2023.

[9] H. Chen et al., "Spatial–temporal graph attention network for video anomaly detection," Pattern Recognition Letters, 2023.

[10] H. Zhu et al., "Spatio-temporal enhanced graph-transformer auto-encoder (STEGT-AE)," IET Computer Vision, 2024.

[11] H. Abduljalil, "Spatio-temporal graph auto-encoder network for skeleton-based anomaly detection," Computers (MDPI), 2024.

[12] A. Belhadi, M. Merniz, A. D. B. Ruíz, and M. A. Ferrag, "Federated contrastive learning and visual transformers," Cognitive Computation, 2024.

[13] S. Seo, J. Lee, and S. Yoon, "Relaxed contrastive learning for federated learning," OpenReview, 2024.

[14] J. Duan, Y. Chen, and J. Yang, "Graph anomaly detection via multi-scale contrastive learning," in Proc. AAAI, 2023.

[15] Y. Han, X. Xu, and W. Hu, "Graph anomaly detection with adaptive node mixup (gADAM)," ACM, 2024

[16] S. Chatterjee and M. K. Hanawal, "Federated learning for intrusion detection in IoT security: A hybrid ensemble approach," arXiv:2106.15349, 2021.

[17] A. Belenguer, J. Navaridas, and J. A. Pascual, "A review of federated learning in intrusion detection systems for IoT," arXiv: 2204.12443, 2022.

[18] Y. Zhang, "Privacy-aware anomaly detection in IoT environments," Journal of Network and Systems Management, 2024

[19] Q. Wang et al., "Federated contrastive learning for cross-domain tasks (Fed-CLR)," IEEE Computer Society, 2025

[20] S.Sangeetha, L.Sudha, "Extreme Gradient Boosted Multifactor Ensemble Relevance Vector Node Classification for Attack Detection in Wireless IoT Network, Neeta Moolani, 2024 Advanced Engineering Science, Volume 56, Issue 01, March, 2024

[21] G. Song, H. Li, and Y. Liu, "STGCN-PAD: A spatial–temporal graph convolutional network for pedestrian anomaly detection," Springer, 2025.

.