EYIB
EDUCATIONAL YIELD
INSIGHTS & BREAKTHROUGHS

TA'LIMGA OID
TUSHUNCHALAR
VA YUTUQLAR

ISSN:3060-4648
Vol.2 №3

https://eyib.uz

Network attack detection using graph neural networks

Axmadaliyev Akramjon Rashidovich

Abstract: This paper presents a graph neural network (GNN)-based approach for network attack detection, emphasizing the representation of hosts and flows as heterogeneous graphs. By leveraging topological and relational dependencies, the proposed models—GraphSAGE, GAT, and temporal GNN—demonstrate superior adaptability and accuracy compared to traditional intrusion detection systems. Evaluations on CIC-IDS2017, UNSW-NB15, and real NetFlow data confirm that GNNs effectively capture multi-stage and evolving attack behaviors while maintaining robustness under dynamic network conditions.

Keywords: Graph Neural Network (GNN), Intrusion Detection, Network Security, Temporal Modeling, Explainable AI (XAI), Heterogeneous Graphs

Proposed Approach: GNN-Based Network Attack Detection

Recent advancements in network security have prompted the adoption of a graph-based perspective for analyzing network traffic in attack detection tasks. Instead of relying solely on raw feature vectors or flat data representations, this novel approach encodes hosts, flows, and connections as nodes and edges in a graph, thereby enabling the capture of complex structural characteristics inherent in modern network environments[2]. Graph Neural Networks (GNNs) process this graph-structured data, allowing models to learn intricate patterns and dependencies that traditional approaches may overlook. By harnessing the topological and relational information extracted from traffic graphs, the detection of multi-stage and stealthy attacks becomes more feasible. Previous studies have shown that such graph-based intrusion detection systems, like Anomal-E, achieve superior accuracy and resilience to evolving attack strategies compared to conventional network intrusion detection techniques[2].

Furthermore, constructing a host-flow heterogeneous graph forms a central component of the proposed network attack detection strategy by representing hosts and network flows as distinct node types. In this graph, hosts are modeled as nodes capturing device-specific attributes, while flow nodes encapsulate communication events, with edges indicating relationships such as source-destination mapping or protocol exchanges. This heterogeneous graph structure allows the model to reflect both direct and indirect associations between actors in network traffic, effectively translating real-world complexity into a trainable computational framework. As a result, the representation enables the learning of structural behaviors that are commonly exhibited by coordinated or multi-stage attacks, which may otherwise evade detection using purely feature-based methods[1]. By leveraging this graph-based modeling, GNNs can extract relational semantics from traffic data, improving recognition of attack topologies and variations across changing environments.

GNN Architectures Explored

Among the range of graph neural network models applied to network attack detection, three architectures were systematically investigated in this study: GraphSAGE, Graph Attention Networks (GAT), and a temporal GNN variant. GraphSAGE operates by aggregating information from a node’s local neighborhood, enabling inductive learning across previously unseen subgraphs and facilitating adaptability to dynamic network environments. In contrast, GAT incorporates an attention mechanism that assigns variable weights to neighboring nodes, thus prioritizing more influential traffic interactions when making node-level inferences—an approach shown to improve both sensitivity and computational efficiency in anomaly detection scenarios[4]. The temporal GNN model extends conventional frameworks by capturing temporal dependencies, allowing the system to account for evolving network behaviors and sequential attack stages as part of the detection process. Collectively, these architectures were selected for their capacities to learn complex graph structures, adapt to dynamic contexts, and address the unique demands inherent in cybersecurity applications.

Additionally, the experimental setup was designed to rigorously assess GNN performance across diverse and realistic network environments. The benchmarking procedure involved three well-established datasets: CIC-IDS2017, UNSW-NB15, and a collection of authentic NetFlow records captured from operational enterprise networks. These datasets encompass a wide array of benign and malicious network activities, ensuring that the evaluation captures the challenges faced in practical deployment scenarios. Care was taken to format each dataset as a host-flow heterogeneous graph, preserving relational and attribute information crucial for meaningful graph-based learning[2]. By using both standardized research benchmarks and real-world traffic samples, the experiments were able to provide a nuanced appraisal of the GNN models’ generalizability and their capacity to identify sophisticated attack behaviors in various operational contexts.

Performance Comparison with Traditional Models

Crucially, the evaluation revealed that GNN-based models deliver superior detection metrics compared to traditional tabu learning approaches, particularly in scenarios involving complex attack structures. Traditional models often employ flat statistical features and lack the capability to incorporate topological dependencies within network data, resulting in diminished efficacy when confronted with advanced multi-stage attack strategies. In contrast, GNN architectures leverage the connectivity and interaction information encoded within host-flow graphs, which translates to enhanced precision, recall, and adaptability across diverse datasets. Empirical results from recent hybrid GCN-GAT studies further underline these strengths by documenting considerably higher recall and F1 scores compared to conventional algorithms, highlighting the acute advantage in both accuracy and detection sensitivity[4]. These findings confirm that the graph-based methodology enables the identification of coordinated attack patterns and complex behaviors that are poorly captured by traditional feature-driven or tabu-based frameworks.

Table 1. Comparison of Traditional and GNN-Based Models

Model	Approach	Key Advantage	Main Limitation	Performance
Traditional ML (RF, SVM)	Feature-based	Simple, interpretable	Fails on complex/multi-stage attacks	Moderate
DNN	Deep feature learning	Learns nonlinear patterns	Ignores topology	High but unstable
GraphSAGE (GNN)	Inductive graph learning	Captures structural context	Sensitive to sparse graphs	High
GAT (GNN)	Attention-based	Focuses on key relations	Higher computational cost	Very high
Temporal GNN	Time-aware	Detects evolving attacks	Requires temporal data	Excellent

Moreover, the drift resistance of the proposed GNN models emerged as a defining factor in their sustained effectiveness within shifting network environments. Adversarial adaptation and the continuous evolution of attack methodologies present persistent obstacles for static or feature-driven detection systems, as they often fail to generalize beyond their training distributions. GNNs, in contrast, inherently model the relational and topological shifts manifested in novel attack traffic, allowing for the dynamic assimilation of unfamiliar patterns without the necessity for frequent retraining. This intrinsic robustness is attributed to GNNs' capacity to generalize from the semantic structure of graphs rather than relying solely on superficial traffic features, enabling them to retain detection accuracy even as malicious behaviors and network usage profiles change[1]. As contemporary evaluations demonstrate, these models are less susceptible to performance degradation under evolving attack tactics, positioning them as a reliable solution for adaptive network security.

Explainable AI (XAI) for Model Interpretation

Consequently, the integration of Explainable AI (XAI) techniques into GNN-based network attack detection addresses the critical need for transparency in automated security decision-making. XAI tools are employed to interpret the complex, often opaque reasoning underlying GNN predictions by providing intelligible explanations of node-level and graph-level outcomes relevant to security analysts. In network security applications, post hoc and self-interpretable XAI approaches can help clarify which host-flow relationships or structural graph features influenced an alert for malicious activity, thereby fostering confidence in the deployment of these advanced models[3]. Such interpretability is not only essential for model validation and compliance in regulated environments, but also for practical incident response, where analysts must understand the rationale behind detection results in real time. By demystifying the decision process of GNNs, XAI methodologies contribute to stronger trust, enabling practitioners to leverage sophisticated detection models while maintaining accountability in critical operational contexts.

Conclusion

The experimental analysis confirms that GNN-based intrusion detection significantly enhances network security by modeling complex relationships within traffic data. Unlike traditional feature-driven techniques, GNN architectures exploit structural and temporal dependencies, yielding improved precision and resilience against evolving attack strategies. The inclusion of Explainable AI further bridges the interpretability gap, empowering analysts to understand model reasoning. Overall, the proposed framework establishes a robust, adaptive, and transparent foundation for next-generation intelligent network defense systems.

References:

[1]. Bilot, T. *et al.* (2023) “Graph neural networks for intrusion detection: A survey,” *IEEE Access*, 11, pp. 49114–49139. doi:10.1109/ACCESS.2023.3275789.

[2]. Caville, E. *et al.* (2022) “Anomal-E: A self-supervised network intrusion detection system based on graph neural networks,” *Knowledge-Based Systems*, 258, p. 110030. doi:10.1016/j.knosys.2022.110030.

[3]. Nandan, M., Mitra, S. and De, D. (2025) “GraphXAI: a survey of graph neural networks (GNNs) for explainable AI (XAI),” *Neural Computing and Applications*, 37, pp. 10949–11000. doi:10.1007/s00521-025-11054-3.

[4]. Yılmaz, A. and Das, R. (2025) “A novel hybrid approach combining GCN and GAT for effective anomaly detection from firewall logs in campus networks,” *Computer Networks*, 259, p. 111082. doi:10.1016/j.comnet.2025.111082.