Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.078546

REVIEW

Graph and Transformer-Based Deep Learning Paradigms for DDoS Detection: A Systematic and Critical Survey

Noor Mueen Mohammed Ali Hayder[1,2], Seyed Amin Hosseini Seno[2,*], Mehdi Ebady Manaa[3,4], Hamid Noori[2] and Davood Zabihzadeh[5]

[1]Faculty of Nursing, Babylon University, Hilla, Iraq
[2]Department of Computer Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
[3]Intelligent Medical Systems Department, College of Sciences, Al-Mustaqbal University, Babylon, Iraq
[4]College of Information Technology, University of Babylon, Babylon, Iraq
[5]Computer Engineering Department, Hakim Sabzevari University (HSU), Sabzevar, Iran

*Corresponding Author: Seyed Amin Hosseini Seno. Email: hosseini@um.ac.ir

Received: 03 January 2026; Accepted: 04 March 2026; Published: 08 May 2026

ABSTRACT: With the rapid expansion of networked systems, Distributed Denial-of-Service (DDoS) attacks have become a major threat to Internet security and service availability. Due to their limited scalability, incapacity to capture temporal and relational relationships, and decreased detection accuracy under dynamic and high-volume network traffic, traditional machine learning algorithms frequently fail in large-scale DDoS scenarios. This encourages the application of deep learning techniques that can simulate intricate relationships. This survey systematically reviews graph-based deep learning and Transformer models for DDoS detection. We categorize methods for transforming network traffic into graph representations and analyze key architectures, including GraphSAGE, GCN, GAT, spatio-temporal Transformers, and hybrid GNN–Transformer models. We summarize the evaluation metrics, datasets, feature extraction strategies, and performance trends reported across existing studies. Results indicate that these approaches effectively capture topological and temporal patterns to detect coordinated attacks. Our comparative review shows that these approaches are capable of capturing both topological and temporal patterns in network traffic, enabling more accurate identification of coordinated DDoS attacks reported in the literature. Remaining challenges include explainability, scalability, data imbalance, and limited generalization. The survey's contributions are a unified taxonomy, comparative analysis, identification of open challenges, and future research directions toward explainable, lightweight, and federated frameworks.

KEYWORDS: Distributed Denial-of-Service (DDoS) detection; Graph Neural Networks (GNNs); transformer architecture; deep learning; network security; graph-based learning

1 Introduction

In recent years, DDoS attacks have become one of the leading threats to the reliability, availability, and overall security of networked systems [1]. With the rapid expansion of online services, IoT ecosystems, and cloud infrastructures, the scale and sophistication of these attacks have increased significantly, often resulting in service outages, reputational damage, and substantial financial losses for organizations worldwide [2]. Conventional detection approaches—including traditional ML and signature-based methods—have long been used to mitigate DDoS threats [3]. However, these techniques often fail to capture the complex spatial

and temporal patterns present in modern network traffic, which leads to lower detection accuracy, slower reaction times, and limited generalization [3].

Recently, graph-based deep learning methods—particularly Graph Neural Networks (GNNs)—have shown strong potential for modeling the relational and topological characteristics of network traffic [4]. By representing communication flows and entity interactions as graph structures, GNNs can effectively capture dependencies among nodes such as sessions, hosts, and IPs, enabling the detection of coordinated attack patterns that many traditional approaches fail to identify [5–7]. In addition, Transformer architectures, known for their attention mechanisms and ability to model long-range dependencies, have increasingly been integrated with graph-based models to enhance feature representation and improve detection performance in large-scale and complex network environments [8].

Despite these advancements, there is still no comprehensive survey that systematically examines graph-based and Transformer-driven deep learning methods for DDoS detection.

This survey particularly covers graph-based and Transformer-enhanced approaches, including hybrid GNN–Transformer architectures, in contrast to previous DL-based IDS surveys that mostly concentrate on conventional neural networks or general deep learning techniques. Additionally, it offers a distinct viewpoint not discussed in other reviews by providing a thorough comparative analysis, practical implementation considerations, and identification of open research issues.

Existing reviews mostly focus on traditional ML techniques or broad DL approaches [9–14], leaving a clear gap regarding the combined strengths, limitations, and design considerations of graph-oriented detection frameworks.

Motivated by this gap, the present survey aims to offer a structured and in-depth review of GNN-based and Transformer-enhanced frameworks for DDoS detection. The primary objectives of this study are:

1. To categorize and analyze existing graph-based deep learning models for DDoS detection, including hybrid GNN–Transformer approaches.
2. To compare their advantages, disadvantages, and practical deployment considerations.
3. To introduce a clear and coherent taxonomy that helps researchers understand the current design space, emerging trends, and methodological directions.
4. To identify open challenges and outline future research opportunities, such as lightweight, explainable, federated, and scalable graph-based detection frameworks.

By providing a structured and comprehensive overview, this survey seeks to serve as a useful reference for researchers, security engineers, and practitioners working on developing advanced graph-based and Transformer-driven DDoS detection systems.

The remainder of this paper is organized as follows. Section 2 reviews the background of GNNs, Transformer models, and relevant concepts. Section 3 introduces the methodological taxonomy used to classify existing approaches. Section 4 provides a comparative analysis and discussion across three categories: Transformer-based models, hybrid GNN–Transformer models, and pure GNN models. Section 5 highlights major challenges and unresolved issues, while Section 6 outlines promising future research directions. Section 7 presents the evaluation metrics used for DDoS attack detection, and Section 8 provides an overview of the datasets employed in the study. Finally, Section 9 concludes the paper.

2 Literature Review

DDoS attacks pose a considerable threat to the new networked systems' accessibility, security as well as reliability. With the proliferation of high-speed networks, cloud computing, and Internet of Things (IoT) devices, attack frequency and sophistication have broadly increased in the last few years [15]. These

attacks aim at overwhelming the targeted servers/systems and networks by flooding them with extensive malicious traffic volumes, thereby rendering legal services inaccessible. New DDoS attacks are sometimes multi-vector and coordinated, integrating protocol-level exploitation, application-layer disruptions, and volumetric flooding. This complexity makes diagnosis particularly challenging for conventional observing systems [16].

Traditional ML techniques have been broadly developed for DDoS threats' mitigation, applying mechanisms like SVM, Artificial Neural Networks (ANNs), Random Forests (RFs), and K-Nearest Neighbors [17]. These models normally depend on manually engineered features taken from traffic volumes/packet statistics and flow-level metrics. While these approaches have demonstrated effectiveness in actual observed scenarios, they sometimes fail to capture the complex spatial and temporal dependencies inherent in network traffic [18]. In addition, models trained on particular sets of data sometimes show weak generalization when used to novel areas of network/evolving attack models, and real-life diagnosis is frequently hindered by the need for aggregated traffic data [19].

In the last few years, graph-driven deep learning strategies have emerged as a promising solution to consider such restrictions. By showing the network entities—like sessions, hosts, IP addresses—as nodes, and their interactions—like traffic flows/network connections—as edges, GNNs can efficiently capture relational and topological structures in network traffic [20]. Variants like GraphSAGE, Graph Convolutional Networks (GCNs), and Graph Attention Networks (GATs) allow for adaptive neighbor weighting, scalable embedding generation, and localized feature aggregation. These models make the coordinated/stealthy attacks detectable, which may remain undetected by conventional ML methods [21].

Therefore, Transformer architectures, basically deployed for natural language processing, have illustrated exceptional ability in modeling long-range dependencies and sequential models. Transformers leverage self-attention algorithms and multi-head attention to learn complicated relations among inputs, making them appropriate to model temporal dependencies in network traffic [22]. While integrated with GNNs, Transformers can improve node feature representations and diagnosis performance, especially in large-scale and active network areas [23]. Spatio-temporal Transformers enhance this capacity by jointly modeling both structural and temporal relations in network traffic [24].

In spite of such advances, a crucial gap exists in the literature: no general study presents that systematically reviews graph- and Transformer-based strategies for DDoS diagnosis. Many of the previous studies concentrate on traditional ML methods/comprehensive DL techniques, with no investigation of unified benefits, restrictions, and considerations for modeling graph-based approaches. The present study targets to fill this gap by presenting the structured state-of-the-art techniques' analysis, highlighting issues, comparing their performance, and categorizing future study directions in this quickly evolving domain.

3 Methodology

To ensure a comprehensive and systematic review, we collected relevant studies from major databases including IEEE Xplore, Scopus, Springer, and arXiv using keywords such as "DDoS detection", "graph neural networks", "transformer", and "hybrid GNN–Transformer". We focused on peer-reviewed articles published between 2021 and 2025. Duplicate and irrelevant papers were excluded based on title, abstract, and full-text screening, resulting in a total of 106 selected studies covering theoretical and applied aspects of graph-based and Transformer-enhanced DDoS detection.

For clarity, we categorize these works based on the dominant architectural characteristic of each model. Pure GNN-based models rely solely on graph neural networks for topological feature extraction,

Transformer-driven models primarily use attention mechanisms to capture temporal or long-range dependencies, and hybrid GNN–Transformer models integrate both to leverage the strengths of each. In cases of complex or multi-component models, categorization is determined by the primary component contributing to detection performance. This taxonomy, along with consideration of feature representation and evaluation strategy, provides a structured framework for understanding current trends, strengths, and challenges in graph-based DDoS detection research. Although the individual categories are not novel, the proposed taxonomy integrates model type, feature representation, and evaluation metrics in a unified framework tailored for DDoS detection, providing practical guidance for both researchers and practitioners. Fig. 1 illustrates our proposed taxonomy, categorizing reviewed DDoS detection approaches.

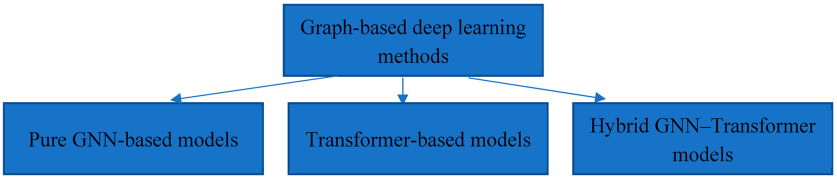

Figure 1: Graph-based deep learning methods for DDoS attack detection.

Although early deep learning and traditional machine learning techniques have been used for DDoS detection, this survey does not concentrate on them. In this article, we examine GNN-based, Transformer-based, and hybrid approaches that provide improved capacity to identify relational, topological, and temporal patterns in large-scale networks.

Pure GNN-based models leverage the intrinsic relational structure of network traffic to capture coordinated and distributed attack patterns. By representing network nodes (such as sessions, IP addresses, hosts) and their interactions as edges, models like GraphSAGE, Graph Convolutional Networks (GCN), and Graph Attention Networks (GAT) enable efficient feature aggregation across neighborhoods. GCNs perform localized convolution operations to propagate information, while GATs apply attention mechanisms to weigh neighboring nodes based on their importance, improving detection accuracy in complex traffic scenarios. GraphSAGE provides scalable embedding methods suitable for large-scale networks, making it practical for real-world applications. Some studies have shown that pure GNN models outperform conventional ML methods in detecting low-volume, stealthy, or coordinated DDoS attacks, primarily due to their ability to capture topological dependencies [25].

Transformer-based models excel at modeling temporal dependencies and sequential patterns in network traffic. Using self-attention mechanisms, these architectures learn relationships among distant traffic events, making them particularly suitable for time-series and flow-level analysis. While Transformers can efficiently model long-range dependencies, they may not fully exploit the structural information inherent in network graphs. Nevertheless, some approaches have successfully applied Transformers to packet orders or aggregated flow features, achieving improved performance compared to traditional recurrent or convolutional models [26].

Hybrid GNN–Transformer models combine the strengths of both paradigms, integrating topological reasoning with temporal and global context modeling. In such approaches, GNNs typically manage the relational network structure, extracting embeddings that represent node interactions, while Transformers capture higher-order dependencies and temporal patterns. This fusion enables more comprehensive feature representations and has been shown to improve detection performance on large-scale, complex network datasets. For instance, studies have applied GCN–Transformer frameworks for DDoS detection using widely

available datasets such as NSL-KDD and CICDDoS2019, achieving higher accuracy and robustness compared to using either technique alone [27–30].

Overall, this taxonomy not only groups the existing techniques but also highlights trade-offs and considerations for each architecture. Pure GNNs offer strong structural modeling but may struggle with temporal dynamics; Transformers excel at sequential modeling but may overlook topological dependencies; hybrid models aim to bridge these gaps. Understanding these distinctions is important for designing effective DDoS detection systems and guiding future research in graph-driven and Transformer-based network security solutions. Table 1 provides a comparative overview of hybrid, Transformer-based, and pure GNN-based models, highlighting their main characteristics, disadvantages, advantages, and typical use cases in DDoS detection.

Table 1: Comparison of DDoS detection models.

Category	Key Characteristics	Advantages	Disadvantages	Typical Use Cases/Remarks
Pure GNN-based Models	Utilize relational structure of network traffic; nodes represent IPs, hosts, or sessions; edges represent interactions. Examples: GCN, GAT, GraphSAGE	-Effectively capture topological dependencies -Detect stealthy, low-volume, or coordinated attacks -Scalable embedding generation (GraphSAGE)	-Limited temporal modeling -May not capture long-range sequential dependencies	Real-time detection in medium to large-scale networks; effective for multi-vector DDoS attacks
Transformer-based Models	Focus on temporal dependencies and sequential patterns using self-attention, applied to packet sequences or aggregated flow features.	-Captures long-range dependencies in time-series data -Improved performance over traditional RNN/CNN for sequential traffic	-Ignores or underutilizes graph/structural information -Computationally expensive for large-scale traffic	Time-series or flow-level analysis; ideal for sequential or streaming traffic patterns
Hybrid GNN–Transformer Models	Combine GNN (for graph structure) and Transformer (for temporal/global patterns); extract node embeddings and capture higher-order dependencies.	-Integrates topological and temporal information -Improved detection performance -Robust on complex, large-scale datasets	-Complex architecture -Higher computational and training costs	Large-scale and complex networks; scenarios requiring both spatial (graph) and temporal modeling; state-of-the-art detection

4 Comparative Analysis and Discussion

4.1 Pure GNN-Based Models

GNN-driven strategies for intrusion detection, with a focus on DDoS, have emerged as a prominent research direction because they naturally model the relational structure of network traffic. Instead of treating streams or packets as independent vectors, GNN techniques represent streams, sessions, hosts, and IPs as nodes and their interactions as edges, enabling the model to learn topological correlations, coordinated behaviors, and propagation patterns. Studies in this subgroup explore graph constructions (host/flow/forwarding graphs), GNN variants (GraphSAGE, GCN, custom relational GNNs, GAT, ST-GCN), and application domains (botnet detection, SDN, IoT, smart grid). Common objectives include improving

detection accuracy for stealthy or multi-vector DDoS attacks, locating attack paths or compromised entities, and enhancing GNN robustness in lossy or noisy network environments. Precise summaries of the 26 selected papers are provided below, followed by a comparative Table 2.

Table 2: Comparative analysis of pure GNN-based models for cybersecurity and network analysis.

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
1	Guo et al. (2022) [31]	DDoS detection (binary & 3-class; source distribution inference)	GLD-Net: flow + topology fusion → GAT → LSTM → FC	First explicit fusion of topology & flow; high accuracy; infers attack source distribution.	Requires careful graph/time modelling; generalization not fully validated dataset split/cross-validation details not clearly reported, high reported accuracy may not generalize
2	Wang et al. (2025) [32]	Intrusion detection in Edge/IoT; multi-class problems	BS-GAT: behavior-similarity graph construction + weighted GAT	Dataset-aware graph construction; strong multi-class metrics	Graph construction complexity; dataset dependence, evaluation setup/train-test split not fully detailed
3	Barsellotti et al. (2023) [33]	DDoS/traffic & flow-level detection	Two-level hierarchical graph (traffic + flow) processed by hierarchical GNN	Removes the need for stateful features; exploits multi-level structure	Complexity in mapping traffic ↔ flow; engineering overhead dataset split/validation details limited
4	Mohan & Kumar (2025) [34]	Botnet detection/mitigation	BotMHG: flow→graph, alternating GNN & GAT layers; preprocessing for imbalance	Focus on bot behaviour; robust statistical validation (multiple tests)	Preprocessing and imbalance handling required; hybrid complexity
5	El Gadal and Ganti (2025) [35]	Intrusion detection & mitigation in SD-IoT (privacy & reporting)	Vision Transformer + GraphSAGE in Federated Learning; MA-DQL mitigation; Flan-T5 reporting	End-to-end (detection + mitigation + reporting); privacy via FL	Very complex multi-component system; integration overhead
6	Le & Park (2024) [36]	Multi-class NIDS (importance of edge features)	Edge-aware GNN encoding ports & flow attributes	Explicit modelling of edge features; strong multi-class performance	Reliance on port/flow features (NAT/encapsulation issues)
7	Saxena et al. (2025) [37]	NIDS for IoT; real-time monitoring + attack graph	GNN-IDS: attack graph + GNN core for dynamic/static features	Practical attack-graph integration; claimed explainability	Dataset scope limited; explainability needs more evidence
8	Ran et al. (2024) [38]	SDN saturation (flow table overflow) detection	TITAN: bi-directional forwarding graph → GCN	SDN-specific forwarding path graph; accurate detection	SDN-centric; needs routing/forwarding path info

(Continued)

Table 2 (continued)

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
9	Hekmati & Krishnamachari (2024) [39]	DDoS in IoT; resilience to lossy links	GCN with multiple graph topologies (correlation-based hybrid)	Robust to connection loss; topology study	IoT focus; dataset details limited
10	Saidane et al. (2025) [40]	DDoS flow detection with heterogeneous elements	CHC-DDoS: Host-Connection Graph + heterogeneous GCN (per edge type)	Handles heterogeneity and multiple edge/node types	Model complexity: heavy per-edge processing
11	Hekmati and Krishnamachari (2024) [41]	DDoS in IoT: comparative graph constructions	GCN with distance-based and correlation-based graphs	Comparative analysis of graph construction: resilient performance	Performance varies by topology; some scenarios are weaker
12	Nagaraj et al. (2021) [42]	DDoS detection + compromised entity identification in Smart Grid (SDN-SGC)	GLASS: supervised GCN for detection + spectral clustering for identification	End-to-end for critical infra; detection + ID pipeline	Smart-grid specific; generalizability to other domains unclear
13	Wang & Wang (2025) [43]	Low-rate DDoS (LDDoS) detection & localization in SDN	Hybrid GCN + GRU with double sliding window; time-freq & QoS features	Online detection + mitigation via OpenFlow; localization of victim switch/port	SDN controller dependency; portability concerns
14	Saunders et al. (2024) [44]	DDoS detection (telecom focus)	3-layer GCN (128 neurons per layer)	Very high reported performance	Extremely high metrics—risk of overfitting or data-split issues
15	Manjula et al. (2023) [45]	Blackhole routing attacks in RPL IoT	GCN on RPL routing topology	Tailored to RPL; effective in LLN contexts	Narrow domain; IoT resource constraints
16	Li et al. (2023) [46]	Multi-vector DDoS multi-classification	GoGDDoS: Graph-of-Graph (packets↔flows) + two-level GNN	Novel GoG representation capturing packet↔flow relations	Complex graph merging; higher computation
17	Xu et al. (2024) [47]	Malicious host detection (imbalance & edge features)	RE-GCN: relational-edge GCN that directly aggregates edge/netflow features	Directly models edge features; addresses imbalance	Uses undersampling (may discard negatives); netflow specificity
18	Cao et al. (2021) [48]	DDoS on SDN data-plane; path tracing & mitigation	ST-GCN on INT-sampled SDN state (spatio-temporal)	Path tracing + targeted mitigation; low CPU/southbound load	Requires INT instrumentation and SDN programmability

(Continued)

Table 2 (continued)

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
19	Li et al. (2023) [49]	DDoS path traceability & algorithm recommendation	AT-GCN: attack traceability KB + intra-domain attack graph + Tracing-Sample	Focus on traceability & recommending trace algorithms	Memory tradeoffs; intra-domain focus limits scope
20	Saunders et al. (2024) [50]	DDoS detection for critical infrastructure	GCN-empowered IDS	Telecom/critical-infra focus; strong reported performance	Replicability & dataset/split details need clarification
21	Venturi et al. (2024) [52]	DDoS Detection	Ensemble GNNs with bagging & boosting across different flow granularities	Robustness, reduced overfitting, strong multi-level flow modeling	Computationally expensive; limited to CICIDS2017/2018 datasets, train/test split not clearly reported
22	Abu Bakar et al. (2024) [53]	DDoS Detection in IoT using FL	GraphFedAI: session-based graphs + Pearson selection + federated GNN training	Privacy-preserving, scalable, dynamic IoT support	Federated learning overhead; lacks real-world deployment validation dataset split or cross-validation not detailed
23	Anjum et al. (2025) [54]	DDoS Detection under partial flow visibility	Graph Isomorphism Network (GIN) on progressively reduced graphs	Studies real-world case of incomplete traffic; useful robustness insights	Simulated flows may not reflect real packet behavior; limited generalizability
24	Holmkvist Bergqvist (2025) [55]	Poisoning attacks in FL-based IoT IDS	AGAT-FL: trust-aware GAT aggregation + CNN-GRU + Mahalanobis filtering	Strong poisoning resistance, use of explainable AI (SHAP, LIME)	High complexity; trust scoring difficult in highly dynamic IoT environments

Guo et al. [31] proposed GLD-Net that fuses topological and flow features through making building graph representations from time-series flow data, applying GAT to mine non-Euclidean correlations, and an LSTM after GAT for neighborhood temporal modelling. Reported high accuracy on NSL-KDD2009 and CIC-IDS2017 (0.993 binary, 0.942 three-class). Advantages: firstly, to clearly fuse topology + flow; robust accuracy and attack-source share inference. Disadvantages: needs careful graph building and order modelling; dataset/generalization questions.

Wang et al. [32] presented BS-GAT, a treat-similarity driven graph construction in addition to a weighted GAT which incorporates edge behavioral relation weights for mitigating overfitting and better mine structural info for Edge/IoT intrusion detection. Reported >99% metrics in binary tasks and >93% multi-class accuracy. Advantages: tailored graph construction to realistic sets of data; robust multi-class performance. Disadvantages: complexity of graph construction; the assessment scope might be restricted to the actual Edge sets of data.

Barsellotti et al. [33] deployed a two-class hierarchical graph representation (traffic-level + flow-level) and a GNN capable of processing the two levels, targeting at exploiting structural info with no stateful features. Tests on CIC-IDS2017 demonstrate comparable state-of-the-art performance just by applying

the traffic structure. Advantages: hierarchical graph opinion decreases dependence on stateful features. Disadvantages: might need careful mapping among traffic/flow levels for various networks.

Mohan & Kumar [34] proposed BotMHG, a hybrid graph model for botnet detection made from network flow graphs, applying alternating GNN and GAT layers to take topological and temporal relations; training contains preprocessing to control imbalance. Validated on CTU-13 and BoT-IoT with low FPR and high accuracy; statistical tests validate importance. Advantages: concentrated on botnet manner and robust assessment. Disadvantages: hybrid complexity and preprocessing needs.

El Gadal and Ganti [35] provided the Federated SD-IoT architecture integrating Vision Transformer (ViT) and GraphSAGE for local/global relation processing, in addition to federated learning and a MA-DQL mitigation agent; contains automatic attack reporting (Flan-T5). Obtained ~98.06% detection with adaptive mitigation. Advantages: end-to-end system (reporting + detection + mitigation), privacy through FL. Disadvantages: high system complexity; several elements enhance the combination attempt.

Le & Park [36] emphasized the edge features' significance (flow features) for multi-class NIDS and proposed the GNN variant, which clearly encodes edge features (packet counts, ports, duration); therefore, the graph better resembles actual networks. Reported large gains (like 98.32% on CIC-IDS2017). Advantages: considers the oft-overlooked point (edge attributes). Disadvantages: dependence on port/stream attributes might be brittle over areas with NAT/encapsulation.

Saxena et al. [37] proposed GNN-IDS, combining real-time observing with an attack graph to take the static and active attributes; the GNN core assesses neighborhood impact on strong detection. Confirmed on CIC-IoT-2023 with robust performance and claimed explainability. Advantages: practical framework with attack-graph augmentation; strength. Disadvantages: assessment restricted to actual IoT datasets; explainability details need elaboration.

Ran et al. [38] proposed TITAN, a bidirectional forwarding-graph construction tailored for SDN saturation attacks; builds bi-directional forwarding graphs (nodes = flows) and uses GCN to diagnose saturation attack flows. Obtained >97% accuracy in SDN adjustments. Advantages: SDN-specific graph which takes forwarding ways; great accuracy. Restrictions: SDN-centric; graph creation relies on routing way data.

Hekmati & Krishnamachari [39] used GCN for DDoS detection in IoT, assessing hybrid graph topologies; defined correlation-driven hybrid graphs and illustrated high F1 (~91%) with resilience to connection loss (≤2% drop under 50% loss). Advantages: strength to lossy areas; topology research. Disadvantages: IoT-specific; might not generalize to non-IoT networks.

Saidane et al. [40] provided CHC-DDoS, applying Host-Connection Graphs (HCGs) and heterogeneous message-passing GNNs (per-edge-type GCN + node-type updates) for flow/node labeling. Concentration on heterogeneity and flow-level predictions. Advantages: controls the heterogeneous components of the graph and edge kinds. Disadvantages: complexity of the model and likely heavy per-edge processing.

Hekmati et al. [41] proposed resilient GCN strategies for IoT with distance-and correlation-driven graph constructions, analyzing performance under differing connection loss; reported up to 85% F1 in aggressive scenarios and restricted performance degradation. Advantages: comparative graph construction research; resilient performance. Disadvantages: The dataset/experimental scope might be restricted.

Nagaraj et al. [42] defined GLASS for SDN-Smart-Grid communications: a 2-step architecture applying supervised GCN for detection and unsupervised grouping for compromised entity recognition; assessed on IEEE 118-bus testbed illustrating via preservation for compromised entities. Advantages: end-to-end detection + recognition in crucial infrastructure. Disadvantages: applicability above smart grid demands' confirmation.

Wang & Wang [43] considered detection of LDDoS in SDN by applying the hybrid GCN-GRU model with double sliding windows; extracts time–frequency and QoS features to recognize victim switches/ports and perform mitigation through OpenFlow. Shown online detection and mitigation. Advantages: integrates spatial (GCN) and temporal (GRU) modeling; actionable mitigation. Disadvantages: SDN controller development specifics might restrict portability.

Saunders et al. [44] proposed the 3-layer GCN DDoS detector and reported very high metrics on CIC-IDS2017 (≈99.95% across metrics). Focuses GCN's capability for taking topological + statistical information among networks of attack and victim. Advantages: robust empirical outcomes. Disadvantages: Broadly high reported metrics warrant careful cross-validation and scrutiny for overfitting.

Manjula et al. [45] used GCN to diagnose blackhole routing attacks in RPL-driven IoT networks, designing methods for nodes in limited IoT topologies and reporting efficient detection. Advantages: considers routing-specific threats in LLN/RPL. Restrictions: narrow scope (RPL networks) and resource limitations.

Li et al. [46] proposed GoGDDoS, a 2-level Graph-of-Graph (GoG) traffic representation that appears graphs of packet and flow relation, after that uses a 2-level GNN for multi-class DDoS classification; reported better performance than baselines. Advantages: new GoG representation to take packet↔flow relations. Disadvantages: complexity in graph merging and higher computational cost.

Xu et al. [47] provided RE-GCN, a Relational-Edge GCN which aggregates and learns directly from edge (netflow) attributes for malicious-host detection; applies time slicing and undersampling for imbalance. Performed better than usual GNN baselines on the NetFlow sets of data. Advantages: directly designs attributes of the edge; considers imbalance. Disadvantages: undersampling might discard useful negatives; netflow specifics might restrict generality.

Cao et al. [48] proposed the ST-GCN over programmable SDN data plane applying In-band Network Telemetry (INT) sampling; diagnoses DDoS ways and makes the aimed mitigation able (developed whitelist + accurate dropping), developing detection accuracy by ~10% over classic techniques. Advantages: spatial-temporal designing with way tracing and mitigation. Disadvantages: depends on SDN programmability as well as INT instrumentation.

Li et al. [49] enhanced AT-GCN, an attack traceability system integrating the attack traceability knowledge base with GCN; models intra-domain attack graphs and a Tracing-Sample subgraph sampling mechanism for remaking ways of DDoS and recommend trace mechanisms, developing recall and decreasing FPR. Advantages: concentrates on traceability and practical recovery. Disadvantages: memory use tradeoffs and intra-domain concentration.

Saunders et al. [50] proposed the GCN-empowered DDoS IDS for telecommunications infrastructure, demonstrating high detection/classification confidence for hybrid DoS variants. Advantages: aimed at crucial infrastructure; high reported performance. Disadvantage: replication details and dataset splits require careful reporting.

Galli et al. [51] investigate adversarial vulnerabilities in GNN-based intrusion detection systems. They introduce the first formal framework for GNN-targeted adversarial attacks, modeling real-world constraints, and experimentally demonstrate that although GNNs resist feature-based attacks, they remain vulnerable to structural perturbations. The main advantage is its novelty and rigorous analysis of attack feasibility. The drawback, however, is the absence of defensive strategies, meaning the work highlights vulnerabilities without proposing robust countermeasures.

Venturi et al. [52] propose a DDoS detection method using GNN ensemble learning, combining multiple GNN models with bagging and boosting to capture relationships between traffic flows at different granularities. Their system achieves high accuracy and reduced overfitting through regularization

techniques. Strengths include improved F1-score, robustness, and the ability to model multi-level flow relationships. Disadvantages include significant computational overhead from maintaining multiple GNNs and limited evaluation across diverse datasets beyond CICIDS2017/2018.

Abu Bakar et al. [53] introduce GraphFedAI, a federated-learning–based GNN framework for privacy-preserving DDoS detection in IoT networks. The method uses adaptive session-based graph modeling, Pearson correlation feature selection, and interpolation-aware GNN training across distributed devices. The major strengths are strong privacy protection, scalability, and good performance on dynamic IoT traffic. Weaknesses include communication overhead typical of federated learning and a lack of real-world deployment validation.

Anjum et al. [54] study how reducing network flow graph complexity affects GNN performance in detecting DDoS attacks. Using simulated flow graphs from CICIDS2017, they test the robustness of a Graph Isomorphism Network (GIN) under reduced node counts. The strength is its focused examination of GNN behavior under incomplete or partial traffic visibility—an important real-world scenario. The weakness is that flow simulation may not fully represent real packet-level traffic, limiting the applicability of the findings.

Holmkvist Bergqvist [55] propose AGAT-FL, a federated learning framework enhanced with graph attention mechanisms for secure intrusion detection in IoT devices. The method uses trust-aware aggregation, CNN-GRU for featurization, and Mahalanobis filtering to suppress malicious client updates. Advantages include strong resistance to poisoning attacks, high accuracy across datasets, and the use of explainable AI (SHAP, LIME). Disadvantages include higher model complexity and reliance on accurate trust scoring, which may be difficult in highly dynamic IoT environments.

A systematic quantitative evaluation of GNN-based network intrusion and DDoS detection techniques in terms of model complexity and classification performance is shown in Table 3.

Table 3: Structured comparison of GNN-based network intrusion and DDoS detection approaches.

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
1	Guo et al. (2022) [31]	99.3% (binary); 95.8% (3-class)	99.5% (binary); 95.2% (3-class)	98.9% (binary); 95.9% (3-class)	99.2% (binary); 95.5% (3-class)	**Heavy** (GAT + LSTM + topology-flow fusion)
2	Wang et al. (2025) [32]	>99% (binary); >93% (multi-class)	>99% (binary); NR (multi-class)	>99% (binary); >91% (multi-class)	>99% (binary); >92% (multi-class)	**Medium–Heavy** (graph construction + weighted GAT)
3	Barsellotti et al. (2023) [33]	99.14%	NR	NR	99.13%	**Heavy** (hierarchical multi-level GNN)
4	Mohan & Kumar (2025) [34]	99.44%	99.53%	99.36%	99.43%	**Heavy** (hybrid GNN/GAT + preprocessing)
5	El Gadal et al. (2025) [35]	≈98.06%	NR	NR	NR	**Very Heavy** (ViT + GraphSAGE + FL + RL + LLM)
6	Le & Park (2024) [36]	98.32% (CIC); 96.71% (UNSW)	97.2%	96.6%	96.8%	**Medium** (edge-aware GNN)
7	Saxena et al. (2025) [37]	92.81%	96.34%	68.36%	79.93%	Medium–Heavy (attack graph + GNN)

(Continued)

Table 3 (continued)

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
8	Ran et al. (2024) [38]	99.42%	99.77%	98.88%	99.32%	**Medium** (GCN on forwarding graph)
9	Hekmati & Krishnamachari (2024) [39]	NR	NR	NR	up to 91%	**Medium** (multi-topology GCN)
10	Saidane et al. (2025) [40]	99.46%	99.40%	99.44%	99.41%	**Heavy** (heterogeneous GCN per edge type)
11	Hekmati et al. (2024) [41]	NR	NR	NR	up to 85%	**Medium** (topology-aware GCN)
12	Nagaraj et al. (2021) [42]	100%	100%	100%	100%	**Medium–Heavy** (GCN + spectral clustering)
13	Wang & Wang (2025) [43]	97.25%	97.30%	97.27%	97.25%	**Heavy** (GCN + GRU + sliding windows)
14	Saunders et al. (2024) [44]	≈99.95%	≈99.95%	≈99.95%	≈99.95%	**Medium** (3-layer GCN)
15	Manjula et al. (2023) [45]	98%	100%	96.59%	98.27%	**Light–Medium** (GCN on RPL topology)
16	Li et al. (2023) [46]	99.33%	99.76%	99.11%	98.93%	**Heavy** (Graph-of-Graph + two-level GNN)
17	Xu et al. (2024) [47]	98.11%	99.63%	98.07%	98.85%	**Medium** (relational-edge GCN)
18	Cao et al. (2021) [48]	NR	NR	NR	NR	**Heavy** (spatio-temporal GCN + INT)
19	Li et al. (2023) [49]	NR	NR	95%	NR	**Medium–Heavy** (attack KB + GCN)
20	Saunders et al. (2024) [50]	99.20%	98.39%	98.57%	98.47%	**Medium** (GCN-based IDS)
21	Galli et al. (2025) [51]	NR	99.3%	98%	98.6%	**N/A** (adversarial analysis study)
22	Venturi et al. (2024) [52]	~99.67%	NR	NR	~99.29%	**Heavy** (ensemble GNNs)
23	Abu Bakar et al. (2024) [53]	99.67%	NR	NR	99.29%	**Heavy** (federated GNN)

(Continued)

Table 3 (continued)

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
24	Anjum et al. (2025) [54]	97.2%	NR	NR	96.9%	**Medium** (GIN under partial visibility)
25	Holmkvist Bergqvist (2025) [55]	~94%	~94%	~94%	~94%	**Very Heavy** (FL + GAT + CNN-GRU + XAI)

The amount of trainable parameters, approximate FLOPs, and memory footprint during inference are used to qualitatively classify model complexity as light, medium, or heavy. Medium models have 1–10M parameters, heavy models have >10M parameters, and light models have <1M parameters.

Overall, the majority of experiments show very high accuracy values (over 97%), especially in binary or carefully chosen multi-class scenarios, demonstrating graph-based learning's great representational capacity for network traffic modeling. The significance of capturing both spatial and temporal dependencies in intrusion detection tasks is confirmed by hybrid architectures that combine GNNs with temporal models like LSTM, GRU, or Transformers (e.g., Guo et al. [31], Mohan & Kumar [34], Wang & Wang [43]). These architectures consistently achieve higher and more balanced Precision–Recall–F1 scores.

The presence of NR (Not Reported) items in Table 3 indicates that, despite the encouraging results, a considerable proportion of studies do not report all conventional evaluation measures. A single metric or simply accuracy is given in a number of instances (e.g., [33,39,41,48]), which restricts fair comparison and hides possible trade-offs between false positives and false negatives. This lack of consistent reporting highlights a significant methodological flaw in the existing literature and underscores the need for standardized evaluation procedures in GNN-based IDS research.

The results show a definite trade-off between computational expense and performance from the standpoint of model complexity. Heavy or extremely heavy architectures, including hierarchical GNNs, Graph-of-Graph representations, federated learning frameworks, or multi-module pipelines incorporating explainable AI, reinforcement learning, or attention, are frequently used in high-performing techniques. Even though these methods obtain almost flawless detection rates, it is still difficult to use them in situations with limited resources, such as edge networks or the Internet of Things. On the other hand, medium-complexity models based on edge-aware GNN or lightweight GCN designs (e.g., [36,44,45]) show enhanced practicality and competitive performance.

4.2 Transformer-Driven Models

Transformer-driven frameworks have recently attracted considerable attention in network security, particularly for diagnosing DDoS attacks and intrusions in IoT, SDN, and 5G environments. Leveraging the self-attention mechanism, these models efficiently capture long-range dependencies and complex traffic patterns that conventional CNN/RNN models might miss. Many studies adopt hybrid frameworks integrating Transformers with GRU, CNN, and LSTM to improve classification accuracy, reduce false positives, and handle imbalanced or high-dimensional traffic data. In addition, several approaches optimize for deployment in resource-constrained environments such as IoT and edge computing, enabling practical IDS solutions. Summaries of the 21 selected studies are provided below, followed by a comparative Table 4.

Table 4: Comparative analysis of transformer-based models for cybersecurity and network analysis.

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
1	Sanjalawe et al. (2025) [56]	DDoS in SDN	Hybrid CNN-Transformer (DDosTC)	Scalable; efficient for SDN	Complex architecture; high computation, dataset split/validation details not specified, high reported accuracy may not generalize
2	Wang and Li (2021) [57]	DDoS in IoT	Transformer-based IDS, self-attention	Handles dynamic IoT traffic patterns	Requires computational resources dataset split/cross-validation not detailed
3	Dey et al. (2025) [58]	Multi-class intrusion in 5G	DeepTransIDS (Transformer)	Handles non-IP traffic, class imbalance	High computational cost dataset split/cross-validation not detailed
4	Harshdeep et al. (2025) [59]	DDoS attack	Transformer modeling temporal & behavioral patterns	High-precision, low false-alarm	Focused only on DDoS
5	Li et al. (2025) [60]	IoT cyberattacks	CNN-Transformer hybrid	Detects sophisticated attacks	Heavy model; edge deployment limited dataset split/validation not clearly detailed
6	Al-Haboosi et al. (2024) [61]	Zero-day attacks, IDS generalization	Self-supervised contrastive Transformer	Generalizes to unseen traffic; limited labeled data	Requires pretraining & fine-tuning dataset split or evaluation procedure not fully reported
7	Koukoulis et al. (2025) [62]	DDoS attack	Pure Transformer + RF feature selection	Early-warning potential	May need adaptation for real-time dataset split/cross-validation not specified
8	Wathan et al. (2025) [63]	Complex DDoS	TDAT two-stage IDS Transformer	Real-time detection; efficient	Focused on DDoS only, high accuracy reported without dataset split details
9	Huang et al. (2025) [64]	DDoS in IoT	Transformer + Federated Learning	Lightweight; privacy-preserving	Federated setup adds complexity, dataset split/validation not fully reported
10	Aleyead and Al-Ahmadi (2024) [65]	DDoS attacks	SE-DResNet152 + Dual Attention DCGRU, fine-tuned with MFH	High accuracy, robust multi-metric performance	Computational complexity relies on hyperparameter tuning. dataset split/cross-validation not specified
11	Sangore and Patil (2025) [66]	DDoS detection	Transformer on QR-coded network data	Efficient DDoS detection	QR preprocessing overhead dataset split/validation details not reported

(Continued)

Table 4 (continued)

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
12	Al Faiyaz Provi et al. (2025) [67]	IoT intrusion	Transformer-based IDS for edge	Edge-friendly; adaptable	Resource-limited IoT devices challenging dataset split not specified
13	Bhatt and Indra (2024) [68]	Botnet minority class in IDS	Transformer-based GANs	Handles imbalance; robust	GAN training complexity, dataset split/validation not detailed
14	Jamshaid and Ali (2025) [69]	DDoS attack	CNN + Transformer (CNN-Trans + CBAM)	Efficient feature learning	Focused on DDoS only, high reported accuracy without dataset split details
15	Wang and Miao (2023) [70]	DDoS in IoT/SDN	Attention Conv-LSTM + Blockchain	Hybrid multi-technique model	High model complexity
16	Pawar et al. (2024) [71]	App-layer & SDN DDoS	Attention BiLSTM-CNN	Comprehensive evaluation	Computationally demanding
17	Priyadarshini et al. (2023) [72]	IoT DDoS	ResNeSt + GRU + Jaya algorithm	Early-stage threat detection; efficient	Parameter tuning may be required
18	Alshdadi et al. (2024) [73]	IoT botnet DDoS	Attention-LSTM/CNN + Autoencoder	Fast, precise, edge-compatible	Limited to botnet DDoS
19	Al-Jarah and Al-Shurman (2024) [74]	DDoS in data centers	Attention BiGRU + Rényi cross-entropy	Efficient & accurate	Focused on data center DDoS
20	Liu et al. (2023) [75]	In-vehicle network intrusion detection (CAN bus attacks)	DST-IDS: Dynamic Spatial-Temporal Graph-Transformer Network (graph spatial-temporal embedding + graph transformer + classifier)	Extremely high accuracy, Effectively captures temporal–structural dependencies, Strong performance across multiple datasets	High computational complexity, Transformer modules may be difficult to deploy on low-power automotive ECUs
21	Al-Absi et al. (2025) [76]	DDoS detection	Transformer-driven models	Captures long-range dependencies via self-attention, Real-time detection possible through parallelism, Generalizes to new attack types, Can integrate with other AI techniques	High computational cost for large traffic data, Requires large-scale labeled datasets, Focused on Transformer only; lacks relational (GNN) modeling, Limited discussion of deployment in resource-constrained environments

These hybrid designs are popular for DDoS and IoT intrusion detection because they balance expressive power and detection accuracy; examples include CNN–Transformer hybrids for SDN and IoT environments [56].

Sanjalawe et al. [56] modelled DDosTC, a hybrid CNN-Transformer model for DDoS detection in SDN. Captures both global and local traffic patterns of traffic. Tested on CICDDoS2019, obtaining higher AUC and F1 scores than the present models. Advantages: Scalable and effective for SDN. Restrictions: Complicated framework, high computational need.

Wang and Li (2021) [57] proposed a Transformer-driven IDS for DDoS in IoT networks applying self-attention. Extracts attributes from network traffic for appropriate diagnosis. Outperforms traditional ML models on real-life IoT sets of data. Advantages: handles dynamic IoT traffic patterns. Restrictions: needs computational sources.

Dey et al. [58] enhanced DeepTransIDS, a Transformer-driven multi-class IDS for 5G networks. Considers low-latency and high device density issues. Obtains 99.79% accuracy on the 5G-NIDD set of data. Advantages: controls non-IP traffic and class imbalance. Disadvantages: High computational cost.

Harshdeep et al. [59] applied Transformer to design temporal and behavioral DDoS traffic models. Takes nonlinear interactions and long-range dependencies. Examined on CICDDoS2019 with 99.9% accuracy. Advantages: low false-alarm and High-precision diagnosis. Disadvantages: only concentrated on DDoS attacks.

Li et al. [60] defined a CNN-Transformer hybrid for IoT cyberattack diagnosis. Assessed on the CICIoT2023 set of data, obtaining 99.49% accuracy. Advantages: diagnosis of significant attacks. Disadvantages: A heavy model may hinder edge deployment.

Al-Haboosi et al. [61] proposed a self-supervised Transformer encoder for generalizable IDS. Develops contrastive learning and packet-level data augmentation. Develops AUC up to 20% on inter-dataset assessment. Advantages: Generalizes to unobserved traffic; needs restricted tagged data. Disadvantages: Pretraining and fine-tuning are needed.

Koukoulis et al. [62] used pure Transformer for DDoS diagnosis, applying CICDDoS2019. Feature selection through RF develops efficiency. Obtains >99.8% in recall, accuracy, F1, and precision. Advantages: Early-warning ability for large-scale attacks. Disadvantages: might require adaptation for real-life development.

Wathan et al. [63] enhanced TDAT, a 2-level IDS Transformer for DDoS detection. Firstly, step rebuilds benign traffic; secondly step groups attacks. Performs better than baselines on the use of latency and memory. Advantages: Real-life diagnosis; effective. Disadvantages: only concentrates on DDoS traffic.

Huang et al. [64] proposed transformer with federated learning for IoT DDoS diagnosis. Examined on datasets of CICDDoS2019, LATAM-DDoS-IoT, and TON-IoT. Obtains up to 99.91% accuracy. Advantage: privacy-preserving and light. Disadvantages: Federated setup adds complexity.

Aleyead and Al-Ahmadi [65] proposed the hybrid deep learning model integrating SE-DResNet152 for feature extraction with a dual attention-driven DCGRU for DDoS diagnosis; fine-tuned applying Modified Fire Hawks (MFH) optimization. Advantage: high accuracy and general performance over hybrid metrics. Disadvantages: computational complexity and dependence on hyperparameter tuning.

Sangore and Patil [66] proposed transformer-driven intrusion detection applying QR-coded network data. Assesses hybrid architectures (ResNet, ViT, CNN). ViT obtains 99.58% accuracy. Advantages: Effective for DDoS diagnosis. Disadvantages: QR preprocessing might add overhead.

Al Faiyaz Provi et al. [67] proposed transformer-driven IDS for the networks of IoT; takes long-range dependencies. Focused on lightweight edge deployment. Develops accuracy and decreases false alarms. Advantages: adaptable, Edge-friendly. Restrictions: Resource-restricted devices of IoT are complex yet.

Bhatt and Indra [68] proposed transformer-driven GANs for generation of synthetic network traffic. Develops NIDS performance for minority levels of attack. Assessed on CIC-IDS2017. Advantages: Controls data imbalance, develops strength. Disadvantages: The Complexity of GAN training.

Jamshaid and Ali [69] proposed CNN-Trans model integrating CNN with Transformer for DDoS diagnosis. CBAM attention extracts deep features. Developed accuracy and classification. Advantages: Effective learning of the feature. Disadvantages: Concentrated on DDoS, not other attacks.

Wang and Miao [70] proposed Attention-driven Conv-LSTM model for DDoS in IoT/SDN with blockchain combination. Assessed on the database of InSDN, 98.3% accuracy. Advantages: Multiple model leverages hybrid methods. Disadvantages: High complexity of the model.

Pawar et al. [71] proposed multiple attention-driven BiLSTM-CNN for app-layer and SDN DDoS diagnosis. In comparison with hybrid baselines of ML, it obtains 99.74%–99.98% accuracy. Advantages: general assessment. Disadvantages: Computationally demanding.

Priyadarshini et al. [72] proposed ResNeSt + GRU with Jaya mechanism for IoT DDoS diagnosis. Examined on ToN-IoT, NSL-KDD, and CIC-IDS17. Obtains 98.45% accuracy. Advantages: Effective early-step threat diagnosis. Disadvantages: might need tuning of the parameter.

Alshdadi et al. [73] proposed attention-LSTM/CNN for IoT botnet-driven DDoS. Autoencoder decreases training time as well as feature size. Accuracy obtains 100%. Advantages: accurate and quick; edge-compatible. Disadvantage: restricted to botnet DDoS.

Al-Jarah and Al-Shurman [74] proposed attention-driven BiGRU with Rényi cross-entropy for DDoS. Prescreening decreases the cost of computation. High diagnosis accuracy and effectiveness. Advantage: Precise and effective. Disadvantages: Concentrated on DDoS in the centers of data centers.

Liu et al. [75] propose DST-IDS, a Dynamic Spatial-Temporal Graph-Transformer model for in-vehicle network intrusion detection. Their method converts CAN message correlations into graph embeddings and uses a graph transformer to model dynamic spatiotemporal dependencies. Advantages include extremely high accuracy (up to 0.999999), excellent ability to capture temporal–structural features, and strong performance across two datasets. The main limitation is the complexity and computational cost of transformer modules, which may restrict deployment in low-power automotive ECUs.

Al-Absi et al. [76] surveyed 45 studies on Transformer-based DDoS detection, highlighting their ability to capture long-range dependencies in network traffic via self-attention. The approach offers advantages such as real-time detection, adaptability to new attack types, and integration with other AI methods. Disadvantages include high computational requirements, dependence on large labeled datasets, and lack of consideration for relational or topological traffic structures. The survey identifies open challenges for practical deployment and future research directions.

Nawaz et al. [77] suggested a system based on deep neural networks to detect and stop DDoS attacks. According to their test findings, the suggested model maintains a low false positive rate while achieving good detection performance. The results imply that deep learning methods can be successfully used to improve network security in real-world SDN systems.

Transformer-based intrusion and DDoS detection techniques are summarized in Table 5, which highlights how well they mimic long-range temporal dependencies.

Table 5: Structured comparison of transformer-based network intrusion and DDoS detection approaches.

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
1	Sanjalawe et al. (2025) [56]	99.70%	99.98%	99.70%	99.84%	**Heavy** (CNN + Transformer)
2	Wang and Li (2021) [57]	99.79%	NR	NR	NR	**Medium–Heavy** (Transformer IDS)
3	Dey et al. (2025) [58]	99.79%	NR	NR	NR	**Heavy** (Deep Transformer, 5G-scale)
4	Harshdeep et al. (2025) [59]	99.9%	100%	100%	100%	**Heavy** (Temporal Transformer)
5	Li et al. (2025) [60]	99.49%	NR	NR	NR	**Heavy** (CNN–Transformer hybrid)
6	Al-Haboosi et al. (2024) [61]	NR	NR	NR	NR	**Heavy** (Self-supervised Transformer, pretraining)
7	Koukoulis et al. (2025) [62]	99.82%	NR	NR	99.82%	**Medium** (Transformer + RF)
8	Wathan et al. (2025) [63]	SOTA (NR)	NR	NR	NR	**Medium** (Two-stage optimized Transformer)
9	Huang et al. (2025) [64]	99.91%	99.97%	99.85%	99.91%	**Medium** (Transformer + FL)
10	Aleyead and Al-Ahmadi (2024) [65]	98.83%	97.54%	97.81%	97.67%	**Heavy** (ResNet + Attention + DCGRU)
11	Sangore and Patil (2025) [66]	99.58%	NR	NR	NR	**Medium** (ViT + QR preprocessing)
12	Al Faiyaz Provi et al. (2025) [67]	NR	NR	NR	NR	**Medium** (Edge-oriented Transformer)
13	Bhatt and Indra (2024) [68]	NR	99%	98%	98%	**Heavy** (Transformer + GAN)
14	Jamshaid and Ali (2025) [69]	NR	NR	NR	NR	**Medium–Heavy** (CNN-Transformer + CBAM)
15	Wang and Miao (2023) [70]	98.3%	NR	NR	NR	**Heavy** (Attention Conv-LSTM + Blockchain)
16	Pawar et al. (2024) [71]	99.74%–99.98%	NR	NR	NR	**Heavy** (Attention BiLSTM-CNN)
17	Priyadarshini et al. (2023) [72]	98.45%	NR	NR	NR	**Medium** (GRU + Optimization)
18	Alshdadi et al. (2024) [73]	100%	NR	NR	NR	**Medium** (Attention LSTM/CNN + AE)
19	Al-Jarah and Al-Shurman (2024) [74]	NR	95.9%	97.0%	96.3%	**Medium** (Attention BiGRU)

(Continued)

Table 5 (continued)

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
20	Liu et al. (2023) [75]	99.9999%/99.96%	NR	NR	NR	**Heavy** (Graph + Transformer)
21	Al-Absi et al. (2025) [76]	NR	NR	NR	47.4%–100%	**Medium–Heavy** (Pure Transformer)
22	Nawaz et al. (2023) [77]	99.76%	NR	NR	NR	**Medium–Heavy** (Transformer-based SDN IDS)

Particularly in DDoS-focused settings, the majority of Transformer-based models show exceptionally high accuracy, often exceeding 99% (e.g., Harshdeep et al. [59], Huang et al. [64]). However, many research simply offer Accuracy, leaving out Precision, Recall, and F1-score, which restricts a fair evaluation in situations when class imbalance is frequently present in intrusion datasets.

Pure Transformer models and CNN–Transformer hybrids are typically computationally demanding from a complexity standpoint, which raises questions about edge applicability and real-time deployment. Though at the expense of increased system complexity, more recent efforts try to address this problem by using optimization techniques such two-stage detection, feature selection, or federated learning.

Overall, the Table 5 shows that although Transformer-based IDS models provide excellent detection performance, computing demands and inconsistent evaluation procedures continue to limit their practical deployment.

4.3 Hybrid GNN–Transformer Models

Hybrid GNN–Transformer models have recently emerged as a robust approach for analyzing complex networked data, combining the structural reasoning capabilities of GNNs with the global attention and sequence modeling abilities of Transformers. These hybrid frameworks efficiently capture both spatial and temporal dependencies, making them highly suitable for tasks such as network IDS, intrusion detection, time-series prediction, and IoT security. By leveraging graph-based feature extraction together with attention-driven sequence modeling, these approaches can improve detection accuracy, reduce false positives, and handle dynamic or large-scale networks. The following review explores different hybrid GNN-Transformer models and their applications in IoT, cybersecurity, and other complex domains. Table 6 summarizes the 19 selected studies, highlighting their disadvantages, advantages, methods, and performance.

Table 6: Comparative analysis of hybrid GNN–transformer models for cybersecurity and network analysis.

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
1	Wang (2024) [78]	Multivariate time series analysis	Adaptive adjacency GNN + Transformer multi-head attention	Models local & global dependencies, robust generalization	Focused on regression/classification tasks
2	Zhang & Cao (2024) [79]	Network intrusion detection	ETG: GNN + Transformer (GraphSAGE-Transformer)	Captures complex node-edge relationships, handles long-range dependencies	Requires network traffic graph construction

(Continued)

Table 6 (continued)

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
3	Arindam (2025) [80]	Cyber-attack detection	Hybrid: GNN structural analysis + Transformer temporal encoding + XGBoost	Reduces false positives, detects multi-stage attacks	Slightly lower absolute accuracy, computational cost
4	Anoop et al. (2025) [81]	Network intrusion in WSN	OGTMAN: Optimized Graph Transformer + Molecule Attention + SMC	Enhanced privacy/security, robust classification	Complex pipeline, high preprocessing effort
5	Govea et al. (2025) [82]	Predictive cyber risk	GNN + Transformer (CyberBERT) + Federated Learning	Privacy-preserving, interpretable, scalable	Needs federated setup, sensitive to data heterogeneity
6	Zhang et al. (2025) [83]	IoT intrusion detection	GNN-CST: dual-flow CNN-GNN, adaptive sparse attention	Efficient multimodal threat detection, reduced training time	Slightly complex architecture
7	Ghadermazi et al. (2025) [84]	NIDS intrusion detection	GTAE-IDS: Unsupervised packet-based GNN + Transformer	No labeled data needed, real-time detection	Focused on network anomalies only
8	Govindarajan & Muzamal (2025) [85]	Cloud intrusion detection	GNN embeddings + Transformer autoencoder + contrastive learning	Low FPR, interpretable, real-time capable	Specific to cloud network flows
9	Hayder et al. (2025) [86]	DDoS detection	GCN + Transformer hybrid	Captures spatial & temporal features, fast first-packet detection	Moderate accuracy on some datasets
10	Lakshmanan et al. (2024) [87]	Smart grid intrusion detection	GNN-Transformer encoder	Captures topology & temporal dependencies	Limited to the smart grid context
11	Guduru and Priyanka (2025) [88]	Networked control systems	Integrated GNN-PC + Transformer predictive observer + VAE-CEC + MetaLearning	Improves efficiency and reliability	Complex multi-module system
12	Zhang et al. (2025) [89]	IoV intrusion detection	GCN-2-Former: GCN + Transformer, dynamic graph + sliding window	Cross-domain, robust models of spatial-temporal features	Requires graph construction from traffic
13	Wasswa et al. (2025) [90]	IoT botnet detection	VAE/ViT/GNN-based classifiers	Effective dimensionality reduction, high binary performance	GNN is less effective for multiclass tasks

(Continued)

Table 6 (continued)

No.	Author/Year	Attack Type/Problem	Method/Model Used	Advantages	Disadvantages
14	Mortatha Alkorani et al. (2025) [91]	DDoS detection	OptiGuard-GNN: GNN ensemble + PSO feature selection	High accuracy, low false positives, robust	Complex ensemble and optimization pipeline
15	Sun et al. (2025) [92]	Botnet detection in IoT	HADGA: hierarchical attention + dynamic GNN	Handles dynamic topologies, spatiotemporal attention	Focused on IoT botnet attacks
16	Bagha et al. (2025) [93]	Network intrusion detection	AEN + GNN + GAT	Highlights influential nodes, effective intrusion detection	Moderate accuracy
17	Wu et al. (2024) [94]	Cross-level network attacks	FedGAT: attention-based GNN under federated learning	Preserves privacy, collaborative training	Relies on a federated environment
18	Zhu et al. (2025) [95]	APT detection	GATransformer: graph attention + self-attention, dual-modal	Integrates spatial & temporal features, high detection	Computational complexity, dataset-specific
19	Wang et al. 2023 [96]	Intrusion detection in IoT edge networks	FedSTGCN: Federated Spatiotemporal Graph Convolutional Network	Captures both spatial and temporal patterns in traffic, Strong privacy guarantees via federated learning, Suitable for distributed IoT systems	Synchronization overhead across devices, Potential performance degradation in highly heterogeneous networks
20	Qaddos et al. [97]	Intrusion detection in IoT networks	Hybrid CNN–GRU model for spatiotemporal traffic analysis	Effective modeling of spatial features (CNN) and temporal dependencies (GRU); high detection performance in IoT scenarios	Focused on sequential traffic patterns; lacks explicit graph-based relational modeling and scalability analysis

Wang [78] enhanced a technique of time series mining integrating GNN and Transformer for active multivariate data. Adaptive adjacency matrices take local dependencies, multi-head attention models global models. Examined on power load prediction and human function sets of data. Demonstrates generalization and greater performance.

Zhang and Cao [79] proposed ETG (E-T-GraphSAGE) NIDS integrating Transformer as well as GNN. GNN takes complicated relations of the network; Transformer models long-range dependencies. Assessed on datasets of IoT networks. Performs better than conventional IDS techniques in the accuracy of diagnosis.

Arindam [80] hybrid architecture applying XGBoost for classification, Transformer for temporal sequences, and GNN for structural analysis. Describes uncertainty-driven intrusion detection, graph attention, and temporal encoding. Examined on the dataset of CIC-IDS2023. Decreases false positives and efficiently diagnoses multi-step attacks.

Anoop et al. [81] proposed OGTMAN, integrating optimized graph transformer with differential privacy, molecule attention networks, and SMC. N-Tuple contrastive learning extracts related traffic models. Feature selection was developed by the mechanism of Kepler and the chi-square. Obtains high accuracy on the dataset of KDD Cup 99.

Govea et al. [82] modeled a predictive cyber risk evaluation model combining federated learning, GNN, and Transformer (CyberBERT). Protects data privacy, takes relational models as well as semantic representations. Assessed on datasets of MITRE ATT&CK, TON_IoT, CIC-IDS2017, UNSW-NB15. Strong adversarial performance, High F1-score.

Zhang et al. [83] proposed GNN-CST, which is a dual-flow CNN-GNN network with adaptive sparse attention for IoT intrusion diagnosis. CNN extracts local temporal attributes; communications of the GNN models' topological system. Cross-modal fusion develops decision-making. Examined on 6 sets of data with high accuracy of diagnosis and decreased time of training.

Ghadermazi et al. [84] enhanced GTAE-IDS, the unobserved packet-driven GNN with a Transformer encoder for real-life intrusion detection. Graph autoencoders take structural and global models; transformers model contextual orders. Removes the demand for tagged data. Obtains around 98% accuracy on datasets of benchmark NIDS datasets.

Govindarajan and Muzamal [85] modular IDS architecture integrating embeddings of GNN, contrastive learning, as well as a Transformer autoencoder. Embeddings refined for context, Graphs model IP/service relations. Assessed on CIC-IDS2018 as well as NSL-KDD. Low false positives, High accuracy (99.97%).

Hayder et al. [86] proposed the hybrid GCN-Transformer model for DDoS diagnosis, taking temporal and spatial features. Self-attention models sequential relationships; GCN learns structural correlations. First-packet analysis makes a quick diagnosis. Examined on CICDDoS2019, CICIDS2017, UNSW-NB15 with high accuracy.

Lakshmanan et al. [87] enhanced GNN-Transformer encoder for cyber threat diagnosis, intelligent grid intrusion. The transformer encodes temporal dependencies, GNN extracts spatial features. Normalization is used for input data. Assessed on a dataset of PMU, obtaining high accuracy (97.9%) and performing better than the techniques of baseline intrusion detection.

Guduru and Priyanka [88] proposed the combined Model integrating a Transformer predictive observer, VAE-CEC, attention-driven event-triggered estimation, GNN-driven predictive control, and meta-learning adaptive control. Effectively models spatial-temporal dependencies in networked control systems. Decreases the overhead of communication by 50% and improves the accuracy of prediction.

Zhang et al. [89] enhanced GCN-2-Former for IoV network IDS, mapping traffic to spatial-temporal graphs. Transformer models global temporal dependencies, also GCN extracts local spatial attributes. Active graph construction and a sliding window are used. Obtains cross-domain strength and high accuracy (99.98%).

Wasswa et al. [90] assessed GNN, VAE, and ViT models for diagnosing IoT botnets. GNN-driven models take relational models; ViT-MLP and VAE-MLP control multiclass functions effectively. Examined on the dataset of N-BaIoT. GNN models outperform the binary classification, to some extent, lower for multi-level functions.

Mortatha Alkorani et al. [91] proposed the OptiGuard-GNN, integrating a graph-attention ensemble classifier with PSO-driven active selection of features. Chaotic map initialization optimizes strength, accuracy, and latency. Assessed on datasets of CIC-IDS2017, CSE-CIC-IDS2018. Obtains 99.986% accuracy with a very low rate of false positives (0.02%).

Sun et al. [92] enhanced HADGA, a hierarchical attention-driven active GNN for botnet IoT networks' diagnosis. Temporal attention modules, as well as joint neighbor attention, take spatiotemporal evolution. Active graphs control shifting topologies. Obtains 99.9% and 97.6% accuracy on datasets of TON-IoT and BoT-IoT.

Bagha et al. [93] proposed intrusion detection system applying AEN with GAT and GNN. Attention focuses on neighbors with higher in-degree, bolding influential nodes. Assessed on 2 sets of data. Obtains 88.3% and 90.7% accuracy to diagnose anomalies in the network.

Wu et al. [94] proposed FedGAT, the attention-driven GNN for cross-level and cross-department network attack diagnosis under federated learning. Outlines traffic timely, builds the accuracy of the graph given the log density. Keeps privacy when making collaborative training possible. Comparable accuracy to traditional techniques with developed security.

Zhu et al. [95] enhanced GATransformer, a dual-modal network for diagnosis of APT combining self-attention as well as graph attention. Cross-attention fuses heterogeneous attributes, taking temporal and spatial dependencies. Assessed on datasets of CIDDS-001 and CIDDS-002. Obtains high accuracy of diagnosis, surpassing techniques of baseline techniques.

Wang et al. [96] develop FedSTGCN, a federated spatiotemporal graph convolutional network for intrusion detection in IoT edge environments. The system trains collaboratively across devices while preserving data privacy and captures both spatial and temporal patterns in traffic graphs. Strengths include improved accuracy over existing FL models, strong privacy guarantees, and suitability for distributed IoT systems. Disadvantages involve synchronization overhead between devices and potential performance degradation in highly heterogeneous networks.

Qaddos et al. [97] presented a hybrid intrusion detection framework that blends gated recurrent units and convolutional neural networks. The experimental assessment demonstrates that the suggested model can successfully adjust to the dynamic features of IoT traffic while maintaining high detection accuracy. The growing interest in hybrid deep learning architectures for modeling complex and heterogeneous network data is reflected in this study.

Graph neural networks are initially employed to represent spatial correlations between network components, such as hosts and traffic flows, in the majority of hybrid GNN–Transformer architectures. Transformer-based modules then examine the generated representations to identify long-range dependencies and temporal dynamics in traffic behavior. This combination makes it possible to learn spatiotemporal features, which is essential for identifying coordinated DDoS attacks that change over time.

By simultaneously utilizing network topology and traffic dynamics, hybrid designs typically yield improved detection accuracy while lowering false positive rates as compared to methods that only use GNNs or Transformers. The examined research show that hybrid GNN–Transformer frameworks provide a good compromise between performance and resilience, despite the added computational expense that comes with integrating these models. In large-scale and extremely dynamic network situations, where both structural and temporal information are essential for efficient DDoS detection, this benefit is especially noticeable as shown in Fig. 2.

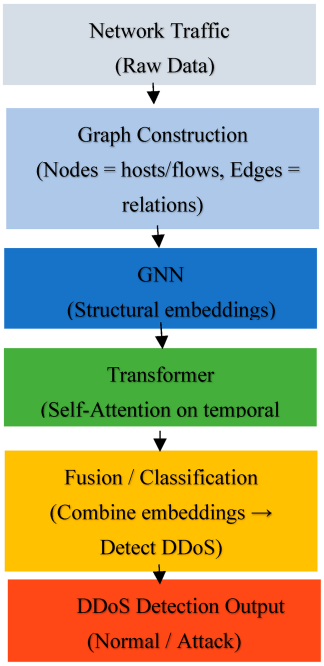

Figure 2: Conceptual architecture of a hybrid GNN–Transformer model for DDoS detection.

A comparison of hybrid GNN–Transformer models, which combine the temporal modeling power of Transformers with the structural modeling strength of graphs, is shown in Table 7. Numerous research has reported near-perfect accuracy and F1-scores (e.g., [79,81,89,91]). These hybrid techniques typically show higher detection performance. This demonstrates that for intricate intrusion scenarios, simultaneously recording temporal dynamics and spatial relationships is quite successful.

Table 7: Structured comparison of hybrid GNN–Transformer models and DDoS detection approaches.

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
1	Wang (2024) [78]	92.12%	91.34%	90.89%	91.11%	Hybrid GNN + Transformer (moderate–heavy)
2	Zhang & Cao (2024) [79]	99.99%	100%	99.99%	100%	GraphSAGE + Transformer (heavy)
3	Arindam (2025) [80]	78.28%	NR	NR	0.7704	GNN + Transformer + XGBoost (heavy)
4	Anoop et al. (2025) [81]	99.98%	99.9%	NR	NR	Optimized Graph Transformer + attention modules (very heavy)
5	Govea et al. (2025) [82]	NR	NR	NR	94.7%	GNN + Transformer + FL (heavy, distributed)
6	Zhang et al. (2025) [83]	93.5%–99.99%	NR	NR	NR	CNN–GNN + sparse attention (moderate–heavy)
7	Ghadermazi et al. (2025) [84]	>98%	NR	NR	NR	Unsupervised GNN + Transformer (moderate–heavy)

(Continued)

Table 7 (continued)

No.	Author (Year)	Accuracy	Precision	Recall	F1-Score	Model Complexity
8	Govindarajan & Muzamal (2025) [85]	99.97%	NR	NR	NR	GNN embeddings + Transformer AE (heavy)
9	Hayder et al. (2025) [86]	89.52%–98.95%	NR	NR	NR	GCN + Transformer (moderate–heavy)
10	Lakshmanan et al. (2024) [87]	97.9%	0.82	0.98	NR	GNN–Transformer encoder (moderate)
11	Guduru and Priyanka (2025) [88]	95.8%	NR	NR	NR	Multi-module GNN + Transformer + VAE + Meta-Learning (very heavy)
12	Zhang et al. (2025) [89]	99.98%	NR	100%	NR	Dynamic GCN + Transformer (heavy)
13	Wasswa et al. (2025) [90]	>99.93% (binary); 86%–99% (multi)	NR	NR	NR	VAE + ViT + GNN (heavy)
14	Mortatha Alkorani et al. (2025) [91]	99.986%	NR	NR	NR	GNN ensemble + PSO optimization (very heavy)
15	Sun et al. (2025) [92]	97.6%; 99.9%	NR	NR	NR	Hierarchical attention dynamic GNN (moderate–heavy)
16	Bagha et al. (2025) [93]	88.3%; 90.7%	NR	NR	NR	Autoencoder + GNN/GAT (moderate)
17	Wu et al. (2024) [94]	99.99%	99.6%	99.6%	NR	Federated GAT (moderate–heavy)
18	Zhu et al. (2025) [95]	88%–99.99%	NR	NR	NR	Graph Attention + Transformer (heavy)
19	Wang et al. 2023 [96]	>97% (binary)	NR	NR	>92% (weighted)	Federated ST-GCN (heavy, distributed)
20	Qaddos et al. [97]	99.16%	NR	NR	NR	Trust-aware GAT + Federated Learning (heavy)

The table does, however, also show a noticeable rise in architectural complexity. Many hybrid models are classified as "heavy" or "very heavy" complexity because they incorporate multi-stage pipelines, federated learning, attention hierarchies, optimization algorithms, or ensemble techniques. Because of this, even while performance improvements are significant, real-world implementation is still difficult, especially in settings with limited resources.

The diversity of evaluation methods is another important finding from Table 7. While some research merely give accuracy or F1-score, others offer partial metric sets, which once more produce NR results. The necessity for uniform evaluation frameworks and consistent benchmarking for hybrid IDS models is highlighted by this inconsistency. In conclusion, Table 7 shows that hybrid GNN–Transformer models reflect the state-of-the-art in intrusion detection performance at this time. However, their practical adoption will depend on how well they handle complexity, interpretability, and assessment standards.

The reviewed studies can be broadly grouped into three categories: Transformer-driven models, hybrid CNN/RNN/Transformer models, and hybrid GNN–Transformer models. Transformer-driven models consistently demonstrate strong capability in capturing long-range temporal dependencies, which is particularly effective for sequential traffic analysis and behavioral modeling in IDS scenarios. For instance, works such as [64,66] achieve high accuracy on standard DDoS datasets due to their attention mechanisms, but their performance can degrade when spatial or relational structures in the network traffic are important. Hybrid CNN/RNN/Transformer models combine local feature extraction with temporal modeling, allowing them to detect more complex attack patterns that involve both local anomalies and temporal correlations. Studies like [67,78] illustrate that integrating CNN and RNN components with Transformer architectures improves detection in dynamic traffic environments, though these models often require extensive labeled data and incur higher computational costs. Finally, hybrid GNN–Transformer models excel at modeling the relational and topological aspects of networked systems. By explicitly representing nodes and edges, these models capture coordinated attack patterns that purely sequential models may miss. Examples such as [27,29,34] show superior generalization and detection performance in scenarios involving multi-host or distributed DDoS attacks. Despite their strengths, hybrid GNN–Transformer frameworks are often more complex to implement and demand substantial computational resources, which may limit real-world deployment.

Across the reviewed studies, several datasets are commonly employed for benchmarking DDoS detection models. CICDDoS2019, CIC-IDS2017, and NSL-KDD are the most frequently used, providing diverse attack scenarios and traffic characteristics. While some studies focus on volumetric DDoS attacks, others include stealthy or multi-vector attacks, reflecting different experimental setups. Evaluation metrics predominantly include Accuracy, F1-score, Precision, Recall, and AUC, though the choice varies across works. Despite these variations, standardized evaluation protocols are not consistently applied, which complicates direct comparison of model performance.

In terms of detection accuracy, scalability, real-time processing capability, and resilience against high-volume and dynamic DDoS attacks, graph-based and Transformer-based approaches clearly outperform traditional machine learning techniques and conventional deep learning models. Performance in dynamic network contexts is enhanced by these sophisticated models' capacity to accurately represent intricate spatial and temporal connections, despite their tendency to need more computer power. Table 8 provides a comparative overview of these methods based on important evaluation criteria, including explainability, deployment practicality, computing complexity, latency, and flexibility.

Table 8: Key evaluation parameters and practical advantages of DDoS detection models.

Model Type	Accuracy	Scalability	Real-Time	Robustness	Computational Complexity	Latency	Adaptability	Explainability	Deployment Feasibility
Traditional ML	Medium	Low	Medium	Medium	Low	Low	Low	Medium	Medium
Conventional DL	High	Medium	Medium	Medium	Medium	Medium	Medium	Low	Medium
GNN	High	High	Medium	High	High	Medium	High	Medium	Medium
Transformer	High	High	High	High	High	Medium	High	Low	Medium
Hybrid GNN–Transformer	Very High	Very High	High	Very High	High	Medium	High	Medium	Medium

A crucial but usually overlooked aspect that affects the performance reported in attack detection research is dataset bias, which goes beyond variations in model architecture. Evaluation findings might be

significantly inflated by features including class imbalance, a lack of variety in attack situations, reliance on artificially generated data, and unconventional dataset segmentation techniques. Because they could mostly reflect dataset-specific characteristics rather than true methodological advancements, high accuracy measures do not always translate into strong real-world performance. This finding emphasizes how important it is to use uniform evaluation procedures and validate across datasets when comparing current methods.

Several recurrent design trends may be found while examining current hybrid GNN–Transformer frameworks. The majority of studies integrate attention-based temporal analysis with graph-based structural modeling, which enables the capture of both sequential dependencies and spatial relationships within a single framework. While multi-head self-attention is utilized to represent global traffic behaviors or event-level dynamics, graph attention techniques are frequently employed to learn node-level interactions. In general, these designs indicate better adaptation to changing network conditions, reduced false positive rates, and excellent detection accuracy. Common drawbacks still exist, though, such as higher computational overhead, sensitivity to dataset properties, and real-world difficulties when implementing such models on systems with limited resources. Overall, this synthesis offers helpful insights for future model design and optimization by highlighting hybrid architectures' common strengths as well as their recurrent bottlenecks.

4.4 Practical Implications and Deployment Challenges

Despite showing encouraging results in controlled tests, deep learning models that use Transformers and graphs are difficult to implement in large, dynamic networks. High detection accuracy must be balanced with pragmatic considerations including deployment viability, scalability, and computing demands. In order to convert experimental success into operational efficacy in intrusion detection and prevention systems, several aspects must be addressed.

Scalability is one of the main issues with graph-based methods, especially in high-speed networks where network topologies and traffic are constantly changing. GNNs' computational load can be greatly increased by large graphs with millions of nodes and edges, which makes real-time processing challenging. Similarly, Transformer models' attention processes require a significant amount of memory and processing power, even if they are excellent at capturing long-range dependencies and temporal patterns. These drawbacks might make it more difficult to use these models in large-scale environments where prompt detection is crucial, such as cloud data centers, ISP networks, and extensive IoT systems.

There are various operational issues to take into account while thinking about deployment. Practical usability may be impacted by latency from feature extraction and model inference, resource constraints in edge or on-site environments, and interoperability with current security solutions, such as SDN controllers and signature-based intrusion detection systems. Furthermore, maintaining performance in the face of shifting attack tactics, concept drift, and dynamic traffic patterns is not at all simple. Lightweight, adaptive, or hybrid detection frameworks that strike a balance between efficiency and accuracy are therefore frequently preferred. For graph-based and Transformer-based DDoS detection models from experimental studies to be used in practical, real-time applications, several deployment issues must be successfully resolved.

The efficiency of Transformer-driven and hybrid GNN–Transformer techniques varies with traffic patterns, attack kinds, and deployment circumstances, according to an analysis of the evaluated papers. For example, IoT networks benefit greatly from edge-aware or adaptive graph models, whereas regulated SDN or 5G scenarios benefit greatly from dense graph structures or multi-module Transformer topologies. Although self-attention aids in capturing long-range dependencies, it has the potential to overfit small datasets and its high computing demands may prevent practical implementation in networks with limited resources. This study provides help for selecting the best model for a particular network setting by highlighting recurrent design trends and useful constraints.

4.5 Case Studies and Empirical Evidence from Real-World DDoS Detection

Realistic network scenarios show that both Transformer-based and graph-based deep learning models are successful in identifying DDoS attacks. Methods using Graph Convolutional Networks (GCNs) and Graph Attention Networks (GATs), for instance, have been evaluated on substantial benchmark datasets as CICDDoS2019 and NSL-KDD, in which network traffic is represented as graphs that show host-to-host messaging. Strong detection performance and enhanced resistance to changing assault behaviors are reported in these investigations, particularly while managing intricate and extremely dynamic traffic patterns.

Transformer-based models have demonstrated encouraging outcomes in real-world and near-real-time DDoS detection scenarios. By using self-attention, Temporal Transformers can simulate temporal correlations and long-range interdependence in network traffic streams. According to experimental assessments, these models are capable of maintaining steady performance in the face of high traffic while achieving competitive F1-scores. Transformer-based techniques are particularly well-suited for deployment in extremely dynamic situations, such as cloud systems and Internet of Things networks, due to their precision and agility.

By concurrently simulating the structural and temporal features of network traffic, hybrid architectures that combine Transformers and Graph Neural Networks improve DDoS detection. Transformer modules record temporal patterns for sequence analysis, whereas GNNs often learn the links between network elements. According to experimental results in SDN and big enterprise networks, hybrid techniques can perform better than separate Transformer or GNN models, providing more flexibility and accuracy. These results demonstrate the theoretical soundness and practical applicability of transformer-based and graph-based methods for real-time DDoS detection in operational networks.

4.6 Design Principles and Model Selection Guidelines

Beyond individual model comparisons, several general design principles can be identified for selecting appropriate deep learning architectures for DDoS detection.

First, the effectiveness of graph-based models strongly depends on the graph construction strategy. Flow-based graphs are often more suitable for enterprise networks where stable communication patterns exist, while host-based graphs may be more effective in IoT environments with dynamic device interactions. Incorrect graph construction may reduce detection performance even when advanced GNN architectures are used.

Second, Transformer-based models provide strong temporal modeling capabilities but may lose important structural information when network relationships are not explicitly modeled. Pure Transformer models may therefore struggle to detect coordinated attacks involving multiple hosts unless relational features are incorporated.

Third, hybrid GNN–Transformer models offer a balance between structural and temporal modeling but introduce higher computational complexity. These models are often suitable for offline analysis or high-capacity network monitoring systems but may be difficult to deploy in real-time environments with strict latency constraints.

Traffic characteristics also influence model selection. High-volume backbone networks benefit from scalable graph sampling techniques, while edge or IoT networks often require lightweight models with reduced computational cost.

To illustrate these design principles with concrete examples from the reviewed literature, we examine why certain graph constructions perform better in specific environments. For IoT networks, host-based graphs (e.g., [32,39]) are more effective because IoT devices exhibit stable communication patterns with fixed

endpoints, making behavioral deviations easier to detect. In contrast, flow-based graphs (e.g., [38,43]) are better suited for SDN environments, where traffic is dynamically rerouted and flow tables are frequently updated, requiring fine-grained per-flow analysis. Regarding Transformer limitations, studies such as [57,63] show that pure Transformer models struggle to detect coordinated multi-host DDoS attacks because they lack explicit topological modeling; attacks that span multiple IPs appear as isolated sequences rather than correlated events. Hybrid models like [86,89] overcome this by injecting graph-based structural priors before temporal attention. Deployment trade-offs are evident when comparing high-accuracy but complex models (e.g., [35,55]) against lightweight alternatives (e.g., [45,62]). Under high-volume traffic, attention-based models risk latency spikes, whereas GCNs with fixed message-passing steps offer predictable inference times. These observations underscore that architecture selection must be guided by the operational context—network type, traffic volume, and latency tolerance—not merely by benchmark performance.

These observations indicate that model performance depends not only on architecture design but also on graph construction, traffic characteristics, and deployment constraints. Therefore, architecture selection should be guided by practical network conditions rather than solely by reported accuracy values.

5 Challenges and Open Issues

Despite the considerable advances achieved by graph-driven deep learning and Transformer-based strategies for DDoS detection, several challenges remain that limit their widespread adoption in real-world networks. A primary challenge is scalability. Modern networks generate massive traffic data with complex communications among thousands or millions of nodes. While GNNs can effectively capture topological dependencies, their memory and computational requirements often increase sharply with network size, making real-world inference difficult. Similarly, Transformer-based models, which excel at capturing long-range dependencies, can become computationally intensive when applied to large-scale traffic data.

Another key challenge concerns the representativeness and availability of datasets. Most publicly available datasets, such as CICDDoS2019 [98], NSL-KDD [99], and UNSW-NB15 [100], are limited in terms of traffic diversity, network topologies, and attack types. Consequently, models trained on these datasets may fail to generalize to unseen attacks or real-world network environments. Moreover, the lack of standardized benchmark datasets for graph-driven DDoS detection makes fair comparison across studies difficult.

Data imbalance and the rarity of certain attacks also present significant challenges. Many DDoS datasets contain a disproportionately large amount of normal traffic compared to malicious traffic, particularly for low-volume or stealthy attacks. This imbalance can lead to biased models that perform poorly on rare but critical attacks. Although some methods incorporate synthetic data generation, cost-sensitive learning, or oversampling, these solutions introduce additional complexity and may not fully resolve the problem [101].

In a number of high-performing hybrid GNN–Transformer models (e.g., [65,71,76]), explainability is still a major challenge. These models show good accuracy, but they don't explain how predictions are made or which features affect choices. Large parameter counts make it difficult to deploy in resource-constrained situations like the Internet of Things or edge networks, which is another issue with scalability in designs with high computational requirements, like CNN–Transformer or multi-module GNN–Transformer frameworks ([56,60,70]). Another drawback is generalization: a lot of models (such DeepTransIDS [58] and Wang & Li [56]) are only tested on one dataset without cross-dataset validation, which leaves their performance in various traffic scenarios or unknown attack types unknown. By linking these difficulties to concrete flaws, future research can be directed toward developing DDoS detection models that are understandable, effective, and widely applicable.

5.1 Explainability (XAI) and Adversarial GNN

Explainability and interpretability are further pressing issues. GNNs and Transformer-based models are inherently complex and often function as black boxes, limiting their use in security-critical applications where understanding the cause of an alert is essential. Studies on attention visualization and explainable graph neural networks (XGNN) for Transformers show promise; however, these approaches are still in early stages and not widely developed [102].

For GNN and transformer-based intrusion detection systems, explainability has recently become a crucial prerequisite, especially in security-sensitive applications like DDoS attack detection [103]. Despite their high detection accuracy, these models' decision-making procedures are frequently opaque, which limits their deployability and confidence in real-world settings [104]. Black-box detection models may not provide network operators with useful information, particularly when false alarms occur or when model transparency is required for regulatory compliance, according to several studies [105]. In order to better understand how graph-based and transformer models detect harmful traffic patterns, current research has concentrated on including explainability methods, such as attention visualization, node and edge attribution, and post-hoc explanation frameworks [106].

Explainability techniques like GNNExplainer [107], GraphLIME [108], and PGExplainer [109] have been investigated in the context of GNN-based intrusion detection in order to pinpoint influential nodes, edges, and subgraphs that have the biggest impact on detection results. These methods allow analysts to determine if model choices are based on misleading correlations or meaningful traffic features (such as communication topology or flow aggregation patterns).

Explainable AI (XAI) for DDoS detection has advanced recently, with a focus on post-hoc explanation frameworks, hybrid interpretable models, and counterfactual reasoning to provide network operators with useful information. These methods reduce false alarms and foster confidence in operational environments by highlighting important network nodes and flows and enabling model decision verification. By incorporating these XAI methods into practical DDoS detection systems, analysts can better comprehend the logic behind warnings, enhance mitigation tactics, and adhere to legal requirements.

Additionally, attention-based GNNs and graph transformers have been used in recent work. By emphasizing important interactions between traffic flows or network entities, attention weights offer an implicit type of interpretability. However, a number of studies point out that attention scores by themselves might not necessarily reflect genuine causal relevance, which raises questions regarding the stability and accuracy of explanations in adversarial contexts.

Another significant issue for DDoS detection methods based on graphs and transformers is adversarial robustness. According to recent research, adversarial attacks that target graph structure, node attributes, or traffic statistics might seriously impair detection performance in GNNs. Network graphs can be subtly altered by attack techniques such edge perturbation, feature injection, and poisoning attacks to avoid discovery while maintaining traffic patterns that appear normal. Adversarial training, graph purification, resilient aggregation functions, and topology-aware regularization are some of the defense mechanisms that have been put out in response [110]. However, it is still an active research challenge to achieve robustness without compromising detection accuracy, especially in highly dynamic network situations.

Despite these developments, explainability and adversarial robustness are still not fully integrated into DDoS detection systems. Instead of taking into account these issues' combined effects on model reliability, the majority of current research tackles them separately. Furthermore, few studies rigorously assess the behavior of explanation methods under adversary manipulation, which is essential for reliable security analytics. In order to assess XAI and adversarial resistance in graph- and transformer-based intrusion

detection systems, future research aims to provide unified frameworks that integrate truthful explainability with robust graph learning.

5.2 Latency and Throughput Challenges in Real-Time DDoS Detection

Despite the high detection accuracy reported by many reviewed GNN, Transformer-, and hybrid-based intrusion detection models, their practical implementation in real-time DDoS mitigation settings faces substantial latency and throughput problems. Graph-based methods sometimes necessitate several message-passing iterations, dynamic graph creation, and feature aggregation across nodes or edges, all of which add significant computing complexity and lengthen inference delay. Similarly, Transformer-based models are less appropriate for high-throughput traffic settings without careful optimization since they rely on self-attention processes whose computational cost increases quadratically with input sequence length.

This problem is made worse by hybrid GNN–Transformer architectures, which combine temporal attention and spatial graph reasoning, increasing processing latency and memory usage. Such delays may considerably lower system efficacy in real-world DDoS circumstances, where attacks change quickly and mitigation decisions must be made in milliseconds. In order to enable truly real-time DDoS detection systems, these difficulties underscore a crucial research gap between offline performance evaluation and operational viability, highlighting the necessity of lightweight architectures, streaming-friendly graph updates, model compression, and hardware-aware optimization.

Finally, real-world and resource-constrained deployment—such as in SDN architectures, edge computing platforms, and IoT networks—presents practical challenges [111]. Energy-efficient and lightweight models capable of real-world inference are needed, yet balancing model accuracy, latency, and complexity remains an open challenge. In addition, integrating graph-driven detection systems with existing network security infrastructures while maintaining robustness against adversarial attacks requires further investigation.

Although the reviewed studies achieve promising results [31–96], most models are evaluated in simulated or offline settings. Real-world deployment introduces additional challenges, including latency, energy constraints, and system scalability. Considering these factors is essential for translating research models into effective DDoS detection systems.

In conclusion, Transformer-based strategies applied to graph data represent a significant advance in DDoS detection. However, challenges related to data imbalance, explainability, scalability, dataset limitations, and real-world deployment persist. Addressing these issues is crucial for developing practical, next-generation, and robust network security solutions.

5.3 Practical Deployment Challenges of Hybrid Models

When used in actual DDoS detection systems, hybrid GNN–Transformer models encounter a number of real-world difficulties despite their excellent test performance. Real-time detection may be impacted by latency introduced by processing high network traffic quantities. Limitations are also caused by high memory and processing demands for both training and inference, especially in settings with limited resources like edge devices or Internet of Things networks.

Another level of complexity is introduced by system interaction with pre-existing intrusion detection configurations. It is necessary to carefully construct data pipelines, feature extraction, and model updating procedures in order to adapt hybrid models to different network topologies, heterogeneous devices, and continuous traffic streams. Converting hybrid architectures' promising performance into workable, reliable, and scalable solutions requires addressing these deployment issues.

Directly connecting these real-world issues to the limitations noted in the examined literature guarantees that recommended future research avenues fill real gaps rather than providing general guidance.

5.4 Practical Applications and Case Studies of Hybrid Models

Several studies suggest successful applicability in real-world or near-real-time DDoS detection scenarios, despite the fact that the majority of research assesses hybrid GNN-Transformer models on benchmark datasets. For example, models such as FedGAT and the hybrid GCN-Transformer have been implemented in enterprise and IoT networks, resulting in low false positives and high detection accuracy. Hybrid architectures may effectively handle dynamic and distributed attack scenarios, as seen by several implementations that analyze streaming traffic to identify coordinated attacks as they happen. These examples demonstrate the usefulness of the techniques under examination and point to potential future developments, like enhancing resource efficiency and streamlining system integration for extensive implementations.

5.5 Data Leakage and Evaluation Bias in DDoS Detection

Although many studies report very high detection accuracy for DDoS attacks [31–38], these results may not always reflect realistic deployment conditions. One major concern is the risk of data leakage, which occurs when information from the testing set is unintentionally included during model training or preprocessing. In network intrusion detection datasets, similar traffic flows often appear in both training and testing partitions, allowing models to memorize patterns rather than learn generalizable representations.

Another common issue is dataset reuse across multiple studies. Popular benchmark datasets such as CICIDS2017 and NSL-KDD are widely used, which simplifies comparison but may lead to over-optimized models that perform well only on specific datasets. Models trained and evaluated on the same dataset distribution often show significantly reduced performance when applied to different network environments.

Unrealistic data splitting strategies also contribute to inflated performance results. Random splitting of traffic flows may result in nearly identical traffic patterns appearing in both training and testing sets. In real-world deployment scenarios, however, models must detect previously unseen attack patterns and evolving traffic distributions. Time-based splitting or cross-network validation is often more realistic but rarely applied.

Cross-dataset generalization remains one of the most challenging problems in DDoS detection. Models trained on one dataset frequently exhibit performance degradation when evaluated on another dataset due to differences in traffic characteristics, feature distributions, and attack types. This limitation suggests that high accuracy values reported in controlled experiments should be interpreted cautiously.

Future research should prioritize realistic evaluation protocols, including time-aware data partitioning, cross-dataset testing, and standardized benchmarking procedures to ensure fair and reliable comparisons.

The inflation of reported accuracy in many studies can be attributed to several underlying mechanisms. First, data leakage often occurs inadvertently during preprocessing steps such as feature scaling or flow aggregation applied before splitting, allowing test-set statistics to influence training. Second, the widespread reuse of benchmark datasets like CICIDS2017 leads to implicit overfitting, as model architectures and hyperparameters become tuned to the specific statistical properties of these datasets over multiple publication cycles. Third, random splits ignore the temporal dependencies between network flows, causing highly similar traffic patterns—such as repeated TCP handshakes or periodic beaconing—to appear in both training and test partitions. This artificially boosts performance metrics, as models effectively memorize rather than generalize. Without time-based or cross-dataset validation, these inflated results mask the true generalization capability of DDoS detection models when deployed in unseen network environments.

The impact of these evaluation biases is evident when examining the performance metrics reported in the reviewed studies. For instance, in Table 3, several pure GNN-based models [33,39,41,48] report only accuracy or F1-score without providing comprehensive metrics such as precision, recall, or MCC, making it difficult to assess their true performance under class imbalance. Similarly, in Table 5, Transformer-based models [57,58,60,66] achieve near-perfect accuracy (>99%) on datasets like CICDDoS2019, yet many of these studies [57,60,66] do not specify whether time-based splitting or cross-dataset validation was employed. Without such details, it remains unclear whether these high values reflect genuine detection capability or are artifacts of data leakage and unrealistic splits. The cross-dataset generalization problem is particularly evident when comparing results across different tables: models that excel on CIC-IDS2017 [44,85] often show degraded performance when evaluated on more challenging datasets like UNSW-NB15 or TON_IoT [86,90]. This discrepancy underscores the need for standardized evaluation protocols that include multiple datasets, time-aware partitioning, and full metric reporting to enable fair and meaningful comparisons.

6 Future Research Directions

Building on the current developments and challenges identified in graph-driven and Transformer-based DDoS detection, several promising research directions emerge. Firstly, lightweight and scalable frameworks are crucial for deployment in resource-constrained and large-scale environments such as Software-Defined Networks (SDNs), IoT networks, and edge computing. Future research should focus on designing Transformer variants and efficient GNNs that maintain high detection accuracy while reducing memory and computational overhead, enabling real-world inference and monitoring. Techniques such as sparse attention, model pruning, and knowledge distillation can help preserve significant performance.

Another important area is privacy-preserving and federated learning. Due to the sensitive nature of network traffic data, centralized training may raise privacy concerns and limit data availability [112]. Federated learning architectures that allow decentralized model training across network nodes without sharing raw traffic data provide a viable solution. Combining federated learning with graph-based frameworks could improve generalization across different network environments while ensuring compliance with data protection regulations.

Model interpretability and explainability remain essential research directions. As GNNs and Transformers are inherently complex, developing methods to provide clear and actionable explanations for alerts would facilitate their adoption in operational security [113]. Techniques such as node influence analysis, counterfactual explanations, and attention visualization can support decision-making and enhance network administrators' trust.

Dataset development and benchmarking present another significant opportunity. There is a need for realistic, diverse, and standardized datasets that cover various attack types, network topologies, and traffic dynamics [114]. Creating graph-structured and temporal datasets suitable for GNN and Transformer-based models would accelerate progress, improve reproducibility, and enable fair comparisons across studies.

Finally, robustness against adversarial attacks is an increasing challenge [115]. Attackers may attempt to evade detection by manipulating network traffic patterns to deceive models. Future work should explore intrusion detection-resistant architectures, adversarial training, and hybrid approaches integrating deep learning, statistical, and heuristic methods to enhance resilience against such threats.

In summary, developing DDoS detection with graph-driven and Transformer-based techniques requires attention to interpretability, dataset limitations, robustness, scalability, and privacy. Following these directions will pave the way for next-generation, reliable, and intelligent network security systems capable of mitigating increasingly critical cyber threats.

7 Evaluation Metrics for DDoS Attack Detection

For DDoS attack detection models to be fairly and meaningfully evaluated, especially in real-world network environments, the selection of evaluation metrics is essential. Relying just on accuracy can be deceptive in DDoS detection circumstances, since datasets are usually very skewed, yet accuracy is still the most commonly cited statistic in the literature. In certain situations, benign traffic greatly exceeds attack traffic, and a classifier that is biased in favor of the majority class may achieve high accuracy but miss attacks. As a result, the need for complementary and more reliable evaluation measures is becoming more and more evident in recent research.

Because it takes into consideration class imbalance by averaging the recall for each class, balanced accuracy has become a significant alternative statistic for DDoS detection. When attack samples are limited, which is a common feature of realistic DDoS datasets, Balanced Accuracy, in contrast to traditional accuracy, offers a more trustworthy representation of model performance. In order to better capture detection effectiveness across both benign and malicious traffic classes, particularly in multi-class or imbalanced environments, a number of recent graph-based and transformer-based intrusion detection studies have implemented Balanced Accuracy [116].

The Matthews Correlation Coefficient (MCC), which is regarded asone of the most informative metrics for binary and multi-class classification in imbalanced conditions, is another frequently suggested metric. In order to assess prediction quality, MCC simultaneously considers true positives, true negatives, false positives, and false negatives. MCC is especially useful for evaluating various detection architectures, such as transformer-based and GNN-based methods, in DDoS attack detection since it offers a single scalar value that represents a model's overall stability and dependability [117].

In addition to detection accuracy, resource-aware assessment measures are becoming increasingly important, particularly for IoT and edge computing environments, according to recent surveys and research [118]. Since many DDoS detection systems are implemented on devices with limited resources, energy cost and computational overhead have become important factors [119]. In real-time edge applications, models with great detection performance but significant energy consumption or inference latency might not be feasible. As a result, in addition to conventional classification measures, a number of recent studies have begun to report energy usage, inference time, or model complexity. In next-generation IoT and edge networks, this trend indicates a move toward comprehensive evaluation frameworks that strike a balance between deployment viability and detection efficacy.

Comparative Analysis of Evaluation Metrics

The nature of the DDoS detection problem and the datasets used have a significant impact on the wide variety of evaluation criteria employed in the evaluated studies. Although accuracy is often praised for its ease of use, in extremely unbalanced environments, where attack traffic is significantly outweighed by routine traffic, it can be deceptive. Accuracy alone runs the risk of ignoring subpar performance on less common attack types, underscoring the necessity for additional insightful criteria in assessment.

Precision and memory are frequently used to give a more insightful assessment. Precision measures a model's ability to reduce false positives, which is crucial in operational settings where security professionals may get overburdened by too many warnings. On the other hand, recall measures the model's capacity to identify real attacks and is frequently given priority in high-risk networks or important infrastructure, where it might be expensive to overlook a DDoS incident. Recall is usually seen as more important than precision in real-time detection settings, particularly in large-scale or high-impact attacks.

Because it offers a fair assessment of performance, particularly in situations when there is a class imbalance, the F1-score—which is the harmonic mean of precision and recall—has become more and more

popular in recent research. As such, it is commonly reported in benchmark evaluations with datasets such as UNSW-NB15, NSL-KDD, and CICDDoS2019. Practical factors like tolerance for false alarms, the need for real-time response, and the possible cost of missed detections should all be taken into account when selecting assessment metrics for real-world deployments. The necessity of interpreting results beyond just performance metrics is shown by this context-sensitive option.

8 Datasets Overview

For the purpose of assessing and contrasting graph- and transformer-based models for DDoS detection, a thorough grasp of datasets is necessary. The research community makes extensive use of a number of benchmark datasets, each with unique features, attack scenarios, and constraints. A selection of frequently used datasets is given in Table 9.

Table 9: Key characteristics of common DDoS detection datasets.

Dataset	Year	Type	Size/Samples	Features/Attributes	Strengths	Limitations
CIC-DDoS2019 [120]	2018	Network Traffic	12.6M flows	80+ flow-based features	Diverse attack types, realistic traffic patterns	Large size, requires preprocessing
CIC-DDoS2022 [121]	2022	Network Traffic	15M+ flows	90+ flow-based features	Updated attacks, IoT traffic inclusion	High imbalance among attack classes
TON_IoT [122]	2020	IoT/Industrial	2M+ flows	Network + system logs	Multi-protocol, realistic IoT scenarios	Smaller scale than network datasets
UNSW-IoT20 [99]	2015	IoT Traffic	1.5M flows	Network & payload features	Modern IoT devices, varied attack scenarios	Limited number of attack classes
NSL-KDD [123]	2009	Network Traffic	125,973 records	41 features	Widely used, well-structured	Outdated, does not reflect modern IoT/DDoS attacks

These datasets make it possible to evaluate models in various scenarios. While TON_IoT and UNSW-IoT20 offer insights into IoT-specific traffic patterns, CIC-DDoS files are useful for assessing high-volume attacks. Although NSL-KDD is still helpful for baseline comparison, its applicability to modern DDoS scenarios is limited by its antiquated attack types and simplified network conditions. Datasets should be carefully chosen by researchers based on their target network environment and model aims.

9 Conclusion

In this paper, we presented a comprehensive overview of graph-driven DL and Transformer-based strategies for DDoS attack detection. By systematically analyzing previous studies, we grouped the techniques into three main categories: Transformer-driven models, hybrid GNN-Transformer architectures, and pure

GNN-driven models. Our review highlights the strengths and limitations of each category, demonstrating how GNNs efficiently capture topological dependencies, Transformers excel in modeling long-range temporal patterns, and hybrid models integrate these capabilities to improve detection performance in complex network environments.

We also discussed the main challenges and open issues in this domain, such as data imbalance, model interpretability, scalability, dataset limitations, and deployment in resource-constrained or real-world settings. By identifying these gaps, we outlined key directions for future research, emphasizing lightweight frameworks, standardized benchmarking, federated learning, explainability, and robustness against adversarial attacks.

Overall, this paper serves as a structured and practical reference for security engineers, researchers, and practitioners aiming to develop next-generation DDoS detection systems. By presenting future research directions, a clear taxonomy, and comparative analyses, it contributes to a deeper understanding of how graph-driven and Transformer-based models can be effectively leveraged to mitigate increasingly critical cyber threats. The insights and perspectives provided here are intended to guide future research and support the development of scalable, efficient, and intelligent network security solutions.

Acknowledgement: Not applicable.

Funding Statement: The authors received no specific funding.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Noor Mueen Mohammed Ali Hay-der, Mehdi Ebady Manaa; methodology, Hamid Noori, Davood Zabihzadeh; software, Noor Mueen Mohammed Ali Hayder; validation, Noor Mueen Mohammed Ali Hayder; formal analysis, Noor Mueen Mohammed Ali Hayder; investigation, Noor Mueen Mohammed Ali Hayder; resources, Noor Mueen Mohammed Ali Hayder; data curation Noor Mueen Mohammed Ali Hayder; writing—original draft preparation, Noor Mueen Mohammed Ali Hayder; writing—review and editing, Seyed Amin Hosseini Seno; visualization, Seyed Amin Hosseini Seno; supervision, Seyed Amin Hosseini Seno; project administration, Noor Mueen Mohammed Ali Hayder; funding acquisition, Noor Mueen Mohammed Ali Hayder. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: Not applicable.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

DDoS	Distributed Denial-of-Service
GNNs	Graph Neural Networks
XGNN	Explainable graph neural networks
SHAP	SHapley Additive exPlanations
LIME	Local Interpretable Model-agnostic Explanations
LSTM	Long Short-Term Memory
CNNs	Convolutional Neural Networks
DL	Deep Learning
IOT	Internet of Things

References

1. Qasim SS, Nsaif SM. Advancements in time series-based detection systems for distributed denial-of-service (DDoS) attacks: a comprehensive review. Babylon J Netw. 2024;2024:9–17. doi:10.58496/bjn/2024/002.
2. Almutairi M, Sheldon FT. IoT-cloud integration security: a survey of challenges, solutions, and directions. Electronics. 2025;14(7):1394. doi:10.3390/electronics14071394.
3. Hnamte V, Ahmad Najar A, Nhung-Nguyen H, Hussain J, Sugali MN. DDoS attack detection and mitigation using deep neural network in SDN environment. Comput Secur. 2024;138(21):103661. doi:10.1016/j.cose.2023.103661.
4. Ponzi V, Napoli C. Graph neural networks: architectures, applications, and future directions. IEEE Access. 2025;13(43):62870–91. doi:10.1109/ACCESS.2025.3558752.
5. Bilot T, El Madhoun N, Al Agha K, Zouaoui A. Graph neural networks for intrusion detection: a survey. IEEE Access. 2023;11:49114–39. doi:10.1109/ACCESS.2023.3275789.
6. Kanca AM, Türker İ. A systematic review of graph-based representation techniques for cyber-attack detection across application domains. Concurr Comput Pract Exp. 2025;37(27–28):e70389. doi:10.1002/cpe.70389.
7. Altaf T, Wang X, Ni W, Yu G, Liu RP, Braun R. GNN-based network traffic analysis for the detection of sequential attacks in IoT. Electronics. 2024;13(12):2274. doi:10.3390/electronics13122274.
8. Shehzad A, Xia F, Abid S, Peng C, Yu S, Zhang D, et al. Graph transformers: a survey. arXiv:2407.09777. 2024.
9. Al-Shareeda MA, Manickam S, Ali Saare M. DDoS attacks detection using machine learning and deep learning techniques: analysis and comparison. Bulletin EEI. 2023;12(2):930–9. doi:10.11591/eei.v12i2.4466.
10. Mittal M, Kumar K, Behal S. Deep learning approaches for detecting DDoS attacks: a systematic review. Soft Comput. 2022;27(18):1–37. doi:10.1007/s00500-021-06608-1.
11. Kumar D, Pateriya RK, Gupta RK, Dehalwar V, Sharma A. DDoS detection using deep learning. Procedia Comput Sci. 2023;218(1):2420–9. doi:10.1016/j.procs.2023.01.217.
12. Ali TE, Chong YW, Manickam S. Machine learning techniques to detect a DDoS attack in SDN: a systematic review. Appl Sci. 2023;13(5):3183. doi:10.3390/app13053183.
13. Zolfagharipour L, Kadhim MH. A technique for efficiently controlling centralized data congestion in vehicular *ad hoc* networks. Int J Comput Netw Appl. 2024;12(2):267–77. doi:10.22247/ijcna/2025/17.
14. Likhar P, Gupta SK, Choudhary J, Singh DP. Delving deep: DDoS attack resilience through deep learning approaches. Knowl Inf Syst. 2025;68(1):19. doi:10.1007/s10115-025-02651-8.
15. Sankara Vadivel SR, Karthikeyan V, Gopalakrishnan K, Dani Reagan Vivek J. Introduction to the Internet of Things (IoT). In: Internet of Things security. Amsterdam, The Netherlands: Elsevier; 2026. p. 3–31. doi:10.1016/b978-0-44-333759-8.00010-1.
16. Adedeji KB, Abu-Mahfouz AM, Kurien AM. DDoS attack and detection methods in Internet-enabled networks: concept, research perspectives, and challenges. J Sens Actuator Netw. 2023;12(4):51. doi:10.3390/jsan12040051.
17. Abiramasundari S, Ramaswamy V. Distributed denial-of-service (DDOS) attack detection using supervised machine learning algorithms. Sci Rep. 2025;15(1):13098. doi:10.1038/s41598-024-84879-y.
18. Sivaroopan N, Silva K, Madarasingha C, Dahanayaka T, Jourjon G, Jayasumana A, et al. A comprehensive survey on network traffic synthesis: from statistical models to deep learning. arXiv:2507.01976. 2025.
19. Zhang W, Lazaro JP. A survey on network security traffic analysis and anomaly detection techniques. IJETAA. 2024;1(4):8–16. doi:10.62677/ijetaa.2404117.
20. Georgousis S, Kenning MP, Xie X. Graph deep learning: state of the art and challenges. IEEE Access. 2021;9:22106–40. doi:10.1109/ACCESS.2021.3055280.
21. Vrahatis AG, Lazaros K, Kotsiantis S. Graph attention networks: a comprehensive review of methods and applications. Future Internet. 2024;16(9):318. doi:10.3390/fi16090318.
22. Huang Z, Yi B. Transformer-based large-scale and intelligent network traffic prediction and optimization. Trans Emerging Tel Tech. 2026;37(1):e70314. doi:10.1002/ett.70314.
23. Yang J, Liu Z, Xiao S, Li C, Lian D, Agrawal S, et al. Graphformers: gNN-nested transformers for representation learning on textual graph. Adv Neural Inf Process Syst. 2021;34:28798–810.

24. Hu C, Wang Y, Zhang X, Zheng M, Feng G, liu J, et al. Dynamic graph neural network-transformer-LSTM based multi-scale spatio-temporal traffic forecasting. Int J Intell Transp Syst Res. 2026;90(3):166. doi:10.1007/s13177-025-00623-4.
25. Xiao Y, Ma Z, Huang W, Qiao C, Zhao B, Zhang D, et al. Pure-GNN: a lightweight purified graph neural network against adversarial attacks. Tsinghua Sci Technol. 2025. doi:10.26599/tst.2025.9010034.
26. Le Duc L, Nguyen Phan Hai P, Hoang T. A novel packet-based preprocessing approach forTransformer models toEnhance DDoS detection accuracy. In: Industrial networks and intelligent systems. Berlin/Heidelberg, Germany: Springer; 2026. p. 246–58. doi:10.1007/978-3-032-02362-9_19.
27. Mutembei LL, Senekane MC, van Zyl T. Deep learning-based network intrusion detection systems: a systematic literature review. In: Artificial intelligence research. Berlin/Heidelberg, Germany: Springer; 2024. p. 207–34. doi:10.1007/978-3-031-78255-8_13.
28. Wasswa H, Abbass HA, Lynar T. ResDNViT: a hybrid architecture for Netflow-based attack detection using a residual dense network and Vision Transformer. Expert Syst Appl. 2025;282(12):127504. doi:10.1016/j.eswa.2025.127504.
29. Yang J, Jiang X, Lei Y, Liang W, Ma Z, Li S. MTSecurity: privacy-preserving malicious traffic classification using graph neural network and transformer. IEEE Trans Netw Serv Manag. 2024;21(3):3583–97. doi:10.1109/TNSM.2024.3383851.
30. Vitulyova Y, Babenko T, Kolesnikova K, Kiktev N, Abramkina O. A hybrid approach using graph neural networks and LSTM for attack vector reconstruction. Computers. 2025;14(8):301. doi:10.3390/computers14080301.
31. Guo W, Qiu H, Liu Z, Zhu J, Wang Q. GLD-net: deep learning to detect DDoS attack via topological and traffic feature fusion. Comput Intell Neurosci. 2022;2022(1):4611331–20. doi:10.1155/2022/4611331.
32. Wang Y, Han Z, Du Y, Li J, He X. BS-GAT: a network intrusion detection system based on graph neural network for edge computing. Cybersecurity. 2025;8(1):27. doi:10.1186/s42400-024-00296-8.
33. Barsellotti L, De Marinis L, Cugini F, Paolucci F. FTG-net: hierarchical flow-to-traffic graph neural network for DDoS attack detection. In: Proceedings of the 2023 IEEE 24th International Conference on High Performance Switching and Routing (HPSR); 2023 Jun 5–7; Albuquerque, NM, USA. doi:10.1109/HPSR57248.2023.10147929.
34. Mohan HG, Kumar J. BotMHG: a hybrid deep learning-based graphical approach to detect botnets using graph neural networks and graph attention networks on topological and temporal features. Neural Comput Appl. 2025;37(23):19303–37. doi:10.1007/s00521-025-11402-3.
35. El Gadal W, Ganti S. Federated secure intelligent intrusion detection and mitigation framework for SD-IoT networks using ViT-GraphSAGE and automated attack reporting. In: Proceedings of the 2025 12th IFIP International Conference on New Technologies, Mobility and Security (NTMS); 2025 Jun 18–20; Paris, France. doi:10.1109/NTMS65597.2025.11076682.
36. Le HD, Park M. Enhancing multi-class attack detection in graph neural network through feature rearrangement. Electronics. 2024;13(12):2404. doi:10.3390/electronics13122404.
37. Saxena S, Grover J, Singhal S. Exploring graph neural networks for robust network intrusion detection. Procedia Comput Sci. 2025;258(10):3630–9. doi:10.1016/j.procs.2025.04.618.
38. Ran L, Cui Y, Zhao J, Yang H. TITAN: combining a bidirectional forwarding graph and GCN to detect saturation attack targeted at SDN. PLoS One. 2024;19(4):e0299846. doi:10.1371/journal.pone.0299846.
39. Hekmati A, Krishnamachari B. Graph-based DDoS attack detection in IoT systems with lossy network. arXiv:2403.09118. 2024.
40. Saidane A, El Kamel A, Youssef H. CHC-DDoS: a DDoS attacks detection scheme using host-connection graph representation and GCN. In: Advanced information networking and applications. Berlin/Heidelberg, Germany: Springer; 2025. p. 317–25. doi:10.1007/978-3-031-87784-1_29.
41. Hekmati A, Krishnamachari B. Graph convolutional networks for DDoS attack detection in a lossy network. In: Proceedings of the 2024 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN); 2024 May 5–8; Stockholm, Sweden. doi:10.1109/ICMLCN59089.2024.10624757.

42. Nagaraj K, Starke A, McNair J. GLASS: a graph learning approach for software defined network based smart grid DDoS security. In: Proceedings of the ICC 2021-IEEE International Conference on Communications; 2021 Jun 14–23; Montreal, QC, Canada. doi:10.1109/icc42927.2021.9500999.
43. Wang J, Wang L. LR-STGCN: detecting and mitigating low-rate DDoS attacks in SDN based on spatial-temporal graph neural network. Comput Secur. 2025;154(3):104460. doi:10.1016/j.cose.2025.104460.
44. Saunders BJ, Kisanga P, Carvalho GHS, Woungang I. A graph convolutional networks-based DDoS detection model. In: Proceedings of the 2024 IEEE International Systems Conference (SysCon); 2024 Apr 15–18; Montreal, QC, Canada. doi:10.1109/SysCon61195.2024.10553611.
45. Manjula HS, Roopa MS, Arunalatha JS, Venugopal KR. Intrusion detection model for IoT networks using graph convolution networks (GCN). In: ICT for intelligent systems. Berlin/Heidelberg, Germany: Springer; 2023. p. 1–12. doi:10.1007/978-981-99-3982-4_1.
46. Li Y, Zhou Z, Li R, Shi F, Guo J, Liu Q. GoGDDoS: a multi-classifier for DDoS attacks using graph neural networks. In: Proceedings of the 2023 IEEE Symposium on Computers and Communications (ISCC); 2023 Jul 9–12; Gammarth, Tunisia. doi:10.1109/ISCC58397.2023.10218316.
47. Xu H, Geng X, Liu J, Lu Z, Jiang B, Liu Y. A novel approach for detecting malicious hosts based on RE-GCN in intranet. Cybersecurity. 2024;7(1):69. doi:10.1186/s42400-024-00242-8.
48. Cao Y, Jiang H, Deng Y, Wu J, Zhou P, Luo W. Detecting and mitigating DDoS attacks in SDN using spatial-temporal graph convolutional network. IEEE Trans Dependable Secure Comput. 2022;19(6):3855–72. doi:10.1109/TDSC.2021.3108782.
49. Li K, Zhou H, Tu Z, Ouyang L, Zhang H. AT-GCN: a DDoS attack path tracing system based on attack traceability knowledge base and GCN. Comput Netw. 2023;236(5):110036. doi:10.1016/j.comnet.2023.110036.
50. Saunders BJ, de Grande RE, Carvalho GHS, Woungang I. Deep graph learning for DDoS detection and multi-class classification IDS. In: Proceedings of the 2024 IEEE International Conference on Cyber Security and Resilience (CSR); 2024 Sep 2–4; London, UK. doi:10.1109/CSR61664.2024.10679447.
51. Galli D, Venturi A, Stabili D, Andreolini M, Marchetti M. Defending network intrusion detection systems based on graph neural networks against structural adversarial attacks. In: Proceedings of the 2025 23rd International Symposium on Network Computing and Applications (NCA); 2025 Nov 5–7; Lisbon, Portugal. doi:10.1109/NCA67271.2025.00043.
52. Venturi A, Stabili D, Marchetti M. Problem space structural adversarial attacks for Network Intrusion Detection Systems based on Graph Neural Networks. arXiv:2403.11830. 2024.
53. Abu Bakar R, De Marinis L, Cugini F, Paolucci F. FTG-Net-E: a hierarchical ensemble graph neural network for DDoS attack detection. Comput Netw. 2024;250(2):110508. doi:10.1016/j.comnet.2024.110508.
54. Anjum M, Dutta AK, Elrashidi A, Shahab S, Aldrees A, Shaikh ZA, et al. GraphFedAI framework for DDoS attack detection in IoT systems using federated learning and graph based artificial intelligence. Sci Rep. 2025;15(1):28050. doi:10.1038/s41598-025-10826-0.
55. Holmkvist Bergqvist E. Graph Neural Network for early DDoS detection: evaluating the impact of progressive node reduction in flow graphs [bachelor's thesis]. Linköping, Sweden: Linköping University; 2025.
56. Sanjalawe Y, Al-E'mari S, Alqurashi T, Alharbi ZH, Makhadmeh SN, Alsharaiah M. Adaptive graph attention-based federated learning for IoT intrusion detection: mitigating poisoning attacks. PeerJ Comput Sci. 2025;11(24):e3281. doi:10.7717/peerj-cs.3281.
57. Wang H, Li W. DDosTC: a transformer-based network attack detection hybrid mechanism in SDN. Sensors. 2021;21(15):5047. doi:10.3390/s21155047.
58. Dey S, Kate PS, Upadhyay V, Vaish A. A transformer-based approach for DDoS attack detection in IoT networks. arXiv:2508.10636. 2025.
59. Harshdeep K, Sumalatha K, Mathur R. DeepTransIDS: transformer-based deep learning model for detecting DDoS attacks on 5G NIDD. Results Eng. 2025;26(16):104826. doi:10.1016/j.rineng.2025.104826.
60. Li Y, Deng X, Yang A, Gao J. A transformer-based framework for DDoS attack detection via temporal dependency and behavioral pattern modeling. Algorithms. 2025;18(10):628. doi:10.3390/a18100628.

61. Al-Haboosi IT, Elbagoury BM, El-Regaily S, El-Horbaty EM. A hybrid-transformer-based cyber-attack detection in IoT networks. Int J Interact Mob Technol. 2024;18(14):90–102. doi:10.3991/ijim.v18i14.50343.
62. Koukoulis I, Syrigos I, Korakis T. Self-supervised transformer-based contrastive learning for intrusion detection systems. arXiv:2505.08816. 2025.
63. Wathan MH, Irawan I, Swengky B, Zain MS, Ramadani A, Riadi S. TransDDoS: transformer-based model for intelligent detection of DDoS attacks. IJICOM. 2025;7(1):243–53. doi:10.35842/ijicom.v7i1.131.
64. Huang Z, Liu S, Zhao K, Xiang Y. TDAT: a real-time two-stage DDoS attacks detector based on anomaly transformer. In: Neural information processing. Berlin/Heidelberg, Germany: Springer; 2025. p. 61–75. doi:10.1007/978-981-96-6591-4_5.
65. Aleyead S, Al-Ahmadi S. A transformer and federated learning techniques for detecting DDoS attacks in IoT environments. In: Proceedings of Ninth International Congress on Information and Communication Technology; 2024 Feb 19–22; London, UK. doi:10.1007/978-981-97-3559-4_1.
66. Sangore RB, Patil ME. DDoS attack detection in blockchain network layer using dual attention based dense convolutional gated recurrent unit. Int J Inf Comput Secur. 2025;27(4):437–64. doi:10.1504/ijics.2025.148108.
67. Al Faiyaz Provi N, Chowdhury MZ, Morshed S, Alam MZ. Transformer-based intrusion detection for securing medical applications in 5G IoMT networks. In: Proceedings of the 2025 2nd International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM); 2025 Jun 27–28; Gazipur, Bangladesh. doi:10.1109/NCIM65934.2025.11160217.
68. Bhatt R, Indra G. Leveraging GANs for adaptive network intrusion detection. In: Proceedings of Fourth International Conference on Computing and Communication Networks; 2024 Oct 17–18; Manchester, UK. doi:10.1007/978-981-96-3244-2_52.
69. Jamshaid A, Ali SH. Leveraging transformer-based GANs to improve intrusion detection in networks. In: Proceedings of the International Data Science and Information Technologies Congress (INFTEC 2025); 2025 May 22–23; Budapest, Turkey.
70. Wang P, Miao X. DDoS attach detection technology based on CNN and transformer. In: Proceedings of the Third International Conference on Green Communication, Network, and Internet of Things (CNIoT 2023); 2023 Jul 14–16; Chongqing, China. doi:10.1117/12.3010685.
71. Pawar PP, Kumar D, Ananthan B, Pradeepa AS, Selvi AS. An efficient DDoS attack detection using attention based hybrid model in blockchain based SDN-IoT. In: Proceedings of the 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT); 2024 May 3–4; Vellore, India. doi:10.1109/AIIoT58432.2024.10574596.
72. Priyadarshini I, Mohanty P, Alkhayyat A, Sharma R, Kumar S. SDN and application layer DDoS attacks detection in IoT devices by attention-based Bi-LSTM-CNN. Trans Emerg Telecommun Technol. 2023;34(11):e4758. doi:10.1002/ett.4758.
73. Alshdadi AA, Ali Almazroi A, Alsolami E, Ayub N, Lytras MD. Enhanced IoT security for DDOS attack detection: split attention-based ResNeXt-GRU ensembler approach. IEEE Access. 2024;12:112368–80. doi:10.1109/ACCESS.2024.3443067.
74. Al-Jarah M, Al-Shurman M. Attention-based deep learning approach for detecting IoT botnet-based distributed denial of service attacks. J Discrete Math Sci Cryptogr. 2024;27(6):1785–815. doi:10.47974/jdmsc-1729.
75. Liu Z, Jiang K, Yan J. DDoS attack detection method using entropy and attention-based BiGRU. In: Proceedings of the 2023 International Conference on Human-Centered Cognitive Systems (HCCS); 2023 Dec 16–17; Cardiff, UK. doi:10.1109/HCCS59561.2023.10452473.
76. Al-Absi GA, Fang Y, Qaseem AA, Al-Absi H. DST-IDS: dynamic spatial-temporal graph-transformer network for in-vehicle network intrusion detection system. Veh Commun. 2025;55(11):100962. doi:10.1016/j.vehcom.2025.100962.
77. Nawaz G, Junaid M, Akhunzada A, Gani A, Nawazish S, Yaqub A, et al. Detecting and mitigating DDOS attacks in SDNs using deep neural network. Comput Mater Contin. 2023;77(2):2157–78. doi:10.32604/cmc.2023.026952.
78. Wang J. Multivariate time series forecasting and classification via GNN and transformer models. J Comput Technol Softw. 2024;3(9):1–6. doi:10.5281/zenodo.14789085.

79. Zhang H, Cao T. A hybrid approach to network intrusion detection based on graph neural networks and transformer architectures. In: Proceedings of the 2024 14th International Conference on Information Science and Technology (ICIST); 2024 Dec 6-9;. Chengdu, China. doi:10.1109/ICIST63249.2024.10805457.
80. Arindam A. Advancing network security through deep learning: a hybrid graph-based and temporal approach to anomaly and threat detection. Int J Res Appl Sci Eng Technol. 2025;13(5):6095–103. doi:10.22214/ijraset.2025.71415.
81. Anoop M, Mary LW, Wilson AJ, Kiran WS. Optimized graph transformer with molecule attention network based multi class attack detection framework for enhancing privacy and security in WSN. Multimed Tools Appl. 2025;84(15):14273–304. doi:10.1007/s11042-024-19516-x.
82. Govea J, Gutierrez R, Villegas-Ch W, Maldonado Navarro A. Hybrid AI for predictive cyber risk assessment: federated graph-transformer architecture with explainability. IEEE Access. 2025;13:122187–206. doi:10.1109/ACCESS.2025.3588076.
83. Zhang Y, Fan Y, Ma H, Wang B, Hou R, Cao J. A dual-stream network architecture based on GNN and CNN for intrusion detection. In: Advanced intelligent computing technology and applications. Berlin/Heidelberg, Germany: Springer; 2025. p. 426–37. doi:10.1007/978-981-96-9911-7_35.
84. Ghadermazi J, Hore S, Shah A, Bastian ND. GTAE-IDS: graph transformer-based autoencoder framework for real-time network intrusion detection. IEEE Trans Inf Forensics Secur. 2025;20(43):4026–41. doi:10.1109/TIFS.2025.3557741.
85. Govindarajan V, Muzamal JH. Advanced cloud intrusion detection framework using graph based features transformers and contrastive learning. Sci Rep. 2025;15(1):20511. doi:10.1038/s41598-025-07956-w.
86. Hayder NM, Seno SA, Noori H, Zabihzadeh D, Manaa ME. Improved DDoS attack detection-based feature selection by using graph convolutional network transformer model. Oper Res Eng Sci Theory Appl. 2025;8(2):22–46. doi:10.5281/zenodo.17160174.
87. Lakshmanan M, Adnan MM, Reddy RA, Vasukidevi G, Aarthy G. A graph neural network and transformer encoder technique for anomaly and cyber threat detection in smart grids. In: Proceedings of the 2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS); 2024 Aug 23–24; Hassan, India. doi:10.1109/IACIS61494.2024.10721753.
88. Guduru J, Priyanka S. A hybrid learning approach combining graph and transformer models for communication efficient distributed control and estimation in networked systems. J Theor Appl Inf Technol. 2025;103(13):4705–23. doi:10.1109/lcsys.2021.3071478.
89. Zhang J, Fan X, Zhao Z. A hybrid intrusion detection model based on dynamic spatial-temporal graph neural network in in-vehicle networks. Sci Rep. 2025;15(1):34736. doi:10.1038/s41598-025-18401-3.
90. Wasswa H, Abbass H, Lynar T. Are GNNs worth the effort for IoT botnet detection? A comparative study of VAE-GNN vs. ViT-MLP and VAE-MLP approaches, arXiv:2505.17363. 2025..
91. Mortatha Alkorani MB, Nuiaa Alogaili RR, Abdulsaeed AA, Dashoor ZA, Alkareem Alyasseri ZA, Alsaeedi AH, et al. OptiGuard-GNN cybersecurity model: leveraging multi-criteria optimization and graph neural networks for enhanced detection of distributed denial of service attacks. Int J Intell Eng Syst. 2025;18(7):792–809. doi:10.22266/ijies2025.0831.50.
92. Sun N, Chen L, Han G. HADGA: hierarchical attention-based dynamic GNN algorithm for IoT botnet detection. IEEE Internet Things J. 2025;12(16):33520–32. doi:10.1109/JIOT.2025.3576710.
93. Bagha AM, Woungang I, Traore I, Rawat DB, Tanwar S, Hassanalizadeh A. Network anomaly detection system using an attention-based GNN. In: Pan-African artificial intelligence and smart systems. Berlin/Heidelberg, Germany: Springer; 2025. p. 164–76. doi:10.1007/978-3-031-94439-0_10.
94. Wu J, Qiu G, Wu C, Jiang W, Jin J. Federated learning for network attack detection using attention-based graph neural networks. Sci Rep. 2024;14(1):19088. doi:10.1038/s41598-024-70032-2.
95. Zhu Q, Zhan X, Chen W, Li Y, Ouyang H, Jiang T, et al. GATransformer: a network threat detection method based on graph-sequence enhanced transformer. Electronics. 2025;14(19):3807. doi:10.3390/electronics14193807.
96. Wang Y, Li J, Han Z, Cheng P, Kumar R. FedSTGCN: a novel federated spatiotemporal graph learning-based network intrusion detection method for the Internet of Things. Front Inf Technol Electron Eng. 2025;26(7):1164–79. doi:10.1631/FITEE.2400932.

97. Qaddos A, Yaseen MU, Al-Shamayleh AS, Imran M, Akhunzada A, Alharthi SZ. A novel intrusion detection framework for optimizing IoT security. Sci Rep. 2024;14(1):21789. doi:10.1038/s41598-024-72049-z.
98. Kumar A. Enhancing DDoS attack detection with multi-layer perceptron algorithms: a machine learning approach using the CICDDoS2019 dataset. J Electr Syst. 2024;20(3):5065–74. doi:10.21203/rs.3.rs-7422019/v1.
99. Moustafa N, Slay J. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS); 2015 Nov 10–12; Canberra, ACT, Australia. doi:10.1109/MilCIS.2015.7348942.
100. Qing Y, Liu X, Du Y. Mitigating data imbalance to improve the generalizability in IoT DDoS detection tasks. J Supercomput. 2024;80(7):9935–60. doi:10.1007/s11227-023-05829-5.
101. Zhou Q, Li R, Xu L, Nallanathan A, Yang J, Fu A. Towards interpretable machine-learning-based DDoS detection. SN Comput Sci. 2023;5(1):115. doi:10.1007/s42979-023-02383-y.
102. Moustafa N, Koroniotis N, Keshk M, Zomaya AY, Tari Z. Explainable intrusion detection for cyber defences in the Internet of Things: opportunities and solutions. IEEE Commun Surv Tutor. 2023;25(3):1775–807. doi:10.1109/COMST.2023.3280465.
103. Almheiri SJ, Ali Shah A, Abbas S, Ahmad M, Khan MA. Smart sustainable cyber security: modelling an interpretable and transparent threat detection with explainable artificial intelligence. Discov Sustain. 2025;6(1):442. doi:10.1007/s43621-025-01280-z.
104. Hoenig A, Roy K, Acquaah YT, Yi S, Desai SS. Explainable AI for cyber-physical systems: issues and challenges. IEEE Access. 2024;12(3):73113–40. doi:10.1109/ACCESS.2024.3395444.
105. Xin R, Wang J, Chen P, Zhao Z. Trustworthy AI-based performance diagnosis systems for cloud applications: a review. ACM Comput Surv. 2025;57(5):1–37. doi:10.1145/3701740.
106. Ying R, Bourgeois D, You J, Zitnik M, Leskovec J. GNNExplainer: generating explanations for graph neural networks. Adv Neural Inf Process Syst. 2019;32:9240–51.
107. Huang Q, Yamada M, Tian Y, Singh D, Chang Y. GraphLIME: local interpretable model explanations for graph neural networks. IEEE Trans Knowl Data Eng. 2023;35(7):6968–72. doi:10.1109/TKDE.2022.3187455.
108. Luo D, Cheng W, Xu D, Yu W, Zong B, Chen H, et al. Parameterized explainer for graph neural network. Adv Neural Inf Process Syst. 2020;33(6):19620–31. doi:10.1145/3729224.
109. Arafat NA, Basu D, Gel Y, Chen Y. When witnesses defend: a witness graph topological layer for adversarial graph learning. Proc AAAI Conf Artif Intell. 2025;39(15):15408–16. doi:10.1609/aaai.v39i15.33691.
110. Hamdan S, Ayyash M, Almajali S. Edge-computing architectures for Internet of Things applications: a survey. Sensors. 2020;20(22):6441. doi:10.3390/s20226441.
111. Yin X, Zhu Y, Hu J. A comprehensive survey of privacy-preserving federated learning: a taxonomy, review, and future directions. ACM Comput Surv. 2022;54(6):1–36. doi:10.1145/3460427.
112. Nandan M, Mitra S, De D. GraphXAI: a survey of graph neural networks (GNNs) for explainable AI (XAI). Neural Comput Appl. 2025;37(17):10949–1000. doi:10.1007/s00521-025-11054-3.
113. Cordero CG, Vasilomanolakis E, Wainakh A, Mühlhäuser M, Nadjm-Tehrani S. On generating network traffic datasets with synthetic attacks for intrusion detection. ACM Trans Priv Secur. 2021;24(2):1–39. doi:10.1145/3424155.
114. Muhammad A, Bae SH. A survey on efficient methods for adversarial robustness. IEEE Access. 2022;10:118815–30. doi:10.1109/ACCESS.2022.3216291.
115. Ali Al-Shukaili N, Kiah MLM, Ahmedy I. Optimizing feature selection and deep learning techniques for precise detection of low-rate distributed denial of service (LDDoS) attack. Discov Internet Things. 2025;5(1):80. doi:10.1007/s43926-025-00182-w.
116. Sujon KM, Hassan R, Choi K, Samad MA. Accuracy, precision, recall, f1-score, or MCC? Empirical evidence from advanced statistics, ML, and XAI for evaluating business predictive models. J Big Data. 2025;12(1):268. doi:10.1186/s40537-025-01313-4.
117. Cajas Ordóñez SA, Samanta J, Suárez-Cetrulo AL, Carbajo RS. Intelligent edge computing and machine learning: a survey of optimization and applications. Future Internet. 2025;17(9):417. doi:10.3390/fi17090417.

118. Abualhassan A, Rashid I, Binbeshr F, Imam M. DDoS attack detection in IoT: a comparative resource and performance analysis of deep learning and machine learning models. IEEE Access. 2025;13(8):116529–47. doi:10.1109/ACCESS.2025.3583855.
119. Hasan MA, Eaman A, Hassan E. Efficient DDoS detection with minimal features: high accuracy using CIC-DDoS2019. Procedia Comput Sci. 2025;265(10):124–31. doi:10.1016/j.procs.2025.07.164.
120. Sharafaldin I, Habibi Lashkari A, Ghorbani AA. Toward Generating a new intrusion detection dataset and intrusion traffic characterization. In: Proceedings of the 4th International Conference on Information Systems Security and Privacy; 2018 Jan 22–24; Funchal, Madeira, Portugal. doi:10.5220/0006639801080116.
121. Mittal M, Kumar K, Behal S. DDoS-AT-2022: a distributed denial of service attack dataset for evaluating DDoS defense system. Proc Indian Natl Sci Acad. 2023;89(2):306–24. doi:10.1007/s43538-023-00159-9.
122. Tareq I, Elbagoury BM, El-Regaily S, El-Horbaty ESM. Analysis of ToN-IoT, UNW-NB15, and Edge-IIoT datasets using DL in cybersecurity for IoT. Appl Sci. 2022;12(19):9572. doi:10.3390/app12199572.
123. Tavallaee M, Bagheri E, Lu W, Ghorbani AA. A detailed analysis of the KDD CUP 99 data set. In: Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications; 2009 Jul 8–10; Ottawa, ON, Canada. doi:10.1109/cisda.2009.5356528.