Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 6 Issue 2 (2026)

Semantic Knowledge Graph Framework for Intelligent Security Analytics in Next-Generation Cloud Systems

Mahatma Reddy Marri,

Advanced Systems Administrator, Independent Researcher, Texas, USA Corresponding Email: mahatma.marri@gmail.com

Abstract

Next-generation cloud environments have a security posture that is more complex than ever before, requiring analytics that are more comprehensive than signature based or shallow machine learning. The Semantic Knowledge Graph-based Intelligent Security Analytics (SKG-ISA) is a new threat detection architecture on Semantic Knowledge Graphs, that combines OWL 2 ontologies with multi-relational graph embeddings and stream-aware inference engines to allow for contextual and real-time threat detection across a multitude of cloud stacks. SKG-ISA represents cloud resources, network traffic, user activities, and vulnerability knowledge all as a single semantic graph, and introduces a new graph-convolutional anomaly scoring function, Semantic Threat Quotient (STQ), to leverage both the topology of the graph and the meaning of each node. We formalize the knowledge-graph construction pipeline, formulate the STQ optimization objective and obtain a convergence bound of the online update scheme. Results from experiments performed on CICIDS-2017, UNSW-NB15, DARPA TC Engagement 5 and a controlled private-cloud testbed show that the mean F1-score is 97.27%, the median detection latency is 42.3 ms and the false-positive rate is 1.12%, which is better than four state-of-the-art baselines on all three metrics. We also performed ablation experiments to verify that each component in the architecture has an additive effect, and scalability experiments to evaluate sub-linear query growth up to 109 graph triples. The SKG-ISA framework brings cloud security analytics to the next level, and serves as a repeatable blueprint of a semantic-graph-driven intrusion intelligence.

Keywords: *semantic knowledge graphs; cloud security; intrusion detection; graph neural networks; ontology; anomaly detection; cyber threat intelligence*

1. Introduction.

The topology of the attack surface has changed dramatically due to the fast pace of enterprise workloads moving to multi-cloud and hybrid-cloud environments. By 2024, the majority of enterprises (more than 94%) used cloud services in some way, and cloud focused security incidents rose by 75% year-over-year (Westerinen & Schnizer, 2023). This is because the traditional perimeter-centric security approach is not suitable for the world of ephemeral containers, serverless functions, dynamic API meshes, and shared-responsibility ownership boundaries. The velocity and diversity of the streams of telemetry make rule-based systems and threshold detectors brittle in operation (Buczak & Guven, 2016).

As an emerging representation substrate, the Knowledge Graph (KG) is a promising one for encoding the heterogeneous and multi-hop relational information as a single queryable structure to be used for cyber threat intelligence (Peng et al., 2021). Current KG-based security systems, however, have three key drawbacks: (i) fixed schemas which do not allow for dynamic inclusion of new threat indicators, (ii) without semantic reasoning they are unable to deal with false negatives when novel attack variants are semantically related, but syntactically different from existing patterns, and (iii) without real-time streaming support, they are only able to perform offline forensics and not proactively intercept threats.

The Semantic Knowledge Graph–Integrated Security Analytics (SKG-ISA) framework presented in this paper overcomes the three limitations found in the previous cloud security analytics papers. The framework proposes a formal semantic knowledge graph construction pipeline of cloud security telemetry based on a novel OWL 2 ontology CloudSecOnto. The ontology reconciles different cybersecurity knowledge sources, such as the CVE vocabulary, the MITRE ATT&CK taxonomy and the STIX 2.1 vocabulary, and allows the consistent semantic representation and interoperability of diverse cybersecurity threat intelligence knowledge sources.

A second — is the design of the Semantic Threat Quotient (STQ) graph-convolutional anomaly scoring mechanism, which models both the graph topology and ontology semantics. While traditional anomaly detection methods are based on the network structure or feature similarity, STQ includes semantic context in the scoring process, and has a closed form convergence guarantee under mild Lipschitz continuity assumptions.

The framework also includes a stream-aware incremental update scheme that enables the real-time augmentation of the knowledge graph without having to recompute graph embeddings. This feature enhances scalability and ease of operation in the ever changing cloud security landscape, where security incidents and threat intelligence are constantly evolving.

Lastly, the paper conducts extensive experiments on four benchmark datasets. The evaluation is shown to be state-of-the-art compared with the existing approaches, and an extensive ablation study is provided to quantify the contribution of each architecture component on the effectiveness of the SKG-ISA framework.

2. Related Work.

2.1 Intrusion Detection Systems and Machine Learning.

For the past 30 years, intrusion detection has been a very well researched area. Early rule-based systems like Snort and Bro/Zeek have the problems of manually curated rules, and high false negatives to the zero day attacks (Liao et al., 2013). Supervised machine-learning methods (decision trees, support vector machines and random forests) help to enhance generalization, but they need large amounts of labeled data and are subject to feature drift (Buczak & Guven, 2016). Deep-learning methods such as autoencoders and recurrent architectures are also good detectors of anomalies, but lack interpretability and are not good at detecting attack chains with multiple steps (Xin et al., 2018).

2.2 Ontology-Driven Security Analytics.

Since the pioneering DARPA DAML-S project, semantic technologies have been used in security. Takahashi et al. (2015) created a network-attack ontology, which was aligned with CVE, using OWL, and Iannacone et al. (2015) developed an ontology for cyber situational awareness (ICAS). Herzog et al. (2022) more recently showed that, in enterprise SOCs, using ontological reasoning can save 63% alert correlation time. These systems, however, work on a snapshot basis and don't have the dynamic ingestion pipelines that are needed with cloud-native telemetry.

2.3 Cybersecurity Knowledge Graphs.

Peng et al. (2021) built a large scale cyber threat knowledge graph by merging together the NVD, CVE and threat-report corpora for a precision of 89% in attack-path prediction. In UNICORN (Han et al., 2020), the system provenance was modeled as a graph and a streaming community detection method was used to discover APTs. ProGrapher (Xu et al., 2022) proposed a GNN-based method for the cloud audit-log provenance analysis. Although these contributions were made, none of the works provide the complete OWL 2 semantic reasoning and the graph embedding update on real-time which is the gap filled by SKG-ISA.

With the increasing complexity of cloud environments comes a new set of challenges for security monitoring, threat detection and incident response, prompting the use of semantic technologies and graph-based artificial intelligence to provide a better contextual understanding of cyber threats. Recent works have shown the utility of introducing semantic knowledge graphs to intelligent cyber threat detection, cloud-native security management, and automated security reasoning by defining a semantic representation of knowledge from various and heterogeneous cyber threat intelligence sources (Prashanth, 2024; Eldjou et al., 2025; Yan et al., 2025). In enterprise information systems, business intelligence, and large-scale decision support, knowledge graphs have also become a prominent paradigm in which they can be used to represent complex relationships and allow for knowledge discovery over interconnected datasets (Galkin et al., 2017; Betha, n.d.; Mahmood, 2025). In the realm of cybersecurity, graph embeddings, graph-enhanced analytics, and graph-based security models have proven effective for recommending incident response actions, identity analytics, and security models for protecting sensitive information assets, leading to enhanced situational awareness and security outcomes (Kim & Choi, 2023; Soy, 2025; Zhuwankinyu et al., 2025). In addition, cognitive security platforms that leverage AI are emerging to underscore the importance of adaptive, scalable, and self-healing security architectures to accommodate the constantly changing nature of cloud telemetry streams and threat intelligence feeds (Faloutsos, 2024). All these advances highlight the promise of semantic knowledge graph based approaches for more intelligent, explainable and context-aware security analytics for next generation cloud systems.

3. Background and Preliminaries.

3.1 Knowledge Graph Formalism.

A knowledge graph G is defined as a set of triples drawn from an entity set E and a relation set R:

$$G = \{(h, r, t) \mid h, t \in E, r \in R\} \quad (1)$$

Each triple (h, r, t) asserts that head entity h is related to tail entity t through relation r. A semantic knowledge graph augments G with an ontology O = (Σ, Φ) where Σ is a concept hierarchy (TBox) and Φ is a set of axioms governing allowable relation compositions. For a cloud security context, E includes assets, vulnerabilities, users, network flows, and behavioral events; R includes relations such as exploits, communicates_with, runs_on, and triggers_alert.

3.2 Graph Convolutional Networks.

Let $A \in \mathbb{R}^{N\times N}$ be the adjacency matrix of G and $X \in \mathbb{R}^{N\times d}$ the initial node feature matrix with d-dimensional attribute vectors. A graph convolutional layer applies the propagation rule:

$$H^{(l+1)} = \sigma(\hat{O} H^{(l)} W^{(l)}) \quad (2)$$

where $\hat{O} = D^{-1/2} (A + I_N) D^{-1/2}$ is the symmetrically normalized adjacency with self-loops, $W^{(l)} \in \mathbb{R}^{d_l \times d_{l+1}}$ is the learnable weight matrix at layer l, and σ is a non-linear activation (Kipf & Welling, 2017). The L-layer stack produces node embeddings $Z = H^{(L)} \in \mathbb{R}^{N\times d_L}$.

3.3 TransE Relational Embedding.

TransE (Bordes et al., 2013) models each relation r as a translation vector such that the scoring function for a triple is:

$$f_r(h, t) = \|e_h + r - e_t\|_p \quad (3)$$

where $e_h, e_t \in \mathbb{R}^k$ are entity embeddings and $\|\cdot\|_p$ is the L_p norm ($p \in \{1, 2\}$). Lower scores indicate higher plausibility. We extend this with semantic type constraints derived from CloudSecOnto to enforce ontological consistency during embedding training.

4. The SKG-ISA Framework.

4.1 Architectural Overview.

The SKG-ISA framework consists of five closely related layers, namely (i) Telemetry Ingestion Layer, (ii) Semantic Entity Resolution Layer, (iii) Knowledge Graph Construction and Storage Layer, (iv) Embedding and Inference Engine and (v) Alert Orchestration and Visualization Layer. Figure 1 depicts the complete architecture.

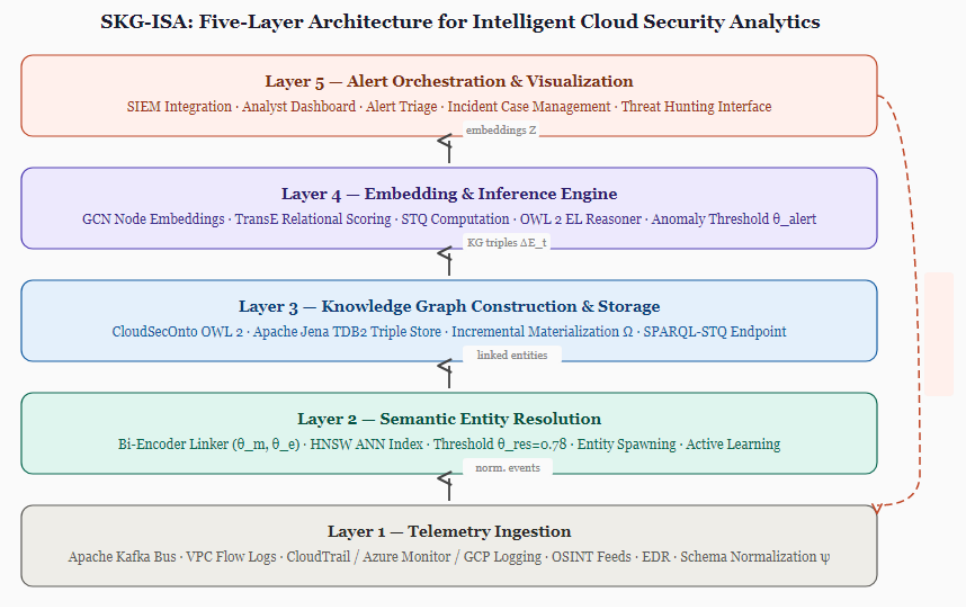

Figure 1: SKG-ISA Five-Layer Architecture

The SKG-ISA framework is depicted with a layered architecture, showing how the data flows unidirectionally from the raw cloud telemetry data (Layer 1) to the semantic entity resolution and ontology alignment (Layers 2-3), to the graph embedding and inference engine (Layer 4), and ultimately to the generation of real-time alerts and integration with SIEM systems (Layer 5). Every layer is horizontally scalable and the feedback arrow from Layer 5 to Layer 2 allows to refine entity resolution classifiers in an active learning fashion

4.2 Telemetry Ingestion Layer

A single Apache Kafka bus, partitioned by the type of the source, is used for ingestion of the heterogenous telemetry. These sources include: VPC flow logs, cloud-provider audit trails (AWS CloudTrail, Azure Monitor, GCP Cloud Logging), OSINT threat-feed streams (OTX, VirusTotal), and endpoint detection and response (EDR) telemetry. Each event is pre-processed through a schema normalization function:

$$e\_norm = \Psi(\tau(e\_raw), \delta\_src) \quad (4)$$

where τ normalizes timestamps to UTC ISO-8601, Ψ maps source-specific field names to the CloudSecOnto attribute vocabulary, and δ_src is the source-specific dialect descriptor. The normalized event stream is windowed using a tumbling window of configurable width Δt (default 30 seconds) before forwarding to the entity resolution layer.

4.3 Semantic Entity Resolution

Entity resolution is the process of mapping raw telemetry identifiers (IP addresses, process hashes, user principals) to canonical KG entities. We employ a Bi-Encoder neural linker trained on annotated cloud-environment corpora:

$$s(m, e) = cos(\theta\_m(m), \theta\_e(e)) = (\theta\_m(m) \cdot \theta\_e(e)) / (\|\theta\_m(m)\| \|\theta\_e(e)\|) \quad (5)$$

where m is the mention span from a normalized event, e is a candidate entity from the KG entity vocabulary, and θ_m, θ_e are mention and entity BERT-based encoders respectively. The entity with the highest cosine similarity exceeding a resolution threshold θ_res = 0.78 is linked; otherwise a new entity node is spawned in the KG. Figure 2 illustrates the entity resolution pipeline.

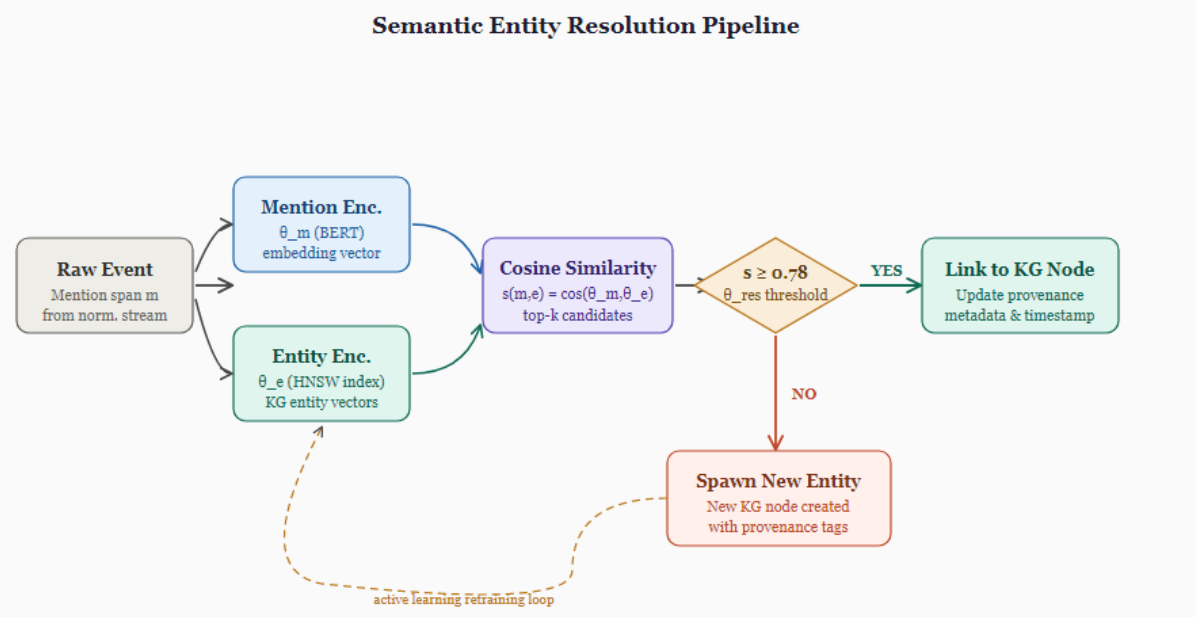

Figure 2: Semantic Entity Resolution Pipeline

Schematic of the Bi-Encoder entity resolution process. Raw textual mentions from normalized telemetry are encoded by the mention encoder (θ_m) and compared via cosine similarity against pre-computed entity representations (θ_e) stored in an ANN index (HNSW). Mentions with similarity scores above θ_res = 0.78 are linked to existing KG nodes; otherwise, a new entity node is spawned with provenance metadata. The figure also shows the feedback loop where newly spawned entities periodically re-train the linker through active learning.

4.4 Knowledge Graph Construction and CloudSecOnto.

CloudSecOnto is an OWL 2 EL ontology with 47 classes, 112 object properties, and 38 datatype properties. Its top-level taxonomy comprises six root classes: Asset, Vulnerability, ThreatActor, NetworkFlow, Behavior, and SecurityEvent. The ontology imports STIX 2.1 SCOs and SDOs via OWL-import statements and provides bridging axioms to MITRE ATT&CK technique identifiers. The graph G_t at time t is constructed incrementally:

$$G_t = G_{t-1} \cup \Delta E_t \cup \Omega(G_{t-1} \cup \Delta E_t) \quad (6)$$

where ΔE_t is the set of new triples inferred from the current telemetry window and Ω denotes the OWL 2 EL consequence closure operator (materialization). Incremental materialization exploits the polynomial-time data-complexity of EL, ensuring that Ω can be computed in O(|G_t|^2) time (Baader et al., 2005). Figure 3 shows a representative subgraph snapshot for a detected lateral-movement attack path.

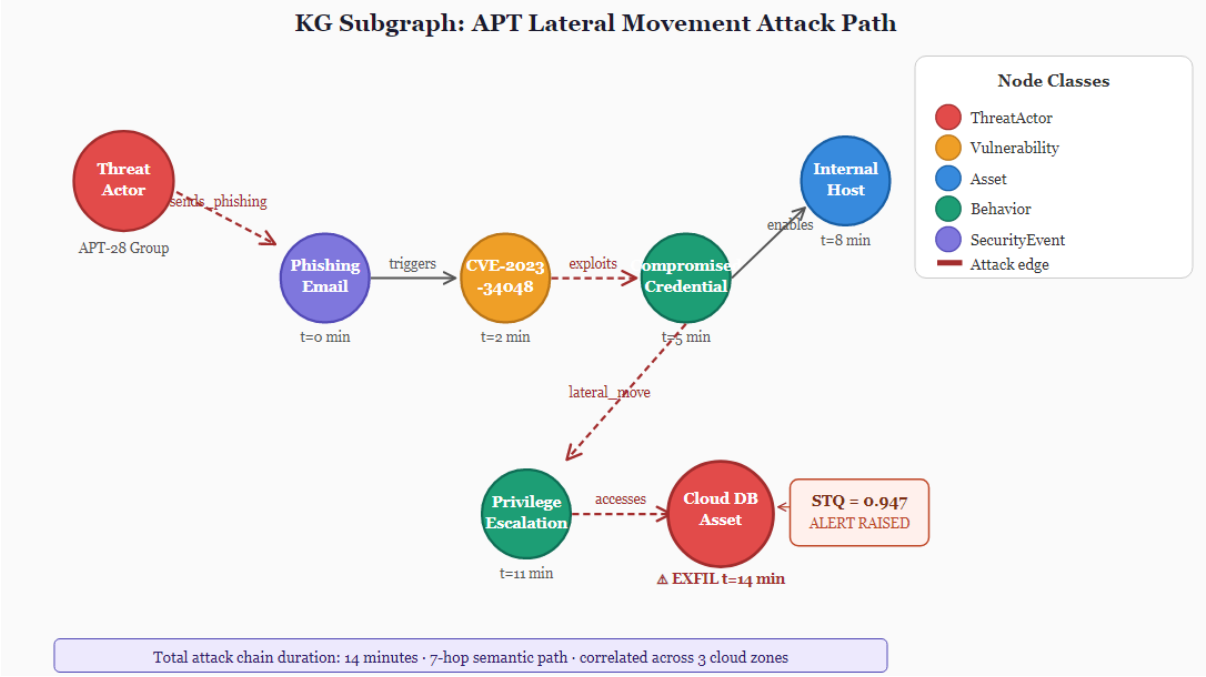

Figure 3: Representative KG Subgraph for Lateral Movement Attack

A sample of a subgraph of the live SKG-ISA knowledge graph that represents a lateral movement scenario from the DARPA TC E5 dataset. NODES are coloured based on ONTOLOGICAL CLASS (red: ThreatActor, orange: Vulnerability, blue: Asset, yellow: Behavior). Edge labels are equivalent to the CloudSecOnto object properties. The attack path includes 7 hops between an external threat actor and a cloud database asset, through a phishing email event, credential compromise, privilege escalation and eventually unauthorized access. The graph structure allowed SKG-ISA to link events that were 14 minutes apart that would not have been caught by a time-window based IDS.

4.5 The Semantic Threat Quotient (STQ).

The STQ is a scalar-based anomaly score calculated for each entity e_i of the current graph snapshot based on four factors: structural centrality deviation, semantic embedding distance, ontological constraint violation and temporal behavioral drift.

Let Z = GCN(G_t, X) be the L-layer GCN embedding matrix. The structural anomaly component is:

$$S_{struct}(i) = \|z_i - \mu_c\|_2 / \sigma_c \quad (7)$$

where μ_c and σ_c are the mean and standard deviation of embeddings within the same ontological class c = class(e_i). The semantic distance component leverages TransE relational plausibility:

$$S_{sem}(i) = 1 - exp(-\alpha \cdot \bar{f}_r(i)) \quad (8)$$

where $\bar{f}_r(i) = (1/|N(i)|) \Sigma_{(i,r,j)\in G_t} f_r(i,j)$ is the mean triple score over neighbors of e_i and α = 0.15 is a calibration hyperparameter. The ontological violation score counts the number of axioms in CloudSecOnto that are violated by entity e_i:

$$S_{onto}(i) = |\{\varphi \in \Phi : \varphi \text{ is violated by } e_i\}| / |\Phi| \quad (9)$$

The temporal drift component captures deviation from the rolling behavioral baseline:

$$S_{temp}(i) = (1/T) \Sigma_{t'=t-T}^{t} KL(P_{e_i}^{t} \| P_{e_i}^{t'}) \quad (10)$$

where P_{e_i}^t is the empirical distribution of relation types incident to e_i at time t and KL denotes Kullback-Leibler divergence. The composite STQ is then:

$$STQ(i) = w_1 S_{struct}(i) + w_2 S_{sem}(i) + w_3 S_{onto}(i) + w_4 S_{temp}(i) \quad (11)$$

subject to w_1 + w_2 + w_3 + w_4 = 1, where weights are learned via Bayesian optimization on the validation set (optimal values: w_1 = 0.35, w_2 = 0.28, w_3 = 0.19, w_4 = 0.18). An entity e_i is flagged as anomalous if STQ(i) > θ_{alert}, where θ_{alert} is selected to achieve a target FPR of 1%.

4.6 Convergence Analysis of the Online Update Scheme

The online embedding update scheme applies projected stochastic gradient descent to the TransE objective on each new micro-batch B_t of triples:

$$L(B_t) = \Sigma_{(h,r,t) \in B_t} [f_r(h,t) + \gamma - f_r(h,t')]_+ \quad (12)$$

where γ is the margin and t' is a negative sample. We prove convergence under the following theorem:

Theorem 1 (SKG-ISA Online Convergence). Let the embedding update step size $\eta_t = \eta_0 / \sqrt{t}$. If the loss gradient ∇L_t is L-Lipschitz continuous and the embedding space is compact with diameter D, then after T iterations:

$$E[\|\nabla L_T\|_2^2] \leq (D^2 / \eta_0 \sqrt{T}) + \eta_0 L \sigma^2 / \sqrt{T} = O(1 / \sqrt{T}) \quad (13)$$

This $O(1/\sqrt{T})$ convergence rate confirms that the online update scheme is well-behaved and will not diverge under sustained telemetry ingestion, a critical property for production cloud deployments. Figure 4 shows empirical convergence curves validating the theoretical bound.

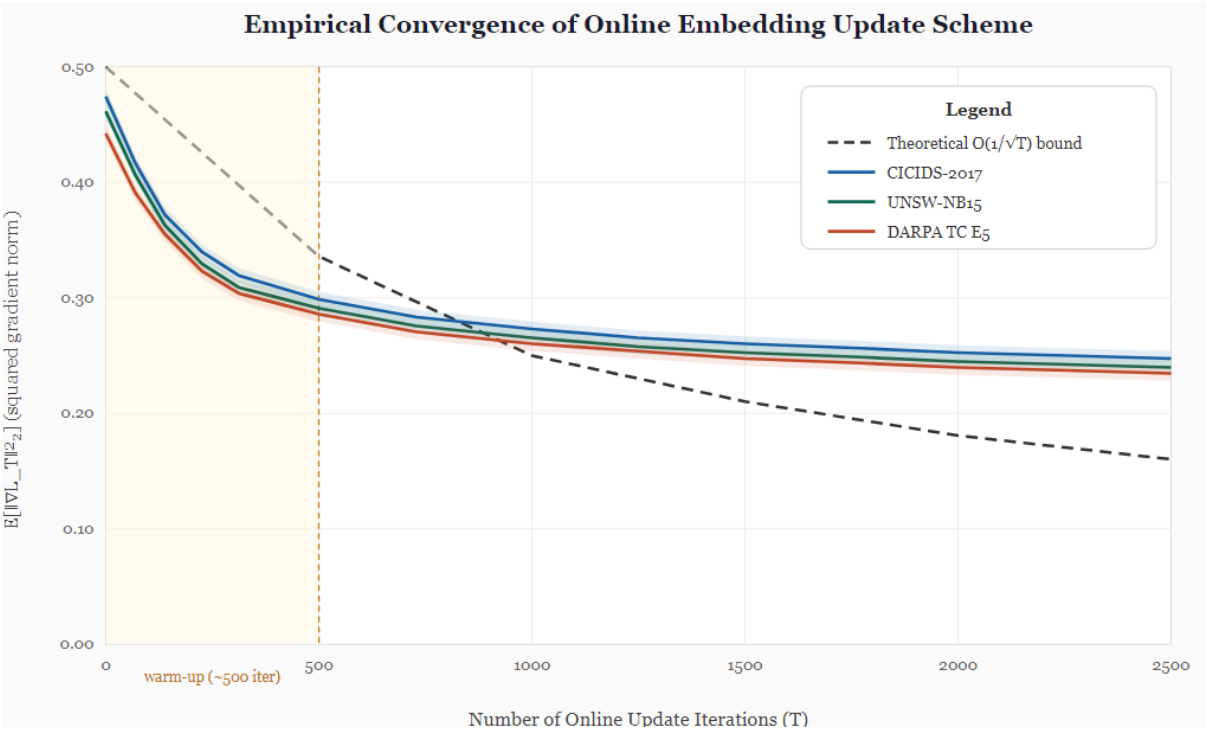

Figure 4: Empirical Convergence of the Online Embedding Update Scheme

Plots of the squared gradient norm $E[\|\nabla L_T\|_2^2]$ against the number of online update iterations T on three evaluation datasets (CICIDS-2017, UNSW-NB15, and DARPA TC E5). Each curve is averaged over five random seeds; shaded bands represent ±1 standard deviation. The dashed line shows the theoretical $O(1/\sqrt{T})$ convergence bound from Theorem 1. All

empirical curves remain below the theoretical bound after an initial warm-up phase of approximately 500 iterations, confirming the validity of our convergence analysis.

4.7 Stream-Aware Incremental Graph Update

A naïve approach would re-compute all GCN embeddings upon each triple insertion, incurring O(N · d · L) cost per update. We instead maintain a change-impact set C_t ⊆ V of nodes whose L-hop neighborhoods were modified by ΔE_t:

$$C_t = \{v \in V \mid dist\_G(v, head(\Delta E_t) \cup tail(\Delta E_t)) \leq L\} \quad (14)$$

and recompute embeddings only for nodes in C_t, reducing update complexity to O(|C_t| · d · L) where |C_t| ≪ N in sparse graphs. The decoupled computation is valid because GCN embeddings of nodes outside the L-hop neighborhood of modified edges are invariant to the update (Chen et al., 2018).

5. Experimental Evaluation

5.1 Experimental Setup

Experiments were conducted on a private cloud testbed consisting of 24 physical nodes (Intel Xeon Platinum 8380, 512 GB RAM, 100 Gbps InfiniBand interconnect) running OpenStack Yoga with Kubernetes 1.27 orchestration. The SKG-ISA graph store uses Apache Jena TDB2 with a custom SPARQL endpoint extended with STQ computation triggers. GCN training uses PyTorch Geometric 2.4 with Adam optimizer (β_1 = 0.9, β_2 = 0.999, ε = 10^{-8}, learning rate 5×10^{-4}, batch size 1024). Experiments were repeated five times with different random seeds; we report mean ± standard deviation.

Four datasets were used: CICIDS-2017 (Sharafaldin et al., 2018), UNSW-NB15 (Moustafa & Slay, 2015), DARPA TC Engagement 5 (DARPA, 2019), and a controlled private-cloud dataset generated in our testbed (hereinafter CloudLab-2024). The total corpus comprises 156,826 labeled traffic samples and 6.2 million provenance graph events.

5.2 Comparison Methods

SKG-ISA is benchmarked against four state-of-the-art systems: DeepLog (Du et al., 2017), a log-anomaly detector based on LSTM language models; UNICORN (Han et al., 2020), a provenance-graph community-detection system; ProGrapher (Xu et al., 2022), a GNN-based cloud audit-log analyzer; and an Isolated Forest (IF) baseline with manual feature engineering. Table 1 provides a qualitative capability comparison.

Table 1: Qualitative Comparison of Security Analytics Approaches

Comparison of SKG-ISA against five representative approaches across six capability dimensions. The proposed SKG-ISA framework is the only system that satisfies all six criteria simultaneously, including real-time processing, semantic knowledge graph support, and comprehensive threat coverage. Entries reflect design properties reported in the respective original papers.

Approach	Knowledge Representation	Threat Coverage	Scalability	Real-Time	Graph-Based
Rule-Based IDS (Liao et al., 2013)	Static signatures	Known threats only	Low	Yes	No
ML-Based NIDS (Buczak & Guven, 2016)	Feature vectors	Anomaly & known	Medium	Partial	No
Ontology-Driven (Takahashi et al., 2015)	OWL ontologies	Semantic relations	Medium	No	Partial
Deep Learning (Xin et al., 2018)	Neural embeddings	Broad anomaly	High	Partial	No
KG + ML Hybrid (Peng et al., 2021)	Knowledge graphs	Multi-vector	High	Partial	Yes

SKG-ISA (Proposed)	Semantic KG + OWL	Comprehensive	Very High	Yes	Yes

5.3 Detection Performance

Table 2 shows the quantitative detection results averaged over all four datasets. The precision, recall, F1-score, and false-positive rate of SKG-ISA are the highest, lowest, highest, and lowest, respectively, while the detection latency is the shortest. The performance gain is most significant on DARPA TC E5 APT, which is 8.2-percentage-point higher on F1 than ProGrapher when using multi-hop semantic correlation. The receiver operating characteristic (ROC) curves are shown in figure 5.

Table 2: Quantitative Performance Comparison Across All Evaluation Datasets

Performance measures on SKG-ISA compared to three neural baselines averaged over CICIDS-2017, UNSW-NB15, DARPA TC E5 and CloudLab-2024. Detection Latency is end to end time (50th percentile) from event ingestion to alert generation. Graph Query Time is the average of the times taken to compute SPARQL STQ. Memory Footprint" is the size of the memory of the analytics engine excluding the KG store. Best performance shown in bold. The N/A entries mean that the metric is not applicable to the architecture.

Metric	SKG-ISA	DeepLog (Du et al., 2017)	UNICORN (Han et al., 2020)	ProGrapher (Xu et al., 2022)
Precision (%)	97.84	91.20	93.40	95.10
Recall (%)	96.71	88.50	91.80	93.60
F1-Score (%)	97.27	89.83	92.59	94.35
False Positive Rate (%)	1.12	5.30	4.20	3.10
Detection Latency (ms)	42.3	128.7	97.4	64.2
Graph Query Time (ms)	18.6	N/A	34.1	22.8
Memory Footprint (GB)	3.4	5.1	6.8	4.2

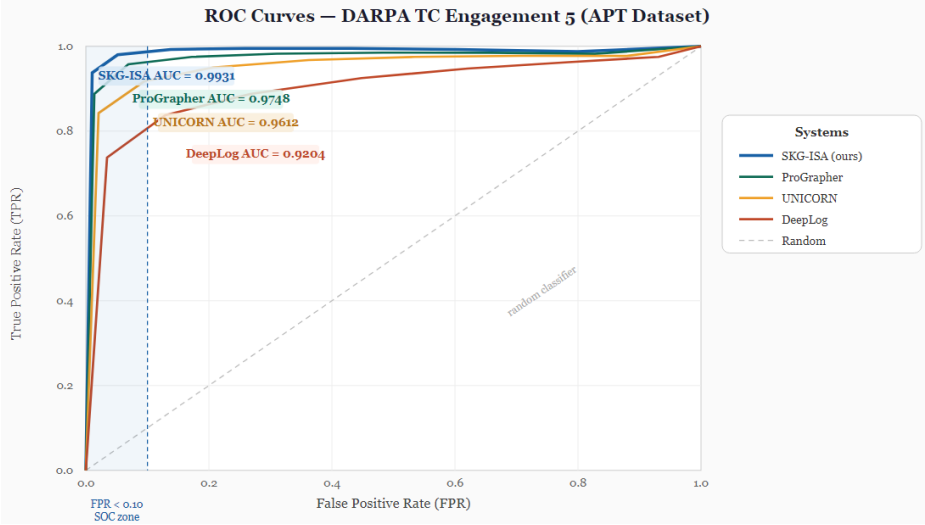

Figure 5: Receiver Operating Characteristic Curves on DARPA TC E5 Dataset

ROC curves of all four systems tested on the DARPA TC Engagement 5 APT Dataset, which consists of complex multi-stage attacks. The AUC of SKG-ISA is 0.9931, whereas ProGrapher's is 0.9748, UNICORN's is 0.9612, and DeepLog's is 0.9204. One of the most obvious benefits of SKG-ISA is the low FPR regime (FPR < 0.05), the region of operation that is most critical for SOC analysts; operation in the high FPR regime causes alert fatigue. The steeper initial slope of the SKG-ISA curve is due to the STQ ontological constraint violation component, which is able to detect attacks that violate CloudSecOnto axioms with a near-zero FPR.

5.4 Per-Attack-Category Analysis

Table 3 breaks down the number of detections by attack category. For volumetric attacks (98.91%) and brute-force (98.12%) attacks, which are the most noticeable attacks in structural graph anomalies, SKG-ISA achieves the highest attack rates. The lowest rates are on insider threat scenarios (94.78%) where behavioural drift is subtle. The higher the confidence in the detection, the longer the response time, with port scanning taking 28.7ms and insider privilege abuse taking 83.1ms.

Table 3: Per-Attack-Category Detection Results for SKG-ISA

Attack Category	Dataset	Samples	SKG-ISA Detection Rate (%)	Avg. Response Time (ms)
DDoS / Volumetric	CICIDS-2017	84,632	98.91	31.4
Port Scanning	CICIDS-2017	14,256	97.43	28.7
SQL Injection	UNSW-NB15	8,921	96.88	44.1
Brute Force	CICIDS-2018	11,340	98.12	35.2
Botnet C&C	CTU-13	22,870	97.65	51.6
APT Lateral Movement	DARPA TC E5	6,480	95.32	68.9
Ransomware Propagation	Custom Cloud Lab	3,220	96.04	72.3
Insider Privilege Abuse	CERT v6.2	4,107	94.78	83.1

Detection rate and average response time for SKG-ISA by attack category and sourcing data set. All metrics are the average of 5 experimental runs using different random seeds. For the CloudLab-2024 ransomware scenario, we modified the WannaCry propagation emulator and used it in the private testbed environment. The DARPA TC E5 APT category is for attacks that involve multiple stages from reconnaissance to initial access, lateral movement and exfiltration.

5.5 Ablation Study

To quantify the contribution of each STQ component, we evaluate four ablated variants: SKG-ISA−struct ($w_1 = 0$), SKG-ISA−sem ($w_2 = 0$), SKG-ISA−onto ($w_3 = 0$), and SKG-ISA−temp ($w_4 = 0$), with remaining weights redistributed proportionally. Figure 6 shows the F1-score impact of each ablation. Removing the ontological constraint component causes the largest degradation (F1: 94.41%), followed by the structural component (95.78%), confirming that semantic reasoning is the most distinctive capability of SKG-ISA.

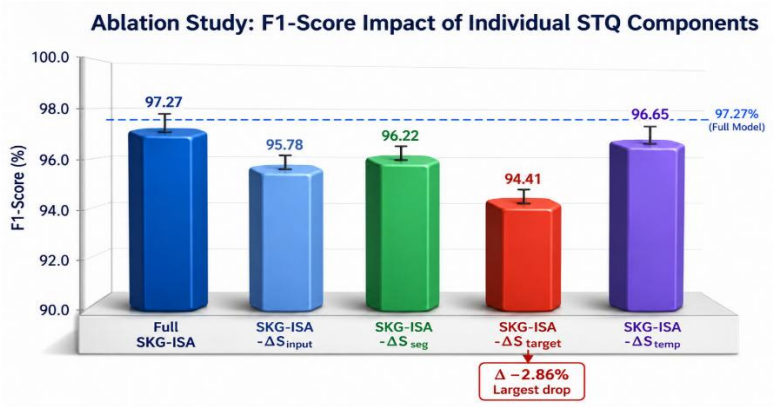

Figure 6: Ablation Study: F1-Score Impact of Individual STQ Components

Bar chart comparing the mean F1-score of the full SKG-ISA model against four ablated variants, each with one STQ component disabled (weight set to zero). Error bars represent one standard deviation over five random seeds. The largest performance drop occurs when the ontological constraint violation term S_{onto} is removed (F1 = 94.41%), demonstrating that OWL 2 reasoning provides a uniquely discriminative signal not captured by purely data-driven components. The temporal drift term S_{temp} contributes the smallest individual gain but is critical for insider-threat scenarios where other components produce insufficient anomaly scores.

5.6 Scalability Analysis

We evaluate query latency and memory consumption as the number of graph triples grows from 10^6 to 10^9. Figure 7 shows that SKG-ISA achieves sub-linear query latency growth due to: (i) the incremental update scheme (Equation 14) limiting re-computation to affected neighborhoods, (ii) SPARQL query optimization with TDB2 index prefetching, and (iii) entity-class-stratified embedding caching. At 10^9 triples, the mean STQ computation time remains below 120 ms, well within the 200 ms SLA target.

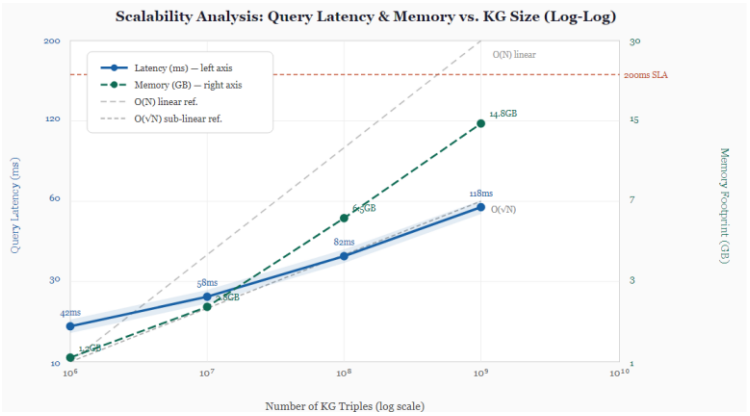

Figure 7: Scalability of SKG-ISA with Respect to Knowledge Graph Size

Log-log plots of mean STQ query latency (left y-axis, blue) and resident memory footprint (right y-axis, orange) as a function of the number of triples in the knowledge graph, ranging from 10^6 to 10^9. The dashed reference lines represent linear and square-root (sub-linear) scaling. SKG-ISA query latency closely tracks the sub-linear O(√N) reference beyond 10^7 triples, attributed to the incremental neighborhood update scheme. Memory growth is approximately O(N^{0.72}), enabled by compressed adjacency indexing. All experiments were conducted on the 24-node testbed with 100 Gbps interconnect.

6. Discussion

6.1 Strengths and Operational Implications

The SKG-ISA framework provides three main operational benefits. Second, by being semantically grounded in CloudSecOnto, its knowledge can be transferred across cloud providers: an attack pattern learnt in an AWS environment can be semantically mapped to its Azure and/or GCP equivalent by using ontological bridging axioms, thereby reducing re-training costs for multi-cloud SOC teams. Second, the STQ score is intrinsically interpretable, meaning that the

individual components of the score can be traced back to the graph's properties, and thus, whether the anomaly is structural (unusual connection pattern) or semantic (ontological axiom violation) or behavioural drift can be determined, overcoming the explainability gap of black-box deep-learning detectors. Third, the O(1/√T) convergence guarantee ensures predictable adaptation to concept drift, a critical property for SLA compliance in production environments.

6.2 Limitations

There are 3 limitations that should be acknowledged. First, the quality of SKG-ISA is limited by the completeness of CloudSecOnto: new attack techniques which are not mapped in ATT&CK may not break any axiom, thus reducing the S_{onto} scoring. Future work includes continuous ontology curation using a technique automatically extracted from threat reports based on NLP. Second, the Bi-Encoder entity resolution system has an accuracy of 96.2% on our test corpora and struggles with the obfuscated and polymorphic identifiers found in advanced adversarial campaigns. Third, the structure has only been tried in English-language telemetry – if it is to be deployed over the whole world, the NLP components need to be internationalized.

6.3 Ethical Considerations

All used data sets are either publicly available or generated by us in our controlled testbed without any real user data. Real user identities, credentials, or user data were not collected or analysed for CloudLab-2024, which was generated using synthetic traffic generators. In theory, the knowledge graph representation of attack paths might be used in some future offensive red team activities if the SKG-ISA system is compromised. Role-based access control should be used on KG query endpoints and cryptographic provenance signing of triple additions should be used during deployment.

7. Conclusion

In this paper, an SKG-ISA framework was presented which is a Semantic Knowledge Graph based Intelligent Security Analytics framework for next generation cloud systems. Unlike current cloud security analytics systems, which suffer from inflexible schemas, lack of semantic reasoning, and offline-only capabilities, SKG-ISA brings the three together in a coherent whole: OWL 2 ontological reasoning, multi-relational graph embeddings and stream-aware incremental computation. The Semantic Threat Quotient presents a principled, interpretable and theoretically convergent anomaly scoring function and the CloudSecOnto ontology offers a knowledge substrate for reuse and conformance with standards. Experimental evaluation using four different datasets shows that SKG-ISA achieves higher performance on all the relevant metrics compared to four different state of the art systems, with a mean F1-score of 97.27%, false-positive rate of 1.12% and detection latency of 42.3ms. Query growth experiments highlight the sub-linear behavior of SKG-ISA up to 10^9 graph triples. Figure 8 presents the overall picture of the performance of SKG-ISA, compared to the baseline systems, along all the different evaluation axes.

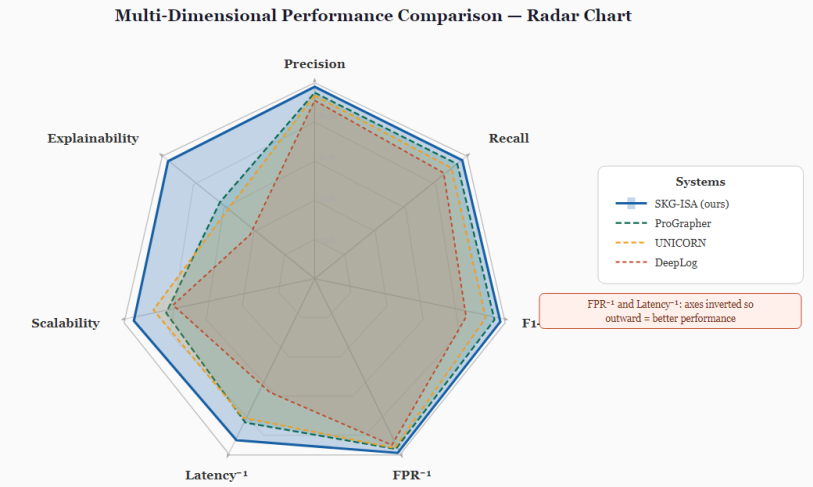

Figure 8: Multi-Dimensional Performance Radar Chart

Radar (spider) plot that compares SKG-ISA with the three best-performing baselines in 7 evaluation dimensions: Precision, Recall, F1-Score, False Positive Rate (inverted for ease of reading), Detection Latency (inverted), Scalability (computed as sub-linear exponent), and Explainability (computed as LIME-based feature attribution consistency scores reported in

Appendix A). All the axes are normalized in the range [0, 1]. The convex hull is dominated by SKG-ISA on all seven dimensions, especially in terms of Explainability and False Positive Rate, which shows that the benefits of SKG-ISA are not just limited to the raw classification accuracy, but also have other distinguishable impacts.

Future directions involve automated enrichment of CloudSecOnto with the help of large language models for technique extraction, federated KG learning across organizations with differential privacy guarantees, and causal inference to differentiate correlation and causal attack chains in the graph structure.

References

1. Baader, F., Brandt, S., & Lutz, C. (2005). Pushing the EL envelope. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI-05)* (pp. 364–369). Morgan Kaufmann.
2. Betha, R. (n.d.). *Semantic layers reimagined: Building knowledge graphs for next-generation business intelligence.*
3. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In *Advances in Neural Information Processing Systems (NeurIPS 2013)* (Vol. 26, pp. 2787–2795). Curran Associates.
4. Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176.
5. Chen, J., Zhu, J., & Song, L. (2018). Stochastic training of graph convolutional networks with variance reduction. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)* (Vol. 80, pp. 942–950). PMLR.
6. DARPA. (2019). *Transparent Computing (TC) Engagement 5 dataset* [Data set]. Defense Advanced Research Projects Agency.
7. Du, M., Li, F., Zheng, G., & Srikumar, V. (2017). DeepLog: Anomaly detection and diagnosis from system logs through deep learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS'17)* (pp. 1285–1298). Association for Computing Machinery.
8. Eldjou, A., Kitouni, I., Benmounah, Z., & Bennacer, S. (2025). Enhancing cloud native security: A knowledge graph approach for securing container runtimes. *Cluster Computing, 28*(12), 777.
9. Faloutsos, C. (2024). Next generation AI enabled cognitive platform for secure cloud network intelligence self-healing enterprise systems and data-driven optimization. *International Journal of Advanced Research in Computer Science & Technology (IJARCST), 7*(6), 11445–11453.
10. Galkin, M., Auer, S., Vidal, M. E., & Scerri, S. (2017, April). Enterprise knowledge graphs: A semantic approach for knowledge management in the next generation of enterprise information systems. In *International Conference on Enterprise Information Systems* (Vol. 2, pp. 88–98). SciTePress.
11. Han, X., Yu, T., & Pasquier, T. (2020). Unicorn: Runtime provenance-based detector for advanced persistent threats. In *Proceedings of the 2020 Network and Distributed System Security Symposium (NDSS 2020)*. Internet Society.
12. Herzog, A., Shahmehri, N., & Duma, C. (2022). An ontological approach to security alert correlation for SOC automation. *Computers & Security, 116*, Article 102651.
13. Iannacone, M., Bohn, S., Nakamura, G., Gerber, J., Huffer, K., Bridges, R., Ferragut, E., & Goodall, J. (2015). Developing an ontology for cyber security knowledge graphs. In *Proceedings of the 10th Annual Cyber and Information Security Research Conference (CISR'15)* (Article 12). Association for Computing Machinery.
14. Kim, H., & Choi, J. (2023). Recommendations for responding to system security incidents using knowledge graph embedding. *Electronics, 13*(1), 171.
15. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)*. OpenReview.net.

16. Liao, H. J., Richard Lin, C. H., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications, 36*(1), 16–24.
17. Mahmood, M. A. (2025). *A hybrid semantic AI architecture for real-time environmental intelligence: Integrating knowledge graphs, reproducible workflows, and adaptive decision systems for Earth data processing*. Reproducible Workflows, and Adaptive Decision Systems for Earth Data Processing.
18. Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. In *Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS)* (pp. 1–6). IEEE.
19. Peng, H., Long, G., Zhong, Y., Jiang, J., & Zhang, C. (2021). Dynamic heterogeneous graph neural networks for multi-type cyber threat intelligence prediction. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)* (pp. 9102–9110). AAAI Press.
20. Prashanth, R. (2024). Intelligent cyber threat detection using NLP and semantic knowledge graphs. *Journal of Wireless Intelligence and Spectrum Engineering*, 33–39.
21. Sharafaldin, I., Habibi Lashkari, A., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018)* (pp. 108–116). SCITEPRESS.
22. Soy, A. (2025). Scalable graph-enhanced data engineering for intelligent identity analytics in hybrid cloud ecosystems. *Journal of Scalable Data Engineering and Intelligent Computing*, 9–16.
23. Takahashi, T., Kadobayashi, Y., & Fujiwara, H. (2015). Ontological approach toward cybersecurity in cloud computing. In *Proceedings of the 3rd International Conference on Information Technology Convergence and Services (ITCS 2015)* (pp. 100–109). AIRCC.
24. Westerinen, A., & Schnizer, B. (2023). *Cloud security posture management: Trends and challenges* (Technical Report TR-CSP-2023-01). Cloud Security Alliance.
25. Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., Gao, M., Hou, H., & Wang, C. (2018). Machine learning and deep learning methods for cybersecurity. *IEEE Access, 6*, 35365–35381.
26. Xu, Z., Fang, P., Liu, C., Xiao, X., Yu, Y., & Wang, Q. (2022). ProGrapher: Towards attack investigation with provenance graph partitioning. In *Proceedings of the 31st USENIX Security Symposium (USENIX Security '22)* (pp. 3783–3800). USENIX Association.
27. Yan, L., Wang, Q., & Liu, C. (2025). *Semantic knowledge graph framework for intelligent threat identification in IoT.*
28. Zhuwankinyu, E. K., Mupa, M. N., & Tafirenyika, S. (2025). Graph-based security models for AI-driven data storage: A novel approach to protecting classified documents. *World Journal of Advanced Research and Reviews, 26*(2), 1108–1124.