

Deep Graph Learning Architecture for Real-Time Cyber Threat Identification and Detection in Cloud Platforms

Muhammad Rizki Pratama¹, Siti Nurhaliza Putri²

¹ Faculty of Computer Science, Universitas Indonesia, Indonesia

² Department of Informatics, Universitas Indonesia, Indonesia

ABSTRACT: The rapid adoption of cloud computing has transformed enterprise information infrastructures into highly dynamic environments characterized by distributed resources, elastic workloads, interconnected services, and continuously changing user and application behavior. These characteristics increase the complexity of identifying cyber threats because malicious activities frequently propagate across multiple entities rather than appearing as isolated events. Conventional machine-learning approaches that treat security observations independently can therefore overlook relational dependencies among users, virtual machines, applications, network flows, and cloud services. This paper proposes a Deep Graph Learning Architecture for Real-Time Cyber Threat Identification and Detection in Cloud Platforms, positioning cloud security monitoring as a graph-based learning problem. The proposed architecture integrates graph construction, deep graph representation learning, adaptive threat classification, continual learning, and real-time alert generation. The theoretical design is informed by research on continual learning, adaptive regularization, domain adaptation, task-aware learning, memory-aware learning, and graph-based cyber-threat identification. Particular emphasis is placed on maintaining detection performance under evolving attack distributions while limiting catastrophic forgetting. The framework conceptualizes cloud entities as graph nodes and their interactions as dynamically updated edges, enabling the model to capture structural and behavioral relationships. Analytical findings indicate that combining relational representations with continual adaptation can improve the suitability of threat detection systems for evolving cloud environments, although computational overhead, graph scalability, concept drift, and uncertainty remain important limitations. The study contributes an integrated research architecture for real-time cloud threat identification and establishes a foundation for adaptive graph-based cybersecurity systems.

KEYWORDS: Deep Graph Learning, Cloud Computing Security, Cyber Threat Detection, Graph Neural Networks, Real-Time Detection, Continual Learning, Concept Drift, Threat Intelligence, Adaptive Learning.

1. Introduction

1.1 Background

Cloud computing environments generate security data from heterogeneous and highly interconnected components, including identities, applications, containers, virtual machines, APIs, network sessions, storage resources, and authentication events. Unlike static computing environments, cloud infrastructures continuously change their topology and operational behavior. Consequently, cybersecurity detection cannot rely exclusively on fixed signatures or independently evaluated events. A suspicious login, for example, may become significantly more informative when it is associated with an unusual API invocation, access to a sensitive resource, and communication with a previously unseen service.

Graph learning provides a suitable theoretical basis for representing these relationships. In a graph-oriented cloud security model, entities can be represented as nodes, while communication, authentication, resource access, process execution, or data-transfer relationships can be represented as edges. The resulting graph provides contextual information that conventional feature-vector approaches may lose. The work of Marri et al. (2025) specifically establishes the relevance of graph-based deep learning for identifying cyber threats in cloud platforms, providing an important foundation for the architecture developed in this study.

A further challenge is that cyber threats are not stationary. Attack strategies, infrastructure configurations, workloads, and legitimate user behavior evolve over time. Continual-learning research

demonstrates that neural models can experience catastrophic forgetting when learning sequentially from changing data distributions (Aljundi et al., 2019b; Aljundi et al., 2019d). Therefore, a practical cloud threat-detection architecture requires not only strong initial classification capability but also mechanisms for adapting to newly emerging threat patterns without destroying previously learned knowledge.

1.2 Problem Statement

Existing cloud threat-detection approaches face three interconnected problems. First, security observations are relational, whereas many detection systems process events as independent samples. Second, cloud environments exhibit distributional changes that can reduce the effectiveness of models trained on historical traffic. Third, continuous retraining can introduce computational costs and catastrophic forgetting.

These problems motivate a deep graph learning architecture that simultaneously models structural dependencies and supports adaptive learning. The central research problem is therefore how to construct a real-time graph representation of cloud activity and use deep learning to identify anomalous or malicious patterns while preserving previously acquired knowledge.

1.3 Research Relevance

The relevance of the proposed approach arises from the intersection of graph representation learning and continual learning. Graph-based cyber-threat identification provides relational context, while continual-learning mechanisms provide adaptability. Marri et al. (2025) demonstrate the importance of graph-based deep learning in cloud threat identification, whereas studies such as Aljundi et al. (2018), Ahn et al. (2019), and Adel et al. (2020) demonstrate mechanisms for controlling knowledge retention and adaptation in sequential learning environments.

1.4 Objectives

The primary objective is to develop a conceptual deep graph learning architecture capable of real-time cyber-threat identification in cloud platforms. The study specifically aims to: establish a graph representation of cloud security events; develop a deep relational representation mechanism; integrate adaptive and continual learning principles; provide a threat-classification and alert-generation layer; and critically analyze the architecture's expected benefits and limitations.

1.5 Scope and Significance

The proposed architecture focuses on cloud-based

cybersecurity monitoring rather than a specific cloud vendor or attack category. It is intended to support network, identity, application, and resource-level security observations. Its significance lies in treating cloud cybersecurity as a dynamic relational learning problem rather than solely as conventional binary classification.

2. Literature Review

Research relevant to the proposed architecture can be divided into three closely related areas: continual representation learning, adaptive model optimization, and graph-oriented cloud threat detection.

Abati et al. (2020) introduced conditional channel gating for task-aware continual learning, demonstrating that neural architectures can selectively activate computational representations according to task requirements. The underlying principle is valuable for cloud security because different traffic contexts may require different feature pathways. A cloud monitoring system encountering authentication anomalies, lateral movement, or abnormal resource access could benefit from selective representation rather than processing every security observation identically.

Achille et al. (2018) examined lifelong disentangled representation learning and emphasized the preservation of transferable latent structures across domains. This concept is relevant to cloud security because cloud environments may differ in workload composition while still sharing fundamental behavioral patterns. A representation that separates persistent security characteristics from environment-specific characteristics could improve transferability.

Adel et al. (2020) proposed Continual Learning with Adaptive Weights (CLAW), highlighting adaptive parameter management as a mechanism for sequential learning. A related concern is addressed by Ahn et al. (2019), whose uncertainty-based continual learning framework uses adaptive regularization. In threat detection, uncertainty is particularly important because security data may contain ambiguous behavior. Instead of forcing every observation into a highly confident classification, uncertainty-aware adaptation can reduce inappropriate model updates.

Ahn et al. (2021) proposed separated softmax for incremental learning, addressing classification interference when new categories are introduced. This principle can be transferred conceptually to evolving threat taxonomies, where newly observed attack behaviors should not necessarily invalidate existing categories.

The domain-adaptation perspective is represented by Ajakan et al. (2014), who introduced domain adversarial

neural networks. Their work is theoretically relevant because cloud environments constitute multiple operational domains. Differences between development, production, private-cloud, and hybrid-cloud workloads can produce distribution shifts. Domain-invariant representations may therefore improve generalization across changing environments.

Few-shot incremental learning is examined by Akyurek et al. (2021), whose subspace regularizers address learning when only limited examples of new classes are available. This is directly relevant to cybersecurity because emerging attacks may initially provide very few labeled observations. A graph threat-detection model should therefore be able to incorporate limited new evidence without requiring large retraining datasets.

Aljundi et al. (2018) introduced Memory Aware Synapses, emphasizing parameter importance when deciding what should be retained during continual learning. Aljundi et al. (2019a) further investigated maximally interfered retrieval for online continual learning, while Aljundi et al. (2019c) studied gradient-based sample selection. These works collectively suggest that continual cloud threat detection should not simply retain all historical events. Instead, strategically selected security observations and important parameters can be preserved to control memory and computational requirements.

Task-free continual learning was investigated by Aljundi et al. (2019b), which is particularly important for cloud environments because attack and workload transitions may occur without explicit task boundaries. Aljundi et al. (2019d) similarly emphasized selfless sequential learning, providing theoretical support for learning under sequentially arriving data.

Most directly, Marri et al. (2025) establish a graph-based deep learning direction for cyber-threat identification in cloud platforms. Building upon this foundation, the present study extends the conceptual architecture by emphasizing real-time graph construction and continual adaptation. The principal research gap is therefore not simply the use of graph learning, but the integration of relational threat representation with adaptive sequential learning under evolving cloud conditions.

3. Methodology

3.1 Proposed Research Architecture

The proposed architecture consists of six interconnected layers: cloud telemetry acquisition, dynamic graph construction, deep graph representation learning, adaptive threat inference, continual-learning management, and real-time alert generation.

The conceptual pipeline is:

Cloud Events → Feature Extraction → Dynamic Security Graph → Deep Graph Representation → Threat Classification → Continual Adaptation → Real-Time Alert

Cloud telemetry may include authentication events, network connections, API requests, resource accesses, process activities, and workload interactions. Each event is transformed into structured attributes before entering the graph-construction layer.

3.2 Dynamic Cloud Security Graph

Let the cloud security environment at time (t) be represented as:

$$\begin{bmatrix} G_t = (V_t, E_t, X_t) \end{bmatrix}$$

where (V_t) represents cloud entities, (E_t) represents interactions, and (X_t) represents node and edge attributes.

Nodes may represent users, devices, virtual machines, containers, applications, services, or storage resources. Edges represent relationships such as:

$$\begin{bmatrix} E = \{(u, v, r, \tau)\} \end{bmatrix}$$

where (u) and (v) are connected entities, (r) denotes the relationship type, and (τ) denotes temporal information.

For example, a user authenticating to a virtual machine creates an identity-to-resource edge. If that virtual machine subsequently communicates with an unfamiliar external endpoint, another edge is created. A sequence of apparently legitimate relationships may therefore produce a suspicious graph structure when evaluated collectively.

This relational representation follows the conceptual direction established by Marri et al. (2025), while dynamic graph updates address the temporal characteristics of cloud environments.

3.3 Deep Graph Representation Layer

For each node (v), an initial representation ($h_v^{(0)}$) is obtained from telemetry attributes. A generic graph aggregation operation can be represented as:

$$\begin{bmatrix} h_v^{(l+1)} = \sigma(W_l h_v^{(l)} + \dots \end{bmatrix}$$

$$\sum_{u \in N(v)} \alpha_{uv} W_{lh_u}^{(l)} \right)$$

]

where $(N(v))$ represents neighboring entities, (W_l) is a trainable transformation, (α_{uv}) represents the relevance of neighboring information, and (σ) denotes a nonlinear activation function.

The important theoretical property is that threat classification becomes context-aware. A node is not evaluated solely according to its own attributes; neighboring behaviors contribute to its representation.

3.4 Temporal and Behavioral Modeling

Cloud attacks frequently unfold through sequences. Therefore, graph representations should preserve temporal information. A node that performs an unusual operation once may be benign, whereas repeated interactions with sensitive services followed by privilege escalation can represent a coordinated attack.

A temporal representation can be formulated as:

[

$$z_v^t = f(G_{\{t-k\}}, X_{\{t-k\}})$$

]

where $(G_{\{t-k\}})$ denotes a sequence of recent security graphs. The resulting embedding captures both structural and temporal information.

3.5 Threat Classification

The threat-inference layer maps learned representations into threat probabilities:

[

$$P(y|z_v) = \text{softmax}(W_{cz_v} + b_c)$$

]

where (y) may represent benign behavior, suspicious activity, credential abuse, unauthorized access, lateral movement, or another operationally defined threat class.

For highly uncertain observations, the model can defer automatic high-severity classification and instead generate an adaptive alert. This approach is consistent with uncertainty-aware continual learning principles proposed by Ahn et al. (2019).

3.6 Continual Learning Mechanism

A major component of the proposed framework is continual adaptation. Assume that cloud observations arrive sequentially as:

[

$$D_1, D_2, \dots, D_T$$

]

The model must learn (D_t) without excessively degrading performance on previous distributions.

A parameter-importance mechanism inspired by Memory Aware Synapses can constrain changes to important parameters (Aljundi et al., 2018). An adaptive regularization objective can be represented as:

[

$$L = L_{\text{new}} + \lambda \sum_i \Omega_i (\theta_i - \theta_i^{\text{old}})^2$$

]

where (Ω_i) measures the importance of parameter (i) . Important parameters receive stronger protection, while less important parameters remain adaptable.

Gradient-based sample selection can additionally prioritize informative security observations (Aljundi et al., 2019c). This is useful because retaining every cloud event is computationally inefficient. Selected historical observations can provide a compact memory for replay and reduce catastrophic forgetting.

3.7 Domain Adaptation

Cloud infrastructures can exhibit substantial domain variation. Domain adversarial learning provides a theoretical mechanism for learning representations that are less sensitive to environmental differences (Ajakan et al., 2014). In the proposed architecture, a domain discriminator can attempt to identify the cloud environment while the graph encoder learns representations that retain security information but suppress irrelevant domain-specific characteristics.

This design is especially useful when a model trained in one cloud deployment must operate in another with different workload distributions.

3.8 Few-Shot Adaptation to Emerging Threats

New attacks may initially have insufficient labeled examples. Few-shot incremental learning principles, including subspace regularization, can help incorporate limited examples without requiring complete model retraining (Akyurek et al., 2021). Similarly, disentangled representations may facilitate transfer of common security features across related environments (Achille et al., 2018).

The practical implication is that a newly emerging threat can initially be treated as an uncertain candidate class. As additional evidence accumulates, the model can progressively refine its representation.

3.9 Real-Time Alert Generation

The final stage transforms threat probabilities into operational alerts. An alert score can combine classification confidence, structural anomaly, temporal deviation, and asset criticality:

$$A = \alpha P_{\text{threat}} + \beta S_{\text{graph}} + \gamma S_{\text{temporal}} + \delta C_{\text{asset}}$$

where (A) is the alert priority score. This permits the system to distinguish between a low-risk anomaly on a noncritical workload and a structurally similar anomaly involving a privileged identity or sensitive resource.

3.10 Training and Evaluation Strategy

The architecture should be evaluated using chronological rather than purely random data partitioning because random splitting can underestimate concept drift. Training should use historical observations, validation should represent subsequent operational periods, and testing should contain temporally later events.

Evaluation should consider detection effectiveness, false-positive behavior, adaptation performance, inference latency, memory consumption, and forgetting. Continual-learning literature indicates that classification accuracy alone is insufficient for sequential learning because a model can perform well on new data while losing knowledge acquired previously (Aljundi et al., 2019a; Aljundi et al., 2019b).

4. Results

The analytical findings of the proposed architecture indicate that graph-based representation is particularly suitable for cloud threat detection because security behavior is inherently relational. A conventional event-level classifier may identify an abnormal login but cannot necessarily determine whether that login forms part of a larger attack chain. By incorporating neighboring identities, workloads, services, and network interactions, graph representations provide contextual evidence. This finding is consistent with the graph-based cloud threat-identification direction established by Marri et al. (2025).

A second finding concerns adaptability. Cloud environments cannot reasonably be modeled as static distributions because legitimate workloads and attack behaviors change continuously. The continual-learning literature indicates that adaptive mechanisms such as importance-based parameter preservation, uncertainty-aware regularization, selective replay, and gradient-informed sample selection can reduce the destructive

effects of sequential learning (Aljundi et al., 2018; Ahn et al., 2019; Aljundi et al., 2019c). Consequently, the proposed architecture treats model adaptation as an intrinsic component rather than as an occasional retraining operation.

A third finding is that emerging threats create an asymmetry between detection requirements and available training data. New attack categories often begin with limited observations. Few-shot incremental learning provides a theoretical response by enabling model adaptation from small samples (Akyurek et al., 2021). Combined with transferable latent representations, this approach can reduce dependence on large labeled datasets.

The fourth finding concerns domain variation. Cloud deployments may differ in workload composition, topology, and operational behavior. Domain-adversarial principles can reduce the impact of non-security-related domain differences (Ajakan et al., 2014). This suggests that the proposed architecture may generalize more effectively when graph representations emphasize invariant security relationships rather than deployment-specific characteristics.

Finally, the architecture introduces practical trade-offs. Dynamic graph construction and neighborhood aggregation can increase computational requirements as the number of cloud entities and relationships grows. Real-time operation therefore requires efficient graph updates, bounded memory mechanisms, and selective processing. Conditional computation principles such as those explored by Abati et al. (2020) provide a potential route toward reducing unnecessary computation. The overall finding is that a graph-based detector is most valuable when its relational expressiveness is combined with efficient adaptive mechanisms rather than implemented as a static deep-learning classifier.

5. Discussion

The proposed architecture contributes theoretically by integrating two dimensions that are often treated separately: relational security modeling and continual adaptation. Graph learning provides the structural mechanism through which cloud entities and their interactions become jointly interpretable, while continual learning provides the mechanism for maintaining relevance as those interactions change. The combination is particularly appropriate for cloud cybersecurity because attacks can evolve at both structural and temporal levels.

The architecture also extends the conceptual contribution of Marri et al. (2025) by emphasizing that graph-based threat identification should not be regarded as a static classification task. A cloud graph is continuously changing. Users are created and removed,

services scale dynamically, workloads migrate, and relationships appear and disappear. Consequently, a graph model trained once may progressively become less representative of operational reality. Continual-learning strategies described by Aljundi et al. (2019b) and Aljundi et al. (2019d) are therefore theoretically important for maintaining model utility without assuming explicit task boundaries.

An important trade-off concerns adaptation versus knowledge preservation. Excessive adaptation can cause catastrophic forgetting, whereas excessive regularization can prevent the model from learning genuinely new attack behaviors. Memory-aware parameter protection provides one response (Aljundi et al., 2018), while adaptive weighting offers another perspective (Adel et al., 2020). The proposed framework consequently requires a dynamic balance between stability and plasticity rather than maximizing either property independently.

Uncertainty represents another critical issue. Cybersecurity datasets frequently contain incomplete, noisy, or ambiguous observations. A highly confident model prediction does not necessarily indicate a reliable security conclusion. Uncertainty-based adaptive regularization (Ahn et al., 2019) therefore provides a useful theoretical basis for deciding when model updates or alerts should receive stronger scrutiny.

The framework also faces scalability limitations. A large cloud deployment can contain millions of nodes and relationships, making full-graph computation expensive. Selective sample retrieval (Aljundi et al., 2019a), gradient-based selection (Aljundi et al., 2019c), and conditional computation (Abati et al., 2020) indicate possible strategies for reducing unnecessary computation. However, these mechanisms may also discard information that later proves important for identifying sophisticated attack chains.

Domain adaptation introduces a further tension. Removing deployment-specific information can improve generalization, but excessive invariance may remove legitimate contextual signals needed for distinguishing benign and malicious behavior. The principles of domain adversarial learning (Ajakan et al., 2014) must therefore be applied selectively.

Overall, the architecture should be interpreted as a research framework rather than evidence of universal superiority over all existing detection systems. Its effectiveness depends on graph quality, telemetry completeness, label availability, computational resources, and the ability of the continual-learning mechanism to distinguish genuine concept drift from temporary anomalies. Future empirical work should therefore benchmark the architecture against graph and non-graph baselines using temporally separated cloud-

security datasets and explicit measures of forgetting, latency, and false-positive cost.

6. Conclusion

This paper proposed a Deep Graph Learning Architecture for Real-Time Cyber Threat Identification and Detection in Cloud Platforms. The architecture models cloud security as a dynamic graph-learning problem in which users, workloads, applications, services, and resources are represented as interconnected entities. Deep graph representations provide contextual information for threat identification, while continual-learning mechanisms support adaptation to changing attack and workload distributions.

The principal contribution is the integration of graph-based threat representation with adaptive learning mechanisms derived from continual-learning research. The framework incorporates memory-aware parameter protection, uncertainty-aware adaptation, selective replay, domain adaptation, and few-shot learning concepts to address the realities of evolving cloud environments. The approach is particularly aligned with the graph-based cloud threat-detection direction of Marri et al. (2025).

Nevertheless, scalability, uncertainty, concept drift, data quality, and catastrophic forgetting remain substantial challenges. Future research should implement the architecture experimentally, investigate temporal and heterogeneous graph neural networks, optimize streaming graph construction, and evaluate the system using realistic evolving cloud workloads. Such developments could establish more adaptive, context-aware, and operationally efficient cybersecurity systems for modern cloud platforms.

References

1. Abati, D., Tomczak, J., Blankevoort, T., Calderara, S., Cucchiara, R., & Bejnordi, B. (2020). Conditional channel gated networks for task-aware continual learning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
2. Achille, A., Eccles, T., Matthey, L., Burgess, C., Watters, N., Lerchner, A., & Higgins, I. (2018). Life-long disentangled representation learning with cross-domain latent homologies. *Advances in Neural Information Processing Systems (NIPS)*.
3. Adel, T., Zhao, H., & Turner, R. (2020). Continual learning with adaptive weights (CLAW). *International Conference on Learning Representations (ICLR)*.
4. Ahn, H., Kwak, J., Lim, S., Bang, H., Kim, H., & Moon, T. (2021). SS-IL: Separated softmax for incremental learning. *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

5. Ahn, H., Lee, D., Cha, S., & Moon, T. (2019). Uncertainty-based continual learning with adaptive regularization. *Advances in Neural Information Processing Systems (NeurIPS)*.
6. Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., & Marchand, M. (2014). Domain adversarial neural networks. *CoRR*, abs/1412.4446.
7. Akyurek, A., Akyurek, E., Wijaya, D., & Andreas, J. (2021). Subspace regularizers for few-shot class incremental learning. *arXiv preprint arXiv:2110.07059*.
8. Aljundi, R., Babiloni, F., Elhoseiny, M., Rohrbach, M., & Tuytelaars, T. (2018). Memory aware synapses: Learning what (not) to forget. *European Conference on Computer Vision (ECCV)*.
9. Aljundi, R., Caccia, L., Belilovsky, E., Caccia, M., Lin, M., Charlin, L., & Tuytelaars, T. (2019a). Online continual learning with maximally interfered retrieval. *Advances in Neural Information Processing Systems (NeurIPS)*.
10. Aljundi, R., Kelchtermans, K., & Tuytelaars, T. (2019b). Task-free continual learning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
11. Aljundi, R., Lin, M., Goujaud, B., & Bengio, Y. (2019c). Gradient based sample selection for online continual learning. *Advances in Neural Information Processing Systems (NeurIPS)*.
12. Aljundi, R., Rohrbach, M., & Tuytelaars, T. (2019d). Selfless sequential learning. *International Conference on Learning Representations (ICLR)*.
13. M. R. Marri, S. B. Kurada, S. Gupta, S. Dey, S. Kumar and R. Chauhan, "Graph-Based Deep Learning Model for Identifying Cyber Threats in Cloud Platforms," 2025 International Conference on Computational Intelligence, Security, and Artificial Intelligence (IntelliSecAI), Al-Khobar, Saudi Arabia, 2025, pp. 1-7, doi: 10.1109/IntelliSecAI66368.2025.11472831.