Towards Intrusion Detection and Trust-Based Slice Access in 5G/B5G Networks

Li-Yu Yang
Department of Computer Science and Information Engineering
National Central University
Taoyuan, Taiwan
113522067@cc.ncu.edu.tw

Jyun-Wei Chen
Department of Computer Science and Information Engineering
National Central University
Taoyuan, Taiwan
113522075@cc.ncu.edu.tw

Li-Der Chou
Department of Computer Science and Information Engineering
National Central University
Taoyuan, Taiwan
cld@csie.ncu.edu.tw

***Abstract*—Network slicing, a key capability in 5G network, enables telecom operators to create multiple end-to-end virtual network slices on shared infrastructure. While offering flexible deployment, multi-tenancy and dynamic resource allocation in slices also increase security risks. To mitigate the vulnerabilities of the network slicing architecture, we propose an Autoencoder-Based Intrusion Detection System with Trust Level Driven Slice Access Control (AEIDS-TLDSAC) framework. Autoencoder-based IDS, leverages reconstruction error to detect anomalies, enabling recognition of unknown attacks and making it suitable for dynamic slicing environments. TLDSAC maintains a trust value for each UE, adjusting it based on IDS detection results. Malicious behavior lowers the trust value and triggers access restrictions, ensuring only UEs meeting trust requirements can access slices. Experiment results show that our Autoencoder-based IDS achieved 97.16% and 97.34% accuracy on two 5G testbed datasets. In addition, Additive-Increase Multiplicative-Decrease (AIMD) and Additive-Increase Additive-Decrease (AIAD) based TLDSAC mechanisms outperformed the baseline in average, median, and tail response time, with AIMD delivering the best reductions—71.23%, 85.26%, and 63.37%, respectively.**

Keywords—5G/B5G, Network Slicing, Autoencoder, Intrusion Detection System, Trust Level, Access Control

I. Introduction

In 5G/B5G networks, network slicing is a foundational capability that enables telecom operators to partition a shared physical infrastructure into multiple, isolated logical slices, each tailored to the needs of a specific service class or application. A slice provides end-to-end connectivity across the User Equipment (UE), Radio Access Network (RAN), and core network, while dynamically allocating compute, storage, and network resources to meet application demands. For example, a slice supporting remote surgery must be prioritized and engineered with ultra-low latency and high reliability to guarantee deterministic performance, whereas a video-streaming slice primarily requires sustained high throughput to ensure stable playback. By elastically orchestrating resources and policies per slice, 5G/B5G systems can more effectively satisfy heterogeneous requirements and provide different services such as enhanced Mobile Broadband (eMBB), Ultra-Reliable and Low Latency Communications URLLC), and massive Machine Type Communication (mMTC) without violating Service Level Agreements (SLAs) or Quality of Service (QoS) constraints.

Despite these benefits, network slicing also expends the attack and abuse surface. Since slices are typically deployed in multi-tenant environments[1], where diverse users and applications coexist and share underlying network functions and infrastructure resources. This makes malicious users become harder to detect, and benign users are more likely to be affected. Moreover, since slice resources can be allocated dynamically[2], a slice under attack may request additional resources after exhausting its own, and indirectly degrading the performance of the overall network system. Such behavior can unintentionally starve the underlying resource pool and degrade other slices' performance, creating cross-slice interference and SLA violations. Consequently, continuous monitoring[3], timely risk assessment, and the ability to restrict malicious or high-risk UEs from accessing sensitive slices are critical to preserving overall system stability and security.

To address these challenges, we propose an Autoencoder-Based Intrusion Detection System with Trust-Level Driven Slice Access Control (AEIDS-TLDSAC) framework. The design integrates two tightly coupled components: Autoencoder-based IDS and TLDSAC. First, we develop an IDS using an autoencoder to identify malicious traffic generated by UE. By deploying this Autoencoder-based IDS, the traffic from diverse UEs can be monitored on the slice, and malicious UEs can be detected as soon as possible. Secondly, we then propose a UE trust value evaluation model that follows a reputation-based mechanism[4]; the trust value is adjusted dynamically according to each UE's historical behavior and IDS outputs and serves as the basis for risk assessment. Finally, we integrate an access-control mechanism that sets slice-specific trust thresholds according to their security requirements; when a UE's trust value falls below the threshold, its access to that slice is restricted and it is redirected to a quarantine slice. By preventing high-risk UEs from accessing unavailable slices, proposed AEIDS-TLDSAC can enhance the overall stability and security of 5G network-slicing architectures.

The proposed approach delivers three key benefits. First, it improves detection coverage and accuracy by integrating Artificial Intelligence (AI) in IDS. Second, it offers a risk assessment mechanism for multi-tenant environments via reputation-based trust value management, ensuring that security actions reflect both short-term evidence and long-term behavior. Third, beyond merely conducting risk assessment, our approach also integrates risk assessment with slice-aware access control. By turning risk signals into concrete access decisions, the system localizes threats and blocks high-risk UEs from accessing critical slices. Overall, by coupling AE-based IDS with trust-driven, slice-aware access control, AEIDS-TLDSAC strengthens the security of 5G network-slicing architectures while preserving their flexibility to meet diverse service requirements.

The remainder of this paper is organized as follows. In Section II, we provide some background of our research and also review some related researches. In section III, we explain proposed AEIDS-TLDSAC in details, including its

979-8-3315-7896-1/26/$31.00 ©2026 IEEE

architecture and workflow. In Section IV, we conduct several experiments to evaluate the effectiveness of our work. Finally, we make a conclusion of our research in Section V.

II. Background and Related Works

This section introduces the research background and related works, including network slicing, intrusion detection system, reputation mechanism, and access control.

A. Network Slicing

By leveraging Software-Defined Networking (SDN) and Network Functions Virtualization (NFV), 5G/B5G networks can be divided into several logically isolated end-to-end slices. To achieve more adaptive orchestration and allocation of slice resources, [5] and [6] introduced a reinforcement learning–based scheme for dynamic slice resource management to deliver more reliable QoS. However, [7] indicated that the high dynamism and flexibility of network slicing make slices more vulnerable to malicious behaviors, with attacks potentially propagating easily across slices. Therefore, stronger slice isolation mechanisms are required, and AI-based IDS should be employed to monitor slice traffic and detect malicious activities at the earliest possible stage[8].

B. Intrusion Detection System

Intrusion Detection Systems monitor and analyze network traffic or system behavior to identify potential malicious activity. According to their detection methodology, IDS can be categorized as signature-based or anomaly-based[9]. The former relies on predefined rules; it is fast but limited because small deviations in traffic features can cause misses. The latter employs AI models to learn complex relationships among traffic patterns, offering greater flexibility, and is more effective against unknown and zero-day attacks. Reference [10] applied supervised learning (SL) models in 5G network slicing environments for intrusion detection, achieving high accuracy when sufficient labeled data are available. Reference [11] adopted signature-based pattern matching for slice-level intrusion detection. However, signature-based or supervised learning models generally struggle to recognize unknown attack types. In highly heterogeneous 5G environment, these limitations are pronounced, and maintaining fully labeled datasets is also costly. Therefore, we adopt an unsupervised Autoencoder–based IDS[12] that leverages reconstruction error for detection[13].

C. Reputation and Access Control Mechanism

Reputation-based security mechanisms quantify the trustworthiness of users by evaluating their historical behaviors and interaction records, thereby estimating a trust value. In multi-tenant 5G network slicing scenarios, the importance of such trust values is amplified. The slice trust model proposed in [14] introduces trust awareness to improve cross-slice security. Likewise, [15] argues that future 5G and beyond networks require quantifiable trust models to overcome vulnerabilities inherent in purely feature-based detection. Nevertheless, reputation-based mechanisms alone remain insufficient: while a trust value reflects potential risk, it often lacks tight integration with existing policy frameworks. To address this, our work draws upon the design principles of rule-based access control, including Role-Based Access Control (RBAC) and Risk-Adaptive Access Control (RAdAC), and combines them with the concept of quarantine slices from [16]. We treat the trust value as a risk signal that directly governs whether a UE may access a given slice, thereby strengthening the isolation of high-risk UEs.

While prior studies have contributed to intrusion detection, reputation mechanisms, and access control for network slicing, most focus on a single aspect. In contrast, proposed AEIDS-TLDSAC framework jointly integrates an Autoencoder-based IDS, dynamic trust evaluation, access control, and quarantine-slice, to more comprehensively enhance the security and isolation of network slices and prevent malicious UEs from degrading slice service quality.

III. System Architecture

In this section, proposed AEIDS-TLDSAC will be introduced in detail, including the overall system architecture and working flow. Section A will introduce the system architecture, Section B will introduce our Autoencoder-based IDS, and Section C will introduce TLDSAC mechanism.

A. System Architecture

AEIDS-TLDSAC consists of two components: The Autoencoder-based IDS and TLDSAC. The IDS employs an autoencoder to identify malicious traffic, thereby locating malicious users connected to a slice. Through the TLDSAC mechanism, it lowers the Trust Level (TL) to restrict the malicious UE's slice access privileges and isolate it. The system architecture of AEIDS-TLDSAC is shown in the Fig. 1. When a UE connects to the 5G core network via the gNodeB (gNB), the Access and Mobility Management Function (AMF), based on the user's trust level stored in the Unified Data Management Function (UDM), queries the Network Slice Selection Function (NSSF) for the slices the UE is allowed to use and then attaches the UE to the corresponding Network Slice Instance (NSI). If the UE's trust level is insufficient to access the requested slice, it is isolated to a quarantine slice to prevent high-risk UEs from entering to that unavailable slice. Each slice instance has its own Session Management Function (SMF) and User Plane Function (UPF), which respectively manage the PDU sessions established between the UE and the Data Network (DN) and forward the UE's packets to the DN. The IDS is deployed on the UPF, and it reports the results to the SMF as the basis for updating the UE's trust level.

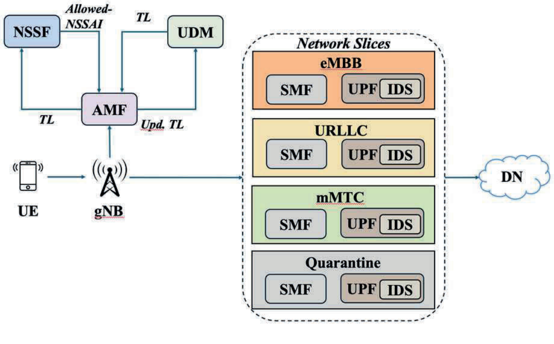

Fig. 1. AEIDS-TLDSAC system architecture

B. Autoencoder-Based Intrusion Detection

In Autoencoder-based IDS, we adopt an autoencoder as the intrusion detection model. By using unsupervised Autoencoder, this IDS can recognize attack types that are absent from the dataset. The core idea is to train the autoencoder solely on normal traffic so that it learns only the distribution of benign behavior. Leveraging the autoencoder's reconstruction property, it becomes proficient at reconstructing normal traffic but not malicious traffic. As a result, when encountering malicious traffic, even if the attack type has not appeared in the dataset, the autoencoder, having not learned how to reconstruct malicious data, usually produces a higher reconstruction error (RE), thereby enabling recognition between benign and malicious traffic. This makes the approach well suited to the highly diverse and dynamic environment of network slicing. The architecture and workflow of IDS are shown in Fig.2.

After network traffic is fed into the IDS, a preprocessing step, including Z-score normalization, label encoding, and outliers handling, transforms it into a form suitable for autoencoder inference. The processed data are then reconstructed by the autoencoder, and the Mean Squared Error (MSE) between the reconstructed data and the original inputs is computed. The RE is then compared against a predefined threshold. For threshold selection, a subset of the normal training data is held out as a validation set; following [17], the 95th percentile of the validation RE is used as the decision criterion. If the RE exceeds this threshold, it indicates that the autoencoder cannot effectively reconstruct the sample, and the flow is classified as malicious; otherwise, it is classified as normal.

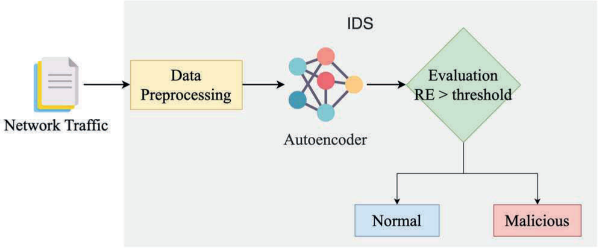

Fig. 2. Autoencoder-based IDS system architecture and workflow

C. Trust Level Driven Slice Access Control

To further enhance the overall security of the network slicing architecture, we design TDLSAC mechanism. The purpose of TLDSAC is to dynamically manage UE slice access permissions through risk-adaptive control and to apply isolation measures, thereby mitigating the impact of malicious UEs and maintaining the QoS of legitimate UEs. TLDSAC workflow is shown in Fig. 3.

When a UE initiates a connection request, if its trust level is below zero, it indicates the presence of excessive malicious behavior and a higher degree of risk; in such cases, the UE's connection will be blocked. If the trust level is greater than zero, the system further checks whether it meets the minimum trust level required by the requested slice. If the UE's trust level is insufficient, the UE will be redirected to a quarantine slice; otherwise, it is allowed to connect to the requested slice. During the connection, a tolerance value is maintained for each UE, representing the amount of malicious traffic that can be tolerated. Whenever malicious traffic is detected, this value will be decreased, and the system will evaluate whether tolerance value has exceeded the threshold. If the threshold is surpassed, the connection will be terminated, and the UE's trust level will be reduced. This design mitigates the impact of occasional IDS detection noise, preventing premature throttling or blocking of legitimate UEs.

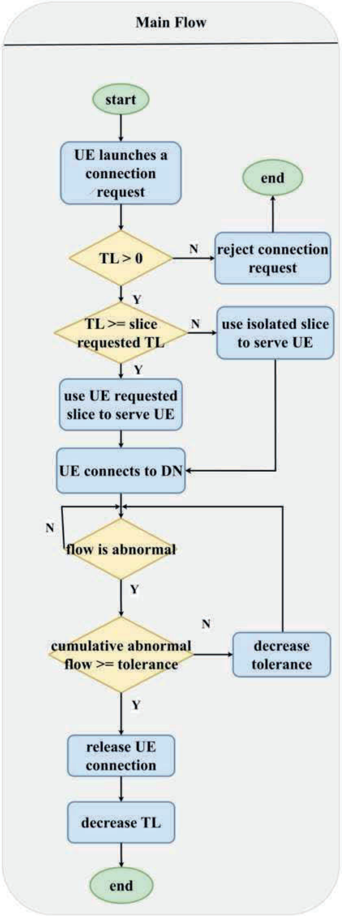
Main Flow
start
UE launches a connection request
end
TL > 0
N
reject connection request
Y
TL >= slice requested TL
N
use isolated slice to serve UE
Y
use UE requested slice to serve UE
UE connects to DN
N
flow is abnormal
Y
N
cumulative abnormal flow >= tolerance
decrease tolerance
Y
release UE connection
decrease TL
end

Fig. 3 TLDSAC workflow

IV. Experiments and Discussions

This section demonstrates the results and discussions of the experiments, including performance evaluation of Autoencoder-based IDS and effectiveness of TLDSAC.

A. Performance Evaluation of Autoencoder-Based IDS

The objective of this experiment is to evaluate the performance of Autoencoder-based IDS in detecting malicious attacks. Since the protocols used in 5G differ from wired network, two datasets constructed on 5G testbeds—5GNIDD[18] and WUSTL-HDRL[19]—were adopted. The results on the testing set are summarized in Table I and Table II. The results demonstrate that our Autoencoder-based IDS achieves 97.16% and 97.34% accuracy in 5GNIDD and WUSTL-HDRL respectively. Besides, this Autoencoder-based IDS also attains high precision, recall, and F1-scores across both normal and malicious classes, indicating that it can effectively detect malicious traffic. The distributions of the data categories of two datasets are shown in Fig. 4 and Fig. 5. In this experiment, 75% of the normal data were used for training, with 10% of the training data further set aside as a validation set. An additional 15% of the normal data were used to compute reconstruction errors, and the remaining 10% served as the testing set.

TABLE I. AUTOENCODER-BASED IDS PERFORMANCE ON 5GNIDD

Class	Metrics		
	Precision	***Recall***	***F1-score***
Normal	0.99	0.95	0.97
Malicious	0.95	0.99	0.97
Accuracy	97.16%		

TABLE II. AUTOENCODER-BASED IDS PERFORMANCE ON WUSTL-HDRL

Class	Metrics		
	Precision	***Recall***	***F1-score***
Normal	0.99	0.95	0.97
Malicious	0.95	0.99	0.97
Accuracy	97.34%		

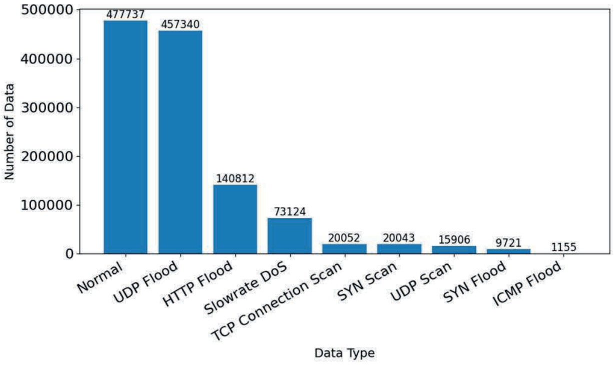

Fig. 4. 5GNIDD data distribution

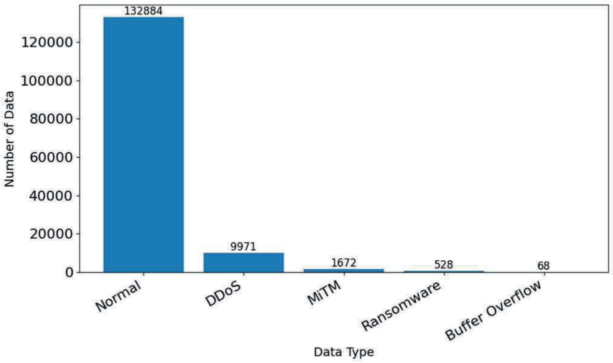

Fig. 5. WUSTL-HDRL data distribution

B. Performance Evaluation of TLDSAC

The objective of this experiment is to evaluate the effectiveness of the proposed TLDSAC in protecting legitimate UEs within network slices. The experiment lasted 150 seconds in total. During the first 30 seconds, 50 legitimate UEs accessed the slice, and 100 malicious UEs launched an attack at 30 seconds. In this experiment, the tolerance value and the trust value are set to 10, minimum required trust level of the slice is set to 8. These values can be adjusted to accommodate different deployment scenarios, parameter tuning is application-specific and is not pursued here. The average response time of legitimate UEs was measured, as shown in Fig. 6. The orange curve represents the baseline without any protection, the blue curve represents the TLDSAC adopting the Additive Increase Multiplicative Decrease (AIMD) tolerance and trust value adjustment strategy, and the green curve represents the Additive Increase Additive Decrease (AIAD) strategy. The results show that, after the attack begins at 30 seconds, the baseline fails to recover to the pre-attack state, whereas both AIMD and AIAD successfully isolate malicious UEs and restore normal service to legitimate UEs. Due to AIMD's faster adjustment compared to AIAD, AIMD recovered by approximately 50 seconds, while AIAD required about 80 seconds.

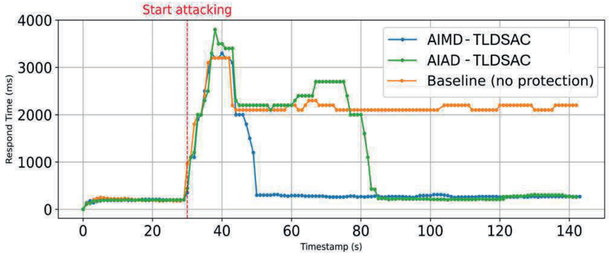

Fig. 6. Legitimate UEs average respond time

Further detailed results are presented in Table III. It shows that AIMD improves the average response time by 71.23% compared to the baseline and by 59.02% compared to AIAD, while AIAD achieves a 29.78% improvement over the baseline. For the average median response time, AIMD improves by 85.26% over the baseline and 46.30% over AIAD, while AIAD improves by 72.56% over the baseline. For average tail response time (99th percentile), AIMD improves by 63.37% over the baseline and 42.89% over AIAD, while AIAD improves by 35.86% over the baseline. While AIMD generally outperforms AIAD, its faster adjustments also result in 4.16% higher false positive rate than AIAD.

TABLE III. TLDSAC AND BASELINE PERFORMANCE COMPARISON

Method	Measurement		
	Average Respond Time (ms)	***Average Median Response Time (ms)***	***Average Tail Response Time (ms)***
AIMD-TLDSAC	268.53	157.16	717.36
AIAD-TLDSAC	655.32	292.64	1256.08
Baseline	933.30	1066.62	1958.45

V. CONCLUSION

The AEIDS-TLDSAC framework is designed to mitigate the vulnerabilities of the network slicing architecture and enhance slice security. Through Autoencoder-based IDS, malicious behaviors within network slices can be effectively monitored. Combined with TLDSAC, which dynamically adjusts each UE's trust value and employs it as the basis for risk assessment, TLDSAC controls UE access permissions to slices and isolates high-risk UEs into quarantine slices.

Experimental results show that our Autoencoder-based IDS achieves detection accuracies of 97.16% and 97.34% on two 5G datasets. The results also demonstrate that the AIMD-based TLDSAC delivers the best performance, with improvements of 71.23%, 85.26%, and 63.37% over the baseline in terms of average, median, and tail response time, respectively. Although AIMD achieves better overall performance, it also results in a 4.16% higher rate of misclassifying legitimate UEs as malicious. Thus, the choice between AIMD and AIAD ultimately depends on the specific application requirements.

Acknowledgment

This research was supported in part by the National Science and Technology Council, Taiwan, under Grant no. 114-2221-E-008-063-MY3, 113-2221-E-008-110-MY3, and 111-2221-E-008-063-MY3.

References

[1] I. Badmus, A. Laghrissi, M. Matinmikko-Blue, and A. Pouttu, “End-to-end Network Slice Architecture and Distribution Across 5G Micro-Operator Leveraging Multi-Domain and Multi-Tenancy,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 94, 2021.

[2] L. K. Mathew, "Dynamic Resource Allocation for Heterogeneous 5G TDD Wireless Networks: Balancing URLLC and eMBB Services," *2025 23rd Mediterranean Communication and Computer Networking Conference (MedComNet)*, Cagliari, Italy, 2025, pp. 1-5.

[3] Y.-C. Yu, C.-Y. Hung, and L.-D. Chou, "Kernel-level hidden rootkit detection based on eBPF," *Computers & Security*, vol. 157, 2025, Art. no. 104582.

[4] I. Ahmad, K.-L. A. Yau, M. H. Ling, and S. L. Keoh, "Trust and Reputation Management for Securing Collaboration in 5G Access Networks: The Road Ahead," *IEEE Access*, vol. 8, pp. 62542–62560, 2020.

[5] C. -C. Liu and L. -D. Chou. "B5G Network Slice Management via Staged Reinforcement Learning," *IEEE Access*, vol. 11, 2023.

[6] C. -Y. Yan, C. -J. Tai, and L. -D. Chou, "Slice Resource Management with MADDPG-Based Traffic Classification in 5G/B5G Networks," *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, Fukuoka, Japan, 2025, pp. 0253-0257.

[7] V. P. Singh, M. P. Singh, S. Hegde and M. Gupta, "Security in 5G Network Slices: Concerns and Opportunities," *IEEE Access*, vol. 12, pp. 52727-52743, 2024.

[8] J. Wang, Y. Li, J. Liu and N. Kato, "Intelligent Network Slicing for B5G and 6G: Resource Allocation, Service Provisioning, and Security," *IEEE Wireless Communications*, vol. 31, no. 3, pp. 271-277, June 2024.

[9] A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges," Cybersecurity, vol. 2, no. 1, pp. 1–22, 2019.

[10] Meshari Huwaytim Alanazi, "Machine Learning-based Secure 5G Network Slicing: A Systematic Literature Review," International Journal of Advanced Computer Science and Applications, vol. 14, no. 12, Jan. 2023.

[11] A. Jain, T. Singh, S. Kumar Sharma, and V. Prajapati, "Implementing Security in IoT Ecosystem Using 5G Network Slicing and Pattern Matched Intrusion Detection System: A Simulation Study," Interdisciplinary Journal of Information, Knowledge, and Management, vol. 16, pp. 001-038, 2021.

[12] H. -H. Lu, H. -L. Yeo and L. -D. Chou, "A Malicious Traffic Detection Model Based on Autoencoder and Multi-Head Attention Mechanism," *IEEE Internet of Things Journal*, 2025.

[13] A. Islam, S. -Y. Chang, J. Kim and J. Kim, "Anomaly Detection in 5G using Variational Autoencoders," *2024 Silicon Valley Cybersecurity Conference (SVCC)*, Seoul, Korea, Republic of, 2024, pp. 1-6.

[14] V. Varadharajan, K. K. Karmakar, U. Tupakula and M. Hitchens, "Toward a Trust Aware Network Slice-Based Service Provision in Virtualized Infrastructures," *IEEE Transactions on Network and Service Management*, vol. 19, no. 2, pp. 1065-1082, June 2022.

[15] J. M. J. Valero, P. M. S. Sánchez, M. G. Pérez, A. H. Celdrán, and G. M. Pérez, "Cutting-Edge Assets for Trust in 5G and Beyond: Requirements, State-of-the-Art, Trends & Challenges," *ACM Computing Surveys*, Nov. 2022.

[16] D. Candal-Ventureira, P. Fondo-Ferreiro, F. Gil-Castiñeira, and F. J. González-Castaño, "Quarantining Malicious IoT Devices in Intelligent Sliced Mobile Networks," *Sensors*, vol. 20, no. 18, Art. no. 5054, 2020.

[17] I. Ortega-Fernandez, M. Sestelo, J. C. Burguillo, and C. Piñón-Blanco, "Network Intrusion Detection System for DDoS Attacks in ICS Using Deep Autoencoders," *Wireless Networks*, vol. 30, no. 6, pp. 5059–5075, 2024.

[18] S. Samarakoon, Y. Siriwardhana, P. Porambage, M. Liyanage, S.-Y. Chang, J. Kim, J. Kim, and M. Ylianttila, "5G-NIDD: A Comprehensive Network Intrusion Detection Dataset Generated Over 5G Wireless Network," *IEEE Dataport*, Dec. 2, 2022.

[19] A. Ghubaish, Z. Yang, and R. Jain, "HDRL-IDS: A Hybrid Deep Reinforcement Learning Intrusion Detection System for Enhancing the Security of Medical Applications in 5G Networks," *2024 International Conference on Smart Applications, Communications and Networking (SmartNets)*, Harrisonburg, VA, USA, 2024, pp. 1-6.