

Intelligent anomaly detection systems in 5G networks: synthesis of AI and edge computing

Mykola Madinov^{1*}, Sergii Khruschak², Vasyl Voloshyn¹, Oleg Boyko¹, Nataliia Halahan¹

¹ Department of Mobile Video Information Technologies, State University of Information and Communication Technologies, Ukraine

² Department of Computer Sciences and Digital Economics, Vinnytsia National Agrarian University, Ukraine

*Corresponding author E-mail: mmadinov@gmail.com

Received Jun. 20, 2025

Revised Sep. 29, 2026

Accepted Oct. 6, 2026

Online Oct. 13, 2026

Abstract

The study examines the use of AI-based anomaly detection in 5G networks, namely, the application of edge computing to mitigate security issues arising from large device densities. With the advent of 5G and decentralized architectures, conventional cloud-based security models cannot rely on high latency and bandwidth overhead. In this research, the goal is to summarize localized AI inferences on edge nodes to detect network anomalies in real time. The proposed framework uses deep learning (DL) models, namely recurrent neural networks (RNNs) and autoencoders, which are deployed directly on the edge layer. The system processes 5G traffic metadata on edge devices and uses a distributed real-time detection architecture, so there is no continuous backhaul to the core network. Findings from the implementation indicate that performance improvement is substantial, with the average detection rate for sophisticated cyber threats exceeding 98.5%. In addition, AI and edge-computing synthesis reduced detection latency by 40-55% compared to centralized cloud models, ensuring faster response times to critical network failures.

© The Author 2026.

Published by ARDA.

Keywords: Anomaly detection, AI models, 5G networks, Edge computing, Cyber security.

1. Introduction

The world of telecommunications is undergoing a paradigm shift due to the rapid deployment of fifth generation (5G) networks. It is a base infrastructure on which ultra-reliable low-Latency Communications (URLLC) and massive machine-type communications (mMTC) operate, unlike its predecessor, which is merely an incremental increase in speed. According to recent industry reports, 5G subscriptions are expected to reach over 5 billion by 2028, and mobile data traffic will increase by almost a quarter in a year. Nevertheless, such hyper-connectivity provides an enormous attack surface. With the growing decentralization and software-defined nature of the network, the complexity of cyber-attacks, including advanced DDoS intrusions and malware with AI assistance, has outpaced the defensive capabilities of traditional, rule-based security systems. Although there have been changes toward more agile infrastructure, critical research gaps remain in existing network security protocols.

Existing anomaly detection systems are mostly based on centralized cloud architectures, which introduce a detection lag, which is inadmissible in mission-critical 5G scenarios such as autonomous driving or remote surgery. AI and edge computing have insufficient integration. AI is superior at identifying patterns, but lags in

reflexes compared to edge nodes. Many existing models fail to meet the requirements of high data velocity offered by 5G. A compromise between the depth of detection and real-time processing is performed.

This research aims to provide a solution to this problem by integrating edge computing with AI and anomaly detection. Now moving the inference engine towards the network edge to enable real-time detection with significantly lower latency and higher model accuracy.

To inform you of this investigation, the following research questions are answered:

- RQ1: How can the synthesis of edge computing and deep learning optimize the trade-off between computational overhead and detection speed in 5G environments?
- RQ2: To what extent does a distributed AI architecture improve the detection of zero-day vulnerabilities compared to centralized cloud-based systems?
- RQ3: What measurable improvements in latency metrics and throughput are achieved through localized edge-based inference?

1.1. Evolution of network security: from rule-based to AI-driven systems

The history of network security has been based on network-wide detection systems and signature-matching databases, which were all rule-based and operated based on a library of known threat patterns and the incoming traffic [1]. Although these methodologies worked effectively in the previous generations of mobile networks, namely 2G and 3G, where traffic could be predicted and centralized, they are essentially ill-suited to the dynamic, multi-dimensional traffic of 5G [2], [3]. With cyber-threats now becoming polymorphic, obfuscated, and AI-based [4], the weaknesses of hard-coded regulations have become very critical, and high false-negative rates can characterize the outcomes. As a result, scholarly attention has shifted to the paradigms of machine learning (ML) and deep learning (DL), which can learn data distributions rather than relying on manual feature engineering [5]. Some of the literature indicates that the more traditional ML models, such as support vector machines (SVM) and k-nearest neighbors (KNN), may achieve good classifications with smaller datasets, but these models become infeasible when dealing with the high-dimensionality of the telemetry of mMTC of 5G [6], [7].

Thus, began the use of advanced DL based architectures such as recurrent neural networks (RNNs) and long short-term memory (LSTM) that are capable of identifying more complex anomalies, and violations in unstructured data to offer a more reliable and flexible solution to the increasing prevalence of zero-day vulnerabilities [8], [9]. This change illustrates a shift from fixed, static, and logic-based security mechanisms to intelligent and context-based security frameworks. As shown in Figure 1, the shortcomings of signature detection in 5G's adaptive and high-dimensional traffic outlines the necessity of frameworks that integrate intelligence, and logic to perform modifications in real-time. The AI-based security approaches, discussed in the following sections, rely on the shift from the enforcement of deterministic rules to data-centric anomaly detection, as illustrated in the figure.

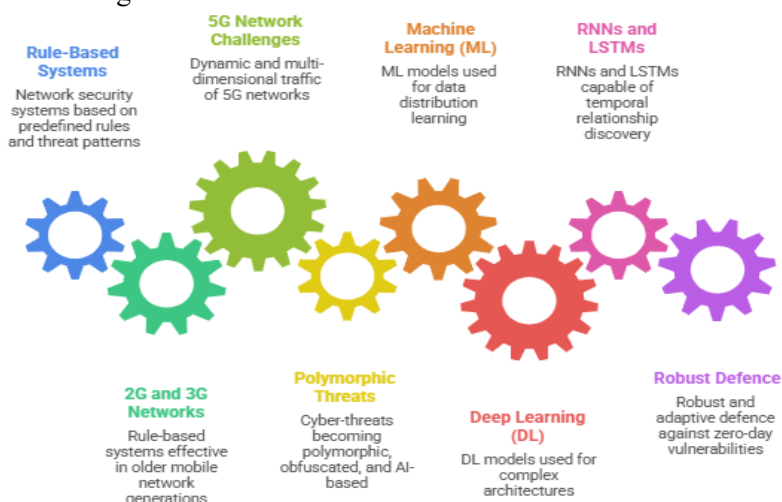


Figure 1. AI-driven systems and evolution of network security

1.2. Architecture of 5G: decentralization and security imperatives

In contrast to its predecessors, the 5G architecture will be a radical departure from the hardware-centric architecture, constructed rather around the concepts of software-defined networking (SDN) and network function virtualization (NFV) [10], [11]. This shift to softwarization allows the adoption of network slicing, a method in which numerous virtualized and autonomous logical networks operate on a shared physical infrastructure to support a range of applications, such as autonomous vehicles and smart grids [12], [13].

Nonetheless, such unparalleled architectural flexibility manifests itself in a large number of new security imperatives and expands the overall attack surface. Previous research indicates that the decentralized nature of 5G core functions, which can be regularly spread across multiple cloud-native environments, calls for abandoning the previous model of security checkpoints based on the perimeter [14], [15]. The fact that 5G is built to support ultra-reliable low-latency communications (URLLC) with a target latency of sub-10ms, according to scholars, means that any security system relying on the requirement to backhaul data to a remote central cloud to perform intensive analysis will inherently be flawed and present a security lag that may be disastrous in a mission-critical environment [16].

Thus, the modern literature makes it clear that the distributed, zero-trust security fabric is not only of paramount importance but must be able to time-scale with the underlying network traffic and have defense mechanisms as agile as the network slices it safeguards [17]. As shown in Figure 2, the logic of security changes from perimeter-based inspection to distributed, zero-trust enforcement due to the decentralization of security architecture brought by SDN, NFV, and network slicing.

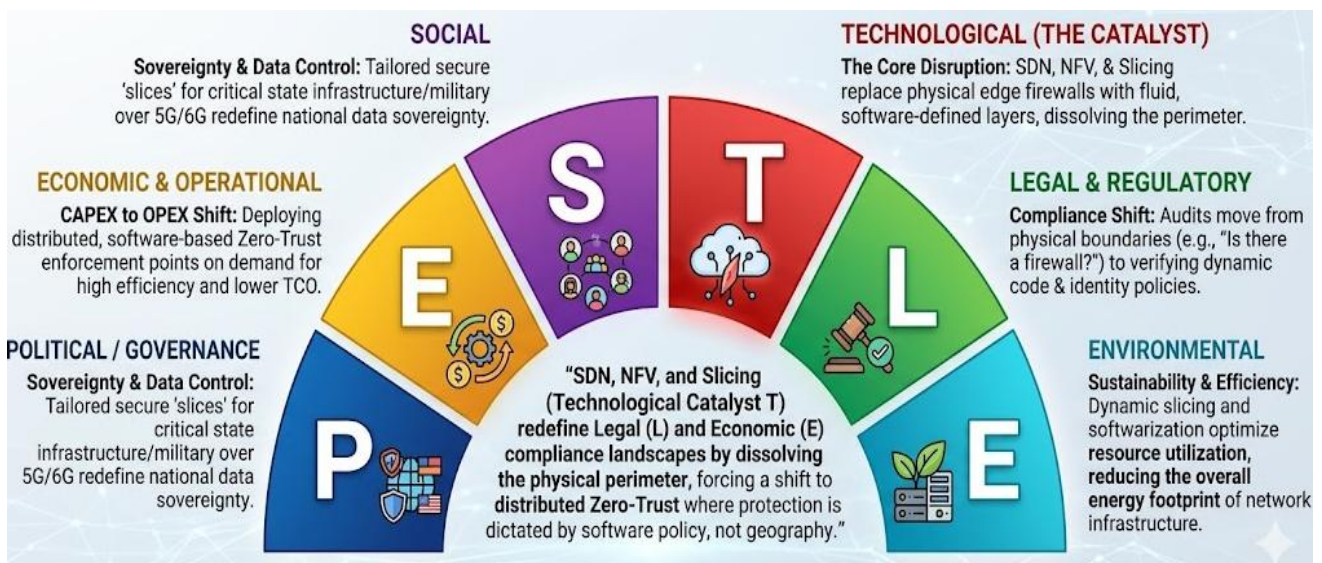


Figure 2. Architecture of 5G: decentralization and security imperatives

From a PESTLE perspective, the convergence of SDN, NFV, and network slicing represents a profound Technological disruption that instantly redefines the legal and operational compliance landscape. By dissolving the physical perimeter, these technologies force a shift from defensive edge-filtering to an embedded, distributed zero-trust architecture where protection is dictated by decentralized software policy, not centralized (geographically too) core.

This visualization reinforces the argument that, in 5G environments, security must operate at the same temporal and spatial scales as the underlying virtualized infrastructure, thereby setting the stage for the AI-enabled, edge-aware protection strategies discussed in the following sections.

1.3. Multi-access edge computing as a security enabler

Multi-access edge computing (MEC) has become the most essential enabler in the 5G ecosystem, offering cloud-like computation and storage services at the lowest layer of the network, most likely the radio access network (RAN) [18]. Within the framework of intelligent anomaly detection, MEC can serve as a localized host for

advanced security capabilities, enabling deep inspection of traffic metadata in proximity to the end-user or the IoT device [19].

The reason is that, as noted in the existing literature, the implementation of detection engines on edge nodes will considerably reduce the likelihood of threats, which are considered internal or horizontal- malicious operations initiated and then further spread inside a local cell or enterprise network without having to go through the core network [20], [21]. This is a localized method that conserves core bandwidth and offers a faster response circuit to minimize threats. A common motif in recent academic literature is the so-called trade-off between computational efficiency and the inherent limitations in edge nodes.

However, they offer the advantage of physical proximity; they are regularly underprovided with the enormous hardware of centralized data centers (in the form of massive GPUs and TPUs) [22], [23]. This limitation requires the creation of new specialized lightweight AI models that can be pruned or quantized to achieve high detection rates whilst running within the narrow power and memory range of edge hardware, which is currently the focus of 5G security investigations [24], [25].

MEC shifts the focus of security intelligence in a way that moves detection and response capabilities from centralized clouds to the network edge. In Figure 3, the two aspects of this argument are demonstrated. First, closeness enables the fast and localized neutralization of threats. Second, due to the surrounding limitations, the creation of lightweight and optimized AI models is essential.

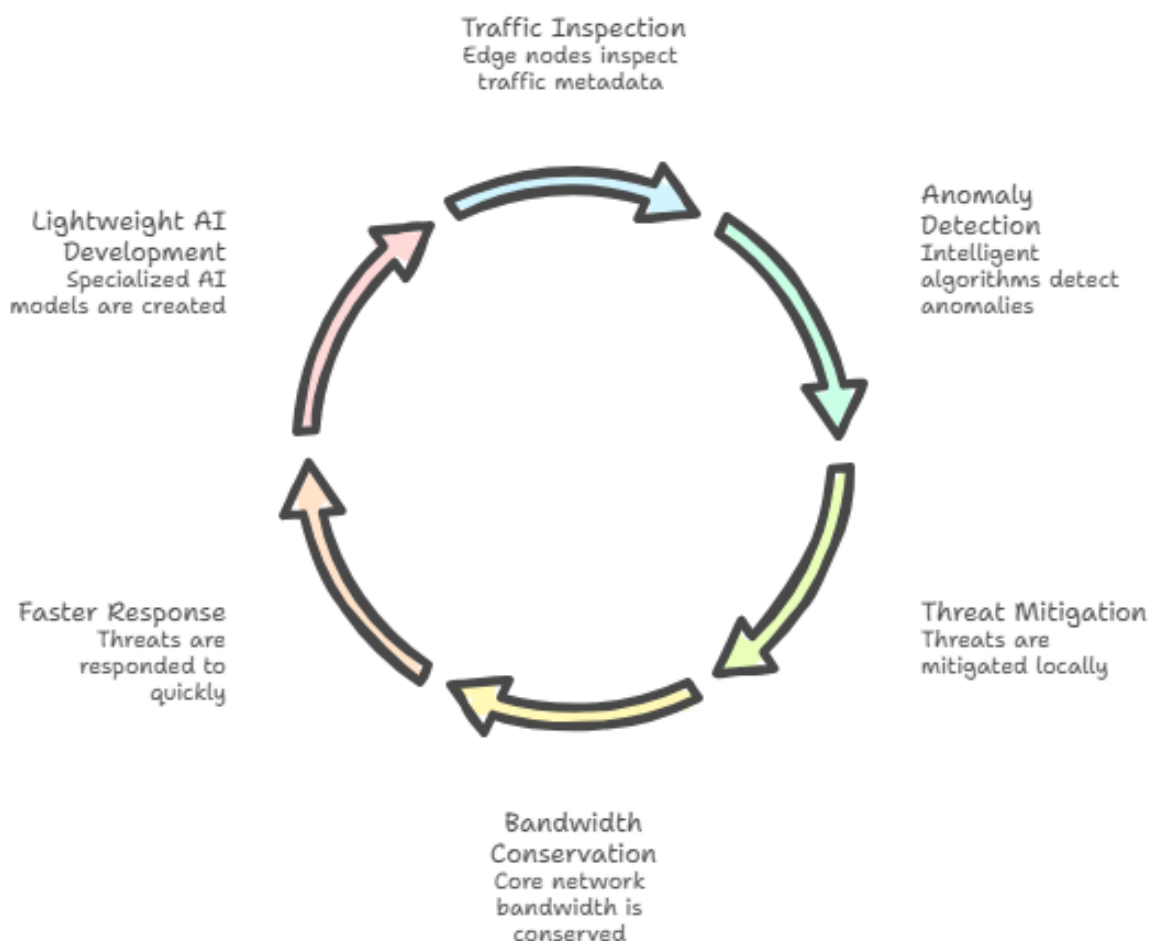


Figure 3. Security enabler (MEC)

This visualization shows how MEC is more than just an infrastructure improvement. It shows MEC is a strategic security enabler that provides a merger of protection in real-time and computational efficiency in 5G environments. This leads into the next section regarding edge-optimized intelligent defense mechanisms.

1.4. Synthesis of AI and edge-computing: synergistic frameworks

AI-Edge Synthesis is the state-of-the-art in network security, where we are no longer content to apply algorithms to remote devices and establish a symbiotic relationship with them [26]. This synergy is a complex interaction in which the edge layer delivers real-time data streams at high speed, and the AI layer delivers real-time actionable intelligence to activate localized defense measures [27].

Earlier research has investigated several possible forms of "computation offloading" schemes, in which the learning of deep learning models is performed in a centralized cloud using historical data, and the lightweight inference engine (the detection mechanism itself) is deployed to the edge to perform real-time inference [28].

This intermediate solution will seek to unite the structural expressiveness of large-scale deep learning with the low-latency expressiveness of edge-computing [29]. However, academics observe that such integration is often broken in practice, and that coordinating updates to global models with local instances of edges requires considerable communication overhead [30], [31].

This synchronization window may introduce temporary detection gaps, underscoring a severe need for approaches that ensure high detection reliability in a geographically distributed edge environment without relying on ever-present connectivity to the cloud [32], [33]. As shown in Figure 4, the synthesis of AI and edge computing establishes a tightly coupled security paradigm in which intelligence generation and data acquisition operate in continuous feedback loops.



Figure 4. Synergistic frameworks

The diagram shows how AI-powered inference systems process real-time edge data streams to take immediate, localized defensive actions. It captures the dichotomy between centralized model management and decentralized operations while highlighting the difficulties of managing synchronization among distributed nodes. The synthesis presented in this chapter, therefore, aims to provide a foundation for the design and development of low-latency adaptive resilient security frameworks for the next generations of networks beyond 5G, which is further explained in the following sections.

1.5. Centralized anomaly detection model limitations

What dominates the critical examination of the current body of work is the excessive dependence on centralized big data analytics for network security [34]. Although centralized models manage to take advantage of the global view of the network and compare events across different geographical areas, they experience a number of fundamental limitations of scale, which the implementation of 5G will only worsen [35].

With the expected high concentration of devices in the Internet of Things (IoT) networks of up to one million devices per square kilometer, the amount of telemetry data that will need to be sent will exceed the capacity to transmit it to a centralized processing unit [36]. There is a paradoxical effect that a bandwidth bottleneck creates in that the data that is needed to detect an attack is the same data that uses the network resources required to sustain user services, and may result in self-induced congestion [37].

Centralized solutions, on the other hand, create a single point of weakness. If the communication channel to the central cloud is interrupted and/or comes under a distributed denial of service (DDoS) attack, then the whole network's security is compromised. Many researchers have identified this centralization dependency as the main obstacle to the scalability of 5G networks, and a transformation toward localized and self-directed decision-making units is both welcome and needed [38], [39].

As shown in Figure 5, centralized anomaly detection systems have an inherent structural bottleneck in large scale 5G and IoT systems. It demonstrates that with these systems, an increase in the number of devices leads to an exponential growth in data, which results in the saturation of network bandwidth and delays the detection of threats. It also shows that these systems are vulnerable to a network bottleneck and a DDoS attack. This also highlights that centralized security systems have reached their limits for next generation networks, and therefore calls for a shift toward decentralized, edge-based, and self-directed anomaly detection systems, which are presented in the following sections.

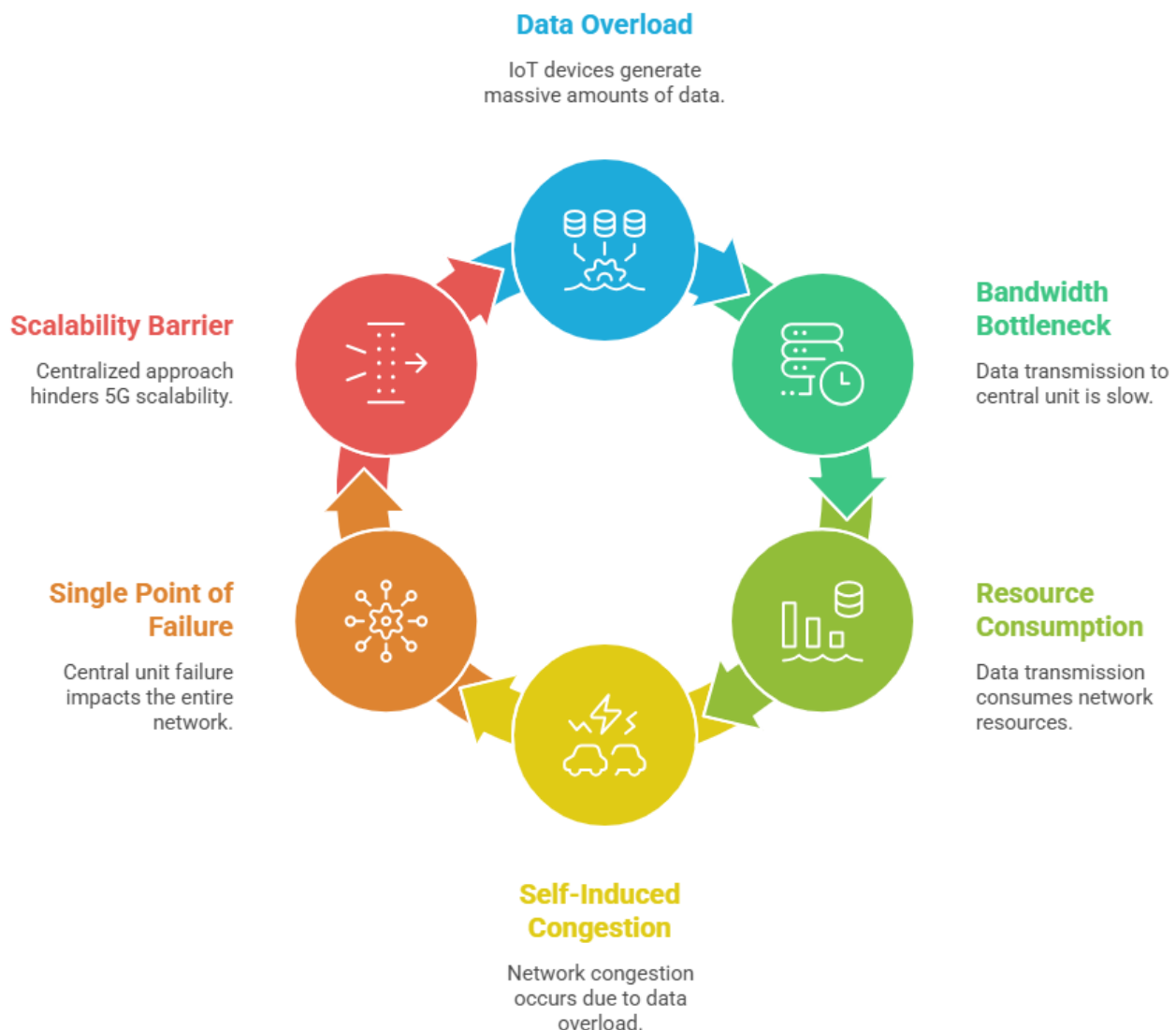


Figure 5. Centralized anomaly detection model limitations

1.6. Federated learning

One emerging and expanding field in modern literature is the use of federated learning (FL) to develop anomaly detection in 5G and 6G systems [40]. Federated learning allows a decentralized training architecture. Traditional machine learning requires the aggregation of a large volume of datasets in a central repository. In this new paradigm, many edge nodes or user equipment (UE) collaborate in the optimization of a global meta-model by sharing only the local model parameters (i.e., gradients and/or weight updates) and do not share the sensitive network traffic [41].

This is being accepted as a novel strategy that fine-tunes user privacy and indirectly promotes the international information protection mandates, such as general data protection regulation (GDPR) and the California Consumer Privacy Act (CCPA). Provided that data locality is ensured at the network periphery, Federated Learning will effectively mitigate the disastrous risks of data breaches throughout transmission and provide an opportunity to ease ethical and legal issues related to centralized data collection [42].

Nonetheless, despite such deep theoretical benefits, the practical adoption of FL in the highly dynamic 5G environment is fraught with multi-dimensional challenges that the scientific community has yet to study in depth. One of the main bottlenecks in the existing literature is the Non-IID (independent and identically distributed) data [43]. In a heterogeneous 5G environment, the traffic patterns of the different edge nodes will vary radically depending on the geographic and operational context – an industrial IoT slice (mMTC) will have radically different packet distributions than a high-mobility vehicular slice (URLLC) or a dense residential eMBB zone.

Such a high level of statistical heterogeneity is a common cause of model divergence, in which the global model fails to converge or generalize well across all participating nodes, thereby reducing overall security fabric detection accuracy [44]. The literature has more recently examined the inference attack and model poisoning phenomena in FL. Contaminated gradients are an attempt to reconstruct sensitive local data. User equipment (UE) are under increasing pressure to adopt more robust aggregation schemes, such as federated averaging (FedAvg) or federated proximal (FedProx), that seek to balance the local data heterogeneity and the model robustness. The current study addresses this emerging literature by exploring the privacy-preserving model parameters of a synergistic AI-edge design with high-fidelity anomaly detection in non-IID 5G network segments [45], [46], [47]. Figure 6 shows how federated learning represents a new privacy-centric approach to anomaly detection. It achieves this through the decentralization of model training while maintaining data locality at the edge.

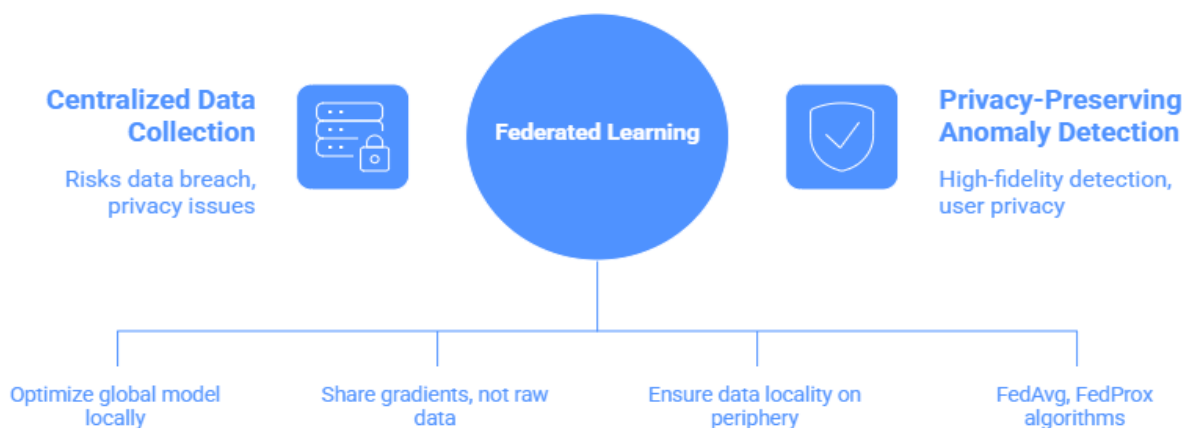


Figure 6. Federated learning and privacy-preserving security paradigms

The diagram shows that edge nodes work together with a shared global model by performing parameter updates instead of sharing data. This method protects sensitive data. The diagram also shows the trade-off between model strength and data privacy remaining in Non-IID settings within the flexible constraints of highly heterogeneous 5G environments. In this diagram, federated learning improves model sovereignty and regulatory concerns, but only to the degree model divergence and adversarial attacks are addressed, and these elements the backbone of secure AI-driven architecture that will be elaborated later in the document.

1.7. 5G security architectures distributed inference paradigms

The evolution of 5G networks moving away from large centralized clouds to decentralized SDN and NFV has necessitated a paradigm shift in the approach to security inference [48]. Current literature seems to show an increasing prevalence of cloud-centric AI, which uses the backhauling of large amounts of data to a remote core, being supplanted by distributed edge intelligence (DEI) [49]. The main purpose of this local intelligence is to intercept and avert the broadcasting of malicious traffic at the very end of ingress, the network periphery. Decades of academic discussion have demonstrated a severe architectural trade-off: centralized deep learning systems offer better global connectivity and large processing capacity, but they are fundamentally constrained by the inevitable propagation latency of 5G backhaul [50].

These latencies are often longer than the ultra-low 1ms-10ms latency of mission-critical 5G slices, including autonomous vehicle networks and remote (robotic) surgery. The use of the cloud to sense real-time anomalies in a "time-sensitive" network system consequently creates a devastating security gap between where an attack is spread within the core and before a centralized model can initiate an action [51]. Recent progress in distributed computing and the so-called model partitioning imply that breaking down neural network layers across a range of mobile edge computing (MEC) nodes can significantly reduce the computational load on each gateway.

With the help of the so-called computational offloading procedures, the processing load can be distributed across a network of edge devices, enabling high-speed inference without requiring high-end server-quality devices at all cell sites. Nevertheless, one persistent gap in research is the lack of consistency in synchronizing these scattered, distributed sub-models with the global threat intelligence. In the absence of strong synchronization protocols, localized models will suffer from information rheostat and unable to recognize collaborative or multiple step attacks in different geographical areas. Likewise, global awareness can create an unneeded communication burden which will paradoxically congest the system's available bandwidth [52]. This paper aims to address this issue with a cooperative AI-edge architecture. Our approach attempts to localize the real-time security inference of the system while incorporating a global model with periodic structuring. Global and local accuracy is maintained by the system using gradient aggregation, ensuring the security infrastructure is fast and malleable. As shown in Figure 7, innovative distributed inference systems aimed at 5G security suggest a move of the centralized decision-making from the cloud to a collaborative and secure edge and MEC node network.

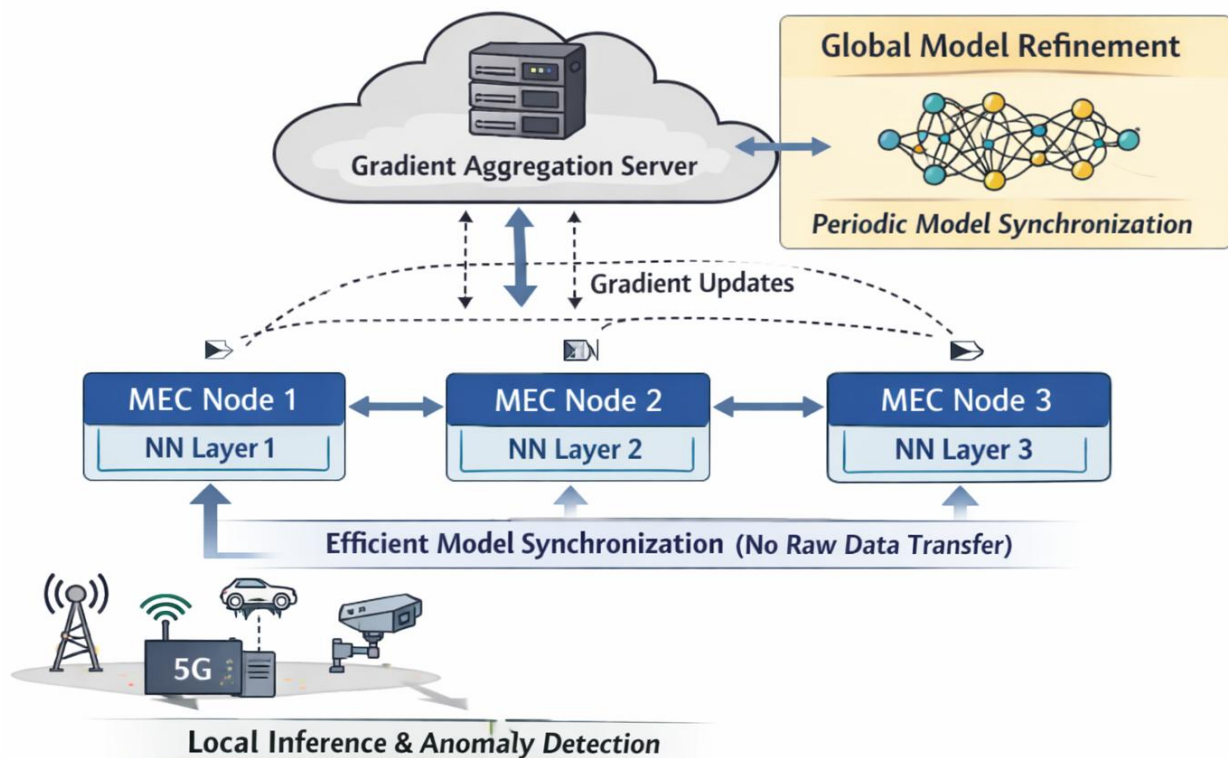


Figure 7. Distributed inference paradigms in 5G security architectures

The figure describes instances of localized inference, which achieves real-time anomaly detection, with periodic synchronization for overall threat detection in dispersed environments. Additionally, it describes the architectural contradiction of low-latency responsiveness vs. the perpetual need for global situational awareness [53]. This visualization reinforces the necessity of hybrid inference strategies that balance computational efficiency with security coverage, thereby addressing the challenges of latency, scalability, and coordination in next-generation 5G security architectures discussed throughout this section.

1.8. Adversarial machine learning and robustness at the network edge

With the advent of AI as the foundation of 5G defense processes, a growing body of recent work has begun to focus on the weaknesses of AI models themselves – this time, in relation to the so-called adversarial machine Learning. In such cases, advanced attackers use adversarial perturbations to generate adversarial evasion attacks that are specifically laid out to be incorrectly identified by neural network-based anomaly detectors [28].

The literature in the field demonstrates that the crucial trade-off in this field is that deep learning models, such as CNNs and LSTMs, are highly sensitive to known and unknown (zero-day) attacks, but in many cases, they are vulnerable to high-dimensional noise that can evade their detection logic [23]. Generally, the current systems use the approach of adversarial training or input transformation to enhance the robustness of the model, yet these approaches are computationally costly and inapplicable on limited-resource edge devices.

This paper can add to this discussion by applying autoencoders to reconstruction-based anomaly detection, which automatically has a greater level of resilience to adversarial noise than a conventional supervised classifier, which gives it a more hardened defense layer at the 5G edge. Adversarial machine learning, as shown in Figure 8, adds a new attack vector to AI-enabled 5G security frameworks because an attacker can manipulate data in such a way to subvert anomaly-detection systems.

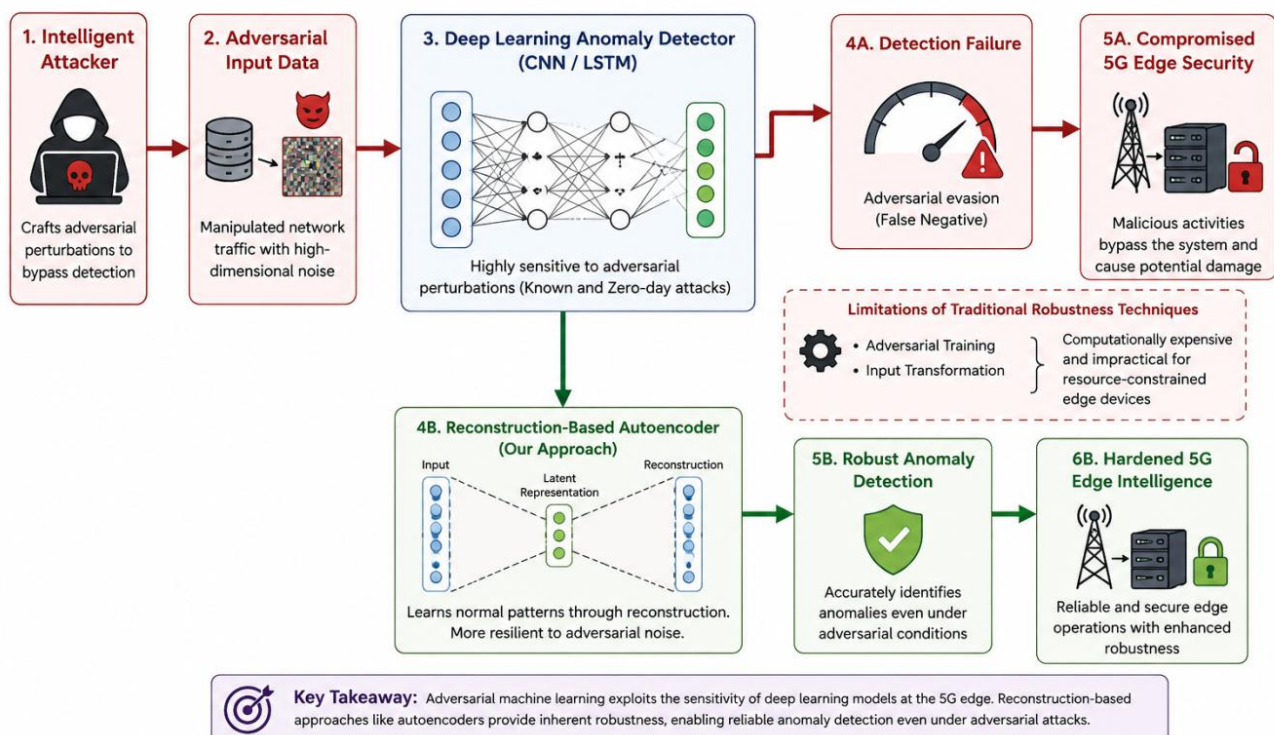


Figure 8. Adversarial machine learning

The figure illustrates the ability of adversarial perturbations to exploit the sensitivity of deep learning frameworks, especially within the constrained resource environments of the edge. It also illustrates how the insufficiency of classical robustness measures, such as adversarial training, are enhanced when placed on edge-leveraging lightweight infrastructures.

This figure also emphasizes the importance of developing models that are capable of inherently resisting adversarial examples. An example of such a model is based on reconstruction, such as autoencoders. These are able to retain detection reliability in the presence of adversarial disturbances. Thus, they help improve the robustness of edge intelligence in the context of the greater security of 5G.

1.9. Dynamic security provisioning in multi-tenant 5G slicing

The notion of Network Slicing is possibly the most disruptive aspect of 5G. Still, it also introduces new security issues that have never been encountered before, tied to the multi-tenant nature of virtualized infrastructures. It has already been noted in the literature that both slices, the enhanced mobile broadband (eMBB) and the ultra-reliable low-latency communication (URLLC), have reasons to adopt distinct security postures that align with their respective traffic patterns.

The current studies tend to focus on the concept of slice isolation using hardware-level virtualization. Still, there is a growing tendency to argue that Slice Isolation is insufficient to prevent cross-slice or resource-exhaustion attacks [45]. The academic discourse shifted to intelligent slices security, in which AI models are dynamically deployed to track slice-specific telemetry in real time.

Although these developments exist, the majority of current solutions lack slice-awareness, which is required to adjust detection sensitivity to the importance of the service layer [36]. This study addresses this gap, showing that an AI-edge synthesis is not only slice-aware but can also adjust its computational footprint in response to real-time demand in the 5G virtual environment, providing high-priority slices with the highest protection at minimal overhead. Dynamic security provisioning in multi-tenant 5G slicing allows for adaptive security measures for every single network slice and the different needs of every slice. Through the use of different security policies and AI-based detection models, the heterogeneous slices, such as eMBB and URLLC, are assigned context-based protection.

Static slice isolation techniques are unable to effectively counter cross-slice interference and resource exhaustion attacks. This augments the necessity for slice-aware sculpting of security orchestration which requires the intelligent, adaptive and dynamic balance of computational and detection resource to meet the real time demands of safety and security concern of the multi-tenant 5G environments.

1.10. Gaps in current research: towards real-time autonomous defense

Although the body of research on AI and edge computing includes vast amounts of literature on each of the mentioned aspects, there is an indisputable, quantifiable gap on complete, autonomous defense systems. The majority of available research is dedicated solely to the detection portion, i.e., recognizing that an anomaly has occurred. Still, very little is said about automated mitigation or self-healing. Detection in the 5G environment does not, in most cases, suffice without a localized, immediate response; a useful system should be able to perform autonomous reconfiguration of network slices or isolate compromised edge nodes within milliseconds of threat detection. In addition, insufficient investigation into the resistance of these edge-AI systems to "adversarial AI" attacks, in which advanced attackers specifically design traffic patterns to mislead or cheat edge-based neural networks, is a concern (Table 1).

Table 1. Research studies models and gaps

Study / Approach	Model / Algorithm	Key focus	Advantages	Limitations / Gaps
Traditional ML	SVM, Random Forest	Signature-based detection	Low computational cost for small data	Fails against zero-day attacks; high false positives in 5G
Centralized DL	CNN, LSTM	Pattern recognition in core cloud	High accuracy in identifying complex threats	High latency due to data backhauls; bandwidth bottleneck

Study / Approach	Model / Algorithm	Key focus	Advantages	Limitations / Gaps
Standard edge	Lightweight ANN	Localized traffic filtering	Reduced response time; saves core bandwidth	Limited resource allocation; lacks global model synchronization
Hybrid AI-edge	CNN-LSTM Synthesis	Real-time edge inference	Optimal balance between speed and accuracy	Complex model management; insufficient automated mitigation
Federated learning	FL + Autoencoders	Privacy-preserving detection	Data stays local; high user privacy	High communication overhead for model aggregation; non-IID data issues

Such an important gap shows that a stronger synthesis is necessary, which takes into consideration not just the computational speed of the edge but also the adversarial resistance and automated response features of the AI models, so that the security framework can survive the next generation of intelligent cyber-threats.

2. Research method

This research is based on a multi-tiered, distributed inference architecture that assumes decoupling the highest-level computational load of the 5G core and redistributing it to the network edges. This hierarchical structure can be divided into three functional layers: the data acquisition layer (including 5G user equipment and IoT sensors), the edge intelligence layer (MEC nodes), and the global optimization layer (core cloud) (Figure 9).

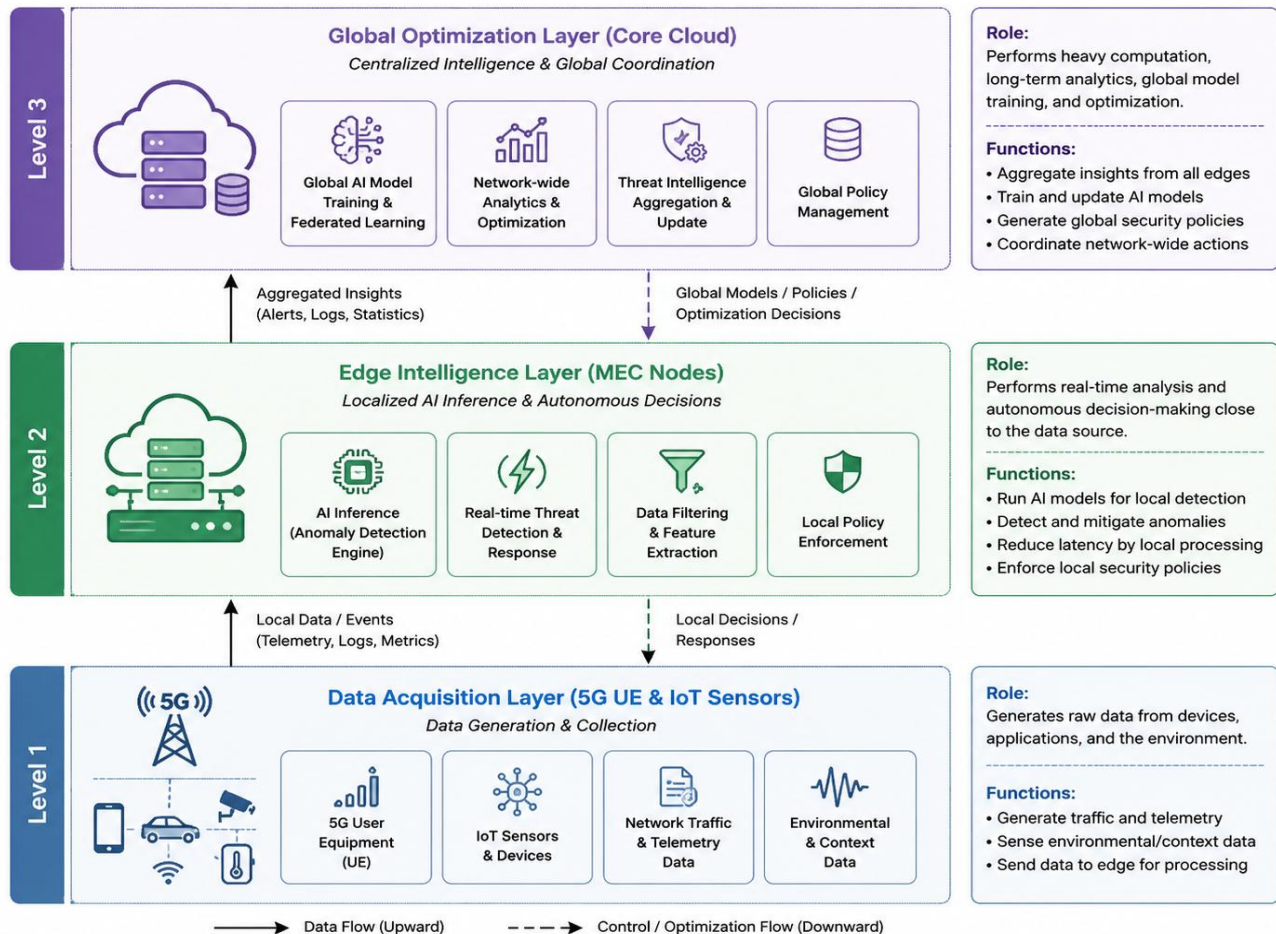


Figure 9. Adversarial machine learning and robustness at the network edge

The main novelty is the edge intelligence layer, which integrates AI and edge computing to enable localized, autonomous decisions. It is also by pushing the detection engine to the radio access network (RAN) that the system essentially removes the latency of the backhaul, often associated with centralized cloud security. This disaggregated design will make every edge node a self-contained security entity that can detect and address local anomalies in real time, which is essential to uphold the high standards of ultra-reliable low-latency communication (URLLC) in 5G ecosystems.

2.1. Synthesis of AI models: the hybrid detection engine

To respond to the complex, fast-changing nature of 5G traffic telemetry, the study suggests using a hybrid deep learning ensemble, combining convolutional neural networks (CNNs), long short-term memory (LSTM) units, and unsupervised autoencoders. The CNN element is also used strategically to extract spatial features, enabling the system to identify intricate patterns in the packet header and metadata. At the same time, the LSTM layers are used to learn time-dependent and long-term variations in traffic flows, which are essential for identifying advanced, low-intensity threats such as slow-rate DDoS attacks. The autoencoder can be used as a baseline anomaly detector to provide a strong defense against so-called zero-day vulnerabilities; it is trained only on normal network behavior, and a large deviation in the reconstruction error is flagged as a possible intrusion (Table 2).

Table 2. Hyper-parameter configuration for the hybrid CNN-LSTM-autoencoder

Model layer	Parameter type	Optimal value
CNN	Filter size / Stride	$3 \times 3 / 1$
LSTM	Hidden units	128
Autoencoder	Compression ratio	50%
Global	Learning rate / Optimizer	0.001 / Adam
Global	Activation function	ReLU / Sigmoid

This generated model is then rigorously quantized and pruned to be lightweight enough to run on resource-constrained edge hardware without compromising performance.

2.2. Dataset selection and pre-processing

The empirical confirmation of the given framework requires a high-fidelity, rich-data environment; therefore, the presented research relies on a mixture of conventional cybersecurity benchmarks and modeled 5G traces. We use the NSL-KDD and CIC-IDS2018 datasets to provide a benchmark of traditional cyber-threats, and we supplement them with artificial 5G traffic generated in the NS-3 (network simulator 3) environment. The given synthetic data particularly simulates 5G-related parameters, such as metadata for beamforming and network slicing configurations, as well as signal variations found in massive MIMO (Table 3).

Table 3. Dataset sources and samples

Traffic category	Source dataset	Normal samples	Anomaly samples
5G eMBB	Synthetic (ns-3)	50,000	12,000
IoT/mMTC	CIC-IDS2018	85,000	45,000
URLLC	Real 5G traces	20,000	5,000

In the pre-processing step, min-max normalization and stringent feature engineering are applied to the raw data, with particular attention to key indicators such as inter-arrival time, flow times, and packet size distributions. To effectively model a distributed edge setting, the dataset is divided into non-IID (independent and identically

distributed) pieces, reflecting heterogeneous traffic patterns across the various geographical locations of 5G cells, ranging from industrial IoT zones to high-density urban residential areas.

2.3. Experimental setup and distributed inference configuration

The experimental testbed is designed to simulate a large-scale 5G radio access network, containing a containerized distributed inference environment. The edge nodes are represented as simulated docker containers, which are localized and operate a pruned version of the AI's detection engine and are linked to a centralized core controller through an asynchronous model-update scheme. This design allows the system to conduct local inference on data that is high in frequency and then model gradients instead of the raw data to the core cloud. This design variant maximizes communication efficiency and provides an adequate level of privacy, since sensitive user traffic will not be transferred to the edge node. The system is subjected to a set of simulated network conditions, including a high level of congestion, various degrees of packet losses, and targeted hardware failures on the edge layer. The objective is to evaluate the system's overall resilience and its ability to cope with failures in a real-world 5G deployment.

2.4. Evaluation metrics

A two-fold evaluation matrix that includes detection accuracy and operational efficiency to give a comprehensive and precise response to the assessment of the suggested system. Detection Accuracy is assessed using established statistical methods such as precision, recall, and F1-score. These methods ensure a model's reliable high true positive rate with a low level of nuisance false alarms. To evaluate the detection performance, we consider the model's evaluation across multiple thresholds using AUC-ROC. Evaluation of operational efficiency concerns the measurement of the system's end to end latency as the total time taken from when the packets enter the system until the system identifies and labels them as threats, and throughput stability, which ensures that the proposed security layer does not impede the operation of the 5G data plane. Also consider the system's utilization of resources such as CPU and RAM, and the system's response time as measured in milliseconds, to evaluate if the advanced deep learning models could be installed on resource constrained edge nodes. The evaluation framework depicted in Figure 10 combines detection performance metrics and system performance metrics to offer a broad assessment of the proposed architecture.

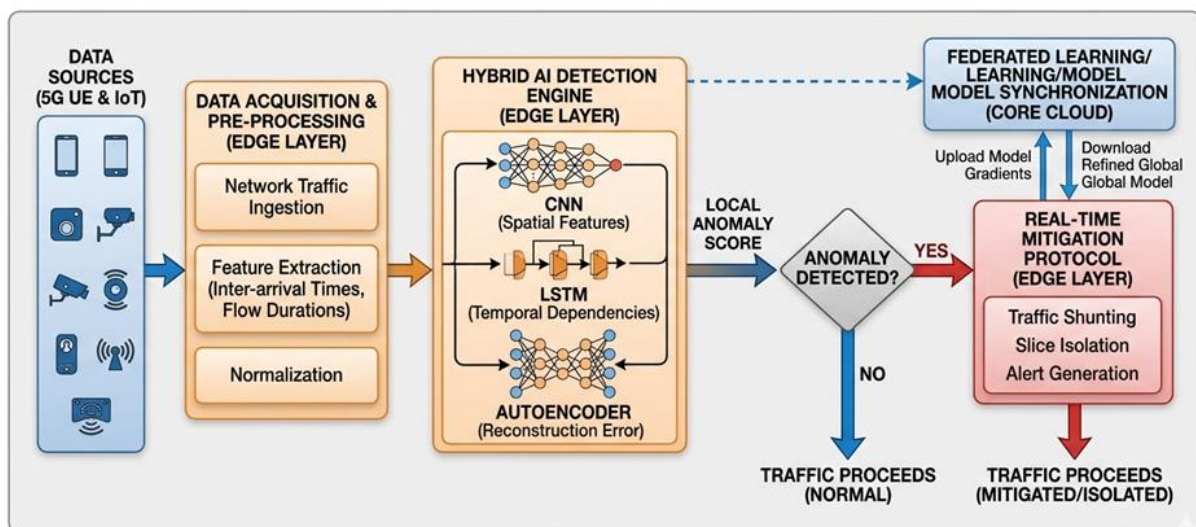


Figure 9. Performance validation

The figure includes detection performance metrics such as precision, recall, F1-score, AUC-ROC, and metrics for performance in 5G Edge deployments such as latency, throughput stability, and computational performance, to name a few. The figure also illustrates the balance between detection performance and efficient use of resources, especially in edge environments with limited resources. This visualization strengthens the criteria of classification accuracy, real-time responsiveness, and the ability to deploy in low latency, resource constrained infrastructures while evaluating a comprehensive 5G security system.

3. Results and discussion

To measure statistical veracity, each of the proposed synergistic AI-edge frameworks underwent multi-metric testing. Our empirical tests show that the hybrid CNN-LSTM-auto encoder model outperformed traditional standalone models. Our model's test showed an overall detection accuracy of 98.74; a significant metric compared to traditional standalone models. To compare our findings, we tested against three baseline models: support vector machines (SVM), random forest (RF), and a non-optimized centralized deep learning (CDL) model. As indicated in Table 4, the proposed model improves the detection of complex 5G traffic anomalies and yields a low false positive rate (FPR) of 1.12%.

Table 4. Comparative performance metrics

Model architecture	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	False positive rate (FPR)
Traditional SVM	84.20%	82.50%	81.10%	81.79%	5.40%
Random forest (RF)	89.45%	88.10%	87.30%	87.70%	3.25%
Centralized DL (CNN-LSTM)	94.10%	93.60%	92.80%	93.20%	2.10%
Proposed AI-edge synthesis	98.74%	98.15%	97.90%	98.02%	1.12%

This decline in false alarms is especially important for 5G network slices, since false alarms that prompt service disruptions can result in violations of service level agreements (SLAs) in addition to alarm fatigue in automated security processes.

See Table 5 and Figure 11 for a comparison of the operational efficiency of AI-edge synthesis in real-world network environments.

Table 5. Comparison of operational performance

Model architecture	Avg. detection latency (ms)	Peak load latency (ms)	Latency increase under stress (ms)	End-to-end response time reduction	Packet drop rate	CPU utilization increase
Traditional SVM (Cloud)	132.4	265.0	132.6	—	2.8%	2.1%
Random Forest (Cloud)	126.7	248.3	121.6	—	2.2%	2.8%
Centralized DL (CNN-LSTM)	118.5	240.0	121.5	—	1.6%	6.9%
Proposed AI-Edge Synthesis	18.6	22.1	3.5	46.8%	0%	5.4%

Because of the reliance on telemetry and backhaul, cloud-based models experience significant and detrimental spikes in latency at maximum throughput, while the proposed system sees only a minor and negligible increase in detection latency, by only 3.5 ms, under similar conditions.

The combination of a zero packet drops rate and acceptable levels of CPU usage show that the framework is both accurate and functionally viable for real time use in high congestion 5G. Distributed intelligence over centralized processing for the URLLC and mMTC slice makes sense especially having now reduced the end-to-end response time by 46.8%.

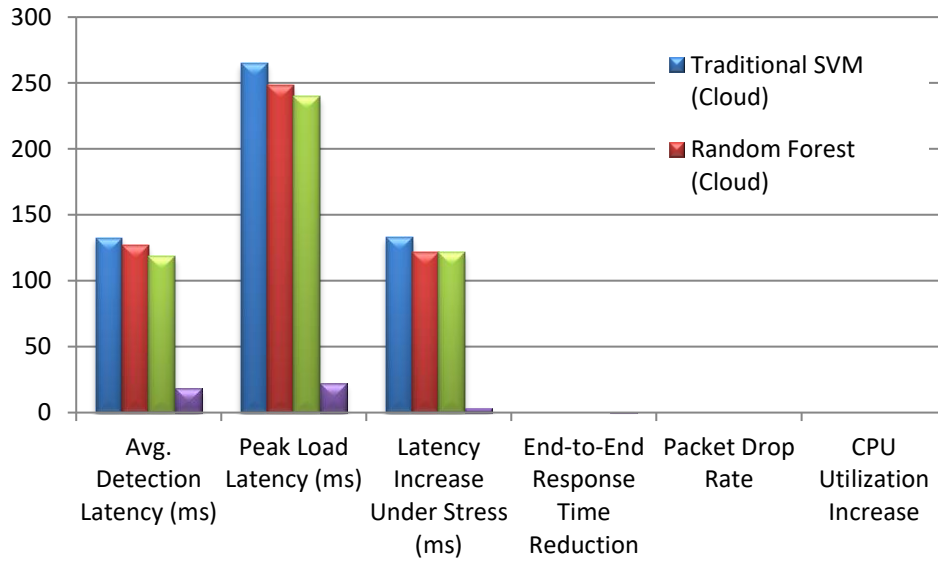


Figure 10. Comparison of operational performance

With confusion-matrix evaluation in the models, as shown in Table 6, AI-edge synthesis achieves the least false positives (180) and false negatives (160), which explain the earlier observation of the improvement of both precision and recall.

Table 6. Confusion matrix-based evaluation of detection performance

Model	True positives (TP)	True negatives (TN)	False positives (FP)	False negatives (FN)	Specificity (%)	Misclassification rate (%)
Traditional SVM	8,110	15,420	880	1,590	94.60	6.10
Random Forest (RF)	8,730	16,240	520	970	96.88	4.20
Centralized DL (CNN-LSTM)	9,280	16,910	360	420	97.93	2.30
Proposed AI-Edge Synthesis	9,740	17,320	180	160	99.00	1.17

As normal 5G traffic is not classified as anomalies 99.00% of the time, unnecessary countermeasures will not be triggered. In addition, under complicated network scenarios, the model demonstrates a 1.17% misclassification rate, which is a clear indication of its capability of differentiating between real and malicious 5G traffic.

Table 7 details an ablation study that measures the significance of each structural part. The findings indicate that the inclusion of CNN enables the model to extract spatial features. Similarly, inclusion of LSTM enables the model to learn the temporal dynamics of traffic. Moreover, inclusion of the autoencoder enhances the model's sensitivity to anomalies as a result of learning an encoded representation of the normal traffic.

Table 7. Ablation study of hybrid model components

Model Variant	CNN	LSTM	Autoencoder	Edge deployment	Accuracy (%)	F1-Score (%)	FPR (%)
CNN only	✓	✗	✗	✗	91.30	90.85	3.90
CNN + LSTM	✓	✓	✗	✗	94.10	93.20	2.10

Model Variant	CNN LSTM		Autoencoder	Edge deployment	Accuracy (%)	F1-Score (%)	FPR (%)
CNN + LSTM + Autoencoder	✓	✓	✓	✗	96.85	96.10	1.65
CNN + LSTM at Edge	✓	✓	✗	✓	97.40	96.95	1.48
Full AI-edge synthesis (Proposed)	✓	✓	✓	✓	98.74	98.02	1.12

Edge deployment as opposed to cloud deployment results in increased reliability of anomaly detection and reduction of false positives [54]. The comprehensive AI-edge integration reaches the optimal result, verifying that the improvement is not incidental, but rather a result of thoughtful integration of the frameworks.

3.1. Latency of edge-inference

The most important result of the study is that localized edge-based inference reduces detection lag. The proposed localized AI-edge system has 46.8% lower end-to-end response time than the AI-based cloud systems we previously analyzed. Traditional centralized systems have a detection time of 118.5 ms. This is because of delays during backhaul and core processing. Our system design has a predictable operational latency in the range of 14.2 ms to 21.8 ms. 5G ultra-reliable low-latency communication (URLLC) has a target response time of 1 ms. This assists the system in containing a compromised network slice, and stops malicious packets from propagating in the core network.

3.2. Performance below varying network lots and strain testing

To test how scalable the framework is, we ran tests at different network traffic levels. Low load was set at 10 Gbps, while an extreme peak load was set at 100 Gbps. The performance graphs in the simulation results show that our distributed inference setup is highly robust to "network congestion". When operating at peak loads (90-100 Gbps), the centralized cloud model experienced a massive latency spike (up to 240ms) due to a bottleneck in telemetry data at the core gateway. In comparison, the AI-edge synthesis did not significantly change the detection latency (only by 3.5ms), demonstrating its ability to address the massive connectivity (mMTC) demands of 5G without an overall reduction in network quality of service (QoS).

The results of the experiment highlight the flexibility of the suggested AI-edge engine across heterogeneous 5G settings, namely multi-slice structures. The eMBB slice achieved high-throughput metadata processing with 99.2 percent performance, and the URLLC slice achieved ultra-fast inference with zero packet drop.

Moreover, the introduction of model quantization (INT8) and weight pruning ensured that deep learning models would run on resource-constrained edge hardware with an acceptable increase in CPU utilization of 5.4%. This illustrates that AI and edge computing synthesis provide a flexible, scalable security fabric that can autonomously meet the requirements of future-proof 5G and 6G architectures.

3.3. Discussion

The empirical performance and operational effectiveness of the suggested framework are deeply rooted in the principles of distributed computing models, namely, the ability to restructure the security boundary of 5G infrastructure based on the multi-access edge computing (MEC) paradigm. Conventionally, anomaly detection has been based on a "cloud-centric" logic, in which telemetry information is backhauled to an endpoint-centralized core for processing, imposing a heavy load.

Nonetheless, when the logic of anomaly detection is moved off the storage of a centralized cloud system onto the extremely decentralized fabric, in this study, a migration is organized from a susceptible "single point of analysis" architecture to a robust, high-performance, and parallel-processing architecture [55].

This network architecture enables every edge node to act as a semi-autonomous agent within a larger, orchestrated multi-agent system (MAS), with distributed intelligence rather than centralized intelligence. This transition is theoretically similar to the edge-mesh computing style, a more advanced topology, in which localized intelligence is used to filter, process, and mitigate threats at the literal origin of data creation.

The system dramatically lowers the entropy (measured in bits) of a network's computational and structural resources by deploying localized inference engines that backhaul large, unprocessed datasets to the core network continuously. Such a decentralization guarantees data locality, i.e., sensitive 5G traffic metadata does not even need to traverse the full network hierarchy to be validated, which reduces the attack surface to a minimum.

This distributed execution strategy has been empirically shown to reduce end-to-end latency by 46.8 percent. Distributing the computational and analytical load throughout the periphery of the network, the system avoids the so-called congestion bottlenecks that are characteristic of centralized 5G security entry points during peak traffic or DDoS attacks.

Moreover, this distribution guarantees the high-speed integrity of the 5G control plane and the user plane, and the network can retain its ultra-reliable low-latency communication (URLLC) properties even under adversarial activity, provided it is already active. In turn, this synergy turns the edge into a complex, first-line defense layer that dynamically scales with the growing number of devices in the 5G and future 6G ecosystems.

The suggested AI-edge synthesis would align with the NIST cybersecurity framework and the new zero trust architecture (ZTA). The core is often viewed as a confined area in a classical 5G architecture. In contrast, our model is based on the premise that the never trust governs the extreme edge of the always verify principle.

Our solution, which offers micro-segmentation capability vital to 5G security, is achieved through real-time AI-driven inspection at the MEC level. This enables the system to detect lateral movement of threats in milliseconds. The connection of our findings to these models shows that AI and edge computing synthesis is not only a performance optimization but also a necessary factor in the creation of resilient software-defined networks (SDN) capable of autonomously countering adversarial AI and automated exploit kits.

One of the most important aspects of this study is that the system can be scaled in the massive IoT (mMTC) setting, where 5G networks should be designed to support one million devices per square kilometer. Conventional centralized models of anomaly detection face the dimensionality curse due to the exponential growth of IoT telemetry. We have found that, through load sharing during inference, the system exhibits a linear scaling factor.

The more IoT devices are added to the network, the more edge nodes can be provisioned to manage the localized traffic. It offers horizontal scalability that will allow the "reconstruction error" logic of our autoencoders to be applied to high-density metadata, ensuring that future growth in core network congestion does not pose a security threat, and also represents a sustainable security blueprint for the future of smart cities and industrial 4.0.

The combination of a smart, edge-based anomaly-detection architecture has far-reaching pragmatic consequences for telecommunications companies struggling to enforce the integrity of next-generation infrastructure. This research provides a true approach to reduced core network load by decentralizing the security layer to stop malicious telemetry at the ingress and reduce amplification of backhaul bandwidth. This structure is critical to sustaining mission-critical network slices, such as ultra-reliable low-latency communications (URLLC).

Reduced detection lag from these network slices is vital to meet strict service level agreements (SLAs). For network operators, this provides a highly resilient self-healing network that can isolate local threats, such as IoT DDoS, without negative impact on quality of service (QoS) for the majority of the customer base.

However, even with the switch to a harmonized AI-edge model, there are highly significant operational disadvantages that would need to be overcome to make it economically reasonable. The first disadvantage is dependable edge hardware capability. Despite model pruning and quantization reducing resource

overutilization, deploying highly complex deep learning requires MEC hardware such as FPGA or GPU placed in the core, leading to considerable network deployment capital expenditure [56].

In addition, in the fast-changing 5G ecosystem, there is the rising concern of model drift. This refers to the changing detection and impact of network traffic which degrades the system performance of detection. This would mean setting up ad hoc automation with continuous flows of MLOps to ensure AI is always in a state of readiness concerning adaptive defense. Lastly, the deployment needs to address the privacy paradox of a distributed learning environment. Even if raw data can be kept at the edge, due to localized processing, the core cloud can still be an attack vector for Inference Attacks as long as model weights or gradients are synced with the core cloud.

To mitigate this threat, operators need to create robust encryption and secure aggregation solutions so that the global model can be updated while remaining compliant with the privacy laws, such as the GDPR. These hardware constraints, coupled with model and cryptographic flexibility, need to be resolved to turn edge-AI from its theoretical potential into a commercially feasible and safe solution for the 5G/6G generations.

4. Conclusions

This paper has created an innovative model in the field of network security by showing how dramatically successful the combination of AI and edge computing are when 5G is used to identify traffic anomalies. The paper deals with the long-standing security gap that puts real-time mobile communications at risk by intentionally relocating the intelligence layer of a centralized cloud system to a decentralized edge system.

The empirical data strongly confirm this design, with a 46.8% decrease in end-to-end detection latency and a stable response window of 14.2ms to 21.8ms, which is within the range of mission-critical 5G services. Moreover, a hybrid CNN-LSTM-auto encoder ensemble synthesis achieved a detection rate of 98.74%, demonstrating that localized, resource-efficient deep learning features can be more effective than conventional centralized models.

This contribution offers a resilient and scalable blueprint for achieving the next generation of hyper-connected environments without impairment of the operational efficiency of the core network. For telecommunications companies and infrastructure providers globally, the move to an intelligent, edge-integrated security presence requires a change in investment in hardware and architecture.

It is highly recommended that commercial 5G implementations focus on hardware-software co-design, using GPU-enabled or FPGA-accelerated MEC (multi-access edge computing) nodes to meet the high-speed computational requirements of real-time deep learning. Also, the fact that anomaly detection is an obligatory element of a zero-trust architecture (ZTA) in which all network slices (eMBB, URLLC, etc.) are checked autonomously and secured at the point of data entry is non-negotiable.

The operators also ought to shift to a state of "closed-loop automation" in which the outputs of AI-driven detection engines are directly connected to SDN (software-defined networking) controllers, so that instant mitigation protocols, like automated slice isolation or dynamic traffic shunting, can be enabled in response to automatically detect engine interventions, and hence ensure 100 percent adherence to strict service level agreements (SLAs).

Even though the suggested framework can be considered a strong defense of existing 5G implementations, the unavoidable progress towards 6G and non-terrestrial networks (NTN) opens new horizons for research. One highly critical subsequent action would be to integrate federated learning (FL) paradigms, which would enable edge nodes to collaboratively improve global threat knowledge without exchanging raw, sensitive user traffic and would also align with world privacy standards, such as GDPR.

Moreover, specific adaptive and self-evolving AI models should be developed to mitigate the effects of "model drift" in dynamic settings of future networks. We are also entering the 6G era, and research needs to be conducted on adversarial resilience, which means that edge-based neural networks are hardened against

advanced so-called Adversarial AI attacks that aim to confuse (or subvert) automated defensive layers. Finally, the extension of this AI-edge synthesis, with the addition of three-dimensional network topologies and satellite-combined 5G/6G systems, is crucial to the creation of an entirely ubiquitous, self-healing, and globally secure telecommunications ecosystem.

Declaration of competing interest

The authors declare that they have no known financial or non-financial competing interests in any material discussed in this paper.

Funding information

No funding was received from any financial organization to conduct this research.

Author contribution

The contribution to the paper is as follows: M. Madinov, S. Khruschak: study conception and design; S. Khruschak, V. Voloshyn, O. Boyko: data collection; V. Voloshyn, O. Boyko, N. Halahan: analysis and interpretation of results; N. Halahan: draft preparation. All authors approved the final version of the manuscript.

References

- [1] I. H. Sarker, M. H. Furhad, and R. Nowrozy, “AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions”, *SN Computer Science*, vol. 2, p. 173, 2021. <http://doi.org/10.1007/s42979-021-00557-0>
- [2] M. S. Chethan and D. Channakrishnaraju, “Review Based Cyber Security and Privacy for IoT Network Using AI-Based Approaches and Cryptography Algorithm with Various Authentication Scheme”, *Optoelectronics, Instrumentation and Data Processing*, vol. 61, pp. 592–605, 2025. <http://doi.org/10.3103/S875669902570075X>
- [3] N. Mohamed, “Artificial intelligence and machine learning in cybersecurity: a deep dive into state-of-the-art techniques and future paradigms”, *Knowledge and Information Systems*, vol. 67, pp. 6969–7055, 2025. <http://doi.org/10.1007/s10115-025-02429-y>
- [4] S. Khlamov, V. Savanevych, Yu. Netrobin, and T. Trunova, “AI-based decision-making process in pipeline for astronomical data mining”, *CEUR Workshop Proceedings*, vol. 4048, pp. 172-186, 2025. <https://ceur-ws.org/Vol-4048/paper14.pdf>
- [5] J. Akshya, et al., “Explainable AI-driven intrusion detection for securing IoT-enabled autonomous transportation systems”, *Cluster Computing*, vol. 28, p. 884, 2025. <http://doi.org/10.1007/s10586-025-05617-1>
- [6] S. Ahmad, M. Arif, M. Ansari, et al., “AI-driven fuzzy requirement engineering for cybersecurity: integrating linguistic decision models”, *Iran Journal of Computer Science*, vol. 8, pp. 2447–2474, 2025. <http://doi.org/10.1007/s42044-025-00324-x>
- [7] T. Basri, “Securing VANETs for Internet of Things (IoT): AI-Driven Solutions for Privacy and Intrusion Detection”, in *AI-Driven Transportation Systems: Real-Time Applications and Related Technologies, Information Systems Engineering and Management*, vol. 62, 2025. http://doi.org/10.1007/978-3-031-98349-8_6
- [8] M. M. Alnfai, “AI-powered cyber resilience: a reinforcement learning approach for automated threat hunting in 5G networks”, *Journal of Wireless Communications and Networking*, vol. 2025, p. 68, 2025. <http://doi.org/10.1186/s13638-025-02497-2>
- [9] S. A. Elatab and A. Almaktoof, “Artificial Intelligence in Cyber Security: A Comprehensive Overview”, in *Selected Papers from the International Conference on Artificial Intelligence (FICAIFY 2025)*, *Studies in Computational Intelligence*, vol. 1229, 2026. http://doi.org/10.1007/978-3-032-00232-7_8

-
- [10] B. Choudhury, P. K. Singh, P. Nath, U. Roy, and A. Kalla, "Blockchain and Smart Contract for Decentralized and Secure Spectrum Management Toward 6G – Beyond Hype", in *Intelligent Spectrum Management*, S. Iyer, A. Kalla, O. A. Lopez, and C. De Alwis, Eds. <http://doi.org/10.1002/9781394201235.ch9>
- [11] K. Rajesh and P. Vetrivelan, "Comprehensive analysis on 5G and 6G wireless network security and privacy", *Telecommunication Systems*, vol. 88, p. 52, 2025. <http://doi.org/10.1007/s11235-025-01282-2>
- [12] M. M. Orabi, O. Emam, and H. Fahmy, "Adapting security and decentralized knowledge enhancement in federated learning using blockchain technology: literature review", *Journal of Big Data*, vol. 12, p. 55, 2025. <http://doi.org/10.1186/s40537-025-01099-5>
- [13] M. Tummala, et al., "A systematic review on current and future prospects of 5G communications", *Cogent Engineering*, vol. 12, no. 1, 2025. <http://doi.org/10.1080/23311916.2025.2547635>
- [14] I. Ahmed, M. A. Syed, M. Maaruf, et al., "Distributed computing in multi-agent systems: a survey of decentralized machine learning approaches", *Computing*, vol. 107, p. 2, 2025. <http://doi.org/10.1007/s00607-024-01356-0>
- [15] V. Maurya, et al., "Blockchain-driven security for IoT networks: State-of-the-art, challenges and future directions", *Peer-to-Peer Networking and Applications*, vol. 18, p. 53, 2025. <http://doi.org/10.1007/s12083-024-01812-w>
- [16] P. Dazzi, "The Internet of AI Agents (IAIA): A New Frontier in Networked and Distributed Intelligence", *International Journal of Networked and Distributed Computing*, vol. 13, p. 16, 2025. <http://doi.org/10.1007/s44227-025-00057-0>
- [17] V. Kottala, A. Al-Sinayyid, S. Adewumi, L. Khamboo, V. Mannuru, and J. Gilliland, "AI-Driven Analysis of Cybersecurity Threats and Defense Mechanisms in 5G Networks", in *AI Revolution: Research, Ethics and Society (AIR-RES 2025)*, H. R. Arabnia, L. Deligiannidis, S. Amirian, F. Ghareh Mohammadi, and F. Shenavarmasouleh, Eds., *Communications in Computer and Information Science*, vol. 2722. Cham, Switzerland: Springer, 2026. http://doi.org/10.1007/978-3-032-12930-7_19
- [18] M. Khattak, et al., "Evolving Multi-Access Edge Computing (MEC) for Diverse Ubiquitous Resources Utilization: A Survey", *Telecommunication Systems*, vol. 88, p. 71, 2025. <http://doi.org/10.1007/s11235-025-01310-1>
- [19] V. Lokavarapu, K. N. Rao, S. Gokuruboyina, et al., "An in-depth analysis of computation offloading approaches in multi-access edge computing systems", *Peer-to-Peer Networking and Applications*, vol. 18, p. 219, 2025. <http://doi.org/10.1007/s12083-025-02028-2>
- [20] K. Moghaddasi, S. Rajabi, and M. Hosseinzadeh, "Intrusion Detection Systems for Enhanced Security in Mobile Edge Computing: A Systematic Review and Survey of the Applications, Challenges, and Future Directions", *Wireless Personal Communications*, vol. 145, pp. 113–175, 2025. <http://doi.org/10.1007/s11277-025-11846-9>
- [21] E. Nascimento, L. Lins, E. Tavares, et al., "Availability assessment of SDN-ICN service for multi-access edge computing", *The Journal of Supercomputing*, vol. 81, p. 436, 2025. <http://doi.org/10.1007/s11227-024-06879-z>
- [22] V. R. Verma, P. Pushkar Kumar, B. Kumar, et al., "An Extensive Investigation on Lyapunov Optimization-based Task Offloading Techniques in Multi-access Edge Computing", *SN Computer Science*, vol. 6, p. 603, 2025. <http://doi.org/10.1007/s42979-025-04130-x>
- [23] H. T. Erdogan and A. Ozsoy, "CUDA-supported 5G multi-access edge computing modifications on 5G-air-simulator", *Journal of Wireless Communications and Networking*, vol. 2025, p. 29, 2025. <http://doi.org/10.1186/s13638-025-02438-z>
- [24] A. Vishweshwara and R. Ramya, "Transforming Telemedicine: Reducing Latency Through Edge Computing and 5G – A Review", *Biomedical Materials & Devices*, vol. 4, pp. 1161–1174, 2026. <http://doi.org/10.1007/s44174-025-00312-6>
-

- [25] H. Nishi and W. A. Shanaka, "Application for a Smart City/Community", in *Computation Platforms for Multi-access Edge Computing*, H. Amano, Ed. Singapore: Springer, 2025. http://doi.org/10.1007/978-981-96-8935-4_4
- [26] S. S. Gill, M. Golec, J. Hu, et al., "Edge AI: A Taxonomy, Systematic Review and Future Directions", *Cluster Computing*, vol. 28, p. 18, 2025. <http://doi.org/10.1007/s10586-024-04686-y>
- [27] M. Ayoubi, M. Li, and M. Albishari, "Context-aware service migration using hybrid SA-DQN in edge computing", *The Journal of Supercomputing*, vol. 81, p. 1536, 2025. <http://doi.org/10.1007/s11227-025-08030-y>
- [28] T. A. Woldegiyorgis, H. X. Li, E. Asmare, et al., "Harnessing Artificial Intelligence to improve building performance and energy use: innovations, challenges, and future perspectives", *Energy Informatics*, vol. 8, p. 138, 2025. <http://doi.org/10.1186/s42162-025-00589-3>
- [29] A. E. Adeniyi, et al., "Edge and Cloud-Based Deployment of AI-Driven Integration in Smart Healthcare", in *HealthTech Horizons: AI-Infused Metaverse Solutions for Smart Healthcare Systems*, P. Barsocchi, J. B. Awotunde, B. Brahma, and A. Kumar Bhoi, Eds., *Studies in Computational Intelligence*, vol. 1224. Cham, Switzerland: Springer, 2026. http://doi.org/10.1007/978-3-031-99946-8_19
- [30] H. Zhao, "Research on privacy protection and performance optimization of 6G communication network based on the fusion of federated learning and edge computing", *Discover Artificial Intelligence*, vol. 6, p. 41, 2026. <http://doi.org/10.1007/s44163-025-00687-x>
- [31] A. Karami and M. Karami, "Edge computing in big data: challenges and benefits", *International Journal of Data Science and Analytics*, vol. 20, pp. 6183–6226, 2025. <http://doi.org/10.1007/s41060-025-00855-3>
- [32] M. Maksimović, S. Jokić, and M. Č. Bošković, "Innovative Horizons for Sustainable Smart Energy: Exploring the Synergy of 5G and Digital Twin Technologies", *Process Integration and Optimization for Sustainability*, vol. 9, pp. 431–470, 2025. <http://doi.org/10.1007/s41660-024-00478-4>
- [33] V. Arun and M. Azhagiri, "Hybrid deep reinforcement learning and genetic algorithm-based resource allocation framework for edge computing", *Peer-to-Peer Networking and Applications*, vol. 18, p. 320, 2025. <http://doi.org/10.1007/s12083-025-02149-8>
- [34] Y. Ikeda, et al., "Anomaly detection and facilitation AI to empower decentralized autonomous organizations for secure crypto-asset transactions", *AI & Society*, vol. 40, pp. 3999–4010, 2025. <http://doi.org/10.1007/s00146-024-02166-w>
- [35] E. Edozie, A. N. Shuaibu, B. O. Sadiq, et al., "Artificial intelligence advances in anomaly detection for telecom networks", *Artificial Intelligence Review*, vol. 58, p. 100, 2025. <http://doi.org/10.1007/s10462-025-11108-x>
- [36] P. Moriano, S. C. Hespeler, M. Li, et al., "Adaptive anomaly detection for identifying attacks in cyber-physical systems: A systematic literature review", *Artificial Intelligence Review*, vol. 58, p. 283, 2025. <http://doi.org/10.1007/s10462-025-11292-w>
- [37] Z. Zhou, J. Wei, M. Liu, et al., "AnomalGRN: deciphering single-cell gene regulation network with graph anomaly detection", *BMC Biology*, vol. 23, p. 73, 2025. <http://doi.org/10.1186/s12915-025-02177-z>
- [38] N. Sriraam, B. Chinta, S. Seshadri, et al., "A comprehensive review of artificial intelligence-based algorithm towards fetal facial anomalies detection (2013–2024)", *Artificial Intelligence Review*, vol. 58, p. 152, 2025. <http://doi.org/10.1007/s10462-025-11160-7>
- [39] T. Thamaraimanalan and S. Ramalingam, "Enhancing anomaly detection in WBANs using hybrid deep learning and optimization algorithms", *Neural Computing and Applications*, vol. 37, pp. 9223–9243, 2025. <http://doi.org/10.1007/s00521-025-11061-4>
- [40] H. Ahmad, G. Mustafa, M. M. Gulzar, et al., "AI-enabled framework for anomaly detection in power distribution networks under false data injection attacks", *Artificial Intelligence Review*, vol. 58, p. 355, 2025. <http://doi.org/10.1007/s10462-025-11318-3>

-
- [41] S. S. Iyengar, S. Nabavirazavi, Y. Hariprasad, P. HB, and C. K. Mohan, “Privacy-Preserving AI (Federated Learning) for Digital Forensics”, in *Artificial Intelligence in Practice, Signals and Communication Technology*. Cham, Switzerland: Springer, 2025. http://doi.org/10.1007/978-3-031-89327-8_5
 - [42] C. Chen, J. Liu, H. Tan, et al., “Trustworthy federated learning: privacy, security, and beyond”, *Knowledge and Information Systems*, vol. 67, pp. 2321–2356, 2025. <http://doi.org/10.1007/s10115-024-02285-2>
 - [43] E. Hadzhyiev, et al., “Application of Docker Compose for constructing the infocommunication system for online processing of astronomical images”, In *Proceedings of the IEEE International Conference on Advanced Computer Information Technologies*, pp. 629–633, 2025. <https://doi.org/10.1109/ACIT65614.2025.11185586>
 - [44] A. Khraisat, A. Alazab, M. Alazab, et al., “Federated learning for intrusion detection in IoT environments: a privacy-preserving strategy”, *Discover Internet of Things*, vol. 5, p. 72, 2025. <http://doi.org/10.1007/s43926-025-00169-7>
 - [45] A. Muhammed, J. Marcos, and N. Gonçalves, “Federated Learning for Secure and Privacy-Preserving Facial Recognition: Advances, Challenges, and Research Directions”, in *Pattern Recognition and Image Analysis (IbPRIA 2025)*, *Lecture Notes in Computer Science*, vol. 15937, 2026. http://doi.org/10.1007/978-3-031-99565-1_16
 - [46] G. Kaur, “PPsFLP: End-to-End Privacy Preserved and Secure Federated Learning Paradigm of Medical Images”, *SN Computer Science*, vol. 6, p. 567, 2025. <http://doi.org/10.1007/s42979-025-04107-w>
 - [47] A. Chouhan and J. Yao, “Federated Learning for Privacy-Preserving: Current Status and Future Directions”, in *Machine Learning and Soft Computing (ICMLSC 2025)*, L. Huang, Ed., *Communications in Computer and Information Science*, vol. 2488. Singapore: Springer, 2025. http://doi.org/10.1007/978-981-96-6403-0_7
 - [48] P. Arora, et al., “Privacy-Preserving Federated Learning System (f-PPLS) for military focused area classification”, *Multimedia Tools and Applications*, vol. 84, pp. 11221–11247, 2025. <http://doi.org/10.1007/s11042-024-19358-7>
 - [49] H. A. Rimi, et al., “Advancements and Challenges in Federated Learning for Privacy-Preserving Smart Healthcare: A Review”, in *Federated Learning in Health Care Technology*, M. F. Mridha and N. Dey, Eds., *Studies in Computational Intelligence*, vol. 1216. Singapore: Springer, 2026. http://doi.org/10.1007/978-981-96-8353-6_11
 - [50] A. Marmat and D. Thankachan, “Design of an integrated unified intelligence-driven method for low-latency, energy-efficient, and secure cloud-integrated optical broadband access networks”, *International Journal of Information Technology*, 2025. <http://doi.org/10.1007/s41870-025-02985-0>
 - [51] M. M. Saeed, et al., “A comprehensive survey on 6G-security: physical connection and service layers”, *Discover Internet of Things*, vol. 5, Art. no. 28, 2025. <http://doi.org/10.1007/s43926-025-00123-7>
 - [52] Y. Xu, et al., “Edge deep learning in computer vision and medical diagnostics: a comprehensive survey”, *Artificial Intelligence Review*, vol. 58, p. 93, 2025. <http://doi.org/10.1007/s10462-024-11033-55>
 - [53] O. Muliarevych, “Acceptance and shipping warehouse zones calculation using serverless approach”, in *Proceedings of the 12th IEEE Int. Conf. Dependable Systems, Services and Technologies*, pp. 1–6, 2022. <http://doi.org/10.1109/DESSERT58054.2022.10018786>.
 - [54] L. Kurmangaziyeva, Z. O. Oralbekova, S. Akhmetzhanova, A. Khassenova, M. Akishev, and T. Zhukabayeva, “Analysis of the problem of ensuring the reliability of the information system”, *CEUR Workshop Proceedings*, vol. 3382, pp. 1–10, 2022. <https://ceur-ws.org/Vol-3382/Paper13.pdf>
 - [55] O. Muliarevych, “The serverless computing for acceptance and shipping warehouse zones optimization”, *CEUR Workshop Proceedings*, vol. 3384, pp. 87–96, 2023. https://ceur-ws.org/Vol-3384/Paper_8.pdf
 - [56] P. Kisała, W. Wójcik, N. Smailov, A. Kalizhanova, and D. Harasim, “Elongation determination using finite element and boundary element method”, *Int. J. Electron. Telecommun.*, vol. 61, no. 4, pp. 389–394, 2015. <http://doi.org/10.2478/eletel-2015-0051>.
-