International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group| Volume 5, Issue 4, 109-121, 2024
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V5I4P111

Original Article

Hybrid AI Models in Network Security: Combining ML, DL, and Rule-Based Systems

Anitha Mareedu
Electrical engineering Texas A&M University - Kingsville 700 University Blvd, Kingsville.

Abstract - *As cyber threats grow in sophistication, traditional rule-based and standalone machine learning (ML) approaches often fall short in ensuring robust network security. This review explores hybrid Artificial Intelligence (AI) models combinations of ML, deep learning (DL), and rule-based reasoning as a promising path for more adaptive and intelligent threat detection systems. The article analyzes how hybrid AI enhances predictive capabilities in network defense, especially when integrated into Security Information and Event Management (SIEM) systems and threat intelligence pipelines. We examine architectural designs such as ensemble models, federated hybrid learning, and AI-assisted policy engines. Through industry case studies including 5G telecom networks, financial fraud detection, and government threat infrastructures we illustrate how hybrid models outperform conventional systems in terms of detection accuracy, response time, and resilience to evasion techniques. The article also addresses critical governance and deployment issues, such as explainability, secure microservice communication, and policy integration in hybrid environments. Emerging research areas are highlighted, including quantum-secure hybrid frameworks, lightweight edge-compatible models, and privacy-preserving federated AI. Despite significant potential, current implementations face limitations in interpretability, scalability, and real-time processing under constrained resources. To address these, we propose practical recommendations for system architects and researchers, emphasizing modular design, auditing mechanisms, and ethical safeguards. This review offers a comprehensive overview of the evolution, benefits, and future of hybrid AI in network security, serving as a guide for both academic inquiry and practical implementation in dynamic threat landscapes*

Keywords - *Hybrid intrusion detection systems (HIDS), machine learning, anomaly detection, AI-driven defense, rule-based security, threat intelligence, adaptive security, real-time monitoring, and cyber threat predicion.*

1. Introduction

The issue of network security has been undergoing a revolutionary change in the last few years owing to complexity and frequency of cyberattacks [1]. The traditional perimeter-based, and statistic firewall and signature-based intrusion detection based architecture no longer suffice to counter dynamic attack vectors like advanced persistent threats (APTs), encrypted communications between malware and its command and control, and lateral movement as well as zero-day exploits. The increase in digital infrastructure, such as cloud computing, 5G, Internet of Things (IoT) devices and remote work infrastructure, has broadened the attack surface excessively and has brought in new opportunities to bad actors[2]. With the advent of software-defined networking (SDN), containerization, and edge computing, organisations are turning towards modern enterprises that require an end to reactive rule-based security. These rapidly changing ecosystems require smart systems that are able to identify new trends, relate varied indicators of threats, and be able to evolve in real-time with no human interaction [3]. The old system of detect and block can no longer effectively offer security security has to be predictive, proactive and adaptive.

The basis of the network security has been rule-based systems that have been in use over the past decades. Firewalls, intrusion detection and prevention systems (IDPS) as well as signature-based antivirus software can all run off of preconstructed patterns of known threats. These methods are good at detecting known attacks, but have difficulties detecting less well-documented attacks, obfuscated and polymorphic attacks, which can bypass them. Moreover, they tend to produce a lot of false positives and cannot cope well with the data streams that might have high rates of scale[4]. Conversely, ML-based security is particularly effective at detecting anomalies in large scale traffic by coming to know the behavioral depths and comparing it to any deviations. Such models can identify possible patterns that would indicate malicious behavior even when the payload or command and control communication is encrypted or the sample is unknown. Those techniques have been proven to be effective in fighting malware, botnets, and insiders.

ML models are not without a problem. They are prone to problems of explainability, model drift, the need to have high-quality labeled data, or manipulation by adversarial inferences. Although they have these shortcomings, their aptitude in pattern detection, predictive and autonomous learning is what makes them very valuable in supplementing the traditional tools [5]. A promising method in the recent security scene is that of a hybrid system of using deterministic rules along with adaptive ML-based solutions. Rule-based systems do come with the advantage of faster and transparent decision making along with regulatory compliance which is absent with ML models but they lack depth and flexibility and are susceptible to specific and intricate attacks. Collectively, these can all produce a complementary architecture that can deliver layered defense and anomaly detection combined with autonomous threat response.

Hybrid systems exhibit a great deal of versatility when compared to high stakes where precision and adaptability are key requirements. As another example, by integrating rule-based detection of the known malware signatures and behavioral ML-based detection, the system will be able to react to both known and zero-day malware. This combination will improve the accuracy of total detections and decrease alert burden since it will minimize false alerts [6]. The uniting of the traditional and AI-based strategies is a more general trend of making intelligent and adaptive cybersecurity ecosystems that will constantly learn, improve, and adapt to emerging threat environments [7]. This survey article discusses the conceptualization and usage of data-driven methodologies and hybrid AI systems in network security. It specializes in features of including ML in threat intelligence management, examining traffic, anomaly hunting, and SIEM. Some of the critical objectives are to provide an analytical overview of recent development trends, focus on practical drawbacks, and propose future alternatives of explainable and autonomous network security systems. The paper will be structured as follows: Section 2 gives a historical background and categorisation. Section 3 deals with ML-based detection models. Section 4 looks at hybrid techniques. The following sections will deal with datasets, implementation case studies, limitations, and new research directions. The conclusion provides a synthesis of the ideas and outlook.

2. Building Blocks of Hybrid AI for Network Security

The evolution of threats assembled in the modern networks necessitates the critical combination of the old rule-based approach with AI to determine and impede complicated multi-vector assaults. Hybrid AI is a collective effort to create a unified intrusion detection and prevention mechanism that can handle familiar and new threats using a combination of fixed rules and adaptive machine learning capabilities. The section describes the key aspects of the hybrid systems such as the rule-based engine, machine learning classifier, and deep neural networks and their complementary nature when used to secure network infrastructures [8].

2.1. Rule-Based Systems in Packet Inspection

Traditional signature-based systems remain a cornerstone in many networks due to their precision in detecting known attack patterns.

2.1.1. Snort and Suricata

- Snort is an open-source intrusion detection/prevention system (IDS/IPS) that uses predefined rules to inspect packet payloads.
- Suricata adds multithreading and improved performance while supporting deeper protocol parsing (HTTP, TLS, DNS)[9].

These systems excel in:

- Detecting known threats with low false-positive rates.
- Real-time alerting for pattern-based attacks like SQL injection or port scans.
- Protocol-aware analysis, making them robust in high-throughput environments.

However, they struggle against:

- Zero-day attacks, where no signature is yet defined.
- Encrypted traffic, which conceals payload data from inspection.

2.1.2. Limitations in Dynamic Environments

Rule-based systems rely on static logic and fail to generalize across evolving attack vectors. This rigidity has motivated the integration of statistical learning approaches that model behavior rather than patterns.

2.2. Machine Learning: Statistical Flow Behavior Modeling

Machine learning (ML) enhances detection by modeling the behavior of network flows rather than relying solely on static patterns [10]. It provides generalization capabilities that help uncover anomalies in encrypted or obfuscated traffic.

2.2.1. Feature-Based Flow Analysis

Features commonly extracted from network flows include:

- Source/destination IP and port.
- Protocol type.
- Byte and packet counts.
- Flow duration and inter-arrival times.

2.2.2. Common Algorithms Used

- Random Forest and SVM for classification of normal vs. malicious traffic.
- K-Means for unsupervised anomaly detection.
- Naïve Bayes for probabilistic modeling of flow patterns.

These models are particularly useful in:

- Detecting polymorphic malware that alters signatures.
- Monitoring lateral movement across internal segments.
- Modeling user and device behavior for anomaly detection.

Machine learning models are data-driven and evolve based on observed patterns, making them a critical complement to rule-based systems.

2.3. Deep Learning: Handling Encrypted, High-Dimensional Traffic

Deep learning (DL) provides an edge in processing raw or high-dimensional data, enabling security analytics without requiring explicit feature engineering [11].

2.3.1. Convolutional and Recurrent Networks

- CNNs are used for packet-level analysis by transforming packet content into image-like matrices.
- RNNs and LSTMs model sequence behavior, such as connection patterns over time.

2.3.2. Use Cases of Deep Learning

- Encrypted traffic classification: DL models can distinguish between benign and malicious encrypted streams using side-channel metadata (e.g., packet size, timing).
- IoT traffic analysis: Lightweight DL models identify anomalies in resource-constrained devices.
- Behavioral fingerprinting: Tracking the identity of attackers based on learned patterns of movement and communication.

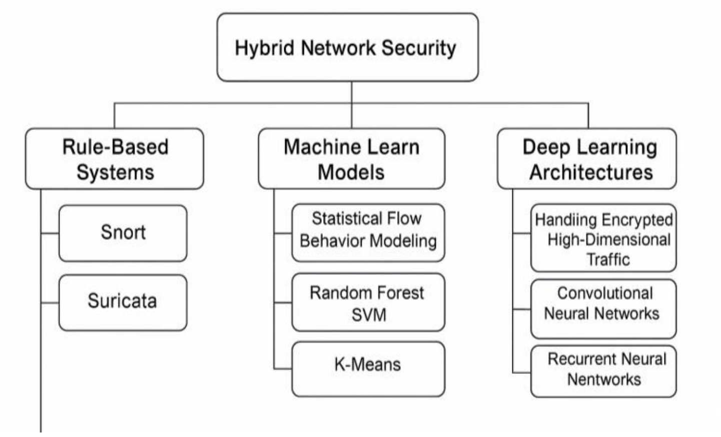
Hybrid Network Security
Rule-Based Systems
Snort
Suricata
Machine Learn Models
Statistical Flow Behavior Modeling
Random Forest SVM
K-Means
Deep Learning Architectures
Handiing Encrypted High-Dimensional Traffic
Convolutional Neural Networks
Recurrent Neural Nentworks

Fig 1: Categorization of AI Components for Hybrid Network Security

Figure 1 provides a visual breakdown of hybrid security components, illustrating how rule-based systems, ML models, and DL architectures interact to form a layered detection system. As shown, these components align across detection layers starting with deterministic pattern matching, progressing through statistical flow learning, and culminating in deep, context-aware analytics.

2.4. Summary: Why Hybrid Works

- Rule-based systems ensure precision and reliability for known threats.
- Machine learning extends adaptability by modeling behavior across flows.
- Deep learning enables context-rich analytics even in encrypted or obfuscated scenarios.

Together, these form a hybrid detection stack that:

- Covers a broad threat spectrum.
- Reduces false positives by correlating multiple models.
- Enhances response time through continuous learning and automation.

The synergy among these components is what enables hybrid AI to outperform traditional models in network security contexts [12].

3. Architectures and Design Patterns

Structure of a hybrid AI-based network security system plays a pivot role to its capability and scalability. A number of different patterns of integration outline how the rule-based engines, machine learning models, and deep learning layers combine in the detection pipeline. This part discusses the fundamental design patterns that facilitate real-time threat detection and low false positives and improved flexibility.

3.1. Hybrid Integration Patterns: Sequential, Parallel, and Layered

Hybrid security networks tend to be built utilizing frameworks of integration that has characterized the interaction of rule and learning engines. Such architectures may summarily be categorized into [13][14]:

3.1.1. Sequential pattern

The first line of defense is rule-based systems like Snort or Suricata that do initial filtering. Qualifying packets or flows then pass through this phase to a machine learning (ML) or a deep learning (DL) classifier which inspects them further.

3.1.2. Parallel pattern

Signature based module and ML based module are done in parallel and the decision is given by a correlation engine or a consensus logic. This pattern is redundant and it enhances accuracy in detection.

3.1.3. Layered pattern

Layers of security checks are arranged in this type of architecture. As an example, Layer 1 may perform traffic screening on the basis of protocol abnormalities, Layer 2 can perform static evaluation and screen and Layer 3 may perform dynamic evaluation using ML/DL. Such designs are flexible and powerful. The concentric pattern, especially, emulates the defense-in-depth approach applicable in enterprise level configurations where various measures are put into effect sequentially.

3.2. Rule-Based Filters → ML/DL Classifiers

A core implementation approach in hybrid architectures is pipelining: where rule-based filters act as the pre-processing phase and ML/DL classifiers serve as the detection core [15].

3.2.1. Pre-filtering

Any malformed packets or known bad signatures are filtered out by static rules. This alleviates the burden on more resource-demanding parts of ML.

3.2.2. ML classes and features filtering

The flows that are passed to it are first subjected to the ML schemes (e.g., Random Forests, SVMs) or DL models (e.g., CNNs, RNNs). These models have the capability to categorize traffic as benign, suspicious or malicious depending on what they are programmed to understand. This strategy optimizes both speed and accuracy while ensuring low false positives.

3.3. Ensemble Techniques: Bagging, Boosting, and Voting

Modern hybrid systems often integrate ensemble learning techniques to improve resilience and generalization [16]:

3.3.1. Bagging (Bootstrap Aggregation)

Several classifiers (e.g. decision trees) come to train on a another set of data, and the results are collected with an average or by voting. This decreases the variance and increases stability.

3.3.2. Boosting

The weak classifiers are sequentially trained based on the models and the subsequent classifiers are trained on the errors of the previous one (e.g., AdaBoost, XGBoost). This goes a long way to improve the detection of complex attack vectors.

3.3.3. Voting mechanisms

In the case of multiple models (rule-based, statistical, and DL) providing conflicting information, voting schemes (hard or soft) are employed to obtain a decision sufficient to be considered the answer that is not corrupted under possible information and time constraints. These ensemble techniques are particularly useful in dynamic environments like cloud networks and IoT-based infrastructures.

3.4. Application in IDS/IPS, Next-Gen Firewalls

Hybrid architectures are already being adopted in:

3.4.1. Intrusion Detection Systems (IDS)

These detect potential breaches through behavioral patterns and static signatures. Hybrid IDSs combine Suricata rules with LSTM-based traffic prediction.

3.4.2. Intrusion Prevention Systems (IPS)

Going beyond detection, IPS actively blocks malicious traffic. ML is used to prevent false positives by learning from historical misclassifications.

3.4.3. Next-Generation Firewalls (NGFWs)

NGFWs integrate traditional firewalling with threat intelligence, application-level control, and ML-based detection of zero-day exploits.

Key Benefits of Hybrid Integration:

- Real-time detection of both known and unknown threats
- Lower false positive/negative rates Scalability across heterogeneous networks
- Compatibility with encrypted traffic analytics

These applications demonstrate how hybrid models enhance both precision and adaptability in enterprise and cloud-centric infrastructures.

Table 1: Hybrid vs. Standalone Model Structures in Network Security

Feature	Rule-Based (Standalone)	ML/DL-Based (Standalone)	Hybrid Architecture
Detection Accuracy	High for known attacks	High for novel patterns	High for both
False Positives	Often high	Moderate	Low
Real-time Performance	Fast	Slower	Balanced
Encrypted Traffic Analysis	Weak	Strong with DL	Strong
Deployment Complexity	Low	High	Moderate
Adaptability to New Threats	Poor	Good	Excellent

Table 1 compares the capabilities of different network security models. Hybrid systems combine the strength of both traditional and AI-based methods. The latter architectural approaches are in line with new design concepts of smart network defense, as observed in state-of-art designs, such as blockchain-based distributed detection schemes and mobile edge networks incorporating AI into the security design [17].

4. Detection in Network Security

The hybrid AI systems have been indispensable in a variety of network security application scenarios where only signature-based systems and only behavior would be deficient. Effectively leveraging both the capabilities of both the non-intelligent and intelligent methods, which organizations can combine intelligently can help them to adequately identify, respond, and even forecast a diverse range of advanced attacks. A few of the fundamental applications are brought out into focus in this section showing the way hybrid systems are used in the real security systems [18].

4.1. DoS/DDoS Detection Using Multi-Layer Models

Denial-Of-Service (DoS) and Distributed Denial-Of-Service (DDoS) attacks remain dynamic in enormity and sophistication, and focus on the availability of life-aspect services. Rule-based firewalls have a tendency of detecting volumetric anomalies but cannot easily detect low rate or protocol DDoS anomalies. The hybrid models solve this problem with multi-level detection pipelines. The first level removes the known attack signatures and dubious traffic amounts through threshold based guidelines whereas the latter ML/DL level examines the time-series information including packet inter-arrival time, connection time, and net protocol entropy. Machine learning techniques like the Random Forest and LSTM networks have proved practical in the identification of the early-phase DDoS activity with enhanced generalization rates[19]. What is more, feedback loops can be employed to tune down rule-based components with ML-beleved false positive with a view to creating more accuracy over a period of time.

4.2. Encrypted Traffic Inspection with DL and Heuristics

Encryption has become standard across the web, improving privacy but also concealing malicious payloads from traditional inspection systems. Legacy intrusion detection tools relying on deep packet inspection (DPI) are increasingly blind to threats embedded in SSL/TLS-encrypted flows. To bridge this gap, hybrid approaches employ deep learning models trained on metadata and flow characteristics, such as packet size distribution, TLS handshake attributes, and timing patterns. These models are often coupled with heuristics derived from domain knowledge for instance, identifying sudden spikes in outbound encrypted traffic or unexpected certificate issuers. Such systems allow passive, non-intrusive detection of encrypted threats without breaking encryption a crucial capability in privacy-sensitive environments [20].

4.3. Lateral Movement and Malware Spread Detection

Lateral movement, often seen in advanced persistent threats (APTs), involves attackers exploiting one compromised host to pivot through the network. Signature-based detection often fails because these activities don't match known patterns and unfold slowly over time. Hybrid architectures can detect these stealthy behaviors through correlation of host-based and network-level indicators. Rule-based policies might flag unusual administrative tool usage or internal file transfers, which are then fed into an ML layer trained on historic movement profiles across subnets. Graph-based anomaly detection and temporal sequence modeling (e.g., using GRUs or temporal convolutional networks) help identify deviation from baseline communication flows [21]. Such systems can even reconstruct an attack chain, enabling rapid containment and forensic analysis.

4.4. IoT Threat Mitigation at the Network Edge

The proliferation of IoT devices introduces a new layer of vulnerability, especially as many devices run outdated firmware or lack proper authentication mechanisms [22]. Edge computing frameworks offer the opportunity to implement low-latency hybrid security pipelines close to the device layer. Rule-based policies at the edge node can instantly block traffic violating protocol behavior (e.g., non-standard ports or malformed packets). Simultaneously, lightweight ML models often decision trees or k-NN classifiers operate on constrained environments to detect deviations from typical device behavior, such as unusual packet rates or sudden external communications. In more advanced settings, compressed neural networks or TinyML models monitor behavioral fingerprints of IoT devices, identifying impersonation, botnet command traffic, or firmware-level exploits.

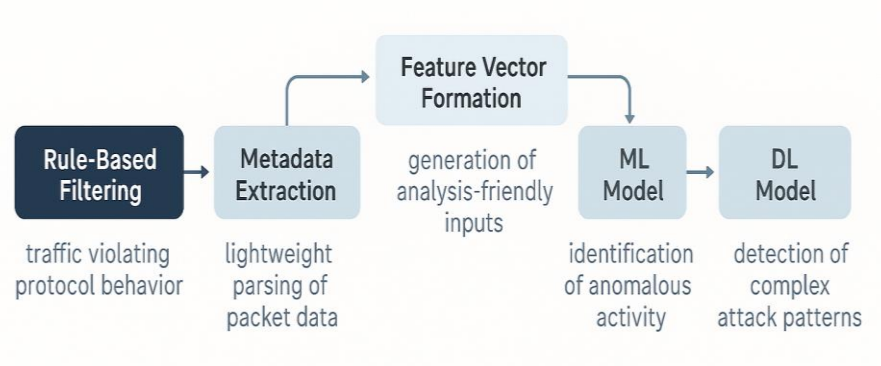
Rule-Based Filtering
traffic violating protocol behavior
Metadata Extraction
lightweight parsing of packet data
Feature Vector Formation
generation of analysis-friendly inputs
ML Model
identification of anomalous activity
DL Model
detection of complex attack patterns

Fig 2: Pipeline of a Layered Hybrid Detection Model

Figure 2 illustrates a generalized pipeline integrating rule-based filtering, metadata extraction, feature vector formation, and a layered ML/DL decision engine for adaptive threat detection. The figure emphasizes the modular nature of hybrid detection systems, where each component contributes to reduced false positives, improved real-time response, and better coverage against evolving threats.

5. Benchmark Datasets and Performance Evaluation

As hybrid AI models continue to evolve for network security applications, the evaluation of their performance across reliable and representative datasets becomes essential. Benchmarking not only validates detection capabilities but also reveals model robustness under diverse conditions, including real-world traffic complexity, encrypted flows, and unseen attack types.

5.1. Public Datasets (NSL-KDD, CICIDS2017, UNSW-NB15)

Hybrid models are typically trained and tested using publicly available labeled datasets, each with distinct advantages and limitations [23].

5.1.1. NSL-KDD

It is an improved version of the outdated KDD'99 dataset, addressing data redundancy and imbalance. It provides labeled flow records for multiple classes of attacks such as DoS, Probe, R2L, and U2R.

5.1.2. CICIDS2017

IT offers modern traffic characteristics and incorporates attack simulations such as Brute Force, Heartbleed, and Web attacks, alongside benign traffic generated using real applications like Skype and HTTP.

5.1.3. UNSW-NB15

Developed using IXIA tools, blending normal and synthetic malicious traffic with a broad set of 49 features across 10 families of attacks. These datasets provide a foundation for training layered models, validating detection accuracy, and comparing against baselines. However, they vary significantly in traffic volume, feature diversity, and recency.

5.2. Traffic Feature Extraction and Flow Labeling

Accurate traffic classification depends on robust feature engineering. Features can be statistical (packet count, byte count, flow duration), time-based (inter-packet arrival times), or content-based (flag values, header lengths). Preprocessing steps such as one-hot encoding, normalization, and dimensionality reduction (e.g., PCA) are often required[24]. Labeling flows correctly is equally critical. In CICIDS2017, labels are timestamp-aligned with the injection of known attack scripts, while in NSL-KDD and UNSW-NB15, labeling is performed by simulation environments. Flow-label integrity is crucial for the validity of supervised learning approaches.

5.3. Evaluation Metrics: Accuracy, TPR, FPR, AUC

To measure detection performance, multiple metrics are employed:

- Accuracy: Measures overall correctness but may be misleading with imbalanced datasets.
- True Positive Rate (TPR) or Recall: Indicates the model's ability to detect actual attacks.
- False Positive Rate (FPR): Represents the rate at which benign traffic is incorrectly flagged.
- Area under the ROC Curve (AUC): Captures trade-offs between TPR and FPR, providing an aggregate performance ***measure across thresholds.***

Precision, F1-score, and confusion matrices are also reported to contextualize performance in multi-class detection tasks. High AUC and low FPR are particularly important for real-time systems where false alarms can overwhelm analysts.

5.4. Challenges in Real-World Benchmarking

Despite their utility, benchmark datasets present challenges:

- Synthetic Traffic Gaps: Simulated environments do not fully capture live internet traffic diversity.
- Feature Drift: Attack patterns and benign usage behavior evolve over time, reducing generalization.
- Label Quality: Inaccurate or inconsistent labels degrade training outcomes.
- Scalability: Some models perform well in lab conditions but falter under high-throughput or encrypted traffic in deployment.

These limitations necessitate the inclusion of live network traces and adversarial testing to supplement benchmark-driven evaluations. Advanced hybrid models should demonstrate adaptability to traffic drift and encrypted payload inspection under low-latency constraints.

Table 2: Dataset Characteristics and Suitability for Hybrid Models

Dataset	Attack Types	Feature Count	Flow Diversity	Suitable for DL
NSL-KDD	4 Classes (DoS, R2L)	41	Low	Limited
CICIDS2017	Multiple (Web, DDoS)	80	High	Yes
UNSW-NB15	10 Attack Families	49	Medium	Yes

These datasets, though essential, are stepping stones toward real-world deployment. Incorporating adaptive feedback from production traffic will remain key in refining future hybrid detection systems [25].

6. Implementation and Deployment Challenges

Hybrid AI-based network security systems hold immense potential, yet their practical implementation in real-world environments is riddled with several technical and operational challenges. As networks grow in speed, complexity, and diversity especially with the proliferation of IoT and edge computing the efficacy and performance of hybrid detection mechanisms are increasingly tested under production conditions.

6.1. Data Imbalance and Concept Drift in Network Traffic

A fundamental challenge in network traffic analysis is the class imbalance problem, where benign traffic overwhelmingly outnumbers malicious instances. This skew can mislead learning algorithms into biasing predictions toward the majority class, severely reducing detection performance for minority (i.e., attack) classes. Moreover, concept drift the gradual or abrupt change in network behavior patterns over time can degrade the accuracy of machine learning (ML) models if not continuously retrained. Static models may quickly become obsolete, especially in adversarial environments where attackers evolve techniques to evade detection. To address this, researchers often integrate online learning or incremental retraining, where models adapt to new traffic distributions. Hybrid systems may combine long-term learned patterns with short-term, adaptive heuristics to stay resilient against emerging threats [26].

6.2. Real-Time Processing in High-Speed Networks

Real-time intrusion detection is a critical requirement in modern enterprise and ISP-grade networks. Processing gigabit-per-second (Gbps) traffic rates with hybrid models especially those involving deep learning (DL) poses substantial performance constraints. Traditional rule-based systems like Snort operate efficiently in line-rate environments but may miss complex threats. On the other hand, ML and DL modules can introduce processing delays due to computational overhead. A practical solution involves parallel pipelines and hardware acceleration (e.g., GPUs or FPGAs) for DL inference, combined with faster pre-filtering mechanisms. For instance, lightweight rule-based modules can act as gatekeepers to pass only suspicious flows to deeper ML layers, thereby maintaining responsiveness in high-speed environments [26].

6.3. Latency and Throughput Bottlenecks

High detection accuracy often comes at the cost of latency. In hybrid systems, where packets are subject to multi-stage analysis e.g., rule matching → feature extraction → classification cumulative latency can disrupt network performance, especially in latency-sensitive applications such as VoIP or industrial control systems. This trade-off demands the optimization of detection pipelines, including reduced feature dimensionality, use of compressed models (e.g., quantized neural networks), and efficient serialization of model inference. Ensemble architectures, while improving robustness, may further increase latency unless parallel execution or approximate computing strategies are employed.

6.4. Edge Resource Constraints in Distributed Networks

The advent of edge computing has also transferred the network protection duties to resource-limited conditions like routers, gateways, and embedded products. The deployment of hybrid AI at the edge has given rise to the issues of the limited volume of memory, CPU, and energy provision. Such setups are often too way too large to use Full DL models. Model partitioning is one such approach, where less significant detection is done at the edge, and any suspicious event is sent to cloud or centralized systems to be analyzed in depth. Then there is the use of federated learning, where each edge device can learn models on common tasks without the transmission of raw data, keeping the data private and addressing the situation outstripping edge resource availability [26].

Figure 3 illustrates the common computational and architectural bottlenecks encountered in deploying hybrid AI systems in live production networks.

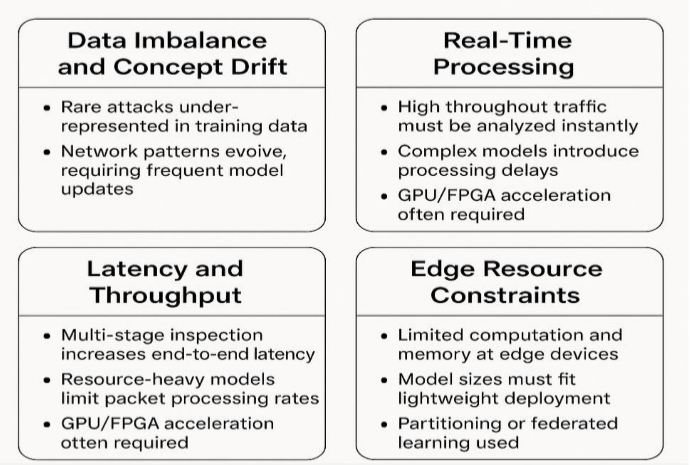
Data Imbalance and Concept Drift
• Rare attacks under-represented in training data
• Network patterns evoive, requiring frequent model updates
Real-Time Processing
• High throughout traffic must be analyzed instantly
• Complex models introduce processing delays
• GPU/FPGA acceleration often required
Latency and Throughput
• Multi-stage inspection increases end-to-end latency
• Resource-heavy models limit packet processing rates
• GPU/FPGA acceleration otten required
Edge Resource Constraints
• Limited computation and memory at edge devices
• Model sizes must fit lightweight deployment
• Partitioning or federated learning used

Fig 3: Key Constraints in Deploying Hybrid AI in Production Networks

Nevertheless, such techniques notwithstanding, there is an unresolved research issue of balancing the quality of the detection and its resources cost, especially in real-time settings and adversarial settings. Engineers would have to compromise among accuracy, explainability and scalability

7. Security Governance and Policy Considerations

Since the control of network traffic and the mitigation of threats begin to be under the impact of AI, this area, security governance, becomes a necessary option. In addition to being accurate in detection, organizations must have measures that would ensure that these systems meet regulatory frameworks, decision transparency and accountability preservation. This paragraph examines the critical governance principles that influence hybrid AI system in cybersecurity.

7.1. Trust and Explainability in Automated Network Filtering

A more questionable issue with the introduction of the machine learning-based filtering systems is the lack of interpretability and reliability of the taken automated decision. In the absence of explainability, system developers might not trust the results of deep neural models, at least not in such high-stakes applications.

7.1.1. Explicable models over Black Boxes

Models such as decision trees, linear classifiers are transparent in the decision path, in contrast to deep neural networks, which are very powerful, but are often described as a black box. The local explanation of complex models is extracted with the help of techniques such as LIME and SHAP.

7.1.2. User confidence and auditability

Explorable outputs inspire trust in the users and help to make human verification possible in reviews of security. These mechanisms are necessary in regulated industries to facilitate compliance as well as the investigation of an incident.

7.2. Security Engineering and Policy Engine Interaction

The enforcement of policy compliance should occur on many tiers of the infrastructure. The hybrid surveys have to interact with existing access controls system and firewalls and network pacification consoles to comply with the policies of the enterprises.

7.2.1. AI Pipelines that are Policy-Aware

Detection paradigms are required to get incorporated in policy-conscious execution chains where policy (block, alert) is subject to enterprise policy. This presents the guarantee that automated responses will not create a clash within the priorities of operation, as well as the requirements of compliance.

7.2.2. Alignment Access Control and SIEM Systems

Constant transmission of information between detection models and the control system such as the SIEM platforms and the access management services, enables the delivery of dynamic changes in policies and makes access decisions based on risks.

7.2.3. Safe Microservice Communication Hybrid Systems

Security Insecurity The microservice security architecture has the capability to distribute their detection and response logic to numerous instances in various services. Safe communication between these services is highly significant in order to maintain the system integrity and confidentiality [27].

7.2.4. Service Authentication and Zero Trust

Zero Trust models mutually authenticate, least privilege and encrypt the traffic between services. Istio and Linkerd are examples of service mesh technologies that assist in enforcement of such policies.

7.2.5. API Gateways and Secure Tokens

API gateways can be used to manage access between services using OAuth tokens or mTLS, providing centralized control and monitoring of service-to-service communications.

7.3. Accountability and Auditing in Hybrid Decision-Making

Hybrid systems introduce complexity into decision-making pipelines, making it essential to implement clear logs and traceability mechanisms that assign responsibility for actions.

7.3.1. Decision Traceability

Logging every input, decision point, and model output helps trace security decisions in hybrid architectures. This supports both incident response and model governance.

7.3.2. Regulatory Compliance and Evidence

Industries such as finance and healthcare require detailed audit trails for all automated decisions. Compliance frameworks like GDPR and HIPAA mandate this level of accountability in AI-driven systems.

Table 3: Governance Layers in AI-Driven Network Defense

Governance Layer	Description
Trust & Explainability	Supports human understanding of AI decisions (e.g., LIME, SHAP)
Policy Integration	Ensures AI actions adhere to enterprise security policies
Secure Communication	Protects inter-service messaging with mTLS, service mesh, or API tokens
Auditing & Compliance	Maintains logs for traceability, regulatory reporting, and accountability

8. Deployment of Hybrid AI Technique

The deployment of hybrid AI techniques has gained traction across several high-risk sectors, including telecommunications, finance, government, and cloud-native environments. These sectors require real-time, scalable, and explainable threat detection strategies. Hybrid models those that blend traditional rule-based methods with advanced machine learning offer tailored solutions that balance performance, interpretability, and operational flexibility. Below are four representative case studies illustrating how hybrid AI has been applied across industries.

8.1. Telecom: Hybrid AI for 5G Core Network Security

The evolution to 5G has introduced a distributed, ultra-low latency architecture, increasing both the attack surface and the demand for intelligent threat detection. In telecom environments, hybrid AI models are increasingly used to monitor signaling protocols, detect session hijacking, and guard against control-plane attacks. For example, a telecom operator deployed a hybrid framework integrating statistical anomaly detection with deep packet inspection (DPI) and deep learning classifiers to flag malicious control-plane behavior in real time. As illustrated in Figure 4, the layered architecture captured packet-based anomalies using temporal sequence models (LSTM) while leveraging rules for known protocol violations. This hybrid approach reduced false positives and improved detection speed, particularly in high-throughput 5G environments (28).

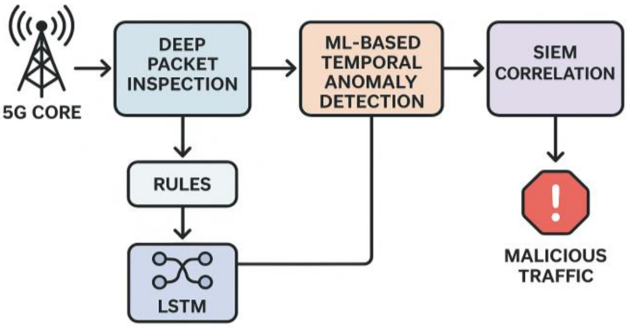

Fig 4: Packet-Based Threat Detection in 5G Network Using Hybrid AI

This diagram shows a layered architecture combining deep packet inspection (DPI), ML-based temporal anomaly detection, and SIEM correlation for detecting malicious traffic in 5G core networks

8.2. Enterprise: Fraud Detection in Financial Network Segments

In the enterprise sector, financial networks are frequent targets for fraud, especially involving business email compromise (BEC), transaction hijacking, and credential abuse. A hybrid AI system was implemented at a global financial firm to secure internal transactional workflows and customer endpoints. The system fused unsupervised learning (isolation forests and autoencoders) with heuristic-based rules on IP behavior, device fingerprinting, and transactional thresholds. Suspicious patterns such as off-hour access or sudden geographical changes triggered alerts, which were routed through a rule-prioritized SIEM. Hybrid models enabled faster risk scoring and adaptive thresholding, improving fraud detection rates while minimizing friction (28).

8.3. Government: Hybrid AI in National Threat Infrastructure

Governments are increasingly deploying hybrid AI models in national critical infrastructure, such as defense, power grids, and cyber-physical systems. These environments often demand explainability and strong policy compliance, alongside real-time analytics. A defense agency in Asia-Pacific adopted a hybrid AI solution that layered decision trees and LSTM models to monitor internal traffic across departments. Rule-based engines were used to enforce access control policies and flag deviations, while deep learning identified insider threats based on behavioral drift. Audit logs and explainable AI techniques (e.g., SHAP values) were incorporated to meet regulatory standards.

8.4. Cloud-native Security Operations with Hybrid Models

Current security operations centers (SOCs) apply hybrid AI in an environment that is cloud-native, especially Kubernetes and containers environments, to provide visibility and response. Such systems need to monitor run-time threats, escape attempts by containers and cross-silo movement, within the cloud cluster. A company offering cloud services used hybrid AI pipelines to incorporate into its DevSecOps system signature-based policies of the known exploits, anomaly detection on the workload behavior that consists of graphs. The hybrid-model allowed updating the policy on the fly, precise finding, and incorporation into the CI/CD pipelines thereby leading to the resilient defense of cloud native environments.

9. Research Trends and Emerging Directions

As Hybrid models have proved to be critical as the field of AI-augmented security transforms to become more adaptive and context-sensitive in terms of identifying threats. Recent work increasingly looks to the possibility of enabling secure, lightweight, and privacy-preserving frameworks acceptable in the new heterogeneous networks of edge and quantum-resilient networks. The following are four aligning directions that seem to be leading the future of hybrid AIs in network security.

9.1. Federated AI for Cross-Domain Network Monitoring

Federated learning gives an answer to collaborative threat intelligence without breaching privacy or jurisdiction regulations through a decentralized approach. It allows the training of the models on multiple organizations or network domains but stores data locally [30].

- Organizations benefit from shared threat models without exposing internal traffic logs.
- Reduces legal and ethical concerns around centralized training on sensitive network data.
- Facilitates adaptive updates to threat detection across global environments.

9.2. Lightweight Hybrid Models for Edge and Fog Computing

Deploying hybrid AI models in edge or fog environments introduces constraints on processing power and memory. To overcome this, recent trends focus on minimizing model complexity while retaining detection accuracy.

- Knowledge distillation and pruning techniques are used to reduce model size.
- Use of decision trees and rule-based systems alongside compact neural nets improves inference speed.
- Facilitates real-time detection in IoT or vehicular networks.

9.3. Quantum-Secure Hybrid Architectures

With the looming threat of quantum decryption, hybrid security systems are being redesigned to include quantum-resistant cryptographic algorithms [29]. This integration ensures secure model deployment and inter-node communication even in a post-quantum era.

- Lattice-based encryption and hash-based signatures are being trialed in hybrid AI deployments.
- Emerging protocols support post-quantum secure exchange of model parameters during federated learning.
- Addresses long-term confidentiality of AI-generated insights.

9.4. Ethical and Privacy Implications of Autonomous Filtering

The use of autonomous AI systems for filtering and blocking network traffic introduces new challenges related to accountability, fairness, and transparency.

- Opaque AI decisions may lead to unjustified packet drops or user access denials.
- Risk of bias if models are trained on unrepresentative or adversarial datasets.
- Need for audit trails and explainability mechanisms to justify decisions.

Recommendations:

- Introduce human-in-the-loop verification for high-impact decisions.
- Embed interpretability layers within neural models used in critical infrastructure.

Table 4: Open Research Problems and Roadmap

Research Area	Open Problem	Roadmap Direction
Federated AI	Secure coordination without data leakage	Trusted execution + privacy-preserving gradients
Lightweight Hybrid Models	Retaining accuracy on constrained devices	Neural pruning + adaptive rule layering
Quantum-Secure Architectures	Post-quantum key exchange in federated learning	Lattice-based comm protocols + hybrid certificates
Ethical Filtering	Ensuring explainability and accountability	Interpretable models + autonomous decision logs

This roadmap, grounded in the insights by [29], highlights a multi-pronged research trajectory spanning computation, cryptography, and ethics for next-generation AI-based network security systems

10. Conclusion

Combination of machine learning, deep learning and rule based logic in hybrid Artificial Intelligence (AI) models has also led to significant enhancement of network security systems. These models provide the flexibility of data-driven studies with the robustness of classical decision rules, and are able to adapt to various categories of threats, anomaly behavior and offer real time mitigation. Hybrid AI models have proved to be useful in areas of telecommunications, finance, government infrastructure, as well as cloud-based environments where the issue of precision, scalability and responsiveness has shown significant improvements. The synergetic efforts of unsupervised and supervised methods are interesting to realize unknown threats and hybrid techniques are more adapted to such dynamic and heterogeneous networks. Just like any other solution, there are a number of weaknesses that limit the maximum potential of these solutions. The fact is that data imbalance tends to bias the models in training, and concept drift caused by a changing attacker behavior lowers the accuracy over time. Moreover, as contemporary networks and networks are high-speed, there are latency and throughput issues, especially in the attribute in edge or resources-limited devices. Externalities are a major problem when it comes to trust and explainability; analysts have to make decisions with security in mind, but many hybrid models are black boxes. In addition, interdependence with existing policy engines, auditing systems, and secure communication systems remains a labor-intensive process that needs area-specific aspects.

In order to proceed further with practical implementations, a number of recommendations are suggested. Researchers have to place an emphasis on the creation of lightweight hybrid models that could effectively work at the edge and in the fog computing networks. Special focus should be given to self-adaptive learning processes that reduce concept drift but also do it without impacting the robustness of models. Second, the network security engineers ought to use modular, explainable hybrid frameworks to promote their integration to existing policy engines, controls of access, and audit. Forming XAI modules is able to make operations transparent and regulatively compliant. Further, security and safety of the federated learning systems can be engineered to be secure and privacy-preserving across organizational boundaries, which is a necessity in collaborative threat intelligence work across the domains. In the future, hybrid AI is the next big leap in intelligent network protection, especially against quantum-era attacks and autonomous attack possibilities. At a younger stage, the hybrid systems are already a paradigm transformation between reactive and predictive security systems. Such ethically-related issues as the effects of automated filtering on personal privacy and due process will need multidisciplinary cooperation in solving. Still, in the end, hybrid AI can provide a robust foundation of next generations cybersecurity but to reach its potential, it is going to need a delicate balance of technological advancement, policy, and institutional responsibility.

References

[1] Y. Li and Q. Liu, “A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments,” Energy Reports, vol. 7, pp. 8176–8186, 2021.

[2] A. Djenna, S. Harous, and D. E. Saidouni, “Internet of things meet internet of threats: New concern cyber security issues of critical cyber infrastructure,” Appl. Sci., vol. 11, no. 10, p. 4580, 2021.

[3] W. Rafique et al., “Complementing IoT services through software defined networking and edge computing: A comprehensive survey,” IEEE Commun. Surv. Tutorials, vol. 22, no. 3, pp. 1761–1804, 2020.

[4] S. Thapa and A. Mailewa, “The role of intrusion detection/prevention systems in modern computer networks: A review,” Proc. Midwest Instruction and Computing Symposium (MICS), vol. 53, 2020.

[5] N. Moustafa et al., “Explainable intrusion detection for cyber defences in the internet of things: Opportunities and solutions,” IEEE Commun. Surv. Tutorials, vol. 25, no. 3, pp. 1775–1807, 2023.

[6] M. A. Adewoyin et al., “A Conceptual Framework for Dynamic Mechanical Analysis in High-Performance Material Selection,” IRE J., vol. 4, no. 5, pp. 137–144, 2020.

[7] J. Jangid and S. Dixit, the AI Renaissance: Innovations, Ethics, and the Future of Intelligent Systems, vol. 1, Technoscience Academy, 2023.

[8] O. A. Agboola et al., “A conceptual model for integrating cybersecurity and intrusion detection architecture into grid modernization initiatives,” Int. J. Multidiscip. Res. Growth Eval., vol. 3, no. 1, pp. 1099–1105, 2022.

[9] W. Park and S. Ahn, “Performance comparison and detection analysis in snort and suricata environment,” Wireless Pers. Commun., vol. 94, no. 2, pp. 241–252, 2017.

[10] S. Wang et al., “Machine learning in network anomaly detection: A survey,” IEEE Access, vol. 9, pp. 152379–152396, 2021.

[11] M. Nasir et al., “Feature engineering and deep learning-based intrusion detection framework for securing edge IoT,” J. Supercomput., vol. 78, no. 6, pp. 8852–8866, 2022.

[12] F. Ahmed, “Cloud Security Posture Management (CSPM): Automating Security Policy Enforcement in Cloud Environments,” ESP Int. J. Adv. Comput. Technol., vol. 1, no. 3, pp. 157–166, 2023.

[13] W. Gan et al., “A survey of parallel sequential pattern mining,” ACM Trans. Knowl. Discov. Data (TKDD), vol. 13, no. 3, pp. 1–34, 2019.

[14] Z. Li, G. Chen, and T. Zhang, “A CNN-transformer hybrid approach for crop classification using multitemporal multisensor images,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 13, pp. 847–858, 2020.

[15] B. Chander, “ML and DL Approaches for Intelligent Wireless Sensor Networks,” in Machine Learning and Deep Learning Techniques in Wireless and Mobile Networking Systems, CRC Press, 2021, pp. 11–40.

[16] M. A. Yaman, F. Rattay, and A. Subasi, “Comparison of bagging and boosting ensemble machine learning methods for face recognition,” Procedia Comput. Sci., vol. 194, pp. 202–209, 2021.

[17] J. Jangid et al., “Enhancing security and efficiency in wireless mobile networks through blockchain,” Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 4, pp. 958–969, 2023. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/7309

[18] M. Agoramoorthy et al., “An Analysis of Signature-Based Components in Hybrid Intrusion Detection Systems,” in 2023 Intell. Comput. Control for Eng. Bus. Syst. (ICCEBS), IEEE, 2023.

[19] S. Ahmed et al., “Effective and efficient DDoS attack detection using deep learning algorithm, multi-layer perceptron,” Future Internet, vol. 15, no. 2, p. 76, 2023.

[20] J. Jangid, “Efficient Training Data Caching for Deep Learning in Edge Computing Networks,” Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol., vol. 7, no. 5, pp. 337–362, 2020. doi: 10.32628/CSEIT20631113

[21] W. Li, W. Meng, and L. F. Kwok, “Surveying trust-based collaborative intrusion detection: State-of-the-art, challenges and future directions,” IEEE Commun. Surv. Tutorials, vol. 24, no. 1, pp. 280–305, 2021.

[22] X. Feng et al., “Detecting vulnerability on IoT device firmware: A survey,” IEEE/CAA J. Autom. Sinica, vol. 10, no. 1, pp. 25–41, 2022.

[23] S. Choudhary and N. Kesswani, “Analysis of KDD-Cup’99, NSL-KDD and UNSW-NB15 datasets using deep learning in IoT,” Procedia Comput. Sci., vol. 167, pp. 1561–1573, 2020.

[24] T. T. Nguyen et al., “Feature extraction and clustering analysis of highway congestion,” Transp. Res. Part C Emerg. Technol., vol. 100, pp. 238–258, 2019.

[25] S. Dixit, “AI-powered risk modeling in quantum finance: Redefining enterprise decision systems,” Int. J. Sci. Res. Sci. Eng. Technol., vol. 9, no. 4, pp. 547–572, 2022. doi: 10.32628/IJSRSET221656

[26] J. Jangid and S. Malhotra, “Optimizing Software Upgrades in Optical Transport Networks: Challenges and Best Practices,” Nanotechnol. Percept., vol. 18, no. 2, pp. 194–206, 2022. [Online]. Available: https://nano-ntp.com/index.php/nano/article/view/5169

[27] Tolba, A., Mostafa, N. N., & Sallam, K. M. (2024). *Hybrid Deep Learning-Based Model for Intrusion Detection. Artificial Intelligence in Cybersecurity*, 1, 1–11.

[28] F. Yashu et al., “Thread mitigation in cloud native application development,” Webology, vol. 18, no. 6, pp. 10160–10161, 2021. [Online]. Available: https://www.webology.org/abstract.php?id=5338s

[29] Jain, M., & Srihari, A. (2024). *Comparison of Machine Learning Algorithm in Intrusion Detection Systems: A Review Using Binary Logistic Regression.* In Hybrid Approaches to Intrusion Detection: Combining Machine Learning and Rule-Based Systems. (Mentioned as reference 1 in context of a hybrid ML + rule-based IDS)

[30] K. Zhang et al., “A cross-domain federated learning framework for wireless human sensing,” IEEE Netw., vol. 36, no. 5, pp. 122–128, 2022.