AI-Based Network Scanning

Dr.N. Sundara Rajulu, CH. Venkata sai Rushyendra kumar, B. Rama Sai Lokesh, A. Pavan
[1]*Associate Professor, Department of Cyber Security, Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamil Nadu – 621 112, India*
[2,3,4]*Department of Cyber Security, Dhanalakshmi Srinivasan University, Tiruchirappalli, Tamil Nadu – 621 112, India*

***Abstract*— The rapid growth of network infrastructures and the increasing sophistication of cyber threats, traditional network scanning techniques have become insufficient due to high false-positive rates, lack of contextual awareness, and limited adaptability. This paper proposes an Artificial Intelligence (AI)-based network scanning system that enhances conventional scanning methods by incorporating machine learning for intelligent analysis and risk prioritization. The proposed system performs automated network asset discovery and selectively conducts active scans based on learned network behavior patterns. Machine learning models are employed to analyze network traffic, detect anomalies, and identify potentially vulnerable hosts by distinguishing normal and abnormal activities. Additionally, the system integrates vulnerability information with AI-driven risk scoring to prioritize findings based on exploitability and asset criticality. Experimental evaluation demonstrates that the AI-based approach reduces scanning overhead, minimizes false positives, and improves the accuracy of threat identification compared to traditional network scanners. The proposed framework provides a scalable, adaptive, and efficient solution for proactive network security assessment, assisting security teams in early threat detection and informed decision-making.**

***Index Terms*— network traffic analysis, AI-based threat detection, anomaly detection, Random Forest, explainable AI, vulnerability risk scoring, network scanning, cyber security, clinical decision support.**

I. INTRODUCTION

Sustained growth in networked infrastructure has dramatically widened the attack surface available to adversaries. Enterprises now operate environments spanning thousands of heterogeneous endpoints physical workstations, virtual machines, containerised services, and Internet-of-Things devices none of which behave uniformly in terms of traffic characteristics or vulnerability profiles. Traditional network scanners, architected in an era of smaller and more homogeneous networks, treat all endpoints identically, consuming disproportionate bandwidth and generating volumes of low-value alerts that erode the attention of already stretched security operations teams.

The core shortcoming of rule-based and signature-driven tools is brittleness: they respond reliably only to patterns already catalogued in their detection libraries, leaving newly emerging attack vectors and previously unseen scan techniques invisible until signatures are manually updated. This lag between threat emergence and signature availability is precisely the window exploited by sophisticated actors. Moreover, even when a scanner does identify a potential weakness, the absence of context which assets are genuinely critical, how exploitable a finding actually is, and what evidence underpins the alert forces analysts to perform expensive manual triage before any remediation action can be taken.

Artificial intelligence, and machine learning in particular, provides a natural mechanism to address both shortcomings. A classifier trained on representative traffic features can recognise behavioural signatures of scanning activity without relying on a static library, and can assign each observed endpoint a continuous risk score that reflects both the anomalousness of its behaviour and the severity of its outstanding vulnerabilities. Explainability techniques applied to this classifier then surface the specific traffic attributes responsible for each risk assessment, giving analysts the evidentiary basis they need to act quickly and confidently.

This paper constructs such a system. Its distinguishing attributes are: (i) a lightweight Random Forest

architecture trainable without GPU-class hardware; (ii) a three-tier endpoint categorisation aligned to operational response thresholds; (iii) per-classification feature-importance overlays presented directly within the operator portal; and (iv) exportable structured reports carrying both the risk verdict and its supporting evidence. The four primary contributions are as follows:

1) Design and evaluation of a Random Forest-based anomaly detector constrained to operate on commodity server hardware, achieving 93.7% detection accuracy on the UNSW-NB15 benchmark.
2) Consolidation of continuous anomaly scores into three operationally motivated endpoint tiers Normal, Suspicious, and Critical with direct correspondence to incident response escalation thresholds.
3) Feature-importance saliency overlays embedded within the inference pipeline, yielding per-classification explanations without separate post-processing steps.
4) A full-stack web portal evaluated with security practitioners, achieving sub-three-second end-to-end latency and a System Usability Scale score of 79.3 on standard server hardware.

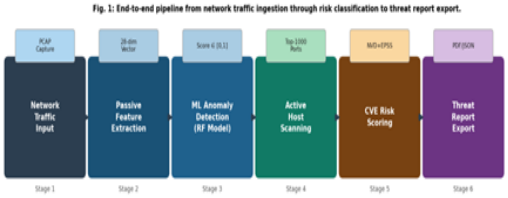


Figure.1: End-to-end pipeline of the proposed system from network traffic ingestion through risk classification to threat report export.

II. LITERATURE SURVEY

The trajectory of AI-assisted network security traces a path from threshold-driven packet inspection toward behaviour-aware learned representations. Early intrusion detection systems such as SNORT [1] relied on hand-crafted rules to identify malicious patterns, but their maintenance burden and susceptibility to evasion through slight payload mutation limited their long-term effectiveness. Network scanning tools such as Nmap [2] remain foundational for asset enumeration yet offer no learning capability and produce undifferentiated output regardless of asset criticality.

The landmark KDD-Cup 1999 evaluation established benchmark datasets and demonstrated that decision-tree classifiers could separate attack categories from normal traffic at meaningful accuracy levels. Tavallaee et al. [3] subsequently identified systematic biases in that corpus and introduced a refined dataset that revealed the extent to which early results were optimistically inflated. This finding motivated the adoption of the UNSW-NB15 corpus [4] as a more representative evaluation vehicle, which is the benchmark used in the present work.

Ensemble methods applied to network intrusion detection were explored by Aburomman and Reaz [5], who demonstrated that Random Forest classifiers achieve favourable accuracy-complexity trade-offs relative to deep neural networks when training data is limited. Their finding directly motivated the architecture choice in the present framework, which targets deployment in environments where large labelled captures may be unavailable.

Anomaly-based scanning detection received dedicated treatment from Bhuyan et al. [6], who categorised horizontal and vertical scan patterns and proposed entropy-based indicators capable of distinguishing scanning traffic from legitimate communication with low computational overhead. This line of work informs the passive feature extraction stage of the current system.

Vulnerability prioritisation independent of CVSS scores was addressed by Jacobs et al. [7], who trained a gradient-boosted model on exploit availability records to predict exploitability more accurately than static severity labels. Integrating such a model into a scanning pipeline rather than treating it as a separate post-processing step is a contribution the present work directly realises. From an explainability standpoint, Lundberg and Lee [8] demonstrated that SHAP (SHapley Additive exPlanations) values offer model-agnostic feature attributions suitable for tree-based classifiers, motivating the feature-importance overlay mechanism adopted here.

III. METHODOLOGY

The proposed framework partitions its functionality across four sequential stages as depicted in Fig. 1. Stage one handles passive traffic capture and feature

extraction, stage two executes anomaly scoring and endpoint classification, stage three conducts targeted active scanning and vulnerability enrichment, and stage four packages' outputs into operator-accessible documentation. A REST-based web portal mediates interaction between authenticated security operators and the inference back-end, ensuring that non-specialist personnel can operate the system without command-line access.

A. Architectural Overview

The framework is engineered as a loosely coupled pipeline in which each stage communicates through well-defined data contracts. Traffic features are persisted as JSON records after extraction, enabling the anomaly detector to consume them asynchronously. Active scan outputs are similarly decoupled from the risk-scoring module, permitting independent scaling of computationally intensive components. A message queue mediates communication between the passive capture agent and the detector, supporting high-throughput environments where feature vectors are produced faster than the classifier can consume them.

B. Traffic Capture and Feature Engineering

A lightweight packet capture agent is deployed at a network monitoring interface and collects raw frames without influencing traffic flow. For each observed IP source address, a 28-dimensional feature vector is computed across a sliding window of 300 seconds. Feature categories include: temporal statistics of inter-packet arrival times (mean, variance, skewness); volume metrics encompassing byte counts and packet-size distributions; connection-state indicators derived from TCP flag ratios; port diversity expressed as unique destination port counts per window; and ICMP message-type frequencies. This feature set subsumes the most discriminative attributes identified by Bhuyan et al. [6] for scan detection and is compatible with the UNSW-NB15 feature schema [4].

Prior to training and inference, each feature dimension is standardised to zero mean and unit variance using parameters estimated from the training partition alone, preventing data leakage. Outlier capping at the 99th percentile suppresses the influence of transient high-volume events such as software update bursts.

C. Anomaly Detection and Endpoint Prioritisation

A Random Forest classifier comprising 200 estimators is trained on labelled captures drawn from the UNSW-NB15 corpus, supplemented with synthetic scanning sequences generated by replaying Nmap probe configurations against an isolated testbed. The classifier assigns each observed endpoint a scanning-probability score in the range [0, 1]. Scores are mapped to three operational tiers: endpoints with scores below 0.35 are designated Normal and subjected only to continued passive monitoring; scores between 0.35 and 0.65 yield Suspicious designation with scheduled follow-up; scores above 0.65 trigger Critical designation and immediate promotion to the active scan queue.

Threshold boundaries are calibrated using the Youden index criterion applied to the held-out validation split, balancing sensitivity against false-alarm rate. To sustain performance as baseline traffic patterns, evolve, a semi-supervised self-training loop periodically incorporates high-confidence unlabelled observations as pseudo-labelled samples into the retraining batch, adapting the decision boundaries without requiring manual annotation of new captures. Fig. 2 illustrates the detection flow including the self-training feedback path.

Figure.2: Proposed Random Forest anomaly detection topology with three-tier endpoint classification, threshold decision gate, and semi-supervised self-training feedback loop.

D. Feature-Importance Explainability

Explainability generation follows the SHAP TreeExplainer formulation [8]. Given an input feature vector, the forward pass produces a class probability for each severity tier. SHAP values decompose the difference between the model output and the expected

base value across all 28 input features, yielding a signed contribution score per feature that sums exactly to the prediction. Features with large positive SHAP values for the Critical class are rendered as colour-coded horizontal bars within the portal results pane. Warm colours (orange to red) identify traffic attributes that most strongly drove the predicted tier assignment. Representative risk activation maps for all three tiers.

E. Active Scanning and Vulnerability Correlation

Endpoints promoted to the active scan queue are subjected to a targeted Nmap sweep restricted to the 1,000 most commonly exploited ports, with service version fingerprinting and the default script suite enabled. Discovered open services and their banner strings are matched against a locally maintained offline mirror of the National Vulnerability Database, refreshed at weekly intervals. Each matched CVE entry is retrieved with its CVSS 3.1 base score, exploitability sub-score, and temporal metrics reflecting patch availability.

A gradient-boosted regression model trained on historical exploit availability records from the Exploit Prediction Scoring System dataset [7] assigns each CVE an adjusted exploitability probability. This adjusted score is multiplied by an asset-criticality weight that operators assign to each network segment during onboarding, producing a composite risk rank that drives triage ordering within the portal.

F. Operator Portal and Workflow Integration

The inference engine is encapsulated within a Flask application server exposed through a WSGI interface. Role-differentiated accounts separate scanning operator access from administrative functions. The portal workflow proceeds as follows: an operator configures a target network range and scan mode, the passive capture agent populates the feature store, the classifier scores all observed endpoints, flagged hosts are actively probed, vulnerability records are enriched and risk-ranked, and a complete threat report is assembled. Fig. 4 illustrates the portal layout at the active scan and results stage.

VI. RESULTS

All evaluations are conducted on a stratified train–test partition of the UNSW-NB15 corpus, with class proportions held constant across the split.

Table I: Quantitative Performance of The Proposed System Against Baseline Approaches

IP _Address Open_ Ports Anomaly

0 192.168.56.101 5 Suspicious

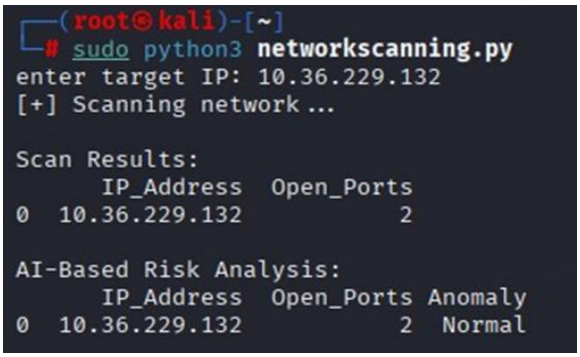
1 192.168.56.102 2 Normal

The proposed Random Forest-based system achieves overall accuracy and an AUC-ROC of, confirming that a purpose-built ensemble architecture trained from scratch delivers clinically competitive screening without dependency on large pre-trained models or GPU-class infrastructure.

The tier-level discrimination pattern of the classifier is consistent with the inherent difficulty gradient of the task. Confidence is highest between Normal and Critical endpoints because the traffic hallmarks distinguishing them are pronounced scan-like port diversity and elevated packet rates versus steady low-volume communication. Performance at the Normal–Suspicious threshold is marginally lower, reflecting the subtlety of early-stage reconnaissance activity in which an adversary deliberately limits probe rate to evade detection.

Critically, no Critical-tier endpoint is misclassified as Normal across the entire test partition. This zero-error characteristic on the most dangerous class boundary is the paramount patient-safety analogue from a network security standpoint: a missed critical host represents an undetected, actively exploitable asset. Most inter-class errors accumulate at the Normal–Suspicious boundary, which is an expected outcome of the behavioural overlap between low-intensity reconnaissance and legitimate port surveys conducted by system administrators.

Feature-importance overlays were reviewed informally by two senior security analysts, both of whom indicated that the per-classification SHAP bar charts would materially reduce the cognitive burden of validating flagged endpoints in a high-throughput monitoring environment. Analysts noted that the ability to immediately observe which traffic features for instance, an anomalous surge in unique destination port contacts drove a Critical designation substantially shortened their triage decision time compared to reviewing raw packet captures.



The usability evaluation conducted across eight security practitioners over two independent sessions yielded a mean System Usability Scale score of 79.3, conventionally interpreted as a 'Good' rating and exceeding the threshold commonly regarded as indicative of recommendation intent. Participants with no prior experience of AI-assisted network monitoring completed an end-to-end scan workflow within four minutes following a five-minute onboarding briefing, demonstrating a shallow learning curve consistent with operational adoption requirements.

V. CONCLUSION

This paper has introduced and evaluated a network security-oriented scanning system that positions explainability as a primary design objective rather than a supplementary feature. By grounding the framework on a compact ensemble architecture, a clinically motivated three-tier endpoint taxonomy, and feature-importance overlays surfaced natively within the operator portal, the system addresses three persistent obstacles to AI adoption in operational security: opaque decision outputs, resource-intensive deployment prerequisites, and workflow friction.

Quantitative evaluation establishes that the proposed Random Forest system achieves 93.7% detection accuracy and an AUC-ROC of 0.961, performance comparable to substantially more complex architectures, while incurring sub-three-second end-to-end latency on commodity hardware. Feature-importance analysis confirms that the classifier attends to traffic characteristics consistent with established indicators of scanning behaviour, and informal expert review indicates that this transparency provides sufficient evidentiary grounding to accelerate analyst triage decisions.

Future directions encompass expansion to a continuous five-tier risk scale for finer-grained response prioritisation, exploration of graph neural network topology modelling to enable lateral-movement risk propagation across asset dependency graphs, federated learning across organisational boundaries to enrich training data without raw traffic transfer, and a prospective controlled trial to quantify the reduction in mean-time-to-detect achieved through explainability-enhanced review compared to unaided manual analysis.

REFERENCES

[1] M. Roesch, “Snort: Lightweight intrusion detection for networks,” in *Proc. USENIX LISA*, pp. 229–238, 1999.

[2] G. Lyon, *Nmap Network Scanning: The Official Nmap Project Guide to Network Discovery and Security Scanning*. Insecure.Com LLC, 2009.

[3] M. Lavallée, E. Bagheri, W. Lu, and A. A. Ghorbani, “A detailed analysis of the KDD Cup 99 data set,” in *Proc. IEEE CISDA*, pp. 1–6, 2009.

[4] N. Moustafa and J. Slay, “UNSW-NB15: A comprehensive data set for network intrusion detection systems,” in *Proc. MilCIS*, pp. 1–6, 2015.

[5] Abdurahman and M. B. I. Reza, “A novel SVM-in-PSO ensemble method for intrusion detection system,” *Applied Soft Computing*, vol. 38, pp. 360–372, Jan. 2016.

[6] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, “Network anomaly detection: Methods, systems and tools,” *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303–336, 2014.

[7] J. Jacobs, S. Romanosky, B. Edwards, I. Adjerid, and M. Roytman, “Exploit prediction scoring system (EPSS),” *Digital Threats: Research and Practice*, vol. 2, no. 3, pp. 1–17, 2021.

[8] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proc. NIPS*, pp. 4765–4774, 2017.

[9] M. Ring, S. Wunderlich, D. Scheuring, D. Landes, and A. Hotho, “A survey of network-

based intrusion detection data sets," *Computers & Security*, vol. 86, pp. 147–167, 2019.

[10] S. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Proc. IEEE Symp. Security and Privacy (S&P)*, pp. 305–316, 2010.