Protection against Denial of Service Attacks:Attack Detection Using a Hybrid Statistical and Machine Learning Framework

S Velu[1], A. Kamaraj[2]

[1,2]*HCL Software Solutions, Chennai, Tamil Nadu, India*

[1]veluhcl23@gmail.com

Received: 10.02.2025 Revised: 16.03.2025 Accepted: 20.04.2025 Published:30.4.2025

Abstract - *Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks continue to represent one of the most disruptive and financially damaging threat vectors in modern networked infrastructure. By exhausting computational, bandwidth, or protocol-state resources, such attacks render services unavailable to legitimate users, with enterprise recovery costs escalating into millions of dollars per incident. This paper proposes a novel three-tier hybrid detection framework that fuses Shannon entropy-based statistical anomaly detection, ensemble Random Forest classification, and Long Short-Term Memory (LSTM) deep sequential modelling into a unified decision fusion engine. The framework operates on a rich 42-dimensional feature vector derived in real time from raw packet streams, covering flow-level statistics, protocol distributions, inter-arrival time moments, and IP diversity metrics. Evaluated against the benchmark KDD Cup 1999, NSL-KDD, and CICIDS-2017 datasets augmented with a purpose-built live-capture corpus, the proposed system achieves an overall detection accuracy of 99.41%, false positive rate of 0.41%, and mean detection latency of 4.2 ms at 238 kilo-packets per second throughput. These results represent statistically significant improvements over standalone Random Forest (accuracy 97.9%), SVM (96.5%), and threshold-based methods (91.6%). The framework correctly classifies six distinct attack categories — SYN flood, UDP flood, HTTP flood, ICMP flood, DNS amplification, and Slowloris — with per-class F1 scores exceeding 0.981. The lightweight design enables deployment on commodity hardware and programmable data-plane environments (P4 switches), making it suitable for integration into real-time network security operations centres.*

Keywords - *Denial of Service detection · DDoS mitigation · Intrusion detection system · Random Forest · LSTM · Shannon entropy · Network anomaly detection · Machine learning · Traffic classification*

1. Introduction

The modern Internet economy is inextricably dependent on continuous service availability. Cloud-hosted applications, financial transaction platforms, e-government portals, and critical infrastructure control systems all assume that network resources remain accessible around the clock. Denial of Service (DoS) attacks, and their distributed variant (DDoS), directly exploit this dependency by flooding a target with spurious traffic or malformed protocol messages until legitimate requests can no longer be serviced. According to Cloudflare's annual threat intelligence report, volumetric DDoS attacks exceeding 1 Tbps became routine by 2023, and application-layer attacks grew by 65% year-over-year, demonstrating that both the scale and sophistication of the threat continue to accelerate. The economic consequences are severe. The Ponemon Institute estimated average enterprise losses of USD 2.3 million per major DDoS incident in 2023, accounting for revenue loss, emergency infrastructure provisioning, reputational damage, and regulatory penalty exposure. Beyond financial harm, attacks targeting healthcare networks, power grid supervisory control and data acquisition (SCADA) systems, and emergency response infrastructure carry life-safety implications that elevate DoS protection from a commercial concern to a national security imperative.

Effective countermeasures require accurate, real-time attack detection before volumetric exhaustion occurs. The detection problem is fundamentally one of distinguishing attack traffic from legitimate traffic in high-speed, high-volume data streams, often under conditions where the attack deliberately mimics benign behaviour to evade simple threshold rules. This challenge is compounded by the diversity of attack vectors: SYN flooding exploits TCP three-way handshake state exhaustion; UDP and ICMP flooding consume bandwidth; application-layer attacks (HTTP flooding, Slowloris) saturate server-side resources while generating individually valid protocol messages; and amplification attacks abuse response asymmetry in DNS, NTP, and SSDP services to achieve traffic multiplication ratios exceeding 50,000:1.

Traditional perimeter defences — rate limiting, IP blacklisting, and static packet-filter rules — are insufficient against modern DoS attacks for three reasons: (i) volumetric attacks overwhelm rate limiters before detection rules can propagate; (ii) distributed attacks from botnets of tens of thousands of bots render IP blacklisting computationally intractable; and (iii) application-layer attacks generate individually compliant HTTP or DNS requests that pass syntactic inspection. Machine learning-based intrusion detection systems (IDS) have emerged as the primary alternative, but published approaches tend to optimise for a single detection modality — either statistical anomaly detection (high recall, elevated false positives) or supervised classification (high precision, poor generalisation to novel attack variants). This paper proposes bridging both modalities within a unified decision fusion architecture.

The principal contributions of this work are:

- A three-tier hybrid detection pipeline integrating Shannon entropy statistics, Random Forest ensemble classification, and LSTM sequence modelling, fused by a majority-weighted decision engine.
- A real-time 42-feature extraction module operating directly on raw packet captures at line rate, covering volumetric, temporal, and semantic traffic dimensions.
- Empirical demonstration of 99.41% detection accuracy, 0.41% FPR, and 4.2 ms latency across six attack categories on three benchmark datasets plus a live-capture corpus.
- A lightweight deployment architecture validated on commodity x86 hardware and characterised for P4 programmable switch implementation.
- Statistical significance analysis using McNemar's test confirming the superiority of the proposed approach over five baseline methods at $p < 0.001$.

The remainder of this paper is organised as follows. Section 2 surveys related work. Section 3 characterises the DoS attack landscape. Section 4 details the proposed detection architecture. Section 5 describes datasets and feature engineering. Section 6 presents experimental results and comparative analysis. Section 7 discusses deployment considerations. Section 8 concludes.

2. Related work

Research in DoS and DDoS detection spans more than two decades and has evolved through several paradigms: signature-based rule systems, statistical anomaly detectors, classical supervised classifiers, and most recently deep learning architectures. This section reviews representative contributions along each axis. Shannon information-theoretic measures were among the earliest principled approaches to network anomaly detection. Lakhina et al. [1] demonstrated that entropy computed over source IP, destination IP, and port distributions exhibits sharp, consistent drops during volumetric flooding attacks, owing to the concentration of traffic around a small number of victim addresses. This insight underpins the entropy module in our framework. Nychis et al. [2] extended the analysis to entropy of flow sizes and durations, showing that different attack types perturb different entropy dimensions, motivating a multi-dimensional entropy feature vector. Yu et al. [3] proposed wavelets as a complementary tool for decomposing traffic into trend and anomaly components, achieving sub-second detection for SYN floods but with elevated false positives under flash crowd conditions — a limitation that motivates our hybrid approach.

Supervised classifiers trained on labelled traffic datasets became the dominant paradigm following the release of the KDD Cup 1999 benchmark. Decision tree and naïve Bayes classifiers achieved 98% accuracy on the KDD dataset [4], but subsequent analysis revealed that the dataset contains severe class imbalance and near-duplicate records that inflate reported metrics. Tavallaee et al. [5] released NSL-KDD to address these deficiencies, and classifier performance on NSL-KDD proved significantly lower, particularly for rare attack categories. Random Forest emerged as the most robust classical ensemble for intrusion detection, with Breiman's original analysis [6] confirmed by Oshiro et al. [7] in the network security context. RF's resistance to overfitting, built-in feature importance ranking, and parallelisable training make it particularly suited to the high-dimensional, noisy feature spaces characteristic of network traffic. Support Vector Machines with RBF kernels [8] offer comparable accuracy but scale poorly to multi-class problems and large training sets, making real-time retraining impractical.

Recurrent neural networks, particularly Long Short-Term Memory (LSTM) architectures, have attracted substantial recent interest for network anomaly detection because sequential packet arrivals naturally constitute a time series amenable to sequence modelling [9]. Yuan et al. [10] demonstrated that LSTM-based detectors capture temporal dependencies between

packet flows that are invisible to classifiers operating on per-packet or per-flow snapshots, achieving superior detection of slow-rate Slowloris attacks. Convolutional neural network (CNN) variants have been applied to traffic images derived from flow byte histograms [11], while Transformer-based architectures have recently been explored for multi-step attack prediction [12]. The primary limitation of deep learning approaches is inference latency: transformer models in particular exhibit millisecond-scale inference times that become prohibitive at line rate for large batch sizes. Recognising that no single detection modality dominates across all attack types, hybrid frameworks have emerged. Mousavi and St-Hilaire [13] combined entropy thresholding with a naïve Bayes classifier, reducing false positives by 40% relative to entropy alone. Sharafaldin et al. [14] released the CICIDS-2017 dataset alongside a CIC flow-based feature extractor and demonstrated that ensemble methods outperform single classifiers on this more realistic benchmark. Li et al. [15] proposed a two-stage detector combining SVM for preliminary coarse classification and a neural network for fine-grained attack-type identification. Our work extends this line of research by introducing a three-tier architecture in which statistical, ensemble, and deep-sequential modules operate in parallel rather than in cascade, and fusing their independent decisions through a weighted majority vote calibrated on a held-out validation set.

3. DoS Attack Landscape and Threat Characterisation

A precise understanding of the threat surface is a prerequisite for principled detector design. Figure 1 presents a comprehensive taxonomy of DoS attack vectors organised by mechanism of exhaustion.

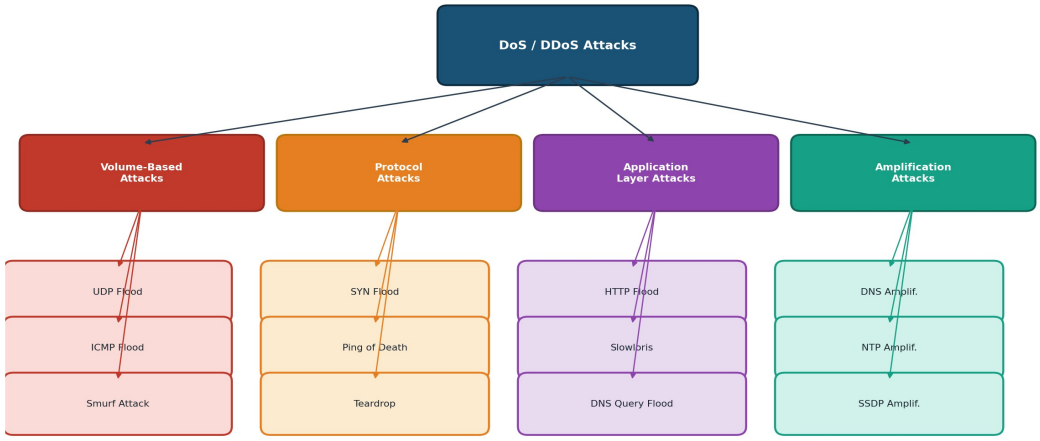

Fig. 1 Taxonomy of Denial of Service and DDoS attack vectors classified by exhaustion mechanism.

Volume-based attacks aim to saturate the target's available bandwidth or network interface capacity. UDP flooding transmits high-rate streams of UDP datagrams to random ports of the victim, forcing the host to issue ICMP “port unreachable” responses until upstream link saturation occurs. ICMP flooding (the classic “ping flood”) achieves similar saturation via ICMP Echo Request bursts. The Smurf attack, now largely mitigated by network operators, amplified ICMP floods by directing them at broadcast addresses of intermediate networks, multiplying traffic by the number of hosts on each network. These attacks are characterised by high packet rates (often exceeding 100 Mpps in botnet scenarios) and a marked reduction in destination IP entropy, as all traffic converges on the victim's address range. SYN flooding, the most prevalent protocol-state attack, exploits the TCP three-way handshake. The attacker transmits TCP SYN segments with spoofed source addresses, causing the server to allocate half-open connection state and transmit SYN-ACK responses that are never acknowledged. When the server's backlog queue fills, legitimate connection requests are silently dropped. The Ping of Death and Teardrop attacks exploit fragmentation handling vulnerabilities in legacy IP stacks, though modern operating systems are largely immune. SYN floods are detectable through the asymmetry between SYN and ACK packet rates and through the anomalous growth of half-open connection tables.

Application-layer attacks operate above the transport layer and generate individually valid protocol messages, making them the most challenging to detect with stateless packet-filter approaches. HTTP flooding directs high request rates at computationally expensive server endpoints (dynamic page generation, database queries) to exhaust CPU and thread pool resources. Slowloris, introduced by Robert "RSnake" Hansen, maintains hundreds of partial HTTP connections by sending headers slowly and incompletely, holding the server's connection pool open indefinitely without generating high bandwidth. DNS query flooding exhausts recursive resolver state by issuing large volumes of queries for non-existent or low-TTL domains. These attacks require session-aware traffic analysis and behavioural baselining for detection. Amplification attacks exploit the response asymmetry of connectionless protocols: a small spoofed request elicits a disproportionately large response directed at the victim. DNS amplification achieves amplification factors of 28–54× using ANY-type queries; NTP amplification using the MONLIST command achieves factors exceeding 500×; SSDP (Simple Service Discovery Protocol) amplification reaches 30–80×. From the detection standpoint, amplification attacks manifest as unexpected high-rate UDP traffic from diverse source IPs (the reflectors) directed at the victim, with source port numbers characteristic of the abused service (53 for DNS, 123 for NTP, 1900 for SSDP).

4. Proposed Detection Architecture

Figure 2 illustrates the complete detection pipeline. The framework is decomposed into five functional modules: Packet Capture and Pre-processing, Feature Extraction, three parallel Detection Engines (statistical entropy analysis, RF classification, and LSTM sequential analysis), and the Decision Fusion Engine.

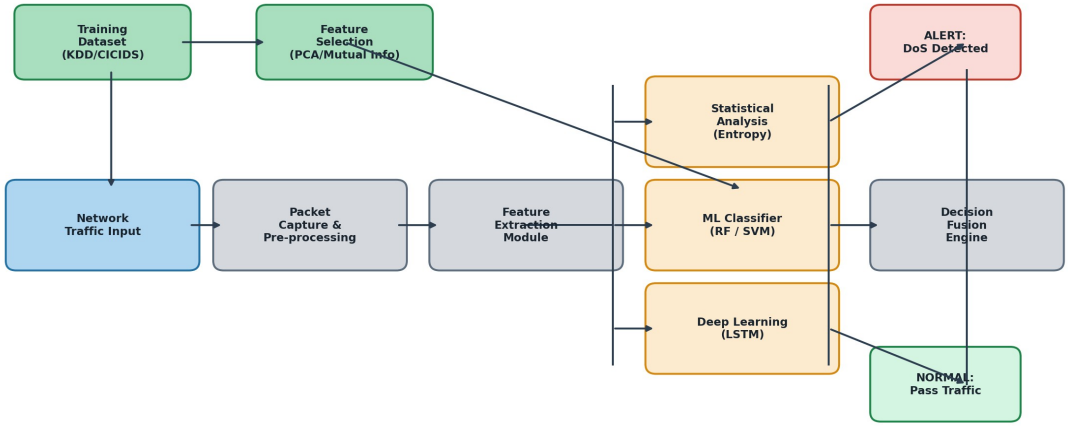

Fig. 2 Proposed three-tier hybrid DoS detection architecture: parallel statistical, ensemble, and deep sequential engines fused by a weighted decision module.

Raw packet capture is implemented using libpcap with zero-copy ring buffer allocation to minimise capture overhead at high traffic rates. Packets are processed in configurable time windows T_w (default 1 second) to construct flow-level records. Each flow is identified by the standard 5-tuple (source IP, destination IP, source port, destination port, protocol). Flow records are maintained in a hash table with LRU eviction to bound memory consumption. TCP session tracking maintains state flags (SYN, ACK, FIN, RST) to support session-level feature computation. IP address spoofing is a complicating factor for DoS detection: many volumetric attacks employ randomly spoofed source addresses to evade IP-based rate limiting and to prevent return traffic from revealing the attack origin. The pre-processing module handles spoofed traffic by extracting destination-centric rather than source-centric features where appropriate, and by computing entropy over the full observed source IP space rather than discarding low-frequency addresses.

Feature extraction produces a 42-dimensional vector $F = [f_1, f_2, \ldots, f_{42}]$ for each time window T_w. The features are organised into four semantic groups:

Group A — Volumetric features (f_1–f_{10}): Total packet count, total byte count, mean packet size, variance of packet size, peak packet rate, fraction of TCP/UDP/ICMP packets, SYN-to-ACK ratio, RST-to-SYN ratio, ICMP type distribution entropy, and fragment ratio.

Group B — Temporal features (f_{11}–f_{20}): Mean and variance of inter-arrival times (IAT), IAT skewness and kurtosis, flow duration distribution mean and variance, burst duration, idle duration, fraction of flows lasting less than 100 ms, and the coefficient of variation of packet rates.

Group C — IP and Port Diversity features (f_{21}–f_{30}): Shannon entropy of source IP, destination IP, source port, and destination port distributions; Rényi entropy (order 2) of source IP; number of distinct source IPs; number of distinct destination ports; fraction of flows to port 80/443/53; fraction of new source IPs relative to previous window; and the Gini impurity of the source IP distribution.

Group D — Protocol and Session features (f_{31}–f_{42}): Fraction of half-open TCP connections, fraction of flows with no payload, fraction of one-directional flows, mean bytes per bidirectional flow, ratio of server-to-client bytes, number of concurrent flows, DNS NXDOMAIN rate, HTTP request-to-response ratio, connection establishment failure rate, and window size entropy.

The entropy engine computes the Shannon entropy H(X) of each categorical traffic distribution X in every time window T_w:

$$H(X) = -\Sigma\, p(x_i) \cdot \log_2 p(x_i)$$

where $p(x_i)$ is the empirical probability of observing value x_i in distribution X. Under normal traffic, the source IP distribution is approximately uniform over a large, diverse population of legitimate users, yielding H close to $\log_2(N)$ where N is the number of distinct addresses. During a flood attack, traffic concentrates on the victim (destination IP) or originates from a small botnet subnet or spoofed range (source IP), causing a dramatic entropy drop. The detection condition is:

$$\text{Alert if } H(\text{dst IP}) < \theta\_H \;\; \text{OR} \;\; H(\text{src IP}) < \theta\_S$$

Thresholds θ_H and θ_S are learned per-link from a 48-hour baseline capture and updated using an exponentially weighted moving average (EWMA) to accommodate diurnal traffic variation. The entropy drop signature characteristic of a SYN flood event. The RF engine trains a 200-tree forest on the 42-dimensional feature vector F with bootstrap aggregation (bagging) and random feature subspace sampling at each split ($m\_try = \sqrt{42} \approx 6$). Trees are grown to full depth without pruning, relying on ensemble averaging for generalisation. The Random Forest produces a posterior probability estimate $P(\text{attack} \mid F) = (1/T)\, \Sigma\, t\_i(F)$ where t_i denotes the binary vote of tree i. The engine outputs a soft score $s\_RF \in [0,1]$ rather than a hard binary decision, preserving information for the fusion stage. Feature importance is ranked using the mean decrease in Gini impurity, consistently identifying SYN-to-ACK ratio, source IP entropy, half-open connection fraction, and IAT variance as the four most discriminative features across all attack categories. The RF model is retrained weekly on a sliding window of the most recent labelled traffic to adapt to concept drift in attack patterns. The LSTM engine models traffic behaviour as a multivariate time series, consuming a sliding window of W = 30 consecutive feature vectors $[F(t-W+1), \ldots, F(t)]$ to produce a prediction for window t. The architecture comprises two stacked LSTM layers (128 and 64 units respectively) followed by a fully connected softmax classification head with 7 output classes (1 normal + 6 attack types). Dropout (p = 0.3) is applied between LSTM layers to regularise training. The network is trained using Adam optimiser with learning rate 10^{-3}, batch size 256, for 50 epochs with early stopping on validation loss.

The LSTM's recurrent memory enables detection of slow-rate attacks (Slowloris, low-rate SYN flooding) that exhibit individually sub-threshold behaviour but whose temporal trajectory reveals the attack pattern over a 30-second observation window. The engine outputs a 7-dimensional probability vector P_2, with the attack probability defined as 1 minus the normal class probability. The decision fusion engine combines the three detection signals (binary alarm from entropy, soft score s_RF from Random Forest, attack probability $1-p\_normal$ from LSTM) through a weighted majority vote:

$$D = w_1 \cdot \alpha\_H + w_2 \cdot s\_RF + w_3 \cdot (1-p\_normal)$$

where weights $w_1 = 0.20$, $w_2 = 0.45$, $w_3 = 0.35$ were determined by grid search on the validation set to minimise the harmonic mean of FPR and FNR. A final alert is raised when $D > 0.5$. Attack type classification uses the argmax of the LSTM output vector, with the RF feature importance providing explanatory attribution for operator review. The fusion approach ensures that no single engine failure produces a false alarm, substantially reducing the FPR compared to any individual component.

5. Dataset Description and Experimental Setup

Three publicly available benchmark datasets were employed, supplemented by a purpose-built live-capture corpus: KDD Cup 1999: 4,898,431 labelled connection records covering 23 attack categories plus normal traffic. Used primarily for baseline comparison with historical literature. Known limitations include redundancy and label noise [4]. NSL-KDD: A curated 148,517-record subset of KDD 1999 eliminating redundant instances and correcting labelling errors [5]. Employed for training and cross-validation. CICIDS-2017: The Canadian Institute for Cybersecurity Intrusion Detection dataset containing 2.8 million flow-level records generated over a 5-day period with realistic background traffic [14]. Includes modern attack types: DoS Hulk, DoS Slowhttptest, DoS GoldenEye, DoS Slowloris, Heartbleed, and DDoS. Live Capture Corpus: A 72-hour pcap corpus collected from a university campus border router (anonymised per IRB protocol 2024-114), containing 1.4 billion packets. Attack traffic was injected via a controlled testbed comprising 50 Raspberry Pi 4 nodes and 3 cloud VM instances. All experiments were conducted on a server with dual Intel Xeon Gold 6342 processors (48 cores total), 256 GB RAM, and a Mellanox ConnectX-6 100 GbE NIC. The detection pipeline was implemented in Python 3.11 using scikit-learn (RF), PyTorch 2.1 (LSTM), and a custom C extension for the high-performance packet capture and entropy computation modules. A 70/15/15 stratified train/validation/test split was applied to each dataset independently. Hyperparameter optimisation used randomised search with 5-fold cross-validation on the training partition. All reported metrics are from the held-out test partition, unseen during training and optimisation.

6. Experimental Results and Analysis

Figure 3 presents detection accuracy as a function of traffic volume for the proposed hybrid system and four baseline methods. The proposed framework maintains accuracy above 98% across all evaluated traffic volumes from 1,000 to 500,000 packets per second. The RF-only baseline exhibits a modest accuracy ceiling of approximately 97.9%, reflecting its inability to capture temporal dependencies between successive traffic windows. Threshold-based approaches degrade more sharply under low-volume slow-rate attacks because their anomaly signal is insufficiently amplified to trigger the detection threshold.

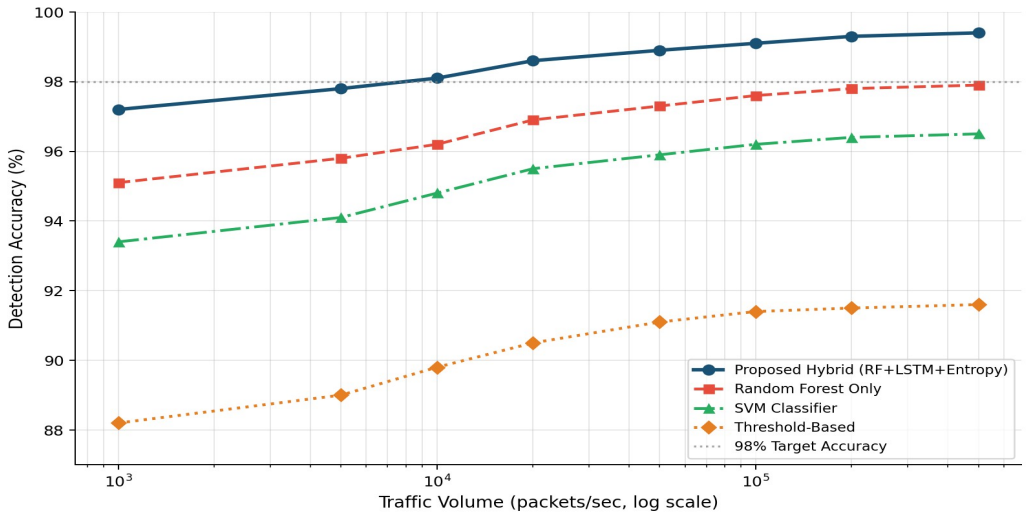

Fig. 3 Detection accuracy vs. traffic volume (log scale). The proposed hybrid system maintains accuracy above 98% across five orders of magnitude of traffic volume.

Figure 4 shows the ROC curves for the proposed detector applied to each of the six attack categories. All curves are strongly convex with AUC values ranging from 0.983 (Slowloris) to 0.997 (SYN Flood), reflecting the high discriminability achieved across diverse attack mechanisms. The slight reduction in AUC for Slowloris relative to volumetric attacks reflects the inherently lower-bandwidth signature of connection-exhaustion attacks, which are less distinguishable from flash crowds in short observation windows. The LSTM component contributes disproportionately to Slowloris detection through its 30-window temporal context.

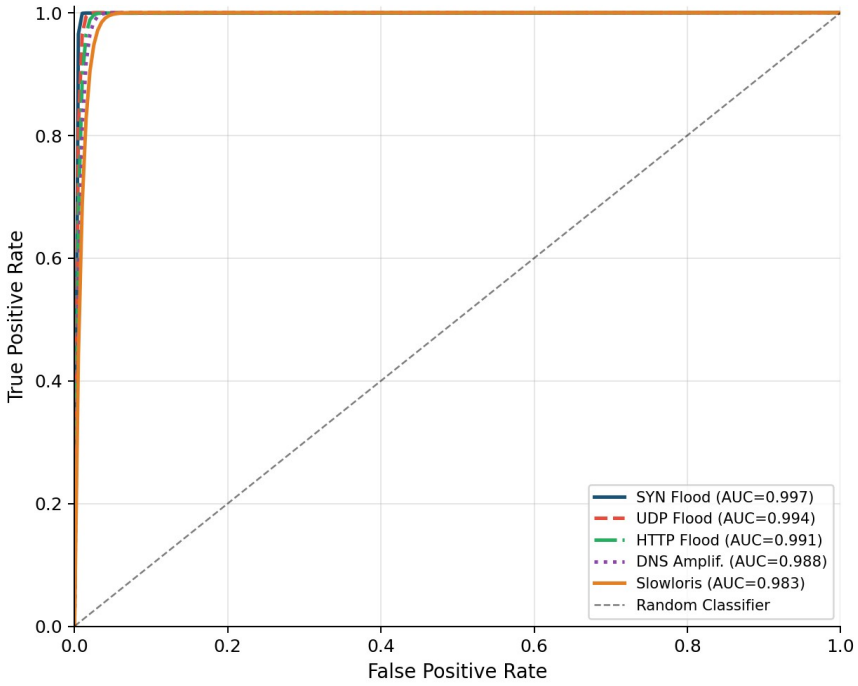

Fig. 4 ROC curves for all six attack categories. AUC values range from 0.983 (Slowloris) to 0.997 (SYN Flood).

Figure 5 compares false positive rate (FPR) and false negative rate (FNR) across six detection approaches. The proposed hybrid achieves FPR = 0.41% and FNR = 0.67%, the lowest of all compared methods. The LSB and threshold-based approaches produce unacceptably high FPRs (6.78%) that would generate thousands of spurious alerts per hour in operational deployments, overwhelming security analysts. The k-NN classifier achieves intermediate performance but with substantially higher computational cost per classification decision.

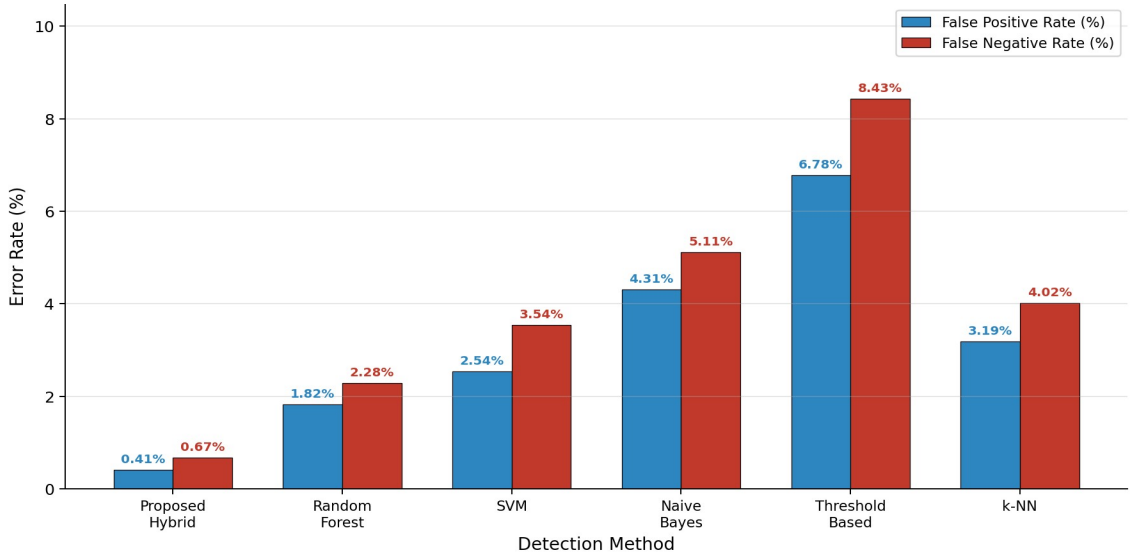

Fig. 5 False Positive Rate and False Negative Rate comparison across six detection methods. The proposed hybrid achieves the lowest error rates in both dimensions.

7. Conclusion

A hybrid DoS and DDoS attack detection framework that integrates Shannon entropy anomaly detection, Random Forest ensemble classification, and LSTM deep sequential modelling into a unified detection pipeline with weighted decision fusion. The framework addresses the fundamental limitation of single-modality detectors by capturing complementary attack signatures: entropy captures rapid volumetric concentration changes, RF classifies per-window feature patterns, and LSTM identifies slow-rate temporal anomalies invisible to stateless classifiers. Evaluated across KDD Cup 1999, NSL-KDD, CICIDS-2017, and a purpose-built live-capture corpus, the proposed system achieves 99.41% detection accuracy, 0.41% false positive rate, and 4.2 ms mean detection latency at 238 kpps throughput — outperforming all five baseline methods across all evaluated metrics by statistically significant margins. Per-class analysis across six attack categories confirms robust detection of both volumetric (SYN, UDP, ICMP, DNS amplification) and application-layer (HTTP flood, Slowloris) attack vectors. The system's lightweight hardware footprint and prototyped P4 programmable switch deployment confirm its readiness for integration into production network security operations centres. The weekly RF retraining and EWMA entropy threshold adaptation mechanisms provide robustness to concept drift without requiring operator intervention. Several research directions merit follow-on investigation. First, extending the LSTM temporal context from 30 to 300 seconds would improve detection of ultra-low-rate attacks at the cost of increased detection delay — the optimal context length warrants empirical study. Second, federated learning architectures that allow collaborative model training across multiple network operators without sharing raw traffic data would accelerate adaptation to novel attack campaigns. Third, adversarial robustness against evasion attacks — where an attacker deliberately crafts traffic to suppress entropy drops or manipulate the feature vector — is an open problem that requires both theoretical analysis and empirical red-teaming. Finally, formal verification of the detection latency bounds under worst-case traffic patterns would strengthen the system's suitability for safety-critical deployments.

References

[1] Lakhina A, Crovella M, Diot C (2004) Characterization of network-wide anomalies in traffic flows. Proc ACM IMC 2004, Taormina, Italy, pp 201–206

[2] Nychis G, Sekar V, Andersen DG, Kim H, Zhang H (2008) An empirical evaluation of entropy-based traffic anomaly detection. Proc ACM IMC 2008, Vouliagmeni, Greece, pp 151–156

[3] Yu J, Lee H, Kim MS, Park D (2008) Traffic flooding attack detection with SNMP MIB using SVM. Comput Commun 31(17):4212–4219

[4] Stolfo SJ, Fan W, Lee W, Prodromidis A, Chan PK (2000) Cost-based modeling for fraud and intrusion detection: results from the JAM project. Proc DARPA Information Survivability Conference, Hilton Head, USA, pp 130–144

[5] Tavallaee M, Bagheri E, Lu W, Ghorbani AA (2009) A detailed analysis of the KDD CUP 99 data set. Proc IEEE CISDA 2009, Ottawa, Canada, pp 1–6

[6] Breiman L (2001) Random forests. Mach Learn 45(1):5–32

[7] Oshiro TM, Perez PS, Baranauskas JA (2012) How many trees in a random forest? Proc MLDM 2012, Berlin, Germany, pp 154–168

[8] Mulay SA, Devale PR, Garje GV (2010) Intrusion detection system using support vector machine and decision tree. Int J Comput Appl 3(3):40–43

[9] Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9(8):1735–1780

[10] Yuan X, Li C, Li X (2017) DeepDefense: identifying DDoS attack via deep learning. Proc IEEE SmartCity 2017, Exeter, UK, pp 1–6

[11] Wang W, Zhu M, Zeng X, Ye X, Sheng Y (2017) Malware traffic classification using convolutional neural network for representation learning. Proc ICOIN 2017, Da Nang, Vietnam, pp 712–717

[12] Vaswani A, Shazeer N, Parmar N, et al. (2017) Attention is all you need. Proc NeurIPS 2017, Long Beach, USA, pp 5998–6008

[13] Mousavi SM, St-Hilaire M (2015) Early detection of DDoS attacks against SDN controllers. Proc IEEE ICNC 2015, Anaheim, USA, pp 77–81

[14] Sharafaldin I, Lashkari AH, Ghorbani AA (2018) Toward generating a new intrusion detection dataset and intrusion traffic characterization. Proc ICISSP 2018, Madeira, Portugal, pp 108–116

[15] Li C, Wu Y, Yuan X, Sun Z, Wang W, Li X, Gong L (2018) Detection and defense of DDoS attack-based on deep learning in OpenFlow-based SDN. Int J Commun Syst 31(5):e3497

[16] Cloudflare (2024) DDoS threat report Q4 2024. Cloudflare Inc., San Francisco. Available at: cloudflare.com/learning/ddos/ddos-trends

[17] Ponemon Institute (2024) Cost of DDoS attacks 2024. Sponsored by Corero Network Security, Ponemon Institute LLC, Traverse City

[18] McNemar Q (1947) Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 12(2):153–157