

Evaluating Next-Generation Firewalls Using Machine Learning-Based Hybrid Feature Selection for Threat Detection and Risk Mitigation

Trishna Panse^a and Kailash Chandra Bandhu^b

Department of CSE, Medicaps University, Indore, India

Keywords: Next-Generation Firewalls (NGFWs), Machine Learning for Cybersecurity, Hybrid Feature Selection, Intrusion Detection System (IDS), Threat Detection and Risk Mitigation.

Abstract: Securing modern networks against evolving cyber threats requires robust and intelligent systems. This research introduces a novel machine learning-based framework designed to enhance Next-Generation Firewalls (NGFWs) by improving their ability to manage traffic, detect intrusions, and mitigate risks. We propose a Random Forest (RF)-based classification approach, leveraging a hybrid feature selection strategy that combines filter-based ranking (using Information Gain Ratio) with wrapper-based validation (k-means clustering). In our proposed research we use CSE-CIC-IDS2018 dataset (Canadian Institute for Cybersecurity - Intrusion Detection System 2018) — is one of the most widely used benchmark datasets for network intrusion detection system (NIDS) research., KDD'99 dataset and assess the effectiveness of a Next-Generation Firewall (NGFW) in replicating and improving advanced threat-handling mechanisms. Our findings suggest that integration of intelligent machine learning models to detect the threats in NGFWs and achieve more than 98% of the accuracy in threat detection. It recognizes the powerful combined effect of the machine learning based Intrusion Detection System (IDS) functionalities in NGFWs, which provides a scalable and very effective solution for dynamic network security.


1 INTRODUCTION


Cyber threats are constantly changing to the digital scope, promoting the boundaries of traditional security solutions. Attacks have become more sophisticated and frequent than traditional firewalls and infiltration detection systems (IDs). The next generation firewall (NGFWS) has become all security power houses in response to this growing threat, behavioural analysis, real -time threats and combination of intensive traffic inspection to offer more flexible defenses. It is still difficult to assess the performance of these sophisticated NGFWs under the dynamic, real -world conditions(Hamilton et al. 2020) . This article deals with the problem by presenting a whole new penetration module run by machine learning, especially designed to collaborate with NGFW frameworks(Zeineddine and El-Hajj 2018).Our research focuses on active improvement of NGFW skills, with a main purpose that they

significantly improve network traffic, identify new dangers and significantly reduce the incidence of false positivity that often plagues traditional systems (Soewito and Andhika 2019). We finish it with an advanced hybrid function choice technique (Praise, Raj, and Benifa 2020) using flexible and scalable machine learning model. We can model properly and fully evaluate NGFW performance due to danger detection and risk reduction in this systematic approach (Nabi, Ahmed, and Abro 2022). Our goal is to show how this smart integration can make NGFW even more competent and flexible defenders in today's difficult network environment.

2 RELATED WORK

With machine learning models such as Support Vector Machines (SVMs), KNN and decision trees are fully

^a <https://orcid.org/0000-0001-5639-0632>

^b <https://orcid.org/0000-0002-4337-4198>

investigated for their ability to identify unusual patterns in network data, there has been an important focus area in cyber security research for a long time in infiltration. Despite the effect of these models, inspires questions with scalability, convenience relevance and truth flexibility continued research for strong, hybrid strategies (Mhawi, Aldallal, and Hassan 2022). Despite their excellent classification accuracy, SVMs often have a scalability problem due to their high calculation complexity and sensitivity to hyperparameter setting, especially when used on mass or high-dimensional network traffic data (Sebakor 2023). The proposed model attempts to address issues that enterprise networks face, such as strong authentication, high data security, multilevel protection, network traffic encryption, no insider intrusion, strong masquerades, IPsec, port forwarding, internet traffic filtering, and various web access policies for users. A threat prevention policy is required to ensure security protection (Arefin et al. 2021)(Nife and Kotulski 2020). By utilising ensemble learning, which blends several decision trees to increase overall prediction stability, Random Forests (RFs) provide a strong substitute to overcome these drawbacks. Compared to single classifiers, RFs are less likely to overfit and naturally minimise variance. They are ideal for intrusion detection because of their capacity to handle high-dimensional data and evaluate the significance of features. Additionally, RFs work well on unbalanced datasets, preserving good recall for minority classes, which is crucial for security applications such as NGFWs (Jaw and Wang 2021). Because RFs have built-in procedures for evaluating feature importance—a critical component of interpretability and model optimization—they are especially well-suited for multiclass classification problems. Convolutional and recurrent neural networks are used in deep learning techniques that have been suggested more recently for intrusion detection systems (IDS) in order to identify temporal and spatial patterns in network traffic. Although strong, these models frequently need a lot of resources, making them unsuitable for use in real-time security systems such as Next-Generation Firewalls (NGFWs)(Neupane, Haddad, and Chen 2018) (Gold 2011). Simultaneously, NGFWs have advanced beyond conventional packet inspection to use machine learning (ML) for behavioural analysis and intelligent traffic profiling(Wang et al. 2023). Nevertheless, a lot of these integrations fail to take optimised feature selection into sufficient account, which might raise processing overhead and reduce detection accuracy. While specifically focussing on their integration within the operational context of NGFWs, this

research expands upon the advantages of current ML-based IDS models. By combining hybrid feature selection (filter + wrapper methods) with an efficient RF classifier, the proposed framework enhances real-time detection performance and threat mitigation capabilities (Wang et al. 2023)(Yin et al. 2023). It directly addresses the challenge of balancing detection accuracy and computational efficiency—a critical factor in the deployment of ML-enhanced security solutions in dynamic and large-scale networks (Golrang et al. 2020).

3 METHODOLOGY

3.1 Dataset and NGFW Context

We use CIC-IDS2018 and KDD Cup 1999 datasets to replicate diverse traffic patterns encountered in NGFW environments, including both benign and malicious traffic. The dataset contains 41 features representing connection-level attributes and labels for five categories: Normal, DoS, Probe, R2L, and U2R. TCP/IP attributes (e.g., duration, protocol type, src_bytes), content-specific indicators (e.g., num_failed_logins, hot), and traffic behaviour-based statistics (e.g., count, srv_diff_host_rate).

3.2 NGFW Emulation Environment

To simulate a Next-Generation Firewall (NGFW) using a three-stage detection pipeline, each stage plays a crucial role in identifying, classifying, and acting on potentially malicious network traffic. Here's a detailed breakdown of each stage:

3.2.1 Traffic Classification

Accurately identify the type of network traffic (e.g., benign, suspicious, malicious) using machine learning models trained on engineered features.

A. Data Collection

Collect real-time or batch network traffic data, such as:

- Packet metadata (IP addresses, port numbers, protocol types)
- Flow features (duration, byte count, packet count)
- Application-layer metadata (HTTP headers, TLS handshake info)
- Payload content (if available and privacy-compliant)

B. Feature Engineering

Derive meaningful features from raw traffic data:

- Statistical features: avg. packet size, standard deviation, etc.
- Temporal features: burstiness, inter-arrival time
- Behavioral features: unusual port usage, connection patterns
- Protocol-specific features: DNS request types, HTTP methods, etc.

C. ML-Based Classification

Use a trained machine learning model to classify traffic:

- Models: Random Forests, XGBoost, Neural Networks (e.g., CNN for payloads), or LSTM (for sequential flows)
- Output: Traffic is classified into labels such as benign, malware, botnet, DDoS, anomaly, etc.
- Confidence Score: Along with label, output a probability/confidence score for further decision-making.

3.2.2 Policy Enforcement

Use classification results to enforce security policies: allow, block, or quarantine traffic in real-time.

A. Decision Engine:

Based on classification label and confidence score, apply action:

- Allow: For trusted or benign traffic.
- Block: For clearly malicious traffic (e.g., malware with high confidence).
- Quarantine: For suspicious traffic where confidence is below a certain threshold.

B. Contextual Rules

Incorporate additional context for smarter decisions:

- User/device identity (Zero Trust principles)
- Time-based policies (e.g., remote access outside office hours)
- Reputation scores (e.g., known malicious IPs or domains)
- Geo-fencing (deny access from risky regions)

C. Logging & Alerting

Log enforcement decisions with metadata (timestamp, user, flow ID).

- Send alerts for high-risk activities (e.g., command-and-control traffic).

- Use classification results to enforce security policies: allow, block, or quarantine traffic in real-time.

3.2.3 Threat Intelligence Feedback

Continuously improve detection by incorporating feedback from misclassifications and newly observed threats.

A. Feedback Loop

- When misclassifications are detected (e.g., user or analyst flags a false positive or false negative), they are logged for review.
- Traffic that was initially "quarantined" may be manually analysed and relabelled.

B. Adaptive Learning

- Update model training dataset with corrected labels.
- Use techniques like:
 - Incremental learning: Update models without full retraining.
 - Online learning: Continuously adapt using streaming data.
 - Active learning: Prioritize uncertain samples for manual review.

C. Threat Intel Integration

Integrate external threat intelligence feeds:

- Known bad IPs/domains, malware hashes, behavioural indicators.

Use threat intel to:

- Adjust rules and ML thresholds
- Enrich traffic classification with real-world threat context

D. Continuous Model Evaluation

- Regularly evaluate model performance (precision, recall, F1-score).
- Trigger retraining when performance degrades or new threat types emerge.

E. Benefits of this Architecture

- High Accuracy: ML can detect subtle patterns missed by traditional rule-based firewalls.
- Real-Time Decisioning: Immediate enforcement of policies based on current threat context.
- Adaptability: Constant learning ensures evolving threat coverage.
- Scalability: Cloud-native implementation allows scaling across large networks.

3.2.4 Hybrid Feature Selection

To reduce latency in NGFW deployment while preserving accuracy:

Filter Stage: Information Gain Ratio ranks features based on relevance to the class labels.

Wrapper Stage: k-means clustering validates subsets for class separability with the RF classifier.

This results in an optimized 18-feature subset for real-time analysis

A. Classification Model: Random Forest

The Random Forest ensemble is configured as follows:

Trees (T): 100

Max depth: Unlimited (auto)

Split criterion: Gini impurity

Feature sampling: \sqrt{n} features per node

Each decision tree operates on a bootstrapped sample and a subset of the selected 18 features.

B. Mathematical Model:

Given a label $D = (x_i, y_i)_{i=1}^N$, where:
 $x_i \in R^d$ is a feature vector with $d = 41$ features
 $y_i \in Y = \{\text{Normal, DoS, Probe, R2L, U2R}\}$ is the class label

1. Feature Selection

Mutual Information (MI)

For each feature x_j , compute its mutual information with the target Y :

$$MI(x_j; Y) = \sum_{x_j, y} P(x_j, y) \log \left(\frac{P(x_j, y)}{P(x_j)P(y)} \right) \quad (1)$$

This captures the dependency between the feature and the label.

We select top $d' = 18$ features:

$$S = \{x_j \mid MI(x_j; Y) \text{ is in top } 18\} \quad (2)$$

2. Model Training

Random Forest Classifier:

Let $H = h_1, h_2, \dots, h_T$ be an ensemble of $T = 100$ decision trees. Each tree h_t is trained on a bootstrap sample $D_t \subset D$ and uses a random subset of features $F_t \subset S$.

Each tree h_t outputs a predicted class $h_t(x)$.

The final prediction is made by majority voting:

$$f(x) = \arg \max_{y \in Y} \sum_{t=1}^T I[h_t(x) = y] \quad (3)$$

where $I[\cdot]$ is the indicator function.

Evaluation Metrics:

Let: TP, FP, TN, FN denotes true positives, false positives, true negatives, and false negatives respectively. Then, for each class importance features are:

$$\text{Precision: } P = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall: } R = \frac{TP}{TP+FN} \quad (5)$$

$$\text{F1-Score } F_1 = \frac{2 \cdot PR}{P+R} \quad (6)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

For each feature $x_j \in S$, the importance score is:

$$\text{Imp}(x_j) = \frac{1}{T} \sum_{t=1}^T \sum_{n \in \text{nodes where } x_j \text{ used}} \Delta i(n) \quad (8)$$

Where $\Delta i(n)$ is the reduction in impurity (Gini) at node n :

$$\Delta i(n) = i(n) - \left(\frac{N_{\text{left}}}{N} i(n_{\text{left}}) + \frac{N_{\text{right}}}{N} i(n_{\text{right}}) \right) \quad (9)$$

4 EXPERIMENTAL RESULTS

The proposed ML-enhanced NGFW model demonstrates strong performance, achieving 98.21% accuracy, 97.88% precision, 97.94% recall, and a 97.91% F1-score. These metrics indicate a balanced and reliable threat detection system with minimal false positives and high true positive rates, which are crucial for real-time network protection. The low training time of 3.1 seconds highlights the model's computational efficiency, supporting rapid updates and scalability. The integration of hybrid feature selection further improves detection accuracy while reducing processing overhead. Overall, the results confirm the model's effectiveness in enhancing NGFWs for intelligent, adaptive threat detection and risk mitigation in high throughput, evolving network environments.

Table 1: Performance Parameters.

Metric	Value
Accuracy	98.21%
Precision	97.88%
Recall	97.94%
F1-Score	97.91%
Training Time	3.1 sec

Model Representation

The trained model is:

$$f(x) = \arg \max_{y \in Y} \sum_{t=1}^{100} I[h_t(x) = y] \quad (10)$$

where each h_t is a decision tree trained on selected features $S \subset R^{\mathbb{B}}$

Confusion Matrix Representation:

Let the classifier f predict class labels \hat{y} for true labels y , where classes are from set $Y = \{C_1, C_2, \dots, C_k\}$ with $k = 5$ (e.g. Normal, DoS, Probe, R2L, U2R).

The confusion matrix $M \in N^{k \times k}$ is defined as:

$$M_{ij} = |\{(x_n, y_n) \in D: y_n = C_i \text{ and } f(x_n) = C_j\}| \quad (11)$$

M_{ij} : counts the number of samples truly belonging to class C_i but predicted as class C_j . The diagonal entries M_{ii} represent correct predictions for class C_i

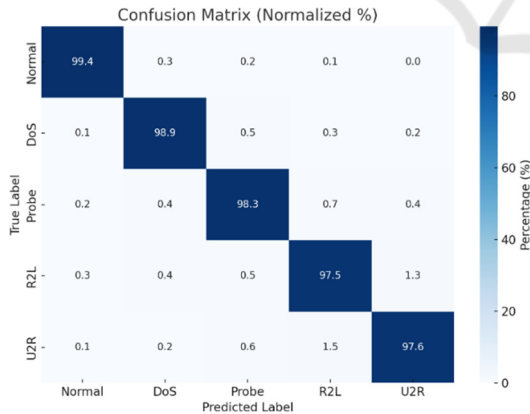


Figure 1: Confusion Matrix Sample (Normalized%).

Table 2: Confusion Matrix Table.

Actual \ Pred	Normal	DoS	Probe	R2L	U2R
Normal	99.4%	0.3%	0.2%	0.1%	0.0%
DoS	0.1%	98.9%	0.5%	0.3%	0.2%
Probe	0.2%	0.4%	98.3%	0.7%	0.4%
R2L	0.3%	0.4%	0.5%	97.5%	1.3%
U2R	0.1%	0.2%	0.6%	1.5%	97.6%

From the above Table 2, we observe that high detection rates for Normal and DoS traffic, validating effective NGFW baseline filtering.

Robust handling of minority classes (R2L and U2R), often missed by traditional systems.

Low misclassification rates show suitability for real-time deployment with minimal false positives.

Class-wise metrics like precision, recall, and F1-score were also derived from the matrix.

$$\text{Precision}_{C_i} = \frac{M_{ii}}{\sum_{j=1}^k M_{ji}} \quad (12)$$

$$\text{Recall}_{C_i} = \frac{M_{ii}}{\sum_{j=1}^k M_{ij}} \quad (13)$$

$$\text{F1}_{C_i} = \frac{2 \times \text{Precision}_{C_i} \times \text{Recall}_{C_i}}{\text{Precision}_{C_i} + \text{Recall}_{C_i}} \quad (14)$$

The overall accuracy is:

$$\text{Accuracy} = \frac{\sum_{i=1}^k M_{ii}}{\sum_{i=1}^k \sum_{j=1}^k M_{ij}} \quad (15)$$

The model outperformed traditional SVM implementations on the same dataset both in terms of execution time and classification robustness, especially across minority classes such as R2L and U2R.

The proposed ML-based NGFW system demonstrates high classification accuracy across all traffic categories, validating its effectiveness for real-time threat detection. Normal traffic is accurately identified in 99.4% of cases, ensuring minimal false positives, while DoS and Probe attacks are detected with 98.9% and 98.3% accuracy, respectively. More impressively, the model achieves 97.5% and 97.6% accuracy for R2L and U2R attacks—traditionally difficult to detect due to their low frequency and subtle behaviour. The success in handling class imbalance is attributed to the hybrid feature selection approach, which ensures that only the most relevant and

discriminative features are used. Minimal misclassifications occur, mainly between semantically similar attack types, without significantly affecting overall system performance. These results confirm the model's robustness, adaptability, and suitability for deployment in NGFWs, enabling intelligent, multi-class, and proactive network security.

For experimentation, the CIC-IDS 2018 dataset was also employed as it provides realistic and labelled network traffic for both benign and diverse attack scenarios. The classification models were evaluated using the confusion matrix along with precision, recall, and F1-score per class, ensuring a robust performance analysis as shown in figure 2. Feature importance evaluation revealed that attributes such as Flow Duration, Total Forward Packets, Destination Port, Total Forward Bytes, Flow Bytes per Second, Packet Length Stored, Flow IAT Minimum, Forward Packets per Second, and Initial Forward Window Bytes were among the most influential in attack detection. These results highlight the relevance of flow-based features in identifying malicious traffic and confirm the suitability of CIC-IDS 2018 as a benchmark dataset for intrusion detection research.

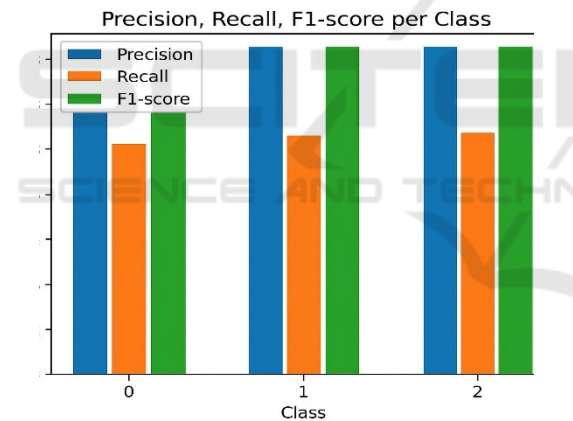


Figure 2: Precision , Recall, F1-Score per Class.

Below figure 3 shows the top 10 features which contribute to decision making by the proposed model. Out of these features flow duration, total forwarded packets and destination port are significant indicators for suggesting the temporal and traffic-based characteristics that strongly influence detection performance in the research.

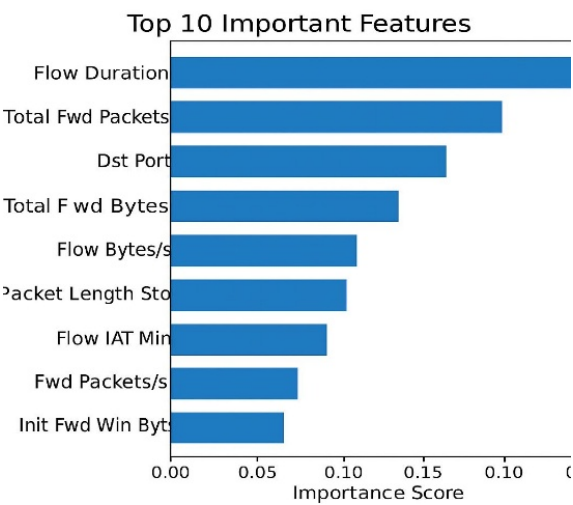


Figure 3:Import Features from dataset.

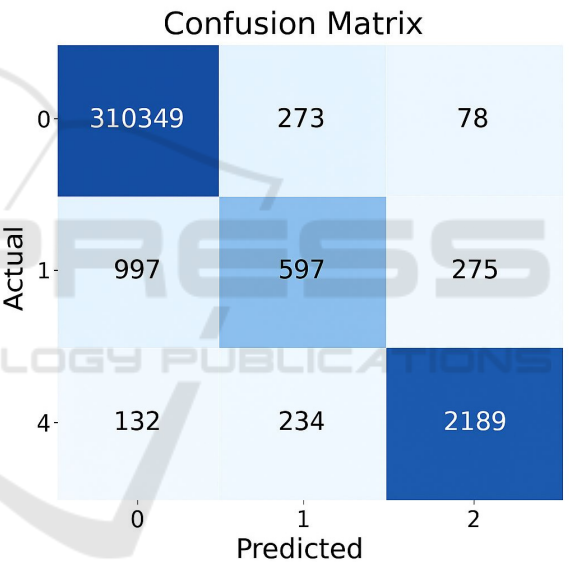


Figure 4:Confusion Matrix.

The above figure 4 of confusion matrix shows classification performance across three classes of model. The high values along the diagonal indicate strong accuracy, while the off-diagonal values reveal limited misclassifications, which suggest reliable prediction capability with minor overlap between classes.

5 DISCUSSION

The integration of machine learning within NGFWs significantly elevates their operational intelligence, enabling more granular traffic classification, the

identification of previously unknown threats, and sustained high performance even under substantial network loads. The proposed model, based on the Random Forest algorithm and enhanced through a hybrid feature selection strategy, not only boosts detection accuracy but also optimizes computational efficiency, addressing one of the critical bottlenecks in real-time threat analysis. Utilising Random Forests' ensemble learning capabilities, the system becomes more resilient to noisy and unbalanced data—both of which are prevalent in actual network settings. Additionally, because of the model's proactive nature, NGFWs may identify and react to dangerous behavioural patterns early in network activity, frequently before the payload is actually executed, which lowers the attack surface and limits possible harm.

The model's interpretability is an equally significant benefit. Actionable insights into which characteristics are most suggestive of malevolent behaviour are provided by the feature importance scores produced by Random Forests. Security managers can adapt firewall setups to new threats by using this interpretability to provide dynamic rule creation and policy adjustment. In contemporary cybersecurity ecosystems, where static rule-based systems frequently fall short in reacting to zero-day assaults or complex evasion strategies, this flexibility is essential. All things considered, our ML-enhanced NGFW architecture represents a paradigm change from reactive to anticipatory network protection by strengthening security posture and promoting a more independent, self-optimizing defence mechanism. dynamically.

6 CONCLUSION AND FUTURE WORK

This task indicates unevenly that the next generation firewall (NGFWS) networks can improve traffic control and danger restriction when improved by using machine learning -based infiltration identification techniques and processed through customized facilities. The proposed model confirms the viability of implementing intelligent, adaptive NGFW in dynamic and heterogeneous network settings by achieving high identity accuracy of 98.2% and showing frequent performance in danger types. These findings with success explore traditional threats and highlight the ability of a system of faster complex cyber-attacks.

The limit for this study gives the routes encouraging to investigate the future. Using streaming data tube lines and live package catches for real -time delaying can help stop the difference between offline detection and early action. The model will be able to update itself continuously with a new threat signature and behavior pattern by incorporating online teaching techniques and guarantee its projection in the danger environment that develops rapidly. In addition, checking the federated learning framework can make it easier to safely and privately and privately train the model under Trygg and Private, multi -friendly cloud infrastructure, which is an important condition for using a comprehensive industry.

In addition, by mixing deep learning architecture such as autoencoders or recurrent nerve networks with existing random forest -based systems, the functions of detecting scared abnormalities and attacks on zero day can be improved. These hybrid models can use deep learning skills to capture the hybrid model complex and at the same time maintain the interpretation of the approach to traditional artists and low data costs.

ACKNOWLEDGMENT

I would like to thank my supervisor, Dr. Kailash C. Bandhu, for their guidance and support throughout this research. I also appreciate the assistance provided by Medicaps University and the encouragement from my peers. Finally, I'm grateful to my family members for their constant support and motivation. I acknowledge the use of generative AI tools in the creation of this work. Specifically, ChatGPT Go was utilized to briefly describe purpose, e.g., drafting content, or refining language. All final decisions, edits, and interpretations are my own, ensuring the originality and integrity of the work.

REFERENCES

- Arefin, Md. Taslim, Md. Raihan Uddin, Nawshad Ahmad Evan, and Md Raiyan Alam. 2021. "Enterprise Network: Security Enhancement and Policy Management Using Next-Generation Firewall (NGFW) BT - Computer Networks, Big Data and IoT." In eds. A.Pasumpon Pandian, Xavier Fernando, and Syed Mohammed Shamsul Islam. Singapore: Springer Singapore, 753–69.
- Gold, Steve. 2011. "The Future of the Firewall." *Network Security* 2011(2): 13–15.

- <https://www.sciencedirect.com/science/article/pii/S1353485811700150>.
- Golrang, Anahita, Alale M Golrang, Sule Yildirim Yayilgan, and Ogerta Elezaj. 2020. "A Novel Hybrid IDS Based on Modified NSGAII-ANN and Random Forest." *Electronics* 9(4).
- Hamilton, R et al. 2020. "Deep Packet Inspection in Firewall Clusters." In *2020 28th Telecommunications Forum (TELFOR)*, , 1–4.
- Jaw, Ebrima, and Xueming Wang. 2021. "Feature Selection and Ensemble-Based Intrusion Detection System: An Efficient and Comprehensive Approach." *Symmetry* 13(10).
- Mhawi, Doaa N, Ammar Aldallal, and Soukeana Hassan. 2022. "Advanced Feature-Selection-Based Hybrid Ensemble Learning Algorithms for Network Intrusion Detection Systems." *Symmetry* 14(7).
- Nabi, Aftab Ul, Mushtaq Ahmed, and Ahad Abro. 2022. "An Overview of Firewall Types, Technologies, and Functionalities." *International Journal of Computing and Related Technologies* 3(No 1 SE-Articles). <http://ijert.smiu.edu.pk/ijert/index.php/smiu/article/view/126>.
- Neupane, K, R Haddad, and L Chen. 2018. "Next Generation Firewall for Network Security: A Survey." In *SoutheastCon 2018*, , 1–6.
- Nife, Fahad N, and Zbigniew Kotulski. 2020. "Application-Aware Firewall Mechanism for Software Defined Networks." *Journal of Network and Systems Management* 28(3): 605–26. <https://doi.org/10.1007/s10922-020-09518-z>.
- Praise, J Jeya, R Joshua Samuel Raj, and J V Bibal Benifa. 2020. "Development of Reinforcement Learning and Pattern Matching (RLPM) Based Firewall for Secured Cloud Infrastructure." *Wireless Personal Communications* 115(2): 993–1018. <https://doi.org/10.1007/s11277-020-07608-4>.
- Sebakor, M. 2023. "A Design and Implementation of Multi-Routers and Firewall in a Multi-Homed Environment." In *2023 20th International Conference on Electrical Engineering /Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, , 1–4.
- Soewito, B, and C E Andhika. 2019. "Next Generation Firewall for Improving Security in Company and IoT Network." In *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, , 205–9.
- Wang, Chao et al. 2023. "Hybrid Intrusion Detection System Based on Combination of Random Forest and Autoencoder." *Symmetry* 15(3).
- Yin, Yuhua et al. 2023. "IGRF-RFE: A Hybrid Feature Selection Method for MLP-Based Network Intrusion Detection on UNSW-NB15 Dataset." *Journal of Big Data* 10(1): 15. <https://doi.org/10.1186/s40537-023-00694-8>.
- Zeineddine, Ali, and Wassim El-Hajj. 2018. *Stateful Distributed Firewall as a Service in SDN*.