Journal of Information Systems Engineering and Management
2025, 10(1s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Cybersecurity in Network Traffic: Integrating Statistical Techniques with AI

Arun Kumar Chaudhary[1], Jitendra Upadhaya[2], Bidur Nepal[3], Murari Karki[4], Madan Kandel[5],Ashok Kumar Mahato[6], Rahul Das[7], Suresh Kumar Sahani*[8],Kameshwar Sahani[9],Garima Sharma[10]

[1]Department of Statistics, Nepal Commerce Campus, Tribhuvan University, Nepal
Email: akchaudhary1@yahoo.com
[2]Nepal Commerce Campus, Tribhuvan University, Nepal Email: jupadhaya@yahoo.com
[3]Department of Statistics, Patan Multiple Campus, Tribhuvan University, Nepal
Email: bnepalpatan027@gmail.com
[4]Department of Statistics, Saraswati Multiple Campus, Tribhuvan University, Nepal
Email: murari.karki@smc.tu.edu.np
[5]Nepal Commerce Campus, Tribhuvan University, Nepal
Email: kandelmadan15@gmail.com
*[6,7,*8]Faculty of Science, Technology and Engineering, Rajarshi Janak University Janakpurdham, Nepal*
[6]ashokkmahato2024@gmail.com
[7]rahulkumardas703732@gmail.com
*[*8]sureshsahani@rju.edu.np*
[9]Department of CivilEngineering, K.U., Nepal
kameshwar.sahani@rju.edu.np
[10]Department of Mathematics, Modi University of Science and Technology, India
sharmagarima2802@gmail.com
Corresponding Authors; sureshsahani@rju.edu.np
jupadhaya@yahoo.com
kameshwar.sahani@ku.edu.np
bnepalpatan027@gmail.com
kandelmadan15@gmail.com
murari.karki@smc.tu.np
sharmagarima2802@gmail.com

ARTICLE INFO

Received: 04 Oct 2024
Revised: 02 Dec 2024
Accepted: 14 Dec 2024

ABSTRACT

This study investigates a hybrid model combining statistical components with AI models to support cybersecurity against attacks on network traffic. To overcome the restrictive nature of classical models such as signature based intrusion detection, this research integrates statistical methods like Z-scores and stack counting with AI techniques including logistic regression and neural network. The hybrid approach is recognized for its ability to identify anomalies, classify network traffic, and respond to changing cyber threats. Results from testing on synthesized data showed that the model was able to detect anomalies to a large extent, as there was considerable accuracy, precision, and recall in targeting malignant traffic. This unified pipeline connects traditional data analytical methods and machine learning in a manner that can expand to cater for real time cyber security needs.

Keywords: Cybersecurity, Network Traffic Analysis, Artificial Intelligence, Statistical Techniques, Anomaly Detection, Hybrid Model, Intrusion Detection, Machine Learning, Network Security.

Introduction:

Data flowing across a computer network at any given time is referred to as network traffic. Often referred to as data traffic, network traffic is transferred over a network in data packets that are then pieced back together by the computer or device receiving it.

To ensure efficient transmission of larger files, data must first be divided into smaller batches before moving over a network or the internet. For the data to be securely transmitted over the network and subsequently opened and read by another user, the network breaks it up, arranges it, and bundles it into data packets. For network traffic to be distributed evenly, each packet uses the most efficient path.

Detecting new or polymorphic threats is difficult for conventional techniques like signature-based intrusion detection systems. Due to the dynamic and ever-changing nature of cyber threats, sophisticated strategies that can adjust and react efficiently are required.

Network traffic security has become a top concern for businesses due to the quick digital transformation and growing interconnectedness of vital assets. Innovative approaches that blend statistical rigor with artificial intelligence's predictive capabilities are needed to meet this problem. With an emphasis on enhancing the identification of unusual activity and reducing cybersecurity risks, this study attempts to investigate the combination of statistical methods and artificial intelligence for network traffic analysis. The study focuses on utilizing machine learning models like Random Forest and Neural Networks in conjunction with statistical variables like entropy and packet size distribution to analyze network traffic. Although the scope emphasizes detection accuracy, encrypted traffic analysis is not included.

This study suggests a hybrid strategy to improve network traffic analysis for cybersecurity by combining conventional statistical techniques with AI-driven solutions. Through the use of statistical methods' interpretability and artificial intelligence's adaptability, the technique provides a scalable, successful, and efficient way to detect and counteract contemporary cyberthreats in practical contexts.

Literature review

Cybersecurity Techniques in Network Traffic:

a) Firewalls: A firewall is a hardware or software program that regulates network traffic by permitting legitimate communication while preventing potentially dangerous information. An untrusted external network, like the internet, can be separated from a trusted internal network by installing a firewall. Analyzing data packets and comparing them to preset security criteria is how firewalls operate. These regulations may be predicated on parameters like specific packet protocols, port numbers, and source and destination IP addresses (see [1-4]).

b) **Intrusion Detection System (IDS):** An intrusion detection system (IDS) is a type of network security technology that monitors devices and network traffic for known hostile conduct, suspicious activities, or violations of security policies.

An intrusion detection system (IDS) can help speed up and automate the process of identifying network risks by alerting security administrators to known or unknown threats or by sending alerts to a centralized security tool. By integrating data from several sources with a unified security solution, such a security information and event management (SIEM) system, security professionals can identify and thwart cyberthreats that might elude other defenses. IDSs can help with compliance initiatives as well. Organizations must have intrusion detection systems in order to comply with requirements like the Payment Card Industry Data Security Standard (PCI-DSS) (see [5-7]).

c) **Traffic Anomaly Detection:** Traffic anomaly detection is the process of identifying deviations from normal network behavior. It can identify possible security threats such as malware infections, cyberattacks, or unauthorized access attempts. This technique looks at network traffic patterns, such as data flow, packet size, and frequency, to establish a baseline of typical activity. When network traffic significantly departs from this baseline, it is flagged as unusual and may be further investigated for possible risks. Anomaly detection techniques might be statistical (based on predefined thresholds or distributions) or AI-driven (using machine learning models to learn and adapt to changing traffic patterns).

d) **Encryption:** One method of data security is encryption, which transforms data into ciphertext. The original plaintext content can only be accessed by authorized people who possess the key to decode the coding. To put it simply, encryption makes information unintelligible to those who are not allowed access to it. This deters fraudsters who may have gained access to a company network using highly advanced methods only to discover that the data is illegible and hence worthless.

Encryption provides authenticity and integrity in addition to guaranteeing the confidentiality of data or messages, demonstrating that the underlying communications or data have not been changed from their initial state (see [8-10]).

e) **Network Segmentation:** By segmenting a network into smaller, discrete sub-networks, network teams can compartmentalize a network and provide each one with its own security controls and services. The technique is known as network segmentation. A physical network is divided into several logical sub-networks as part of the network segmentation process. Controls are applied to the distinct, segregated components of the network after it has been split up into smaller, easier-to-manage sections (see [11-13]).

Statistical Methods in Network Traffic Analysis:

a) **Clustering:** Clustering is an unsupervised machine learning technique where the algorithm group's data points into clusters without any predefined labels based on the number of clusters the developer selects. A well-liked method called K-means randomly initializes cluster centers before repeatedly redistributing data points and recalculating cluster centers until they settle. Clustering is useful in incident response because it allows analysts to uncover relationships in data that were previously unknown. Because the results can change with each run, a number of repetitions with different cluster numbers may uncover anomalous data points that need further investigation (see [14-15]).

Example:

Imagine we have a dataset containing network traffic details (e.g., packet size, frequency of packets, and destination IP). We will use a **K-Means clustering** algorithm to group this data into 3 clusters (representing normal, suspicious, and attack traffic).

Sample Data (Packet Size and Frequency):

Packet Size (KB)	Frequency (Packets per Minute)
50	100
60	120
70	150
200	300
220	310
240	320

K-Means Clustering (Assuming K=3 Clusters):

- **Cluster 1 (Normal Traffic)**: 50, 60, 70 (small packets, normal frequency)
- **Cluster 2 (Suspicious Traffic)**: 200, 220 (larger packets, medium frequency)
- **Cluster 3 (Attack Traffic)**: 240 (large packet size, high frequency)

Because of its great size and frequency, any packet from the attack traffic cluster (Cluster 3) may be identified as possibly malicious based on this clustering.

b) **Stack Counting:** Stack counting is a method for finding abnormalities in a dataset by classifying data points into bins based on their value for a specific property. It's important to keep in mind that benign events are common, but malignant ones are uncommon. For instance, a Sqrrl examination of a collection of online traffic showed that the destination ports for most traffic were 80, 443, or 25. However, only one hit was found in four ports, indicating that further investigation is necessary. For benign samples, it is crucial to choose a feature whose values fall into one or a small number of bins.

Example:

Consider keeping an eye out for login attempts from various IP addresses on a server. To monitor the volume of login attempts from different IPs, we'll employ stack counting.

Sample Data (Login Attempts per IP):

IP Address	Login Attempts (Count)
192.168.0.1	5
192.168.0.2	3
192.168.0.3	2
192.168.0.1	4
192.168.0.2	1
192.168.0.1	7

Stack Counting Result:

- **IP 192.168.0.1**: 5 + 4 + 7 = 16 login attempts.
- **IP 192.168.0.2**: 3 + 1 = 4 login attempts.
- **IP 192.168.0.3**: 2 login attempts.

If a threshold is set (e.g., 10 login attempts in an hour), **IP 192.168.0.1** would be flagged for potentially brute-force activity because it exceeds the threshold. By monitoring the frequency of events (like login attempts), stack counting makes it possible to spot unusual activity (like possible brute-force assaults).

AI in Network traffic Analysis:

Network traffic analysis is the process of examining and monitoring data as it flows through a network. It provides network administrators with crucial data about the behavior of users, devices, and apps, allowing them to identify irregularities that may indicate security breaches or operational issues. Traditional methods of network traffic analysis identify suspicious activity using predefined rules and signatures. However, these techniques sometimes lag behind evolving cyberthreats and may result in false positives or overlook subtle indications of compromise.

AI-powered network traffic analysis is a paradigm shift in cybersecurity that uses machine learning algorithms to identify patterns and abnormalities in vast amounts of network data. AI systems can automatically identify deviations that can point to malicious activity by using past data to learn typical network behavior. AI algorithms are different from rule-based systems in that they can adjust to dynamic threats and shifting network conditions, increasing the efficacy and accuracy of intrusion detection.

Use of AI in Network traffic:

Anomaly detection: Artificial intelligence (AI) solutions for network monitoring can promptly spot odd trends or departures from typical network behavior, which could point to a system malfunction or security compromise.

Traffic Classification: By examining characteristics like protocol and packet size, AI divides network traffic into categories like streaming and browsing. This facilitates efficient network resource management by optimizing bandwidth and detecting fraudulent or illegal activity.

Encryption Traffic Analysis: Without compromising privacy, AI examines encrypted traffic metadata, such as packet size and time, to identify harmful activities. By following encryption guidelines and activating secure traffic monitoring, this method can detect dangers such as encrypted malware.

Automated Threat Response: AI makes it possible for systems to respond to security risks by automatically banning malicious IP addresses, isolating compromised devices, or sending out alerts. This ensures real-time defense by decreasing damage and reaction time during strikes.

Proposed Approach

The Hybrid Model for Network Traffic Analysis improves network traffic identification and classification by combining statistics and artificial intelligence techniques. In order to filter out irregularities, the mean, variance, and interquartile range are calculated as part of the preconceptioning process. AI models like Decision Trees and Neural Networks are then used to process such abnormalities in order to determine whether or not they pose a hazard. The tools are a trustworthy "on demand" resource for handling cybersecurity challenges since they change over time in tandem with statistical boundaries and artificial intelligence education.

Statistical Techniques in the Hybrid Model:

1. **Outlier detection:** Statistical anomaly detection uses measures such as mean, standard deviation, and z-scores to identify anomalous data values. When a data point's Z-score is high, it indicates that it deviates greatly from the mean and may be an outlier.

$$\mathbf{Z} = \frac{X-\mu}{\sigma}$$

Where, X = Observed value

μ = Mean of the dataset

σ = Standard deviation

2. **Stack counting:** It does the binning of data points based on features such as packet size, flow duration, and arranges the bins according to frequency. Bins which are rare have a small count and should be under scrutiny.

AI Techniques in the Hybrid Model:

A number of network traffic analysis hybrid approach heavily relies on artificial intelligence (AI) methods. They complement the system in classifying traffic patterns and detecting outliers through including machine learning (ML) techniques. These solutions augment statistical techniques by utilizing high-dimensional datasets, forming intelligent decisions and responding to changing conditions of the network.

Feature Extraction: A feature vector is created by extracting features such mean packet size, flow time, and port usage.

Feature Vector: [Packet Size, Flow Duration, Port Type]

Classification: The feature vector will be analyzed by AI methods such as logistic regression, decision trees, or neural networks to classify the traffic. For example, multi-class logistic regression determines a flow's probability.

$$P(y = k \mid x) = \frac{e^{(\theta_k)^T x}}{\sum_{j=1}^{k} e^{(\theta_j)^T x}}$$

Where,

P(y = k | x) = Probability of class k

θ_k = Parameters for class k

x = Input feature vector

Adaptation: AI models are routinely retrained using fresh data, improving the system's accuracy and adaptability. Essentially, in order to adjust to the changing network conditions, AI techniques retrain on more current tagged data. During the learning process, a loss function, like cross-entropy for classification, is minimized:

Hybrid Approach Workflow:

The hybrid methodology for network traffic analysis combines the benefits of AI-driven classification with statistical methodologies to deliver dependable, accurate, and adaptable anomaly identification. Each of these steps in the workflow has a specific function, as explained below:

Preprocessing Data: Gather unprocessed traffic information, such as ports, packet size, and duration. Determine the data's statistical characteristics (Z-scores, count in bin).

Finding Anomalies: Use statistical techniques to identify potentially unusual behavior. The AI component receives the outliers and uncommon occurrences.

AI Categorization: Give the AI model data that has been statistically highlighted. Sort traffic into three categories: benign, suspicious, and malicious.

Loop of Feedback: To improve adaptability, the AI is retrained using new data and statistical thresholds are adjusted on a regular basis.

Methodology

By fusing artificial intelligence (AI) and statistical methods, the hybrid model of network traffic analysis effectively examines and categorizes network traffic. Combining statistical preprocessing techniques with machine learning models can provide a thorough strategy to detecting abnormalities and categorizing communications as benign, suspicious, or malicious. This approach expands on the advantages of statistical analysis in data preprocessing, feature extraction, and outlier detection by using AI models for classification.

Data Collection and Preprocessing:

Data collection: Compile raw information about network traffic, including protocols, ports, source and destination IP addresses, packet size, and flow time. Consider the dataset D = {d_1, d_2,...,d_n}, where each d_i denotes a data packet with attributes such as protocol type, size, and duration.

Data cleaning: Fill in the gaps and get rid of duplicates. For instance, if there are missing values in the dataset's packet size, utilize mean imputation to impute the missing values.

$$X=\frac{\sum_{i=1}^{n} X_i}{n}$$ where,

X_i=observed packet size

X=missing value

To compute key statistical features for each data flow:

- Mean of packet size(μ)= $\frac{1}{n}\sum_{i=1}^{n} x_i$
- Variance (σ^2)= $\frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2$
- Interquartile range (IQR)= Q_3-Q_1

  Where, Q_3 and Q_1 are the 75th and 25th percentiles, respectively.

Statistical Analysis for Anomaly Detection:

Outlier Detection: To find possible outliers, the Z-score is utilized to discover data points that substantially differ from the mean. Utilizing each data point X, determine the Z-score:

$$Z = \frac{X-\mu}{\sigma}$$

Example: if the packet size X=5000 bytes, with μ=1500 bytes and σ=1200 bytes,

$$\text{Z-score (Z)} = \frac{5000-1500}{1200} = 2.92$$

Z-score of 2.92 is abnormal since it is higher than the 3 threshold.

Stack Counting for Feature Binning:

First, the dataset is arranged according to packet size or flow duration. The number of data points that fall into each bin could be determined by, for instance, binding the packet size as follows: 0 – 500, 501 – 1000, 1001 – 2000, and so on. Less than 5 out of 1000 packets fit into this bin, which is an unusual condition where a small number of packets surpass 3000 bytes.

AI Model for Classification:

Feature Vector Construction: Construct x = [Mean Packet Size, Variance, Flow Duration, ...] to create a feature vector for each flow. For instance, a packet flow with length = 2.5 seconds, variance = 500, and mean packet size = 1500 bytes may have the following feature vector:

$$x = [1500, 500, 2.5]$$

Model selection: Choose a model for classification. We will use a logistic regression model to classify the traffic to keep things simple. The logistic regression model predicts the following odds for each class.

$$P(y = k \mid x) = \frac{e^{(\theta_k)^T x}}{\sum_{j=1}^{k} e^{(\theta_j)^T x}}$$

Where,

P(y=k | x) = Probability of class k

θ_k = Parameters for class k

x= Input feature vector

Labeled data is used to train the model, with y ∈ {0, 1, 2} denoting the traffic class (Benign, Suspicious, Malicious).

Training the Model: Apply the Maximum Likelihood Estimation (MLE) technique to fit the logistic regression model using a training dataset:

$$\theta = arg \max_{\theta} \mathbf{G}_{i=1}^{n} P(y_i \mid x_i; \theta)$$

Where y_i is the true label for the i^{th} data point and x_i is the feature vector for that point.

Hybrid Integration:

- **Preprocessing for Statistics:** Use statistical techniques (such as stack counting and Z-score) to eliminate insignificant or harmless data flows. The AI model is only sent anomalous flows for classification.
- **AI Classification:** The AI model classifies the filtered data. For example, the model may receive the following output for the feature vector of a flow being x = [5000, 1200, 5.0]

  P(Benign) = 0.05, P(Suspicious) = 0.85, P(Malicious) = 0.10

  Given the probability of suspicion, the traffic is alerted for further investigation.

- **Retraining and Feedback:** The AI model is updated with new labeled traffic data that is added to the training set, retraining it on a new batch of misclassifications.

Evaluation and Validation:

The model's performance is assessed using measures like F1-scor, recall, accuracy, and precision.

Accuracy:

$$\text{Accuracy} = \frac{\text{True positives}+\text{True Negatives}}{\textit{Total Sample}}$$

Precision:

$$\text{Precision} = \frac{\text{True positives}}{\text{True Positives} + \text{False Positives}}$$

Recall:

$$\text{Recall} = \frac{\textit{True Positive}}{\textit{True Positive}+\textit{False Negative}}$$

Deployment:

- **Real-Time Monitoring**

  Use an operational network infrastructure that regularly examines traffic flows to install the hybrid model. Following real-time statistical preprocessing of the streaming data, the AI model classifies the data.

- **Autonomous Procedures**

  This lessens the hazard because malicious activity will prompt the system to automatically ban IP addresses or ports.

- **Constant Education**

  New labeled data is used to retrain the model on a regular basis to ensure that it adapts to evolving network threats and behaviors.

Experimental Result

The following outcomes are obtained from a simulated dataset of 1000 network traffic flows using the Hybrid Model. Accuracy, precision, recall, and F1-score—all of which are obtained from the confusion matrix are used to assess this model's performance. The ability of a model to identify and categorize anomalies in network data may be assessed using these criteria.

Dataset Overview

Total Traffic Flows (N) = 1000
Benign= 700 flows
Suspicious= 200 flows
Malicious= 100 flows
Example;

Flow id	Packet Size(bytes)	Flow Duration(s)	Packet Count	Category
1	1500	2.5	10	Benign
2	300	0.8	1	Malicious
3	2000	1.0	20	Suspicious

Statistical Preprocessing

Z-Score Calculation:

Mean (μ) = 1200 bytes
Standard Deviation (σ) = 400 bytes

For Flow ID 3 (X_3 = 2000):

$$Z_3 = \frac{2000-1200}{400} = 2.0$$

Anomaly Detection:

If | Z_i | > 3, the flow is flagged as anomalous.

Flow ID	Packet Size (bytes)	Z-Score	Anomalous
1	1500	0.75	No
2	300	-2.25	No
3	2000	2.0	No

Stack Counting:

We group flows into **bins** based on **Flow Duration** and count occurrences

Bin (Duration Range)	Count
0.0 - 1.0	500
1.0 - 2.0	300
2.0 - 3.0	200

Flows in bins with fewer than 10 samples are flagged as anomalies.

AI Classification:

We use labeled data to train a logistic regression model. The equation defines the decision boundary.

$$P(y = 1 \mid X) = \frac{1}{1+ e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3)}}$$

Example:

Here, $\beta_0 = -1.5$, $\beta_1 = 0.01$, $\beta_2 = 0.5$, $\beta_3 = 0.1$ and $X_1 = 1500$, $X_2 = 2.5$, $X_3 = 10$

$$P(y = 1|X) = \frac{1}{1+ e^{-(-1.5+0.01*1500+0.5*2.5+0.1*10)}}$$

$$P(y = 1|X) = \frac{1}{1+ e^{-(-1.5+15+1.25+1)}}$$

$$= \frac{1}{1+ e^{-(15.75)}}$$

$$P(y = 1|X) \approx 0.9999$$

The model predicts this flow as Malicious

Evaluation Metrics:

From predictions on the dataset

Category	TP	TN	FP	FN
Benign	680	315	30	5
Suspicious	170	855	30	10
Malicious	85	925	15	5

F1-Score for Malicious Traffic

$$\text{Precision} = \frac{\text{True positives}}{\textbf{True Positives + False Positives}}$$

$$= \frac{85}{\mathbf{85 + 15}} = 0.85$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive+False\ Negative}$$

$$= \frac{85}{85+5} = 0.94$$

$$\text{F1} = 2.\frac{Precision\ .\ Recall}{Precision\ +Recall}$$

$$= 2.\frac{0.85 * 0.94}{0{,}85\ +0.94} = 0.89$$

Conclusion

A significant step forward in tackling the difficulties of cybersecurity is the hybrid model for network traffic analysis, which blends statistics and artificial intelligence techniques. We may assist in separating different types of traffic and lowering the complexity of the required data, which in turn lowers the effort required of the AI classifiers, by using statistical pre-processing techniques like Z-score for identifying traffic outliers and counting for stacks. With excellent accuracy and F1-scores for malicious traffic identification, the machine learning component in this example, derived from logistic regression is legitimate. These techniques bridge the gap between traditional data analysis and deep machine learning, offering a scalable solution for real-time traffic inspection in intricate network environments. The best features of both methodologies are combined in this hybrid methodology, which makes it possible for large-scale application by improving anomaly detection capabilities and ensuring resource utilization. The method has limitations, such as sensitive statistical parameters and dataset imbalance, despite its possible advantages. Extending enhancements to deep learning models and processing adaptive statistics thresholds for related challenges might yield better performance and flexibility.

The significance of hybrid techniques in modern cybersecurity was underlined, and further study provided a practical framework that is effective and scalable for a range of real-world applications. Results: It was verified that a proposed combination of statistical rigor and AI innovation was effective in shielding networks from ever-evolving cyberthreats.

References

[1] Powell, M. S., & Drozdenko, B. M. (2024). SSOLV: Real-Time AI/ML-Based Cybersecurity via Statistical Analysis. *IEEE Access, 12*, 114786–114794. https://consensus.app/papers/ssolv-realtime-aimlbased-cybersecurity-via-statistical-powell-drozdenko/38efcd0b1df458fdb42cd088a163277f

[2] Alkhudaydi, O. A., Krichen, M., & Alghamdi, A. D. (2023). A Deep Learning Methodology for Predicting Cybersecurity Attacks on the Internet of Things. *Information*. https://consensus.app/papers/a-deep-learning-methodology-for-predicting-cybersecurity-alkhudaydikrichen/37a0ec56cf2d59ff8382cef7544fa2c3

[3] Changala, R., Kayalvili, S., Farooq, M., Rao, M., Vuda, T., & Rao, S. (2024). Using Generative Adversarial Networks for Anomaly Detection in Network Traffic: Advancements in AI Cybersecurity. *2024 International Conference on Data Science and Network Security (ICDSNS)*, 1–6. https://consensus.app/papers/using-generative-adversarial-networks-for-anomaly-changala-skayalvili/88d67155ffad575db06af810b964363b

[4] Agrafiotis, G., Kalafatidis, S., Giapantzis, K., Lalas, A., & Votis, K. (2024). Advancing Cybersecurity with AI: A Multimodal Fusion Approach for Intrusion Detection Systems. *2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, 51–56. https://consensus.app/papers/advancing-cybersecurity-with-ai-a-multimodal-fusion-agrafiotis-kalafatidis/dca38823215e5ec1bee7783b11cb2995

[5] Peter, I., Ijiga, M., Idoko, P., Isenyo, G., Olajide, F. I., Olatunde, T. I., & Ukaegbu, C. (2024). Harnessing Adversarial Machine Learning for Advanced Threat Detection: AI-Driven Strategies in Cybersecurity Risk Assessment and Fraud Prevention. *Open Access Research Journal of Science and Technology*. https://consensus.app/papers/harnessing-adversarial-machine-learning-for-advanced-peter-ijiga/734e9a596a9a5631a69dbe5481b6cdba

[6] Huang, J., Chen, J., Lu, X., Mo, B., Zeng, C., & Qiu, S. (2023). Research on Detection Techniques for Scanning Attacks in Software-Defined Network Environments. *2023 4th International Conference on Computer*

Engineering and Application (ICCEA), 115–118. https://consensus.app/papers/research-on-detection-techniques-for-scanning-attacks-in-huang-chen/a68a19860b6b5e0797a6b5ef7ac27650

[7] Camacho, J., García-Teodoro, P., & Maciá-Fernández, G. (2017). Traffic Monitoring and Diagnosis with Multivariate Statistical Network Monitoring: A Case Study. *2017 IEEE Security and Privacy Workshops (SPW)*, 241–246. https://consensus.app/papers/traffic-monitoring-and-diagnosis-with-multivariate-camacho-garcía-teodoro/ac3b4dbbb1045118a86b365eac972c08

[8] Nayana, S. P., & Aradhya, M. (2020). Review on Machine Learning Algorithm for Cybersecurity. *ArXiv*. https://consensus.app/papers/review-on-machine-learning-algorithm-for-cyber-security-nayana-aradhya/191f786e220e5b4f8d8d96da35b31ef0

[9] Hassan, S. E. H., & Duong-Trung, N. (2024). Machine Learning in Cybersecurity: Advanced Detection and Classification Techniques for Network Traffic Environments. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 11*. https://consensus.app/papers/machine-learning-in-cybersecurity-advanced-detection-and-hassan-duong trung/afb6a986e1885f86b9640aedf853f484

[10] Rustam, F., Ranaweera, P. S., & Jurcut, A. (2024). AI on the Defensive and Offensive: Securing Multi-Environment Networks from AI Agents. *ICC 2024 - IEEE International Conference on Communications*, 4287–4292. https://consensus.app/papers/ai-on-the-defensive-and-offensive-securing-rustam-ranaweera/9da99237f565503581e2928eb56776e3

[11] Rodríguez, M., Iglesias, Á. A., Mehavilla, L., & García, J. (2022). Evaluation of Machine Learning Techniques for Traffic Flow-Based Intrusion Detection. *Sensors (Basel)*. https://consensus.app/papers/evaluation-of-machine-learning-techniques-for-traffic-rodríguez-iglesias/6e9fa74eb67455cb8b03938fa678fff9

[12] Saadi, C., Belghiti, I. D., Atbib, S., & Radah, T. (2024). Contribution to Threat Management Through the Use of AI-Based IDS. *Revista de Gestão Social e Ambiental*. https://consensus.app/papers/contribution-to-threat-management-through-the-use-of-saadi-belghiti/45d01bf0710e5f2592a3e6b8cd0f9240

[13] Ozkan-Okay, M., Akin, E., Aslan, Ö., Kosunalp, S., Iliev, T., Stoyanov, I., & Beloev, I. H. (2024). A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions. *IEEE Access, 12*, 12229–12256. https://consensus.app/papers/a-comprehensive-survey-evaluating-the-efficiency-of-ozkan-okay-akin/5b65d6daa22f5eb090d155242b5baaeb

[14] Ammara, D. A., Ding, J., & Tutschku, K. (2024). Synthetic Data Generation in Cybersecurity: A Comparative Analysis. *ArXiv*. https://consensus.app/papers/synthetic-data-generation-in-cybersecurity-a-comparative-ammara-ding/69fe4fa1715959578 46c8df1c08d5e11

[15] Diallo, A., & Patras, P. (2021). Adaptive Clustering-Based Malicious Traffic Classification at the Network Edge. *IEEE INFOCOM 2021*, 1–10. https://consensus.app/papers/adaptive-clusteringbased-malicious-traffic-diallo-patras/732f31499a5052e6b79c75aa06b8a7e8