Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 5s (2025)

AI/ML Based Detection of Unknown Compromises on ICT Devices

Manikrao Dhore[1], Deepak Mane[2], Shivam Pandagale[3], Om Soni[4], Digvijay Patil[5], Niraj Patil[6]

Vishwakarma Institute of Technology, Bidwewadi, Pune-411037, Maharshtra, India.

[1]manikrao.dhore@vit.edu, [2]deepak.mane@vit.edu, [3]shivam50x20@gmail.com, [4]omsoni2807@gmail.com, [5]nirajpatil1306@gmail.com, [6]digvijaypatil1511@gmail.com

Correspondence Author: deepak.mane@vit.edu

Article History:

Received: *07-10-2024*

Revised: *27-11-2024*

Accepted: *06-12-2024*

Abstract:

The growing dependence on online communication and increased cyberattacks need more reliable techniques for detecting network anomalies. While successful, traditional signature-based detection systems have trouble detecting zero-day attacks and cannot analyse encrypted traffic, which today makes up over half of all internet traffic. The goal of the machine learning-based method for network anomaly detection in this research is to find malicious activities in encrypted and unencrypted network traffic. This research uses machine learning methods to detect abnormal behaviours that suggest possible security breaches. Several classification methods were investigated and tested using datasets that were made available to the public. Even in encrypted network environments, the results demonstrate that machine learning techniques, especially anomaly-based detection, can improve detection accuracy and lower false positives.

The proposed model's novelty lies in its integration of tailored feature selection with attack-specific and unified datasets, ensuring a comprehensive approach to identifying network anomalies.

Keywords: • Network Anomaly Detection, Machine Learning, Cybersecurity, Intrusion Detection System, Zero-Day Attacks, Feature Selection, Benign Traffic, Naïve Bayes, Random Forest regressor, Dimensionality Reduction

Introduction:

Every day, hundreds of thousands of organisations and millions of individuals communicate using the Internet. With over 4 billion users today, the number of Internet users has grown dramatically over the past 20 years and continues to expand at a rapid rate. Online attacks are increasing constantly in step with these developments. Signature-based identification is the primary detection method used to prevent these attacks and ensure data security. Signature-based strategies leverage a database of known attack patterns to identify threats. Despite their effectiveness, these databases are vulnerable to zero-day (unknown) assaults and still require frequent updates to include new attack data. Signature-based methods cannot prevent zero-day attacks since they are not yet

Additionally, nearly half of all web traffic is encrypted today via SSL/TLS (Secure Sockets Layer /Transport Layer Security) protocols, and this percentage is constantly increasing. Since signature-based methods cannot view the contents of encrypted internet communication, they are useless. By basing its analysis on general attributes like size, connection duration, and packet count, the anomaly-based approach, on the other hand, may assess encrypted protocols without requiring access to message content. Because of these advantages, the anomaly-based detection technique is increasingly being

used to detect and thwart network attacks. This project aims to further the field by developing a system that quickly and correctly detects network anomalies using machine learning techniques.

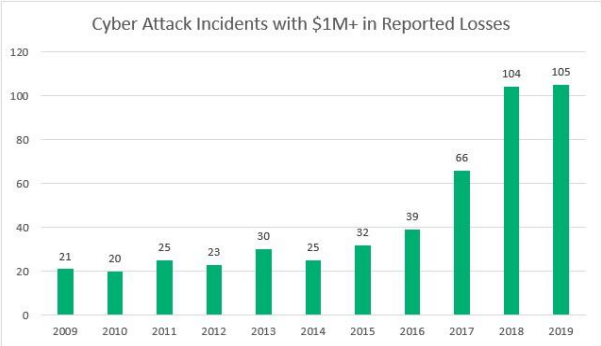

Fig 1: Cyber Attack numbers

By determining the most pertinent characteristics for different attack types, the suggested approach seeks to improve machine learning-based network anomaly detection. By ensuring accurate and effective classification, this improves network system security.

The significance of network security and the necessity of efficient threat detection are covered in the introduction to provide a summary of the inquiry. The literature study that follows examines past research on network intrusion detection systems, feature selection, and machine learning models. Methodology describes the steps involved in gathering data, choosing features, putting machine learning models into practice, and evaluating the results. The paper's findings are examined. The model performance findings are presented in the paper using a variety of metrics, such as recall, accuracy, and precision. In 2019, some researchers used machine learning approaches for the pattern classification [7][8].

Objectives:

To develop a machine learning-based network anomaly detection system capable of identifying malicious activities, including zero-day attacks, in both encrypted and unencrypted network traffic. The system focuses on:

- Enhancing detection accuracy.
- Reducing false positives.
- Using feature selection to optimize the classification of network anomalies.

Methods:

This paper describes a machine learning-based approach to network anomaly detection. It utilises the CICIDS 2017 dataset, which contains malicious and benign traffic as well as a range of attack strategies. We have used different papers and came to the conclusion that our model KNN performs well and it has F- measure of 86 that is highest in comparison to other models that we have compared. The research focuses on feature selection and the application of machine learning approaches to identify significant patterns that classify specific assault behaviours. The goal is to develop a reliable

machine learning model that has a high degree of accuracy and few false alarms while detecting network anomalies. KNN performs well and it is best to counter the attacks.

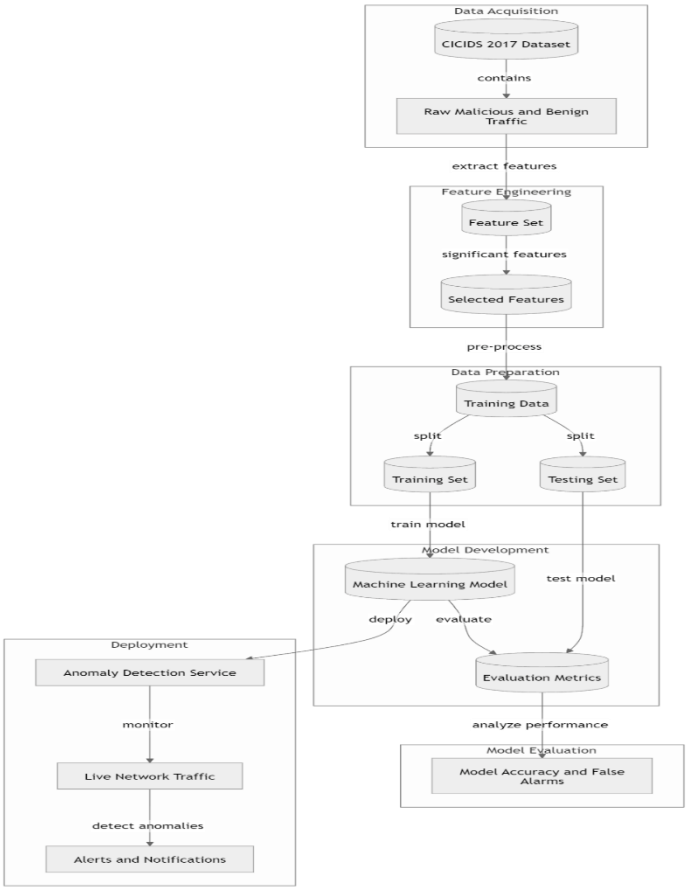

Fig 2: System Architecture

The design of the suggested system is organised into multiple stages: gathering features, data preparation, model selection and training, and evaluation matrix, deployment of the detection, analysis as shown is fig 2. Developing a trustworthy machine learning model that can detect network irregularities with a high degree of accuracy and few false alarms is the aim. Then the model analyses performance of the dataset and makes sure it is robust. Another thing that this system architecture has is the detection of anomaly and alerting system and notifications sent to the user when the attack is happening. We have the K-Nearest Neighbour algorithm for detecting the attacks. With an F- measure of 86 that is the highest. We have included 16 attacks in this project.

A. Dataset Description

This Dataset is addressed in the CICIDS 2017 (Intrusion Detection Evaluation Dataset) that was created by the Canadian Institute for Cybersecurity at the University of New Brunswick. This dataset consists of a 5-day (3rd July- 7th July 2017) data stream on a network created by computers using up-to-date operating systems. They have used all major attacks that happen in the industry. Below figure 3 shows dataset description.

Flow Recording Day (Working Hours)	pcap File size	Duration	CSV File Size	Attack Name	Flow Count
Monday	10 GB	All Day	257 MB	No Attack	529918
Tuesday	10 GB	All Day	166 MB	FTP-Patator, SSH-Patator	445909
Wednesday	12 GB	All Day	272 MB	DoS Hulk, DoS GoldenEye, DoS slowloris, DoS Slowhttptest, Heartbleed	692703
Thursday	7.7GB	Morning	87.7 MB	Web Attacks (Brute Force, XSS, Sql Injection)	170366
		Afternoon	103 MB	Infiltration	288602
Friday	8.2GB	Morning	71.8 MB	Bot	192033
		Afternoon	92.7 MB	DDoS	225745
		Afternoon	97.1 MB	PortScan	286467

Fig 3: Collected Dataset for major Attacks

This CICIDS data set has the following advantages given below:

- The dataset contains real-world data from a testbed of actual computers. Data streams were gathered from machines running the latest operating systems, with diversity across Mac, Windows, and Linux platforms for both attackers and victims OS (operating system) that are used.
- This dataset is labelled, enabling its use in machine learning applications. A critical step of feature extraction was performed, and the end result 85 features.
- The raw data (p-cap files containing captured network packets that are usually sent during communication) and processed data (CSV files with comma-separated values) are provided for analysis and detection.
- This selection of attacks included in the dataset was informed and verified by the 2016 McAfee Security Report and Norton security report, ensuring a comprehensive and current variety of attack types. The dataset is richer than others in terms of the protocols it encompasses, including HTTPS, TP, SSH.

B. Dataset Preprocessing

This section describes the processes performed to resolve concerns in the dataset, including as error correction, missing value management, and data type transformations for optimal model performance and dataset verification.

The CICIDS2017 dataset initially contains almost 3,119,345 stream records that were used, which are distributed as indicated in Fig below. Some records were found to be incomplete or unlabelled, particularly 288,602 records saved as empty placeholders or different values. These records were eliminated to improve the dataset's quality and efficiency. Following Fig 4:

Label Name	Number
Benign	2359289
Faulty	288602
DoS Hulk	231073
PortScan	158930
DDoS	41835
DoS GoldenEye	10293
FTP-Patator	7938
SSH-Patator	5897
DoS slowloris	5796
DoS Slowhttptest	5499
Bot	1966
Web Attack – Brute Force	1507
Web Attack – XSS	652
Infiltration	36
Web Attack – SQL Injection	21
Heartbleed	11

Fig 4: Data Points for various attacks

C. Feature Selection:

For feature selection, two approaches were used according to attack specific and a combined dataset.

Approach 1: Feature Selection according to Attack Types

Approach, each attack of type data was isolated into a separate file and for every attack type, the file created had 30% attack data points and 70% benign data points. Each file used to attack specific relevant features based on the feature selection phase. And this approach was employed to understand relation between attack specific features and how different algorithms respond to them helping in understanding at a more granular level.

Approach 2: Feature Selection for Combined Attack and Benign Labels

In this approach, all the dataset was combined into one dataset. The data points were divided into "attack" and "benign" classes creating a binary classification. Here, for each attack type top 4 features were selected. Based on 12 attack types, 48 features were selected from which after removing the duplicates resulted in 18 unique features. Following Fig 5 shows features selected.

Bwd Packet Length	Std Bwd Packet Length	Flow Duration
Max Flow IAT	Std Flow IAT	Fwd Packet Length Max
Mean Fwd Packet Length	Std Total Backward Packets	Total Length of Bwd Packets
Min Bwd Packet Length	Flow Bytes/s	Flow IAT Max
Mean Flow IAT	Fwd IAT Total	Fwd Packet Length Mean
Min Fwd Packet Length	Total Fwd Packets	Total Length of Fwd Packets

Fig 5: Combined 18 features selected

Another approach was employed based on the approach above. features that contributed most were selected from the features selected above. Seven features were selected that contributed to the 97% of the dataset. The threshold was set at 0.8% for selecting features.

This strategy was employed to reduce the selected features and use only the most relevant features to reduce the computational efficiency. The following feature sets were applied across models to assess which yielded optimal performance

- First Feature Set: Manually defined, consisting of 18 relevant features.
- Second Feature Set: The top 4 features for each attack type, resulting in a total of 18 unique features.
- Third Feature Set: The 7 most important features selected according to the 0.8% threshold weight, covering 97% of total feature importance.

These selected feature sets enable a focused and efficient evaluation of model accuracy in detecting network anomalies across individual attack types and generalized attack pattern

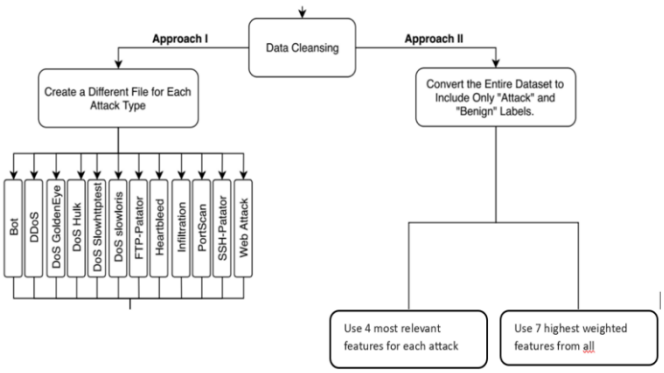

Fig 6: Feature selection model

D. Model Selection and training:

Seven machine learning algorithms—Naïve Bayes, Random Forest, K-Nearest Neighbours (KNN), ID3 (Iterative Dichotomiser 3), AdaBoost, Multi-Layer Perceptron (MLP), and Quadratic Discriminant Analysis (QDA)—were chosen for training and testing. These models were selected for their proven capability in classification tasks and their diverse methodological foundations. These models were selected due to the stability to handle prediction tasks which was necessary for the approach.

There are Two distinct approaches that were utilized to apply machine learning algorithms to the dataset.

In the first approach, files generated during the Feature Selection phase, along with the selected attributes, were used. These files consist of 30% attack data and 70% benign data, each named according to the specific attack type it contains. Seven different machine learning methods were applied to each file 10 times, producing separate results for each attack type. This approach aimed to evaluate the effectiveness and performance of various machine learning methods across different attack types.

In the second approach, the entire dataset is treated as a single file, where all attack instances are grouped under a single common label, "attack." As a result, the data in this file is categorized into two labels: "attack" and "benign." The feature set is constructed by combining the top four most significant features (based on importance weights) identified for each attack type in the first approach. This

process yields a pool of 48 features derived from the 12 attack types. After eliminating duplicate features, the final feature set is reduced to 18 unique attributes mentioned above in fig 5.

The alternative implementation of this study involves selecting features based on their high-performance scores, as determined in the "Feature Selection of Attack" section, rather than using all 18 features mentioned earlier. A threshold value of 0.8% is applied to the feature importance scores. By using this threshold, 97% of the total feature importance weight is captured with just 7 features, while the remaining 13 features account for only 3% of the total weight. When features with an importance weight of 0.8% or higher are selected, the following fig 7 features are used and their %:

Feature Name	Importance Weight	Percentage
Bwd Packet Length Std	0.246627	38.97%
Flow Bytes/s	0.178777	28.25%
Total Length of Fwd Packets	0.102417	16.18%
Fwd Packet Length Std	0.063889	10.10%
Flow IAT Std	0.009898	1.56%
Flow IAT Min	0.006946	1.10%
Fwd IAT Total	0.005121	0.8 %

Fig 7: Weights for the most relevant features

All the seven algorithms were applied to the dataset created in both the approaches again. This was done to give us more generalized knowledge regarding the relation and response of algorithms to the generalized dataset giving us a broader perspective of the working of algorithms.

E. Evaluation:

The evaluation assessed the performance of seven machine learning models (Naïve Bayes, Random Forest, KNN, ID3, AdaBoost, MLP, QDA) using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. A 10-fold cross-validation ensured generalization.

- Approach 1: Models trained on attack-specific datasets were evaluated for each type of attack.
- Approach 2: Models trained on a unified dataset were tested for overall performance across all attacks.

AdaBoost and Random Forest excelled, while simpler models like Naïve Bayes and QDA struggled with precision due to data assumptions

Experiment Results:

We have taken the CCIDS 2017 Dataset; our dataset doesn't have any dedicated training and test data but we have a the whole unbundled data, we have taken the configuration of 80:20 ,that is 80% training data and rest 20% testing data, we have make the selection random while created the sets which called as cross-validation , we have used the hardware configuration as our laptop which is ASUS TUF 15 , intel i7 12th gen , and a Nvidia graphics card of 8GB RTX-3050. which can handle large data for the training and testing because it requires high computational power. We have seven machine models on the 12 attacks and the primary metric is F-Measure our evaluation criteria. We have given all the F-measures of our algorithms in the following Fig 8. The bold values imply that for each attack the height OF-measure is achieved and the one which underlined ones are the smallest f-measures. Following Fig 8 results of models

Attack Names	NB	RF	KNN	ID3	AB	MLP	QDA
Bot	*0.54*	0.96	0.95	0.96	**0.97**	0.64	0.68
DDoS	0.77	0.96	0.92	**0.96**	0.96	0.76	*0.34*
DoS GoldenEye	0.81	0.99	0.98	**0.99**	0.99	*0.64*	0.71
DoS Hulk	*0.23*	0.93	0.96	**0.96**	0.96	0.95	0.36
DoS Slowhttptest	*0.35*	0.98	0.99	0.98	**0.99**	0.78	0.38
DoS slowloris	*0.37*	0.95	0.95	**0.96**	0.95	0.74	0.46
FTP-Patator	**1.00**	1.00	1.00	1.00	1.00	1.00	1.00
Heartbleed	**1.00**	0.99	1.00	0.95	0.93	*0.66*	1.00
Infiltration	0.78	0.92	0.88	0.89	**0.92**	*0.52*	0.83
PortScan	*0.39*	1.00	1.00	**1.00**	1.00	0.61	0.85
SSH-Patator	*0.33*	0.96	0.95	**0.96**	0.96	0.83	0.41
Web Attack	0.74	0.97	0.93	**0.97**	0.97	*0.60*	0.84

Fig 8. Results of models on different attacks

In following figure, we have provided the comparison of different matrices on the different algorithms, in all of the algorithms we have found that the K Nearest Neighbour algorithm have the highest F-measure with the 0.97 which is written in Bold letter which conveys that there is a complete balance between recall and precision. which also conveys that it is able to minimize the false positives and able to detect true positives, so it is the best model. The ones which are written in red are one of the lowest F-measures scores which conveys that they are not able to balance between recall and precision. Among them MLP has the lowest scores and Random Forest, ID3, AdaBoost are the one of the strongest with F-measure of 0.94-0.95.but the K Nearest Neighbours is still outperforming them and all the data we have provided in the following Fig 9.

Machine Learning Algorithms	Evaluation Criteria				
	F-Measure	Precision	Recall	Accuracy	Time
Naive Bayes	0.86	0.86	0.87	0.87	**1.8255**
QDA	0.86	0.87	0.88	0.88	2.3696
Random Forest	0.94	0.94	0.94	0.94	19.0899
ID3	0.95	0.95	0.95	0.95	9.5107
AdaBoost	0.94	0.94	0.94	0.94	135.2455
MLP	*0.83*	0.82	0.87	0.87	59.6933
K Nearest Neighbours	**0.97**	0.97	0.97	0.97	*1626.833*

Fig 9: Classification Report for Algorithms

Conclusion:

The paper successfully identified a set of optimal features and demonstrated the potential of using advanced machine learning techniques for detecting compromised devices on IoCs. The MLP algorithm, in particular, showed the highest performance, indicating the value of neural networks in this context. This study shows that detecting the attacks is very efficient for this model. The ICT devices being compromised before now are not because of the model for detecting the attacks. Finally, the 7 algorithms used in this project, the most efficient and robust was KNN algorithm with f-measure of 0.86.

References:

[1] P. Kostus, "Anomaly Detection in Networks Using Machine Learning," Research Proposal, 23 Mar 2018

[2] “Internet Growth Statistics,” Miniwatts Marketing Group, 2 Mar 2018.[Online].Available:https://www.internetworldstats.com/emarketin g.htm. [Accessed 26 Aug 2018].

[3] C. Leckie and K. Leung , "Unsupervised anomaly detection in network intrusion detection using clusters," in Proceedings of the Twenty-eighth Australasian conference on Computer Science-Volume 38, 2005, pp. 333-342:

Australian Computer Society, Inc.

[4] A. Gharib, A. H. Lashkari, and A. A. Ghorbani, "Towards a reliable intrusion detection benchmark dataset," Software Networking, vol. 1, no. 1, pp. 177200, 2017.

[5] "1998 DARPA Intrusion Detection Evaluation Data Set," Lincoln Laboratory, Massachusetts Institute of Technology, [Online]. Available: https://www.ll.mit.edu/rd/datasets/1998-darpa-intrusion-detection-evalu ation-data-set. [6]C. Thomas, V. Sharma, and N. Balakrishnan, "Usefulness of DARPA dataset for intrusion detection system evaluation," in Data Mining, Intrusion Detection, Information Assurance, and Data Networks Security 2008, 2008, vol. 6973, p. 69730G: International Society for Optics and Photonics.

[6] A. Gharib, I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "An evaluation framework for intrusion detection dataset," in Information Science and Security (ICISS), 2016 International Conference on, 2016, pp. 1-6: IEEE.

[7] G.D. Upadhye, . U. V. Kulkarni, D. T. Mane (2021). Improved Model Configuration Strategies for Kannada Handwritten Numeral Recognition. Image Analysis & Stereology, Vol. 40, Issue 3, pp. 181-191. DOI: 10.5566/ias.2586 (SCI & SCOPUS)

[8] Chaitanya Bhagat and D. T. Mane “Survey on Text Categorization Using Sentiment Analysis,” International Journal of Scientific & Technology Research, Vol 8, Issue-8, pp 1189-1195, 2019

[9] "KDD Cup 1999 Data," University of California, Irvine, [Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.