A Novel Machine Learning Model for Early Detection of Advanced Persistent Threats Utilizing Semi-Synthetic Network Traffic Data

Nadim Ibrahim[1]*, N.R. Rajalakshmi[1], Karam Hammadeh[1]
[1]*Department of Computer science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India*

KEYWORDS:
APTGuard,
Advanced Persistent Threats,
APTs,
Machine learning,
LSTM,
SVM,
KNN,
LR

ARTICLE HISTORY:
Received 19.04.2024
Accepted 16.05.2024
Published 13.06.2024

DOI:
https://doi.org/10.31838/jvcs/06.02.04

Abstract

Advanced Persistent Threats are not merely a buzzword, these highly sophisticated and stealthy cyber threats are characterized by their ability to infiltrate and persistently operate within target systems for extended periods, often remaining undetected until significant damage has been done. APTs have emerged as a formidable adversary, and it frequently attack important infrastructure, government entities, and private businesses. This research paper embarks on an examination of APTs, shedding light on their characteristics and strategies. The proposed model APTGuard in the paper presents vital way to detect and counter this menace effectively, employing a methodology that involves utilizing a Semi-Synthetic dataset using 6.8 Million samples of processed network flows for training and testing. Orchestrating four pivotal phases: data collection, feature selection, data pre-processing and applying of machine learning algorithms. Encompassing the application of the algorithms: long short-term memory (LSTM), logistic regression (LR), support vector machine (SVM), and k-nearest neighbours (KNN), with comparing the results, the paper emphasizes that APT Guard, achieves a notable accuracy of 99.89 % using 83 features. The paper makes a substantial contribution to create effective method for detecting and resist hidden and malicious APTs.

Author e-mail: nadimibrahimcs@gmail.com

How to cite this article: Ibrahim N, Rajalakshmi NR, Hammadeh K. A Novel Machine Learning Model for Early Detection of Advanced Persistent Threats Utilizing Semi-Synthetic Network Traffic Data, Journal of VLSI Circuits and System Vol. 6, No. 2, 2024 (pp. 31-39).

1. INTRODUCTION

Cyber-attacks have evolved into sophisticated and pervasive threats, capable of inflicting significant damage on individuals, businesses, and nations alike. These attacks encompass a wide range of tactics, techniques, and procedures, each designed to jeopardize the integrity, availability and confidentiality of digital assets.[1]

(APTs) are among the most dangerous types of cyber-attacks, as they constitute a highly focused and covert approach to infiltration and espionage. Unlike opportunistic assaults, which frequently employ indiscriminate tactics and target known weaknesses, APTs are painstakingly planned, executed, and continually maintained to achieve their long-term objectives.

APTs utilize sophisticated methods to breach networks, searching sensitive data, and maintain secret access for a long time. Standard security solutions often fail to keep up with the constant characteristics of APTs. requiring creative methods to improve defense strategies.[2]

APTs are carefully prepared, often supervised by well-resourced suppliers with specific objectives such as spying, destruction, or financial gain. As such, traditional signature-based detection techniques are often useless against APTs, as they are able to bypass static patterns and signatures. Machine learning, a kind of artificial intelligence, has become recognized as a powerful instrument in the arsenal of cybersecurity experts for detecting and mitigating such difficult threats.

These techniques can analyze various datasets including network traffic, system logs, user behavior, and endpoint activities to detect variations from normal patterns and mark potential security incidents in real-time. By applying supervised, unsupervised, and reinforcement learning techniques, machine learning models can constantly improve their understanding of new threats, enhancing their efficacy in APT tracking.[3]

In addition, machine learning allows organizations to advance above a reactive approach to cybersecurity by enabling proactive risk hunting and predictive analysis.[4] By evaluating historical records and identifying hidden indicators, machine learning algorithms are able to predict potential APT incursions, allowing security teams to earlier establish their defenses and reduce risks before they grow into full-blown violations.

The paper explores the influence of machine learning algorithms. in boosting APT detection abilities within cybersecurity frameworks. Along with explore the key challenges caused by APTs, through scenarios, empirical analyses, and real-life examples, we point out the practical applications of machine learning in decreasing APT risks and protecting important resources in today's digital environments.

By employing the predictive and analytical strengths of machine learning, organizations can obtain more insight into their cybersecurity standing, identify emerging threats proactively, and build a more effective protection.

The researchers in agreed that APTs are defined by several essential points:

- Persistence: APTs use complex strategies to remain undetected in targeted networks for months or even years.
- APT attackers use advanced techniques: including zero-day exploits, social engineering, and unique malware to avoid detection and overcome traditional security mechanisms.
- APTs targets: like businesses, industries, or geopolitical entities for strategic aims including intellectual property stealing, espionage, or disruption.

APT attacks can have severe consequences, including economic losses, ruined reputations, national security hazards, and geopolitical crisis. As a result, detecting and mitigating APTs pose a significant problem for cybersecurity professionals and companies worldwide.

By utilizing the strength of artificial intelligence, ML-based solutions offer the potential to identify subtle patterns, anomalies, and indicators of compromise associated with APT activities.

After reviewing existing literature, analyzing real-world case studies, and discussing practical considerations this research paper is guided by three main objectives. Initially, this paper aims to increase awareness over organizations and individuals about the significant threats and significant effects related to APTs. Second, it seeks to create a robust early detection system for APTs. Finally, the most important objective of this study is to build and perform an effective model to prevent the evolving threat posed by APTs. By achieving these goals, this work makes significant contributions to overall realm of cybersecurity by enhancing the defenses against the aggressive cyber threat.

2. Related Work

Since Advanced Persistent Threats (APTs) are so hidden, early detection is critical. This section discusses the previous studies about APTs detection.

Detecting beaconing is crucial in guarding against APTs.[5] particularly in the industrial security field. APTs are cyber intru sions carried out by skilled and well-resourced adversaries who target specific information in high-profile organizations and governments, frequently as part of a multi-phase long-term operation. One of the phases of the APT process is the command-and-control (C&C explores AI algorithms for APT detection, analyzing datasets and discussing strengths and challenges of detection methods. It also outlines cybersecurity vendor projects categorized by their detection approach for APT or beaconing operations.

The research[6] Introduces a graph convolutional neural network (GCN)-based model for detecting APTs. Traditional methods struggle with long-term relationships in these attacks. The proposed model utilizes an understanding of APT activities, extracted from threat intelligence and software security entities, converted into a homogeneous graph. Using GCNs, the method achieves a 95.9% detection accuracy.

The research[7] countries face a multitude of electronic threats that have permeated almost all business sectors, be it private corporations or public institutions. Among these threats, advanced persistent threats (APTs addresses the pervasive threat of APTs in electronic era, proposing a multi-stage framework for automated APT detection using time series data. Aiming to enhance real-time detection, the approach surpasses previous models, leveraging standardized techniques and diverse datasets. The study introduces a composition-based decision tree (CDT) system.

The research of[8] examines behavior-based detection techniques, including heuristics that assess processes, network communications, and file interactions. This research offers insights into the adaptability of behavior-based detection in response to the evolving nature of APTs. Unlike traditional methods, Behavior-based detection is not dependent on existing patterns; instead, it profiles activities using statistical and machine learning methods, whereas signature-based detection systems compare system actions with identified threats patterns and trigger predefined responses when a match is found.

References[4] and.[9] present an overview of machine learning strategies, including unsupervised as well as supervised learning, and discuss their strengths and weaknesses in detecting APTs.

The study of.[10] Introduces a graph heuristic algorithm using belief propagation, which utilizes relationships across domains during different phases of an APT attack to infer other compromised hosts and malicious domains from known entities.

Moreover, deep learning (DL), especially neural networks, has demonstrated potential in detecting APTs through its capacity to recognize complex patterns.

Research by.[11] governments, and businesses. The approaches of using machine learning or deep learning algorithms to analyze signs and abnormal behaviors of network traffic for detecting and preventing APT attacks have become popular in recent years. However, the APT attack detection approach that uses behavior analysis and evaluation techniques is facing many difficulties due to the lack of typical data of attack campaigns. To handle this situation, recent studies have selected and extracted the APT attack behaviors which based on datasets are built from experimental tools. Consequently, these properties are few and difficult to obtain in practical monitoring systems. Therefore, although the experimental results show good detection, it does not bring high efficiency in practice. For above reasons, in this paper, a new method based on network traffic analysis using a combined deep learning model to detect APT attacks will be proposed. Specifically, individual deep learning networks such as multilayer perceptron (MLP investigates the use of DL approaches, which includes deep neural networks as well as convolutional neural networks, for detecting APT threats. The research assesses the advantages and challenges of integrating DL into current detection frameworks.

Further,[12] proposes a DL framework designed to tackle APT attacks, which it views as multi-vector and multi-stage processes requiring a comprehensive approach. This framework utilizes entire network flows, particularly raw data, to detect specific types of anomalies and behaviors. Following the success of AlphaGo, DL technologies have proven their extensive capabilities in artificial intelligence.

The work by.[13] offers a thorough examination of multi-layered defense tactics, integrating signature-based detection, anomaly detection, and behavioral analysis. This study underscores the importance of a varied and stratified approach to effectively detect APTs across multiple attack vectors.

In the research. The methods CompreX, OD, OC3, FPOF and AVF have proven effective about detecting anomalies, showcasing their utility in identifying potential threats. However, it’s common for these techniques to also flag hundreds or thousands of anomalous processes that aren’t linked to any malicious activity.

According to[15] An remarkable level of detection accuracy was reached by using only 12 of the dataset’s 65 features. This result demonstrates the effectiveness of the feature selection approach in improving the detecting system’s performance. Using a multi-label method could allow for a more detailed identification of complex, multiple threats, thereby offering greater insights and boosting network security.

The approach to detection outlined in[16] is highly effective, especially in the early stages of APTs. It analyzes DNS logs and Packet Capture (Pcap) data obtained from the data center. This approach makes use of critical indicators found in DNS logs and network traffic to quickly identify suspicious actions suggestive of APT activity. This strategy improves the ability to detect and respond to possible attacks in their early phases by focusing on data from the data center, where APTs frequently breach networks.

The study[17] about the application of data dimensionality reduction, especially through a combination of SVM (Support Vector Machines) and PCA (Principal Component Analysis), achieved notable findings. While the reduction in data dimensionality had no significant effect on detection accuracy, it did significantly enhance detection speed. The strategy improved computational efficiency while maintaining detection efficacy by combining SVM for classification and PCA for dimensionality reduction.

As stated by,[18] a correlation framework was created to build links between alerts issued during the first phase and possible APT assaults, with the goal of minimizing false positive rates. Synthetic data was employed to test this system because there were few suitable publically available data sources. Despite this restriction, the framework’s capacity to connect alerts to genuine APT attacks was demonstrated, highlighting its potential to reduce false positives and improve overall threat detection precision.

As of, the proposal introduces a semi-supervised learning approach that relies on an improved Self-Organizing Neural Network (SNN)-based clustering algorithm. This approach tries to improve APT detection accuracy by combining labeled and unlabeled data. However, a fundamental issue connected with this strategy is the substantial computational overhead it incurs. Processing enormous volumes of data and training neural networks can take significant computer resources.

The research[19] has been useful in detecting and alerting to Command and Control (C&C) server activity using the Random Forest machine learning approach, as well as identifying anomalous behavioral patterns in network data.

But this approach has disadvantages when APT assaults use encryption to relay information. Encryption can disguise harmful activity, making typical detection methods less effective.[20]

Also in,[21] the multi-feature spatial weighted combination SVM classification detection algorithm outperforms standard single classification approaches in detecting disguised APT attacks, achieving higher accuracy and reduced rates of false alarms. However, using this advanced algorithm has a disadvantage: it increases temporal complexity. The increased computational costs of processing multiple information and adopting spatial weighting contribute to lengthier processing times. This situation emphasizes the importance of additional refinement and analysis to improve the accuracy of these detection systems, reducing the likelihood of false positives, which could result in inefficiencies and alert weakness within security operations.[22]

3. METHODOLOGY

Fig. 3 depicts the APTGuard model's methodology, which comprises a structured approach to detecting and classifying Advanced Persistent Threats (APTs) utilizing machine learning techniques. The procedure is divided into various stages: data preparation, feature selection, model training, validation, testing to the evaluation stage. This methodology ensures the detection model's robustness and usefulness.

3.1 Dataset Description

For this research, we used the Unraveled dataset,[23] which is a semi-synthetic dataset designed to imitate the traffic patterns and attack vectors associated with APTs. This dataset mixes real network traffic data with synthetically generated attack scenarios that mimic the tactics and techniques employed by APTs. The dataset contains a wide range of APT signatures and normal traffic, creating a balanced environment for training and testing our models. The collection includes 97,416 APT alarms for 6,877,330 network flow packets.

"Unraveled" dataset presents several compelling reasons for using it to detect APTs: Realism, Comprehensive Information, Emulation of Real-world Scenarios, Diverse Attack Skills and Sources, Employee Behavior Generation Model, Challenge for Detection Models. The dataset has been processed into CSV files containing network flow information, extracted from pcap files using NFStream. We chose it after compare it with NLS-KDD,[24] HERITRIX.[25] DAPT 2020[26] and SCVIC-APT-2021[27] datasets because it's covering the attacks for more stages that is mean for longer period as known about APTs from the Reconnaissance stage through the lateral moving stage then to cover up stage as illustrated in **Table 1.** The dataset after processing has 2,618,568 rows and 89 columns of network flows data.

3.2 Data Preparation

The first step involves preparing the Unraveled dataset for analysis. This includes:

Data Cleaning: Removing corrupt or inaccurate records from the dataset, addressing outliers, and filling missing values.

Table 1. Comparing several datasets and Unraveled.

APTs stage	Dataset				
	NSL-KDD	HERITRIX	DAPT 2020	SCVIC-APT-2021	Unraveled 2023
Conducted Over	-	-	5 days	5 days	6 weeks
Number Of Features	8	-	85	84	89
Reconnaissance Stage	Yes	-	Yes	Yes	Yes
First Establish Foothold	Yes	Yes	Yes	Yes	Yes
During Lateral Movement	-	-	Yes	Yes	Yes
Data Exfiltration Stage	-	-	Yes	Yes	Yes
Cover Up Stage	-	-	-	-	Yes

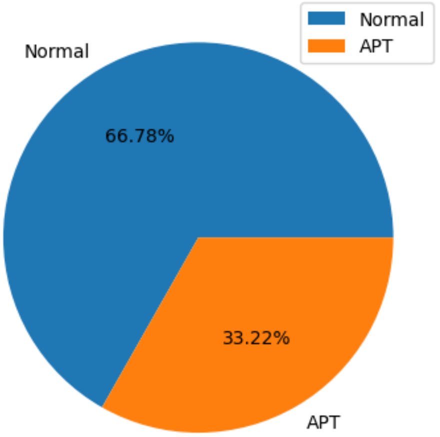

Fig. 1: Percentage of APTs in the processed data.

Data Transformation: Normalizing data to bring all features to a similar scale, which is crucial for models like SVM and KNN that are sensitive to the range of data values. The **Fig. 1** presents the Percentage of APTs in the processed data.

Data Segregation: Dividing the dataset into training, verifying, and testing sets. Approximately 60% of data utilized during training, 20% over validation, and the balance of 20% for testing.

As we can see in **Fig. 2** the data took for 6 weeks to mentioned the importance of the persistent for APTs that is visible how each stage started from reconnaissance to Cover up and what happened between them during the 6 weeks and the percentage of the threats in each week how it increased till the last week that the attacks happened fully.

3.3 Proposed Model

The input dataset of APTs undergoes a comprehensive preparation and processing phase to eliminate irrelevant data. Label Encoding is employed to convert categorical text data into integer values. Feature selection is conducted based on Information Value (IV) along with Weight of Evidence (WoE) metrics. Subsequently, the selected features are utilized in the application with four algorithms. The architecture of APTGuard model for APTs detection is depicted in **Fig. 3** and it outlines a technique for using unraveled semi-synthetic dataset. This approach consists of many critical stages like: data gathering, feature selection, data processing, and the use of machine learning algorithms.

Data collection entails obtaining a full set of network flow data, which includes both normal traffic and harmful APT activity. This dataset forms the cornerstone for our analysis and is sourced from a number of environments and timeframes throughout six weeks, ensuring diversity and relevance.

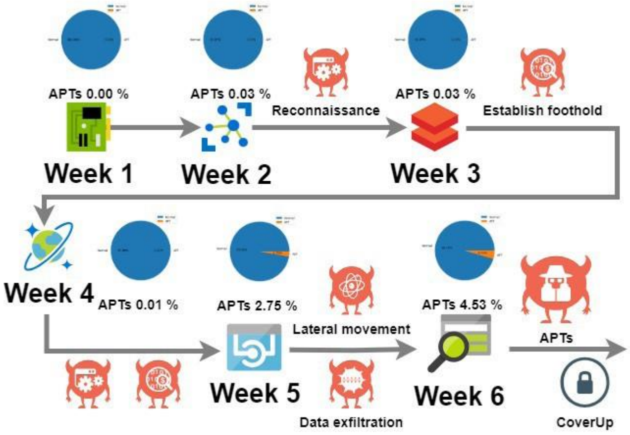

Fig. 2: APTs stages and percentage in the processed dataset.

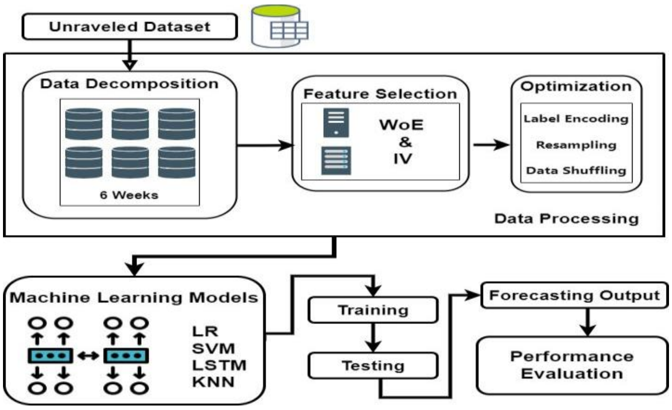

Figure. 3 APTGuard model

Feature Selection analyzes the derived network flow characteristics to select the most relevant features for APT detection. Techniques such as WoE and IV are used to assess the predictive power of the features and determine which ones are most important for classification.

Data processing comprises a variety of preparatory methods that improve the accuracy of machine learning. Normalization, dimensionality reduction, data reorganization, and feature scaling are all steps taken to optimize the dataset for effective analysis.

The processed data is prepared for use in testing and training. Four machine learning implemented to detect and classify APTs:

1. Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) that can learn order dependency in sequence prediction tasks.
2. Logistic Regression (LR) is a statistical model that employs a logistic function to compute probabilities for a dependent variable with a binary value.

3. Support Vector Machine (SVM) - A powerful classification algorithm that locates a hyperplane in an N-dimensional space and classifies the data points. SVM: a supervised learning algorithm designed for classification tasks, utilizing hyperplanes to establish decision boundaries between two classes of data.
4. K-Nearest Neighbors (KNN) - A non-parametric approach to both regression and classification that measures distance from the nearest K number of samples.

The proposed model utilizes this methodology to accurately classify and detect APTs, leveraging the comprehensive and meticulously processed dataset to achieve high levels of accuracy and reliability.

4. Evaluation and Results

In our feature selection process for the classification model, we utilized Weight of Evidence (WOE) and Information Value (IV) to rank and choose variables. These methods help identify the most predictive variables. **Fig. 4** show the 20 most important feature, ensuring that those with the highest IV values are prioritized for inclusion. This approach reduces dimensionality, making the model more efficient and interpretable. By choosing the features with high predictive potential, we enhance the overall performance of our model. This targeted selection process is crucial for improving accuracy and reliability in our classification tasks.

Many assessment criteria have been frequently employed to evaluate the performance of classification models. The metrics consist of the F1 score, recall, accuracy, and precision. The confusion matrix, displayed in **Table 2**, is employed to assess the classifier's major indicators of performance.

According to Eq. (01), accuracy is the percentage of accurately classified instances in the dataset as a whole. Precision—also referred to as the Detection Rate and determined by Eq. (2), it calculates the ratio of correctly classified cases to all cases that have been classified. Eq. (3) defines recall as the proportion of correctly identified instances to the total number of true positive instances. Eq. (4) mentions the False Alarm Rate, which indicates the frequency with which attacks are incorrectly classified or misidentified. At last, F1 Score, often known as the F-measure, as given by Eq. (5), offers a thorough assessment of a test's accuracy by integrating precision and recall into a single measure. This measure facilitates comparison and analysis of different models and methodologies and aids in understanding the classifier's key performance metrics.

Table 2. CONFUSION MATRIX

Threat		Predicted	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

$$\text{Accuracy} = \frac{TP + TN}{TP + TN+FP +FN} \quad (1)$$

$$\text{Precision}(P) = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall }(R) = \frac{TP}{TP + FN} \quad (3)$$

$$\text{False\_Alarm\_Rate} = \frac{FP}{P + TN} \quad (4)$$

$$\text{F1\_score} = 2 + \frac{P + R}{P + R} \quad (5)$$

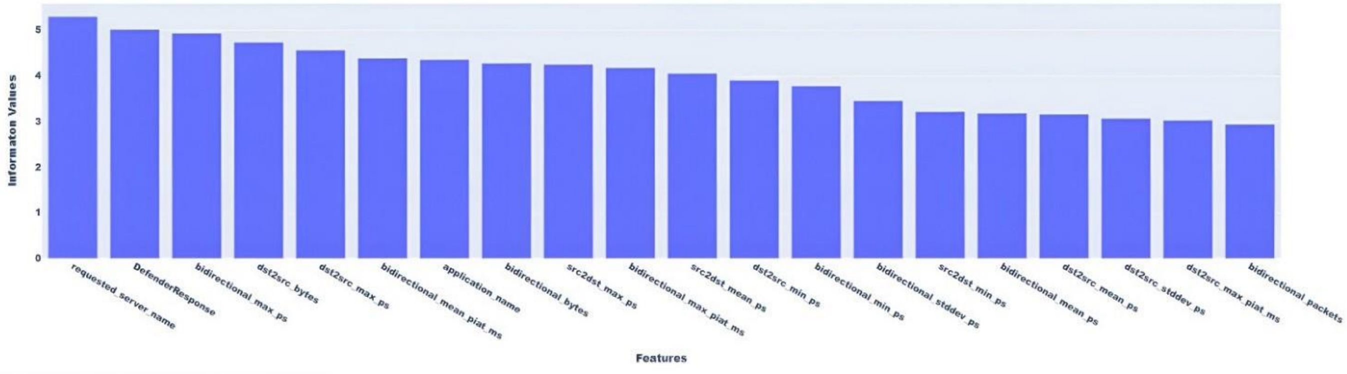

Fig. 4. The 20 most important features according to Information Value

Table 3 and **Fig. 5** provide a summary of the performance outcomes of the algorithms, including LSTM, LR, SVM and KNN.

The evaluation results demonstrate that among the tested models, LSTM algorithm achieves the highest performance with an accuracy of 99.89%. Comparatively, the Logistic Regression (LR) model attains an accuracy of 96.37%, SVM algorithm achieves 95.74%, and KNN Algorithm reaches 98.65%.

Table 4 highlights the accuracy achieved by various models. It is a critical reference point for understanding the relative performance of each algorithm and used datasets to identifying the most effective approach.

The evaluation underscores the strengths and limitations of each model, focusing in particular on the APTGuard model's better performance Shown in **Fig. 6**.

The enhanced functionality of the APTGuard model as shown in **Fig. 7** can be attributed to its ability to capture and learn from sequential patterns in the network traffic data, which is critical for identifying the stealthy and persistent nature of APTs.

Additionally, LSTM's proficiency in handling long-term dependencies makes it particularly effective in distinguishing between benign and malicious activities over extended periods. Therefore, LSTM is identified as the most effective algorithm for detecting APTs in this study.

Table 3. Experimental Results

	LSTM	LR	SVM	KNN
Precision	98.90	96.37	95.74	95.77
Recall	99.24	98.15	97.98	97.18
False_Alarm_Rate	99.96	97.51	98.23	98.26
F1_score	98.55	93.15	96.82	95.45
Accuracy	**99.89**	**96.37**	**95.74**	**98.65**

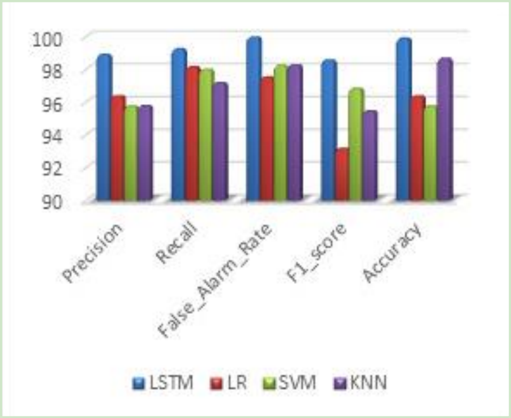

Fig. 5: Outcomes of the applied algorithms.

Table 4. Comparative on various APTs detection models.

Model	**Ref.**	**Dataset**	**Accuracy (%)**
CNN-BiLSTM	[30]	NSL-KDD	83.58
SVM-Naïve Bayes (NB)	[31]	NSL-KDD	99.35
AE-Triplet Network	[32]	UNSW-NB15	92.4
AE-SVM-GO	[33]	NSL-KDD	99.6
CRNN	[34]	CSE-CIC-DS2018	97.6
LSTM	[10]	Generated	99.08
APTGuard	**-**	**Unraveled 2023**	**99.89**

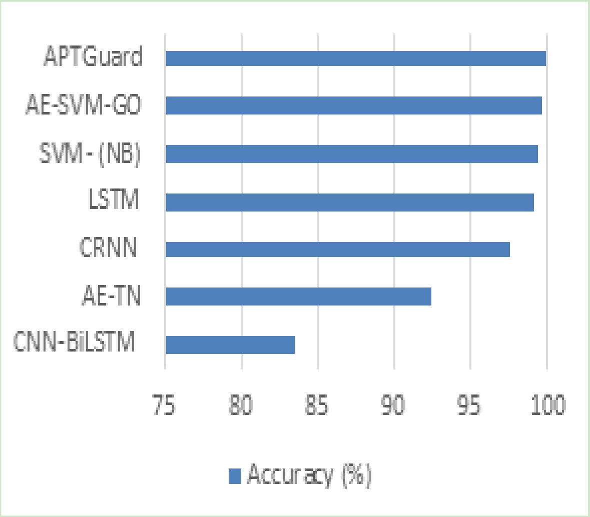

Fig. 6 Comparative the accuracy of various models

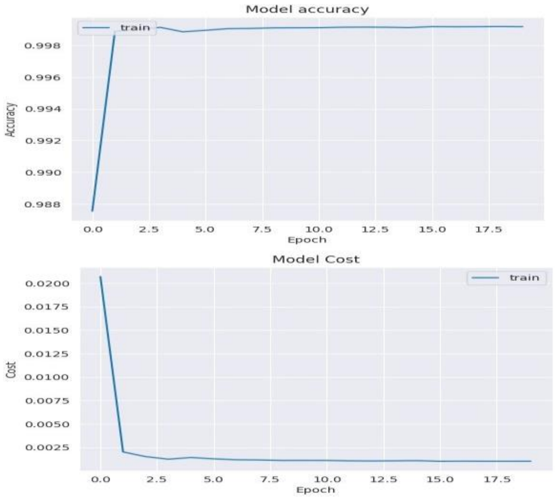

Fig. 7 APTGuard-LSTM (Accuracy & Loss mode).

The APTGuard scored a remarkable 0.9989, this illustrates how well it detects APTs and shows that it can identify complex patterns correlations in the data. APTGuard model is a reliable option for APT detection due to its excellent accuracy and constantly modifies the parameters it contains during training in order to reduce the loss, which eventually leads to improved predictions and increased accuracy.

5 Conclusion And Future Work

Machine learning models and optimizations on the semi-synthetic Advanced Persistent Threats APTs dataset and utilizing Weight of Evidence (WOE) and Information Value (IV) to rank and choose variables are combined to create the model used in this paper. Reducing the dimension of features utilized in the models is the foundation of the optimization technique. Despite requiring more time for training and testing, APTGuard, the best performance model, achieved an outstanding 99.89% accuracy with the processed Unraveled dataset. The benefit of the suggested model is that it will give the researcher a dominant input to the detection model and enhance its performance by combining optimization techniques with machine learning. Additionally, it will provide the system and pertinent administrators a heads-up so they may take appropriate action, such incident response, to lessen the threat's impact. While the suggested model has demonstrated noteworthy results in terms of forecast performance, it might be challenging to ensure that a proposed model operates well across different real datasets or time periods. The model might need to be adjusted or retrained when applied to new networks since it might be susceptible to changes in the distribution of the data. When compared to existing models, the performance of the proposed APTGuard is impressive. It is recommended to use a real-time dataset in the future for more precise and timely detection.

Conflicts of Interest

Authors declare no conflict of interest.

Author Contributions

Mr. Nadim Ibrahim played a key role in crafting the research inquiries, gathering data, structuring the study, devising the methodology, conducting data analysis, and drafting the manuscript.

Dr. N.R. Rajalakshmi Offered crucial perspectives, aided in result interpretation.

Mr. Karam Hammadeh Contributed to the literature survey, verified the data.

References

1. S. K. Chenniappanadar, S. Gnanamurthy, V. K. Sakthivelu, and V. K. Kaliappan, "A Supervised Machine Learning Based Intrusion Detection Model for Detecting Cyber-Attacks Against Computer System," *Int. J. Commun. Networks Inf. Secur.*, vol. 14, no. 3, pp. 16-25, 2022, doi: 10.17762/ijcnis.v14i3.5567.
2. I. M. Technologies, "Abnormal Event Detection for 5G-IoT Devices," vol. 17, no. 17, pp. 59–71, 2023.
3. S. Vinoth Kumar, H. Shaheen, A. Christopher Paul, M. Shyamala Devi, R. Aruna, and S. Sangeetha, "Information-Based Image Extraction with Data Mining Techniques for Quality Retrieval," in *Lecture Notes in Networks and Systems*, R. P. Mahapatra, S. K. Peddoju, S. Roy, and P. Parwekar, Eds., Singapore: Springer Nature Singapore, 2023, pp. 175–188. doi: 10.1007/978-981-19-8825-7_16.
4. M. Imran, H. U. R. Siddiqui, A. Raza, M. A. Raza, F. Rustam, and I. Ashraf, "A performance overview of machine learning-based defense strategies for advanced persistent threats in industrial control systems," *Comput. Secur.*, vol. 134, no. August, p. 103445, 2023, doi: 10.1016/j.cose.2023.103445.
5. N. Ibrahim, N. R. Rajalakshmi, and K. Hammadeh, "Exploration of Defensive Strategies, Detection Mechanisms, and Response Tactics against Advanced Persistent Threats APTs," *Nanotechnol. Perceptions*, vol. 20, no. S4, pp. 439–455, 2024, doi: 10.62441/nano-ntp.v20iS4.33.
6. M. Abu Talib, Q. Nasir, A. Bou Nassif, T. Mokhamed, N. Ahmed, and B. Mahfood, "APT beaconing detection: A systematic review," *Comput. Secur.*, vol. 122, no. July 2023, 2022, doi: 10.1016/j.cose.2022.102875.
7. W. Ren *et al.*, "APT Attack Detection Based on Graph Convolutional Neural Networks," *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, 2023, doi: 10.1007/s44196-023-00369-5.
8. A. S. AL-Aamri, R. Abdulghafor, S. Turaev, I. Al-Shaikhli, A. Zeki, and S. Talib, "Machine Learning for APT Detection," *Sustainability*, vol. 15, no. 18, p. 13820, 2023, doi: 10.3390/su151813820.
9. F. Shen, Z. Liu, and L. Perigo, "Strategic Monitoring for Efficient Detection of Simultaneous APT Attacks with Limited Resources," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 3, pp. 19–24, 2023, doi: 10.14569/IJACSA.2023.0140303.
10. K. Hammadeh and M. Kavitha, "Unraveling Ransomware: Detecting Threats with Advanced Machine Learning Algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 9, pp. 484–491, 2023, doi: 10.14569/IJACSA.2023.0140952.
11. A. Oprea, Z. Li, T. F. Yen, S. H. Chin, and S. Alrwais, "Detection of Early-Stage Enterprise Infection by Mining Large-Scale Log Data," *Proc. Int. Conf. Dependable Syst. Networks*, vol. 2015-Septe, pp. 45–56, 2015, doi: 10.1109/DSN.2015.14.
12. T. Bodström and T. Hämäläinen, "A novel deep learning stack for APT detection," Appl. Sci., vol. 9, no. 6, 2019, doi: 10.3390/app9061055.

13. C. Do Xuan and M. H. Dao, "A novel approach for APT attack detection based on combined deep learning model," Neural Comput. Appl., vol. 33, no. 20, pp. 13251–13264, 2021, doi: 10.1007/s00521-021-05952-5.

14. P. R. Brandao and V. Limonova, "Defense Methodologies Against Advanced Persistent Threats," Am. J. Appl. Sci., vol. 18, no. 1, pp. 207–212, 2021, doi: 10.3844/ajassp.2021.207.212.

15. G. Berrada et al., "A baseline for unsupervised advanced persistent threat detection in system-level provenance," Futur. Gener. Comput. Syst., vol. 108, pp. 401–413, 2020, doi: 10.1016/j.future.2020.02.015.

16. J. Al-Saraireh and A. Masarweh, "A novel approach for detecting advanced persistent threats," Egypt. Informat- ics J., vol. 23, no. 4, pp. 45–55, 2022, doi: 10.1016/j. eij.2022.06.005.

17. D. X. Cho and H. H. Nam, "A method of monitoring and de- tecting APT attacks based on unknown domains," Procedia Comput. Sci., vol. 150, pp. 316–323, 2019, doi: 10.1016/j. procs.2019.02.058.

18. W. L. Chu, C. J. Lin, and K. N. Chang, "Detection and classification of advanced persistent threats and attacks using the support vector machine," Appl. Sci., vol. 9, no. 21, 2019, doi: 10.3390/app9214579.

19. I. Ghafir et al., "Detection of advanced persistent threat using machine-learning correlation analysis," Futur. Ge- ner. Comput. Syst., vol. 89, pp. 349–359, Dec. 2018, doi: 10.1016/j.future.2018.06.055.

20. A. Zimba, H. Chen, Z. Wang, and M. Chishimba, "Modeling and detection of the multi-stages of Advanced Persistent Threats attacks based on semi-supervised learning and complex net- works characteristics," Futur. Gener. Comput. Syst., vol. 106, pp. 501–517, 2020, doi: 10.1016/j.future.2020.01.032.

21. C. Do Xuan, L. Van Duong, and V. N. Tisenko, "Detecting C&C server in the APT attack based on network traffic using machine learning," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 5, pp. 22–27, 2020, doi: 10.14569/IJAC- SA.2020.0110504.

22. K. Hammadeh, M. Kavitha, and N. Ibrahim, "Enhancing Cybersecurity in Software- Defined Networking : A Hybrid Approach for Advanced DDoS Detection and Mitigation," vol. 4, pp. 514–529, 2024.

23. X. Wang, Q. Liu, Z. Pan, and G. Pang, "APT attack detection algorithm based on spatio-temporal association analysis in industrial network," J. Ambient Intell. Human- iz. Comput., no. 0123456789, 2020, doi: 10.1007/s12652- 020-01840-3.

24. M. S. Devi, S. V. Kumar, P. S. Ramesh, A. Kavitha, K. Jayasree, and V. S. S. Rajesh, "Feature Reduced Anova Element Over- sampling Elucidation Based Categorisation for Hepatitis C Virus Prognostication," Lect. Notes Networks Syst., vol. 600, pp. 375 – 385, 2023, doi: 10.1007/978-981-19 8825-7_32.

25. S. Myneni et al., "Unraveled — A semi-synthetic dataset for Advanced Persistent Threats," Comput. Networks, vol. 227, no. March, p. 109688, 2023, doi: 10.1016/j.com- net.2023.109688.

26. M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set in Computational Intelligence for Security and Defense Applications," Comput. Intell. Secur. Def. Appl., no. Cisda, pp. 1–6, 2009.

27. Y. Wang, W. D. Cai, and P. C. Wei, "A deep learning approach for detecting malicious JavaScript code," Secur. Commun. Networks, vol. 9, no. 11, pp. 1520–1534, 2016, doi: 10.1002/sec.1441.

28. S. Myneni et al., "DAPT 2020 - Constructing a Benchmark Dataset for Advanced Persistent Threats," Commun. Com- put. Inf. Sci., vol. 1271 CCIS, pp. 138–163, 2020, doi: 10.1007/978-3-030-59621-7_8.

29. J. Liu et al., "A New Realistic Benchmark for Advanced Persistent Threats in Network Traffic," IEEE Netw. Lett., vol. 4, no. 3, pp. 162–166, 2022, doi: 10.1109/lnet.2022.3185553.

30. K. Jiang, W. Wang, A. Wang, and H. Wu, "Network Intrusion Detection Combined Hybrid Sampling with Deep Hierarchi- cal Network," IEEE Access, vol. 8, no. 3, pp. 32464–32476, 2020, doi: 10.1109/ACCESS.2020.2973730.

31. . J. Gu and S. Lu, "An effective intrusion detection approach using SVM with naïve Bayes feature embedding," Comput. Secur., vol. 103, p. 102158, 2021, doi: 10.1016/j. cose.2020.102158.

32. G. Andresini, A. Appice, and D. Malerba, "Autoencod- er-based deep metric learning for network intrusion detection," Inf. Sci. (Ny)., vol. 569, no. xxxx, pp. 706–727, 2021, doi: 10.1016/j.ins.2021.05.016.

33. S. R. Chikkalwar and Y. Garapati, "Autoencoder – Support Vector Machine – Grasshopper Optimization for Intrusion Detection System," Int. J. Intell. Eng. Syst., vol. 15, no. 4, pp. 406–414, 2022, doi: 10.22266/ijies2022. 0831.36.

34. M. A. Khan, "HCRNNIDS : Hybrid Convolutional Recurrent Neural," Multidiscip. Digit. Publ. Inst., 2021.