Detecting Cyber Threats using Machine Learning: A Proactive Defence Approach

Kunal Chimurkar
Department of Master in Computer Application, K.D.K. College of Engineering and Management Nandanvan, Nagpur, Maharashtra, India

Darshan Khirekar; Shweta Choudhary
Sanjana Jenekar; Anuradha Muttemwar
Department of Master in Computer Application, G H Raisoni College of Engineering and Management Nagpur, Maharashtra, India

Abstract

In this paper, the author is aimed at presenting how machine learning enhances detection of cyber threats with appropriate feature engineering and quality datasets. The author employs commonly used benchmarking datasets such as KDDCup99, NSL-KDD, and CICIDS2017 that provide a good platform for intrusion detection system training and evaluation. Other attributes such as IP address, type of protocol, and other network flow attributes are essential to learning and improve threatdetectionaccuracy.

However, the writer identifies severe issues in this space, including the issue of unbalance within datasets where traffic is benign-centric, and advanced evasions being used to make evasion from detection systems. They are constraints which bring down performance in conventional machine learning models and call for devising more resilient and adaptive models.

To answer all of this, the author feels it is feasible to discuss sophisticated techniques such as federated learning, where model training could be conducted on a decentralized platform while data are secured. some of which are being positively developed include threat response automation and feature extraction, leveraging secure logging mechanisms through the implementation of blockchain technology, and adaptive models having the ability to learn how to adapt to future threats. The accuracy and flexibility of cyber threat detection systems can be significantly enhanced by following these steps.

The author wishes that this work contributes to the pool of research that attempts to enhance cybersecurity mechanisms by utilizing machine learning-based techniques that are smart, scalable, and secure.

Keywords

Machine Learning, Cyber Threat Detection, Feature Engineering, Federated Learning, Blockchain Technology

1. Introduction

Cybersecurity is more important now than ever as today people, businesses, and governments all around the world are interconnected by computers. The need to keep sensitive information safe and make sure that systems are secure has increased as reliance on the digital world has increased. However, cybercrime has grown and expanded in complexity and frequency at an incredible rate due to this growth. Malware, phishing, ransomware, and Distributed Denial of Service (DDoS) attacks, for example, are all examples of cyber threats that are not only becoming more prevalent but are also becoming more sophisticated and can defeat traditional defences more easily.

In the past, cybersecurity systems relied on signature-based or rule-based detection methods. These methods are based on predefined patterns of malicious activities and rules of identification. These methods are efficient at spotting recognized threats but ineffective at detecting the new rapidly growing cyber threats. In contrast, machine learning can dynamically learn from a significant amount of data in detecting new unknown threats.

In the rapidly changing world of threats, it is important to promptly identify cyber threats at

DOI: https://doi.org/10.5281/zenodo.18398325

the preparatory stage. This helps to avoid possible consequences and threats that may be dangerous. Machine learning helps to identify in a timely manner, since it can recognize risks and see any anomalies. Therefore, by means of machine learning technologies in cybersecurity and their integration into the entire defence infrastructure, the organization can move from reactive protection to proactive protection.

This paper examines how Machine Learning is used in cybersecurity in the prevention of cyber threats, offering a more advanced, smart, and forward-thinking security solution.

2. Literature Review

Over the past years, various types of research have been conducted in the field of cyber threat detection. This trend has been necessitated by the need to protect computer systems from cyber threats over the years. Initially, the field of cybersecurity used signature-based and heuristic approaches, which were effective when used on known threats. These strategies were not effective with zero-day attacks and rapidly changing malware. Null. This led to various advancements in the field of cyber threat detection to solve these challenges.

Cyber threats come in various forms and thus require different methods of detection. Common threats include malware, phishing attacks, distributed denial-of-service (DDoS) attacks, and ransomware. Threats are hard to detect due to the sheer volume of traffic, the stealth of the attackers, the use of encryption, and the sophistication of the evolving threats. Traditional systems have a high rate of false positives and are not updated often enough to keep up with new threats.

In order to provide a solution to these challenges, a variety of ML techniques have been implemented in the field of cyber threat detection. The techniques include supervised learning methods such as Support Vector Machines (SVM) and decision tree algorithms for classifying normal and malicious activities based on a labelled dataset. Moreover, random forests have been applied to cyber threat detection to effectively cover complex feature spaces. More recently, neural networks and deep learning models have advanced the detection of cyber threats, by identifying various patterns within large datasets.

Several Machine Learning (ML) algorithms and systems are being implemented in the field of cybersecurity. Systems like Snort and Suricata are using ML algorithms. Moreover, the many research prototypes and commercial products demonstrate that there is rapid development in the field of combining ML with cybersecurity to develop proactive, adaptive, and automated security systems. This shows the importance of Machine Learning in combating the problem of ever-evolving cyber threats.

3. Machine Learning Techniques For Cyber Threat Detection

Machine learning presents diverse means that are effective in detecting and preventing cyber threats and these multiple methods are quite useful and are very powerful based on the data type in place.

3.1 Supervised Learning:

The cybersecurity field has widely adopted supervised learning systems for categorizing data as either bad or not bad. The model training process involves using a model supervised with labelled data from yesterday. The classification one needs to use various models, for example, tree-based used to identify some network attacks that repeat. To detect these attacks, the model is then used to identify threats and for offering protective solutions.

3.2 Unsupervised Learning:

It can be difficult to find realistic datasets that contain labelled data. This is an unsolvable problem especially in the domain of cybersecurity. However, unsupervised learning methods, such as clustering, and anomaly detection, can be utilized to understand the normal behaviour of a network and detect any anomalies that may exist in it. Anomaly detection methods are especially useful for detecting both zero-day and insider threats. For example, techniques such as K-means, and autoencoders, have high accuracy in detecting new attacks and zero-day vulnerabilities.

3.3 Reinforcement Learning:

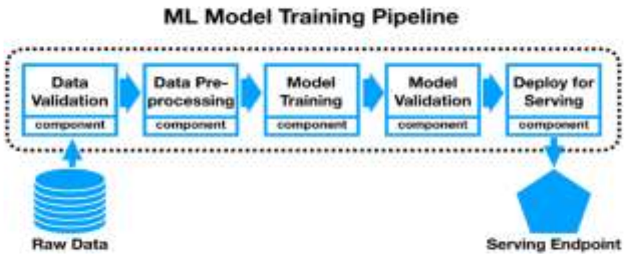
The reinforcement learning (RL) is previously an exciting new choice for cyber protection since it will allow the machine to learn the best responses to cyber threats by conducting the interactions in the environment. In this case, the machine learning agent learns what is successful such as allowing or blocking the

DOI: https://doi.org/10.5281/zenodo.18398325

network traffic and adjusts its strategy to maximize the system defence. The use of RL benefits the adaptive security system that it must have to learn how to operate the defender against the evolving cyber threats.

3.4 Deep Learning:

Techniques like deep learning, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), can identify intricate and intricate patterns of malware in vast amounts of data. These networks can determine very elaborate attacks that won't be identified by other methods, drawing on high-level features from raw data. Their capacity to deal with large datasets and perform computations rapidly makes them excellent for real-time identification and known malware identification on a wide scale.
Using these methods in protection can help shift security systems from being static or rule-based systems to a more adaptive, intelligent, and predictive system.



Source: medium.com

The success of Machine Learning Cyber Security Models in detecting cyber threats depends on the availability of useful data sets as well as the right feature extraction methods. For this reason, several benchmark datasets have been developed over the years in the field of cybersecurity.

3.5 Commonly Used Datasets:

- KDDCup99: It is one of the earliest and most widely used datasets for intrusion detection. It includes a wide range of simulated attacks and normal network traffic.
- NSL-KDD: A modified version of the KDDCup99 dataset that addresses some of its limitations in order to provide a more balanced dataset for classification evaluation, and avoid redundancy.
- CICIDS2017 presents a contemporary and extensive database that incorporates genuine network operations and the latest modes of cybercrimes such as DDoS, web attacks, brute force, and infiltration.

4. **Dataset and Features**

Performance of Machine Learning models to identify cyber threats primarily relies on the quality and number of datasets available and good feature selection. A number of benchmark datasets have been developed over time to train and test cyber security models.

4.1 Popular Datasets:

KDDCup99: Most popular and oldest intrusion detection dataset. It includes enormous variability of created attacks and normal network traffic.
NSL-KDD: A sanitized version of KDDCup99, with duplicate samples eliminated and some of the natural skewing alleviated, thus an actionable and representative dataset upon which to train models against.
CICIDS2017: A comprehensive and modern dataset with real-world network traffic, including normal usage and all types of attack scenarios one encounters in current times like DDoS, brute force, infiltration, and web attacks.

4.2 Detection Key Features

Effective cyber threat identification is founded upon the tracking of a set of network attributes, such as
IP Address: Allows the source and destination of the network traffic to be identified.
Port Number: Allows the service being accessed to be identified, from which anomalous access patterns may be identified.
Protocol Type: Provides the protocol type as TCP, UDP, ICMP, etc., and allows threats to be identified.
Packet Size: Abnormal packet sizes may be a sign of attacks like DDoS or data exfiltration.
Traffic Behaviour: Connection duration, failed login attempts, and frequency of access can signal malicious behaviour.

4.3 Effect of Feature Engineering:

Feature engineering is the most crucial step in increasing the performance of ML models. All of them have something to do with feature selection, feature transformation, and feature creation, which are closer to the original data structure. Good feature engineering

DOI: https://doi.org/10.5281/zenodo.18398325

significantly affects whether or not the model can distinguish normal vs. attack traffic, reduce false positives, and generalize to novel attack types. Normalization, dimensionality reduction, and feature selection are techniques widely used in trying to simplify learning and produce good cyber threat detection.

5. **System Architecture**

A typical machine learning-driven cyber threat detection system is structured in a well-organized pipeline, aimed at processing huge streams of data efficiently and actively detecting security threats. A solid architecture enhances both detection performance and system robustness. The key phases of such a system are Data Collection, Preprocessing, Feature Extraction, Model Training, Threat Detection, and Response [1][2].

5.1 **Data Collection**

The first step is the gathering of data from diversified sources, for example, network traffic logs, server logs, endpoint systems, intrusion detection systems (IDS), honeypots, and firewall logs. Standard benchmark datasets such as KDDCup99 [3], NSL-KDD [4], and CICIDS2017 [5] are regularly employed in academic papers for machine learning model training and testing. Real-time collectors such as Zeek (formerly Bro) or SIEM tools are utilized for gathering rich and diverse telemetry data in production scenarios.

5.2 **Preprocessing**

Raw collected data is generally incomplete, noisy, and inconsistent. Preprocessing transforms this data into a normalized and structured format. This includes handling missing values, feature removal of unwanted features, encoding categorical features, and numerical feature normalization [6]. Flow-based aggregation techniques, session reconstruction, and time-windowing are some of the other preprocessing operations that help identify temporal patterns in cyber activity [7].

5.3 **Feature Extraction**

Feature engineering is an essential step that has a direct impact on detection. Important features typically include IP addresses, port numbers, connection durations, types of protocols, packet size distribution, and login attempt behaviour [8]. Algorithmic feature selection methods such as PCA and RFE help in reducing dimensionality and maintaining useful patterns beneficial to the detection of intrusions [9].

5.4 **Model Training**

In this stage, machine learning models are trained to detect or classify abnormal activities based on the engineered features:

- Supervised learning algorithms like Support Vector Machines (SVM), Random Forests, and XGBoost are used when there is labelled attack data [10].
- Unsupervised learning techniques like clustering (e.g., K-Means) or anomaly detection with autoencoders are used in cases where there is not much labelled data [11].
- Deep learning methods, such as Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM), are more often employed to identify sophisticated, non-linear attack patterns from big data sets [12].

5.5 **Threat Detection**

After deployment, the trained model regularly scans real-time traffic or logs and tags events as malicious or benign. Stream processing libraries such as Apache Flink and Apache Kafka are usually used in operational systems to support scalable real-time threat detection [13].

5.6 **Response**

When malicious activity is detected, the system invokes proper response mechanisms. These are:

- Issuing alerts for security teams to analyze manually.
- Automatically blocking suspected IP addresses or closing malicious sessions.
- Triggering automatic playbooks via SOAR (Security Orchestration, Automation, and Response) platforms.

A quick and automated response function is crucial to reduce possible damages and contain incidents efficiently.

6. **Challenges**

6.1 **Data Imbalance**

In cybersecurity datasets, attack traffic usually constitutes an extremely small portion of

DOI: https://doi.org/10.5281/zenodo.18398325

normal network traffic. Such class imbalance makes it difficult for machine learning models to accurately classify attacks since models are skewed toward the majority (benign) class. As a result, most attacks escape detection, profoundly affecting the detection system's performance. Data imbalance is still one of the serious challenges in constructing machine learning solutions to identify cyber threats.

6.2 **Evasion Techniques by Attackers**

Cyber attackers repeatedly revise their attacks in an effort to evade security controls. Adversarial attacks, where input data are tailored specifically to trick machine learning models, are particularly undesirable. Malware can dynamically adapt its behaviour (polymorphism and metamorphism) to evade detection as well. Building models that can learn new, unknown attack channels is necessary but a complex and on-going research activity.

6.3 Real-Time Detection vs. Accuracy Trade-offs

One of the main challenges is to achieve a balance between speed and high accuracy of detection and real-time detection. Light models will detect rapidly but with less precision and more false alarms, typically. Highly accurate deep learning models typically consume enormous amounts of compute resources and introduce delay and are thus not suitable for real-time applications. Model optimization to achieve a balance between speed and trustworthiness remains an open problem for real-world applications.

6.4 Need for Explainability

Another significant challenge is balancing real-time detection with high accuracy. Light models will detect fast but typically at the expense of reduced precision and increased false positives. Highly accurate deep learning models, in turn, typically consume large computational resources and introduce delay, making them unsuitable for real-time applications. Model optimization to trade-off between speed and reliability is still a pressing issue for real-world implementations.

7. **Future Scope**

7.1 **Combining Machine Learning with AI-driven Automated Incident Response**

Going beyond passive detection, the cybersecurity systems will harness machine learning coupled with automated response to threats in the future. Together, they will facilitate automated responses in real time like isolating vulnerable systems, blocking malicious traffic, and automatically patching vulnerabilities [2]. Automation lowers response time, enables reduced damage by attacks, and optimizes the workload of security professionals.

7.2 Use of Federated Learning to Enhance Privacy

Training robust machine learning models typically requires a lot of diverse data, but data sharing between organizations that is sensitive raises privacy concerns. Federated learning addresses this by enabling collaborative model training without sharing data centrally [6]. This approach keeps the sensitive data local but assists in optimizing global threat detection models.

7.3 Real-Time Adaptive Learning

Cyber-attacks keep changing at a fast pace, and hence static models become outdated over time. Real-time adaptive learning offers the feature of allowing models to adapt and update themselves dynamically every time they encounter new kinds of attacks. This can be done through online learning paradigms to enable systems with high detection rates for zero-day attacks [10][12]. This adaptability is very important in next-generation intrusion detection systems.

7.4 **Integration with Blockchain for Secure Data Logging**

Immutable, tamper-resistant logging is required for forensic and reporting needs. Blockchain might enable an interface to make system logs tamper-evident and verifiable using a distributed, encrypted ledger [5][8]. Blockchain and machine learning integration will enhance trust, accountability, and data integrity in cybersecurity procedures.

8. **Conclusion**

Hence, based on this argument, the author opines that machine learning now forms the building block of today's cybersecurity as it offers substantial enhancement in a timely and effective identification of cyber threats over customary rule-based security systems. Thanks

DOI: https://doi.org/10.5281/zenodo.18398325

to its capabilities of processing voluminous amounts of network traffic, detecting poor patterns, and addressing new attack mechanisms, machine learning speeds up and optimizes identification of vast attacks like malware, phishing, and intrusion attempts [2][6].

The writer concedes, however, that constant work is being done to most fully exploit machine learning in cybersecurity. Class-imbalanced data sets, smart evasion techniques, detection limits under real-time conditions, and other chronic issues continue to require constant address and innovation.

The author therefore sums up that though machine learning is an extremely potent engine of cybersecurity progress, it becomes realizable only through years of relentless research, inter-disciplinary approaches, and constant model enhancement to accommodate the fast-evolving and changing cyber threat environment.

9.**Reference**

[1] Santos, I., Brazo, F., Ugarte-Pedrero, X., & Bringas, P. G. (2013). Opcode sequences as representation of executables for data-mining-based unknown malware detection. Information Sciences, 231, 64-82.

[2] Buczak, A. L., & Guven, E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176.

[3] Stolfo, S. J., Fan, W., Lee, W., Prodromitids, A., & Chan, P. K. (2000). Cost-based modelling for fraud and intrusion detection: Results from the JAM project. DARPA Information Survivability Conference and Exposition.

[4] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defence Applications.

[5] Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. ICISSP, 108-116.

[6] Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. IEEE Symposium on Security and Privacy.

[7] Moustafa, N., & Slay, J. (2015). UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Military Communications and Information Systems Conference (MilCIS).

[8] Ring, M., Wunderlich, S., Scheuring, D., Landes, D., & Hotho, A. (2019). A Survey of Network-Based Intrusion Detection Data Sets. Computers & Security, 86, 147-167.

[9] Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. Computers & Electrical Engineering, 40(1), 16-28.

[10] Zhang, J., & Zulkernine, M. (2008). Anomaly Based Network Intrusion Detection with Unsupervised Outlier Detection. Proceedings of the 2006 IEEE International Conference on Communications.

[11] Shiravi, A., Shiravi, H., Tavallaee, M., & Ghorbani, A. A. (2012). Toward developing a systematic approach to generate benchmark datasets for intrusion detection. Computers & Security, 31(3), 357-374.

[12] Kim, G., Lee, S., & Kim, S. (2016). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.

[13] Apache Kafka. (n.d.). Distributed Event Streaming Platform.

DOI: https://doi.org/10.5281/zenodo.18398325