

Exploring methods to enhance network security through artificial intelligence

Yanni Tang

School of Optoelectronic Information and Computer Engineering, University of Shanghai for Science and Technology, Shanghai, 200093, China

yannitang@hhu.edu.cn

Abstract. In the dynamic landscape of cybersecurity, traditional rule-based systems find themselves frequently outstripped by the intricacy, variety, and mutable nature of cyber threats. This paper explores the capabilities of Machine Learning (ML) in detecting cyber-attacks, offering a fresh perspective to fortify cyber defense mechanisms. Through its unparalleled strengths in data analysis, pattern discernment, and outcome prediction, Machine Learning emerges as a promising ally in grappling with the multifaceted challenges posed by cyber adversaries. The exploration zeroes in on the potential of utilizing machine learning for cyber-attack detection, spotlighting supervised learning algorithms such as SVM and Random Forest. Experimental findings robustly underscore the value of Machine Learning in identifying potential cyber threats. In conclusion, the transformative potential of Machine Learning in the domain of cyber-attack detection is evident. Equipped with the prowess to derive insights from vast data sets, swiftly adapt to changing parameters, and preemptively recognize threats, Machine Learning promises to redefine the paradigms of cybersecurity. As the digital expanse continues to evolve, defense mechanisms must also evolve, with Machine Learning serving as a pivotal tool in this endeavor.

Keywords: Machine learning, Cybersecurity, Cyber-attack detection.

1. Introduction

The dawn of the digital era has ushered in a multitude of advantages, spanning from heightened connectivity and access to information to augmented efficiencies in various sectors. Yet, this digital evolution has also birthed new challenges, most prominently in the sphere of cybersecurity. Cyber-attacks have burgeoned into paramount concerns for individuals, enterprises, and governmental bodies alike. Conventional defensive measures, which heavily rely on rule-based systems, find themselves increasingly at odds with the intricate and ever-evolving landscape of contemporary cyber threats. This landscape signals an immediate call for more agile and efficacious solutions, highlighting the pivotal role of Machine Learning.

Machine Learning, a subset of artificial intelligence, embodies a mode of data scrutiny that streamlines the creation of analytical constructs. Through the adoption of algorithms that learn in iterative patterns from data, Machine Learning equips computers with the capability to unearth concealed insights without any explicit directional programming. Such characteristics render it a potent instrument for the detection of cyber-attacks, which frequently encompass advanced methodologies and

swift tactical shifts. Several modalities exist for integrating Machine Learning into cyber-attack detection. A prevalent technique involves supervised learning, wherein the model receives training on a dataset with established labels to categorize or forecast outcomes. To illustrate, a model might undergo training using network traffic data designated as 'normal' or 'malicious'. This training subsequently empowers the model to discern potential threats in novel, unlabeled datasets. Yet, integrating Machine Learning for the detection of cyber-attacks does not come without its set of hurdles. Dilemmas surrounding data privacy, the imperative for voluminous quantities of top-tier training data, the peril of overfitting, and the necessity to stay abreast with the swift metamorphoses of cyber threats present formidable challenges. Additionally, the enigmatic "black box" demeanor of a myriad of Machine Learning algorithms complicates the understanding of their decision-making processes, an aspect that becomes gravely pertinent in high-stakes security scenarios. However, the prospective magnitude of Machine Learning in amplifying our prowess in cyber threat detection and mitigation cannot be understated. As research and development in this domain persist, Machine Learning stands poised to solidify its indispensable position in our armory against cyber onslaughts.

2. Correlation Theory

2.1. Network Traffic

To analyze network traffic and find out intrusions, we need firstly learn about the composition of network traffic.

2.1.1. Network model. Network model is consisted with 7 layers, is also known as 7-layer network model. Each layer responsible for specific functions and tasks. Physical Layer. The physical layer manages the tangible elements of data transport, including the electrical and mechanical characteristics of the physical medium, the physical connectors, and signal operations. It is tasked with the direct transmission and receipt of unprocessed sequences of bits over the physical medium. Data Link Layer. The data link layer ensures consistent and reliable one-to-one data transmission between different nodes in the network. It is responsible for dividing data into frames, detecting and correcting errors, and managing access to physical media. This layer also handles flow control and establishing reliable links between neighboring nodes. Network Layer: The network layer oversees the process of directing and moving network packets from one network or node to another. This layer identifies the best path for data transfer, conducts logical addressing, and manages the procedures for packet switching and routing. Transport Layer. The transport layer provides reliable end-to-end data transmission between hosts. It segments data received from higher layers into smaller units (segments) and provides mechanisms for error detection, flow control, and retransmission as needed. This layer is responsible for end-to-end communication and can establish either connection-oriented (TCP) or connectionless (UDP) communication. Session Layer. The session layer sets up, controls, and ends communication sessions or links between applications. It facilitates session creation, synchronization, and checkpoint services, enabling applications to initiate and sustain continuous communications. Presentation Layer. The presentation layer focuses on data formatting, encryption, compression, and representation. It ensures that data exchanged between systems is in a compatible format and handles tasks such as data encryption, data compression, and data conversion. Application Layer. The application layer, being the highest layer in the model, directly engages with applications used by the end-user. It provides a means for users to access network services and supports various application protocols, such as HTTP (web), SMTP (email service), FTP (file transmission), and DNS (domain name resolution). The Internet network layer is considered "unstable," as it doesn't assure the complete transmission of data from one end to another [1]. This research may focus on 2.1.3 Network layer.

2.1.2. Transport layer to analyze network traffic. Network packet structure. This research will analyze network packet, so it is essential to take a brief introduction to the packets and learn about what feature is useful to detect attack. IP data packet structure. IP data packet is transferred in network layer.

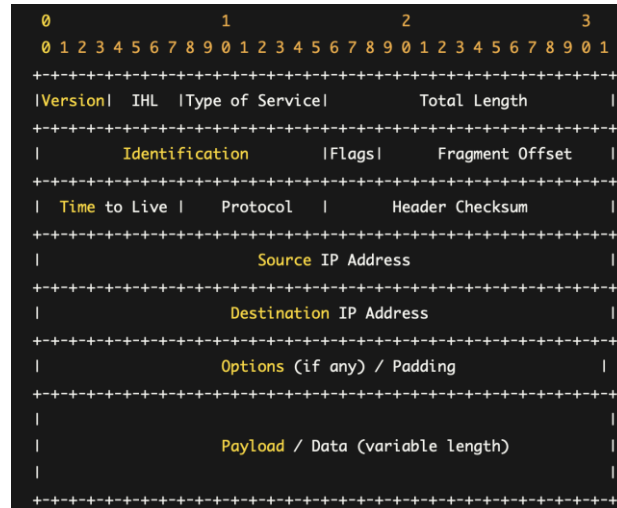


Figure 1. Ip packet structure (Photo/Picture credit: Original).

Figure 1. illustrates the structure of ip data packet. Many features in this structure are easy to understand by its name. This research will just take a brief introduction about two of them. Type of Service is used to prioritize different types of traffic. Payload / Data refers to the actual data being transmitted within the IP packet. TCP packet structure.

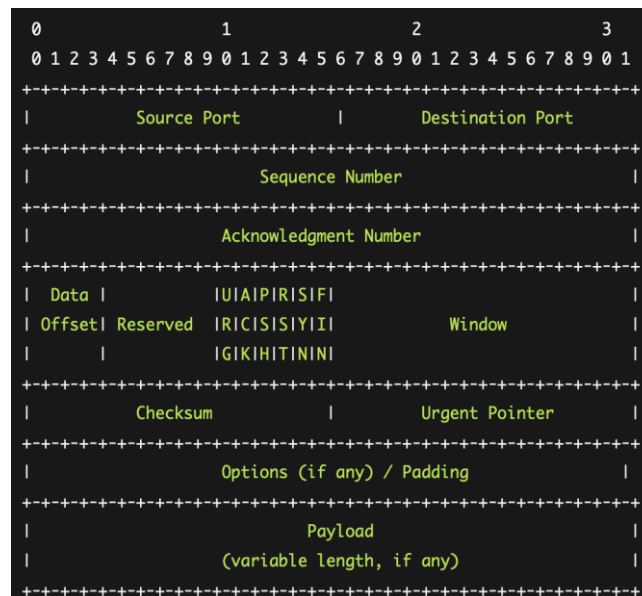


Figure 2. Tcp packet structure (Photo/Picture credit: Original).

Figure 2. shows the structure of tcp packet. It is a simplified representation. An actual TCP packet structure may include additional fields or options depending on specific requirements and options being used. UDP packet structure. UDP is a simple, connectionless protocol that operates independently of TCP. It provides a lightweight, low-overhead transport mechanism for sending datagrams over an IP network. Unlike TCP, UDP does not guarantee reliable delivery or ordered data transmission. It is often

used in applications that prioritize speed and efficiency over reliability, such as real-time streaming, DNS, DHCP, and certain gaming applications. Its packet structure is shown as Figure 3. below.

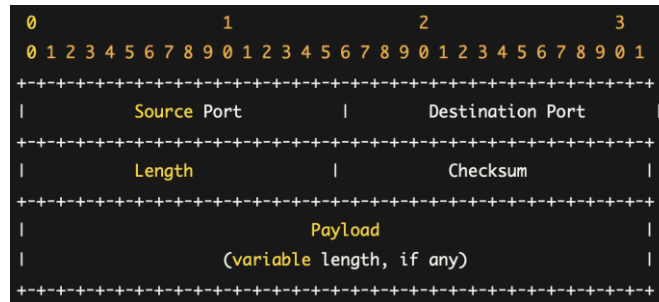


Figure 3. UDP packet structure (Photo/Picture credit: Original).

2.2. Network intrusions

Network intrusions are unauthorized actions performed by malicious entities intending to harm a network, its services, or its users. Network attacks may lead to financial loss, data breaches, repetitional damage, operational disruption and so on. As there is a increasing population surfing on the internet, it becomes crucial for us to get familiar with these attacks so as to mitigate the harm.

2.2.1. Types of attacks. Malware Attacks: Malware, or malicious software, encompasses various harmful programs including viruses, worms, Trojans, and ransomware. These are typically used to cause damage to systems, steal sensitive data, or gain unauthorized access. Phishing Attacks. Phishing is a cyber attack where the attacker masquerades as a trustworthy entity to trick victims into disclosing sensitive information such as usernames, passwords, and credit card details. Denial-of-Service (DoS) and Distributed Denial-of-Service (DDoS) Attacks. A DoS or DDoS attack aims to make a machine or network resource unavailable to its intended users by overwhelming the target's bandwidth or resources. In February 2000, cyber intruders targeted numerous prominent websites, such as Amazon.com, Buy.com, CNN Interactive, and eBay. They perpetrated this by dispatching a multitude of fake packets with the purpose of hampering or discontinuing the services provided [2]. Man-in-the-Middle (MitM) Attacks. In a MitM attack, the perpetrator secretly captures and potentially modifies the exchange of information between two entities without their awareness.

2.2.2. The Activities of Crackers. It's important for us to learn about how crackers to hack into our computer and get control of it. Knowing this, finding a good way to mitigate intrusion may become easier. Software Cracking. One of the most common activities associated with crackers is software cracking, i.e., modifying software to remove or disable features such as copy protection and serial number requirements. This allows the software to be used more widely than intended by the original creators. Creating and Spreading Malware. Crackers often create and spread various forms of malware, including viruses, worms, and trojans, to gain unauthorized access to systems, steal data, or cause damage. Exploiting System Vulnerabilities. Crackers often exploit vulnerabilities in software and hardware to gain unauthorized access to systems. This could include exploiting known vulnerabilities that have not yet been patched, or discovering and exploiting new vulnerabilities.

3. Detect cyber attack

3.1. Intrusion detection systems (IDS)

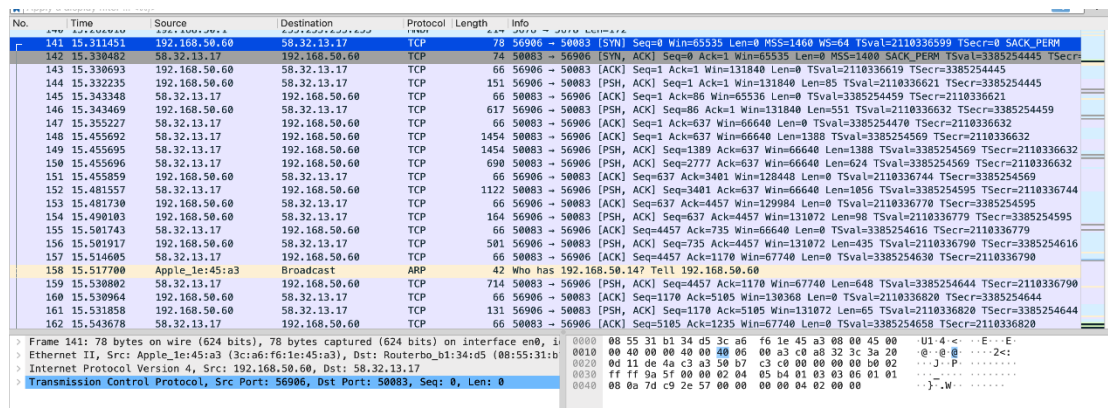
Intrusion Detection Systems serve as integral components of contemporary cybersecurity frameworks. Broadly, IDS are bifurcated into two primary categories: Host-based IDS (HIDS) and Network-based

IDS (NIDS) [3]. HIDS operates intrinsically within a host system, diligently monitoring a myriad of parameters such as processing actions, system logs, and resource consumption to pinpoint deviations or anomalies [4]. In contrast, NIDS casts a wider net, closely observing network interactions for any trace of malevolent activities. Marked shifts or abnormalities in network communication patterns might be telltale signs of impending threats or ongoing cyber-attacks [5]. This study is poised to zoom in on the intricacies and applications of NIDS in the realm of attack detection.

3.2. Detect cyber attack with machine learning

Although it has been a long time since first network intrusion's presence, individuals and organizations still haven't found an excellent way to avoid it. Nowadays, people implement such security measures to fight against these cyber attacks as strong passwords, regular software updates, firewall protection, and user education about potential threats. These methods have certain limits. When encountering an attack that has not been identified and stored in the firewall database, the attack is likely to be misidentified. And in other words, cyber attacks can easily escape from recognition by wrapping themselves to conceal the feature that has been noted in the firewall db. Therefore, a more efficient way to detect network intrusions is desired. Detecting attacks with machine learning can avoid attack escape, because what a detection system base on ml uses to detect attack is not a certain feature of a network traffic packet, but rules about what attacks always are.

3.2.1. Machine learning models. Random forest Algorithms. Since 1995, Ho from Bell Labs proposed the concept of Random Decision Forests, promoting the idea of enhancing prediction accuracy by aggregating multiple classifiers [6]. Then random forest algorithms is widely used in all walks of life. The fundamental principle of the random forests algorithm involves the amalgamation of multiple decision tree classifier models. It essentially merges Bagging and random subspace methods for decision making, and determines the final outcome through a voting decision process [7]. SVM: Support Vector Machine (SVM) is a robust and flexible Machine Learning method applied for classification and regression problems. Essentially, it operates by identifying an optimal hyperplane that most effectively differentiates between various classes within the feature set. It is effective in high dimensional spaces and best suited for problems where there is a clear margin of separation. However, it doesn't perform well when the classes overlap. It is also not suitable for large datasets because of its high training time [8]. This research will compare the classification result of these two models to find a possible solution to cyber attack detection.



No.	Time	Source	Destination	Protocol	Length	Info
141	15.311051	192.168.50.60	58.32.13.17	TCP	78	56906 → 50083 [SYN] Seq=0 Win=65535 Len=0 MSS=1460 WS=64 TSval=2110336899 TSecr=0 SACK_PERM
142	15.339482	58.32.13.17	192.168.50.60	TCP	74	50083 → 56906 [SYN, ACK] Seq=0 Ack=1 Win=65535 Len=0 WS=1400 SACK_PERM TSval=3385254445 TSecr=
143	15.336693	192.168.50.60	58.32.13.17	TCP	66	56906 → 50083 [ACK] Seq=1 Ack=1 Win=131840 Len=0 TSval=2110336619 TSecr=3385254445
144	15.332235	192.168.50.60	58.32.13.17	TCP	151	56906 → 50083 [PSH, ACK] Seq=1 Ack=1 Win=131840 Len=85 TSval=2110336621 TSecr=3385254445
145	15.343348	58.32.13.17	192.168.50.60	TCP	66	50083 → 56906 [ACK] Seq=1 Ack=86 Win=65536 Len=0 TSval=3385254459 TSecr=2110336621
146	15.343469	192.168.50.60	58.32.13.17	TCP	617	56906 → 50083 [PSH, ACK] Seq=86 Ack=1 Win=131840 Len=551 TSval=2110336632 TSecr=3385254459
147	15.355227	58.32.13.17	192.168.50.60	TCP	66	50083 → 56906 [ACK] Seq=1 Ack=637 Win=66640 Len=0 TSval=3385254478 TSecr=2110336632
148	15.455692	58.32.13.17	192.168.50.60	TCP	1454	50083 → 56906 [ACK] Seq=1 Ack=637 Win=66640 Len=1388 TSval=3385254569 TSecr=2110336632
149	15.455695	58.32.13.17	192.168.50.60	TCP	1454	50083 → 56906 [PSH, ACK] Seq=1389 Ack=637 Win=66640 Len=1388 TSval=3385254569 TSecr=2110336632
150	15.455696	58.32.13.17	192.168.50.60	TCP	698	50083 → 56906 [PSH, ACK] Seq=2777 Ack=637 Win=66640 Len=624 TSval=3385254569 TSecr=2110336632
151	15.455859	192.168.50.60	58.32.13.17	TCP	66	56906 → 50083 [ACK] Seq=637 Ack=3401 Win=128448 Len=0 TSval=2110336744 TSecr=3385254569
152	15.481557	58.32.13.17	192.168.50.60	TCP	1122	50083 → 56906 [PSH, ACK] Seq=3401 Ack=637 Win=66640 Len=1056 TSval=3385254595 TSecr=2110336744
153	15.481730	192.168.50.60	58.32.13.17	TCP	66	56906 → 50083 [ACK] Seq=637 Ack=4457 Win=129984 Len=0 TSval=2110336770 TSecr=3385254595
154	15.490103	192.168.50.60	58.32.13.17	TCP	164	56906 → 50083 [PSH, ACK] Seq=637 Ack=4457 Win=131072 Len=90 TSval=2110336779 TSecr=3385254595
155	15.501743	58.32.13.17	192.168.50.60	TCP	66	50083 → 56906 [ACK] Seq=4457 Ack=735 Win=66640 Len=0 TSval=3385254616 TSecr=2110336779
156	15.501917	192.168.50.60	58.32.13.17	TCP	501	56906 → 50083 [PSH, ACK] Seq=735 Ack=4457 Win=131072 Len=435 TSval=2110336790 TSecr=3385254616
157	15.514605	58.32.13.17	192.168.50.60	TCP	66	50083 → 56906 [ACK] Seq=4457 Ack=1170 Win=67740 Len=0 TSval=3385254630 TSecr=2110336790
158	15.517700	Apple_1e:45:a3	Broadcast	ARP	42	Who has 192.168.50.14? Tell 192.168.50.60
159	15.536802	58.32.13.17	192.168.50.60	TCP	714	50083 → 56906 [PSH, ACK] Seq=4457 Ack=1170 Win=67740 Len=648 TSval=3385254644 TSecr=2110336790
160	15.536964	192.168.50.60	58.32.13.17	TCP	66	56906 → 50083 [ACK] Seq=1170 Ack=5105 Win=130368 Len=0 TSval=2110336820 TSecr=3385254644
161	15.531858	192.168.50.60	58.32.13.17	TCP	131	56906 → 50083 [PSH, ACK] Seq=1170 Ack=5105 Win=131072 Len=65 TSval=2110336820 TSecr=3385254644
162	15.543678	58.32.13.17	192.168.50.60	TCP	66	50083 → 56906 [ACK] Seq=5105 Ack=1235 Win=67740 Len=0 TSval=3385254658 TSecr=2110336820

> Frame 141: 78 bytes on wire (624 bits), 78 bytes captured (624 bits) on interface en0, i
> Ethernet II, Src: Apple_1e:45:a3 (3c:a6:f6:1e:45:a3), Dst: Routerbo_b1:34:d5 (08:55:31:b
> Internet Protocol Version 4, Src: 192.168.50.60, Dst: 58.32.13.17
> Transmission Control Protocol, Src Port: 56906, Dst Port: 50083, Seq: 0, Len: 0

0000 08 55 31 b1 34 d5 3c a6 f6 1e 45 a3 08 00 45 00 U1:4<...E...<
0010 00 00 00 00 00 00 00 00 a3 c0 a8 32 3c 3a 20 @...P...<2<
0020 00 11 0e 4a c3 a3 58 b7 c3 c0 00 00 00 00 02 ...J...P...<2<
0030 ff ff 5a 5f 00 00 02 04 05 b4 01 03 03 06 01 01 ...W...<2<
0040 08 0a 7d c9 2e 57 00 00 00 00 04 02 00 00 ...<2<

Figure 4. Using Wireshark to capture packets (Photo/Picture credit: Original).

3.2.2. Detecting attacks with Machine Learning. The study involved the collection of a week's worth of network traffic packets to obtain raw data. As demonstrated in Figure 4, tools such as Wireshark can be instrumental in capturing and analyzing this traffic. Subsequently, data points deemed irrelevant, such as timespan and IP address, were discarded. Categorical data were then converted into numerical format

utilizing a one-hot encoder. Infinite and null values in the dataset were removed, a decision justified by the abundance of usable data relative to these flawed entries [9]. The initial objective was to utilize machine learning to differentiate between DDoS attacks and regular network activities. Classification was attempted using models like DummyClassifier, SVMLinearSVC, RandomForest, and KNN. As indicated in Figure 6, the Dummy model, with a score of 0.5664569580435072, was found inadequate. In contrast, both LinearSVC (with a score of 0.9156458365638084, shown in Figure 5) and KNN (scoring 0.9985379793101921, depicted in Figure 7) exhibited commendable performance.

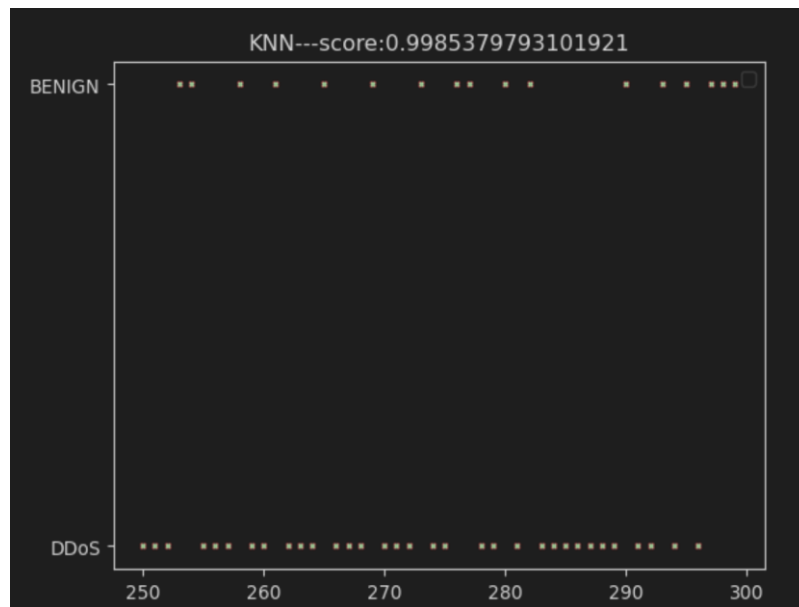


Figure 5. Result of KNN classifier with 50 pieces of data (Photo/Picture credit: Original).



Figure 6. Result of Dummy classifier with 50 pieces of data (Photo/Picture credit: Original).

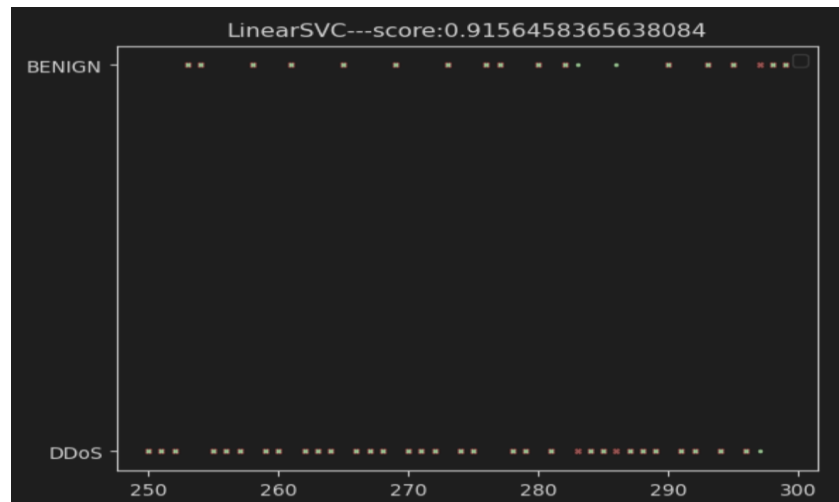


Figure 7. Result of linear svc classifier with 50 pieces of data (Photo/Picture credit: Original).

RandomForest owns the highest score (0.9998670890281993) Figure 8.

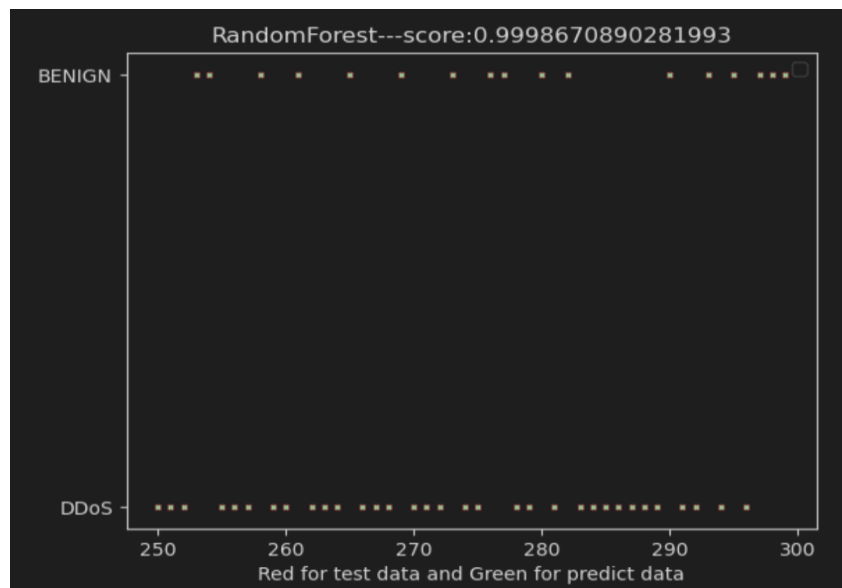


Figure 8. Result of RandomForest Classifier with 50 pieces of data (Photo/Picture credit: Original).

It seemed that RandomForestClassifier can do the job [10]. Then this research tried to use this model to distinguish between different types of attacks, attack types are shown in Figure 9.

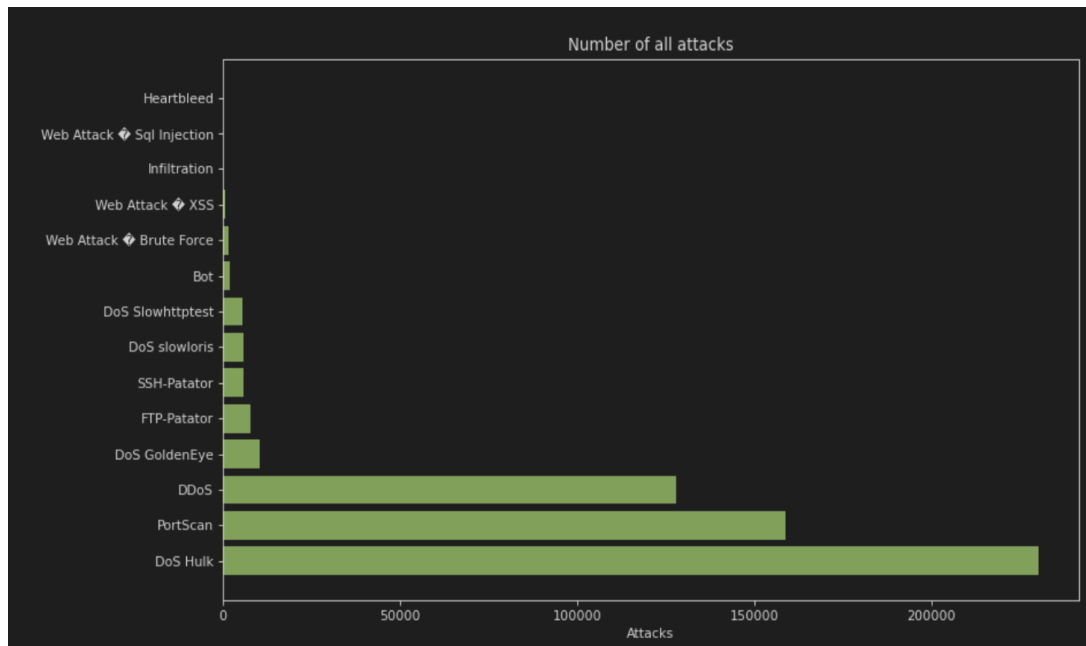


Figure 9. Numbers of all attacks in multiple dataset (Photo/Picture credit: Original).

After training the model, it can be found out that the result was not so good. So some improvement was concerned. A pipeline was used to deal with the data by filtering out some poorly related features. After retraining, the score increased a bit, and the accuracy improved to 0.99853(Figure 10.). A score of 0.998539 turned out that machine learning can distinguish cyber attack from normal network traffic, and it is really an efficient way.

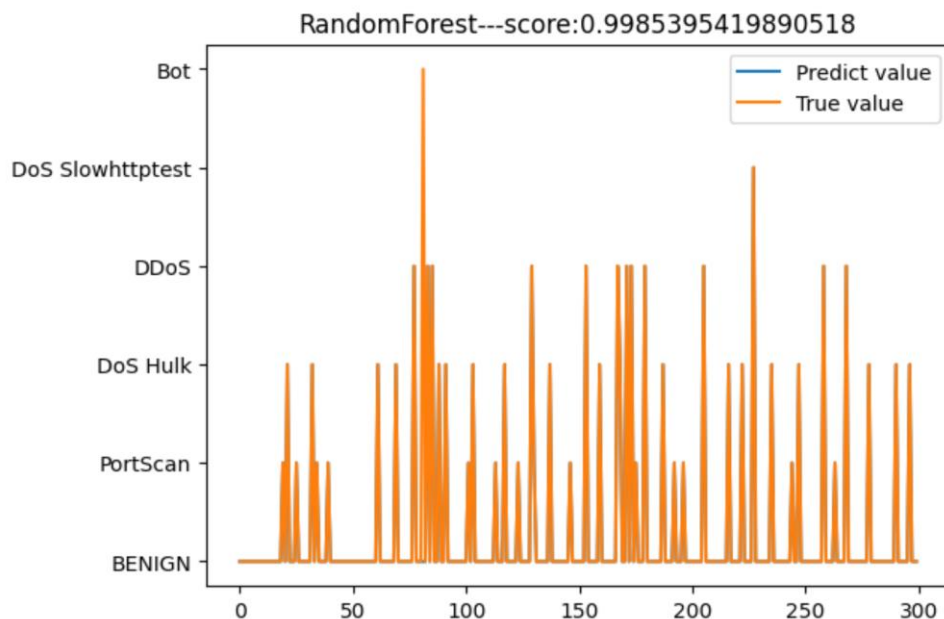


Figure 10. Result of Random Forest Classifier with multiple dataset (Photo/Picture credit: Original).

4. Challenges

While the integration of machine learning promises to revolutionize cybersecurity frameworks, it's not without its challenges. One paramount concern is the quality and availability of training data. The

efficacy of machine learning models hinges on the caliber of their training data. To adeptly detect cyber-attacks, this data must be extensive, varied, and mirror real-world situations. Yet, sourcing high-quality, appropriately labeled data often proves to be an uphill battle. Another intricacy lies in striking the right balance between sensitivity and specificity. Ensuring a model accurately pinpoints attacks, while simultaneously minimizing the pitfalls of false positives (erroneously identifying normal activities as malicious) and false negatives (overlooking genuine threats), remains a daunting task. Additionally, the sheer volume and complexity of cybersecurity data, rife with high-dimensional attributes, demand scalable machine learning algorithms, presenting another layer of challenge. Lastly, the issue of interpretability looms large. Many advanced machine learning models, notably neural networks, operate as enigmatic "black boxes." Such opacity in decision-making can be a stumbling block in cybersecurity, where discerning the rationale behind a specific alert can be instrumental.

5. Improvements

This study primarily centered on Network Intrusion Detection Systems (NIDs), delving deep into their intricacies and capabilities. However, determining a clear victor among various intrusion detection paradigms remains challenging. Each system, with its unique set of strengths and vulnerabilities, offers distinct advantages in specific scenarios. Given this, a more holistic approach might be the fusion of multiple systems – crafting a hybrid intrusion detection system (hybrid-IDs). Such a hybrid model would ideally amalgamate the robustness of multiple systems, offering a more comprehensive solution that capitalizes on the collective strengths while mitigating individual weaknesses. Hybrid models are often seen as the next frontier in intrusion detection, promising enhanced accuracy, adaptability, and scalability. They have the potential to provide a more nuanced view of threats, thereby improving the overall security posture. Future research endeavors can pivot towards the in-depth exploration of these hybrid systems. By dissecting the merits of HIDs and understanding their underlying mechanisms, there's potential to harness a new wave of cybersecurity solutions. This deepened perspective could lead to the design of systems that not only react to known threats but proactively anticipate and thwart new, emerging challenges in the ever-evolving cyber landscape. In essence, the journey towards a fortified cyber defense might very well lie in the integration and synergy of various intrusion detection paradigms.

6. Conclusion

The high result score of 0.99853 from the Random Forest model validates the initial hypothesis: machine learning offers immense potential for fortifying cybersecurity measures. With the integration of machine learning algorithms, cybersecurity infrastructures can evolve to be more adaptable, proficient in forecasting, and addressing threats in real-time. As the digital era unfolds, Machine Learning is poised to assume an increasingly pivotal role in safeguarding information systems against the continuously changing threat landscape. For future advancements, the development of even more resilient algorithms is essential to address the fluid nature of cyber threats. There is also a heightened need to gather and refine high-quality datasets to train these predictive models. Despite inherent challenges, the magnitude of Machine Learning's potential in cyber-attack detection remains immense. Equipped with the capacity to discern from data, adjust to evolving situations, and foresee forthcoming threats, Machine Learning stands on the brink of transforming the cybersecurity domain.

Merging machine learning with other nascent technologies such as Big Data and the Internet of Things paves the way for devising even more robust and insightful cybersecurity frameworks. In summation, while there remains ample scope for enhancement, a persuasive argument exists for the integration of Machine Learning in the domain of cyber-attack detection. By honing these methodologies and navigating the pertinent challenges, a future emerges where digital domains are safeguarded by astute, flexible, and forward-thinking defense mechanisms.

References

- [1] Marin, G. A. (2005). Network security basics. *IEEE Security & Privacy*, 3(6), 68-72.

- [2] Bonisteel, S. (2003). Yahoo DoS Attack Was Sophisticated. ComputerUser.com. Retrieved from www.computeruser.com/news/00/02/14/news1.html
- [3] Bace, R. G., & Mell, P. (2001). Intrusion Detection Systems. NIST Special Publication.
- [4] Bsed, G., Hanumanthappa. (2022). Dual Mode Host-Based and Cloud-Based Smartphone Intrusion Detection System. 2022 International Conference on Futuristic Technologies (INCOFT), Belgaum, India, 1-5.
- [5] Ho, T. K. (1995). Random decision forests. J. Economic Problem, 8, 278-282.
- [6] Lan, H., & Pan, Y. (2019). A Crowdsourcing Quality Prediction Model Based on Random Forests. 2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS), Beijing, China, 315-319.
- [7] Hosseini, M. M., & Parvania, M. (2021). Artificial intelligence for resilience enhancement of power distribution systems. The Electricity Journal, 34(1), 106880.
- [8] Alsuwian, T., Shahid Butt, A., & Amin, A. A. (2022). Smart Grid Cyber Security Enhancement: Challenges and Solutions—A Review. Sustainability, 14(21), 14226.
- [9] Liu, Y., Wang, T., Zhang, S., et al. (2020). Artificial intelligence aware and security-enhanced traceback technique in mobile edge computing. Computer Communications, 161, 375-386.
- [10] Mosteanu, N. R. (2020). Artificial Intelligence and Cyber Security—A Shield against Cyberattack as a Risk Business Management Tool—Case of European Countries. Quality-Access to Success, 21(175).