2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS) | 979-8-3503-8749-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICTACS62700.2024.10840935

Guarding Against IoT Threats: An Analysis of Intrusion Detection with the Kitsune Attack Dataset

Anshika Sharma[1], Himanshi Babbar[2]
[1,2]*Department of Computer Science,*
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

***Abstract*—The extensive adoption of Internet of Things (IoT) devices has resulted in previously unheard-of levels of connectedness and ease, but it has also created new cybersecurity challenges. Hackers are turning an increasing number of IoT devices into their targets. They take advantage of security flaws to execute a variety of attacks, from botnet malware attacks and distributed denial-of-service (DDoS) attacks to surveillance, reconnaissance (recon.) and man-in-the-middle(MITM) attacks. Innovative detection techniques are required because traditional security systems typically cannot keep up with the dynamic and varied nature of IoT environments. Using the Kitsune attack dataset, a large collection of network traffic captures that have been carefully chosen for use in network intrusion detection system (NIDS) research, the effectiveness of machine learning (ML)-based techniques have been examined for identifying IoT attacks in this study. Utilising characteristics taken from network traffic data, such as communication patterns, payload attributes, and packet headers, investigate how well-supervised learning algorithms like Support Vector Machines (SVM), Random Forest (RF), Logistic Regression (LR), and K-Nearest Neighbours (KNN) perform. Each algorithm's detection performance has been assessed using metrics including accuracy, precision, recall, and F1-score. Their respective advantages and disadvantages in terms of scalability, computational efficiency, and detection accuracy have also been compared. Additionally, the effects of feature selection and ensemble learning approaches have been evaluated on detection performance, offering recommendations for constructing durable and dependable IoT-IDS. The results show that, at 98.99%, the LR model has the highest accuracy. On the other side, the accuracy rates of the KNN, SVM and RF models are 79.42%, 92.75%, and 97.98%, respectively.**

Keywords: Machine Learning, Network Intrusion Detection System, IoT Attacks, Security, Kitsune Attack Dataset

I. Introduction

An era of remarkable interaction and smart automation across multiple industries has been brought about by the swift growth of IoT devices, offering the promise of improved user experiences, higher efficiency, and innovative applications [1]. The global deployment of millions of IoT devices has created a massive, networked system that makes data sharing and communication easy [2]. Though the increased attack surface provides malicious parties with new avenues to exploit weaknesses and endangers the reliability and safety of these devices, interconnectivity also brings with it previously unheard-of cybersecurity issues [3]. The increasing complexity and variety of cyber threats aimed at IoT networks and devices is one of the major obstacles facing the IoT industry [4]. From classic network-based intrusions like MITM, and recon to more complex and orchestrated operations like DDoS attacks and the spread of botnet malware, these threats cover a broad range [5]. The distinctive features and limits of IoT environments, such as resource shortages, a variety of transmission protocols, and the enormous scope of deployment, are difficult for traditional security mechanisms, which are frequently created for conventional computing systems, to adjust to [6].

The application of ML techniques to NIDS for the IoT has emerged as a viable response to this dynamic threat landscape [7], [8]. ML offers a continually evolving way to strengthen security in IoT networks because of its capacity to identify intricate patterns and abnormalities within massive information [9]. The Kitsune attack dataset is a fundamental resource for examining and developing ML-based solutions for IoT security [10]. It was created specifically for research purposes in NIDS. The Kitsune dataset contains a wide range of network traffic transcriptions that include both malicious and benign activity [11]. The dataset provides a rich basis for training and as- sessing machine learning models since information including headers of packets, payload features etc. are extracted from this network traffic data. To identify and detect different types of IoT attacks, this study explores the Kitsune dataset using a range of supervised learning methods, such as SVM, RF, LR, and KNN [12], [13].

This study aims to evaluate how well ML approaches can strengthen the safety measures of IoT environments [14]. Determine the advantages and disadvantages of each algorithm in recognising distinct attack types by doing a thorough examination of detection performance, including measures like accuracy, precision, recall, and F1-score [15]. Examine the effects of critical elements such as feature

selection, and ensemble learning techniques to provide practical insights that will improve the resilience and practical applicability of IoT- NIDS.

Afterwards, the remaining components of the paper are classified into the categories of various groups mentioned below: Section II contains a comprehensive summary of the pertinent previous research on the use of ML techniques to detect threats in the IoT industry. A detailed explanation of the dataset and ML approaches used is provided in section III. Multiple ML algorithms are combined in this strategy, which is crucial for studying and analysing any kind of data. A few instances of the different kinds of statistical procedures that are used are feature selection, data preprocessing, and the collection of previously gathered data. Its precision, recall, accuracy, and F1-score evaluations are examined in detail in Section IV. For this evaluation and analysis, numerous ML algorithms have been provided. The written paper's conclusion is provided in V.

II. Related Work

Alabdulatif et al. [1] provide a thorough analysis aimed at determining the best ML model or models for Kitsune. Many ML algorithms have been chosen for this inquiry. According to their research, several tree algorithm variations, including Simple Tree, Medium Tree, RUSBoosted, Coarse Tree, and Bagged Tree, have demonstrated comparable efficacy albeit with somewhat different levels of inefficiency. Finally, the competition was won by Coarse Tree, which is the optimal method for detecting malware attacks using the Mirai botnet.

Mirsky et al. [2] present Kitsune, a plug-and-play NIDS sys- tem that can quickly and effectively identify online threats on a local network without human intervention. An ensemble of neural networks known as autoencoders are used by Kitsune's core algorithm (KitNET) to collectively distinguish between typical and anomalous traffic patterns. An attribute extraction framework that effectively monitors each network channel's trends underpins KitNET. Our tests demonstrate that, even on a Raspberry PI, Kitsune can identify a wide range of threats with a performance similar to offline anomaly detectors. This proves that Kitsune is a useful and affordable NIDS.

Pranto et al. [3] have shown that ML approaches may be utilised for categorising incoming data from networks into two categories: normal and abnormal. The NSLKDD dataset has been used to assess several classifiers. Tests were carried out to assess the KNN, DT, NB, LR, and RF. A basic feature extraction technique has reduced both the size of the data set and the level of computational complexity.

Shahid et al. [4] give a short overview of the research articles that have been published recently about the use of both conventional and state-of-the-art ML to identify XSS. This paper aims to assist researchers in selecting a course of action when creating novel approaches for their investigations and to encourage them to focus on the relationship between internet safety and ML. It will discuss various ML techniques that have been employed, understand the main findings from each study, highlight their positive outcomes, and highlight any persistent negative aspects.

The study proposed by Raza et al. [5] aims to precisely and promptly detect network attacks to prevent detrimental losses. They used the CICIDS 2017 dataset to create state-of- the-art AI-based ML approaches. They recommend a unique technique known as Class Probability Random Forest (CPRF) to enhance the effectiveness of network attack detection. Using the proposed CPRF method, they generated a distinct feature set. The CPRF approach anticipates class probabilities utilising a network attack dataset, which are then used as characteristics in the creation of applied machine learning approaches.

III. Materials and Methods

A. Dataset

Investigations and development in the field of cybersecurity and NIDS can benefit from the publicly available Kitsune network assault dataset. The dataset was designated after the study team's ensemble-based method for detecting in- trusions in online networks. Gathered from both real-world and simulated network environments, the Kitsune network attack dataset includes traffic captures from several sources. It includes both legitimate and malicious network traffic, such as that which is generated by port scans, malware infections, DDoS attacks, reconnaissance operations, and virus infections with 16 features as shown in Table I. Providing realistic statistics on network traffic and annotated real-world labels, the Kitsune network attack dataset is a great tool for cybersecurity academics and practitioners. Researchers can use this information to build and test intrusion detection systems that can detect and handle different kinds of network threats instantly. These four network attacks i.e. Recon., MITM, DDOS, and Botnet Malware were captured via either an IP- based corporate monitoring system or an internet connection filled with IoT devices, as described below.

1) Reconnaissance

Gathering information about the in- tended network or system is the first step in a cyber attack. As part of their reconnaissance, attackers use a variety of methods to gain insight into the target organization's systems, services, and security measures, as well as to find any loopholes in their defences.

2) Man-in-the-Middle

Intercepting and maybe altering ex-changes between two individuals with their permission or consent is a kind of cyber assault. Attackers using an MITM technique can eavesdrop on conversations, alter data in transit, or even assume the identities of the people involved in a conversation.

3) Distributed Denial of Service

Computers, servers, routers, and IoT devices are common targets for DDoS attacks because they allow attackers to take control of a network of vulnerable devices. The perpetrator then directs the bot-net to launch a coordinated deluge of traffic at the victim, overwhelming its application layer, computing resources, or network capacity.

4) Botnet Malware

Attackers can gain remote control over infected devices, or ”bots,” after they join the botnet. Malicious botnets can distribute more malware, steal information, mine cryptocurrencies, launch DDoS attacks, and run spam email campaigns, among other things.

B. Data Preprocessing

The cleaning, transformation, and preparation of raw data for additional analysis or modelling constitute data preprocess- ing, an essential stage in the data analysis process. Data pre-processing entails the following steps, which are summarised here:

Table I: Features of IoT Botnet Traffic Dataset

Feature	Feature
Src IP Address	HTTP Status Code
Dest IP Address	Connection Frequency
Src Port	Connection Duration
Dest Port	Pkt Payload
Protocol	Connection Patterns
Pkt Size	Payload Length
Pkt Timing	Payload Content
HTTP Method	Attack Label

- Imputing missing values, deleting duplicate records, and fixing mistakes are all part of data cleaning, which also includes finding and addressing incorrect or missing data.
- Improving the data’s format so it can be used for analysis or modelling is part of data transformation. Standardi- sation, growing, encapsulation of categorical variables, and feature creation to generate novel characteristics or extract more data from current ones are common transformation techniques.
- When doing analysis or modelling, feature selection is the process of identifying which features or variables are most important and eliminating those that aren’t. Boost model performance, and lower dimensionality with feature selection strategies.
- Data integration is the process of gathering and standar- dising data formats for analysis from many sources or datasets. Data compatibility and resolving shortcomings or discrepancies between data may be necessary steps in the data integration process.
- Reducing the quantity or complexity of a dataset while keeping all of the necessary information is known as data reduction. To reduce data, one can utilise methods like sampling, feature extraction, or principal component analysis (PCA).
- To make sure that numerical aspects are consistent and comparable, data must be normalised, which means that they are scaled to a standardised range or distribution. Examples of common normalisation approaches are z-score normalisation and min-max scaling.

C. Machine Learning Approaches

1) K-Nearest Neighbour

To determine which class to assign to the test point, it locates the K training set data points that are closest to the test data point. A lazy learner, KNN retains the entirety of the training dataset instead of acquiring explicit models. When dealing with nonlinear decision bound- aries or complex data distribution, it proves to be efficacious in classification problems. Large datasets may make it extremely expensive, and the distance analysed and the K value that is selected can have a significant impact.

2) Support Vector Machine

The goal of SVM is to identify the best hyperplane with the largest margin that divides each of the points of several classes. It is efficient with datasets that have a distinct margin of separation and performs well in spaces with high dimensions. SVM is capable of handling both nonlinear as well as linear classification by utilising a variety of kernel-based functions. SVM is appropriate for datasets with vast feature spaces since it only employs a selected group of training points to build the decision boundary, which saves memory.

3) Random Forest

During training, it constructs several DTs and then uses a technique known as bagging to integrate the predictions of each tree. To lessen overfitting and enhance generalisation performance, each DT is trained using a random selection of features and training data. RF is resilient to noise and anomalies in the data and can perform jobs including regression and classification. It has a reputation for being very accurate, scalable, and capable of handling datasets with many dimensions.

4) Logistics Regression

It uses the sigmoid function to estimate the likelihood that an input corresponds to a specific class. Because of its interpretability and processing efficiency, LR is a good fit for real-time applications and huge datasets. It is especially helpful when there is roughly a linear connection between the independent factors and the dependent variable’s log odds.

IV. Results and Discussion

A. Accuracy

The general correctness of an identification model is gauged using a popular evaluation metric called accuracy. It is com- puted as the ratio of all instances in the dataset to the number of accurately predicted instances i.e. true positives (TP) and true negatives (TN) combined.

Table II: Analysing and Comparing the Performance Metrics Using Machine Learning

ML Models	Parameters(%)			
	Accuracy	Precision	Recall	F1-Score
KNN	79.42	81.85	82.67	82.25
SVM	92.75	94.79	91.84	93.29
RF	97.98	98.12	98.41	98.26
LR	98.99	98.97	97.99	98.47

Table II illustrates how the accuracy, recall, precision, and F1-score are used to identify and predict attacks in the IoT space using the Kitsune attack dataset based on NIDS. Ad- ditionally, this table shows how the ML models KNN, SVM, RF, and LR are evaluated for performance. Fig 1 illustrates the comparison of the accuracy of different ML techniques, including KNN, SVM, RF, and LR. The results show that, at 98.99%, the LR model has the highest accuracy. On the other side, the accuracy rates of the KNN, SVM and RF models are 79.42%, 92.75%, and 97.98%, respectively.

B. Precision

The percentage of properly anticipated positive instances, or TP, among all instances projected as positive, or TP and FP, is measured by a performance indicator called precision. It is especially helpful when the expense of FP is significant because it concentrates on the reliability of positive predictions.

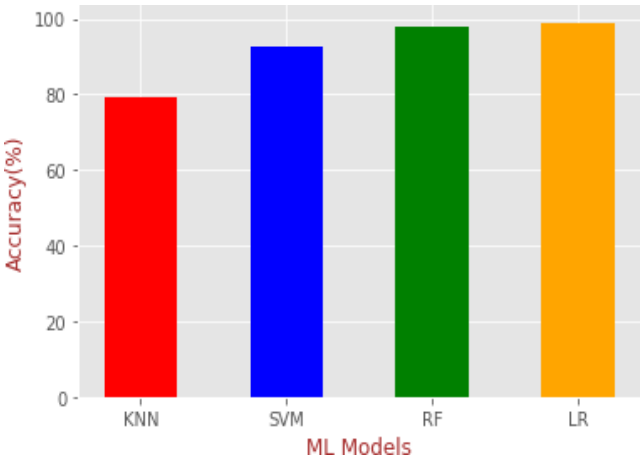

Fig. 1: An Analysis of Machine Learning Algorithms' Detection Accuracy for IoT Attacks

Fig 2 illustrates the comparison of the precision of different ML techniques, including KNN, SVM, RF, and LR. The results show that, at 98.97%, the LR model has the highest precision. On the other side, the precision rates of the KNN, SVM and RF models are 81.85%, 94.79%, and 98.12%, respectively.

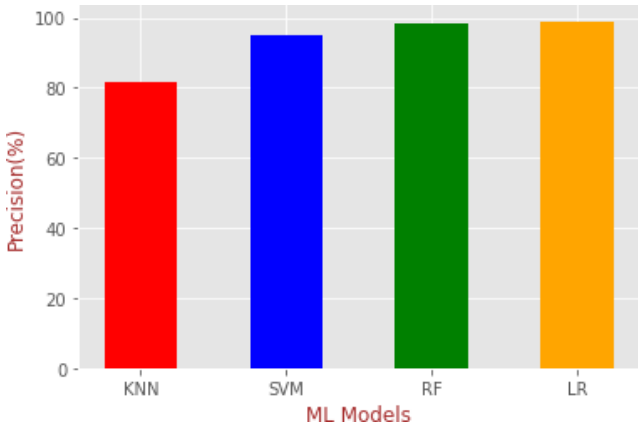

Fig. 2: An Analysis of Machine Learning Algorithms' Detection Precision for IoT Attacks

C. Recall

The recall of a dataset is defined as the amount of TP or positively predicted cases, amongst all real positive instances. It concentrates on the model's capacity to include all positive cases, which is crucial in situations where the absence of good instances could have detrimental effects.

Fig 3 illustrates the comparison of the recall of different ML techniques, including KNN, SVM, RF, and LR. The results show that, at 98.12%, the RF model has the highest recall.

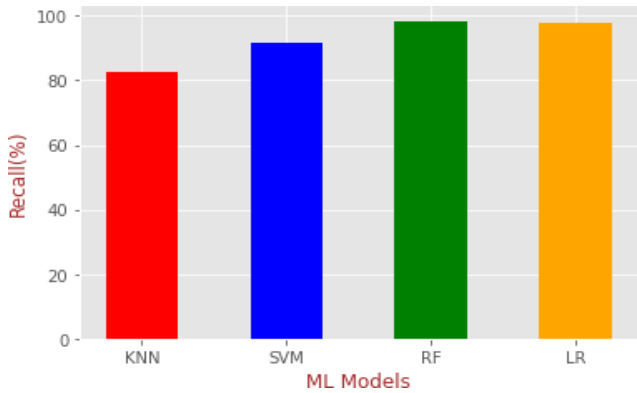

Fig. 3: An Analysis of Machine Learning Algorithms' Detection Recall for IoT Attacks

On the other side, the recall rates of the KNN, SVM and LR models are 82.67%, 91.84%, and 97.99%, respectively.

D. F1-score

The F1-score is an integrated metric that provides a single assessment of the accuracy of a model by balancing recall and precision. It offers a fair evaluation of a model's capacity to accurately make equally positive and negative predictions. It is the harmonic mean of accuracy and recall. Higher values in the F1-score range from 0 to 1, signifying superior performance.

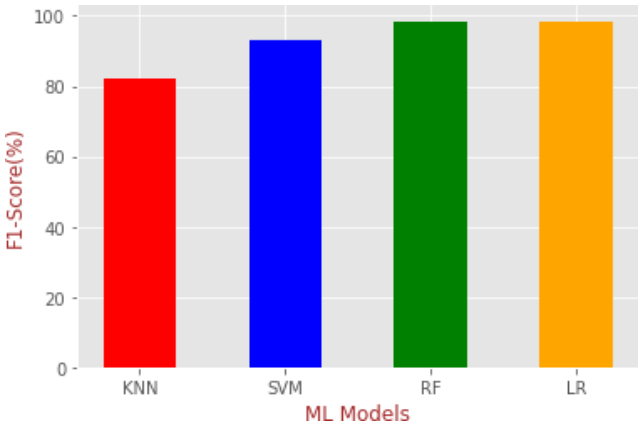

Fig. 4: An Analysis of Machine Learning Algorithms' Detection F1-score for IoT Attacks

Fig 4 illustrates the comparison of the F1-score of different ML techniques, including KNN, SVM, RF, and LR. The results show that, at 98.47%, the LR model has the highest F1- score. On the other side, the F1-score rates of the KNN, SVM and RF models are 82.25%, 93.29%, and 98.26%, respectively.

V. Conclusion

Based on the Kitsune attack dataset, the incorporation of machine learning techniques into network intrusion detection for IoT contexts offers a promising path for improving cy-bersecurity in the face of changing threats. The effectiveness of machine learning-based techniques in identifying a variety of IoT attacks, such as recon, MITM, DDoS, and botnet malware attacks, has been demonstrated by the investiga- tion of supervised learning approaches like SVM, RF, LR, and KNN. The potential to strengthen the security stance of IoT ecosystems, secure sensitive data, and maintain the dependability and trustworthiness of IoT installations exists with the ongoing development of ML-based NIDS. Organ- isations may reduce risks and strengthen the durability of connected devices against attackers by proactively identifying and mitigating emerging cyber threats by utilising the insights gathered from the research. The study is a step in the right direction towards addressing the cybersecurity issues that IoT environments provide. One may create a more resilient and secure IoT environment where the advantages of automation and connectivity can be realised without sacrificing security by utilising ML and datasets like Kitsune. The results show that, at 98.99%, the LR model has the highest accuracy. On the other side, the accuracy rates of the KNN, SVM and RF models are 79.42%, 92.75%, and 97.98%, respectively.

References

[1] A. Alabdulatif, S. S. H. Rizvi, and M. A. Hashmani, "Optimal machine learning models for kitsune to detect mirai botnet malware attack," *Journal of Hunan University Natural Sciences*, vol. 48, no. 6, 2021.

[2] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: an ensemble of autoencoders for online network intrusion detection," *arXiv preprint arXiv:1802.09089*, 2018.

[3] M. B. Pranto, M. H. A. Ratul, M. M. Rahman, I. J. Diya, and Z.-B. Zahir, "Performance of machine learning techniques in anomaly detection with basic feature selection strategy-a network intrusion detection system," *J. Adv. Inf. Technol*, vol. 13, no. 1, 2022.

[4] M. Shahid, "Machine learning for detection and mitigation of web vulnerabilities and web attacks," *arXiv preprint arXiv:2304.14451*, 2023.

[5] A. Raza, K. Munir, M. S. Almutairi, and R. Sehar, "Novel class probability features for optimizing network attack detection with machine learning," *IEEE Access*, 2023.

[6] F. Alwahedi, A. Aldhaheri, M. A. Ferrag, A. Battah, and N. Tihanyi, "Machine learning techniques for iot security: Current research and future vision with generative ai and large language models," *Internet of Things and Cyber-Physical Systems*, 2024.

[7] A. Saini, K. Guleria, and S. Sharma, "Machine learning approaches for an automatic email spam detection," in *2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)*. IEEE, 2023, pp. 1–5.

[8] M. Azeem, D. Khan, S. Iftikhar, S. Bawazeer, and M. Alzahrani, "Analyzing and comparing the effectiveness of malware detection: A study of machine learning approaches," *Heliyon*, vol. 10, no. 1, 2024.

[9] A. Sharma and H. Babbar, "Detecting botnet attacks in iot healthcare systems through iot technology," in *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelli- gence and Telecommunication Engineering (RMKMATE)*. IEEE, 2023, pp. 1–6.

[10] P. Shourie, N. Kumar, R. K. Kaushal, and S. Verma, "Iot-based hybrid fitness monitoring device," in *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE, 2023, pp. 1527–1532.

[11] V. Tanwar and K. Ramkumar, "A survey on the role of reverse engineer- ing in security attacks," in *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*. IEEE, 2023, pp. 1–6.

[12] A. V. Turukmane and R. Devendiran, "M-multisvm: An efficient feature selection assisted network intrusion detection system using machine learning," *Computers & Security*, vol. 137, p. 103587, 2024.

[13] L. Cui, S. Yang, F. Chen, Z. Ming, N. Lu, and J. Qin, "A survey on application of machine learning for internet of things," *International Journal of Machine Learning and Cybernetics*, vol. 9, pp. 1399–1417, 2018.

[14] M. Bagaa, T. Taleb, J. B. Bernabe, and A. Skarmeta, "A machine learning security framework for iot systems," *IEEE Access*, vol. 8, pp. 114 066– 114 077, 2020.

[15] S. Vashisht, B. Sharma, and S. Lamba, "Using support vector machine and generative adversarial network for multi-classification of pneumonia disease," in *2023 4th International Conference for Emerging Technology (INCET)*. IEEE, 2023, pp. 1–6.