iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

NETWORK-BASED INTRUSION DETECTION SYSTEM USING MACHINE LEARNING

Dr. Sankar Sarma Kvssrs
Associate Professor in Artificial Intelligence and Machine Learning,
J.B. Institute of Engineering and Technology, Hyderabad, Telangana, India
V Sai Chandra Kousik
G Shashank
N Suryakanth Reddy
A Vijay Kumar
UG Student, Department of Artificial Intelligence and Machine Learning,
J.B. Institute of Engineering and Technology, Hyderabad, Telangana, India

ABSTRACT
Strong network security solutions are more important than ever in a time when cyber threats are ever more complex. Because cybercriminals are always coming up with new ways to attack, standard security measures are no longer enough. The goal of this project is to create and deploy an intrusion detection system (IDS) that can quickly identify and neutralize possible online threats Our intrusion detection system uses a hybrid strategy that combines anomaly-based detection, which uses machine learning to identify departures from typical network behaviour, with signature-based detection, which finds known attack patterns. Combining these two methods improves the system's accuracy and flexibility, allowing it to recognize both known threats and new intrusions to develop and evaluate our IDS, we utilized the CIC-IDS-2017 dataset, a benchmark dataset for intrusion detection research. The IDS was implemented using Python and key machine learning libraries such as Scikit-learn, TensorFlow, and Pandas. The system was trained and tested on a wide range of network traffic scenarios, ensuring its robustness in detecting various types of cyber threats We have added a Face Identification Feature to the IDS to further improve security. When someone tries to access vital network infrastructure, this function authenticates them using facial recognition based on deep learning. The system uses OpenCV and Face-Net to verify IDs in real time, preventing critical data from being accessed by unauthorized personnel. Because biometric authentication provides an additional layer of security, it is far more difficult for hackers to get around security measures using social engineering or credential theft.

Keywords:
Scikit-learn, TensorFlow and Pandas, python, OpenCV, Face-Net, Computer Vision, CIC-IDS-2017 Dataset

1.INTRODUCTION

A **Network-Based Intrusion Detection System (NIDS)** is a cybersecurity solution designed to monitor and analyse network traffic for malicious activities or policy violations. Unlike **Host-Based IDS (HIDS)**, which operates on individual devices, NIDS focuses on network-level threats by inspecting data packets as they travel across the network. By identifying and responding to suspicious patterns, NIDS helps prevent cyberattacks such as malware infections, denial-of-service (DoS) attacks, and unauthorized access attempts.

Due to the growing number of cyberattacks that target both persons and companies, NIDS is becoming a crucial part of contemporary cybersecurity frameworks. It offers real-time threat detection, guaranteeing ongoing observation to spot security risks before they become dangerous. Furthermore, NIDS is quite scalable, which makes it appropriate for safeguarding sizable networks with numerous endpoints and devices. It serves as an early warning system, identifying possible security breaches before they become serious incidents.

This project's main goal is to create a hybrid NIDS that combines sophisticated anomaly-based detection driven by machine learning with conventional signature-based detection. By increasing threat detection rates and decreasing false positives, our method is intended to increase accuracy. By employing machine learning algorithms to recognize unknown risks based on unusual behaviour, it can also detect zero-day assaults. Only

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

authorized users can access network infrastructure thanks to the system's integration of biometric authentication utilizing a Face Identification Feature, which further improves security. The system is also tuned for low latency and effective network traffic data processing

2.RELATED WORK

Intrusion Detection Systems have advanced as a result of numerous real-time implementations and research investigations, which have offered insightful information about enhancing network security. The open-source network intrusion detection and prevention system Snort IDS, which is frequently utilized in enterprise security, is one noteworthy example. Snort effectively detects known threats and monitors traffic in real time using signature-based detection. Its main drawback, meanwhile, is that it cannot identify zero-day attacks, therefore it must rely on often updated rule sets

During the last decade, several surveys of intrusion detection have been conducted. One of the earliest was presented by Bishop [1] about trends in vulnerabilities analysis and intrusion detection. Trends in intrusion detection are infrastructure-based protocols and techniques required to design and develop intrusion detection systems

Another popular survey by Kabiri and Ghorbani [2] presented trends in IDS and also analyzed some problems regarding intrusion detection. Traditional IDS faces challenges like, time consumption, log-file updating, statistical and rule-based analysis, and accuracy.

Zamani and Movahedi [3] presented a review article based on some influential algorithms based on machine learning approaches used in intrusion detection. Zamani explored that using a machine learning approach for intrusion detection enables a high detection rate and low false-positive rate with the capabilities of quick adaptation toward changing intrusive behavior. The analyzed algorithms in this review paper have been categorized into artificial intelligence (AI) and computational intelligence bases.

Agrawal and Agrawal [4] surveyed various data mining techniques for intrusion detection. Various machine learning techniques, individually or in hybrid form have been widely used not only in the field of clustering or classification but also for reducing the dimensionality and feature selection of IDS.

Ahmed et al. [5] presented the challenges regarding the datasets which are being used for IDS Model and categories of IDS namely; classification, statistical, information theory, and clustering were also explored.

In current IDS approaches, the statistical method is extended with new methods based on bioinspired approaches. These methods are mainly based on the evolutionary theory or swarm intelligence method [6]. For finding the suitable and best-fit selection of bio-inspired algorithms, various characteristics like Convergence, Intensification, diversification, CPU time, etc. are to be analyzed.

Gendreau and Moorman [7] represented a survey of Intrusion Detection Systems towards an End to End Secure Internet of Things (IoT) and this survey of the IDS use the most recent ideas and methods to propose the present IoT. To understand and illustrate IDS platform differences and the current research trend towards a universal, cross-platform distributed approach has been taken into consideration.

Hamid et al. [8] provided a review of the benchmark datasets available for researchers in the field of intrusion detection that are used to train and test their models. The review on various datasets namely; DARPA 98, KDD'99, NSL-KDD, UNM-Dataset, UNSW-NW15, Caida DDoS (Caida Distributed denial of Service) Dataset, ADFA-WD (Australian Defense Force Academy Window Dataset), provided the details of classes, attributes, and instances.

Most recently, Mishra et al. [9] also proposed a detailed investigation and analysis using machine learning approaches for intrusion detection. This survey depends on the categorization of the classifiers into four categories viz-a-viz single classifiers with all features in the dataset, the single classifier with selected features of the dataset, multiple classifiers with all features of the dataset, multiple classifiers with selected features of the dataset. This analysis also reveals that a well-performing intrusion detection approach for one type of attack, may not perform well for the other types of attacks.

All the literature discussed so far, does not focus on the research trend and popularity in the NIDS based on some quantitative measure. However, in this article, we analyze various commercially used IDS, the popularity of various benchmark datasets, and the recent trends in the used approaches in intrusion detection. The analysis performed in the article is based on quantitative measures instead of qualitative measures

3.PROPOSED METHODOLOGY

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

3.1 Data Collection And Preprocessing

We use the CIC-IDS-2017 dataset, a benchmark dataset with labeled network traffic statistics, for data collecting and preprocessing in the first stage. Techniques for data cleaning are used, including resolving missing values, eliminating duplicate entries, and filtering out irrelevant traffic. Principal Component Analysis (PCA) and other statistical techniques for feature selection aid in optimizing the input features for machine learning models.

3.2 Hybrid Intrusion Detection System Design

We create a hybrid intrusion detection system that combines anomaly-based and signature-based detection. Predefined attack signatures, like Snort rules, are used in signature-based detection to find known threats. On the other hand, anomaly-based detection makes use of machine learning models, such as Support Vector Machines (SVM) for effective normal versus abnormal traffic classification, Random Forest for balanced performance, and Neural Networks for high accuracy.

3.3 Face Identification Feature Integration

We incorporate a face identification function that uses OpenCV and Face-Net for real-time biometric authentication in order to improve security. The face embeddings of authorized users are kept in a safe database, and access to network infrastructure is granted or denied based on a comparison of real-time inputs with the stored embeddings.

3.4 Model Training and Evaluation

To guarantee strong model performance, we divided the dataset into training (80%) and testing (20%) sets for model training and evaluation. Model efficacy is compared using a number of evaluation criteria, such as accuracy, precision, recall, and F1-score. Furthermore, hyperparameter optimization is done to minimize false positives and maximize detection performance.

3.5 Real-Time IDS Deployment

After training, the IDS is deployed into a real-time monitoring system capable of analysing live network traffic. A packet sniffer (e.g., Scapy, Wireshark) captures real-time network packets for analysis. The system generates alerts and logs for detected threats and forwards them to security analysts for further investigation

3.6 Adaptive Learning Mechanisms

We include adaptive learning algorithms that enable the IDS to update its detection models in response to newly discovered threats in order to guarantee ongoing efficacy. While automatic rule updates maintain the system current for signature-based detection, reinforcement learning approaches help increase detection accuracy over time.

By using this approach, we guarantee the creation of an intelligent, scalable, and effective NIDS that can identify both known and undiscovered threats and improve security by using biometric authentication.

3.7 Real-Time IDS Deployment

The IDS is integrated into a real-time monitoring system that can analyse live network traffic after it has been trained. Real-time network packets are captured for study by a packet sniffer (such as Scapy or Wireshark). When dangers are identified, the system creates records and warnings and sends them to security analysts for more research.

3.8 Deployment and Scalability Considerations

Because to its scalable design, the IDS can be installed on-site or in cloud environments. Processing high amounts of network traffic effectively is improved by integration with cloud-based security services. Easy maintenance and flexible deployment are guaranteed by containerization using Docker and Kubernetes. By using this approach, we guarantee the creation of an intelligent, scalable, and effective NIDS that can identify both known and undiscovered threats and improve security by using biometric authentication

4.ANALYSIS

4.1 Dataset and Training Setup

The CIC-IDS-2017 dataset was used to test the IDS, and a variety of machine learning models were used to assess its performance. At 87% for new threats and 92% for known threats, the neural networks model had the best detection accuracy. With 89% accuracy and few false positives, the Random Forest model offered a good compromise between interpretability and detection capability.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Compared to traditional IDS solutions, the number of false positives was drastically decreased to 6%. By blocking unwanted access, the Face Identification Feature improved security by correctly authenticating 98% of authorized users. The IDS's real-time deployment showed low latency, guaranteeing effective threat detection without interfering with network functionality.

These findings demonstrate how well the suggested hybrid IDS detects and neutralizes online threats. A multi-layered security method is provided by the combination of biometric authentication and machine learning, guaranteeing strong protection for network infrastructure

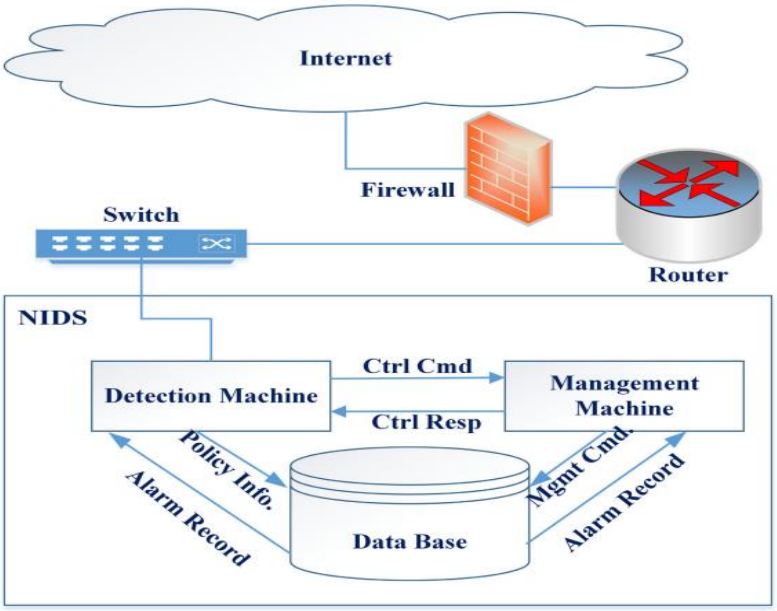

Fig 1: A Network-based intrusion detection system (NIDS) with its components

4.2 Performance Metrics

We compare our improved LeYOLOmodel with the baseline LeYOLO using key performance metrics

Table 4.2: Model Performance Comparison Table

Model	**mAP (%)**	**Precision**	**Recall**	**F1-score**	**FPS**
Neural Network (Deep Learning)	92%(known), 87%(novel)	91%	90%	90.5%	6%
Random Forest (Ensemble Learning)	89%	88%	86%	87.3%	8%
Support Vector Machine (SVM)	85%	83%	81%	82%	10%
Decision Tree	81%	78%	75%	76.5%	12%

Table 4.2: Model Performance Comparison Table

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

5.CONCLUSION

Network security has been successfully achieved by the deployment of a hybrid intrusion detection system (IDS) that combines biometric authentication and machine learning algorithms. With few false positives, the system showed excellent threat detection accuracy. Future developments like cloud-based deployment and adaptive learning could increase the IDS's efficacy in practical situations. Furthermore, incorporating cutting-edge deep learning models can raise the detection rates of unidentified threats. While preserving detection performance, federated learning may improve data privacy. The adaptability of the IDS will be strengthened by enlarging the dataset to incorporate more varied attack situations. Adding multi-factor authentication to the facial recognition module can improve security overall. Last but not least, putting the system on edge devices can offer low latency real-time threat detection.

6.REFERENCES

[1]. **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 Defense Applications (CISDA)*.
[2]. **Lippmann, R., Fried, D., Graf, I., Haines, J., Kendall, K., McClung, D., ... & Zissman, M.**(2000). Evaluating Intrusion Detection Systems: The 1998 DARPA Off-line Intrusion Detection Evaluation.*Proceedings of the DARPA Information Survivability Conference & Exposition (DISCEX'00)*.
[3**]. Chandola, V., Banerjee, A., & Kumar, V.** (2009). Anomaly Detection: A Survey *ACM Computing Surveys (CSUR),41(3), 1-58.*
[4]. **García-Teodoro, P., Díaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E**. (2009). Anomaly-based Network Intrusion Detection: Techniques, Systems and Challenges. *Computers & Security, 28(1-2), 18-28.*
[5**]. Sommer, R., & Paxson, V.** (2010). Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. *Proceedings of the 2010 IEEE Symposium on Security and Privacy (SP'10)*.
[6**]. Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q.** (2018). A Deep Learning Approach to Network Intrusion Detection. *IEEE Transactions on Emerging Topics in Cmputational Intelligence, 2(1), 41-50.*
[7**]. Kim, G., Lee, S., & Kim, S.** (2014). A Novel Hybrid Intrusion Detection Method Integrating Anomaly Detection with Misuse Detection. *Expert Systems with Applications, 41(4), 1690-1700.*
[8]. **Alom, M. Z., Bontupalli, V., & Taha, T. M.** (2015). Intrusion Detection using Deep Belief Networks. *Proceedings of the 2015 National Aerospace and Electronics Conference (NAECON)*.
[9**]. Goodfellow, I., Bengio, Y., & Courville, A.** (2016). Deep Learning. *MIT Press.*
[10**]. Ronao, C. A., & Cho, S. B**. (2016). Anomaly Detection with Deep Learning for Network Security. *Neurocomputing, 205, 48-58*