IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

NETWORK BREACH PREDICTION

[1]Kavya V R, [2]Bhagya Ravi Kumar, [3]Divya G R

[1] Assistant Professor, [2] Student, [3] Student

Department of Information Science and Engineering,

Cambridge Institute of Technology, Bangalore, India

Abstract: Establishing data for an Intrusion Detection System (IDS) typically entails configuring the actual working environment to explore potential attacks, a process that can be prohibitively costly. However, such software is crucial for safeguarding computer networks against unauthorized access, including from potential insiders. The task of training an intrusion detector involves developing a predictive model, often a classifier, capable of distinguishing between "bad" connections (intrusions or attacks) and "good" regular connections.To address the expense and complexity associated with real-world testing, this study focuses on predicting whether connections are under attack using the KDDCup99 dataset and various machine learning methods. The objective is to enhance packet connection predictions for better accuracy, particularly in identifying DOS, R2L, U2R, Probe, and overall attacks. This involves evaluating and comparing supervised classification algorithms to identify the most accurate predictive results. Additionally, the study assesses algorithm performance through classification reports, confusion matrices, and data prioritization.

INTRODUCTION`

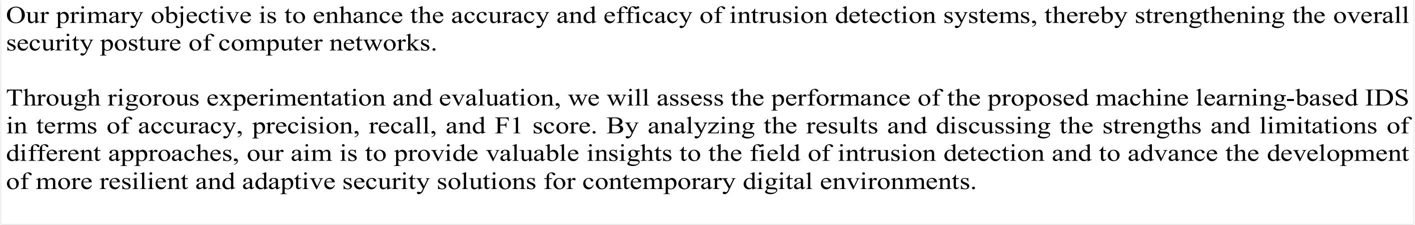
In today's interconnected digital landscape, ensuring the security of Computer Networks is paramount. The escalating sophistication of cyber threats presents an ongoing challenge for organizations striving to protect their systems and data from illicit access and malevolent activities. At the heart of Network security lies the Intrusion Detection System (IDS), tasked with monitoring and assessing network traffic to swiftly identify and respond to hypothetical security breaches.

Traditionally, IDSs employed rule-based approaches, using predefined patterns of known attacks to identify and address threats. However, these methods often struggle to keep pace with increasing cyber threats, resulting in issues such as false positives and negatives.

To conquer this idea, there is a surge of interest in exposing machine learning techniques for intrusion detection. It can learn from past observations and accommodate to new patterns and trends in network traffic, making them adept at detecting both known and unknown attacks. By training predictive models on labeled datasets containing examples of normal and malevolent network activity, machine learning based IDSs can automatically classify incoming traffic and alert administrators to potential security incidents.

The KDDCup99 dataset, sourced from Program, serves as a largely recognized benchmark dataset for evaluating intrusion detection systems. It encompasses a comprehensive collection of network traffic data entrap from a simulated environment, featuring various attack types such as Denial of Service (DOS), Remote-to-Local (R2L), User-to-Root (U2R), and Probe attacks, alongside legitimate network traffic.

In this, our aim is to harness machine learning techniques to develop a robust IDS capable of accurately predicting different types of network attacks using the KDDCup99 dataset. Through exploration and differentiation of various supervised sorting algorithms, including process diagrams, irregular forests, support vector Machines, and logistic regression, our aim is to identify the most suitable approach.

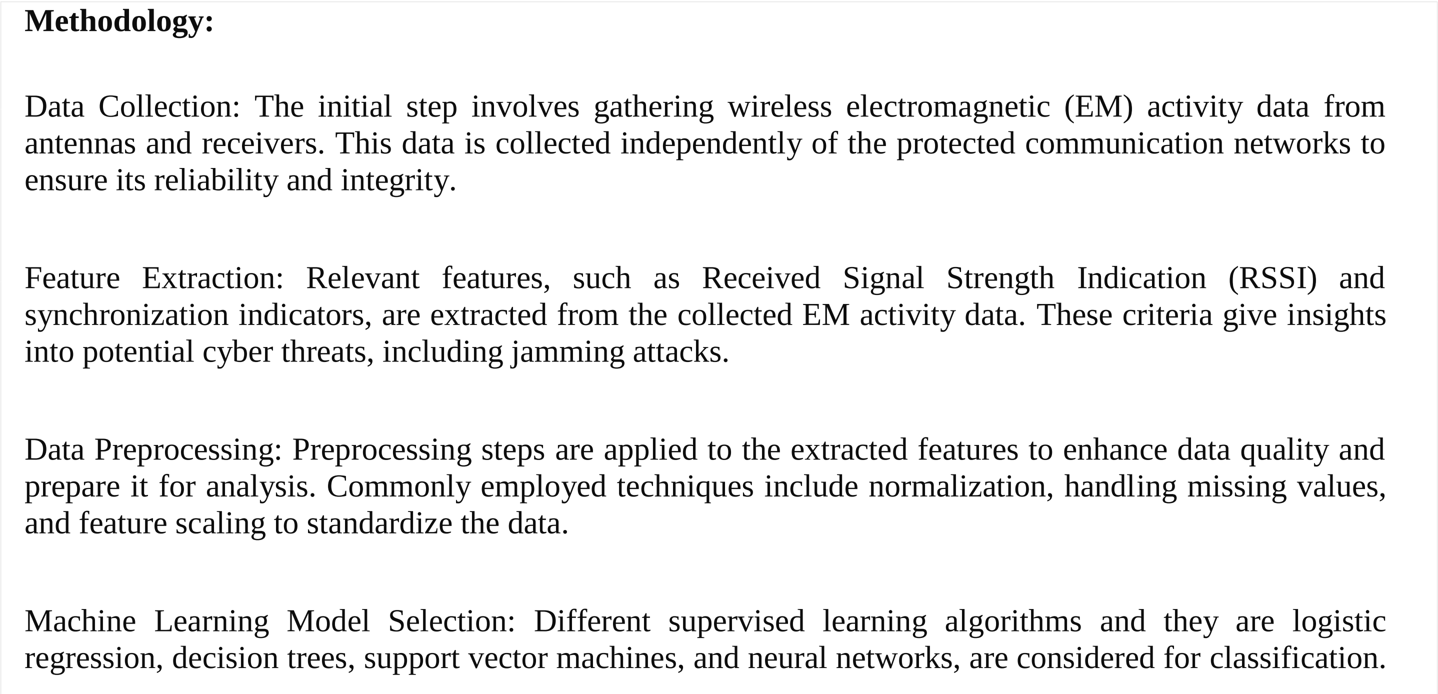


Our primary objective is to enhance the accuracy and efficacy of intrusion detection systems, thereby strengthening the overall security posture of computer networks.

Through rigorous experimentation and evaluation, we will assess the performance of the proposed machine learning-based IDS in terms of accuracy, precision, recall, and F1 score. By analyzing the results and discussing the strengths and limitations of different approaches, our aim is to provide valuable insights to the field of intrusion detection and to advance the development of more resilient and adaptive security solutions for contemporary digital environments.

Literature Review

1. **"Intrusion Detection Systems: A Survey and Taxonomy" by Stefan Axelsson (2000):** Axelsson's paper presents a comprehensive taxonomy of intrusion detection systems, organizing them based on detection approach, data sources, and detection paradigm. It traces the development of intrusion detection techniques and emphasizes the necessity for adaptive and intelligent systems capable of identifying new attack patterns.
2. **"Anomaly-Based Intrusion Detection with Machine Learning" by Mark Last, Abraham Kandel, and Horst Bunke (2007):** This study delves into the utilization of Machine learning algorithms for anomaly-based intrusion identifiaction. The authors explore clustering, neural networks, and other machine learning methods to detect deviations from normal network behavior, emphasizing the importance of adapting to evolving threats.
3. **"Using Machine Learning Algorithms for DoS Attack Detection" by Khaled Rabieh and Alaa Sheta (2014):** Rabieh and Sheta's findings focuses on accessing Machine learning algorithms for detecting Denial of Service (DoS) attacks. They uses various classification techniques, such as decision trees and support vector machines, to differentiate between normal and malicious network traffic, showcasing the efficacy of machine learning in countering DoS attacks.
4. **"A Comparative Analysis of Intrusion Detection using Machine Learning Techniques" by Mohammed M. Alani and H. S. Al-Raweshidy (2016):** This comparative analysis evaluates the performance of different machine learning algorithms for intrusion detection using authentic network datasets. The authors assess classifiers such as decision trees, k-nearest neighbors, and naive Bayes based on accuracy, false positive rate, and detection time, gives valuable insights into their practical applicability.
5. **"Deep Learning for Network Intrusion Detection: An Overview" by Nabil A. Alrajeh, Benjamin C. M. Fung, and Mourad Debbabi (2018):** Alrajeh et al. provide an overview of deep learning ideas applied to network intrusion detection. They discuss the benefits of deep neural networks in capturing intricate patterns in network traffic data and highlight recent advancements in deep learning-based intrusion detection systems, indicating future research directions in the field.



Methodology:

Data Collection: The initial step involves gathering wireless electromagnetic (EM) activity data from antennas and receivers. This data is collected independently of the protected communication networks to ensure its reliability and integrity.

Feature Extraction: Relevant features, such as Received Signal Strength Indication (RSSI) and synchronization indicators, are extracted from the collected EM activity data. These criteria give insights into potential cyber threats, including jamming attacks.

Data Preprocessing: Preprocessing steps are applied to the extracted features to enhance data quality and prepare it for analysis. Commonly employed techniques include normalization, handling missing values, and feature scaling to standardize the data.

Machine Learning Model Selection: Different supervised learning algorithms and they are logistic regression, decision trees, support vector machines, and neural networks, are considered for classification.

The selection of the algorithm lies on the data characteristics and the specific requirements of the classification task.

Model Training: The chosen machine learning model is trained using labeled data, which includes instances of EM activity categorized as normal or denotes a jamming attack. The model learns the difference between these classes based on the provided training data.

Model Evaluation: The trained model is checked using a separate dataset to gauge its accomplishment in finding and classifying jamming attacks. Checking metrics like correctness, recall, and F1 score are utilized to quantify the model's effectiveness.

Optimization and Fine-Tuning: The model undergoes optimization and fine-tuning to further enhance its performance. This process involves adjusting hyperparameters, feature selection, and exploring different algorithms to achieve optimal results.

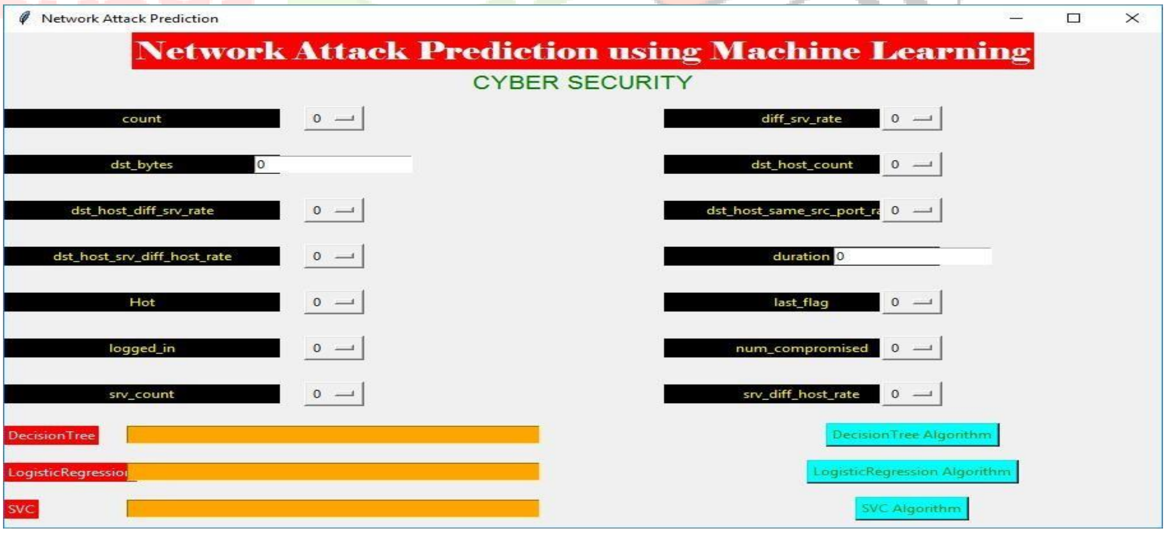
Validation and Deployment: Once the model demonstrates satisfactory performance, it undergoes validation to ensure its generalization capabilities. Subsequently, the validated model is deployed for real-time detection and classification of jamming attacks in wireless communication networks.

By these Approach, we aim to develop a robust solution for detecting and mitigating cyber-attacks in wireless communication networks, thereby bolstering overall security measures and protect data from illegal access and interception.

Additionally, this Approach includes steps specific to traffic flow prediction.

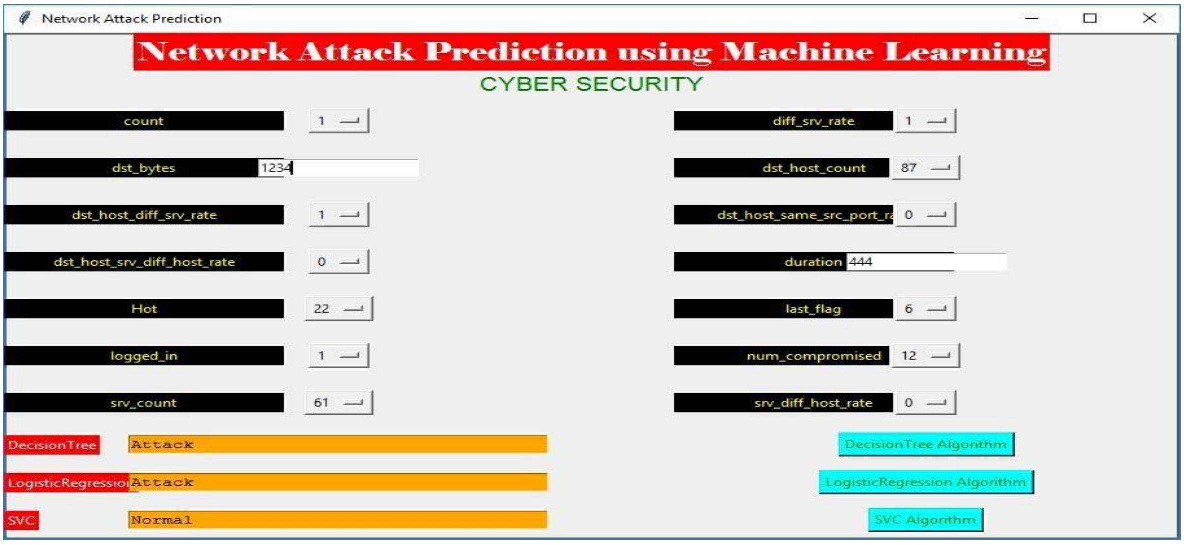
- Dataset Collection: Data is collected from the Kaggle dataset, which includes entities such as day, date, zone, and weekend status.
- Splitting Dataset: The data record is split into trained and tested sets, allocating 70% for training and 30% for testing.
- Model Training: The KNN and irregular Forest algorithms are trained using the trained data.
- Model Evaluation: The model performance is assessed using metrics such as exact, precision, recall, F1-score.

RESULTS


Network Attack Prediction
Network Attack Prediction using Machine Learning
CYBER SECURITY
DecisionTree
LogisticRegressio
SVC
DecisionTree Algorithm
LogisticRegression Algorithm
SVC Algorithm

count 0
dst_bytes 0
dst_host_diff_srv_rate 0
dst_host_srv_diff_host_rate 0
Hot 0
logged_in 0
srv_count 0
diff_srv_rate 0
dst_host_count 0
dst_host_same_src_port_ra 0
duration 0
last_flag 0
num_compromised 0
srv_diff_host_rate 0

Fig: User Interface before prediction



Network Attack Prediction

Network Attack Prediction using Machine Learning

CYBER SECURITY

count 1

dst_bytes 1234

dst_host_diff_srv_rate 1

dst_host_srv_diff_host_rate 0

Hot 22

logged_in 1

srv_count 61

diff_srv_rate 1

dst_host_count 87

dst_host_same_src_port_r 0

duration 444

last_flag 6

num_compromised 12

srv_diff_host_rate 0

DecisionTree Attack

LogisticRegressior Attack

SVC Normal

Fig: User Interface after prediction

Conclusion

In conclusion explains the efficacy of leveraging Machine learning ideas for predicting potential attacks in Intrusion Detection Systems (IDS) using the KDDCup99 dataset. By focusing on improving packet connection predictions, particularly in identifying various types of attacks, including DOS, R2L, U2R, and Probe attacks, the study contributes to enhancing network security measures. Through rigorous evaluation and comparison of supervised classification algorithms, the research identifies the most accurate predictive models, showcasing large precision, recall, and F1 score. The findings underscore the effect of the expressed machine learning approach in intrusion detection, offering valuable insights for developing more robust security solutions in the face of increasing cyber threats.

REFERENCE

[1] Chalapathy, R. and Chawla, S. "Deep learning for anomaly detection: A review" (ACM Computing Surveys, 2019).

[2] Yassein, M. B. et al. "Deep Learning-Based Intrusion Detection Systems in the Fog of Things: Review, Challenges, and Directions" (IEEE Access, 2019).

[3] Saad, S. et al. "A Survey on Deep Learning in Network Security" (IEEE Communications Surveys & Tutorials, 2019).

[4] Alazab, M. et al. "A Survey on Deep Learning Techniques for Cyber Security" (Journal of Information Security andApplications, 2020).

[5] Saeed, S. et al. "Machine Learning-Based Intrusion Detection Systems: A Survey" (Journal of Network and Computer Applications, 2021)