IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

Machine Learning Based IoT Network Intrusion Detection Classification

Dr. Jyoti Deshmukh[1], Pooja Hargude[2], Divya Ghate[3], Sacchidanand Linge[4], Rahul Mahajan[5]
Faculty, Department of Computer Engineering / Information Technology[1]
Students, Department of Computer Engineering / Information Technology[2,3,4,5]
G H Raisoni Institute of Engineering and Technology, Pune, Maharashtra, India

Abstract: *IoT network is a promising technology, IoT implementation is growing rapidly but cybersecurity is still a loophole, detection of attacks in IOT infrastructures is a growing concern in the field of IoT. With the increased use of Internet of Things in different areas, cyber-attacks are also increasing proportionately and can cause failures in the system. IDS become the leading security solution. Anomaly based network intrusion detection (IDS) detection plays a major role in protecting networks against various malicious activities. Improving the security of loT networks has become one of the most critical issues. This is due to the large-scale development and deployment of loT devices and the insufficiency of Intrusion Detection Systems (IDS) to be deployed for the use of special purpose networks. In this article, the performance of several machine learning models has been compared to accurately predict attacks on IoT systems, the case of imbalanced classes was subsequently treated using the SMOTE technique.*

Keywords: IoT, Network Intrusion Detection

I. INTRODUCTION

Internet of Things (IoT) and its applications are the most popular research areas at present. The characteristics of IoT on one side make it easily applicable to real-life applications, whereas on the other side expose it to cyber threats.
Due to open nature, global connectivity and resource constrained nature of smart devices and wireless networks the Internet of Things is susceptible to various routing attacks.
Internet of Things (IoT) integrates billions of self-organized and heterogeneous smart nodes that communicate with each other without human intervention.
Most of the latest IDS are based on a machine learning algorithm for the detection of cyber-attacks in the network. currently their use is in all areas of human life. The IoT network consists of connections between different types of smart objects ranging from supercomputers to small devices which can have very low computing power, so securing this type of network is difficult and therefore cybersecurity is a big loophole Statista has estimated the impressive number of connected loT devices in 2020 and this number will double by 2025.

1.Title: Feature Extraction for Machine Learning-based Intrusion Detection in IoT Networks
Link: https://arxiv.org/abs/2108.12722
Author: Siamak Layeghy
Published Date: 28 August 2021
The tremendous numbers of network security breaches that have occurred in IoT networks have demonstrated the unreliability of current Network Intrusion Detection Systems (NIDSs). Consequently, network interruptions and loss of sensitive data have occurred which led to an active research area for improving NIDS technologies. During an analysis of related works, it was observed that most researchers aimed to obtain better classification results by using a set of untried combinations of Feature Reduction (FR) and Machine Learning (ML) techniques on NIDS datasets.

2.Title: Supervised Machine Learning Based Network Intrusion Detection System for Internet of Things
Link: https://ieeexplore.ieee.org/abstract/document/9225340
Author: Deepa Rani, Narottam Chand Kaushal
Publish Date: 15 October 2020

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

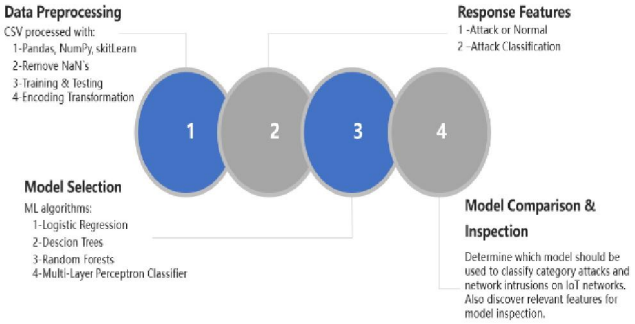
This paper mainly proposes an efficient method with uniform detection system based on supervised machine learning technique by using Random Forest classifier. Also, two different datasets, NSL-KDD and KDDCUP99 with minimal feature sets have been used that give lightweight attack detection strategy for IoT network. Simulation of proposed method with these datasets has 99.9 percentage accuracy in intrusion detection with less amount of time and energy.

3.Title: Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using UNSW-NB15 data-set
Link: https://jwcneurasipjournals.springeropen.com/articles/10.1186/s13638-021-01893-8
Author: Muhammad Ahmad, Qaiser Riaz, Muhammad Zeeshan, Hasan Tahir, Syed Ali Haider
Published Date: 21 January 2021
They apply supervised Machine Learning (ML) algorithms, i.e., Random Forest (RF), Support Vector Machine and Artificial Neural Networks on the clusters. Using RF, we, respectively, achieve 98.67% and 97.37% of accuracy in binary and multi-class classification. In clusters-based techniques, we achieved 96.96%, 91.4% and 97.54% of classification accuracy by using RF on Flow & MQTT features, TCP features and top features from both clusters.

4. Title: Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset)
Link:https://www.researchgate.net/publication/348206258_Machine_Learning_Based_IoT_Intrusion_Detection_System_An_MQTT_Case_Study_MQTT-IoT-IDS2020_Dataset
Author: Miroslav Bures
Publish Date: January 2021
In this paper, the effectiveness of six Machine Learning (ML) techniques to detect MQTT-based attacks is evaluated. Three abstraction levels of features are assessed, namely, packet-based, unidirectional flow, and bidirectional flow features. An MQTT simulated dataset is generated and used for the training and evaluation processes. The dataset is released with an open access license to help the research community further anal-yse the accompanied challenges.

5.Title: Internet of Things: A survey on machine learning-based intrusion detection approaches
Link: https://www.sciencedirect.com/science/article/abs/pii/S1389128618308739
Author: KeltonA.P.da,Celso O.Lisboaa,RobertoMunoz ,Victor Hugo C.de Albuquerque
Publish Date: 14 March 2019
This research has noticed that intrusion detection within the Internet of Things context still presents a challenge. As the Internet evolves into IoT, the focus shifts from connectivity to data. This work, therefore, focused on the newest studies in intrusion detection and intelligent techniques applied to IoT to keep data secure.
This section includes some of the researches of various ML algorithms and classifiers integrated IDSs to detect intrusions in IoT networks. Roy et al. has introduced a Bi-LSTM recurrent neural network (Bi-LSTM RNN) approach for intrusion detection aiming to identify a binary classification of normal and attack patterns. The implemented model has been trained using the UNSW-NB15 dataset and it achieves over 95% accuracy in IoT attack detection.

II. METHODOLOGY



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

2.1 Date Processing

Data Processing is the task of converting data from a given form to a much more usable and desired form i.e., making it more meaningful and informative. Using Machine Learning algorithms, mathematical modeling, and statistical knowledge, this entire process can be automated.

2.2 Response Features

Choosing informative, discriminating and independent features is a crucial element of effective algorithms in pattern recognition, classification and regression. Features are usually numeric, but structural features such as strings and graphs are used in syntactic pattern recognition. The concept of "feature" is related to that of explanatory variable used in statistical techniques such as linear regression.

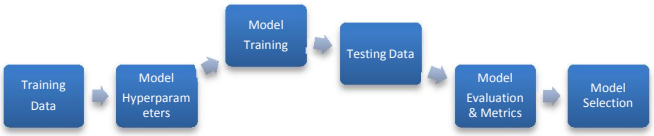
2.3 Model Selection

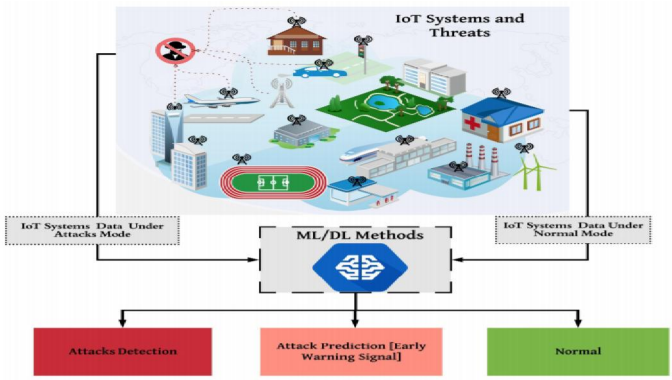
Model selection refers to the process of choosing the model that best generalizes. Training and validation sets are used to simulate unseen data. Overfitting happens when our model performs well on our training dataset but generalizes poorly.

2.4 Model Comparison & Inspection

Determine which model should be used to classify category attacks and network intrusions on IoT networks. Also discover relevant features for model inspection.

III. SYSTEM ARCHITECTURE





3.1 Proposed Algorithm

A. Logistic Regression

Logistic regression is one of the most popular Machine Learning algorithms,which comes under the Supervised Learning technique.

It is used for predicting the categorical dependent variable using a given set of independent variables.

B. Decision Trees

Decision Tree is a Supervised learning technique that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems.

It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

C. Random Forests

Random Forest Algorithm is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.
The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

D. Multi-Layer Perceptron Classifier

A multilayer perceptron (MLP) is a feed forward artificial neural network that generates a set of outputs from a set of inputs. An MLP is characterized by several layers of input nodes connected as a directed graph between the input nodes connected as a directed graph between the input and output layers.

IV. ACKNOWLEDGEMENT

We take this opportunity to thank our project guide and Head of the Department Prof. Yogesh Mali for their valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this project report. We are also thankful to all the staff members of the Department of Information Technology of G.H. Raisoni Institute of Engineering & Technology, Pune for their valuable time, support, comments, suggestions and persuasion. We would also like to thank the institute for providing the required facilities, Internet access and important books.

V. CONCLUSION

Intrusion detection systems is an inevitable processing unit in recent wireless networks due to lack of security and increased number of intruders. IoT is a heterogeneous network which severely faces security threats similar to wireless networks and it is essential to develop an intrusion detection system to avoid performance degradation in IoT networks.
Proposed research work analysis the different types of attacks in IoT and proposed a hybrid convolutional neural network module by incorporating long short-term memory process.
Proposed model is experimentally verified and compared with conventional recurrent neural network and attains better detection accuracy of 98% which makes the application suitable for different IoT environments.

REFERENCES

[1]. Moustafa, Nour, and Jill Slay. "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)." Military Communications and Information Systems Conference (MilCIS), 2015. IEEE, 2015.
[2]. Moustafa, Nour, and Jill Slay. "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 dataset and the comparison with the KDD99 dataset." Information Security Journal: A Global Perspective (2016): 1-14.
[3]. Moustafa, Nour, et al. "Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks." IEEE Transactions on Big Data (2017).
[4]. Moustafa, Nour, et al. "Big data analytics for intrusion detection system: statistical decision-making using finite dirichlet mixture models." Data Analytics and Decision Support for Cybersecurity. Springer, Cham, 2017. 127-156.
[5]. Feature Extraction for Machine Learning-based Intrusion Detection in IoT Networks https://arxiv.org/abs/2108.12722
[6]. Supervised Machine Learning Based Network Intrusion Detection System for Internet of Things https://ieeexplore.ieee.org/abstract/document/9225340
[7]. Intrusion detection in internet of things using supervised machine learning based on application and transport layer features using UNSW-NB15 data-set https://jwcn-eurasipjournals.springeropen.com/articles/10.1186/s13638-021-01893-8
[8]. Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study (MQTT-IoT-IDS2020 Dataset)https://www.researchgate.net/publication/348206258_Machine_Learning_Based_IoT_Intrusion_Detection_System_An_MQTT_Case_Study_MQTT-IoT-IDS2020_Dataset

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, February 2023
Impact Factor: 7.301

[9]. Internet of Things: A survey on machine learning-based intrusion detection approaches https://www.sciencedirect.com/science/article/abs/pii/S1389128618308739

[10]. Towards Machine Learning Based IoT Intrusion Detection Service https://link.springer.com/chapter/10.1007/978-3-319-92058-0_56