SRM Institute of Science and Technology,

Ramapuram Campus, Chennai-89

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

A Flexible Hybrid Behavior Dynamic Methodology for Network Intrusion Detection using CNN

Batch No: A10

DETAILS OF THE PROJECT MEMBERS

Viswanath VS	RA1911003020012
Surendharan TG	RA1911003020021
Naveen PK	RA1911003020054

SUPERVISOR DETAILS

Ms. P.Jayalakshmi
AP/CSE

OBJECTIVE

The integration of Artificial Intelligence (AI) technology into the development of Network Intrusion Detection Systems (NID) aims to improve the performance and effectiveness of these systems in detecting and preventing network intrusions. Machine Learning algorithms enable the processing of vast amounts of data, enabling the systems to identify patterns and anomalies in real-time. The objective of this integration is to enhance the overall cybersecurity posture of organizations by providing them with real-time protection against cyber-attacks.

SCOPE

The scope of AI technology integration into the development of Network Intrusion Detection Systems (NID) includes the following main areas:

- **Massive data analysis:** AI methods, such as machine learning, allow NID systems to analyze massive volumes of data and spot trends, making them more successful at detecting network breaches.
- **Real-time anomaly detection:** NID systems integrated with AI algorithms are capable of identifying abnormalities in real-time, providing enterprises with the essential cyber security protection.
- **Learning from previous encounters:** AI algorithms enable NID systems to learn from previous events and adapt to new threats, making them more successful at detecting intrusions.

ABSTRACT

Artificial Intelligence (AI) has significantly impacted the development of Network Intrusion Detection Systems (NIDS). AI algorithms, such as machine learning, have enabled NID systems to analyze vast amounts of data, identify patterns, and detect anomalies in real-time. This helps in detecting and preventing cyber attacks, which are becoming increasingly sophisticated and challenging to detect. NID systems that use AI algorithms are capable of learning from past experiences and can adapt to new threats, making them more effective in detecting intrusions. Furthermore, AI algorithms also help in reducing false positive alarms, which can be a significant hindrance in the effective functioning of NID systems. In conclusion, the integration of AI into NID systems has significantly improved their performance and effectiveness in detecting network intrusions, making them an essential tool in the fight against cybercrime.

INTRODUCTION

- Networks have taken on a crucial role in today's linked world and in contemporary commercial operations.

- Network Intrusion Detection Systems (NIDS) are essential for safeguarding networks in the linked world of today, when cyber dangers are pervasive.

- Security systems known as Network Intrusion Detection Systems (NIDS) are used to monitor network traffic and find unusual activity or intrusions.

- NIDS is used to examine network packets, protocols and traffic patterns to find possible risks.

SYSTEM REQUIREMENTS

Hardware Requirements:

- Intel i7 9th Gen core Processor
- Nvidia GTX 1660ti GPU
- 8gb RAM
- 1 TB HDD

Software Requirements:

- Windows 10
- Jupyter Notebook
- Anaconda
- TensorFlow
- Keras
- Pandas
- Numpy

LITERATURE SURVEY PAPER - 1

Title	**Author**	**Published Year**	**Methodology Used**	**Advantages**	**Disadvantages**
Generating Labeled Training Datasets Towards Unified Network Intrusion Detection Systems	R. Ishibashi, K. Miyamoto, C. Han, T. Ban, T. Takahashi and J. Takeuchi	**2022**	Supervised Learning	• Novel Method to create new Datasets • Highly effective to real-life problems	• Computational Burden • Less Labelled Data

LITERATURE SURVEY PAPER - 2

Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
Immune System Based Intrusion Detection System (IS-IDS): A Proposed Model	Inadyuti Dutt , Samarjeet Borah and Indra Kanta Maitra	**2020**	Statistical Modeling based Anomaly Detection	• Ability To Deliver High Quality Results •Its not difficult to see what is Impacted	• Poor Application Performance • Cannot be implemented real time

LITERATURE SURVEY PAPER - 3

Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
HML-IDS: A Hybrid-Multilevel Anomaly Prediction Approach for Intrusion Detection in SCADA Systems	Izhar Ahmed Khan , Dechang Pi , Zaheer Ullah Khan , Yasir Hussain and Asif Nawaz	**2019**	Multi-level Hybrid based Anomaly detection	• Less resource used to meet demands •Simple to understand and interpret	• Difficult to be used in large-scale parallel computing. • This system is eager and uncontrollable

LITERATURE SURVEY PAPER - 4

Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
Applying big data based deep learning system to intrusion detection	Wei Zhong , Ning Yu and Chunyu Ai	**2020**	Big Data	• Quick Calculation Time •Lowering the Complexity Threshold	• Big payloads • Heavyweight

LITERATURE SURVEY PAPER - 5

Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
A Linear Systems Perspective on Intrusion Detection for Routing in Reconfigurable Wireless Networks	Jaime Zuniga-Mejia , Rafaela Villalpando-Hernandez , Cesar Vargas-Rosales and Andreas Spanias	**2019**	Linear Systems Theory	•Achieve sub-optimal performance. •Improve the operational efficiency.	• It cannot be implemented real time • It cannot meet current network business demands

LITERATURE SURVEY PAPER - 6

Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
Decentralized Intrusion Prevention (DIP) Against Co-Ordinated Cyberattacks on Distribution Automation Systems	Jennifer Appiah-Kubi and Chen-Ching Liu	**2019**	Multi-Agent System	•Simplicity and easy to understand •Simple, fast and less complex.	• Unsuitable for large scale scenarios. • Big payloads

LITERATURE SURVEY PAPER - 7

Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
Multi-Source Multi-Domain Data Fusion for Cyberattack Detection in Power Systems	Abhijeet Sahu , Zeyu Mao , Patrick Wlazlo , Hao Huang , Katherine Davis , Ana Goulart and Saman Zonouz	**2021**	Data Fusion Framwork	•Proving High Robustness and imperceptibility •Provides the integrity and nontransferablity.	• Difficult to be used in large-scale parallel computing. • High complexity of installing and maintaining

LITERATURE SURVEY PAPER - 8

Title	**Author**	**Published Year**	**Methodology Used**	**Advantages**	**Disadvantages**
Intrusion Detection System Based on Integrated System Calls Graph and Neural Networks	F. J. Mora-Gimeno , H. Mora-Mora , B. Volckaert and A. Atrey	**2021**	Neural Networks	•Excellent empirical performance •Fast and efficient, but also as accurate	• Large Payloads • Approach is time-consuming

LITERATURE SURVEY PAPER - 9

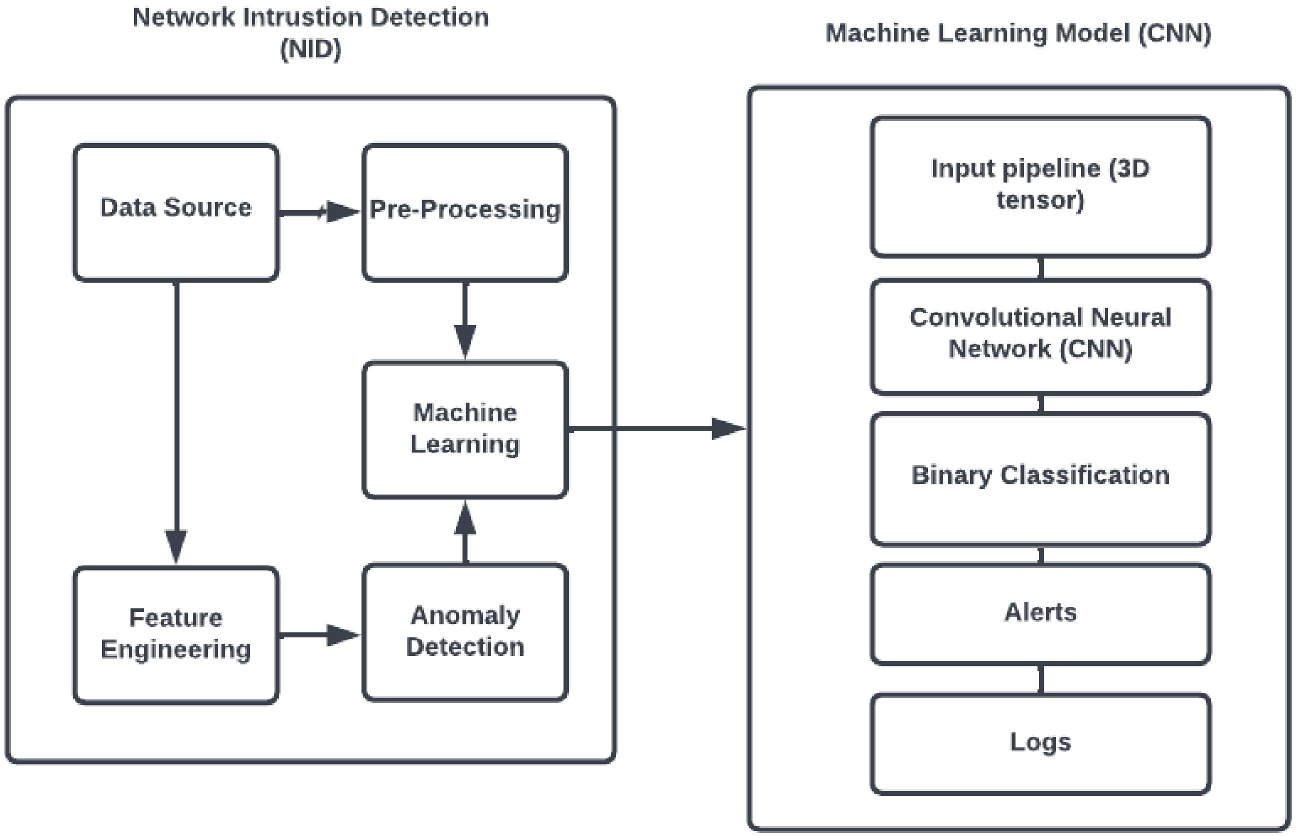
Title	Author	Published Year	Methodology Used	Advantages	Disadvantages
Securing the Smart Grid: A Comprehensive Compilation of Intrusion Detection and Prevention Systems	Panagiotis I. Radoglou-Grammatikis and Panagiotis G. Sarigiannidis	**2019**	Smart Grid paradigm	•Simple to understand and interpret •May meet the real-time requirement.	• High complexity of installing and maintaining • Difficult to be used in large-scale parallel computing.

ISSUES

There are several issues associated with Intrusion Detection Systems (NIDs), which can affect their effectiveness. Some of these issues are:

- **Complexity**: NID systems can be complex and require significant resources to implement and maintain.
- **False positives**: One of the most significant issues with NIDs is the risk of false positives, which are alerts for events that are not actual attacks. False positives can be caused by misconfigurations, network noise, or anomalies.
- **Integration**: NID systems need to integrate with multiple security tools and platforms, such as firewalls, intrusion prevention systems, and security information and event management (SIEM) systems. This can be complex, particularly if the organization is using different vendors for these tools.
- **Training and Expertise**: To effectively use NID systems, organizations need staff who have the skills and expertise to configure, maintain, and monitor the system.

Architecture Diagram


Network Intrustion Detection (NID)
Machine Learning Model (CNN)
Data Source
Pre-Processing
Machine Learning
Feature Engineering
Anomaly Detection
Input pipeline (3D tensor)
Convolutional Neural Network (CNN)
Binary Classification
Alerts
Logs

Algorithm Used

CNN (Convolutional Neural Network):

- **Highly accurate**: CNNs have proven to be highly accurate in image classification and other pattern recognition tasks.
- **Reduced feature engineering**: In traditional machine learning algorithms, the data preprocessing phase involves feature engineering, which can be a time-consuming and labor-intensive process. However, with CNNs, feature engineering is reduced as the algorithm automatically learns the relevant features from the input data.
- **Robust to noise**: CNNs are more robust to noise in the data than traditional machine learning algorithms, making them well-suited for applications where the data is noisy or incomplete.

Proposed Methodology

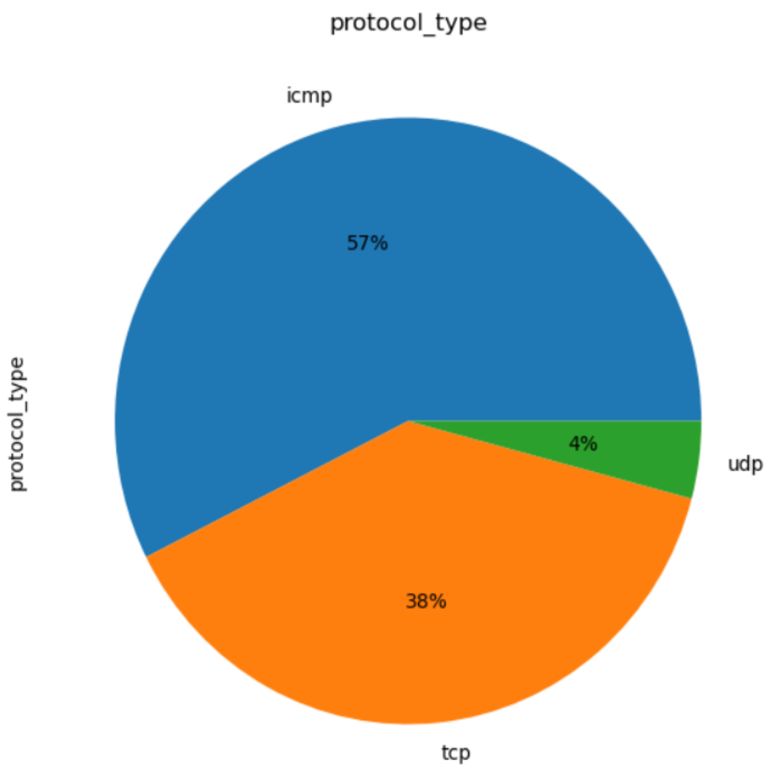
- The purpose of this work is to create an effective machine learning model for network intrusion detection system using the CNN algorithm.
- For this purpose, the Public dataset KDD Cup 1999 is used which contains various attacks that have happened over a period to understand the different forms of vulnerabilities in the system.
- Using this dataset, a model can be created based on which the Network Intrusion Detection system will detect anomalous connections form normal connections.
- Thus, the proposed model is split into three main categories:
    - Data Pre-Processing
    - Model Training
    - Testing and Analysis

Module 1 : Data Pre-Processing

- To improve the NID Network Intrusion Detection System, first step is to identify many datasets which can provide various types of cyberattacks.

- For this, KDD Cup 1999 Dataset which contains various types of attacks namely- DOS attacks such as Neptune, smurf, teardrop and common types of scanning attacks as well as unauthorized access using rootkits, backdoors.

- A dataset that is large enough can provide various types of attacks through which the model can improve in detection of such attacks.



Module 2 : Model Training

- Model Training represents the usage of a machine learning algorithm namely Convolutional Neural Network Algorithm (CNN) to train the model for identifying normal connections from malicious attacks
- The CNN is commonly used for Image classification, but also can be used for text-based analysis where it is converted into sequence of word vectors which is then fed into the one –dimensional convolutional neural network
- Compared to traditional machine learning algorithms that rely on manually defined features, CNNs can automatically learn features from the input data, making them more effective at handling large and complex datasets.
- Additionally, CNNs can capture the context and dependencies between words in the input text, which can be important for accurate classification.

Module 3 : Testing And Analysis

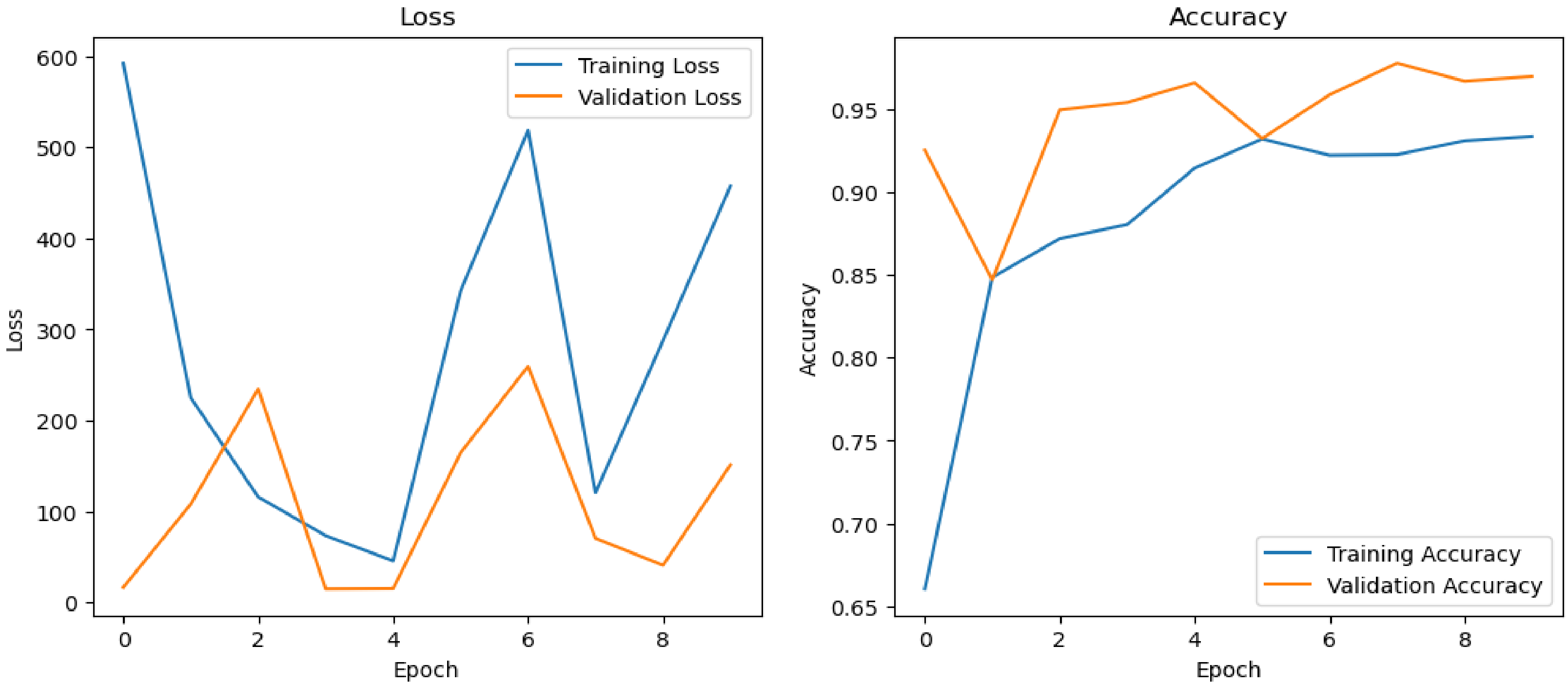
- To evaluate the correctness of the results using the testing data and plotting them, the dataset has already been divided into training and testing data.

- This testing dataset will contain various data in order to generalize the performance of the model

- This data is then pre-processed by tokenizing the text and applying necessary data transformations to be used for testing

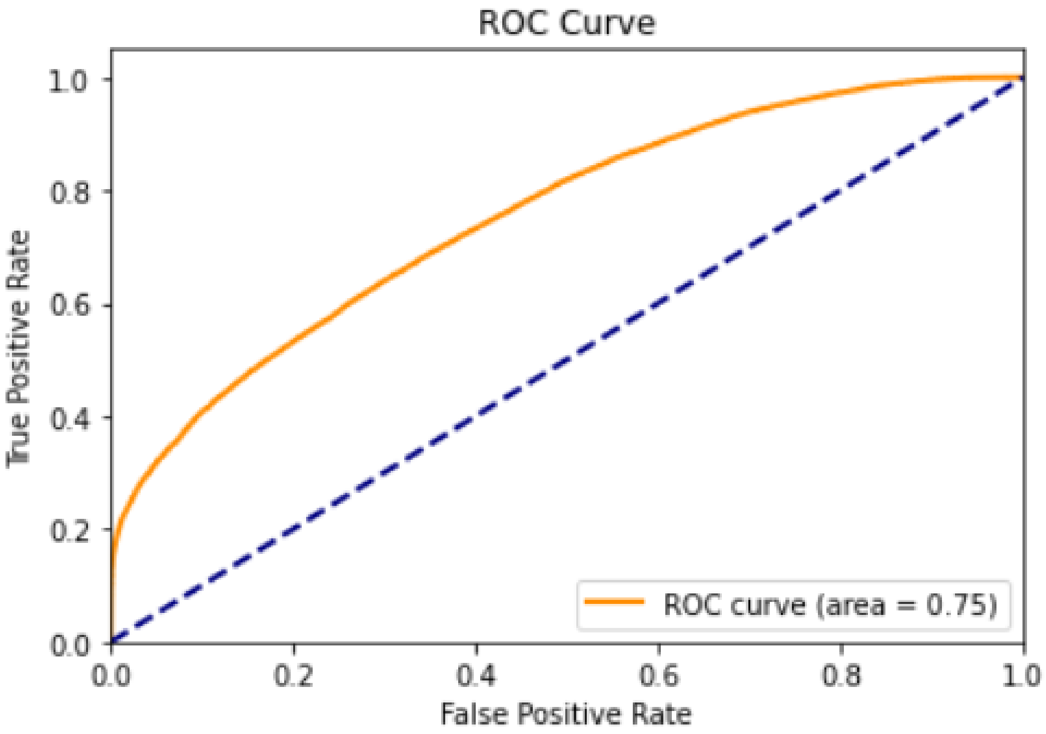
- Thus, it can evaluate using the model by calculating performance metrics such as accuracy and precision

Results



- The Above shows the loss and accuracy graph for the Training model. This graph also shows the difference in accuracy from the training and testing model

Results



- The ROC curve gives a more accurate reading on the training model. The above graph shows curve area of 0.75 which shows an accurate model.

CONCLUSION

The application of artificial intelligence, namely the convolutional neural network (CNN) algorithm, has substantially increased the efficiency and accuracy of network intrusion detection systems. These systems can discover patterns and detect anomalies in network traffic in real time by harnessing the power of machine learning. There are various advantages to applying AI in network intrusion detection systems, including faster threat detection and reaction times, less false positives, and the capacity to identify previously undisclosed or zero-day assaults. Furthermore, AI-based intrusion detection systems can adapt to new and developing threats, making them a significant tool for enterprises and organizations that rely on secure networks to function. Overall, the enlistment of AI in network intrusion detection systems has provided and advantage to network security, delivering more advanced and dependable protection against cyber-attacks.

REFERENCES

[1] Ryosuke Ishibashi, Kohei Miyamoto, Chansu Han , Tao Ban Takeshi Takahashi and Jun’ichi Takeuchi “Generating Labeled Training Datasets Towards Unified Network Intrusion Detection”, Systems May 2022

[2] S. Aljawarneh, M. B. Yassein, and M. Aljundi, ``An enhanced J48 classication algorithm for the anomaly intrusion detection systems,"*Cluster Comput.*, vol. 22, no. S5, pp. 1054910565, Sep. 2017

[3] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, ``Toward generating a new intrusion detection dataset and intrusion trafc characterization,‘’ in *Proc. 4th Int. Conf. Inf. Syst. Secur. Privacy*, 2018

REFERENCES

[4] R.-H. Hwang, M.-C. Peng, V.-L. Nguyen, and Y.-L. Chang, ``An LSTM - based deep learning approach for classifying malicious trafc at thepacket level," *Appl. Sci.*, vol. 9, no. 16, p. 3414, Aug. 2019

[5] H. Hindy, D. Brosset, E. Bayne, A. Seeam, C. Tachtatzis, R. Atkinson,and X. Bellekens, ``A taxonomy of network threats and the effect of current datasets on intrusion detection systems," *IEEE Access*, vol. 8, pp. 104650104675, 2020

[6] Guide to Intrusion Detection and Prevention Systems (IDPS) - Karen Scarfone, Peter Mell , February 2007

REFERENCES

[7] Importance of intrusion detection system(IDS)- A. S. Ashoor and S. Gore ,January-2011

[8] Deep learning approach for network intrusion detection in soft_x0002_ware defined networking- T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, M. Ghogho ,October 2016

[9] Anomaly detection in ad-hoc networks based on deep learning model: A plug and play device -F. Feng, X. Liu, B. Yong, R. Zhou, Q. Zhou ,March 2019

[10] A deep learning method with filter based feature engineering for wireless intrusion detection system -S. M. Kasongo, Y. Sun, March 2019