Available Online at www.ijcsmc.com/journal

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 11, Issue. 9, September 2022, pg.29 – 33

IMPLEMENTATION OF DDOS ATTACKS USING K-MEANS CLUSTERING BASED ON SEMISUPERVISED ALGORITHMS

[1]Dr. D.J. Samatha Naidu; [2]S. Vijay Kumar

[1]Principal, APGCCS, Rajampet, Annamayya(dist) 516216, A.P, India
[2]MCA Department & APGCCS, Rajampet, Annamayya (dist) 516216, A.P, India
[1] samramana44@gmail.com; [2] seethakavijay1999@gmail.com

DOI: https://doi.org/10.47760/ijcsmc.2022.v11i09.003

ABSTRACT— Distributed denial of service (DDoS) attack is an attempt to make an online service unavailable by overwhelming it with traffic from multiple sources. Presents a semi-supervised weighted k-means detection method. Hybrid feature selection algorithm to find the most effective feature sets and propose an improved density-based initial cluster centers selection algorithm to solve the problem of outliers and local optimal.
KEYWORDS— Black box testing, Unit testing, Data Flow Diagrams, Semi supervised, k-means algorithm

INTRODUCTION

This project introducing new concept called Semi-Supervised to detect DDOS attack using K-Means and Hybrid features selection algorithm. In existing techniques supervise algorithms require class labels to predict attack and unsupervised clustering technique require no class labels but its detection rate is low and to overcome from above problem author using below concept 1) Hybrid Features Selection Algorithm: in this module author using Hadoop MapReduce algorithm to rank all features and the features with high rank will be selected and using this technique we will drop some attributes from dataset which solve 'curse of dimensionality error'. All features will be normalized before applying features selection. 2) Semi-Supervised Weighted K-Means algorithm: In this algorithm we will assigned weight to each class label so algorithm can predict class label accurately from large dataset. Each class will have weight value and then test data weight values will be calculated and based on matching weight class label will be predicted.

PURPOSE

To develop machine learning semi-supervised weighted k-means detection method. Specially, we present a Hadoop-based hybrid feature selection algorithm to and the most effective feature sets and propose an improved density-based initial cluster centers selection algorithm to solve the problem of outliers and local optimal.

SCOPE

A denial-of-service attack (DoS attack) is a cyber-attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indenitely disrupting services of a host connected to Internet. DoS is typically accomplished by coding the targeted machine or resource with super requests in an attempt to overload systems and prevent some or all legitimate requests from being fuiled. In a distributed denial-of service attack (DDoS attack), incoming traffic coding the victim originates from many different sources.

RELATED WORK

1 Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection

Intrusion detection has become essential to network security because of the increasing connectivity between computers. Several intrusion detection systems have been developed to protect networks using different statistical methods and machine learning techniques. This study aims to design a model that deals with real intrusion detection problems in data analysis and classify network data into normal and abnormal behaviours.

2 Detection and offense mechanism to defend against application layer DDoS attacks

Application layer DDoS attacks, which are legitimate in packets and protocols, gradually become a pressing problem for commerce, politics and military. We build an attack model and characterize layer-7 attacks into three classes: session flooding attacks, request flooding attacks and asymmetric attacks.

3 DDoS detection using modified K-means clustering with chain initialization over landmark window

Denial-of-service is a common form of network attack that affect user access right by preventing legitimate user from accessing certain information, thus giving great, disadvantage to the user and service provider. This paper present a method of denial-of-service detection using clustering technique with k-means algorithm which available to be modified and developed in many possible way.

EXISTING WORK

There are many DDoS attack detection methods, while this paper mainly focuses on machine learning based detection methods. Machine learning methods mainly include unsupervised learning and supervised learning. Unsupervised learning techniques deal with learning tasks with unlabeled or untagged data, and clustering is the most popular unsupervised learning technique K-means algorithm, as a clustering method, has been successfully.

Limitations

- The last decades have witnessed a revolution of Internet usage in many domains like e-commerce, e-government and so on.
- The expansion of the Internet is accompanied by the intensification of security violation issues.

PROPOSED WORK

There are many features used for DDoS detection. While, too many features in the learning process cause ‘‘the curse of dimensionality’’, and unreasonable feature sets lead to poor detection performance. This paper proposes a hybrid feature selection method for DDoS detection. The method includes three steps, namely data normalization, feature ranking and feature subset searching. The input of this method is the candidate feature set and the output is the selected feature set for detection model.

Contribution Work

- Some traffic features in CICDS are ineffectual, leading to degradation of learning quality, more memory consumption and an increase in computational time.
- Feature selection methods can properly solve these issues. In this paper, a machine learning method, RF, is used to collect more important features.

ALGORITHM

1. Hybrid Features Selection Algorithm

   In this module author using to rank all features and the features with high rank will be selected and using this technique we will drop some attributes from dataset which solve ‘curse of dimensionality error’. All features will be normalized before applying features selection

2. Semi-Supervised Weighted K-Means algorithm

   In this algorithm we will assigned weight to each class label so algorithm can predict class label accurately from large dataset. Each class will have weight value and then test data weight values will be calculated and based on matching weight class label will be predicted.

Sample Screens

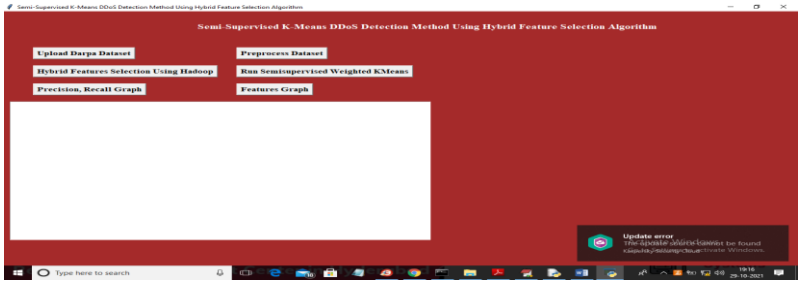

FIG:1

In above screen click on ‘Upload Darpa Dataset’ button to upload dataset and to get below screen

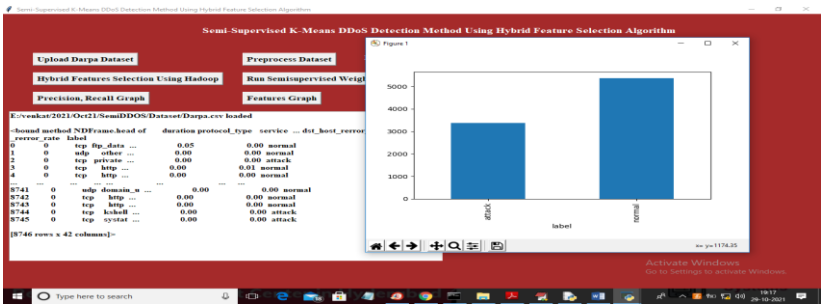

FIG:2

In above screen dataset loaded and we are displaying some values from dataset which contains records as 'normal' and 'attack' and in graph x-axis represents record type as 'normal' and 'attack' and y-axis represents number of records for normal and attack and now close above and then click on 'Pre process dataset' button to replace empty and non-numeric values with numeric data.

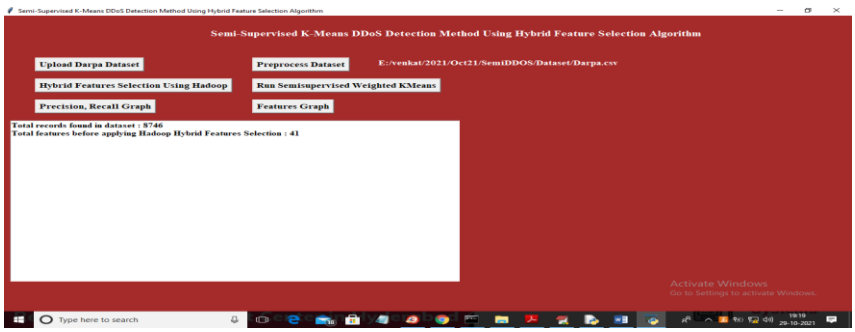

FIG:3

In above screen dataset pre process and we can see dataset contains 41 features before applying hybrid feature selection and now click on 'Hybrid Features Selection Using Hadoop' button to select features from dataset

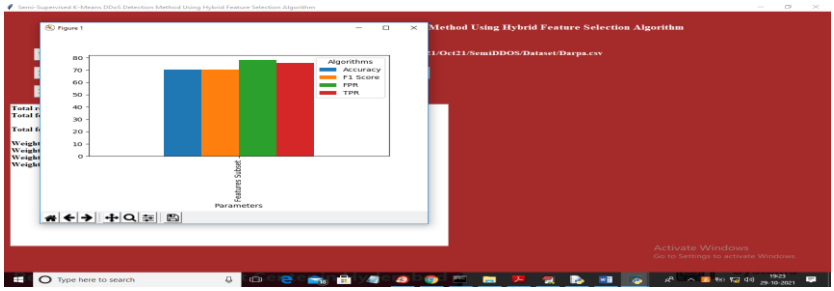

FIG:4

In above screen different bars represents different values such as FPR, TPR, Accuracy and FSCORE and now click on 'Features Graph' button to get below graph.

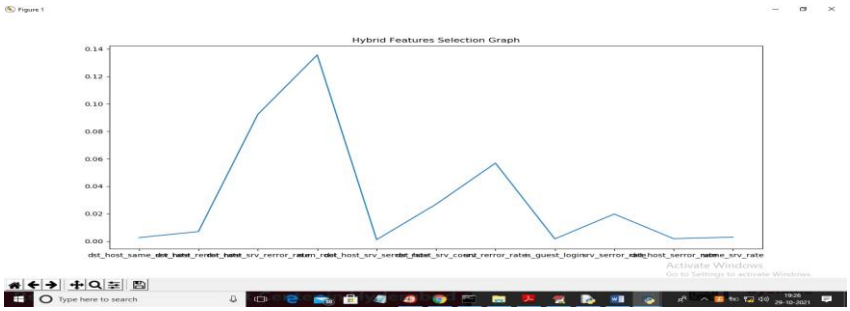

FIG:5

In above graph x-axis represents features names and y-axis represents ranking of that features

Conclusion

In order to tackle the issues of supervised and unsupervised based DDoS detection methods, this paper presents a semi-supervised weighted k-means detection method. Specially, we firstly provide a Hadoop-based hybrid feature selection method to find the most effective feature set. Secondly, we present an improved density-based initial cluster centers selection method to solve the problem of outliers and local optimal of k-means clustering. Then, we propose a semi-supervised weighted k-means method using hybrid feature selection algorithm (SKM-HFS) to achieve better detection performance.

REFERENCES

[1] K. Graves, Ceh: Official certified ethical hacker review guide: Exam 312-50. John Wiley & Sons, 2007.

[2] R. Christopher, "Port scanning techniques and the defense against them," SANS Institute, 2001.

[3] M. Baykara, R. Das¸, and I. Karado ˘gan, "Bilgi g ¨uvenli ˘gi sistemlerinde kullanilan arac¸larin incelenmesi," in 1st International Symposium on Digital Forensics and Security (ISDFS13), 2013, pp. 231–239.

[4] Rashmi T V. "Predicting the System Failures Using Machine Learning Algorithms", International Journal of Advanced Scientific Innovation, vol. 1, no. 1, Dec. 2020, doi:10.5281/zenodo.4641686.

[5] S. Robertson, E. V. Siegel, M. Miller, and S. J. Stolfo, "Surveillance detection in high bandwidth environments," in DARPA Information Survivability Conference and Exposition, 2003. Proceedings, vol. 1. IEEE, 2003, pp. 130–138.

