Available online at: http://www.advancedscientificjournal.com
http://www.krishmapublication.com
IJMASRI, Vol. 1, issue 1, pp. 304-307, Apr. -2025
https://doi.org/10.53633/ijmasri

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY ADVANCED SCIENTIFIC RESEARCH AND INNOVATION (IJMASRI)

ISSN: 2582-9130 **IBI IMPACTFACTOR 1.5** **DOI: 10.53633/IJMASRI**

RESEARCH ARTICLE

DATA MINING APPROACH TO ANALYSING INTRUSION DETECTION OF WIRELESS SENSOR NETWORK

Monisha S[1] N Vennila S[2] and Sathya R[3]

[1]*PG & Research Department, Department of Computer Science, St.Ann's College of Arts and Science, Tindivanam.*
Email: monishasuresh19@gmail.com

[2]*Principal & Head of the Department, PG & Research Department of Computer Science, St.Ann's College of Arts and Science Tindivanam*
Email: sathiyaat@gmail.com

[3]*Assistant Professor & Head, Department of Computer Science, St.Ann's College of Arts and Science Tindivanam.*
Email: moulikrishna19@gmail.com

Abstract

A wireless sensor network (WSN) is a collection of wireless sensor nodes and a base station that are distributed in nature. The distributed nodes are used to monitor and collect physical data of the environment, and then these files are organized into repositories. Its wide range of applications has expanded from critical military applications such as combat surveillance to transportation, medical care, industrial areas and human detection, migration, security and surveillance. Wireless Sensor Networks (WSNs) are vulnerable to various external attacks due to their many features. To prevent these attacks, an Intrusion Detection System (IDS) is essential to prevent attackers from stealing or modifying data. Data mining is a process that helps discover patterns in large amounts of data. This paper presents a data mining approach for various classification algorithms to detect four types of Denial of Service (DoS) attacks. They are Gray Hole, Black Hole, Flood and TDMA. Several data mining techniques such as KNN, Naive Bayes, Logistic Regression, Support Vector Machine (SVM) and ANN algorithms are applied to the data and their performance in finding the output is analyzed. The analysis reveals the applicability of the algorithms in detecting and predicting attacks and can be recommended to cyber experts and analysts

Keywords: Wireless Sensor network, DoS

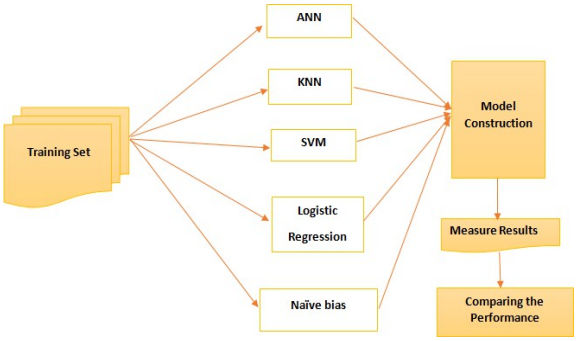
Introduction

Wireless Sensor Networks (WSNs) stands out as one of the newest and largest sites in the world [2-3]. It was initially designed to accelerate and facilitate military operations, but its applications later expanded to include healthcare, transportation, commerce and threat intelligence [4]. Wireless Sensor Networks have special devices and communication systems that use radio waves to monitor and record physical or environmental conditions [5]. It consists of local electronic devices with low detection and transmission rates and is transmitted to the monitoring are a via wireless communication [6]. Sensors are responsible for exchanging environmental information to create models of the monitored area. Each sensor node consumes some power to transmit data overthe network. Therefore, the lifetime of the network depends on the energy consumed in each transmission. To extend the lifetime of wireless sensor networks, application protocols have been developed that reduce the power consumption of sensors. Some of the best-known ones are LEACH, PEGASIS, TEEN, APTEEN, and HEED. LEACH is a cooperative, scalable, and self-managing protocol. The goal of the LEACH protocol is to create clusters of nodes to distribute power among all network nodes. Each policy has a Cluster Head (CH) node that is responsible for collecting data from cluster nodes and sending them to the Base Station (BS) or sink. The location of the BS is far from the sensor nodes. The aim of the LEACH protocol is to reduce the energy consumption required to manage the clusters and thus extend the lifetime of the WSN. The limitations of the sensor nodes, such as processing power, memory and battery life, also affect the security of the wireless sensor networks. Most of the attacks on the Wireless Sensor Networks aim to limit or eliminate the ability of the network to perform its basic tasks. One of the attack types is the DoS attack. This attack is carried out through hardware failures, errors, resource usage, poor broad casting of high- power signals, and degrading network performance. Protection mechanisms are not sufficient to protect network packets and maintain WSN services. They also cannot protect against all attacks on the Internet. Therefore, it would be more effective to combine detection-based technologies with prevention-based technologies. Therefore, intrusion detection tools are indispensable to monitor network connections for suspicious activities and raise alarms when such activities are detected. By accessing the detection capabilities of the system, we can detect attacks early and protect the network against various malicious attacks. Authors inpresented a methodology to investigate four types of DoS attacks in Wireless Sensor Networks. Random forest distributions were used to analyze the attacks on data and achieved the best performance against black holes, floods, gray holes, time division multiple access (TDMA) attacks and Good character. M. Ahsan Latif and M. Adnan in developed three different ANN-based models to explore the behavior of network links. They introduced a model that uses intelligent agents to monitor traffic patterns at the station level. By using appropriate data mining techniques, access to the network can be effectively discovered as it helps in finding patterns in large data sets. Different algorithms then help in classifying and predicting access. Using data mining, we can see the attack patterns and even predict whether an attack is happening when a new request arrives. After using various mining algorithms, the performance of intrusion detection can be analysed and various methods can be compared to determine the best intrusion detection algorithm. And some solutions have been prepared to control and solve some attacks. They classified some DoS attacks according to layer techniques, identifying sink, hello flood, wormhole, selective routing attacks at network layer and flood attacks at transport layer. Luigi Coppolino et al , proposed a unified detection system for Wireless Sensor Networks that uses both error detection and fault detection. They use decision trees as the classification algorithm in the search process .

Methodology

After collecting the data, we did the first thing that would help our experiments. First, replace the attack names with numerical values. The values 0, 1, 2, 3, and 4 are used for normal (non-closing), black hole, gray hole, flood, and TDMA stops, respectively. A binary classification is also performed for attack and non-attack (normal) modes, assigning 1 and 0, respectively.AndcheckthedataApply10-fold cross-validation. Then, we select some of the most frequently used classifications in many data sets and apply them to the data to detect all types of attacks and binary classifications of the inputs. In this paper, we

apply KNN, Naive Bayes, Support Vector Machine (SVM) and Logistic Regression algorithms to see which algorithm can detect more attacks. ANN is also applied to compare with the data in [4].Various measurement models are used to evaluate the obtained results. Analysing the experimental results can reveal important information about the types of algorithms and their suitability in detecting various types of attacks.

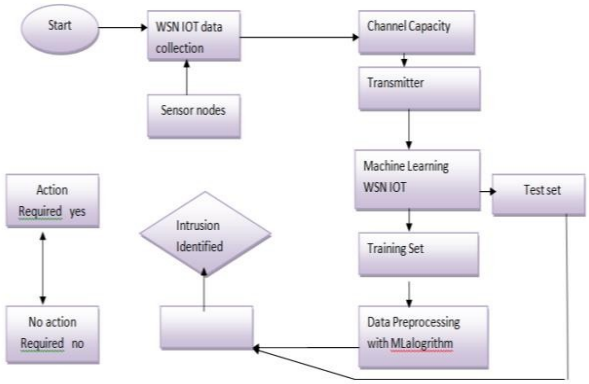


Overview of Dataset

Data is collected from the study in [4]. It is called WSN-DS. In order to create a unique database for WSN access detection, the authors use the LEACH protocol. This IDS file is based on the characteristics of WSN and is suitable for detecting and classifying four types of DoS attacks. This file contains 374661 samples and 23 features. Of these 23 features, RSSI, maximum distance to CH, average distance to CH, and current power are not used to detect DoS attacks. 4.

Implementation

The machine used for this study is a Core i3 CPU at 2 GHz and uses Jupyter workbooks. In order to predict teach attack type, the string data of the attack type features type, the string data of the attack type features are converted to numerical values as specified in the plan and various algorithms are applied to the dataset. In this paper, we first trained the model and then estimated the accuracy from the test data by training the model using the fit () function and extracted the confusion matrix. KNN works well for large datasets. Naive Bayes Classifier is a probabilistic machine learning model. Logistic Regression: It is a distribution algorithm used to estimate the probability of a binary value depending on one or more variables. Implemented Machine Learning Algorithms Four widely used algorithms were used in the experiment. Classification performance is measured using general metrics such as accuracy, precision, recall, F1Score, and error. KNN clustering is done to detect all attacks and not to struggle. In the test data, there are 4077 black hole attacks, KNN can detect 3794 attacks out of 5938 gray hole attacks, 1006 attacks out of 1310 flood attacks, while TDMA can detect 2143 attacks out of 2651 predicted attacks. Out of 135889 normal cases,the detected number is 135323.



Results and Analysis:

This paper describes the use of various machine learning methods on training and testing data. The learning algorithm is implemented in two stages. The first stage is the multi-class classification stage, where all algorithms are used to classify all attacks in the data. In the second stage, a binary classification is performed, treating each attack as one category (attacklevel) and non-attackas another category (normal).

S.No	Attack	Dataset	Percentage
1	Grayholes	14596	3.90%
2	flood	3312	0.88%
3	Blackholes	10049	2.68%
4	TDMA	6638	1.77%
5	Normal	340066	90.77%

There are only 879 black holes, cannot predict gray holes and floods, only2651 attackspredict411 TDMA attacks,135883 normal data are tested from 135889 attacks. 4055 attacks from 4077 black hole attacks, 930 attacks from 5938 Gray Hole attacks can be detected, 905 attacks from 1310 Flood Protection attacks can be detected, 2194 TDMA455 attacks from 2651 predictions,3rd olddata .However ,in most cases, LR performs better.

Conclusion

1. I. F. Akyildiz and M. C. Vuran, "Wireless sensor networks," John Wiley & Sons, vol. 4, 2010.
2. A. H. Farooqi and F. A. Khan, "Intrusion Detection Systems for Wireless Sensor Networks: A Survey,"In International Conference on Future Generation Communication and Networking, Springer, Berlin, Heidelberg, pp. 234–241, 2009.
3. Ghosal, A., & Halder, S., "Intrusion detection in wireless sensor networks: Issues, challenges and approaches,"In Wireless networks and security. Springer, pp. 329-367. 2013.
4. A. H. Farooqi and F. A. Khan, "A surveyof Intrusion Detection Systems for Wireless Sensor Networks," Int. J. Ad Hoc Ubiquitous Computer., vol. 9, no. 2, pp. 69-83, 2012.
5. Ardiansyah, A. Y., & Sarno, R., "Performance analysis of wireless sensor network with load balancing for data transmission using xbee zb module," Indonesian Journal of Electrical Engineering andComputerScience(IJEECS), vol.18,no.1, pp.88-100, 2019.
6. Q. Liao and H. Zhu, "An Energy Balanced Clustering Algorithm Based on LEACH Protocol," Appl.Mech.Mater.,vol.341–342,
   1. pp. 1138–1143,Jul.2013

******