ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

Intrusion Detection Using Machine Learning Algorithm - Decision Tree

Dr.Vijay Reddy Madireddy

Assistant Professor, Swami Ramananda Tirtha Institute of Science and Technology, Nalgonda, India

ABSTRACT: Data is most significant resource for association and they require legitimate administration and assurance. These days PC assault has become exceptionally normal. In spite of the fact that there are many existing systems for Intrusion detection, however the significant issues is the security furthermore, accuracy of the framework. In this paper we research and assess the decision tree information mining procedures as an intrusion detection component. Our exploration shows that Decision trees gives better generally overall performance.

KEYWORDS: Decision tree, machine learning

I. INTRODUCTION

Because of expanded number of internet users there is an issue because of intrusion which might harm data and data put away in PC server or data base server. So we want a channel which can channel malicious data and typical data. Intrusion detection is the most common way of checking and examining the occasions happening in a PC framework inrequest to recognize indications of security issues [1].There are two sorts of intrusion detection procedures: Abuse and Irregularity. Abuse indicators dissect framework movement, searching for occasions or sets of occasions that match a predefined example of occasions that portray a known assault. As the examples comparing to realize assaults are called marks, abuse detection is in some cases called "signature-based detection [2]." Oddity identifiers distinguish strange surprising way of behaving (peculiarities) on a host or organization. They capability with the understanding that assaults are unique in relation to "ordinary" (authentic) movement and can accordingly be identified by frameworks that distinguish these contrasts.To help the preparation and testing the NSL-KDD dataset is utilized, which comprises of various sorts of organization associations marked with the class[3]. A model with high accuracy will be attempted to create .Model will be prepared and tried on the typical and known assaults.

II. SYSTEM ARCHITECTURE

There are many existing instruments for Intrusion detection framework, however the significant issues is the security and accuracy of the framework [4]. To work on the issue of accuracy and the productivity of the system,exceptionally normal order approach for example decision tree is utilized. Proposed research work acquaints a system with foster a classifier based on data mining procedures as displayed in figure.1:

System Implementation:

Proposed research work acquaints a system with foster a classifier based on data mining strategies . In this system NSL-KDD[5] dataset is given to Preprocessing stage which order in C4.5 algorithm furthermore, lessen contemptuous highlights from the data set so that data with less number of component will expect to take care of to the classifier and will give productivity to the classifier.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

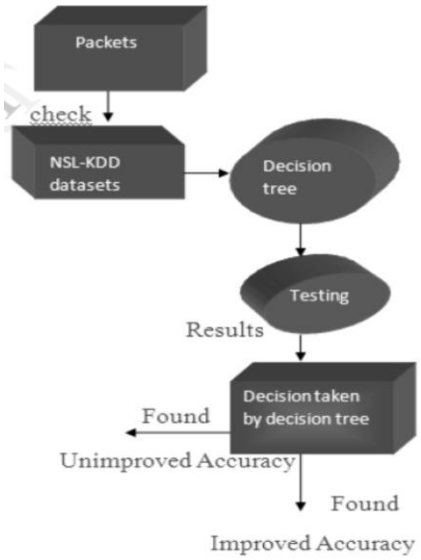

Figure 1:**System Architecture**

III. METHODOLOGY

Procedure continued in this exploration incorporates data sets and characterization strategy C4.5 algorithm . The depiction of these systems are described below.

1. Data Description

NSL-KDD is a dataset proposed to tackle some of innate issues of KDD99 datasets which are referenced. Albeit this new rendition of KDD data set actually experiences some of issues and may not be a delegates of existing genuine organization in light of absence of public dataset for network based intrusion detection system[6].The preparation dataset comprises of 25,192 records and contains 42 qualities and its class is named as either normally or irregularity, in which peculiarity is with precisely one explicit assault type[8]. The assaults types are gathered into four classifications

- DOS: Denial of service - for example syn flooding
- Examining: Reconnaissance and other testing, for example port checking
- U2R: unapproved admittance to neighborhood super client (root)honors, for example support flood assaults
- R2L: unapproved access from a remote machine,for example secret phrase guessing.

2. C4.5 Algorithm

Very much like Classification and Regression Tree, the C4.5 algorithms recursively visits every hub, choosing the ideal split, until no further parts are conceivable[9]. The steps of C4.5 algorithm for growing a decision tree is given below

- Pick property for root node by utilizing quality determination measure Gain Ratio [10].
- Make branch for each worth of that trait Divide cases as indicated by branches[11].
- Rehash process for each branch until all cases in the branch have similar class or all credits are processed.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

IV. RESULTS

For training the system a part of the 20% of NSLKDD dataset is considered which consists of 25, 192 records of the network connection out of which 13449 records are of normal non-malicious category, 01 connections of land, 8282 connections of Neptune, 181 connections of warez client, 710 connections of ipsweep,188 connections of teardrop,587 connections of portsweep,38 connections of pod,10 connections of guess_passwd,301 connections of nmap,691 connections of satan,529 connections of smurf,2 connections of multihop,196 connections of back,1 connections of ftp_ write ,5 connections of imap,2 connections of phf,4 connections of rootkit,7connections of ware master[12]. Once the system has been trained, it can be tested for it's performance. The data sets include whole training set itself, cross validation is applied on the training set, splitting the training dataset and providing a completely different test dataset[13].

Datasets used for testing	Correctly classified instances	Incorrectly classified instances	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC
K=2	99.468%	0.531%	0.995	0.005	0.995	0.995	0.995	0.995
K=4	99.579%	0.420%	0.996	0.004	0.996	0.996	0.996	0.997
K=6	99.555%	0.444%	0.996	0.004	0.996	0.996	0.996	0.997
K=8	99.579%	0.420%	0.996	0.004	0.996	0.996	0.996	0.997
K=10	99.559%	0.440%	0.996	0.004	0.996	0.996	0.996	0.998

Attack Types	Correctly classified instances	Incorrectly classified instances
DOS	99.898%	0.101%
PROBE	90.207%	9.792%
R2L	99.714%	0.285%
U2R	99.918%	0.0817%

Table 2:**Testing the system for all the attack system**

V. CONCLUSION

we have carried out methods for intrusion detection which gives better execution. In this examination we have explored in signature based intrusion detection which recognize just known attacks. Anyway this is one of the significant disadvantage of this framework that it can't recognize obscure and attacks. The future upgrade of this framework is, it eliminates its disadvantage by executing a framework that identify both obscure and known attack.

REFERENCES

[1]. Adithya Vuppula, "OPTIMIZATION OF DATA MINING AND THE ROLE OF BIG DATA ANALYTICS IN SDN AND INTRADATA CENTER NETWORKS" International Journal of Scientific Development and Research (IJSDR), Volume 1 Issue 4, April 2016.
[2]. Kola Vasista, "ROLE OF A STOCK EXCHANGE IN BUYING AND SELLING SHARES", International Journal of Current Science (IJCSPUB), Volume 12, Issue 1, ISSN: 2250-1770.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

[3]. I. Ahmad and K. Pothuganti, "Smart Field Monitoring using ToxTrac: A Cyber-Physical System Approach in Agriculture," International Conference on Smart Electronics and Communication (ICOSEC), pp. 723-727, doi: 10.1109/ICOSEC49089.2020.9215282.
[4]. Satya Nagendra Prasad Poloju, "An Overview on Cloud Computing Technologies", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 4, Issue 10, October 2015.
[5]. Ramana, solleti, "A Two-Level Authentication Protocol for Secure M-Commerce Transactions using AMQP Protocol – Design Engineering, Issue: 6, ISSN Number 0011-9342
URL:http://www.thedesignengineering.com/index.php/DE/article/view/2047

[6]. Peddyreddy. Swathi. A Study On The Restrictions Of Deep Learning. Journal of Artificial Intelligence,Machine Learning and Neural Network (JAIMLNN) ISSN: 2799-1172, 2(02), 57–61. Retrieved from http://journal.hmjournals.com/index.php/JAIMLNN/article/view/444
[7]. Kola Vasista, "TYPES AND RISKS INVOLVED TOWARDS INVESTING IN MUTUAL FUNDS", International Journal of Current Science (IJCSPUB), Volume 12, Issue 1, ISSN: 2250-1770.
[8]. Peddyreddy. Swathi. Industry Applications of Augmented Reality and Virtual Reality. Journal of Environmental Impact and Management Policy(JEIMP) ISSN:2799-113X, 2(02), 7–11. Retrieved from http://journal.hmjournals.com/index.php/JEIMP/article/view/453
[9]. Satya Nagendra Prasad Poloju, "DATA MINING AS A SUPPORT FOR BUSINESS INTELLIGENCE APPLICATIONS TO BIG DATA" , International Journal of Creative Research Thoughts (IJCRT), Volume 7, Issue 2.
[10]. S. Ramana, S. C. Ramu, N. Bhaskar, M. V. R. Murthy and C. R. K. Reddy, "A Three-Level Gateway protocol for secure M-Commerce Transactions using Encrypted OTP," International Conference on Applied Artificial Intelligence and Computing (ICAAIC), pp. 1408-1416, doi: 10.1109/ICAAIC53929.2022.9792908.
[11]. K. Pothuganti, B. Sridevi and P. Seshabattar, "IoT and Deep Learning based Smart Greenhouse Disease Prediction," International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), pp. 793-799, doi: 10.1109/RTEICT52294.2021.9573794.
[12]. Peddyreddy. Swathi. Implications For Research In Artificial Intelligence. Journal of Electronics,Computer Networking and Applied Mathematics(JECNAM) ISSN : 2799-1156, 2(02), 25–28. Retrieved from http://journal.hmjournals.com/index.php/JECNAM/article/view/447
[13]. Adithya Vuppula, "A Study on Minnesota Intrusion Detection System (Minds)" International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), Volume 1, Issue 1, November.
[14]. Kola Vasista, "A REVIEW ON THE VARIOUS OPTIONS AVAILABLE FOR INVESTMENT", International Journal of Creative Research Thoughts (IJCRT), Volume 7, Issue 2, ISSN: 2320-2882.
[15]. Satya Nagendra Prasad Poloju, "BIG DATA ANALYTICS: DATA PRE-PROCESSING, TRANSFORMATION AND CURATION", International Journal of Creative Research Thoughts (IJCRT), Volume 5, Issue 2 ,2017
[16]. Kola Vasista, "Regulatory Compliance and Supervision of Artificial Intelligence, Machine Learning and Also Possible Effects on Financial Institutions", International Journal of Innovative Research in Computer and Communication Engineering, Volume 9, Issue 6, June .
[17]. K. Pothuganti, B. Sridevi and P. Seshabattar, "IoT and Deep Learning based Smart Greenhouse Disease Prediction," International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), pp. 793-799, doi: 10.1109/RTEICT52294.2021.9573794.
[18]. Adithya Vuppula, "EFFICIENCY AND SCALABILITY OF DATA MINING ALGORITHMS", International Journal of Scientific Development and Research (IJSDR), Volume 4 Issue 9, September 201.
[19]. Kola Vasista, "Scope for the Usage of Ai and Machine Learning in Portfolio Management and Possible Effects on Consumers and Investors", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, Issue 2, February 2016
[20]. Ramana, solleti, A Two-Level Protocol for Secure Transmission of Image using IOT Enabled devices Webology, Volume 18, Issue 5, ISSN Number: 1735-188X
URL: https://www.webology.org/abstract.php?id=2194
[21]. Satya Nagendra Prasad Poloju, "Privacy-Preserving Classification of Big Data", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 2, Issue 4, April 2013.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

[22]. Adithya Vuppula, "Integrating Data Mining with Cloud using Four Levels of Data Mining Services" International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), ISSN: 2582-7219, Volume 4, Issue 5, May 202.

[23]. Satya Nagendra Prasad Poloju. "Relevant Technologies of Cloud Computing System". International Journal of Engineering Research and Applications, ISSN: 2248-9622, Vol. 4, Issue 4, (Version-3) April 2014, pp. 74-78

[24]. Ramana, solleti, "A Two-Level Authentication Protocol for Secure M-Commerce Transactions using Encrypted OTP– International Journal of Mechanical Engineering, Volume 7, Issue: 3, ISSN Number 0974-5823