International Journal of Science Engineering and Advance Technology

A Supervised Filter Based Feature Selection Algorithm For Intrusion Detection-FMIFS

D.S.S.Madhuri[*1], R.Anusha[*2]
[1]M.Tech Student, [2]Assistant Professor
Dept of CSE , Srinivasa Institute of Engineering and Technology,Amalapuram, AP.
dssmadhuri@gmail.com, anusha.rudraraju@gmail.com

ABSTRACT:

We propose a shared data based algorithm that analytically chooses the optimal feature for classification. This common information based component choice calculation can manage straightly and nonlinearly subordinate data features. Its reasonability is surveyed in the cases of framework intrusion distinguishing proof. An Intrusion Detection System (IDS), named Least Square Support Vector Machine based IDS (LSSVM-IDS), is manufactured using the segments picked by our proposed include choice calculation. The execution of LSSVM-IDS is surveyed using three interruption location appraisal datasets, to be particular KDD Cup 99, NSL-KDD and Kyoto 2006+ dataset.

KEYWORDS: Detection System, dataset,classification.

I. INTRODUCTION:

A lot of research has been led to create intelligent intrusion detection systems, which help accomplish better system security. Packed away boosting-in light of C5 decision trees [2] and Kernel Miner [3] are two of the most punctual endeavors to build intrusion detection plans. Strategies proposed in [4] and [5] have effectively connected machine learning systems, for example, Support Vector Machine (SVM), to group organize activity designs that don't coordinate normal network traffic. The two frameworks were outfitted with five distinct classifiers to identify normal traffic and four unique sorts of attacks (i.e., DoS, testing, U2R and R2L). Exploratory outcomes demonstrate the viability and robustness of utilizing SVM in IDS.

LITERATURE SURVEY:

[1],this tends to utilizing an outfit approach of various delicate computing and hard registering methods for intrusion detection. Because of expanding episodes of digital attacks, building viable intrusion detection frameworks are fundamental for ensuring data frameworks security, but then it remains a subtle objective and an awesome test. We considered the execution of Artificial Neural Networks (ANNs), Support Vector Machines (SVMs) and Multivariate Adaptive Regression Splines (MARS).

[2],this proposes a technique for applying Support Vector Machines to network-based Intrusion Detection System (SVM IDS). Support vector machines(SVM) is a learning procedure which has been effectively connected in numerous application ranges. Intrusion detection can be considered as two-class grouping issue or multi-class order issue. We utilized dataset from 1999 KDD Intrusion detection challenge. SVM IDS was found out with triaing set and tried with test sets to assess the execution of SVM IDS to the novel attacks. What's more, we likewise assess the significance of each element to enhance the general execution of IDS.

PROBLEM DEFINITION

A lot of research has been directed to create intelligent intrusion detection methods, which help accomplish better system security. Packed away boosting-in light of C5 choice trees and Kernel Miner are two of the most punctual endeavors to construct intrusion detection plans.

Mukkamala et al. examined the likelihood of collecting different learning strategies, including Artificial Neural Networks (ANN), SVMs and Multivariate Adaptive Regression Splines (MARS) to detect intrusions.

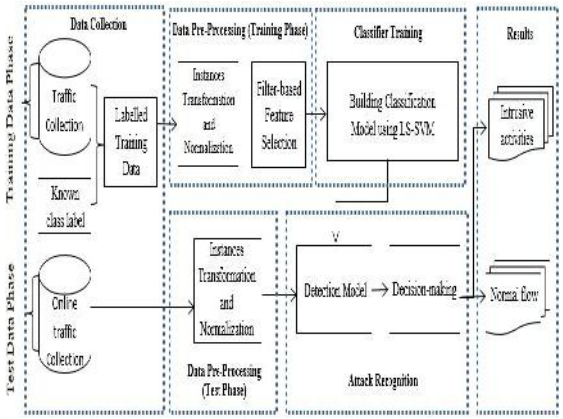
PROPOSED APPROACH

We have proposed a hybrid feature selection algorithm (HFSA). HFSA comprises of two stages.

The upper stage directs a preparatory hunt to dispense with insignificant and repetition highlights from the first information. This helps the wrapper strategy (the lower stage) to diminish the looking extent from the whole unique component space to the pre-chosen highlights (the yield of the upper stage). The key commitments of this paper are recorded as takes after.

This work proposes another channel based component determination technique, in which hypothetical examination of shared data is acquainted with assess the reliance between

elements and output classes.**SYSTEMBLOCK DIAGRAM:**



PROPOSED METHODOLOGY:

DATA PRE-PROCESSING:

The information acquired amid the period of information gathering are first prepared to create the essential highlights, for example, the ones in KDD Cup 99 dataset. The prepared classifier requires each record in the information to be spoken to as a vector of genuine number. Hence, every emblematic element in a dataset is first changed over into a numerical esteem. For instance, the KDD CUP 99 dataset contains numerical and in addition emblematic highlights.

FILTER BASED FEATURE SELECTION:

Obviously, a straight measure can't uncover the connection between two nonlinearly subordinate factors. Consequently, we require a measure equipped for dissecting the connection between two factors regardless of whether they are directly or nonlinearly needy. Therefore, this work expects to investigate a methods for choosing ideal highlights from a component space paying little heed to the kind of relationship between's them.

ATTACK CLASSIFICATION & RECOGNITION:

In general, it is simpler to build a classifier to distinguish between two classes than considering multiclasses in a problem. This is because the decision boundaries in the first case can be simpler. The first part of the experiments in this paper uses two classes, where records matching to the normal class are reported as normal data, otherwise are considered as attacks.

PERFORMANCE EVALUATION:

The majority of the IDS experiments were performed on the KDD Cup 99 datasets. In addition, these datasets have different data sizes and various numbers of features which provide comprehensive tests in validating feature selection methods.

ALGORITHM:

MUTUAL INFORMATION BASED FEATURE SELECTION:

INPUT:FSET,f1,f2,S

STEP1: starting initlize the feature set empty.

STEP2: calculate the each feature based on all features.

STEP3: selecting the relevent features based on mutual dependency.

STEP4: sorting the features according to relevent features.

STEP5: returns features subset.

INTRUSION DETECTION USING LEAST SQUARE SUPPORT VECTOR MACHINE:

INPUT:CLASSIFIER,SUBSET FEATURE,DATA ITEMS

STEP1: classification of data items are stores it in class label variable.

STEP2:if class label variable==”normal” then

STEP3: return normal label

STEP4: else run attack classification algorithm to find class of attack.

ATTACK CLASSIFICATION ALGORITHM:

INPUT: CLASSIFIER,SUBSET FEATURE,DATA ITEMS

STEP1: classification label stored in class variable.

STEP2:if class variable==”dos” then

STEP3: return dos atack classification variable.

STEP4:else

STEP5: class variable==”probe” then

return probe atack classification variable

STEP6:else

Classification of data items based on R2L class

STEP7:if class variable==”R2L” then

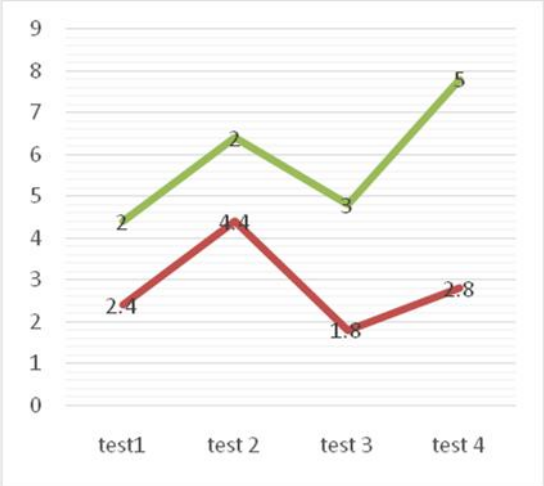
return r2l atack classification variable

else

return U2R atack classification variable

STEP8:classification labels for all data items.

RESULTS:



This demonstrates the proposed approach shows compelling execution as far as security and correspondence and calculation overhead contrasted with before strategy.

EXTENSION WORK:

Including a more efficient deep neural network based on a three-layer stacked auto-encoders for network traffic data. For each auto encoder we design a specific and effective learning strategy to capture useful features from input data.

CONCLUSION:

The proposed LSSVMIDS + FMIFS has been assessed utilizing three well known intrusion detection datasets: KDD Cup 99, NSL-KDD and Kyoto 2006+ datasets. The execution of LSSVM-IDS + FMIFS on KDD Cup test information, KDDTest+ and the information, gathered on 1, 2 and 3 November 2007, from Kyoto dataset has shown better arrangement execution regarding classification precision, detection rate, false positive rate and F-measure than a portion of the current detection approaches.

REFERENCES:

[1] S. Pontarelli, G. Bianchi, S. Teofili, Traffic-aware design of a high speed fpga network intrusion detection system, Computers, IEEE Transactions on 62 (11) (2013) 2322–2334.

[2] B. Pfahringer,Winning the kdd99 classification cup: Bagged boosting, SIGKDD Explorations 1 (2) (2000) 65–66.

[3] I. Levin, Kdd-99 classifier learning contest: Llsoft's results overview, SIGKDD explorations 1 (2) (2000) 67–75.

[4] D. S. Kim, J. S. Park, Network-based intrusion detection with support vector machines, in: Information Networking, Vol. 2662, Springer, 2003, pp. 747–756.

[5] A. Chandrasekhar, K. Raghuveer, An effective technique for intrusion detection using neuro-fuzzy and radial svm classifier, in: Computer Networks & Communications (NetCom), Vol. 131, Springer, 2013, pp. 499–507.

[6] S. Mukkamala, A. H. Sung, A. Abraham, Intrusion detection using an ensemble of intelligent paradigms, Journal of network and computer applications 28 (2) (2005) 167–182.

[7] A. N. Toosi, M. Kahani, A new approach to intrusion detection based on an evolutionary soft computing model using neurofuzzy classifiers, Computer communications 30 (10) (2007) 2201–2212.

[8] Z. Tan, A. Jamdagni, X. He, P. Nanda, L. R. Ping Ren, J. Hu, Detection of denial-of-service attacks based on computer vision techniques, IEEE Transactions on Computers 64 (9) (2015) 2519–2533.

[9] A. M. Ambusaidi, X. He, P. Nanda, Unsupervised feature selection method for intrusion detection system, in: International Conference on Trust, Security and Privacy in Computing and Communications, IEEE, 2015.

[10] A. M. Ambusaidi, X. He, Z. Tan, P. Nanda, L. F. Lu, T. U. Nagar, A novel feature selection approach for intrusion detection data classification, in: International Conference on Trust, Security and Privacy in Computing and Communications, IEEE, 2014, pp. 82–89.

[11] R. Battiti, Using mutual information for selecting features in supervised neural net learning, IEEE Transactions on Neural Networks 5 (4) (1994) 537–550.

[12] F. Amiri, M. RezaeiYousefi, C. Lucas, A. Shakery, N. Yazdani, Mutual information-based feature selection for intrusion detection systems, Journal of Network and Computer Applications 34 (4) (2011) 1184–1199.

[13] A. Abraham, R. Jain, J. Thomas, S. Y. Han, D-scids: Distributed soft computing intrusion detection system, Journal of Network and Computer Applications 30 (1) (2007) 81–98.

[14] S. Mukkamala, A. H. Sung, Significant feature selection using computational intelligent techniques for intrusion detection, in: Advanced Methods for

Knowledge Discovery from Complex Data, Springer, 2005, pp. 285–306.

[15] S. Chebrolu, A. Abraham, J. P. Thomas, Feature deduction and ensemble design of intrusion detection systems, Computers & Security 24 (4) (2005) 295–307.

D.S.S. Madhuri is a student of Srinivasa Institute of Engineering and Technology, Cheyyeru. Presently she is pursuing her M.Tech [Computer Science And Engineering] from this college. She received her B.Tech from RGUKT Nuzvid.

R.Anusha working as Assistant Professor in the Department of CSE in Srinivasa Institute of Engineering and Technology, Cheyyeru affiliated to JNTUK. She received her M.Tech from SRKR Engineering College, Bhimavaram.