INTRUSION DETECTION SYSTEM-VIA FUZZY ARTMAP IN ADDITION WITH ADVANCE SEMI SUPERVISED FEATURE SELECTION

Swati Sonawale and Roshani Ade

Department of Computer Engineering, Savitribai Phule Pune University, India

ABSTRACT

Outstanding to the promotion of the Internet and local networks, interruption occasions to computer systems are emerging. Intrusion detection systems are becoming progressively vital in retaining appropriate network safety. IDS is a software or hardware device that deals with attacks by gathering information from a numerous system and network sources, then evaluating signs of security complexities. Enterprise networked systems are unsurprisingly unprotected to the growing threats posed by hackers as well as malicious users inside to a network. IDS technology is one of the significant tools used now-a-days, to counter such threat. In this research we have proposed framework by using advance feature selection and dimensionality reduction technique we can reduce IDS data then applying Fuzzy ARTMAP classifier we can find intrusions so that we get accurate results within less time. Feature selection, as an active research area in decreasing dimensionality, eliminating unrelated data, developing learning correctness, and improving result unambiguousness.

KEYWORDS

Feature Selection, Intrusion Detection, Redundancy, Fuzzy ARTMAP

1. INTRODUCTION

In today's computing world security is main issue. There are many solutions available to protect the network structure and communication over the Internet, amongst them we can use firewalls, encryption, and virtual private networks. Intrusion detection is advance step for these methods.

IDS protect a system from various types of attack, misuse, and compromise. It is also used for observing activity of networks. Network traffic watching and measurement is gradually considered as a critical task for understanding, increasing performance & safety of various cyber organization. Feature base uses plays an important role in reducing development lead periods improving quality of product [1].

By using dimensionality reduction technique we can achieve efficiency as it will reduce data Hence, in a number of cases it can be useful or even necessary to first reduce the dimensionality

of the records to a convenient size, by keeping as it original info as probable, and then give this reduced dimension data as a input to system [2].

In formativeness coherence that capture relevant properties of the constraint sets coherence measure is independent of any learning algorithm. Training time for learning fuzzy art map classifier is less as compare to other classifier [3].

By using this technique we can detect intrusions to find whether there is attack or not in system as information is valuable asset for any organization it can cause millions of harm for any organization within a few seconds [4].

1.1 INTRUSION DETECTION

An intrusion detection system (IDS) inspects all inbound and outbound network activity and identifies suspicious patterns that may indicate a network or system attack from someone attempting to break into or compromise a system. IDS is a system which is designed for security purpose. It will protect computers or network of computer. It collects information from numerous sources to detect security breaches. IDS generates various types of alert. Alert can be From change in log files to email messages or sms[5].

There are several ways to categorize an IDS:

1.1.1. Misuse detection vs. Anomaly detection

In misuse detection, the IDS analyses the information it gathers and compares it to large databases of attack signatures. Essentially, the IDS looks for a specific attack that has already been documented. Like a virus detection system, misuse detection software is only as good as the database of attack signatures that it uses to compare packets against. In anomaly detection, the system administrator defines the baseline, or normal, state of the networks traffic load, breakdown, protocol, and typical packet size. The anomaly detector monitors network segments to compare their state to the normal baseline and look for anomalies [6].

1.1.2. Network-based vs. Host-based system

In a network-based system, or NIDS, the individual packets flowing through a network are analysed. The NIDS can detect malicious packets. They work basically on a “eavesdropping idea,” information is collected from the network traffic stream, as data travels on the network segment [7].

Drawback of NIDS is that it is unable to scan protocol if data is in other formats means if data is in encrypted format. There are also problems in case of networks having high bandwidth. It is hard for implementation.

In a host-based system, the IDS examines at the activity on each individual computer or host [8].Host-based intrusion detection systems are required for numerous causes because host-based systems can screen access to information in terms of “who accessed what,” these systems can trace spiteful or incorrect actions to a particular user ID. It has been observed that due to HIDS false positive rate has been decreased [9].

One of the drawback of host base system is portability. This is because it run on a single platform so it must be compatible with the system on which it is running and it cannot monitor network traffic because it is not design for this purpose. They uses local resources. HIDS usually includes an agent fixed on every system, monitoring and informing on local OS and application action [10].

1.1.3. Passive system vs. Reactive system

In a passive system, the IDS identifies a possible security break, logs the data and signals an aware. It is designed in such way that it will give only alert. It cannot perform any function on its own [11].

In a reactive system, the IDS replies to the doubtful activity by logging off a user or by reprogramming the firewall to block network traffic from the doubted malicious cause [12]. However they both related to network security, an IDS contrasts from a firewall in that a firewall looks out for intrusions in order to stay them from happening. The firewall bounds the access amongst networks in order to stop intrusion and does not signal an attack from inside the network. An IDS estimates a doubted intrusion once it has taken place and signals an alarm. An IDS also watches for attacks that originate from inside of a system [13].

2. BACKGROUND

2.1 FEATURE SELECTION

Feature selection (FS) methods have commonly been used as a main way to select the relevant features. Author has proposed a novel unsupervised FS method, i.e., locality and similarity preserving embedding (LSPE) for feature selection. Specifically, the nearest neighbour graph is firstly constructed to preserve the locality structure of data points, and then this locality structure is mapped to the reconstruction coefficients such that the similarity among these data points is preserved. Moreover, the sparsity derived by the locality is also preserved. Finally, the low dimensional embedding of the sparse reconstruction is evaluated to best preserve the locality and similarity [14].

Energy efficiency is a key issue in wireless sensor networks where the energy resources and battery capability are very inadequate. In this author has introduced a new pattern recognition based creation for reducing the energy usage in wireless sensor networks. The proposed scheme encompasses an algorithm to rank and select the sensors from the most important to the least, and followed by a nave Bayes classification.[15].

An efficient fuzzy classifier is depend on the ability of feature selection based on a fuzzy entropy technique. It is used to evaluate the information of pattern distribution in pattern space. With this info, we can divide the pattern space into non overlapping decision regions for pattern classification. Then the decision sections do not overlap, both the difficulty and computational burden of the classifier is decreased and thus the training and classification time are very short. Although the judgment areas are divided into non overlapping subspaces, we can obtain good classification performance meanwhile the decision areas can be properly identified via proposed fuzzy entropy method. The feature selection procedure reduces the dimensionality of a problem as well as rejects noise-corrupted, redundant and insignificant features [16].

Today it is very vital need of provision of a high level security to guard highly sensitive and confidential information. In Network Security Intrusion Detection System is an essential technology. Nowadays scientists have interested on intrusion detection system using Data mining procedures as an devious skill. IDS is a software or hardware device that deals with attacks by gathering information from a variation of system and network sources, then analysing symptoms[17] The main aim of feature selection is to select number of features from the set of thousands of features. It is applied to reduce the dimensionality & improving learning performance [18].

Feature selection has been active research area in pattern recognition statistics and data mining community. The main idea of feature selection is to remove redundant features select features which are needed. Feature selection can considerably increase the clarity of the resulting classifier models and repeatedly form a model that simplifies better to unseen points. Further, it is frequently the event that identifying the particular subset of predictive features is an significant problem in its own way [19].

In supervised learning Feature selection has been well measured, where the key aim is to discover a feature subset that forms greater classification accuracy have planned feature selection and clustering together with a only or joint measure[20].

For feature selection in unsupervised learning, learning algorithms are aimed to find natural grouping of the instances in the feature space. Thus feature selection in unsupervised learning aims to find a good subclass of features that is needed to build high value of clusters from a given number of clusters. Though, the traditional approaches to feature selection with particular evaluation criterion have displayed limited capability in terms of knowledge detection and judgment provision. This is as decision makers should take into consideration multiple, conflicted ideas immediately. In particular no single condition for unsupervised feature selection is best for each use and simply the decision maker can choose the relative weights of criteria for any application [21].

Algorithms for feature selection fall into two broad classes namely wrappers that use the learning algorithm itself to estimate the effectiveness of features and filters that calculate features according to heuristics based on overall features of the data. In machine learning and statistics, feature selection, also well-known as variable selection, attribute selection or variable subset selection, is the method of selecting a subset of related features are used in model creation. The principal hypothesis while using a feature selection method is that the data contains several duplicate or unrelated features. Redundant features are those inappropriate features deliver useless info in any perspective. Various feature selection techniques are consider as a subdivision of the additional universal field of feature extraction Feature extraction creates new features from purposes of the unique features, whereas feature selection yields a subclass of the features. Feature selection techniques are frequently used in areas where there are numerous features and relatively limited samples [22].

The typical case is the usage of feature selection in examining DNA microarrays, where there are several thousands of features, and a less number of samples. Feature selection method offers three key advantages when making analytical models [23].

Feature selection is used to eliminate irrelevant and redundant features, which improves prediction accuracy and reduces the computational overhead in classification. This paper presents comparison of 3 methods namely fast correlation based feature selection (FCBF), Multi thread based FCBF feature selection and decision dependent -decision independent correlation (DDC-DIC). These approaches are concerning the relevance of the features and the pair wise features correlation for redundancy checking in order to improve the prediction accuracy and reduce the computation time. The experimental results are tested in weka tool for C4.5 decision tree construction algorithm, which provide better performance for lung cancer, Tic 2000 Insurance company data and breast cancer data sets [24].

Features are representative characteristics of data sets. Identifying such features in a high dimensional dataset play an important role in real world applications. Data mining is best used to determine important features. Selecting important features from a subject of identified features can help in making expert decisions. However, efficient identification of such feature subset and selection is a challenging problem. Recently Song et al. proposed a solution that is capable of selecting subset of features with good quality. They used clustering approach before selecting representative features for final selection. Similar work is carried out in this paper which demonstrates the proof of concept. The proposed solution makes use of clustering for achieving the goal of the system. The empirical results reveal that the application is useful. The results are compared with many existing algorithms like C4.5, Naïve Bayes, IB1 and RIPPER [25].

2.2 INTRUSION DETECTION SYSTEM

Today it is very vital need of provision of a high level security to guard highly sensitive and confidential information. In Network Security Intrusion Detection System is an essential technology. Nowadays scientists have interested on intrusion detection system using Data mining procedures as an devious skill. IDS is a software or hardware device that deals with attacks by gathering information from a variation of system and network sources, then analysing symptoms. An overview of intrusion detection system introduces the reader to some fundamental concepts of IDS methodology. The primary intrusion detection techniques has been discussed. In this we focus on data mining algorithms for implementation of IDS such as Support Vector Machine It contains Kernelized of security problems support vector machine, Extreme Learning Machine and Kernelized Extreme Learning Machine [26].

The Internet environment has become more complex and untrusted Over the past several years,. Enterprise networked systems are unsurprisingly showing to the increasing threats posed by hackers as well as malicious users' internal to a network. IDS technology is one of the important tools used now-a-days, to counter such threats. Various IDS techniques has been proposed, which identifies and alarms for such threats or attacks. Data mining provides a wide range of techniques to classify these attacks. The author provides a comparative study on the attack detection rate of these existing classification techniques [27].

Present intrusion detection methods mostly effort on determining irregular system measures in computer networks and distributed communication schemes. Clustering methods are usually used to decide a probable attack. Due to the doubt nature of intrusions, fuzzy sets show an significant part in identifying hazardous actions and decreasing false alarms level. The author suggests a dynamic method that attempts to determine well-known or unfamiliar intrusion forms. A dynamic fuzzy boundary is established from labelled data for dissimilar levels of security necessities [28].

We introduces the network intrusion detection system (NIDS), which uses a suite of data Mining techniques to automatically detect attacks against computer networks and systems. It consist of two specific contributions:

1. An unsupervised anomaly detection technique that assigns a score to each network connection that reflects how anomalous the connection is, and
2. An association pattern analysis based module that summarizes those network connections that are ranked highly anomalous by the anomaly detection module. Experimental results show that our anomaly detection techniques are successful in automatically detecting several intrusions that could not be identified using popular signature-based tools .Furthermore, given the very high volume of connections observed per unit time, association pattern based summarization of novel attacks is quite useful in enabling a security analyst to understand and characterize emerging threats[29].

Proposed Intrusion Detection System is an intrusion detection system (IDS) proposed by analysing the principle of the intrusion detection system based on host and network. Here we are concentrating and analysing overall performance as well as security of the proposed IDS. Moreover the proposed IDS approve the effectiveness of the proposed method, and presented results shows advantages of host based as well as network based security. The proposed model of hybrid IDSs offers several advantages over alternative systems. First of all it provided higher security, it supported high availability and scalability, and most important thing it produced good results in terms of normal and abnormal behaviours of captured packet. The proposed model includes integration of individual components to produced better results [19].Author has proposed novel approach for intrusion detection and diagnosis. The proposed approach uses Sequential Backward Floating Search for feature selection and fuzzy ARTMAP for detection and diagnosis of attacks. The optimal vigilance parameter for the fuzzy ARTMAP is chosen using a genetic algorithm. Due to this reduced features computation time is saved by 0.789 s[30].

The reliance on information technology became serious and IT infrastructure, sensitive data, intangible intellectual property are susceptible to attacks threats. Organizations mount Intrusion Detection Systems (IDS) to alert suspicious traffic or action. IDS produce a big number of alerts and most of them are false positive as the activities interpret for partly attack pattern or lack of background knowledge. Monitoring and identifying risky alerts is a major concern to security administrator. The present diagnosis stage, respectively work is to design an operational model for minimization of false positive alarms, including periodic alarms by security administrator. The construction, design and performance of model in reduction of false positives in IDS are discovered [31,32].

The Intrusion detection system deals with huge amount of data which contains irrelevant and redundant features causing slow training and testing process, higher resource consumption as well as poor detection rate. Feature selection, therefore, is an important issue in intrusion detection. In this paper we introduce concepts and algorithms of feature selection, survey existing feature selection algorithms in intrusion detection systems, group and compare different algorithms in three broad categories: filter, wrapper, and hybrid. We conclude the survey by identifying trends and challenges of feature selection research and development in intrusion detection system[33].

3. ALGORITHIM

3. 1 CONSTRAINT SELCTION ALGORITHIM

Step 1: Select Dataset from System.
D = {KDD99, WAVE,}
Step 2: Create and Initialize Instances.
X =x1, x2,…, xN}
Step 3: Divide instances into Categories i.e. Labeled Instance and
Unlabeled Instances.
XL = {x1, x2,…, xN}
XU = {x1, x2,…, xN}
Step 4: Create and Initialize Constraints.
Ω= {(x1, x2), (x1, x3) … (L*(L-1)/2) }
Step 5: Divide constraints into Categories i.e. Must-Link Constraints and Cannot-Link Constraint.
Ω'ML = {Must-Link Constraints}
Ω'CL = {Cannot -Link Constraints}
Step 6: Calculate Global coherence of constraint Coh(Ω)
Step 7: for i=1 to Ω
if Coh(Wi) >= Coh(Ω) then
ΩS = ΩS U {Wi}
end if
end for
Step 8: Selected Constraints are
ΩS = Ω'ML U Ω'CL

3.2 FUZZY ARTMAP ALGORITHM

1. Initially, all neuron values should be normalized to guarantee that they are in the range 0- Encode the vectors of ARTa and ARTb modules.
2. Initialize the weights and parameters of ARTa, ARTb and Inter-ART.
3. Choose the category for ARTa and ARTb. If more than one module is active, take the one that has the highest ordering index.
4. Test the vigilance of ARTa and ARTb. If the vigilance criterion is met, then the resonance (Match) takes place.
5. Match tracking between ARTa and ARTb: check if there was matching between the input and output. If not, search another index that satisfies it.
6. Repeat steps 4 through 7 for every pair of vectors to be trained. Algorithm Of Fuzzy ARTMAP [20].

3.3 FAST ALGORITHIM

INPUT: D (F_1,F_2.....F_m,C)
Θ=Threshold
OUTPUT:S=Set of selected feature subset

Step 1:

1. For i=1 to m
2. T-relevance=SU(F_i ,C)
3. If T-relevance > Θ then
4. S=S U { F_i }
5. Generate spanning tree using prims algorithm
6. Make partition of tree to select representative features.

Step 2: We get selected feature set.

4. FUZZY ARTMAP NEURAL NETWORKS

ART stands for "Adaptive Resonance Theory" developed by Stephen Grossberg in 1976.ART signifies family of neural networks. Real world is faced with a situations where data is continuously changing. In such situation, every learning system faces plasticity-stability dilemma. This dilemma is about :"A system that must be able to learn to adapt to a changing environment (i.e. it must be plastic) but the constant change can make the system unstable, because the system may learn new information only by forgetting everything it has so far learned[34].

This phenomenon, a contradiction between plasticity and stability, is called plasticity - stability dilemma. The back-propagation algorithms suffer from such stability problem. The basic ART system is unsupervised learning model.

It typically consists of

- Comparison field and a recognition field composed of neurons
- Vigilance parameter, and
- Reset module

In comparison field it takes an input vector (a one-dimensional array of values) and transfers it to its best match in the recognition field. Its best match is the single neuron whose set of weights most closely matches the input vector.

Neural networks have been used widely for identifying intrusions in computer networks , which confirms that the paradigm of learning by sampling in training IDSs are becoming more and more popular these days. There are various methods available for learning classifier such as incremental learning given in [37,38].

In particular, the fuzzy ARTMAP neural network represents a valuable supervision learning system that classifies input data into stable categories to respond to random input patterns [35].

Fig. 2 depicts the architecture of the fuzzy ARTMAP neural network. It comprises two modules: fuzzy ARTa and fuzzy ARTb. Both modules use the same structure of the neural network. These two modules are interconnected by a third module called inter-ART.

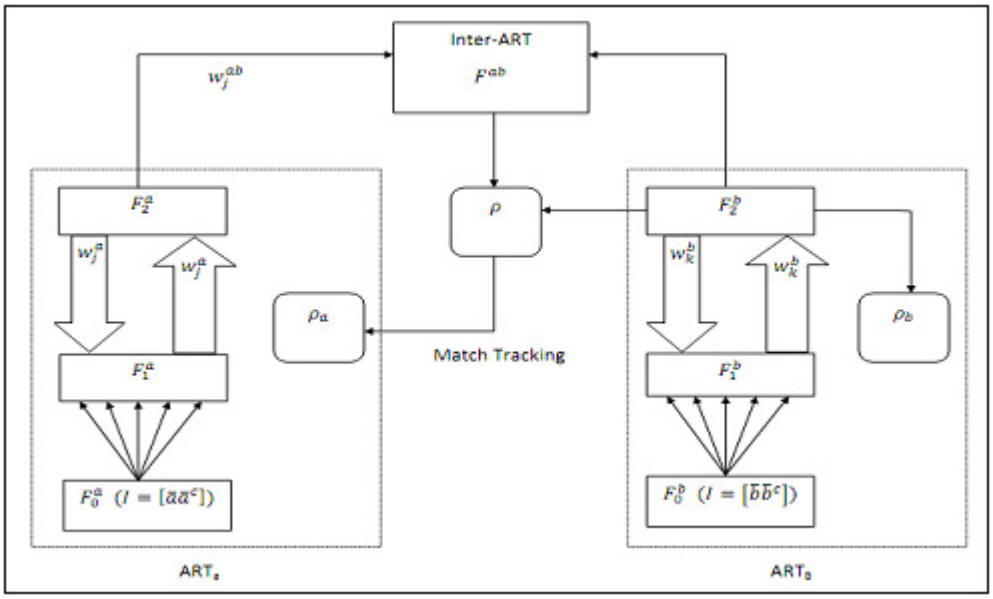

Figure 1. Architecture of the fuzzy ARTMAP neural network [20].

5. ARCHITECTURE OF SYSTEM

The main component in the architecture diagram Dataset process of feature selection IDS and final output. Semi supervised feature selection has constraint selection, Feature relevance, Dimension Reduction Method. Following process occur in case of feature selection process. Then applying fuzzy ARTMAP algorithm we can detect intrusion has occurred or not.

5.1 KDD99 AS A DATASET

For the purpose of this study we have used kdd99 as a dataset. The main goal is to use this dataset is that this database is freely available. Since 1999, KDD'99 has been the most wildly used data set for the estimation of anomaly detection techniques. This dataset is prepared by Stolfo et al.and is made based on the data captured in DARPA'98 IDS estimation Program.

DARPA'98 is around 4 gigabytes of compacted raw (binary)tcp dump data of 7 weeks of network traffic, which can be managed into about 5 million connection records, each with Around 100 bytes. The two weeks of test data have around 2 million link records. KDD training dataset comprises of nearly 4,900,000 particular connection vectors each of which contains 41 features and is labeled as either normal or an attack, with just one specific attack type. The simulated attacks comes under one of the following four attacks:

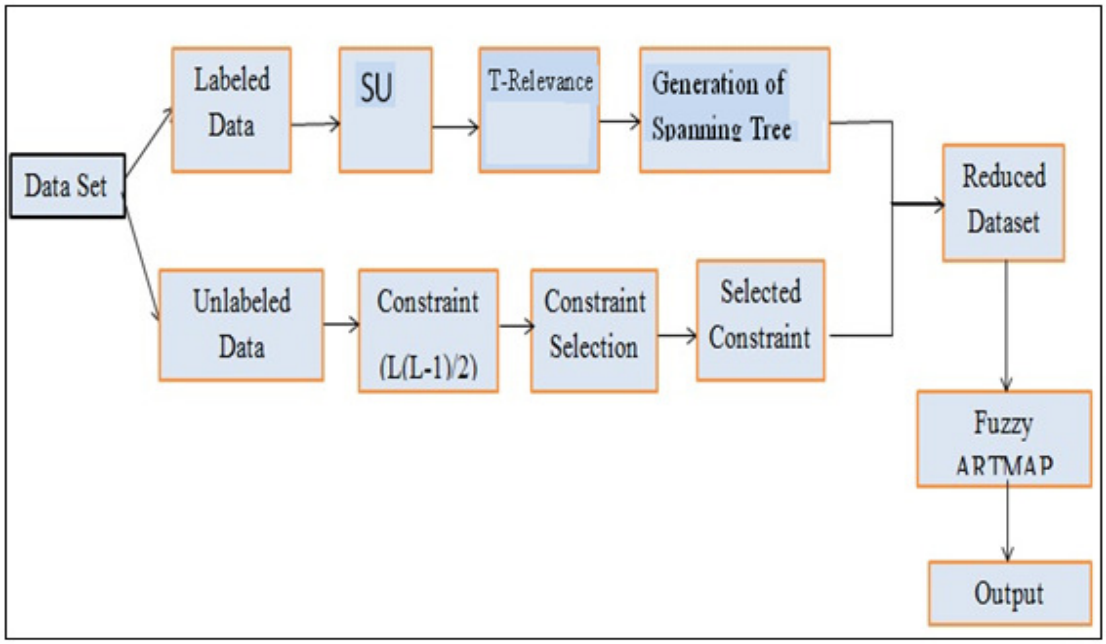

Figure 2. Architecture of the Proposed system

1) Denial of Service Attack (DoS): is an attack in which the attacker creates some computing or memory resource too busy or too full to handle valid requests, or denies valid users access to a machine.

2) User to Root Attack (U2R): is a class of action in which the attacker starts out with entry to a normal user account on the system (perhaps gained by sniffing passwords, a dictionary attack, or social engineering and is capable to exploit some susceptibility to gain root access to the system.

3) Remote to Local Attack (R2L): happens when an attacker who has the ability to send packets to a machine over a network but who does not have an account on that machine exploits specific weakness to gain local entrée as a user of that machine.

4) Probing Attack: is an task to gather information about a network of computers for the apparent resolution of circumventing its security controls[36].

Table 5.1: Basic Characteristics of the KDD 99 Intrusion Detection Datasets In Terms of Samples

Sr.No	Dataset	DOS	Probe	U2r	R2l	Normal
1	KDD10	391458	4107	52	1126	97277
2	Whole KDD	3883370	41102	52	1126	972780

6. RESULTS

6.1 CONSTRAINT SELECTION ALGORITHIM

By applying constraint selection algorithm we get following results as shown in table 1 & features which are selected are marked as a status true and features which are rejected are marked as a false. For selecting features following condition must be satisfied.

i.e Coh(Wi) >= Coh(Ω) then
ΩS = ΩS U {Wi}

Table 1. Result of constraint selection algorithm

Sr.No	Coh(Wi)	Coh(Ω)	Status
1	0.777303234	0.640956891	True
2	0.73947529	0.640956891	True
3	0.730933496	0.640956891	True
4	0.749847468	0.640956891	True
5	0.640591966	0.640956891	False
6	0.765100671	0.640956891	True
7	0.573542736	0.640956891	False
8	0.584113561	0.640956891	False
9	0.653881003	0.640956891	True
10	0.776082977	0.640956891	True
11	0.761439902	0.640956891	True
12	0.761439902	0.640956891	True
13	0.55602537	0.640956891	False
14	0.572032619	0.640956891	False
15	0.785845027	0.640956891	True

Total Constraint=4950
Selected constraint=3928

6.2 FAST ALGORITHIM

Figure 3 shows how the features are selected by using symmetric uncertainty. Here we have implemented FAST algorithim to select features. Features which are marked as a blue are selected features.

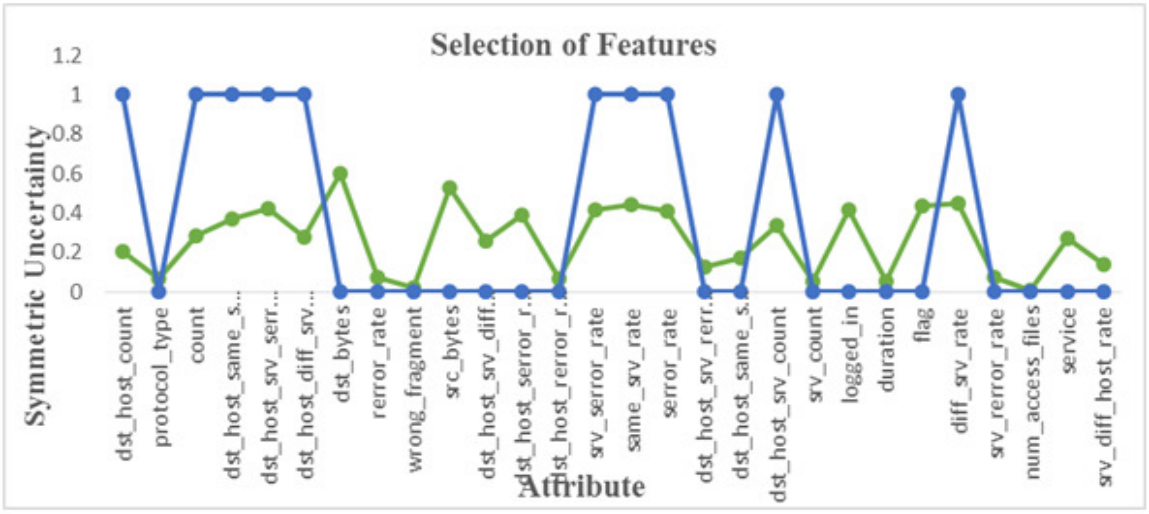

Figure 3. Selection of features by using symmetric uncertainty

6.3 FUZZY ARTMAP ALGORITHM

Figure 4 shows comparison between ARTMAP & Improved ARTMAP algorithms in terms of classification accuracy. Intrusion detection system will identified attacks as a normal or anomaly after applying fuzzy ARTMAP classifier.

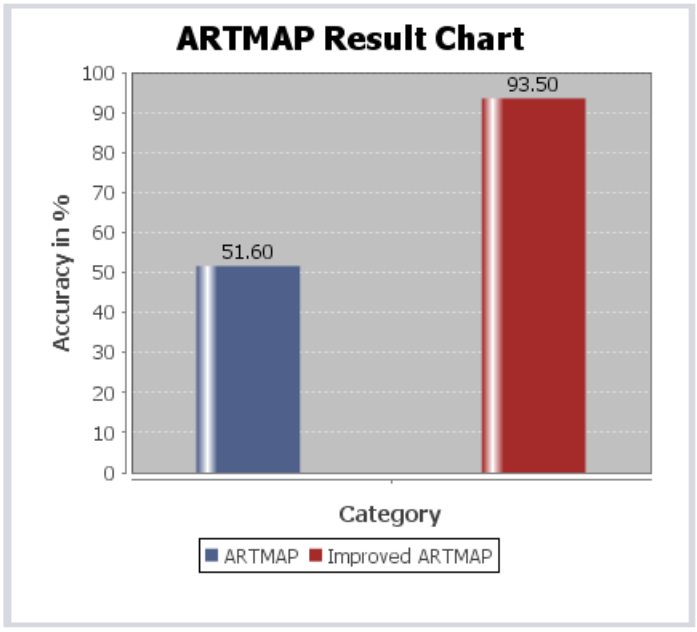

Figure 4. Comparison of ARTMAP and improved ARTMAP after reducing features

7. CONCLUSION

Security tools installation, monitoring to ensure security is the responsibility of the Security administrator in an organization. IDS generate a large number of alerts (false positives). Most of these alerts demand manual intervention from Administrator. Continuous monitoring of alerts and there by evolving judgment for improving security is the major concern. Thus, Feature relevance analysis is performed on KDD 99 training set, which is widely used by machine learning researchers.. This framework for feature selection is based on constraint selection and redundancy elimination for semi-supervised dimensionality reduction. A new score function was developed to evaluate the relevance of features based on both, the locally geometrical structure of unlabelled data and the constraint preserving ability of labelled data. Learning time of fuzzy ARTMAP classifier is very less as compare to other classifier.

In future we can use another classifier for classification of intrusion detection system so that it will save more time. Thus, Feature selection is an essential step in successful data mining applications which can effectively reduce data dimensionality by removing irrelevant features if we use that data our processing time will reduced.

ACKNOWLEDGMENT

Sincerely thank to all anonymous researchers for providing us such helpful opinion, findings, conclusions and recommendations. Also thank to guide Prof. Roshani Ade, HOD Prof Arti Mohanpurkar, Principal Dr.Uttam Kalwane & colleagues for their support and guidance.

REFERENCES

[1] L. Yu and H. Liu, Efficient feature selection via analysis of relevance and redundancy, J. Mach. Learn. Res., vol. 5, pp. 12051224, Oct 2004.
[2] M. Kalyani and M. Sushmita, Clustering and its validation in a symbolic framework, Pattern Recognition. Lett., vol. 24, no. 14,pp. 23672376, 2003.
[3] M. Kalakech, P. Biela, L. Macaire, and D. Hamad, Constraint scores for semi-supervised feature selection: A comparative study, Pattern Recognit. Lett., vol. 32, no. 5, pp. 656665, 2011
[4] K. Benabdeslem and M. Hindawi, Constrained Laplacian score for semi-supervised feature selection, in Proc. ECML-PKDD, Athens, Greece, 2011, pp. 204218
[5] I. Davidson, K. Wagstaff, and S. Basu, Measuring constraintset utility for partitional clustering algorithms, in Proc. ECML/PKDD, 2006
[6] K. Allab and K. Benabdeslem, Constraint selection for semisupervised topological clustering, in Proc. ECML-PKDD, Athens, Greece, 2011, pp. 2843.
[7] Z. Zhao, L. Wang, and H. Liu, Efficient spectral feature selection with minimum redundancy, in Proc. AAAI, 2010
[8] C. Ding and H. C. Peng, Minimum redundancy feature selection from microarray gene expression data, in Proc. IEEE CSB, 2003, pp. 523528. Intrusion Detection using Fuzzy ArtMAP
[9] H. Peng, F. Long, and C. Ding, Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy, IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 8, pp. 12261238, Aug. 2005
[10] B. Auffarth, M. Lopez, and J. Cerquides, Comparison of redundancy and relevance measures for feature selection in tissue classification of CT images, in Proc. 10th ICDM, Berlin, Germany,2010, pp. 248262.
[11] M. Hindawi, K. Allab K. Benabdeslem, Constraint selection based semi-supervised feature selection, in Proc. IEEE ICDM,Vancouver, BC, Canada, 2011, pp. 10801085

[12] J. B. MacQueen, Some methods for classification and analysis of multivariate observations, in Proc. 5th Symp. Math. Statist. Probab., Berkley USA, 1967, pp. 281297
[13] E. Elhamifar and R. Vidal, Sparse manifold clustering and embedding, in Proc. NIPS, 2011, pp. 5563.
[14] J. H. Ward, Hierarchical grouping to optimize an objective function, J. Amer. Statist. Assoc., vol. 58, no. 301, pp. 236244,1963 .
[15] Z. Zhao and H. Liu, Spectral Feature Selection for Data Mining (Data Mining Knowledge Discovery Series). Boca Raton, FL, USA:Chapman and Hall-CRC, 2012
[16] X. He, D. Cai, S. Yan, and H. Jiang Zhang, Neighbourhood preserving, in Proc. 10th IEEE Int. Conf. Computer Vision, Beijing, Germany, 2005, pp. 12081213
[17] Robust ensemble feature selection for high dimensional data sets Ben Brahim, A. ; LARODEC Univ. of Tunis, Tunis, Tunisia ; Limam .July 2013
[18] Stephen Guo et al.,To Link or Not to Link? A Study on End-to-End Tweet Entity Linking,NAACL-HLT,AAAI,2013 Department of Computer Engineering Intrusion Detection using Fuzzy ArtMAP
[19] Yegin Genc et al.,Discovering Context: Classifying Tweets through a Semantic Transform Based on Wikipedia,Pages 482-492,Springer,2011
[20] Carpenter, G.A., Grossberg, S., Markuzon, N., Reynold, J.H., Rosen, D.B.: FuzzyARTMAP: A neural network for incremental supervised learning of analog multidimensional maps. IEEE Transactions on Neural Network 3(5), 689713,1992
[21] Vilakazi, C.B., Marwala,T.:Application of feature selection and fuzzy ARTMAP to intrusion detection. In Proceeding of 2006 IEEE International Conference on Systems, Man Cybernetics, pp. 48804885 ,2006
[22] Jun Yan, Benyu Zhang, Ning Liu, Shuicheng Yan, Qiansheng Cheng, Weiguo Fan, Qiang Yang, Wen si Xi, and Zheng Chen, Effective and Efficient Dimensionality Reduction for Large-Scale and Streaming Data Preprocessing, Mar 2006.
[23] Y. Saeys and etal. A review of feature selection techniques in bioinformatics. Bioinformatics, 23(19):25072517, 2007.
[24] X. He, D. Cai, and P. Niyogi. Laplacian score for feature selection. In Advances in Neural Information Processing Systems 18, 2005.
[25] L. Song, A. Smola, A. Gretton, J. Bedo, and K. Borgwardt. Feature selection via dependence maximization. Journal of Machine Learning Research, 2007.
[26] L. Yu and H. Liu. Feature selection for high-dimensional data: A fast correlation based filter solution. In Proceedings of the ICML, 2003.
[27] Z. Zhao and H. Liu. Spectral feature selection for supervised and unsupervised learning. In Proceedings of the ICML, 2007.
[28] H. Li, T. Jiang, and K. Zhang, Efficient and Robust Feature Extraction by Maximum Margin Criterion, Proc. Conf. Advances in Neural Information Processing Systems, 2004.
[29] Jun Yan, Benyu Zhang, Ning Liu, Shuicheng Yan, Qiansheng Cheng, Weiguo Fan, Qiang Yang, Wen si Xi, and Zheng Chen, Effective and Efficient Dimensionality Reduction for Large-Scale and Streaming Data Preprocessing, Mar. 2006.
[30] G. Forman. An extensive empirical study of feature selection metrics for text classification. Journal of Machine Learning Research, 3:12891305, 2003.
[31] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. Annals of Statistics, 32:40749, 2004.
[32] J. G. Dy and etal. Unsupervised feature selection applied to content-based retrieval of lung images. Transactions on pattern Analysis and Machine Intelligence,25(3):373378, 2003.
[33] X. He, D. Cai, and P. Niyogi, Laplacian score for feature selection, in Proc. NIPS, Vancouver, Canada, 2005.
[34] O. Chapelle, B. Scholkopf, and A. Zien, editors. Semi- Supervised Learning. MIT Press, Cambridge, 2006.
[35] K. wagstaff, c. cardie, s. rogers, and s. schroedl, Constrained k means Clustering with background knowledge, in ICML01, Williamstown, MA, 2001, pp. 577584.
[36] T. ZHANG AND R. K. ANDO, Analysis of spectral kernel design based semi supervised Learning, in NIPS 18, MIT Press, Cambridge, MA, 2006, pp. 16011608.

[37] Pallavi Kulkarni and Roshani Ade., 2014 Article: Prediction of Student's Performance based on Incremental Learning. International Journal of Computer Applications 99(14):10-16.

[38] Ade, Ms RR, et al.,2013, "Methods for Incremental Learning: A Survey "International Journal of Data Mining & Knowledge Management Process 3.4 (2013).