Volume 5, Issue 4, April 2015 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

A Meta-Analysis on Different Attack Detection Techniques

Sana Warsi
M.Tech Student, Comp. Science,
SIST, Bhopal, India

Yogesh Rai
Asst. Prof., Computer Science,
SIST, Bhopal, India

Santosh Kushwaha
HOD, Computer Science,
SIST, Bhopal, India

Abstract—Intrusion detection (ID) is an attack identification technique by which possible threats like DoS, U2R, R2L and Probe are identified. With the massive transformation of data through internet or network enhance the security requirements also. So security with respect to controlling and prediction is very important. In this paper we have concentrated on better classification and prediction strategy. By this paper a proper insight has been presented in the direction of attack detection. KDD database is the basis of the study. Data mining and evolutionary algorithms are being used for this study. Based on the study and analysis some future scope has been suggested in this paper.

Keywords –ID, Data Mining, Evolutionary algorithm, DoS, U2R, R2L and probe

I. INTRODUCTION

The Association for Computing Machinery (ACM) has a particular vested party on Knowledge Discovery and Data mining (KDD) [1] for the information mining understudies and analysts. They gave set KDD Cup99 information sets for intrusion discovery [2]. This collection is use for intrusion detection and several researchers had considered this as the benchmark dataset for result comparison.

Recently, numerous researchers are centering to utilize information digging ideas for Intrusion Detection [3]. This is a procedure to concentrate the verifiable data and learning.

Intrusion discovery is the methodology of pernicious assault in the framework and system when we are presently correspondence or separating information in the constant environment [4][5]. Since its innovation, interruption location has been one of the key components in accomplishing data security. It goes about as the second-line barrier which supplements the entrance controls. At the point when the controls fizzled, the interruption identification frameworks ought to have the capacity to recognize it constant and caution the security officers to take provoke and suitable activities [5][6].

Interruption recognition framework manage overseeing the episodes happening in PC framework or system situations and looking at them for indications of conceivable occasions, which are encroachment or inescapable dangers to PC security, or standard security hones Intrusion identification frameworks (IDS) have risen to distinguish activities which jeopardize the uprightness, privacy or accessibility of are source as a push to give an answer for existing security issues [7].

So in the above bearings we study a few perspectives in the consequent segments. We likewise talk about information mining and advancement strategies, on the grounds that it can be utilized as a part of shaping the structure which delivers better identification framework.

As we are examined this study toward a superior system with the mix of information mining and streamlining. These strategies are helpful and has been utilized as a part of diverse methodologies like [8][9][10][11][12][13]. So the utilization of these calculations can improve an effect.

II. LITERATURE SURVEY

In 2012, LI Yin–huan [14] concentrates on an enhanced FP-Growth calculation. As per creator Preprocessing of information mining can expand proficiency on looking the normal prefix of hub and diminish the time unpredictability of building FP-tree. In view of the enhanced FP Growth calculation and other information mining systems, an interruption location model is completed by creators. Their exploratory results are successful and doable.

In 2012, P. Prasenna et al. [15] proposed that in ordinary system security just depends on numerical calculations and low counter measures to taken to avert interruption identification framework, albeit the majority of this methodologies as far as hypothetically tested to execute. Creators propose that as opposed to producing substantial number of principles the advancement improvement procedures like Genetic Network Programming (GNP) can be utilized .The GNP is in view of coordinated diagram. They concentrate on the security issues identified with send an information mining-based IDS in a continuous situation. They sum up the issue of GNP with affiliation principle mining and propose a fluffy weighted affiliation guideline mining with GNP system suitable for both constant and discrete qualities.

In 2011, LI Han [16] concentrates on interruption discovery in light of grouping examination. The point is to enhance the recognition rate and abatement the false caution rate. An adjusted element K-implies calculation called MDKM to identify inconsistency exercises is proposed and relating reenactment analyses are introduced. Firstly, the MDKM calculation channels the commotion and segregated focuses on the information set. Also by ascertaining the separations between all example information focuses, they acquire the high-thickness parameters and group part parameters, utilizing element iterative methodology we get the k bunching focus precisely, then an oddity discovery model is displayed. They utilized KDD CUP 1999 information set to test the execution of the model. Their outcomes demonstrate the framework has a higher recognition rate and a lower false caution rate, it attains to hopeful point.

In 2011, Z. Muda et al. [17] talk about the issue of current irregularity identification that it not able to distinguish a wide range of assaults effectively. To beat this issue, they propose a half breed learning approach through blend of K-Means bunching and Naïve Bayes characterization. The proposed methodology will be grouping all information into the comparing gathering before applying a classifier for order reason. An examination is done to assess the execution of the proposed methodology utilizing KDD Cup '99 dataset. Result demonstrate that the proposed methodology performed better in term of exactness, location rate with sensible false caution rate.

In 2014, Deshmukh et al. [18] presents a Data Mining system in which different preprocessing techniques will be included such as Normalization, Discretization and Feature choice. With the help of these techniques the information will be preprocessed and obliged highlights are chosen. They utilized NaIve Bayes system in directed learning strategy which groups different system occasions for the KDD cup'99 Dataset.

In 2014, Benaicha et al. [19] present a Genetic Algorithm (GA) approach with an enhanced starting populace and choice administrator, to proficiently identify different sorts of system interruptions. They utilized GA to enhance the look of assault situations in review documents, thanks to its great offset investigation / misuse; as per the creators it gives the subset of potential assaults which are display in the review document in a sensible preparing time. The testing period of the Network Security Laboratory Knowledge Discovery and Data Mining (NSL-KDD99) benchmark dataset has been utilized to identify the abuse exercises. Their methodology of IDS with Genetic calculation increments the execution of the identification rate of the Network Intrusion Detection Model and decreases the false positive rate.

In 2014 Kiss et al. [20] recommend that Modern Networked Critical Infrastructures (NCI), including digital and physical frameworks, are presented to keen digital assaults focusing on the steady operation of these frameworks. To guarantee abnormality mindfulness, their watched information can be utilized as a part of agreement with information mining procedures to create Intrusion Detection Systems (IDS) or Anomaly Detection Systems (ADS). They proposed a grouping based methodology for identifying digital assaults that cause peculiarities in NCI. Different bunching strategies are investigated to pick the most suitable for grouping the time-arrangement information highlights, consequently characterizing the states and potential digital assaults to the physical framework. The Hadoop execution of MapReduce standard is utilized to give a suitable preparing environment to extensive datasets.

In 2014, Thaseen et al. [21] proposed a novel technique for coordinating vital segment investigation (PCA) and bolster vector machine (SVM) by upgrading the piece parameters utilizing programmed parameter determination system. Their methodology lessens the preparation and testing time to distinguish interruptions consequently enhancing the exactness. Their proposed strategy was tried on

KDD information set. The datasets were painstakingly partitioned into preparing and testing considering the minority assaults, for example, U2R and R2L to be exhibit in the testing set to distinguish the event of obscure assault. Their outcomes show that the proposed strategy is effective in recognizing interruptions. Their exploratory results demonstrate that the order exactness of the proposed system outflanks other arrangement strategies utilizing SVM as the classifier and other dimensionality decrease or highlight choice systems.

In 2014, Wagh et al. [22] proposed Network security is an essential part of web empowered frameworks in the present world situation. As per the creators because of perplexing chain of PCs the open doors for interruptions and assaults have expanded. In this way it is need of great importance to locate the most ideal routes conceivable to secure our frameworks. So the creators propose interruption identification framework is assuming basic part for PC security. The best strategy used to tackle issue of IDS is machine learning. Thy watched that the rising field of semi regulated learning offers a guaranteed route for corresponding exploration. So they proposed a semi-managed system to diminish false alert rate and to enhance discovery rate for IDS.

In 2014, Masarat et al. [23] presented a novel multistep structure taking into account machine learning procedures to make a proficient classifier. In first step, the highlight choice technique will execute taking into account pick up proportion of highlights by the creators. Their technique can enhance the execution of classifiers which are made taking into account these highlights. In classifiers mix step, we will exhibit a novel fluffy gathering technique. In this way, classifiers with more execution and lower expense have more impact to make the last classifier.

III. PROBLEM IDENTIFICATION

The observation bases on the study are following:

1) Intrusion detection technique based on the single method is not sufficient. Combinatorial approach will provide better classification.
2) Evolutionary algorithm along with the data mining techniques can provide better attack detection.
3) Neuro-Fuzzy Combination can be used as the distributed classifier.
4) The possible threats are DoS, U2R, R2L and Probe, it not guarantee that if the method which provides good result in any one type can provide better detection with other types also.
5) K-Means with K-Nearest Neighbor (KMKNN) approach for better interruption recognition. These methodologies demonstrated a sensible location rate contrast with our methodology. Sadly, a potential downside of this strategy is the rate of false alerts.
6) The detection rates in comparison to U2R and R2L are less so need to improve the rates in these cases also.

7) Evolutionary Soft Computing based Intrusion Detection System (ESC-IDS) which centers to recognize and characterize interruption has proposed. This methodology has genuine inadequacies in its low exactness rate and in addition the inclination to deliver high false alert rate.

IV. ANALYSIS

The results comparison based on the previous methodology are shown below.

Table 1: Analysis

S.no	Approach	Accuracy (%)
1	Random Forest [25]	92.93 %
2	JRip [26]	92.30 %
3	SVM [25]	92.18 %
4	PSO[2]	95.46 %
5	JRip [25]	92.30 %
6	NBTree [25]	92.28 %
7	LBK [25]	92.22 %
8	SVM [25]	92.18 %
9	J48 [25]	92.06 %
10	NB Training [16]	DOS 94.3
11	KM + NB Training [16]	DOS 99.5
12	NB Testing [16]	DOS 82.5
13	Km + NB Testing [16]	DOS 99.6
14	FSVM [17]	97.14
15	Rule based [17]	89.90
16	FSVM [17]	95.23
17	Rule based [17]	91.34
18	FSVM [17]	95.12
19	Rule based [17]	92.22
20	FSVM [17]	97.13
21	TANN [27]	96.91
22	SOM[28]	97.13
23	FP-Growth and Dynamic Rule Generation with Clustering [29]	Efficient in non-candidate generation
24	Hierarchical Clustering and SVM [30]	95.7
25	ESC-IDS [31]	65.48

V. CONCLUSION AND FUTURE WORK

The study shows good results have been achieved by using combining methods. It can be data mining and evolutionary techniques. The future suggestions are as under:

1) Data mining techniques with evolutionary algorithm like PSO, ACO, ABC and Cuckoo search algorithm can provide good results.
2) Separate classification can be done in each intrusion cases.
3) Different evolutionary task can be applied simultaneously.
4) Rule can be classified dynamically with positive and negative association.

REFERENCES

[1] Alexander O. Tarakanov, Sergei V. Kvachev, Alexander V. Sukhorukov ," A Formal Immune Network and Its Implementation for On-line Intrusion Detection", Lecture Notes in Computer Science Volume 3685, pp 394-405, 2005.

[2] Ranjna Patel, Deepa Bakhshi and Tripti Arjariya, " Random Particle Swarm Optimization (RPSO) based Intrusion Detection System " , International Journal of Advanced Technology and Engineering Exploration (IJATEE), Volume-2, Issue-5, April-2015 ,pp.60-66.

[3] Meng Jianliang,Shang Haikun,Bian Ling," The Application on Intrusion Detection Based on K-means Cluster Algorithm", International Forum on Information Technology and Applications, 2009.

[4] Lundin, E. and Jonsson, E. "Survey of research in the intrusion detection area", Technical Report, Department of Computer Engineering, Chalmers University of Technology, Göteborg, Sweden. January 2002.

[5] R.Venkatesan, R. Ganesan, A. Arul Lawrence Selvakumar, " A Comprehensive Study in Data Mining Frameworks for Intrusion Detection " , International Journal of Advanced Computer Research (IJACR), Volume-2, Issue-7, December-2012 ,pp.29-34.

[6] S.Devaraju, S.Ramakrishnan:,"Analysis of Intrusion Detection System Using Various Neural Network classifiers, IEEE 2011.

[7] Moriteru Ishida, Hiroki Takakura and Yasuo Okabe," High-Performance Intrusion Detection Using OptiGrid Clustering and Grid-based Labelling", IEEE/IPSJ International Symposium on Applications and the Internet, 2011.

[8] S. T. Brugger, "Data mining methods for network intrusion detection",pp. 1-65, 2004.

[9] W. Lee, S. J. Stolfo, "Data Mining Approaches for Intrusion Detection",Proceedings of the 1998 USENIX Security Symposium, 1998.

[10] Kamini Nalavade, B.B. Meshram, " Mining Association Rules to Evade Network Intrusion in Network Audit Data " , International Journal of Advanced Computer Research (IJACR), Volume-4, Issue-15, June-2014 ,pp.560-567.

[11] W. Lee, S. J. Stolfo, "Data mining approaches for intrusion detection" Proc. of the 7th USENIX Security Symp.. San Antonio, TX, 1998.

[12] Reyadh Naoum, Shatha Aziz, Firas Alabsi, "An Enhancement of the Replacement Steady State Genetic Algorithm for Intrusion Detection", International Journal of Advanced Computer Research (IJACR), Volume-4, Issue-15, June-2014, pp.487-493.

[13] Aditya Shrivastava, Mukesh Baghel, Hitesh Gupta, " A Review of Intrusion Detection Technique by Soft Computing and Data Mining Approach " , International Journal of Advanced Computer Research (IJACR), Volume-3, Issue-12, September-2013 ,pp.224-228.

[14] LI Yin–huan , "Design of Intrusion Detection Model Based on Data Mining Technology", International Conference on Industrial Control and Electronics Engineering, 2012.

[15] P. Prasenna, R. Krishna Kumar, A.V.T Raghav Ramana and A. Devanbu "Network Programming And Mining Classifier For Intrusion Detection Using Probability Classification", Pattern Recognition, Informatics and Medical Engineering, March 21-23, 2012.

[16] LI Han, "Using a Dynamic K-means Algorithm to Detect Anomaly Activities", Seventh International Conference on Computational Intelligence and Security, 2011.

[17] Z. Muda, W. Yassin, M.N. Sulaiman, N.I. Udzir," Intrusion Detection based on K-Means Clustering and Naïve Bayes Classification", 7th International Conference on IT in Asia (CITA), 2011.

[18] Deshmukh, D.H.; Ghorpade, T.; Padiya, P., "Intrusion detection system by improved preprocessing methods and Naïve Bayes classifier using NSL-KDD 99 Dataset," Electronics and Communication Systems (ICECS), 2014 International Conference on , vol., no., pp.1,7, 13-14 Feb. 2014.

[19] Benaicha, S.E.; Saoudi, L.; Bouhouita Guermeche, S.E.; Lounis, O., "Intrusion detection system using genetic algorithm," Science and Information Conference (SAI), 2014 , vol., no., pp.564,568, 27-29 Aug. 2014.

[20] Kiss, I.; Genge, B.; Haller, P.; Sebestyen, G., "Data clustering-based anomaly detection in industrial control systems," Intelligent Computer Communication and Processing (ICCP), 2014 IEEE International Conference on , vol., no., pp.275,281, 4-6 Sept. 2014.

[21] Thaseen, I.S.; Kumar, C.A., "Intrusion detection model using fusion of PCA and optimized SVM," Contemporary Computing and Informatics (IC3I), 2014 International Conference on , vol., no., pp.879,884, 27-29 Nov. 2014.

[22] Wagh, S.K.; Kolhe, S.R., "Effective intrusion detection system using semi-supervised learning," Data Mining and Intelligent Computing (ICDMIC), 2014 International Conference on , vol., no., pp.1,5, 5-6 Sept. 2014.

[23] Masarat, S.; Taheri, H.; Sharifian, S., "A novel framework, based on fuzzy ensemble of classifiers for intrusion detection systems," Computer and Knowledge Engineering (ICCKE), 2014 4th International eConference on , vol., no., pp.165,170, 29-30 Oct. 2014.

[24] J. Zhang, M. Zulkernine, and A. Haque, "Random-forestsbased network intrusion detection systems," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions, vol. 38, pp. 649-659, 2008.

[25] H. A. Nguyen and D. Choi, "Application of data mining to network intrusion detection: classifier selection model," in Challenges for Next Generation Network Operations and Service Management, ed: Springer, 2008, pp. 399-408.

[26] T. Ambwani, "Multi class support vector machine implementation to intrusion detection", Proceedings of the International Joint Conference on Neural Networks, 2003, pp. 2300-2305.

[27] C. F. Tsai, and C.Y Lin, “A triangle area-based nearest neighbors approach to intrusion detection,” Pattern Recognition, 2010, 43(1):222-229.

[28] Deepak Rathore, Anurag Jain, “Design Hybrid method for intrusion detection using Ensemble cluster classification and SOM network”, International Journal of Advanced Computer Research (IJACR), Volume-2, Issue-5, September-2012, pp.188-194.

[29] Manish Somani, Roshni Dubey, “Design of Intrusion Detection Model Based on FP-Growth and Dynamic Rule Generation with Clustering”, International Journal of Advanced Computer Research (IJACR), Volume-3, Issue-10, June-2013, pp.146-150.

[30] S-J Horng, M-Y Su and Y-H Chen, “A novel intrusion detection system based on hierarchical clustering and support vector machines,” Expert Systems with Applications, 2011, 38:306–313.

[31] M. Toosi, “A new approach to intrusion detection based on an evolutionary soft computing model using neuro-fuzzy classifiers,” Computer Communications, 2007, 30: 2201-221.