iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** XII **Month of publication:** December 2015

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 3 Issue XII, December 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

An Intelligent System For Intrusion Detection By Svm And Ant Colony Using Neural Network

Prachi Agrawal[1] , Akhilesh Pandey[2]
[1]M.Tech Scholars , [2]Assistant Professor, Department of Computer Science & Engineering
Suresh Gyan Vihar University, Jaipur

Abstract*— Intrusion detection systems in wireless, mobile ad hoc and sensor networks are determined greatly by classification argued in the previous chapter. Although, in contrast with wired networks, it is required to be noted into records the extra difficulties because of their distributed character and ad hoc architecture. Rule-based systems, like as expert systems normally cover a thorough working knowledge of engineering, the development of knowledge experts called intrusive rules. As a secondary method, rule induction technology, find and clean data sets to naturally create such rules.
Keywords-Manet, Intrusion, Neural Netwok, Ant Colony

I. INTRODUCTION

Intrusion detection detects conflicting behavior of users in the computer system with network by analyzing and monitoring all the occurrences of events into the system that are different than the intended events. For this purpose, techniques to identify intrusion behavior in a computer system are developed by modelling and recognizing such behaviors by Intrusion Detection System (IDS). When a system behavior differs from normal or expected usage of network and system, it is generally called as intrusion behavior. For the detection of an intrusion, many challenges like fault detection, localization and fault management come into face. When these are not taken into consideration, a natural overlap in between these domains is developed for event correlation.
Surveillance/probing stage : Vulnerabilities in software and configurations are scanned by intruders and potential targets are identified gathering information from their computers which includes password cracking.
Activity (exploitation) stage: Administrator rights are targeted to be obtained once weakness has been identified in previous stage from the selected host(s). This can provide attacker free access to exploit the system. Denial of Service (DoS) attacks are also included in this stage as explained below.
Mark stage: This stage marks the achievement of intruder's goal (Asaka et al. 1999). As after activity stage or exploitation stage , attacker can steal essential information from the system, destroy data that may be important for tracking attackers intrusion through log files, hide a virus or spyware in the system or software, or exploit attacked host for conducting further attacks on new vulnerable host(s).
Masquerading stage: This is the final stage. The intruder attempts to destroy all traces of the attack for example, all the log entries that can reveal the intrusion time and location.

A. Intrusion Detection Systems

The architecture of IDSs keeps on changing due to the diversity and evolution of intrusion behavior. However, Verwoerd and Hunt (2002) identified and generalized the common building blocks of IDS :
Sensor probes: Collect data from the system being checked.
Monitor: from many sensors and forwarding suspicious content received events to a "resolver."
Resolver: to determine an appropriate response for content that is suspicious
Controller: Provides management capabilities.

Volume 3 Issue XII, December 2015
ISSN: 2321-9653

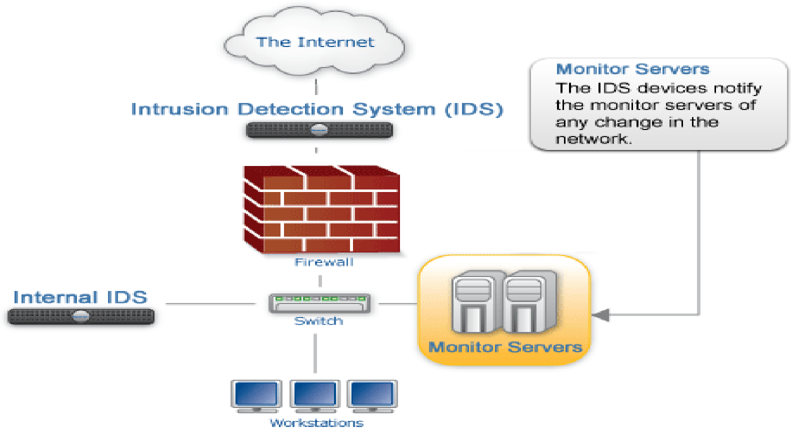

Figure 1: An Intrusion Detection Method

There are two methods to do detection, known as misuse detection and anomaly detection (Endler1998, 2006 Gollmann, Kemmerer cowpea and 2002, Lee and Cheung 2001). These terms also called on the basics of intrusion detection and behavior (Debar, et al., 1999, Debar 2000). The former try coding known intrusion (misuse) of knowledge, generally as a rule, use the screening event (also called signature IDS). The latter tries to "learn" features of occurence patterns constituting normal behavior, and by observing the deviation from established norms patterns, detects intrusion occurred (Denning 1987). Some of IDS provides two functions, usually through a variety of hybridization techniques, see for example (Depren et al., 2005,1998Endel). However, the system can also be based on the normal data and intrusion, which has become a recent study using machine learning techniques (common modeling methods Bouzida and Cuppens of 2006 a; B, Depren et al., 2005, panda and Patras2007 ; in 2009, Sabhnani and 2003 Serpen, Gordon et al., 2008, Zhang and annual 2006 Zulkernine). Misuse detection is accepted for commercial intrusion detection, according to Gollman (2006, p. 252-253), who pointed out that "misuse detection is the ground of all commercial IDS products at the time of creation [2005]." However, this method cannot detect the attack has not been inbuilt for it, and therefore, it is liable to generate false negative issue, if the system is not updated with the trending intrusion (Gollmann 2006, pp. 251-252, Lewis 1993). On contrary, misuse detection systems typically generate some false positives (Kruegel et al., 2004).

General views on misuse detection, as described above, are not anymore entirely correct. In latest years, researchers have developed technology making the abuse detection system more adaptable, more able to detect changes in the attack. This could only happen with machine learning programs, like artificial neural networks, which are created to be possible to make a common category for known attacks carried out under the classification of unknown circumstances. This is even a scenario for the rule-based system, which wereconsidered in the past that could not detect attacks, even minor changes due to the case of mandatory guidelines follow up (Esmaili et al., 1996, 1993 Lewis, Owens and Levary 2006). Rule-based systems are efficient enough now to detect changes in the attack, and also can be used for anomaly detection, mainly because of the researchers implying fuzzy logic rule building (see details in section 3.2 on page 19).

One advantage of anomaly detection is the characteristic to find new attacks, because the system is constructed as per the normal behavior. The word "behavior" implies that host-based IDS, its anatomy of user behavior, besides it can even be an IDS anatomizing network traffic. In both cases, shaping up of normal behavior / traffic is an exquisite task, which leads this method easy to send false positives (Dokas et al. 2002, Kruegel et al., 2004).

The act or process behavior-based anomaly detection by the host can be dedicated on the user / program. For the former, it is on assumption that the user will use the system in an expected manner; and some map track can be derived seeing the behavior lead by the habits. Therefore, predicting that it is doable to get the user based on the data, and if the usage of system is not in accordance from each user's habitual ways, then it is determined to be a potential intrusion (Debar et al. 1992, Ryan et al., 1998).

Debar and so on. On (1992) to consider many levels of data source simulation anomaly detection, they are categorized as follows:

Keyboard levels: key that was pressed, since last playing time,etc.

Command level: usage of commands along with their trail of actions. Now, researchers also keep remembrance of output parameters and system calls (Micarelli and 2007 Sansonetti) parameter.

Session Level: Keeping track of terminal session event, which can generate data, such as "the length of the meeting, the names of all

Volume 3 Issue XII, December 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

the CPU, memory, and input and output use, the use of the terminal, login time, day ......" (Debar et al., 1992). However, as Debar like. State, the data derived is due only if the user has fulfilled a session, till this time, the user may have welcomed the invasion. Hence this method is doubtful to be actually catching any real-time intrusion.

Group level: User gathered in group formation. Based on any of these groups, the exception detecting system can create a plurality of user profiles (therefore also referred to as user profile). This can be executed as a consideration, or as a stage for each user profile, for further details on the system user groups specific privileges, for example, administrators, programmers, secretaries, etc., see MUTZ like. (2006).

A difficulty to the host-based anomaly detection system is constantly be updated with the varying environment. Retraining or regular updates are needed to escape false alarms growth called as behavioral drift (Balajinath and ECLAC 2001). It is feasible the modeling / training of abnormal system through the time, however, the presence of this one specific problem : learning invasive behavior, as well as the risk (Kruegel et al.2004). If user realizes that training anomaly detection system is been conducted, he / she may slowly change his / her in this way, a decided attack will not be caught through the difference of behavior (2006 Gollmann, p. 253).

II. PAST STUDY

The KDD Cup '99 datasets (UCI KDD Archive of 1999) was a knowledge discovery and building up of Data mining applications to contest and win in related meetings in 1999 (Elkan 2000), and has been of help widely to verify network-based intrusion detection system prototypes. Although, as analyzed in the first chapter, there are differences between the results generated in the literature along with this set of data. Hence, this surveys this chapter.

Another question the survey found differences related to KDD Cup '99 datasets, and in the data where it is coming from. It was developed by the change of the original tcpdump DARPA98 / 99 data (Lippmann et al 2000 a; b) a set of characteristics to be benchmarked appropriate for machine learning Group's technology (Lee and Stolfo 2000). After you make a data set briefly DARPA98 with Chen suggested assessment for present intrusion detection systems, McHugh (2000) printed critical findings of the estimation project. Other scholars have expressed more reviews for KDD Cup '99 and DARPA dataset (Bouzida and Cuppens 2006 a; B, Brugger 2007 a, Mahoney and 2003 Chen, Sabhnani and 2004 Serpen), which causedBrugger (year 2007 b) to declare data set are basically blemish and results dependent only on the data are irrational.

Maximum of the living reviews focus DARPA's data (Brugger 2007 a; B, O'Neill and Chen In 2003, McHugh 2000), but Brugger (year 2007 b) can be extended to KDD Cup'99 data sets also, which is not completely right. At the same time, on the basis of some findings (Bouzida and Cuppens 2006 a; B, Sabhnani and 2004 Serpen), KDD '99 Cup dataset includes 'problems' those does not exist in DARPA data. The term "problem" is taken with caution here, because there is no adequate analysis and arguments in the existing study resulting that methodological factors have data sets problems. There are numeral methodologies factor in this paper that have been found as largely influencing the outcome. Additionally, this article tries to find that these are actual problems with the data set, or if they are just the difficulties of intrusion detection, in which common machine learning methods can be futile.

McHugh (2000) wished that through critiquing the present hard work and open arguments of issues will help likewise efforts in the next times. Although, ten years later, no such measures have been implemented. However Brugger (year 2007 b) demoralize scholars to use KDD Cup '99 datasets, he does understands that researchers keep on using it, for scarcity of improved alternatives due openly reachable. The findings in last chapter illustrates this point clearly. Henceforth, in order to take charge of these criticisms is very crucial, and methodological issues, to analyze and finalize if the KDD Cup '99 data sets can be utilized for future work to give meaningful outcomes.

A. Contradictory Results

There are three kinds of KDD Cup '99 datasets openly in reach, a subset (All) training set, 10% of this training set version, and test set. All the training and test sets used by some researchers, such as in the competition, while others use 10% of the training set and test set. Also, the use of only some of the training set, or even smaller subset is also done.

It is anticipated that the use of different subsets of the data will bring in different results, especially when the original test set has 17 new attacks. At the same time, the result is not only different, they are in some scenarios opposite. For example, Pan and so on. (2003) declares that artificial neural network (artificial neural networks) is not competent to detect U2R R2L intrusion, while Mukkamala and Sung (2003) presented a relatively high detection rate (48% U2R and 95% R2L). Likewise, for the decision tree

Volume 3 Issue XII, December 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

(DTS), Sabhnani and Serpen (2004) presented that high detection U2R (99.18%) and R2L (99.18%), although Benferhat and concrete (2005) told the detection rate to be very poor, 10.09% and 0.56% vice versa. This difference is a problem, because recent research in machine learning intrusion detection targeted to observe a combination of classification based on different categories for different categories of invasion (Anuar et al., 2008, Gharibian and annual 2007 Ghorbani, Pan et al. In 2003, Peddabachigari like 2007, Sabhnani and 2003 Serpen, and Hunan. 2008). As because of these findings, we cannot finalize that a technology is more perfect for a specific class of intrusion. Although, there are great differences between these methods of research, which is to be exposed going deeply into the results when they are present. For example, Pan and so on took only one type of intrusion for every U2R and R2L, where asMukkamala and Sung utilized a very small data set (for instance training, 5092 and 6890 for testing).

Analysis of present research results show, selection of a subset of data is the main reason for differences. Data sets can be categorized as any of the below ones:

Choosing a few species of invasion.

Create a new, smaller version of the data set.

Utilizing the initial training set only.

Utilizing the initial training set and test set.

The combined training and testing sets to produce a new data set.

The cleaning of data to meet the criteria and assumptions about data distribution invasion.

The above classification is greatly self-explanatory, but for the avoidance of doubt, it should be kept in mind that the research only utilized the training set (# 3) Do use authentication methods like trapping or cross-validation. The method number 6 is not within the scope of this article, because it is only unsupervised anomaly detection study observed (Eskin) who use cluster technology, in 2002, Liang and 2005 Leckie.

Study	Technique	U2R %	R2L %	Data subset
Pan *et al.* (2003)	ANN	0	0	#1
Bosin *et al.* (2005)	NB	52.40	94	#2
Chebrolu *et al.* (2005)	DT	48	90.58	#2
Mukkamala and Sung (2003)	ANN	48	95	#2
Peddabachigari *et al.* (2007)	DT	68	84.19	#2
Depren *et al.* (2005)	Hybrid	80	98.02	#3
Ben Amor *et al.* (2004)	DT NB	7.89 11.84	0.52 7.11	#4
Benferhat and Tabia (2005)	DT NB	10.09 11.40	0.56 8.66	#4
Bouzida and Cuppens (2006a;b)	DT ANN	7.02 0	2.58 ~27	#4
Sabhnani and Serpen (2003)	DT ANN	1.80 13.20	4.60 5.60	#4
Shafi *et al.* (2009)	DT ANN NB	33.33 44.00 60.00	0.31 35.34 8.89	#4
Sabhnani and Serpen (2004)	DT ANN	87.50 89.28	99.18 99.44	#5

Table 2.1: The U2R R2L and intrusion detection rate of ANN, DT and NB classification overview

III. PROPOSED MEHODOLOGY

Neural networks have agreed to neurons, synapses add with each other. Neuron Realization easy task involves, usually a common prescribed or accommodation by correlating the relationship between process and result between synapses. The functioning of synapses is the material bond and produce it with its end connected to a single neuron. Neurons are handy composite materials easily obtained with respect to single or multiple manufacturers synapses and synaptic single or multiple results. Design with a variety of neurons and synapses media interconnected neural networks will be described. At this moment, we have a common discoing, in explaining the structure of the neural network of the way, open their own ideas, in this method as a network is to perform the calculation and the actual neural network in other situations, ideas. The actual neural network built up neurons. Interconnected cells send electrician supply curls. Results prices neurons through these synapses relationship controlled gapped at various neurons. Methods relationship with a general increase in synaptic worth straight road through the network.Never actual nerve cells through the vocal part of the producer as our eyes and fragrant bloom relevant.

Network familiar, fear and muscle relationship based on the results. With manufacturers and come out there are several types of management procedures that we are not really connected by neural network .Ours neural networks can be connected to sensors, motors, servo system, and it may be market or weather events database, through this program. No we we are faced with what the

Volume 3 Issue XII, December 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

issue is resolved, our neural networks to obtain and agreed to write natural numbers as producers and agree to use the actual number of places like the best input output neurons connecting the profits of manufacturers plan to merge Buyers and manufacturers of neurons supplementary matters.

With a variety of neurons through synapses. Usually at the level of neurons are methodized relationship - neurons combine to become familiar with the components mounted around the first producer of primary neurons become like the rest of the battle by the right-to-long-term value of a diversified assorted layer layerlayerproducer . Final layer neural network admission layer acts like the result, and the result is worth it in the final results of the network layer. The production line is a straight sketch genius galvanized, organized like onion skin beautifying the living nerve cells. Concert by the various designs by the results of a network of different people. Few people accurately the value of research lies, as a result of occupying the result by the number of rows of neurons among neurons get help. They both nerves and synapses price results is premature given the value of its producers. Synaptic imitators distance between neurons EEG action may also be competent. Rather than the actual outcome of synaptic worth as much as it is formed producer of different types of neurons in a small trade-off worth .Bases endure engender specific result multiplied. Some general neuron-like all indexed-

input neurons: This is a website to teach the value of the product may be given to the network. The product is designated as a result of neurons neurons.

hidden neurons: teach this network through smart implementation of liquidation. Results hidden layer neurons is the same a production passenger Hide update.

output neurons: This is a Web site, the network can learn the results of worth.

Bias neurons: This is the result of a neuron, which is worth of real estate. It may have a variety of other neurons implanted bias immovable relationship.

IV. IMPLEMENTION

ALGO

Read a dataset for training the network.
Divide the dataset into the zero and one form.
Selection of the dataset from each class randomly.
Calculate the total throughput of network.
Use the SVM classifier for recognizing of intruders in the network.
Add the sub-intruders into the training dataset
Use the super wise learning method to train the dataset.
Add the some more intruders into the training dataset.
To recognize the dataset by using of back propagation algorithm.

A. The Ant System

In early 1991, the three di_erent of AS (. Rodrigo, M., Maniezzo, five, Colorni, A (1991a)) version was developed: they are called ant density, number of ants and ant in the ant density and cycle.Whereas The version number of ants pheromone ant city directly after moving from one city to the adjacent cycle version of the ant pheromone updating ant after all, just do a build tourism and pheromone amount deposited is set by each ant function quality of the tour. Because of its poor performance density and number of versions ant ants abandoned and AS algorithm only refers to the actual ants cycle version.

The two main stages in the AS algorithm ants build solutions and pheromone update. In the initialization pheromone is a pheromone deposited more than the number of ants are expected to be slightly higher in the first iteration; by setting 8 to obtain a rough estimate of the value of (i ; j) Article; _ij = _0 = M = nn, where m is the number of ants, and C nn is the length of trips generated by the nearest neighbor heuristic. The reason for this choice is, if the initial pheromone value _0 are too low, the search is rapidly produced by the ants _rst tour, which is generally biased toward the lead in the search space exploration bad area. On the other hand, if the initial pheromone value is too high, then the number of iterations will wait until the pheromone evaporation loss reduction Enough pheromone evaporation, so that ant pheromones are added to start the search bias.

Volume 3 Issue XII, December 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

V. RESULT

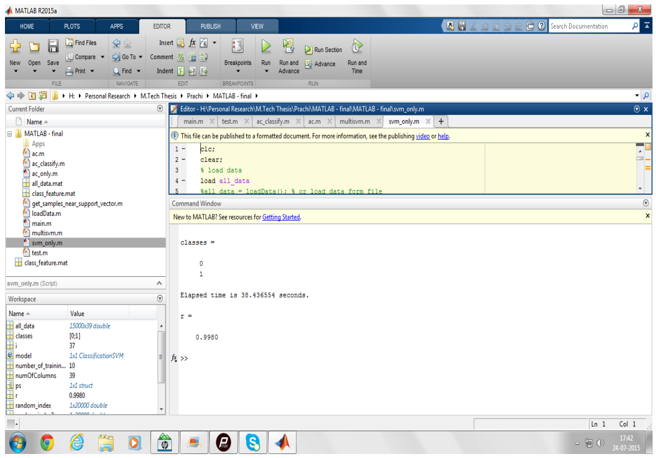

Figure 2: Implementation in MatLab

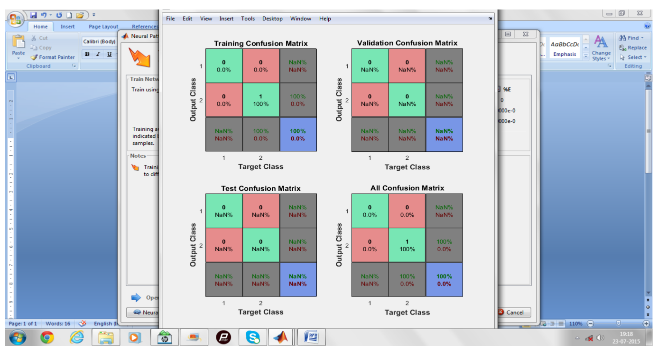

Figure 3: Result using Bachpropogation

REFERENCES

[1] Rajiv Misra and C.R.Manda, ‘‘Performance Comparison of AODV/DSR On-demand Routing Protocols for Ad Hoc Networks in Constrained Situation’’, Indian Institute of Technology, Kharagpur (India), july 2012.

[2] Opnet Technologies, Inc. “Opnet Simulator,” Internet: www.opnet.com, date last viewed: 2009-12-21

[3] Tanu Preet Singh, Dr. R K Singh and Jayant Vats. Article: “Routing Protocols in Ad Hoc Networks: A Review. International Journal of Computer Applications 25(4):30-35”, July 2011. Published by Foundation of Computer Science, New York, USA.

[4] O. K. Tonguz and G. Ferrari, “Ad Hoc Wireless Networks: A Communication-Theoretic Perspective”, John Wiley and Sons, June 2004.

[5] Jyu-Wei Wang Hsing-Chung Chen Yi- ping Lin Dept. Of Inf. & Comm., Asia Univ., Taichung, Taiwan INC, IMS and IDC, 2009. NCM '09. Fifth International Joint Conference on 2079-2084

[6] Maan, F. ; Nat. Univ. Of Sci. & Technol. (NUST), Islamabad, Pakistan ; Mazhar, N. Ubiquitous and Future Networks (ICUFN), 2011 Third International Conference on June 2011

[7] Md. Anisur Rahman, Md. Shohidul Islam, Alex Talevski, ‘‘Performance Measurement of Various Routing Protocols in Ad-hoc Network’’. Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong

[8] Http://www.ietf.org/rfc/rfc2501.txt , date last viewed: 2009-12-21

[9] Samuel Chellathurai, E.George Dharma Prakash Raj. "A Strategic Review of Routing Protocols for Mobile Ad Hoc Networks", International Journal of Engineering Trends and Technology (IJETT), V10(8),390-395 April 2014. ISSN:2231-5381.