International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

A General Study of Associations rule mining in Intrusion Detection System

Vikas Markam [1], Lect. Shirish Mohan Dubey [2]

[1]M.Tech Scholar, Department of Computer Science Engineering, SRIT,Jabalpur, M.P.INDIA
[2]Department of Master of Computer Applications ,SRIT,Jabalpur,MP,INDIA

[1]Vikas.markam@gmail.com
[2]dubey78@gmail.com

***Abstract*—Intrusion Detection Systems (IDSs) can easily create thousands of alerts per day, up to 99% of which are false positives (i.e. alerts that are triggered incorrectly by benign events). This makes it extremely hard for to analyze and react to attacks. Data mining generally refers to the process of extracting models from large stores of data. The intrusion detection system first apply data mining programs to audit data to compute frequent patterns, extract features, and then use classification algorithms to compute detection models. The most important step of this process is to determine relations between fields in the database records to construct features. The standard association rules have not enough expressiveness. Intrusion detection system can extract the association rule with negations and with varying support thresholds to get better performance rather than extract the standard association rule. This paper presents a novel method for handling IDS alerts more efficiently some important features of association rule mining to IDS. In this paper, we integrate fuzzy association rules to design and implement an abnormal network intrusion detection system. Since the association rules used in traditional information detection cannot effectively deal with changes in network behavior, it will better meet the actual needs of abnormal detection to introduce the concept of fuzzy association rules to strengthen the adaptability.**

***Keywords*— internet; intrusion detection using fuzzy rule data mining; data pre-processing; association analysis.**

I. INTRODUCTION

An intrusion detection system (IDS) is a component of the computer and information security framework. Its main goal is to differentiate between normal activities of the system and behaviour that can be classified as suspicious or intrusive [1]. IDS's are needed because of the large number of incidents reported increases every year and the attack techniques are always improving.

IDS approaches can be divided into two main categories: misuse or anomaly detection [1]. The misuse detection approach assumes that an intrusion can be detected by matching the current activity with a set of intrusive patterns (generally defined by experts or "underground" web sites). Examples of misuse detection include expert systems, keystroke monitoring, and state transition analysis. Anomaly detection systems assume that an intrusion should deviate the system behaviour from its normal pattern. This approach can be implemented using statistical methods, neural networks, predictive pattern generation and association rules among others techniques. In this survey i am studying using fuzzy data mining techniques to extract patterns that represent normal behaviour for intrusion detection. In this survey I am describing a variety of modifications that I will have made to the data mining algorithms in order to improve accuracy and efficiency. We are using sets of fuzzy association rules that are mined from network audit data as models of "normal behaviour." To detect anomalous behaviour, I will generate fuzzy association rules from new audit data and compute the similarity with sets mined from "normal" data. If the similarity values are below a threshold value, an alarm will issued. This survey is organized as follows. Section II presents literature review, Section III describes proposed work, discusses some experiments which will produce after implementation of proposed working model, and presents our conclusions.

II. LITERATURE SURVEY AND PROBLEM ANALYSIS

Here a newly developed technique named, "The Research on the Application of Association Rules Mining Algorithm in Network Intrusion Detection" [6] is discussed. In this paper author have describe about Network Intrusion Detection System (IDS), as the main security defending technique, is second guard for a network after firewall. Since it can discern and respond to the hostile behavior of the computer and network resource, it is a hot area for research network security nowadays.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Furthermore author descried about Data mining technology which can be applied to the network intrusion detection, and Precision of the detection will be improved by the superiority of data mining. Basically In this paper, author have presents the study of an example running to contract two algorithms. Presented results have shown that the fuzzy rule mining algorithm is more convenient than Apriori algorithm to mine mass network log database.

In [24] paper, authors have integrated two technique data mining and fuzzy technique. Where fuzzy association rules have applied to design and implement an abnormal network intrusion detection system. Here author presents that when the association rules used in traditional information detection cannot effectively deal with changes in network behavior, it will better meet the actual needs of abnormal detection to introduce the concept of fuzzy association rules to strengthen the adaptability. Basically in This paper author mainly focused on the study of Denial of Service (DOS). According to the author's experimental results, they have found that their system can correctly identify all DOS attacks on test after appropriate adjustment of system parameters. Moreover, they have also proved, in the experiment, that their system would not result in false positives under such circumstances as a large amount of instantaneous FTP normal packet flow. In addition, if source of an attacker can be determined, the system will also be able to promptly inform the firewall to alter its rules and cut off the connection. According to another research network security is becoming an increasingly important issue, since the rapid development of the Internet. Network Intrusion Detection System (IDS), as the main security defending technique, is widely used against such malicious attacks.

Data mining and machine learning technology has been extensively applied in network intrusion detection and prevention systems by discovering user behavior patterns from the network traffic data. Association rules and sequence rules are the main technique of data mining for intrusion detection. Considering the classical Apriori algorithm with bottleneck of frequent item sets mining, author presented a Length-Decreasing Support to detect intrusion based on data mining, which is an improved Apriori algorithm. Experiment results indicate that the author presented method is Efficient [25]. Here another newly developed technique named, "A Study of Intrusion Detection System Based on Data Mining" [26] is discussed. In this paper authors have present classifications of intrusion detection and methods of data mining applied on them were introduced.

Then, intrusion detection system design and implementation of based on data mining were presented. Such a system used APRIORI algorithm to analyze data association, which is the most influencing algorithm in mining Boolean association rules continuity item muster, with recurrence arithmetic based on idea of two period continuity item muster as core. In this paper authors showed that new type of attack can be detected effectively in the system and knowledge base can be updated automatically, so the efficiency and accuracy of the intrusion detection were improved, and security of the network was enhanced.

Here another paper [27] author presents IDS using fuzzy data mining techniques to extract patterns that represent normal behavior. Basically In this paper they have described a variety of modifications that they have made to the data mining algorithms in order to improve accuracy and efficiency. They have used sets of fuzzy association rules that are mined from network audit data as models of "normal behavior." To detect anomalous behavior, they have generated fuzzy association rules from new audit data and compute the similarity with sets mined from "normal" data. If the similarity values are below a threshold value, an alarm is issued. Furthermore In this paper they have described an algorithm for computing fuzzy association rules based on Borgelt's prefix trees, modifications to the computation of support and confidence of fuzzy rules, a new method for computing the similarity of two fuzzy rule sets, and feature selection and optimization with genetic algorithms. They have demonstrated that they can achieve better running time and accuracy with these modifications.Most of the existing commercial NIDS products are signature-based but not adaptive.

In [28], an adaptive NIDS using data mining technology is developed. Data mining approaches are used to accurately capture the actual behavior of network traffic, and portfolio mined is useful for differentiating "normal" and "attack" traffics. On the other hand, most of the current researches are using only one engine for detection of various attacks; the proposed system is constructed by a number of agents, which are totally different in both training and detecting processes. Each of the agents has its own strength on capturing a kind of network behavior and hence the system has strength on detecting different types of attack In addition, its ability on detecting new types of attack as well as a higher tolerant to fluctuations were shown. The experimental results showed that the frequent patterns mined from the audit data could be used as reliable agents, which outperformed from traditional signature-based NIDS [28]

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

In [29], author have presented a hybrid misuse based IDS, using combined structure of an association rule mining algorithm and a connectionist model. The key idea of authors is to take advantage of different classification abilities of knowledge-based and machine learning approaches for different attacks. To lower the computational load of association rule mining, the inputs of rule mining algorithm are selected based on the results of a feature relevance analysis. Experimental results show that the proposed hybrid model, in which knowledge-based section of the system reports hard recognizable attack categories, can improve classification results, especially for remote-to-local (R2L) and user-to-root (U2R) attack classes. This hybrid system also offers better detection rate (DR) and cost per example (CPE) compared to neural-based IDS. False alarm rate (FAR) of the proposed model is comparable with other intrusion detection systems, as well [29].

Here another paper [37] have presents that NAIDS apply the association rule and classification techniques into detecting intrusion behavior among network audit record from a new perspective. Aspect to association rules mining, constructing two mining modes: static mining and dynamic mining; implementing two level mining: single-level mining and domain-level mining. About classification engineering, the mainstream classification techniques were compared through thoroughly experiments, and some improvement was made to decision tree toward the concrete problem, which make NAIDS detect some new type attacks and this kind of capability embodies the advantage of anomaly detection over misuse detection; incremental mining approach was put forward which detect one window data amount, instead of batch of tcp/ip record, which was very suitable to on-line mining and make NAIDS be high real-time performance. Research work on data mining based intrusion detection approaches which has been done belongs to the field of misuse detection in nature, association rules and frequent episodes mining aim to describe the intrusion signature, the ruler classifier was used to mainly detect intrusion behavior. NAIDS is the first data mining based anomaly detection system, the first intrusion detection system which lower false positive rate by classification engineering, the first intrusion detection system which put forward sliding windows techniques to carry out incremental, on-line mining. In principal, dynamic sliding window make NAIDS have the ability of real-time detection; classification engineering make NAIDS keep lower false positive rate.

A large amount experiments on DARPA 1998, 1999 was carried out and the validation and effectiveness of their approach were verified.

Now here I am going to be present some week point of previous research and how I will overcome to these week point. I am also suggesting simple model of proposed Intrusion Detection System Using efficient data mining approach in the next chapter

III. Analysis Of Previous Research:

In research paper titled "The Research on the Application of Association Rules Mining Algorithm in Network Intrusion Detection" [6] has using simple and basic association algorithm like apriori and apply fuzzy set concept to enhance efficiency of association rule mining which is not a good approach to generate frequent item set because there is so many method to produce frequent item set in better way. Another problem we have analyzed that in fuzzy set theory, it is often requires several analysis before the number of frequent item set produce. It can be very sensitive to the choice of initial analysis. Another disadvantage is that it does not yield the same result with each run, since the resulting frequent item set depend on the initial random assignments. Another disadvantage it's not for intrusion detection system because run time efficiency.

Observation on the state of the art

Here I make the following four observations about contemporary data mining efforts in intrusion detection form [26, 27, 28 and 29]:

- Most research concentrates on the construction of operational IDSs, rather than on the discovery of new and fundamental insights into the nature of attacks and false positives.
- It is very common to focus on the data mining step, while the other data mining steps are largely ignored.
- Much research is based on strong assumptions that complicate practical application. Up to now, data mining in intrusion detection focuses on a small subset of the spectrum of possible applications

IV. Proposed Work

In this section we are going to be present simple model of proposed Intrusion Detection System Using efficient data mining approach with fuzzy logic.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Future research directions:

We have discussed these observations in a critical manner, which has leaded us to the following recommendations for further research:

- Future research should pay closer attention to the data mining process.
- Either more work should address the (semi-automatic) generation of high-quality labeled training data, or the existence of such data should no longer be assumed.
- Future research should explore novel applications of data mining that do not fall into the categories feature selection and anomaly detection.
- To deal with some of the general challenges in data mining, it might be best to develop special-purpose solutions that are tailored to intrusion detection.

Proposed Technique: Here I am going to be present general idea on a new proposed technique for intrusion detection system which will enhance efficiency as compare existing intrusion detection system. In the proposed technique Based on the detailed and comprehensive study on data mining based intrusion detection techniques, Proposed Network-based Anomaly Intrusion Detection System (PNAIDS) apply the association rule, Clustering and fuzzy techniques into detecting intrusion behavior among network audit record from a new perspective.**Proposed Model for DM:** Here figure 1 is showing model diagram of proposed IDS

It is important thing that a Data Mining does not identify for the intrusions or anomalies. What it does is identify patterns within the data that it is processing. It is then up to analyze the patterns and come up with the necessary flags to highlight any anomalies.

The different steps involved in the proposed system for anomaly-based intrusion **detection** (shown in figure 2) are described as follows:

(1) Classification of training data
(2) Data pre-processing
(3) Association Rule Mining
(4) Clustering Mining
(5) Strategy for fuzzy rules

Finding an appropriate classification for a test input

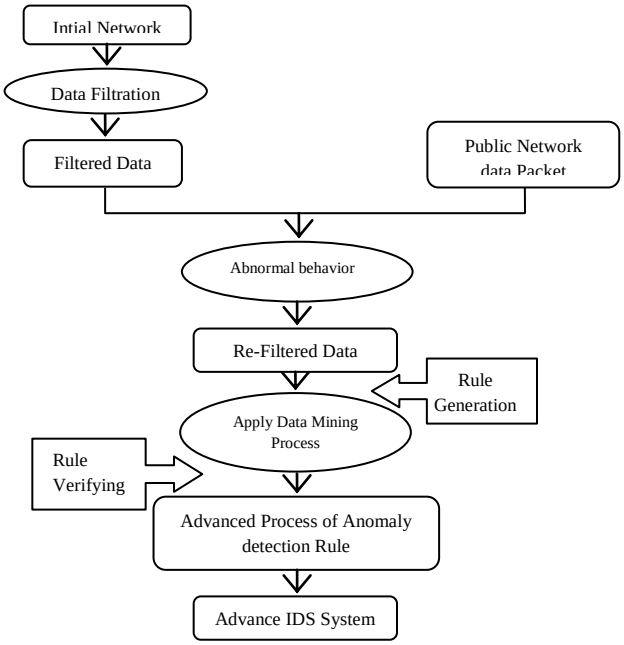


Figure 1: - Proposed Data Mining Model for anomaly Detection System

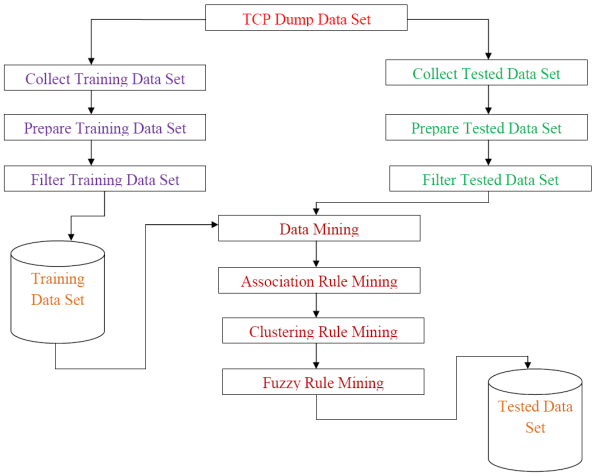


Figure 2:- Proposed IDS Architecture

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Classification of training data and Data pre-processing
Selecting and generating the data source: First the acquisition of data was done. In the case of this research, Sample datasets from DATA- BASE were used. The DATABASE contained high volume network traffic data, and a subset of data ranging a period of 2 – 5 days was selected

Data scope transformation and pre-processing: For the purpose of the research, the scope of the data was limited to tcp/ip packets. **Only** six intrinsic features were extracted from each packet within the dataset. These were timestamp, Source IP, Source Port, Destination IP, Destination Port, and Service. The table 1 below shows the scope of the input dataset.

Table- 1

Flags	**Source**	**Destination**
TCP Port Number	Source TCP Port	Destination TCP Port
IP Address	Source IP Address	Destination IP Address
Timestamp		
Service		

For the purpose of reporting for this paper the data was extracted from the DATABASE set using Excel. But several data extraction tools are available in Public Domain. Excerpts from the data set are included. We extracted the necessary features and saved the data as a comma delimited (db) flat file, so that each line within the file represented one data packet within the dataset. The data was read into a .Net ArrayList, which had as its elements an ArrayList. As the number of data packets in each three minute period can vary considerably, the .Net ArrayList structure was utilized since you don't need to preset its boundaries. Once the data was loaded into the pre-processor, it was prepared for use by the Data Mining approach.

Association Rule Mining and Clustering Mining: Here I am using apriori algorithm for association rule mining technique to produce association rule and to produce cluster on that rule I am using K-Mean clustering technique. There are different types of algorithms used to mine frequent item sets. Some of them, very well known, started a whole new era in data mining. They made the concept of mining frequent item sets and association rules possible. Others are variations that bring improvements mainly in terms of processing time. I will go through some of the most important algorithms apriori algorithm.

It is by far the most important data mining algorithms for mining frequent item sets and associations. It opened new doors and created new modalities to mine the data.

K-Means clustering is a data mining/machine learning algorithm used to cluster observations into groups of related observations without any prior knowledge of those relationships. The k-means algorithm is one of the simplest clustering techniques and it is commonly used in medical imaging, biometrics and related fields.

The k-means Algorithm: The k-means algorithm is an evolutionary algorithm that gains its name from its method of operation. The algorithm clusters observations into k groups, where k is provided as an input parameter. It then assigns each observation to clusters based upon the observation's proximity to the mean of the cluster. The cluster's mean is then recomputed and the process begins again. Here's how the algorithm works:

- The algorithm arbitrarily selects k points as the initial cluster centers ("means").
- Each point in the dataset is assigned to the closed cluster, based upon the Euclidean distance between each point and each cluster center.
- Each cluster center is recomputed as the average of the points in that cluster.
- Steps 2 and 3 repeat until the clusters converge. Convergence may be defined differently depending upon the implementation, but it normally means that either no observations change clusters when steps 2 and 3 are repeated or that the changes do not make a material difference in the definition of the clusters.

Strategy for Fuzzy Rules: This section describes the designed strategy for automatic generation of fuzzy rules to provide effective learning. In general, the fuzzy rules given to the fuzzy system is done manually or by experts, who are given the rules by analyzing intrusion behavior. But, in our case, it is very difficult to generate fuzzy rules manually due to the fact that the input data is huge and also having more attributes. But, a few of researches are available in the literature for automatically identifying of fuzzy rules in recent times. Motivated by this fact, we make use of mining methods to identify a better set of rules. There are four rule I using which is following:

- Rule 1-: If Flag contain SYN & FIN then Abnormal Packet
- Rule 2-: If Flag == NULL Then Abnormal Packet
- Rule 3-: If Flag contain combination of SYN, SYN ACK, & ACK Then Normal Packet
- Rule 4-: If REJ Packet comes many times then Abnormal

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Proposed Algorithm-: Proposed algorithm is derived in phases first phase is for mining technique where I am using K-Mean clustering technique and in second phase I am using fuzzy rule.

Applying K-Mean Clustering

- Step 1: Place randomly initial group centroids into the 2nd space.
- Step 2: Assign each object to the group that has the closest centroid.
- Step 3: Recalculate the positions of the centroids.
- Step 4: If the positions of the centroids didn't change go to the next step, else go to Step 2.
- Step 5: End.

Applying Rules-:

- Rule 1-: If Flag contain SYN & FIN then Abnormal Packet
- Rule 2-: If Flag == NULL Then Abnormal Packet
- Rule 3-: If Flag contain combination of SYN, SYN ACK, & ACK Then Normal Packet
- Rule 4-: If REJ Packet comes many times then Abnormal

Explanation-: Working of fuzzy rule in proposed concept is defined following with explanation

- Rule 1-: If flag contain syn and fin attribute in the TCP frame format then it is abnormal because syn attribute is used to start the connection while fin attribute is used to end an existing connection. So it does not make any séance to perform both actions.
- Rule 2-: Some packet has absolutely no flags attribute set at all they are referred as null packet it is illegal to have packet with no flags set.
- Rule 3-: Syn , syn ack & ACK all these four attribute can be find in TCP frame format. If these entire attribute are present in TCP frame format then this combination will represent normal behavior of received packet because it is used in 3 way handshaking.
- Rule 4-: If Rej flags attribute find in TCP frame format then this will represent as a abnormal behavior of received packet**.** This type of attribute can be occur many time in TCP frame format so carefully scan TCP frame format.

V. RESULTS COMPARISON

I am using .Net implementation to present an evaluation system. For timing evaluation of the fuzzy rule algorithm and my suggested algorithm, it is necessary to describe the detailed evaluation method, as illustrated in Figure-3 and Figure-.4

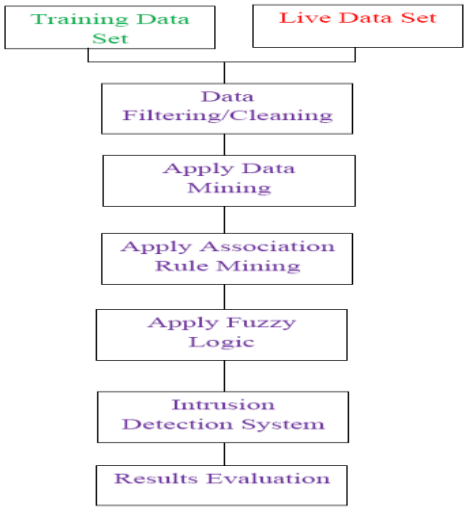


Figure 3: Evolution Mode of Fuzzy Logic

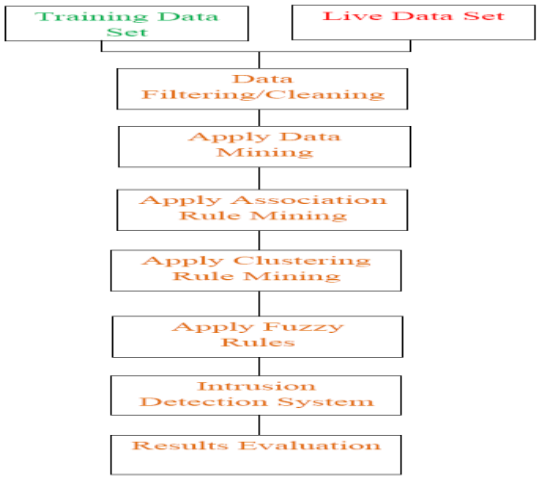


Figure 4: Evolution Mode of Proposed Concept

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

For our experiment, we use a laptop Pentium® Dual-Core CPU T4400 @2.20Ghz and 32-bit operating system, in which performance data is collected. In the experiments, the laptop executes different file size ranges from 2 K to 14 K record data sets. Several performance metrics are collected:

- Packet Performance
- Execution time
- CPU Utilization time
- Memory Utilization

Packet Performance: Packet performance is considered

Execution Time: - The execution time is considered the time that an algorithm takes to produce results. Execution time is used to calculate the throughput of an algorithm. It indicates the speed of algorithm.

Memory Utilization: - The memory deals with the amount of memory space it takes for the whole process of Intrusion Detection System.

CPU Utilization: - The CPU Utilization is the time that a CPU is committed only to the particular process of calculations. It reflects the load of the CPU. The more CPU time is used in the execution process, the higher is the load of the CPU.

Here I am doing compare packet performance, time-consuming, memory utilization and CPU utilization of known algorithm on different size of record sets. During processing, the record sets are coming from data base. For evaluation mode, there are two parameters: the number of evaluated record set and the size of evaluated record set, where the number of evaluated record sets is the number of record set that are generated randomly and the size of evaluated record sets can be chosen from database. In this mode, I do n cycles (that is, the number of the evaluated record sets). In each cycle, record sets are respectively executed by fuzzy logic and our proposed concept by copying them. Finally, the outputs of the evaluation system are packet performance, execution time, and the execution time is measured in seconds. Actually, for an algorithm, the time-consuming of execution not only depends on the algorithm's complexity, but also the size of record sets. The evaluated results are illustrated as in Table 2 – 5

Table 2 is showing normal packet performance compression on the basis of different record sizes between fuzzy logic and proposed concept.

S .No	FILE SIZE	Fuzzy Logic	Proposed Algo.
1	2000 records	89.61%	90.13%
2	5000 records	90.13%	94.86%
3	10000 records	96.55%	97.04%
4	14000 records	96.94%	97.28%
5	20000 records	92.3%	95.2%
6	30000 records	96.1%	97%
7	40000 records	95%	95.7%
8	50000 records	96.6%	98%

Here table-3 is representing execution time which is measuring in second. In this table, the evaluated mode is Different Size of Record Sets.

S .No	Name of Concept	Available Memory	Used Memory
1	Fuzzy Logic	1008436	617184
2	Proposed Concept	1008436	619012

Here table-4 is representing Memory Utilization by the executed algorithm at run time which is measuring in KB. In this table, the evaluated mode is Different Record Sets.

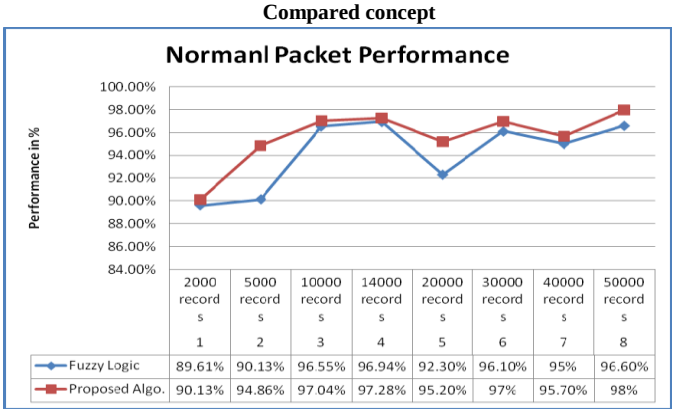
S .No	FILE SIZE	Fuzzy Logic	Proposed Algo
1	2000 records	16 Sec	7 Sec
2	5000 records	20 Sec	10 Sec
3	10000 records	24 Sec	13 Sec
4	14000 records	29 Sec	20 Sec

Here table-5 is representing CPU Utilization by the executed algorithm at run time which is measuring in %. Here I am drawing the graph-1 from table 2 to reveal it. In these graph the evaluated mode is different size of record sets ranging from 2 to 14 thousand.

S .No	Name of Concept	CPU at Initial Level	CPU During Execution
1	Fuzzy Logic	0%	54%
2	Proposed Concept	0%	54%

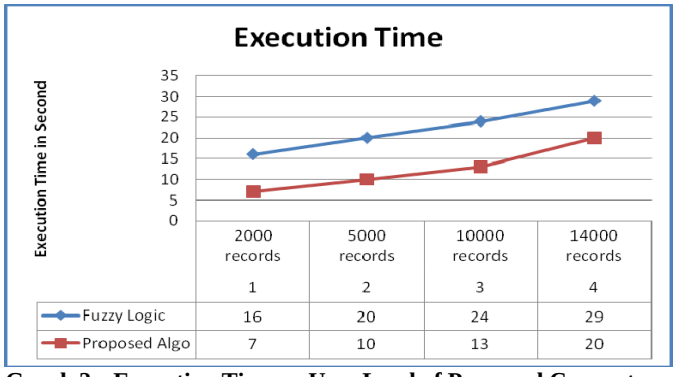
International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Graph 1:- Packet Performance between Proposed Concept with Compared concept



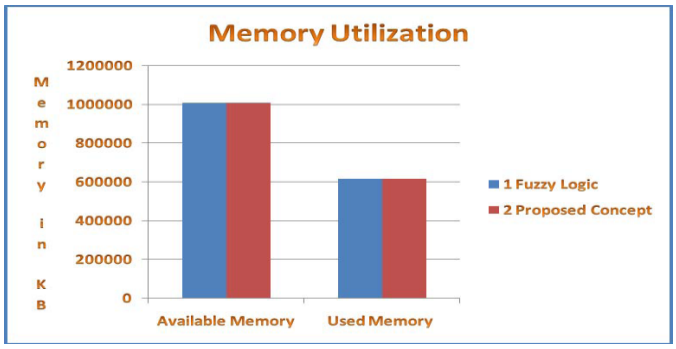
	2000 records 1	5000 records 2	10000 records 3	14000 records 4	20000 records 5	30000 records 6	40000 records 7	50000 records 8
Fuzzy Logic	89.61%	90.13%	96.55%	96.94%	92.30%	96.10%	95%	96.60%
Proposed Algo.	90.13%	94.86%	97.04%	97.28%	95.20%	97%	95.70%	98%

Here I am drawing the graph-2 form Table-3 to reveal it. In this graph, the evaluated mode is different size of record sets ranging from 2 to 14 thousands



	2000 records 1	5000 records 2	10000 records 3	14000 records 4
Fuzzy Logic	16	20	24	29
Proposed Algo	7	10	13	20

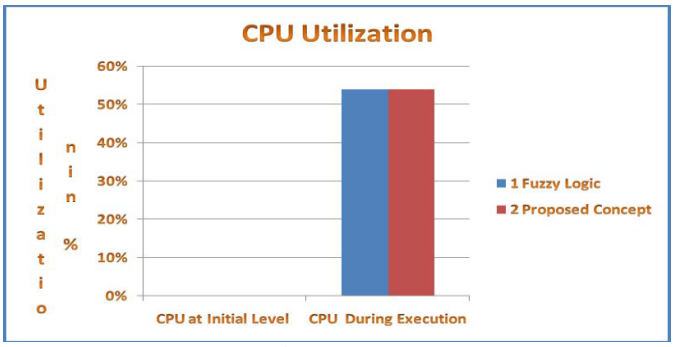
Graph 2:- Execution Time vs User Load of Proposed Concept and Compared Concept in Linear way



Compared Concept in Linear way

Here I am drawing the graph-3 forms Table-4 to reveal it. In this graph, the evaluated mode is fixed size of record sets (14528).

Here I am drawing the graph-4 forms Table-4 to reveal it. In this graph, the evaluated mode is fixed size of record sets (14528).



Graph 4: CPU Utilization of Proposed Concept and Compared Concept

Experimental results for this compassion point are shown Table 2 to 4 at execution stage. The results show the superiority of our suggested algorithm over existing algorithms in terms of the packet performance, processing time, Memory Utilization and CPU Utilization.

Some typical results obtained by the evaluation system can be found in Table 2 and Graph 1. The results illustrated in Table 2 show that our suggested concept is 98% better than compared concept in different record sets. Finally, it is not difficult to find that, in contrast with these tables, the larger the data record sets, the bigger execution time is. Besides, in contrast with these tables, it is not difficult to find that the increasing data length can lead to the significant increment of execution time as well as memory utilization and CPU Utilization.. Generally speaking, the time-consuming of known algorithm usually depends on the size of record sets of

Reasons for Supremacy over other algorithms:-
My proposed concept is better then compared concept to find normal packet performance

- My proposed concept is faster then compared concept in terms of execution time.
- Proposed concept is much smaller than the compared concept and easy to understand and implement.

It does not contain complex structure, control flow is well defined and looping structure is minimized. Due to the following facts it takes very less time for execution

VI. CONCLUSION

This paper improve detecting speed and accuracy as a goal, and proposed a more efficient association rules mining method as comparing algorithm to abnormal detecting experiment based on network, finally I am comparing and analyzing the experiment result. The experiment result shows, using the proposed algorithm raised by this paper, the detecting effectiveness and accuracy will improve shows satisfying result. How to further improve the detection performance and opimize the algorithum is our future work

References

[1] J. Allen, A. Christie, W. Fithen, J. McHugh, J. Pickel, and E. Stoner.2000. State of the Practice of Intrusion Detection Technologies. CMU/SEI- 99-TR-028. Carnegie Mellon Software Engineering Institute

[2] R.Agrawal,, and R. Srikant. 1994. Fast algorithms for mining association rules. In Proceedings of the 20 th international conference on very large databases held in Santiago, Chile, September 12-15, 1994, 487-99.

[3] C. Borgelt, 2001. Association Rule Induction. http://fuzzy.cs.unimagdeburg. de/~borgelt.

[4] S.M. Bridges, and Rayford M. Vaughn. 2000. Fuzzy data mining and genetic algorithms applied to intrusion detection. In Proceedings 23 rd National Information Systems Security Conference, Oct. 16-19, 2000, Baltimore, MD, pp. 13-31.

[5] J. Han, and M. Kamber. Data mining: Concepts and techniques. Morgan Kaufmann. San Francisco, CA. 2001.

[6] Ye Changguo, Wei Nianzhong, Wang Tailei "The Research on the Application of Association Rules Mining Algorithm in Network Intrusion Detection" Published in 2009 First International Workshop on Education Technology and Computer Science. 978-0-7695-3557-9/09 IEEE.

[7] Liu, Bing: Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Springer, 2007

[8] Hansen, Hans Robert; Neumann, Gustaf: Wirtschaftsinformatik I. Lucius & Lucius, 2001.

[9] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic: Knowledge Discovery and Data Mining: Towards a Unifying Framework. In Proceeding of The Second Int. Conference on Knowledge Discovery and Data Mining, pages 82--88, 1996.

[10] Hipp, Jochen; Guentzer,Ulrich; Nakhaeizadeh, Gholamreza: Algorithms for Association Rule Mining - A General Survey and

[11]http://wires.wiley.com/WileyCDA/WiresArticle/wisId-WICS161.html

[12] http://www.infosyssec.net/infosyssec/security/intdet1.htm

[13] Hansen, Hans Robert; Neumann, Gustaf: Wirtschaftsinformatik I. Lucius & Lucius, 2001.

[14] Liu, Bing: Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Springer, 2007.

[15] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic: Knowledge Discovery and Data Mining: Towards a Unifying Framework. In Proceeding of The Second Int. Conference on Knowledge Discovery and Data Mining, pages 82--88, 1996.

[16] Frawley, William J.; Piatetsky-Shapiro, Gregory; Matheus, Christopher J.: Knowledge Discovery in Databases: an Overview. AAAI/MIT Press, 1992.

[17] Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic: The KDD Process for Extracting Useful Knowledge from Volumes of Data. Communications of the ACM, Volume 39, Issue 11 Pages: 27 - 34, 1996

[18] Sumathi, S.; Sivanandam, S. N.: Introduction to Data Mining and its Applications. Springer, 2006.

[19] Fayyad, Piatetsky-Shapiro, Smyth: From Data Mining to Knowledge Discovery in Databases. AI MAgazine, 1996.

[20] Roiger, Richard J.; Geatz, Michael W.: Data Mining: A Tutorial-Based Primer. Addison Wesley, 2003

[21] Agrawal, Rakesh; Imielinski,Tomasz; Swami, Arun: Mining Association Rules between Sets of Items in Large Databases. Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, 1993.

[22] Zadeh, Lofti A.: Fuzzy Logic. University of California, Berkeley, 1988.

[23] Hipp, Jochen; Guentzer,Ulrich; Nakhaeizadeh, Gholamreza: Algorithms for Association Rule Mining - A General Survey and Comparison. ACM SIGKDD Explorations Newsletter, Volume 2 , Issue 1, 2000.

[24] Ma yanchun "The Intrusion Detection System Based on Fuzzy Association Rules Mining" published in IEEE Conferences in 2010

[25] Lei Li, De-Zhang Yang, Fang-Cheng Shen "A Novel Rule-based Intrusion Detection System Using Data Mining" published in IEEE Conferences in 2010

[26] Chunyu Miao and Wei Chen "A Study of Intrusion Detection System Based on Data Mining" published in IEEE Conferences in 2010

[27] German Florez, Susan M. Bridges, and Rayford B. Vaughn "An Improved Algorithm for Fuzzy Data Mining for Intrusion Detection" This work was supported in part by the National Science Foundation Grant CCR-9988524 and the Army Research Laboratory Grant DAAD17-01-C-0011

[28] Cheung-Leung Lui , Tak-Chung Fu and Ting-Yee Cheung "Agent-based Network Intrusion Detection System Using Data Mining Approaches" published in Proceedings of the Third International Conference on Information Technology and Applications (ICITA'05) 2005 IEEE

[29] Mansour Sheikhan and Zahra Jadidi "Misuse Detection Using Hybrid of Association Rule Mining and Connectionist Modeling" published in World Applied Sciences Journal 7 (Special Issue of Computer & IT): 31-37, 2009 ISSN 1818-4952 © IDOSI Publications, 2009

[30] Anshu Veda, Prajakta Kalekar and Anirudha Bodhankar "Intrusion Detection Using Datamining Techniques"

[31] Marin J., Ragsdale D., Surdu J.: A Hybrid Approach to the Profile Creation and Intrusion Detection. Proceedings of the DARPA Information Survivability Conference and Exposition – DISCEX 2001,June2001.http://www.itoc.usma.edu/Documents/Hybrid_DISCEX_AcceptedCopy.pdf

[32] Fan W., Miller M., Stolfo S., Lee W., Chan P.: Using Artificial Anomalies to Detect Unknown and Known Network Intrusions. In Proceedings of the First IEEE International Conference on Data Mining, San Jose, CA, November 2001, http://www.cc.gatech.edu/~wenke/papers/artificial_anomalies.ps

[33] Lee W., Stolfo S, Mok K.: Adaptive Intrusion Detection: a Data Mining Approach**.** Artificial Intelligence Review, 14(6), December 2000,pp.533-567, http://www.cc.gatech.edu/~wenke/papers/ai_review.ps

[34] Lee W. i inni: A data mining and CIDF based approach for detecting novel and distributed intrusions. Recent Advances in Intrusion Detection, Third International Workshop, RAID 2000, Toulouse, France, October 2-4, 2000, Proceedings. Lecture Notes in Computer Scienc**e** 1907 Springer, 2000, pp. 49-65. http://www.cc.gatech.edu/~wenke/papers/lee_raid_00.ps

[35] Bass T.: Intrusion Detection Systems Multisensor Data Fusion: Creating Cyberspace Situational Awareness. Communication of the ACM, Vol. 43, Number 1, January 2000, pp. 99-105**,** http://www.silkroad.com/papers/acm.fusion.ids.ps.

[36] Manganaris S., Christensen M., Zerkle D., Hermiz K.: A data mining analysis of RTID alarms. Computer Networks, 34, 2000, pp. 571-577.

[37] http://www.it-paper.com/study-on-data-mining-based-intrusion-detection-approaches-and-system.html

[38] L.A. Zadeh, Fuzzy Sets, Information and Control, **8**(3), 338-353, 1965.

[39] J.E. Dickerson, J. Juslin, O. Loulousoula, and J. A. Dickerson, Fuzzy Intrusion Detection, *IFSA World Congress and 20th North American Fuzzy Information Processing Society (NAFIPS) International Conference*, 2001, pp1506-1510.

[40] Susan M. Bridges and Rayford B. Vaughn "INTRUSION DETECTION VIA FUZZY DATA MINING" Accepted for Presentation at The Twelfth Annual Canadian Information Technology Security Symposium June 19-23, 2000, The Ottawa Congress Centre

[41] http://www.seattlerobotics.org/encoder/mar98/fuz/flindex.html

[42] Chris Clifton, Gary Gengo. "Developing Custom Intrusion Detection Filters Using Data Mining". MILCOM 2000. 21st Century Military Communications Conference Proceedings, Volume: 1, 22-25 Oct. 2000. Pagees: 440 – 443 Vol. 1

[43] Heikki Mannila, Hannu Toivonen, and A. Inkeri Verkamo, "Discovery of Frequent Episodes in Event Sequences", Data Mining and Knowledge Discovery l(3): 259-289 (1997)

[44] Daniel Barbara, Julia Couto, Sushil Jadodia, Ningning Wu. "ADAM: A Testbed for exploring the Use of Data Mining in Intrusion Detection". *ACM SIGMOD RECORD*