International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-6 Issue-5, June 2017

Packet Dropping and Intrusion Detection using Forensic and Flow Based Classification Techniques

Shyju S., Prathibha S Nair

Abstract— Internal Intrusion detection is one of the serious problems in the computer network areas. Most of the computer system uses username and password as login pattern to enter in to the system. This is one of the weakest points of computer security. Some studies claimed that analyzing system calls (SCs) generated by commands can identify these commands and obtains the features of an attack. This paper propose a security system, named the Internal Intrusion Detection and Protection System(IIDPS) to detect insider attacks at SC level by using data mining and forensic techniques in networked data. The IIDPS creates users' personal profiles to keep track of users' usage habits as their forensic features and determines whether a valid login user is the account holder or not by comparing users current computer usage behaviors with the patterns collected in the account holder's personal profile. The idea behind the inside attacker detection in wireless sensor network by exploiting the spatial correlation between the packet ratio, which help to detecting dynamic attacking behaviors The routing is performed to identify the shortest path between each source node and their destination address and residual energy is calculated for each node in the network.

Index Terms— Insider attacks, intrusion detection, Flow based classification and System calls.

I. INTRODUCTION

Security has been a prime aspect to be taken care of in the computer network domain. The most difficult attack to be detected is insider attacks among pharming attack, distributed denial-of-service attacks (DDoS), eavesdropping attack and spear-phishing attack. The insider attacks are a greater malware where security is concerned. Today, most systems use user ID and password as a login pattern. Most Intrusion detection systems monitor network and system activities in order to avoid attacks and malicious activities that can come from within a network domain. Typically, the intrusion detection systems will monitoring the network and notify the network administrator when any suspicious activity has been detected. In some cases, the intrusion detection systems can take corresponding actions when any problems are detected such as barring a user or IP address from accessing the system in the networked domain.

Hence, in the designed system, the Internal Intrusion Detection and Protection System detect important intruder behaviors launched at system level.

Manuscript published on 30 June 2017.
* Correspondence Author (s)
Shyju S, M.Tech, Department of Computer Science and Engineering, Mohandas College of Engineering and Technology Trivandrum (Kerala), India, E-mail: shyjus15@gmail.com
Prathibha S. Nair, Department of Computer Science and Engineering, Mohandas College of Engineering and Technology, Trivandrum (Kerala), India. Email: prathibhanlaiju@gmail.com

The forensic features of the user, defined by the particular system call pattern they follow are recorded for reference of the identity of the user and are determined from the user‘s log history. After identifying user‘s usage habits, the corresponding System calls (SC) are analyzed to enhance the accuracy of attack detection. Intrusion Detection rate are handled based on the category in which any anomaly belongs to. If the anomaly is from the networked data, then that should be prevented from entering in to the node and from forwarding to another node in the network. In these situations we cannot delete that packet because the packet is created by some other node in the network .If any abnormal anomaly is found inside the node that should be deleted immediately and should prevent from forwarding to other nodes. For avoiding the entry of anomalies from the network to any packet all the packets should satisfy no backtracking property in the network. The Port scanning details can be availed by continuously monitoring all the open ports in the system. Each individual computer will be runs on different ports. The Port Scanning is the name for the technique used to identify open ports and services available on a network host in the network. Most hackers utilize port scanning because it is the simplest way in which they can quickly discover services they can break into. The hackers can even open the different ports themselves in order to access the computer nodes. At any time, there are open ports on all personal computers, there is potential for the loss of data and services, the occurrence of a virus, and at times, even complete systems will be compromise. It is essential for one to protect user’s virtual files and data’s, as new security risks in personal computers are discovered and every system protection must be the number one priority for those who use computers. Port scanning method is considered as a serious threat to one's PC, as it can occur without producing any outward signs to the owner that anything dangerous is taking place. If any foreign node is continuously trying to access any open port in a node then that should be dismissed. In this paper the main idea behind the inside attacker detection in wireless sensor network by exploiting the spatial correlation between the packet ratio, which help to detecting dynamic attacking behaviors The routing is performed to identify the shortest path between each source node and their destination and residual energy is calculated for each node in the network.

II. RELATED WORK

Most of the intrusion detection techniques focus on how to find malicious network domain behaviors [1], [2] and acquire the characteristics of attack packets and attack patterns, based on the histories recorded in user’s log files [3], [4].

Qadeer *et al.* [5] used self developed packet sniffer to collect network packets with which to discriminate network attacks with the help of network nodes states and packet distribution. O' Shaughnesy and Gray [6] acquired network intrusion and attack patterns from computer log files. These files contain traces of system misuse. In this from the synthetically generated log files, the traces or patterns of misuse can be more correctly reproduced. Wu and Banzhaf [7]. applying methods of computational intelligence including artificial neural network, fuzzy system, evolutionary computation, artificial immune systems, and swarm intelligence to detects malicious behaviors in the system. The authors summarized and compared different intrusion detection methods and thus allowing us to clearly view those existing research challenges.

These techniques and applications truly contribute to network domain security. But, they cannot easily authenticate remote login users and detect intrusions, e.g., when an unauthorized user logs in to a computer with a valid user id and password. Hu *et al.* [8] proposed intelligent lightweight IDS that use a forensic technique to profile user behaviors and a data mining technique to carry out different attacks. The authors claimed that the systems could detect intrusions effectively in real time. The Model based anomaly detection systems prevent program execution by a predefined method of allowed system call sequences. These methods are useful only if they detect actual attacks in the systems. These methods are manually constructed mimicry attacks that avoided detection by hiding a malicious series of system calls within a valid system call sequence allowed by the model. Jonathon T. Giffin, Somesh jha, and Barton P. Miller [9] proposed automated discovery of Mimicry Attacks. It is an example for integrating computer forensics with a knowledge based systems. The system uses a predefined model which allowing system call sequences to be normally executed and employed by a detection system to restrict program execution to ensure the security of the protected system. This paper contributes two functions like automated discovery of mimicry attacks and a system design where attack sequences and obfuscations need not be known. This is helpful in detecting applications that issue a series of malicious system calls and identifying attack sequences having been collected in knowledge bases. When an undetected attack is presented the system finds the attack sequence in 2 s as its computation overhead. The drawback of this model is that the attackers have more freedom in program models that do not constrain system call arguments.

Ugo Fiore, Francesco Palmieri, Aniello Castiglione and Alfredo DeSantis [10] proposed a method that explored the effectiveness of a detection approach based on machine learning using the discriminative restricted Boltzmann machine to combine the power of generative model with more classification accuracy. It expresses the capabilities to infer part of its knowledge from partial training data so that the network anomaly detection can provide an adequate degree of protection from both external and internal menaces in the network. The Discriminative restricted Boltzmann machine has been chosen for its ability to combine the generative power to capture the aspects of the normal traffic class and classification accuracy. The detection capability in the network is directly associated with the correctness of the underlying self learnt traffic model. In this the training set does not accurately represents the real network normal traffic. It may overestimate or underestimate anomalous phenomena in the system.

The Advanced metering infrastructure (AMI) is responsible for collecting, measuring, analyzing energy usage data and transmitting the information from a smart meter to a data concentrator and then to a head end system in the utility side the security of AMI is of great concern in smart grid's deployment. Mustafa Amir Faisal, Zeyar Aung, John R. Williams, and Abel Sanchez analyzed the possibility of using data stream mining to enhance the security of advanced metering infrastructure through IDS. The advanced metering infrastructure, which is one of the most crucial components of smart card, serves as a bridge for providing bidirectional information flow between the user domain and the utility domain. Several obvious issues like characteristic of traffic in AMI, co-ordination among the IDSs, registering dynamic device to smart meter, designing special mining techniques for meeting specific requirement for each component in AMI, etc. come in light from this methods. The methods used as an IDS as a second-line security measure after the first line of primary advanced metering infrastructure security techniques such as encryption, authorization, and authentication.

Karen A. Garc´ıa, Ra´ul Monroy, Luis A. Trejo, Carlos Mex-Perera, and Eduardo Aguirre proposed a novel approach for postmortem intrusion detection, which factors out repetitive behavior, thus speeding up the process of locating the execution of an intrusion. Central to our intrusion detection mechanism is a classifier, which separates abnormal behavior from normal one. The method used a novel approach to host based postmortem intrusion detection method which factors out spurious, repetitive behavior to quickly locate the execution of an exploit. Postmortem intrusion detection method is very valuable for computer postmortems because it speeds up the process of gathering evidence of an intrusion in the system. It speeds up the process of building an attack signature. An attack signature is very valuable for intrusion detection system construction, especially in the context of commercial IDS. Postmortem intrusion detection is complex, given both the overwhelming length of a standard log file and the difficulty of identifying exactly where the intrusion has occurred.

Security attacks against wireless LAN can be classified as active and passive [11–13].The passive attacks are silent in nature and are used to extract important information from the network domain. The passive attacks do not harm the network node or network resources. The active attacks are used to misdirect or drop packets. The unique characteristics such as wireless medium, contention based medium access; multi hop nature and random deployment of such networks make them more vulnerable to security attacks at various layers of the network.

III. PROPOSED SYSTEM

The block diagram of the proposed system is shown in fig.1

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

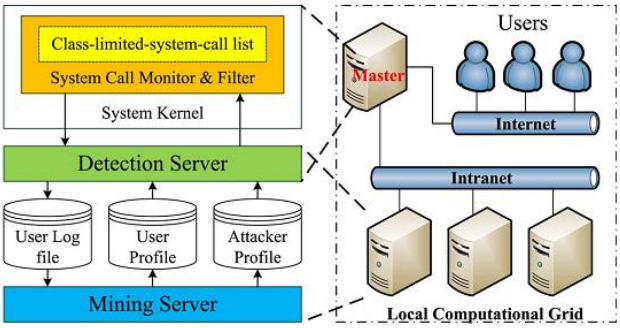

Fig.1. Block Diagram of proposed system

The system consists of an SC monitor and filter, a mining server, a detection server, a local computational grid, and three repositories, including user log files, user profiles, and an attacker profile.

A System Call is an interface between a user application and services provided by the system kernel. The statistical model of term frequency-inverse document frequency (TFIDF) is used to sort the required system calls collected in the user's log file.

$$TF_{i,j} = \frac{n(i,j)}{\sum_{k=1}^{k=h} n_{k,j}} \qquad (1)$$

$$IDFi = \log |D| / |\{j : t \in dj\}| \qquad (2)$$

$$(TF\text{-}IDF)i,j = TFi,j \times IDFi. \qquad (3)$$

The mining server extracts SC-sequence generated by the user's log file. The similarity weights of SC Patterns are matched to remove commonly used SC patterns. Further, the result is matched with other users' habit file to identify user's specific behavior in the SC-patterns. The SCs which are collected in user's log file are processed by the internal intrusion detection and protection system with a sliding window, named as log sliding window (L-window). Another window of same size is used called compared sliding window (C-window) is used to identify other patterns in the user's log file.

Input: u's log file where u is a user of the underlying system

Output: u's habit file.

- G = |log file| - |sliding window|

  /*|sliding windows|=|L-window|=|C-window|/*

- for(i = 0; i <= G-1; i ++){
- for(j=i+1; j <= G; j++){
- for (each of $\sum_{k=2}^{|\text{sliding window}|}$(|sliding window| - k+1) k-grams in current L-window){
- for (each of $\sum_{k'=2}^{|\text{sliding window}|}$(|sliding window| - k'+1) k'-grams in C-window){
- Compare the k-grams and k'-grams with the longest common Subsequence algorithm.
- if (the identified SC-pattern already exists in
- the habit file)
- Increase the count of the SC-pattern by one;
- else
- Insert the SC-pattern into habit file with count = 1;}}}}

Fig.2. Algorithm to generate user's habit file

The detection server captures the system calls sent by the user to the server when the user is executing shell commands and are stored in user's log file. The server tries to check whether the user is actual account holder or not by comparing the similarity scores [14] between newly generated system calls and the user's usage habits files.

$$Sim(u,j) = \sum_{i=1}^{p} F_{iu} \cdot W_{ij} \qquad (4)$$

$$W_{ij} = \frac{f_{ij}}{fij+0.5+\frac{1.5*nsj}{nsavg}} * \frac{\log\left(\frac{N+0.5}{Mi}\right)}{\log(N+1)} \qquad (5)$$

The concept of the detection Server is same as the mining server the only difference is that the comparison between L-window and C-window is from the back to front each time when a system call is input by the user.

Input: - user u's current input SCs, i.e NSC_u (each time only one SC is input) and all user's user profiles

Output: - u's is suspected as an internal intruder

- $NCS_u = \Phi$;
- while(receiving u's input SC, denoted by h) {
- $NCS_u = NCS_u \cup \{h\}$;
- if (| NCS_u| > |sliding window|) {
- L-window = Right(NCS_u; |sliding window|;
- for(j=|NCS_u| - |sliding window|; j>0; j--) {
- C-window = Mid(NCS_uj, |sliding window|);
- Compare k-grams and k'grams by using comparison logic employed in Algorithm 1 to generate NHF_u}
- for (each user g, 1<=g<=N)
- Calculate the similarity score Sim(u, j) between NCS_u and g's user profile by invoking equation (4).
- if((|NCS_u| mod paragraph size) == 0) { /*paragraph size = 30, meaning we judge whether u is an attacker or the account holder for every 30 input SCs*/
- Sort similarity scores for all users;
- if (((the decisive rate of u's user profile < threshold;) or (the decisive rate of attacker profile > threshold;)){
- Alert system manager that u is a suspected attacker, rather than u himself ;}}}}

Fig. 3 – Detection server detects whether user is a possible intruder.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

The different modules in the proposed system consist of

- Network Monitoring
- Packet Monitoring
- System Monitoring
- Signature Based Packet Filtering
- Policy Monitoring
- Remote Monitoring
- Intrusion Detection

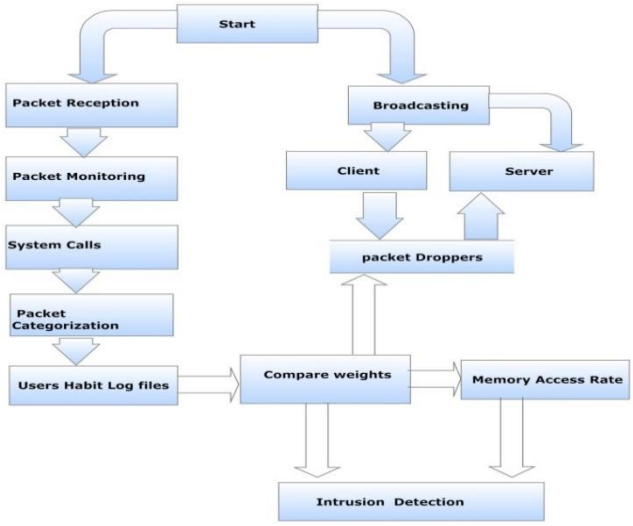

Fig.4. Design flow of proposed system

A. Network Monitoring System

In network monitoring it collects the information about the user details and peer details. In user details it collects the IP address, user name and login time. In peer details it collects the IP address of the system and its connection status like success or failed.

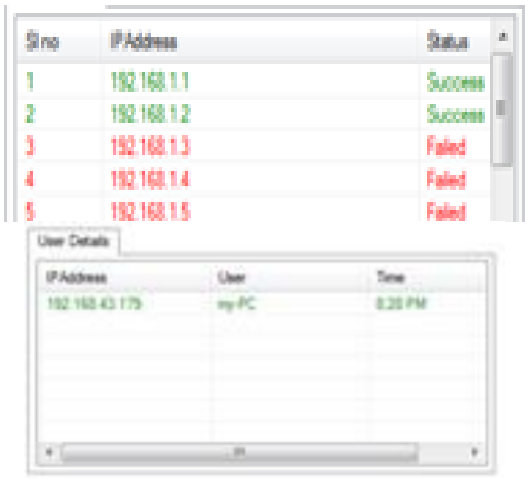

Fig : User details and peer details

B. Packet Monitoring System

In packet monitoring it collects the packet information's. In this the packet classification is performed based on the different protocols. The packet classification is performed on the basis of IP Packet, TCP packet and UDP packets. It also collects the packet information's like source IP, destination IP, protocol type etc. In packet monitoring it also collects the information's about the ports like its port number, protocol, and its status. The user habit file is generated based on the corresponding IP addresses. The Deterministic Packet Marking (DPM) is used for packet monitoring.

The main goal of DPM, that issue was that each packet in a datagram network is being routed individually, so even if the sources and the destinations of the packets are the same, they may be routed along different paths. This feature of packet networks may prevent the attack path reconstruction by the victim, using the PPM algorithm. Since each packet may travel a different route from the same source to the same destination, the only address in the network path that is surely the same for all packets is the ingress interface IP address of the closest router to the source of packets. The main idea behind the DPM approach is that the ingress interface IP address of the closest router to the source of the packet is enough to find the attacker network. It should be noted here that in the current Internet network, packet routing is mostly stable. However, there is still this potential to route the packets from different paths. Only the ingress interfaces of the edge router marks the packets, and the rest, including the backbone routers, are exempt. DPM uses 17 bits of the IP header, including the 16-bitsIdentification field and the 1-bit reserved flag, to embed the marking information to every packet.

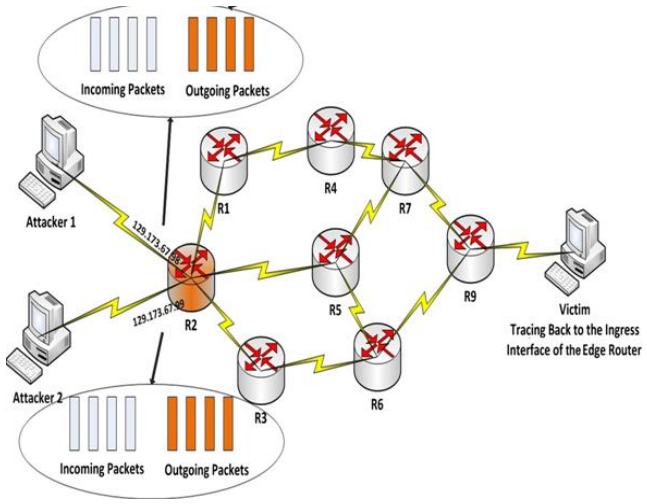

Fig: Deterministic Packet Marking

The 32-bit ingress interface IP address is split into two segments, 16 bits each: segment 0-bits 0 through 15, and segment 1 − bits 16 through 31. When a packet passes through an edge router, one segment is selected with equal probability and inserted to the Identification field. The victim maintains a table matching the source addresses to the ingress addresses When the victim gets both segments of an edge router, then it is able to reconstruct the whole ingress interface IP address of that router. The 1-bit reserved flag plays the role of a sign for the victim to identify which part of the IP address is carried by the current packet.

C. System Monitoring

In system monitoring it collects the information's like client information's, Memory information's, Drive information's and installed programs in the system.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

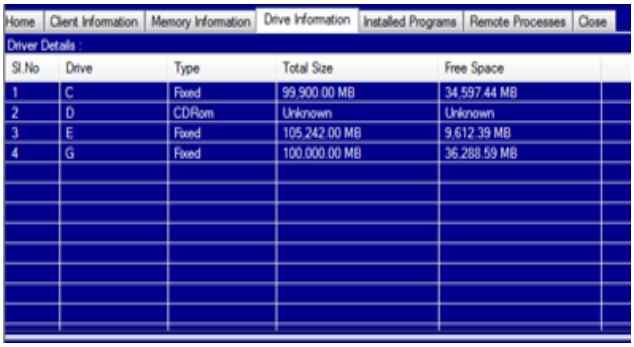

Fig : System monitoring details

D. Signature Based Packet Filtering

In signature based packet filtering user activity monitoring is performed. In this broadcasting is performed within the client-server systems. In this the message transferring is performed. Hash coding is performed for authentication in the wireless network.

Fig : Signature based packet filtering

This module also identifies the dropped packets in the network. All the information's will be monitored by the administrator.

E. Intrusion Detection

The intrusion detection will be performed in different ways. In Protocol based intrusion detection is performed on the basis of IP, TCP and UDP information's. It identifies the count of dropped packets. The packet dropper is a type of intrusion on the basis of network errors. Another type of intrusion detection is on the basis of system level i.e. File checking. In file checking it checks the c drive in the system. When any viruses are affecting the system, it normally changes the c directory information's. File change rate is measured and it detects the anomaly.

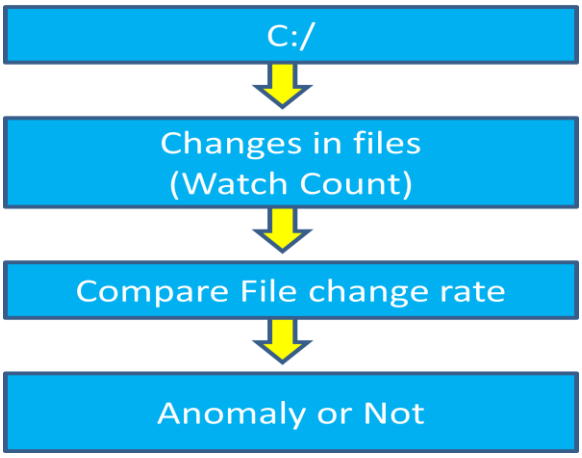

Fig. : File checking intrusion detection

F. Policy Monitoring

In policy monitoring, we can set different rule's based on the IP addresses and we can obtain the corresponding rule set information's.

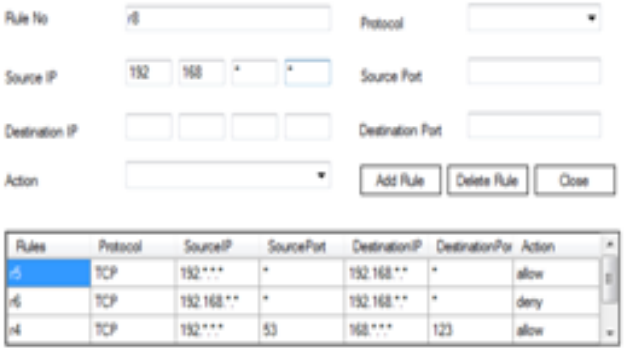

Fig. : Policy rule set and monitoring

G. Remote Monitoring

In remote monitoring, remote system monitoring is performed. For this it connects with the destination system with the help of the IP address of the system. It consists of the following operations

- ✓ Client details
- ✓ Memory Information's
- ✓ Drive Information's
- ✓ Installed Programs
- ✓ Remote Process
- ✓ Screen Capture
- ✓ Messaging
- ✓ CD Drive Open
- ✓ CD Drive Close
- ✓ Remote System Logoff
- ✓ Remote System Restart
- ✓ Remote System Shutdown

IV. EXPERIMENTAL RESULTS

The following figures show the results of the proposed methods.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

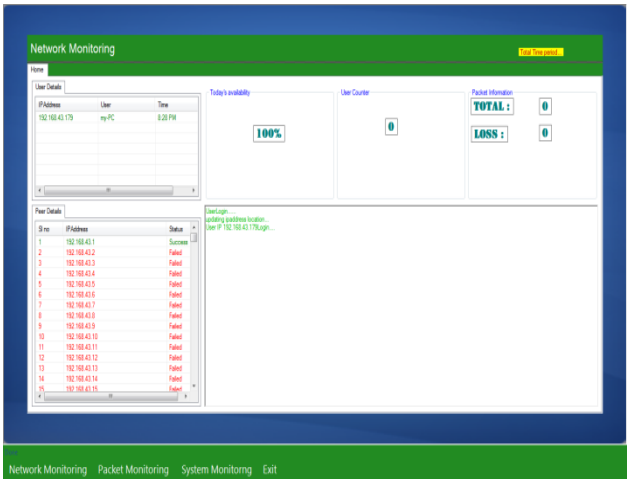

Fig.5. Network Monitoring

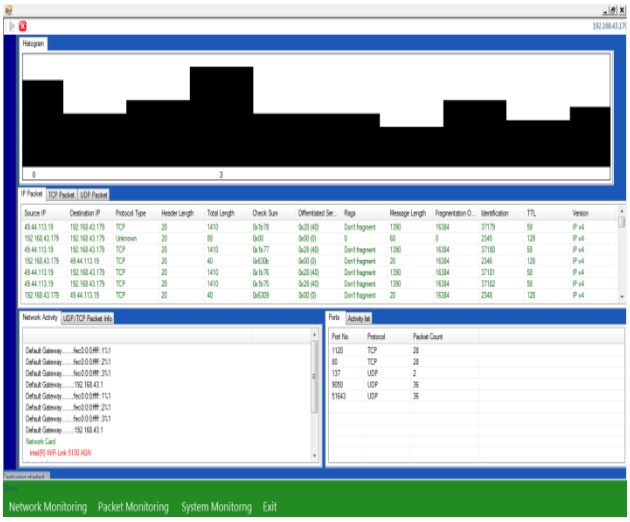

Fig.6. Packet classification

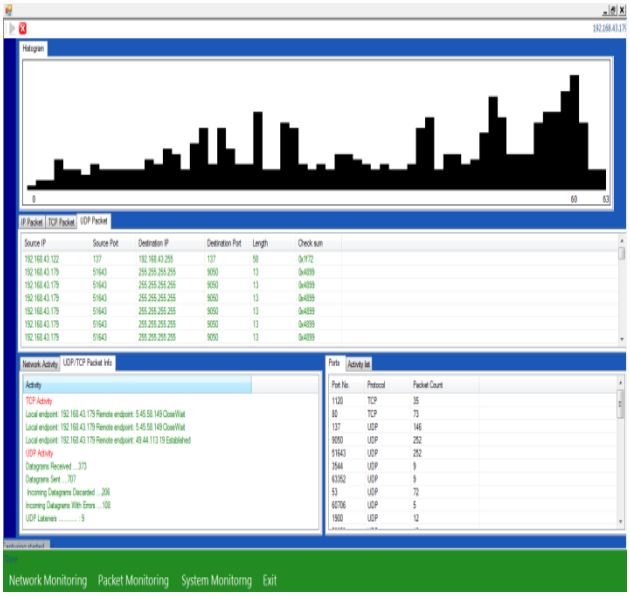

Fig.7. Packet classification and port details

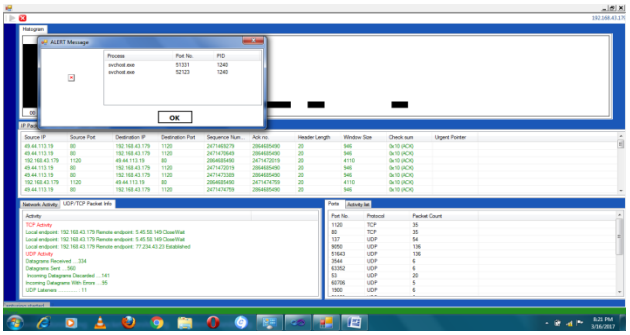

Fig.8. Port monitoring and intrusion detection

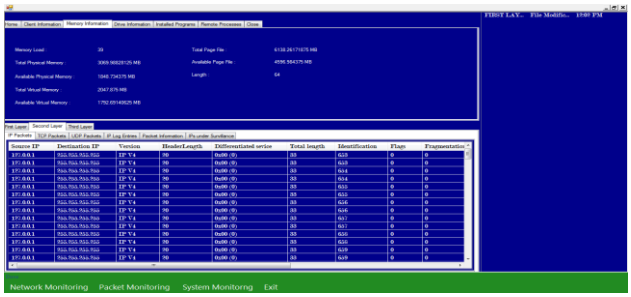

Fig.9. System monitoring and intrusion detection

V. PERFORMANCE ANALYSIS

The performance can be measured using the total number of packets, average packet rate and different type's protocols.

Table.1. Comparison table

	Existing Method	Proposed method
Trace Length (seconds)	3600	3600
Number of packets	874613	1074132
Avg packet rate(per second)	242.9	298.3
TCP Packets	303142	403433
UDP Packets	571471	670699
Anomaly Detection rate	147	320

The graph shows that the anomaly detection rate will be higher with compared to other existing methods. When number of packet increase the rate of Intrusion Detection rate will increase.

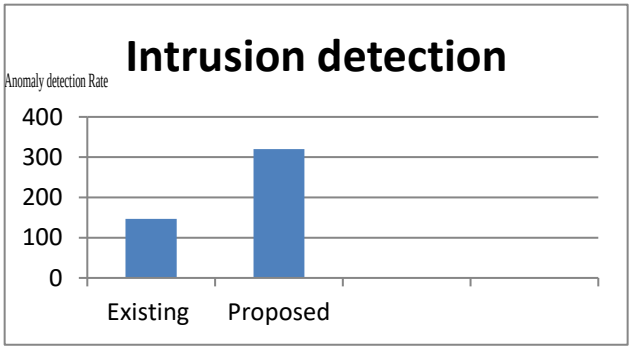

Fig.10. Anomaly detection rate

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

VI. CONCLUSION

In this proposed system, internal intrusion detection and protection system employs Data Mining and Forensic Techniques to identify the representative system call patterns for a user. The time that a habitual system call pattern appears in the user‘s log file is counted the most commonly used SC patterns are filtered out and then a user‘s profile is established. The intrusion detection is performed based on the IP address and different ports in the system. These methods are also used within the broadcasting in the network.

REFERENCES

1. Q. Wang, L. Vu, K. Nahrstedt, and H. Khurana, “MIS: Malicious nodes identification scheme in network-coding-based peer-to-peer streaming,” in Proc. IEEE INFOCOM, San Diego, CA, USA, 2010, pp. 1–5.
2. Z. A. Baig, “Pattern recognition for detecting distributed node exhaustion attacks in wireless sensor networks,” Comput. Commun., vol. 34, no. 3, pp. 468–484, Mar. 2011.
3. S. Kang and S. R. Kim, “A new logging-based IP traceback approach using data mining techniques,” J. Internet Serv. Inf. Security, vol. 3, no. 3/4, pp. 72–80, Nov. 2013.
4. K. A. Garcia, R. Monroy, L. A. Trejo, and C. Mex-Perera, “Analyzing log files for postmortem intrusion detection,” IEEE Trans. Syst., Man,Cybern., Part C: Appl. Rev., vol. 42, no. 6, pp. 1690–1704, Nov. 2012.
5. M. A. Qadeer, M. Zahid, A. Iqbal, and M. R. Siddiqui, “Network traffic analysis and intrusion detection using packet sniffer,” in Proc. Int. Conf.Commun. Softw. Netw., Singapore, 2010, pp. 313–317.
6. S. O’Shaughnessy and G. Gray, “Development and evaluation of a data set generator tool for generating synthetic log files containing computer attack signatures,” Int. J. Ambient Comput. Intell., vol. 3, no. 2, pp. 64–76, Apr. 2011.
7. S. X. Wu and W. Banzhaf, “The use of computational intelligence in intrusion detection systems: A review,” Appl. Soft Comput., vol. 10, no. 1, pp. 1–35, Jan. 2010.
8. Z. B. Hu, J. Su, and V. P. Shirochin “An intelligent lightweight intrusion detection system with forensics technique,” in Proc. IEEE Workshop Intell. Data Acquisition Adv. Comput. Syst.: Technol. Appl., Dortmund, Germany, 2007, pp. 647–651.
9. T. Giffin, S. Jha, and B. P. Miller, “Automated discovery of mimicry attacks,” Recent Adv. Intrusion Detection, vol. 4219, pp. 41–60, Sep. 2006.
10. U. Fiore, F. Palmieri, A. Castiglione, and A. D. Santis, “Network anomaly detection with the restricted Boltzmann machine,” Neurocomputing, vol. 122, pp. 13–23, Dec. 2013.
11. M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, “Data-streambased intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study,” IEEE Syst. J., vol. 9, no. 1, pp. 1–14,Jan. 2014.
12. S. Khan, N. Mast, and J. Loo, “Denial of service attacks and mitigation techniques in IEEE 802.11 Wireless mesh networks,” Information,vol.12,pp.1–8,2009.
13. S. Khan and J. Loo, “Cross layer secure and resource-aware ondemand routing protocol for hybrid wireless mesh networks,” Wireless Personal Communications,vol.62,no.1,pp.201–214,2010.
14. S. Khan, N. Mast, K.-K. Loo, and A. Silahuddin, “Passive security threats and consequences in IEEE 802.11 wireless mesh networks,” International Journal of Digital Content Technology and Its Applications,vol.2,no.3,pp.4–8,2008.
15. S. E. Robertson, S. Walker, M. M. Beaulieu, M. Gatford, and A. Payne, “Okapi at TREC-4,” in Proc. 4th text Retrieval Conf., 1996, pp. 73–96.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)