Advanced Intrusion Detection System with Prevention Capabilities

A.B. Pawar
Comp Dept., S.R.E.S COE,
Kopargaon, SPPU, Pune (M.S)
JJTU, Rajasthan, India

D.N. Kyatanavar, Ph.D.
S.R.E.S COE, Kopargaon,
SPPU, Pune (M.S)
JJTU, Rajasthan, India

M.A. Jawale
IT Dept., S.R.E.S COE,
Kopargaon, SPPU, Pune (M.S)
JJTU, Rajasthan, India

ABSTRACT

Today, with the advent of internet, everyone can do information exchange and resource sharing. Even business organization and government agencies are not behind in this move to reach users for their decision making and for business strategies. But at the same time, with ease of use and availability of various software tools, breaching and penetrating into other's network and confidential credential can be done by any individual with little knowledge expertise and hence the internet attacks are rise and are main concerns for all internet users and business organizations for internal as well as external intruders. Even, existing solutions and commercial Intrusion Detection Systems (IDSs) are developed with limited and specific intrusion attack detection capabilities without any prevention capabilities to secure vital resources of the information infrastructure. So, this paper explores the details about the implementation and experimental analysis of Advanced Intrusion Detection System (AIDS) with its prevention capabilities to provide detection of known as well as unknown intrusions in the computer system and also automatic alerts are given to the network administrator for applying prevention capabilities. Further, this system is intended to generate new intrusion signatures from unknown intrusions and store them back into signature database to speed up detection capabilities of this AIDS in next iterative computation. Data mining approach is used to handle the large amount of data captured in the Internet to improve its execution time and to give fast response to the network administrator for prevention of data resource with minimal user intervention. In experimental analysis, this proposed system gives improved and effective intrusion detection rate up to 91% in comparison with existing research IDSs Snort and PHAD with minimization in false positive rate up to 11%.

Keywords

Attack, Data Mining, Intrusion Detection, Intrusion Prevention

1. INTRODUCTION

The current scenario in information security shows that the novel attacks are evolving at very fast rate on the internet. With increased level of automation in attack tools, the expertise requires to breach the security is minimizing, the complexity also increases proportionally, making the tasks of security professional very challenging.

Every business is depending on network to meet their business needs. Enterprises and government agencies have developed their own information networks, including technologies like distributed data storage systems, encryption techniques, Voice over IP (VoIP), remote and wireless access, and Web services. These networks have become preliminaries as business partner's access services via extranets, Customer Relationship Management (CRM) processes and employees use company systems through Virtual Private Networks (VPN) [14].

For attackers, these paths make networks more vulnerable than ever before. With relative little expertise, hackers have significantly impacted the networks of leading brands or government agencies. Cyber-crime is also no longer right of lone hackers or random attackers. Today, disgruntled employees, unethical corporations, even terrorist organizations all look to the internet as a portal to gather sensitive data and initiate economic, social and political disruption. With networks more vulnerable and hackers equipped to cause havoc, it's no surprise that network attacks are on the rise [1], [14].

A joint report published by CSI and FBI in year 2010 indicates that hacking and malware are the most popular attack methods. Malware was a factor in about half of the year 2010 caseload and was responsible for loss of 80 percent of data. The most common kinds of malware found in the caseload were like involving sending data to an external entity, opening backdoors, and key logger functionalities. At the same time, stolen passwords and credentials are found out of control. Ineffective, weak or stolen credentials continue to cause on enterprise security [2].

In order to secure enterprise and government networks against complete spectrum of threats and vulnerabilities, integration of existing methodologies of intrusion detection must be integrated i.e. Signature Detection and Anomaly Detection. Also, Intrusion Detection System (IDS) must be able to do more than detecting attacks and it should give accurate detection to prevent attacks from reaching and damaging critical network resources and data [3]. From this, it's clear that enterprises and government agencies need to step up and deliver innovative solutions that effectively protect their networks from malicious attacks and misuse [14]. The proposed research work is intended to research and develop such innovative solution to provide computer security with the advantages of data mining techniques and sentiment analysis in the field of intrusion detection system [15].

2. MOTIVATION AND RELATED WORK

The motivation to bring out this research work was with the advent of internet technology and massive growth of attacks. Today, not a single intrusion detection system can meet the requirement of high intrusion detection rate with minimum false positive rate. Also, during the literature review, it is observed that significant separate research is done on broad categories of intrusion detection models namely, misuse or signature intrusion detection and anomaly intrusion detection. Also, most of available IDSs just detect certain types of attacks, but no immediate prevention strategies are given for the network administrator, so that administrator can perform immediate actions to prevent the resources. It leads towards use of Intrusion Prevention System (IPS) security measures in the implementation of IDS to provide another layer of protection with great computational power of today's computer technology and tools. As stated earlier, data analysis for intrusion detection is the major challenge for all IDSs because of ample and large availability of this network data with the growth of internet and ways used by the intruders with modern technology to generate attacks and penetrate into the information networks of the organization. At the same time, the Data Mining (DM) technology is used by various research fields to handle large amount of data [6]. With the use of Knowledge Discovery in Database (KDD), it is possible to enhance the decision making process effectively. This motivated towards the use of DM even into the field of intrusion detection to speed-up the data processing to identify intrusion in reduced analysis. The following works have also given an opportunity to select proposed research work. It is known an intrusion takes place when an unauthorized access of a host computer system is attempted. Whereas anomaly is observed at the network connection level and when the observed behavior diverges from expected behavior, an anomaly is raised. Unfortunately, they are prone to false positives which can be triggered as novel, but by non-malicious traffic. Both known and unknown intrusion or attack types may compromise valuable hosts, disclose sensitive data, deny services to legitimate users, and pull down network based computing resources. Malicious intrusions on these systems may destroy valuable hosts, network, and storage resources. Network anomalies cause even more damages. Internet anomalies found in routers, gateways, and distributed hosts may hinder the acceptance of grids, clusters, and public-resource networks [4].

Existing IDSs are built based on either signature-based or anomaly-based detection models. Signature matching is based on a misuse model, whereas anomaly detection is based on a normal use model. The design philosophies of these two models are quite different, and they were rarely mixed up in existing IDS. Anomaly detection discovers temporal characteristics of network traffic based on DM. This system can detect unknown attacks and handles multi-connection attacks well as studied by [5]. However, anomaly detection may result in higher false alarms. Both signature-based and anomaly-based IDSs are sensitive to the attack characteristics, system training history, services provided, and underlying network conditions.

Other attempts to solve the intrusion detection and response problem can be found in [7], [8]. Whereas, [9] proposed intrusion detection using sequential pattern mining in the field of information security. It first introduces several common sequential pattern mining algorithms, and then expands its current development with comparisons about the merits and shortcomings with the current mainstream technologies. At the same time, the comprehensive analysis for intrusion behaviors from multiple angles by introducing other data mining techniques with the sequential pattern and implementing multi-level mining is inspected. It is concluded that providing more valuable intrusion information to security administrators and reducing false alarm rate will be also the goal of future research. Further, [10] stated that in spite of the significant role of databases in information systems, not enough attention has been paid to intrusion detection in database systems [16].

With above survey of various kinds of IDSs and their implementation strategies, it become necessary to understand that network-based computer systems play increasingly vital roles in modern society and become the targets of intruders. According to the reports studied by [13], due to network security the United States caused economic losses amounting to tens of billions of dollars every year. Much network management center is connected to the Internet and it has been inside and outside hackers or invasion, there have been some vandalism and theft of information network of criminals, it has been on the internal computer system and information network pose the great threat. Regular contact with internal staff within the information and any information security are not careful, and then it can have both the threat. Therefore, the information network must have adequate security measure to ensure that the network of information is confidential, integrated and secured. Therefore, there is a need to find the best ways possible to protect computer network systems [16], [17]. So, from this review, it gives the sense of development of single IDS to deal with these threats, so the proposed work intends to solve these problems with building integrated intrusion detection system which will deal with known and unknown kinds of intrusions with enhanced effectiveness of it [16].

During this reviews, it is observed that, intrusion prevention has been used to protect computer systems as a first line of defense. Intrusion prevention is not sufficient because as systems become ever more complex, there are always exploitable weaknesses in the systems due to design and programming errors, or various “socially engineered” penetration techniques [14].

For the same, one needs to set another wall of protection. So, to set prevention from the detected intrusions or attacks, the proposed research work will provide the system which will automatically and systematically build adaptable and extensible intrusion detection system based on data mining concepts and will provide in-built prevention policies in the detection system so, that it will reduce network administrator's system re-configuration efforts. At the same time, effectiveness of proposed system will be evaluated based on false positive or false negative reviews detection [12].

3. ECONOMIC CONSEQUENCES OF CYBER ATTACKS

3.1 Attack Aspects

The economic consequences of cyber-attacks have many aspects as stated in [17]. They are different at societal level, organizational level, and individual level. United States spends a major chunk of resources like cash flow and intelligence on developing the weapons, and training and maintaining large army and military bases. On the contrary, these resources have to be mobilized and shared with Information Technology (IT) risks and cyber-attacks. This means that along with the security of this country's nationals, now humongous amount of resources have to be invested in cyber security as it is a threat of an equivalent magnitude of risk.

3.2 Cyber Attack Trends

Before drilling down into the data for the past year, it's worth to have a look to the trend of the last three years (with the caveat that data for 2011 are incomplete as it was consolidated into a form comparable with year 2012 and year 2013 only starting from September). Apparently in year 2012 and 2013 have a very different shape: Year 2012 shows a constant trend (with a high activity between May and June), while, after an initial peak, the line for 2013 experiences a progressive decrease, reaching a stable state. This is probably due to the minor influence of attacks motivated by hacking activities throughout the year as shown in Fig. 1. A closer look to 2013 allows understanding the influence of the motivations throughout the different months as shown in Fig. 2. It is observed that the initial part of the year is characterized by hacktivism. Cyber Crime is quite constant and ends up dominating the second half of respective year. This trend does not mean a decrease of hacktivism, but rather a different connotation throughout the year: the global-scale operations executed by anonymous have progressively been replaced by local phenomena (for instance the cyber-attacks in India and Pakistan). Also the first months of the year are influenced by the DDos attacks of Cyber Fighters against US Banks [13], [16]. It is studied that DDoS leads the chart of known Attack Techniques (23%) ahead of SQLi (19%) and Defacements (14%) given in the Fig.3.

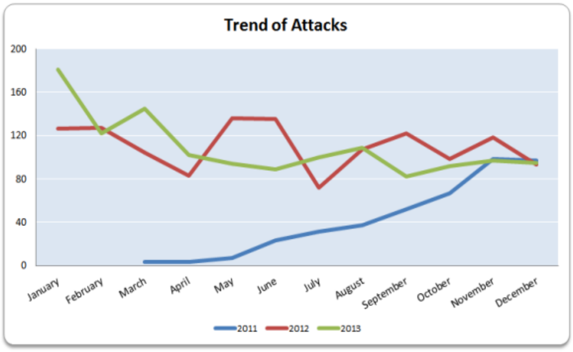


Fig. 1: Trend of Attacks (Data for Duration Sept 2011 to Dec2013) [13]

These attack techniques are mostly used to attack all the business enterprises, government agencies, commercial industries and individual system too. The Fig. 4 shows the top targets of the cyber attackers where governments and industries have been the most preferred targets for Cyber Attackers with similar values (respectively 23% and 22%) and targets belonging to finance rank at number three (14%), immediately ahead of News (6%) and Education (5%) [13].

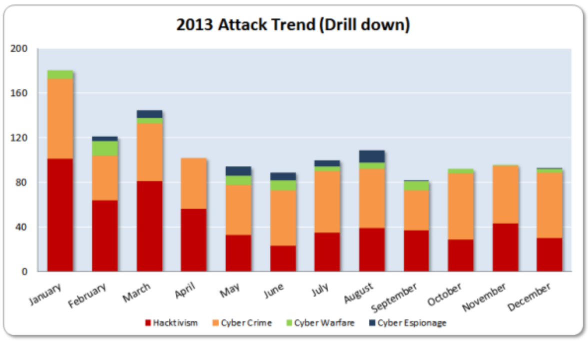


Fig. 2: Attack Trend with the Drill-down of M0otivations for Year 2013 [13]

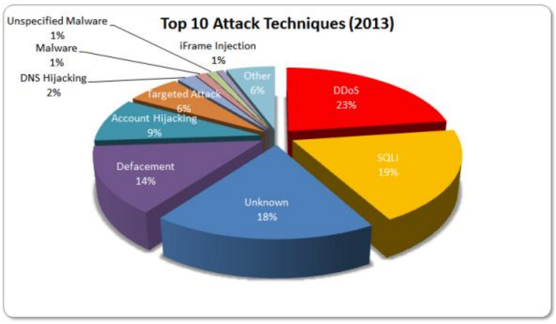


Fig. 3: Top Attack Techniques Used by Intruders during Year 2013 [13]

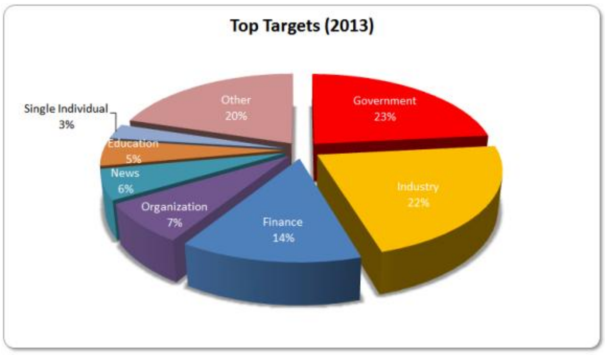


Fig. 4: Top Targets of the Cyber Attackers [13]

4. SYSTEM MODEL

4.1 Objective

Main objective of the proposed research work was set as: "How to automatically and systematically build adaptable and extensible advanced intrusion detection system using Data Mining techniques and how to provide in-built prevention policies in the detection system so that it will reduce network administrator's system re-configuration efforts and application of sentiment analysis to enhance its performance".

4.2 Proposed Model

For proposed research work, the system architecture is modeled and designed with the two major components: Intrusion Detection Phase and Prevention Phase. Intrusion Detection Phase includes input data / testing dataset, training attack dataset, Existing Attack Database, Signature-based intrusion detection process, Anomaly-based intrusion detection process; Detected attacks types either 'known' or 'unknown', New attack signature generation process, Attack Detection Correlation, Attack database updating process and generation of alerts process. Prevention Phase consists of list of prevention policies provided by the developed system, prevention policy selection strategies for system administrator and graphical user interface component.

It is observed that not a single architecture can be made as generic since each one is having their own merits and demerits. So, the following section gives the details about the system architecture designed and used for implementation for the research work. Initially, the theoretical formulation is done which gives an overview about the steps carried out while designing the architecture of research work. Subsequently, the mathematical model for this architectural design is introduced in next section.

4.3 Mathematical Model

In set theory representation, Let *S* represents the system, *I* represent the input set of the system, *P* is set of processes of the system, *R* is the rule set applied on the system during its processing and *O* is the output set of expected outcomes of the system. Therefore, the system representation in set form is

$$S \in \{ I, P, R, O\} \text{.. (1)}$$

Input

The input is considered as online as well as offline network traffic captured data. The online network traffic data is captured through Smart Sniff tool which was further preprocessed and then utilized for intrusion detection. The offline data is obtained from KDD cup 1999 intrusion attack dataset from the MIT/LL which is preprocessed further for the system evaluation. Therefore, set *I* represented as:

$$\mathrm{I} = \{\mathrm{I1}, \mathrm{I2}\} \text{... (2)}$$

Where, *I1* is online captured input network traffic data; *I2* is KDD cup 1999 standard intrusion dataset.

Processes

These are functions of the system to represent the complete flow of execution from initial process to final step of execution. For proposed system, the main processes are described below:*P1* is Data preprocessing process used to make data compatible for further processing;*P2* is Known attack detection process based on misuse detection;*P3* is Unknown attack detection process based on anomaly detection;*P4* is Data correlation process used to perform known and unknown attack correlation;*P5* is Attack alert generation process for known as well as unknown attacks; *P6* is Prevention policies application process to provide layer of protection and *P7* is the management console process used to administer the whole system. Therefore, set *P* is represented as:

$$\mathrm{P} = \{\mathrm{P1}, \mathrm{P2}, \mathrm{P3}, \mathrm{P4}, \mathrm{P5}, \mathrm{P6}, \mathrm{P7}\} \text{......................... (3)}$$

Output

Outputs of the developed system are the final outcomes of the system i.e. type of known and unknown attacks detection i.e. detection of classes of attacks which are mentioned earlier in this section, applied prevention policies. Generally, this set represents only the final expected outcomes and does not provide the intermediate outcome details. The output set *O* can be written as

$$\mathrm{O} = \{\mathrm{O1}, \mathrm{O2}, \mathrm{O3}, \mathrm{O4}, \mathrm{O5}, \mathrm{O6}, \mathrm{O7}\} \text{......................... (4)}$$

Where, *O1* represents DOS attack and its subtype’s detection; *O2* represents R2L attack and its subtype’s detection; *O3* represents U2R attack and its subtype’s detection; *O4* represents Probe attack and its subtype’s detection; *O5* represents unknown or anomalous attack detection; *O6* represents generated alerts once known or unknown attack is detected; *O7* represents applied prevention policies.

The mapping of developed system input, process and output set can be represented with the help of Venn diagram as shown in Fig. 5. For the developed system, to bring out completeness in the model presentation; the process state diagram is shown in Fig.6.

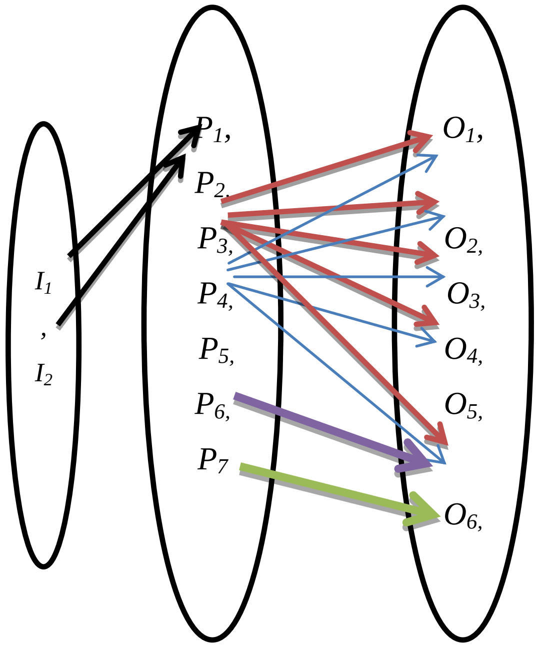


Fig. 5: Venn diagram Representation for *S = {I, P, O}*

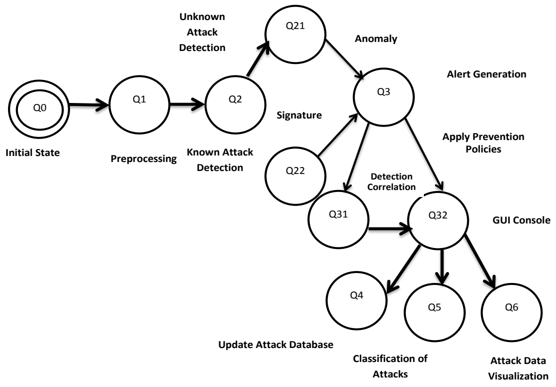


Fig. 6: Process State Diagram of Proposed System to Represent Execution Flow of the System

4.4 Design Architecture

The proposed research design architecture can be divided into three phases of development namely, data collection and preprocessing; Known and unknown attack detection; and Prevention as shown in Fig. 7 and Fig. 8.

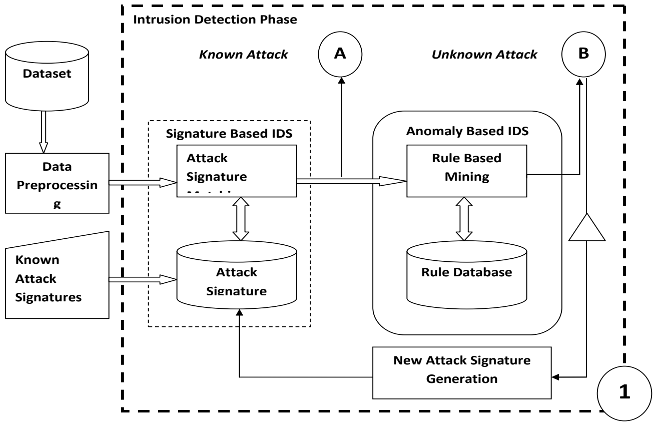


Fig. 7: Intrusion Detection Phase of System Architecture

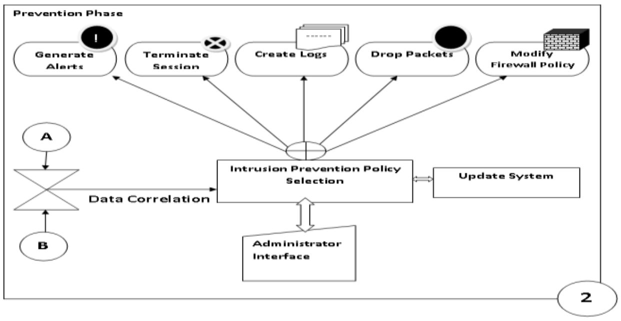


Fig. 8: Prevention Strategies Used in Proposed System

Data Collection and Preprocessing

The data obtained from the standard dataset is preprocessed to get the data in the format which is acceptable for the detection. The outcome of this step will give formatted dataset which will be the input for detection phase. KDD Cup 1999 IDS evaluation data set is used as input for proposed system which is Massachusetts Institute of Technology/Lincoln Laboratory traffic files collection and is available online.

The original KDD Cup 1999 dataset contains 41 features. Each row contributes for these 41 features of the network traffic packet. Additionally, these feature attributes are categorized into further sub-classification based on the features extracted from the parts of packet i.e. Packet Header, Packet Payload Contents and Packet Traffic features .Once, this KDD Cup 1999 dataset with feature definition file is obtained, it is possible to process complete file for the further modules of the proposed system as shown in Fig. 9. It gives more clearly the details about each feature attribute of packet data.

Duration	Protocol_type	Service	Flag	Src_bytes
0	Tcp	Smtp	SF	829
0	Udp	Smtp	SF	105
0	Udp	Private	SF	105
0	Udp	Private	SF	105

Fig. 9: Preprocessed File Contents

Known Intrusion Detection

Here, the main aim is to filter out the known attack traffic through by Snort like signature detection based on signature matching with the existing or provided signature database prepared by human experts at initial stage. The processing of this signature based intrusion detection consists of signature matching algorithm, signature database retrieval and response for the detected attack and forwarding of remaining traffic to next step of the system. Finally, the outcome is detected known attack.

Unknown Intrusion Detection and New Intrusion Signature Generation

The remaining traffic containing unknown or burst attacks is forwarded to the profile pattern mining engine to generate frequent pattern rules with different levels of support threshold. The provision is to provide threshold level as user defined to detect unknown attacks. This leveling provision allows the detection of few rare patterns, which will be later declared as anomalies or unknown intrusions in the given input data. The frequent patterns are then compared with precompiled frequent patterns from normal traffic database. The profile patterns that do not match normal profiles or match them with unusually high frequency are labeled as anomalous and unknown intrusions.

The detected unknown intrusions are used to generate new attack signatures based on captured anomalous behavior using a weighted frequent item set mining scheme algorithm. These newly generated signatures are then added to the signature database for future detection of similar attacks. Generally, this step detects unknown, burst, or multi-connection attacks from the given network traffic data.

The detection architecture enables multi modes of operation that allow the system to capture malicious traffic provides thorough attack analysis methodologies and implements a complete set of intelligent Signature Detection, Anomaly Detection. The Detection Correlation layer connect the system's Signature, Anomaly, and at some extent, Denial of service detection functionality and this interdependence and cross checking of suspicious traffic gives more accurate attack detection than using individual intrusion detection system.

Lastly, the output of this step is detected unknown attack detection along with input to new signature generation module. This is the last and most important phase of this system. In this phase, attacks detected by detection phase as shown in Fig. 8 are given as input (Shown with labels 'A' and 'B' in Fig. 8) and after their detection correlation, the suitable prevention policy can be applied to avoid the intrusion on the system.

Prevention Phase

Here, the attempt is, a single system should provide comprehensive protection by monitoring public, private, and segments of the network with firewall can offer correlation among these segments to give more accurate picture of network attacks that were either blocked by the firewall or made them possible to enter into the private network. So, the proposed system gives the prevention policy responses which can be used by network administrator during prevention phase of the system as shown in Fig. 8. Depending on detected attack and its nature of severity, the prevention phase of the system model enables the system to do: Drop Packets; Terminate Session; Modify Firewall Policies; Generate Alerts; Log Packets.

5. EXPERIMENTAL ANALYSIS

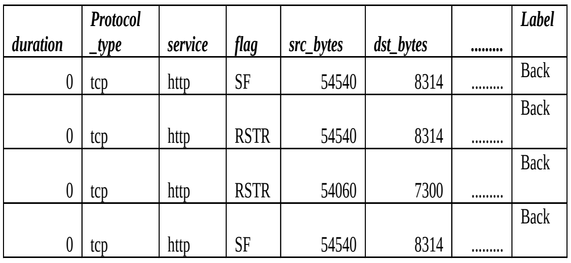
5.1 Proposed System Evaluation

The usefulness and the reason behind wide support for KDD Cup dataset in the IDS evaluation can be found in [5]. The 10% KDD Cup Dataset is considered for the proposed system experimental analysis. The proposed system is implemented with the help of following run time architecture. It includes Smart sniff, Java programming and Database technology. The developed system is tested on the following minimum hardware configuration in its offline mode: RAM: 256MB; Processor: Intel Core series @ 2.27 GHz; Disk Space: 2 GB for standalone machine connected to internet.

To Identify Known Attacks

Once the data is fed to next intrusion detection phase, the known attacks are identified based on matching of provided trained data of each attack definition with input data chosen for the testing. The trained data contains total 22 attack definition files to cover the attacks and their sub attack types under the major classes of intrusion attacks. The following Fig. 10 shows the one of the definition in training file for the *back* sub attack(Since, the *back* attack is mainly dependent on *src_bytes* and *dst_bytes* feature variations, so only those details are highlighted here, instead of all 41 feature values for each captured packet data.)

Once the signature based IDS module is executed, the trained file is compared with input file and gives the known attack detection and provision to test the remaining input data for the unknown attack detection through anomaly detection module.



duration	Protocol _type	service	flag	src_bytes	dst_bytes	....	Label
0	tcp	http	SF	54540	8314	....	Back
0	tcp	http	RSTR	54540	8314	....	Back
0	tcp	http	RSTR	54060	7300	....	Back
0	tcp	http	SF	54540	8314	....	Back

Fig. 10: Training files contents for *Back* attack

The calculated percentage of each attack for the first selected 1000 records of KDD Cup input dataset is shown in the following Fig. 11.

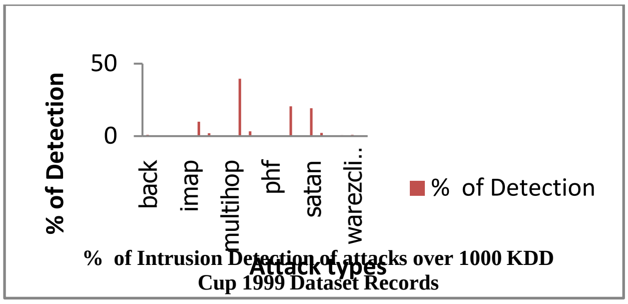


Fig. 11: Percentage of Detection of Attacks in Input Dataset of 1000 Records

To Identify Unknown Attacks as anomaly

During this phase, the anomaly is detected over the intrusive traffic, since it is important to identify the new unknown attacks as well as potential variation in the certain type of intrusion attack also. The existing research work considers the anomalous intrusion identification over the normal profile deviation. On the basis of percentage of attack detection variation in comparison with expected input threshold value, the final anomalous traffic records are identified with attack type too. So, following Table 1 shows the relationship of major attacks and their sub attack types in input dataset.

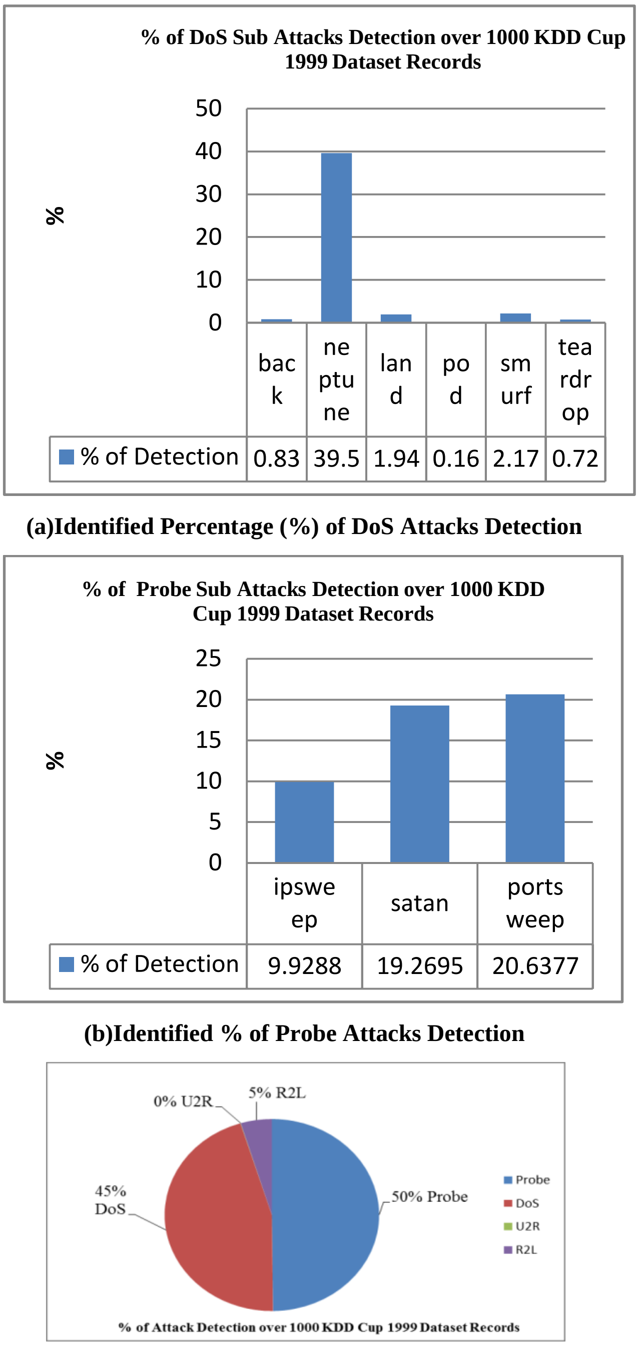
Table 1. Major Attack Classification

Attack	Sub Attack Types
Probe	Ipsweep, portsweep, satan,nmap
DoS	Back,land,Neptune,pod,smurf,teardrop
U2R	Laodmodule,perl,buffer_overflow,rootkit
R2L	ftp_write,imap,phf,guess_passwd, multihop,spy,warezclient,warezmaster

With the same scenario, the attack correlation is done to get exact and reduced attack dataset for the further processing. For above case, the percentage of anomalous traffic for Probe and DoS records is shown in the following Fig. 12 (a) ,(b) and (c) (The % of detection for U2R and R2R is found very less based on their sub attack types identified here, so they are not visualized here).

Apply Prevention

Based on the alert generated from the developed system can drop the packets containing the above two attack data. Even, can reset the firewall configuration settings to block these attack packets potentially before entering into the network as well as can reset or terminate session immediately for this current traffic. Additionally, network administrator can send this data for the further analysis with the log file contents to generate new and powerful signature to detect faster for similar attacks in the future.



(a)Identified Percentage (%) of DoS Attacks Detection

(b)Identified % of Probe Attacks Detection

(c) Summary of DoS, Probe, U2R, and R2L attacks Detection

Fig. 12: Attack Detection over 1000 KDD Dataset Records Test File by Proposed System

5.2 Results based on Precision, Recall and F-Score Measures

With intrusion detection rate, Precision, Recall and F-score are efficient and accurate performance measures used to evaluate proposed system. Each of these parameters is defined in the following manner for the proposed system evaluation. Consider, $D= \{d_1, d_2, d_{3,} \ldots ,d_n\}$ is set of detected attacks by IDSs and $I=\{i_1, i_2,i_3,\ldots,i_n \}$ is set of expected detection of attacks from IDSs , then values of Precision(p), Recall(R) can be calculated as Precision, $P= Tp/(Tp+Fp)$. Fig. 13 shows the precision values calculated for IDS Snort, PHAD and proposed advanced IDS where, it is found that IDS Snort had given precision 77% by precision equation, where value of T_p was 7 as actual detected attacks and value of F_p was 2 as number of false attacks detected, IDS PHAD had given 85 %, where value of T_p was 6 as actual detected attacks and value of F_p was 1. Similarly, proposed advanced IDS was shown precision value 91%, where value of T_p was 21 as actual detected attacks and value of F_p was 02 from the total number of attacks to be detected by simulated system and which was found better than other two IDSs considered for the performance evaluation. The improvement in precision value of proposed advanced IDS is possible because of immediate signature generation from the undetected attack traffic and its updating into existing attack signature database. However, in IDS Snort and PHAD, updating signature database is possible with manual intervention and need extensive knowledge to update them.

Similarly, Fig. 14 shows the recall values calculated for IDS Snort, PHAD and proposed advanced IDS where, it is found that IDS Snort had given recall 53% by recall equation, $R=Tp/(Tp+F_N)$ where, value of T_p for IDS PHAD was 7 as actual detected attacks and value of F_N was 6 as number of undetected attacks, IDS PHAD had given 60 %, where value of T_p was 6 as actual detected attacks and value of F_p was 10. Similarly, proposed advanced IDS was shown the recall value 90%, where value of T_p was 18 as actual detected attacks and value of F_p was 02 from the total number of attacks to be detected (23 attack types in KDD Cup 1999 Data set training file) by simulated system and which was found greater and better than other two IDSs considered for the performance evaluation.

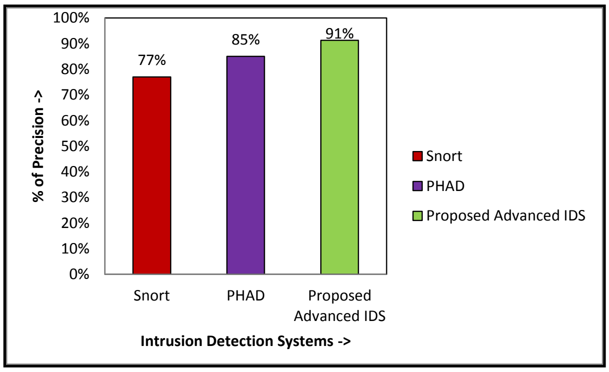


Fig. 13: Precision value (in %) for IDS Snort, PHAD and Proposed Advanced IDS

Finally, F-Score is used to represent the balance between precision and recall. It is used to measure the accuracy of a test carried out.

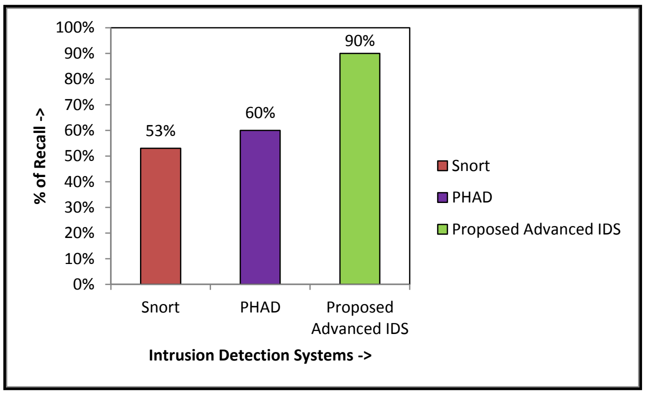


Fig. 14: Recall value (in %) for IDS Snort, PHAD and Proposed Advanced IDS

The F-score is the harmonic mean value of recall and precision, it is calculated as: $F\text{-}Score= (2*P*R)/(P+R)$.

Precision, Recall and F-Score are the standard measures are derived based on probability theory and allows one to take into account the intrinsic variations of performance estimation of any system. Hence, for the performance evaluation of proposed advanced IDS with other IDSs Snort and PHAD, F-Score is considered as another parameter to evaluate accuracy of the IDSs taken into consideration. Fig. 15 depicts the variation of F-Score measure values of evaluated IDSs.

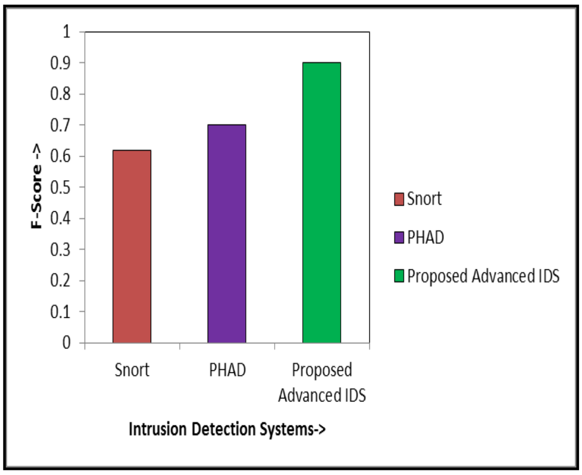


Fig. 15: F-Score values for IDS Snort, PHAD and Proposed Advanced IDS

6. CONCLUSION AND FUTURE SCOPE

The proposed system is evaluated with the empirical values and modeling help to understand proposed system and can be used to further design and research in IDSs. This knowledge can be used by security analyst to understand and respond more effectively and instantly to response future intrusions to avoid further damages and vital resource information. It is also observed that as the intrusion detection performance improves with time, the slope of F-score is positive, which causes the effect of attacks to disappear. However, it is not possible to get this type of a growth rate with single IDS (like IDS Snort, IDS PHAD). Mostly, effect of attacks is not felt and observed during the use of the information systems, it is important for the IDS performance to rise and reach towards F-score value of 1. During the review of literature, it is found that no single IDS can achieve this, so it is quite essential to integrate multiple IDSs, with merging and benefiting from the advantages of each one of them as done in the proposed system. The proposed system modeling is implemented in real environment of local network with multiple IDSs with protection strategies, looking at the system as a whole, instead of separate deployment of IDS and IPS components.

7. REFERENCES

[1] A.B.Pawar, Dr.D.N.kyatanavar, M.A.Jawale 2013, Development of Advanced Intrusion Detection System: Review In Proc. of ICRTET 2013, Feb 2013 & In IJCA.

[2] A.B.Pawar, Dr.D.N.kyatanavar, M.A.Jawale 2013, Design of Advanced Intrusion Detection System In Proc. Series 01 of AIM 2013& CPE 2013, Elsevier.

[3] A.B.Pawar, Dr.D.N.kyatanavar, M.A.Jawale 2014 Implementation of Advanced Intrusion Detection System to improve Detection and Prevention Capabilities based on Data Mining In Proc. of ICRTET'14,Elsevier.

[4] K. Hwang, Y. Kwok, S. Song, M. Cai, Y. Chen, and Y. Chen2006 DHT-Based Security Infrastructure for Trusted Internet and Grid Computing In Int'l J. Critical Infrastructures, vol. 2, no. 4, pp. 412- 433.

[5] M.V. Mahoney and P.K. Chan 2003 An Analysis of the 1999 DARPA/ Lincoln Lab Evaluation Data for Network Anomaly Detection In Proc. Int'l Symp. Recent Advances in Intrusion Detection pp. 220-237.

[6] Muamer N. Mohammad, Norrozila Sulaiman, Osama Abdulkarim Muhsin 2011 A Novel Intrusion Detection System by using Intelligent Data Mining in Weka Environment In Science Direct, Procedia Computer Science, pp. 1237–1242.

[7] Adeeb Alhomoud, Rashid Munir,Jules Pagna Disso,Irfan Awan,A. Al-Dhelaan 2011 Performance Evaluation Study of Intrusion Detection Systems In Procedia Computer Science pp.173–180.

[8] S.Sathya Bama, et al. 2011 Network Intrusion Detection using Clustering: A Data Mining Approach, In International Journal of Computer Applications (0975 – 8887) Volume 30– No.4, pp.14-17

[9] Yuanqin Wu, Liang Shi, Beizhan Wang, Panhong Wang, Yangbin Liu 2011 Research on Intrusion Detection Based on Sequential Pattern Mining Algorithms In Science Direct Energy Procedia , pp. 505 – 511.

[10] Rezk, H. Ali, M. El-Mikkawy and S. Barakat 2011 Minimize the false positive rate in a database intrusion detection system In International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 5, pp.29-38.

[11] Xiangyang Zheng, Qian He 2011 Research on Distributed Intrusion Detection System Model, In Energy Procedia,pp.1480-1485.

[12] Bing Liu 2010 Sentiment Analysis: A Multi-Faceted Problem, In IEEE Intelligent Systems, pp.1-5.

[13] http://hackmageddon.com/2014/01/19/2013-cyber-attacks-statistics-summary,2013.

[14] https://www.mcafee.com/japan/products/pdf/IntruVert-NextGenerationIDSWhitePaper_en.pdf

[15] http://shodh.inflibnet.ac.in:8080/jspui/bitstream/123456789/1000/1/1.introduction.doc.

[16] http://shodh.inflibnet.ac.in/bitstream/123456789/1000/2/2.literature%20review.doc.

[17] http://qenru.blogspot.in/p/blog-page.html