Development of Advanced Intrusion Detection System: Review

A.B.Pawar
Research Scholar, JJTU
SRES, College of Engineering, Kopargaon

D.N.Kyatanavar, Ph.D
Research Guide & Principal
SRES, College of Engineering, Kopargaon

M.A.Jawale
Research Scholar, JJTU
SRES, College of Engineering, Kopargaon

ABSTRACT
In this paper, we have been explored the brief review about the intrusion detection system. This review emphasizes about 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. Intrusion detection and prevention is really widely researched filed and still there is a scope for its advancements. This review gives the requirement of advancement in current intrusion detection systems based on data mining technique in its introduction section. In related work, it focuses on the growth and research contributions made in the field of security with intrusion detection and prevention. In the section of objectives, it concludes the current research requirements and use of possible techniques to step forward in intrusion detection and prevention. Finally, the possible applications of the proposed research work are highlighted to make its sense in society and conclusion provides the actual research direction based on the review.

General Terms
Anomaly Detection, Data Mining, Intrusion Detection, Signature Detection, Intrusion Prevention.

Keywords
Anomaly, Attack, Intrusion, Misuse, Signature, Prevention Policy.

1. INTRODUCTION
Today, every business is depending on network. Mostly, because of business needs, enterprises and government agencies have developed sophisticated, complex information networks, incorporating technologies as diverse as distributed data storage systems, encryption techniques, Voice over IP (VoIP), remote and wireless access, and Web services. These networks have become more permeable as business partners access services via extranets, customers interact with the network through e-commerce transactions or Customer Relationship Management (CRM) processes and employees tap into company systems through Virtual Private Networks (VPN).

For hackers, these well-traveled paths make networks more vulnerable than ever before and with relative little expertise, hackers have significantly impacted the networks of leading brands or government agencies. Cyber-crime is also no longer the prerogative 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 instigate 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.

A joint report by [6] 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 almost 80 percent of lost data. The most common kinds of malware found in the caseload were those involving sending data to an external entity, opening backdoors, and key logger functionalities. At the same time, stolen passwords and credentials are out of control. Ineffective, weak or stolen credentials continue to wreak havoc on enterprise security. Failure to change default credentials remains an issue, particularly in the financial services, retail and hospitality industries.

In order to robustly protect enterprise and government networks against the complete spectrum of threats and vulnerabilities, all three methodologies of intrusion detection must be employed—Signature Detection, Anomaly Detection, and Denial of Service Detection and Prevention. Also, Intrusion Detection System (IDS) must do more than detect attacks: it should enable accurate detection to prevent attacks from reaching and damaging critical network resources and data. Without this range of detection methods—and the performance to accurately prevent attacks—many IDS products are no more than a digital Maginot Line: while they may offer the illusion of protection, when real attacks come, defenses can be circumvented or overrun.

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. 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.

2. RELATED WORK
Intrusions and anomalies are two different kinds of traffic events in an open network environment.

In [6], it is stated that Anderson, while introducing the concept of intrusion detection in 1980, in technical report of Computer Security: Threat Monitoring and Surveillance, defined "an intrusion attempt or a threat to be the potential possibility of a deliberate unauthorized attempt to access information, manipulate information, or render a system unreliable or unusable". Since then, several techniques for detecting intrusions have been studied.

[14] Illustrates 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 by novel, but 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.
This paper [15] states that 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.
[11] Describes IDS offer intelligent protection of networked computers or distributed resources much better than using fixed-rule firewalls. Firewalls are widely deployed as a first level of protection in a multi-layer security architecture, primarily acting as an access control device by permitting specific protocols (such as HTTP, DNS, SMTP) to pass between a set of source and destination addresses. Integral to access policy enforcement, firewalls usually inspect data-packet headers to make traffic-flow decisions. In general, firewalls do not inspect the entire content of the packet and can't detect or thwart malicious code embedded within normal traffic. It should be noted that routers also offer some rudimentary protection through packet-filtering processes. Firewalls and router-based packet filtering are necessary components of an overall network security topology; they are insufficient on their own to detect and prevent intrusions.
Existing IDSs are built with either signature-based or anomaly-based detection models [28]. 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.
L. Ertoz et al. [17] studied that traditional methods for intrusion detection are based on extensive knowledge of attack signatures that are provided by human experts. The signature database has to be manually revised for each new type of intrusion that is discovered. A significant limitation of signature-based methods is that they cannot detect novel attacks. In addition, once a new attack is discovered and its signature developed, often there is a substantial latency in its deployment.

The [4] described Signature-based IDS like SNORT employs a priori knowledge of attack signatures. The signatures are manually constructed by security experts analyzing previous attacks. The collected signatures are used to match with incoming traffic to detect intrusions. These are conventional systems that detect known attacks with low false alarms. However, the signature-based IDS cannot detect unknown attacks without any recollected signatures or lack of attack classifiers.

In [9, 16], it is concluded that signature matching performs well only for single-connection attacks. With the sophistication of attackers, more attacks involve multiple connections. This limits the detection range by signature matching. On the other hand, an anomaly-based system uses a different philosophy. It treats any network connection violating the normal profile as an anomaly.

Anomaly detection tracks events that are inconsistent with or deviate from events that are known or expected [27]. For example, in intrusion detection, anomaly detection system observed activities that deviate significantly from established normal usage profiles. Additionally, a network anomaly is revealed if the incoming traffic pattern deviates from the normal profiles significantly.

The [18] also focused on Network-level signature detection systems generally distinguish attack signature from legitimate traffic patterns by using certain detection thresholds, which are determined by trading off signature detection rates against false alarm rates.

[2] Focuses on a detailed comparative study of several anomaly detection schemes for identifying different network intrusions. Several existing supervised and unsupervised anomaly detection schemes and their variations are evaluated on the DARPA 1998 data set of network connections as well as on real network data using existing standard evaluation techniques as well as using several specific metrics that are appropriate when detecting attacks that involve a large number of connections. Here, the experimental results indicate that some anomaly detection schemes appear very promising when detecting novel intrusions in both DARPA'98 data and real network data.*This comparative study concluded that data generated from network traffic monitoring tends to have very high volume, dimensionality and heterogeneity, making the sense of data mining algorithm usage in intrusion detection .

The paper [19] illustrates that through a data mining approach, anomaly detection discovers temporal characteristics of network traffic. This system can detect unknown attacks and handles multi-connection attacks. 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.

Data mining is the latest introduced technology of intrusion detection [20]. Its advantage lies in the fact that it can withdraw the needed and unknown knowledge and regularities from the massive network data and host log data. It is a new attempt to use data mining in achieving network security, both at home and abroad.

Data mining techniques are also used to build classification models from labeled attacks [24]. Based on the in-depth study of existing intrusion detection systems, for the shortcomings of low precision, test result instability and the high false positive rate existed in the current intrusion detection, data mining techniques are applied to intrusion detection system, and according to common intrusion detection framework, an intrusion detection system based on data mining is designed. Experimental results show that this detection scheme can greatly improve the detection rate of intrusion detection system.

It stated earlier, the existing IDS are implemented using either misuse based or anomaly based intrusion detection models. SNORT and Bro are two widely used IDSs that are based on the misuse model [8, 26]. Even along with these IDS implementation models, a method developed for using Bayesian multiple hypothesis tracking to classify intrusion detection system events into attack sequences.

This may be used to reorganize data that is already being collected from intrusion detection systems in order to provide security analysts with a better situational view of what is occurring on their networks [12, 13]. By doing so, the actions of individual attackers are made clear so that the proper steps to minimize the potential damage and losses due to attack may be taken as rapidly as possible.

[10] Presents the approach with designing the correlation function of CRIM, a cooperative module for intrusion detection systems. After specifying an attack base in Lambda, the offline correlation process analyzes these attack descriptions to automatically generate a set of correlation rules. The online correlation process then applies these correlation rules on the alerts generated by the IDS to recognize more global attack scenarios.

Other attempts to solve the intrusion detection and response problem are described in [1]. Also describes a system that is able to detect the network intrusion using clustering concept. This unsupervised clustering technique for intrusion detection is used to group behaviors together depending on their similarity and to detect the different behaviors which are then grouped as outliers.
It [33] proposed Intrusion detection using sequential pattern mining in the field of information security. This paper first introduces several common sequential pattern mining algorithms, and then expounds 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, this [22] 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. A limited number of techniques have been proposed in the last few years for the detection of intrusion in databases. Therefore, there is still an urgent need to exert more effort to improve the performance of those systems.

With above survey of various kinds of IDSs and their implementation strategies, it is 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 [30], 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. The security of these computer systems is compromised when an intrusion takes place.

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.

During these reviews, it is observed that, Intrusion prevention has been used to protect computer systems as a first line of defense [31,32]. Intrusion prevention alone 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.

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 applied for false positive or false negative reviews detection [5].

3. OBJECTIVES

It is known that no single technique or technology is the "magic bullet" to guarantee protection against current or future attacks. In order to robustly protect enterprise and government networks against the complete spectrum of threats and vulnerabilities, all three methodologies of intrusion detection must be employed: Signature Detection, Anomaly Detection, and Denial of Service Detection and Prevention. IDS must do more than detect attacks: it should enable accurate detection to prevent attacks from reaching and damaging critical network resources and data. Without this range of detection methods and the performance to accurately prevent attacks, many IDS are no more than a digital Maginot Line. So to handle this, the main objectives of proposed research work are set as:

It gives requirement of IDS system which should do detection of all kinds of known, unknown attacks with minimization of false positives and false negatives rate of detection.
Even it is clear that no single technology can provide solution to this problem. The proposed work will integrate Signature Detection, Anomaly Detection methodologies with their merits to address this issue.
To prevent computer and network resources from data loss and damages, provision of prevention steps will be invoked automatically, once known or unknown attack is detected.
This is definitely a challenging task and to make this system effective, there is need of techniques to handle large amount of data for any kind of attack detection automatically instead of manually. Data mining (DM) techniques are useful to do such kind of effective knowledge discovery in automated way.
Additionally, it is proposed to make use of sentiment analysis techniques to enhance the performance of proposed system as innovative idea for false positive or false negative opinion detection.
In summarized way, the objective of the proposed work is: "How to automatically and systematically build adaptable and extensible advanced intrusion detection system using DM 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. USEFULNESS

In academic organizations, the proposed research system will be applicable effectively with only permissible software access to registered users in existing organization campus and can prevent important online document access to the outside world.

In IT industries, the proposed system will be useful to isolate the outside world internet access within organization by their employees for applications like e-mail, chat, etc. during working hours and even track of malicious data traffic can be traced and blocked.

For cyber cafe systems too it will be effective system to assign the legal usages of internet to all kinds of users and to identify intruders for their own network.

If it will be integrated with the military security applications, it will be very effective system for nation to trace out cyber-attacks and their impacts rapidly.

From the network administrator's point of view, it would be very user friendly system to admin their network through the GUI based interface instead of having command-based interface which is almost integral part of today's existing IDS administration.

This will open up new field / avenue for researchers and technocrats.

5. CONCLUSION

Since the intrusion detection and prevention is such a large subject, there is plenty of scope for its advancements. The proposed system intends to speed up the attack data detection and its prevention could be improved by applying the attack data inference detection. If proposed system would be implemented with this detection capability, then this will be major improvement in security field. Further the system can be improved from intrusion detection and prevention to the generalized cyber security system like existing anti-virus systems. The initial intention of the proposed system is for the small Internet and LANs. It has wide scope to its expansion as well as its application in the diverse field where it could play vital role in the intruder identification process to maintain the security for wired as well as for wireless environment too.

6. REFERENCES

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

[2] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava(2003), "A Comparative Study of Anomaly Detection Schemes in Network Intrusion Detection", Proc. Third SIAM Conf. Data Mining,pp.1-12.

[3] Aurobindo Sundaram (1996), "An Introduction to Intrusion Detection", pp.1-10.

[4] B. Casewell and J. Beale (2004), SNORT 2.1, Intrusion Detection, Syngress Pub, Second Edition.

[5] Bing Li (2010), "Sentiment Analysis: A Multi-Faceted Problem", IEEE Intelligent Systems, pp.1-5.

[6] CSI and FBI (2010), "CSI & FBI Report 2010", pp.1-2.

[7] D.J. Burroughs, L.F. Wilson, and G.V. Cybenko (2002), "Analysis of Distributed Intrusion Detection Systems Using Bayesian Methods Performance", Proc. IEEE Int'l Computing and Comm. Conf., pp. 329-334.

[8] D.J. Ragsdale, C.A. Carver, J. Humphries, and U. Pooch (2000), "Adaptation Techniques for Intrusion Detection and Response Systems," Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics, pp. 2344-2349.

[9] E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo (2002), "A Geometric Framework for Unsupervised Anomaly Detection: Detecting Intrusions in Unlabeled Data", Applications of Data Mining in Computer Security, Kluwer Academic Publishers, pp.1-20.

[10] F. Cuppens and A. Miege (2002), "Alert Correlation in a Cooperative Intrusion Detection Framework", Proc. 2002 IEEE Symp. Security and Privacy, pp. 187-200.

[11] Fengmin Gong (2003), "Next Generation Intrusion Detection Systems (IDS)", McAfee Network Protection Solutions, pp.1-5.

[12] Hesham Altwaijry, Saeed Algarny (2012), "Bayesian based intrusion detection system", Journal of King Saud University – Computer and Information Sciences, pp. 1–6.

[13] Hesham Altwaijry, Saeed Algarny (2011), "Multi-Layer Bayesian Based Intrusion Detection System", Proceedings of the World Congress on Engineering and Computer Science 2011 Vol IIWCECS 2011 ISBN: 978-988-19251-7-6, pp.1-5.

[14] K. Hwang, Y. Chen, and H. Liu (2005), "Defending Distributed Computing Systems from Malicious Intrusions and Network Anomalies", Proc. IEEE Workshop Security in Systems and Networks (SSN '05) held with the IEEE Int'l Parallel & Distributed Processing Symp, pp.1-8.

[15] K. Hwang, Y. Kwok, S. Song, M. Cai, Y. Chen, and Y. Chen(2006), "DHT-Based Security Infrastructure for Trusted Internet and Grid Computing", Int'l J. Critical Infrastructures, vol. 2, no. 4, pp. 412- 433.

[16] K.S. Killourhy and R.A. Maxion (2002), "Undermining an Anomaly-Based Intrusion Detection System Using Common Exploits", Proc.Int'l Symp. Recent Advances in Intrusion Detection (RAID '02), pp. 54-73.

[17] L. Ertoz, E. Eilertson, A. Lazarevic, P. Tan, J. Srivastava, V. Kumar, and P. Dokas(2004), "The MINDS—Minnesota Intrusion Detection System", Chapter 3:Next Generation Data Mining, MIT Press,pp.1-21.

[18] M. Cai, K. Hwang, J. Pan, and C. Papadopoulos (2007), "WormShield: Fast Worm Signature Generation Using Distributed Fingerprint Aggregation", IEEE Trans. Dependable and Secure Computing, pp.1-35.

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

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

[21] P. Ning, S. Jajodia, and X.S. Wang (2001), "Abstraction-Based Intrusion Detection in Distributed Environments," ACM Trans. Information and System Security, vol. 4, no. 4, pp. 407-452.

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

[23] R.P. Lippmann and J. Haines (2000), "Analysis and Results of the 1999 DARPA Off-Line Intrusion Detection Evaluation", Proc. Third Int'l Workshop Recent Advances in Intrusion Detection (RAID '00), H. Debar, L. Me, and S.F. Wu, eds., pp. 162-182.

[24] S. Noel, D. Wijesekera, and C. Youman (2002), "Modern Intrusion Detection, Data Mining, and Degrees of Attack Guilt", Applications of Data Mining in Computer Security, D. Barbara` and S. Jajodia, eds., Kluwer Academic Publishers, pp.1-29.

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

[26] V. Paxson (1998), "Bro: A System for Detecting Network Intrusions in Real Time," Proc. Seventh USENIX Security Symp., pp.1-18

[27] W. Fan, M. Miller, S. Stolfo, W. Lee, and P. Chan (2001), "Using Artificial Anomalies to Detect Unknown and Known Network Intrusions", Proc. First IEEE Int'l Conf. Data Mining, pp.123-130.

[28] W. Lee, S.J. Stolfo, and K. Mok(2000),"Adaptive Intrusion Detection: A Data Mining Approach", Artificial Intelligence Rev., vol. 14, no. 6, pp. 533-567, Kluwer Academic Publishers, pp.1-40.

[29] W. Lee and S. Stolfo (2000), "A Framework for Constructing Features and Models for Intrusion Detection Systems", ACM Trans.Information and System Security (TISSec), pp.227-261.

[30] Xiangyang Zheng, Qian He (2011), "Research on Distributed Intrusion Detection System Model", Energy Procedia, pp.1480-1485.

[31] Yang Lan (2011), "Design and Implementation of Intrusion Detection System Based on Data Mining", Energy Procedia 13, pp. 5645-5651.

[32] Yanjie Zhao (2011), "Research of Network Intrusion Detection System Based on Data Mining", Energy Procedia, pp. 1126 – 1132.

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