International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Impact Factor (2020): 6.733

An Examination of Big Data Analytics Frameworks for Targeted Cyber-Attack Detection

Karan Chawla

Abstract: ***Traditional intrusion detection systems isolate alarms and concentrate on low-level threats. Due to the numerous alerts received each day. Since human users must put in so much effort, it is nearly difficult to thoroughly investigate each alarm message. Analyzing historical data and looking for abnormalities that depart from the norm constitute anomaly detection. The benefit is that it can identify unidentified assaults. The drawback is that because network users' unpredictable behavior makes false alerts possible. Many systems for detecting anomalies rely on data mining methods. However, these abilities rarely keep up with the various forms of assaults and the rapidly evolving technologies. However, taking into account human aspects in anomaly identification gives us the chance to enhance the current algorithms and provide better outcomes. SNORT is the de facto industry standard and a reliable, proven system technology. In past research, SNORT log data was not used to compare various anomaly detection methods. SNORT log data analysis software is already widely available; however these programmes are purely visualization tools and do not use data mining techniques. Using Big Data Analytics, HeteMSD is a framework for identifying targeted cyber attacks. The recommended framework's name is heterogeneous multisource data. There has to be a strong framework that can aid security analysts in order to reduce the blindness of data analysis from many data sources without decreasing the level of digital security assurances. A correlation engine can reduce alert volume while just analyzing one log resource by grouping several warnings that are a part of an ongoing assault. Alert threading is the term for this procedure. In the case of heterogeneous log resources, a correlation engine should be able to determine if reports from several log resources relate to the same incident. This review paper analyses three types of big data techniques such as SNORT, Heterogeneous Data and Alert-Intrusion Detection technique for targeted cyber-attack detection.***

Keywords: Intrusion Detection, SNORT, HeteMSD, Cyber-Security, Attack, Threat, Big Data

1. Introduction

All forms of computer network assaults are based on a set of general concepts that also apply to physical security breaches as well as computer security breaches. Can we identify patterns of misuse using this information to identify behavioral models of anomalies? is the question that this study aims to resolve. This necessitates a paradigm shift away from reinforced defense and towards cybersecurity as a people issue.

In order to create data-driven judgements and ascertain client preferences, big data analytics is frequently employed. But in today's contemporary corporate world, big data analytics is applied to enhance cybersecurity. Businesses may use big data and data analytics to determine what is "normal" and, depending on the results, tighten cybersecurity standards. After researching an attack, big data analytics also enables businesses to see how an attacker may get into their systems using data analytics and machine learning. Regular data analysis enhances the ability to identify assaults before they are manually identified. Computers can do intricate analyses often, practically in real-time, to guarantee the security of your systems. They search via many different data sources. from user actions and network events to server and application logs.

Big data analytics is rapidly expanding in cyber security networks as a result of business executives placing a high priority on the rapid and accurate identification of contemporary cyber security threats. Massively vast amounts of data may be quickly handled thanks to the use of big data analytics in cyber defense. In turn, this enables the early identification of weaknesses and abnormalities, significantly increasing overall resilience. The standards of corporate intelligence and cyber security networks are improved by big data analytics' statistical, machine learning, predictive modeling, and computing capabilities.

Armed with historical data extracted from a wide range of sources, cyber security analysts and defense engineers can create statistical models or AI-based algorithms. Experts are able to swiftly identify weaknesses by setting up a baseline for typical activities. As a result, it can be claimed that using big data analytics in cyber security networks has enabled analysts to anticipate cyberattacks. Big data analysts are now able to identify deviations from the norm in order to anticipate impending assaults by combining technology and cutting-edge solutions like artificial learning, data mining, machine learning, natural language processing, and statistics.

Big data analytics offers automated monitoring and threat detection solutions that enable continuous environment monitoring and real-time detection.

It is important to provide credentials to authorized users as more organizations experience cyber security issues brought on by insider threats and carelessness. Automated monitoring allays this worry by gathering and evaluating user behavior information. Automated monitoring generates an alarm promptly in the event that any odd or potentially hazardous action is discovered. The most recent game-changing inventions, such Security Information and Event Management (SIEM) systems, originated with intrusion detection systems (IDS). In reality, SIEM is still developing because of the unrivaled power of machine learning, which helps cyber security analysts and cyber defense professionals manage a variety of massive, unstructured data sets more effectively and accelerate consolidation, correlation, and insights.

All firms, no matter how big or little, may use big data analytics solutions to improve their cyber security. Businesses of all sizes can now replace traditional security tools with big data analytics in cyber defense in an effort to overcome cyber security challenges thanks to the combination of machine learning and AI, real-time detection, automated monitoring, and data intelligence.

SNORT

There are universal principles that underlie all types of computer network attacks; these principles apply to both computer security breaches and physical security breaches. Can this information be used to spot behavioral models of anomalies where we can spot patterns of abuse? This calls for a paradigm change to see cybersecurity as a people problem rather than a matter of fortified defense.

Since computer networks have proliferated over the past ten years, technology has advanced to the point where networks are now interconnected with every aspect of the lives of regular people. Computing systems are increasingly vulnerable to assaults, exploitation, and misuse as they grow more common and networked. Network intrusions are a new societal issue brought on by the development of computer technology. Today's internet has a growing selection of tools and techniques that can be used to break into and attack private networks. Network intrusions are growing more common and more significant, making them a truly sensitive matter at all levels, including the government, small businesses, and personal life.

To safeguard networks and computers, there is a high requirement for efficient tools and detecting techniques. There are several Intrusion Detection Systems (IDS) for networks that have been created and are usable (Rehman 2003). There are generally two types of IDS: abuse detection systems and anomaly detection systems. The majority of commercial systems adopt mistreatment tactics that recognise common sorts of incursions. These are additionally known as signature-based invasions.

One such well-known signature-based detection system that has seen extensive application is SNORT. Researchers are now concentrating on using data mining and social network algorithms to analyze the alert records. These articles describe several intrusion detection techniques that the authors have developed. They also reveal hidden patterns that are not apparent from simply analyzing system communications.

One of the most widely used IDS systems worldwide is SNORT. More than 400, 000 registered users are utilizing SNORT to secure their systems, claims the website www.snort. org. SNORT is a robust and established system technology and is the de facto industry standard. SNORT log data was not used in earlier studies to evaluate different anomaly detection techniques. There are various software programmes already available to analyze SNORT log data, however these programmes are merely visualization tools and do not contain any data mining methods.

We may be able to learn more about attack patterns and relationships by examining IDS alarm signals, which may help us find system vulnerabilities. The signature database can utilize these discovered patterns to create new system rules.

Understanding IDS's setup and the data in the alarm log it produces is one of the main challenges of employing IDS alerts. Sometimes the information about log variables in IDS user manuals is sparse or spread out over several chapters. Although SNORT has a very thorough user manual, it might be challenging to locate all the log variable information because it is dispersed over many chapters.

Parsing the log data into a format that may be used is a significant problem to solve. SNORT creates alarm log data in text format, which makes it difficult to utilise as a database for analysis. Reading the text data into a database is the initial step in parsing these text alerts. This work is difficult in a number of ways. The text data is semi-structured, to start. Each alert has a different set of data pieces or an alert structure. Therefore, when we import the whole data into a database, we must extract all of the missing variables and account for them for other entries. Second, each warning or data record is divided into numerous lines rather than being contained on a single line.

So, in order to read all the components of the same record into a single data record, we identify patterns. Third, the names of the data variables are part of the records and may be the same across alerts. As a result, we must be aware of them and transpose the variable names into the header data rather than the data lines. Fourth, there is a lot of data. Therefore, we handle them using database approaches.

An attacker will select a target in a cyberattack before starting an attack. An attacker will need to attempt more than once and switch up their tactics till they are successful or give up, before the attack succeeds or fails. This describes the nature of cyberattacks, which are essentially all of the same kind. This presumption eliminates assaults by insiders because they are already familiar with the system and do not need to scan it or try several unauthorized access techniques.

So, in order to read all the components of the same record into a single data record, we identify patterns. Third, the names of the data variables are part of the records and may be the same across alerts. As a result, we must be aware of them and transpose the variable names into the header data rather than the data lines. Fourth, there is a lot of data. Therefore, we handle them using database approaches.

The sensors of the network must be positioned inside the network to gather data that is intended to uncover such assaults in order to detect insider or expert attacks when there are no scanning or probing stages. We see these assaults as having been effective in obtaining the required access after the first hurdle. In this study, the more specific characterization of assaults on the obtaining access stage was investigated. The destructive phase, in which successful attackers attempt to either take the information or disrupt the network, was not our main emphasis.

In most cases, a cyberattacker won't be successful on their first attempt. To access the target, the attacker will use a

variety of techniques. The target address is often unique or scarce when an attack occurs. Attackers will keep going for the special targets whether they succeed or fail. A target IP address is being attacked if it receives access from a disproportionately large number of distinct IP addresses in a brief period of time.

Massive efforts made in a short period of time to learn the password are one of the most popular attack strategies. Therefore, if a single or a small number of IP Sources are sending a lot more traffic to an IP Target than usual, the IP Target is under assault.

Attackers often do not carry out slow and persistent attacks manually once they have been launched. They'll create robots or software to automate these processes. These automated programmes will have characteristics in common that might be used as hints when designing models to find such assaults. The length of time between such strikes is one of their characteristics. These attacks pass for regular site visits, yet they happen at precisely the same intervals over a longer duration.

These might be regarded as false positive rates if we assume that the alarm data in SNORT logs is fully correct. However, this is due to the SNORT log's restriction that only allows for individual instance labels in the alert messages. Collectively, they are unable to recognise assaults that are still in the probing and scanning stages. Therefore, the occurrences found by Model 1 in the Priority 3 category might possibly represent concealed assaults that SNORT Alerts missed.

Heterogeneous Data

Network assaults in the modern network environment are more slick and organized. Governments, businesses, and other organizations are subject to severe financial losses and security threats as a result of advanced persistent threat (APT). Until weeks, months, or even years later, network security administrators are unaware that their network has been infiltrated. Targeted cyber-attacks that include complexity and customization-related qualities can't be effectively handled by traditional detection systems. Cyberattacks with several steps that are specifically targeted are becoming the biggest threat to network security. There are two reasons why rapid detection and analysis of targeted cyberattacks are important. First, prompt identification of intrusion behaviors allows for the quick cleanup and recovery of compromised system functions following an assault. This can lessen losses brought on by less regular service hours. Second, because targeted cyberattacks are frequently carried out in phases, the present aberrant behaviors may not constitute a complete reconstruction of the assault process. We can learn more about an attack's purpose by promptly identifying its preattack procedures. We can take precautionary steps to stop follow-up assaults from causing more harm by anticipating and detecting them in time.

Host detection and network detection are two categories under which intrusion detection technology is currently being researched. The malicious code is represented by the host-based intrusion detection approach, which analyzes process behavior to discover the payload. In order to recognise unidentified network intrusions, network-based intrusion detection primarily monitors network flow. A more thorough understanding of security is provided by the combination of threat intelligence with the big data analysis technique of enormous logs and traffic data.

1) Heterogeneous multi source security data is challenging and has a greater variety of expression semantics. For diverse data sources, there are several attack detection techniques. However, it is not quite evident how the research questions combine. Academic research is deficient in structured discussion.
2) Network attack-related abnormalities cannot be found fast and effectively using current approaches. Data correlation is still a challenging topic. The advancement of targeted cyberattacks detection is still hampered by the data association approach in heterogeneous multisource.
3) The automation and intelligence of the current detecting technologies are unsatisfactory. Secure data semantic information is not well conveyed. Attack recognition still primarily relies on manual analysis.

The suggested methodology makes use of correlation analysis for attack investigation, which effectively handles the detection of targeted cyberattacks in a huge data setting. This concept serves as the foundation for a suggested inner-layer and cross-layer analytical strategy for targeted cyberattacks. On the basis of this, follow-up researchers can do more study.

HeteMSD is a Big Data Analytics Framework for Targeted Cyber-Attacks Detection that Uses Heterogeneous Multisource Data is the name of the suggested framework. In order to minimize the blindness of data analysis from diverse data sources without lowering the degree of digital security assurances, there needs to be an effective framework that can help security analysts. There are still some notable differences between the study on targeted cyber-attacks detection based on heterogeneous multisource data, and both practical and theoretical efforts for this purpose are still in the exploratory stage. The relevance of our suggested approach lies in its goal to resolve the tensions between sophisticated network attack defense and heavy diverse data loads. The following are the primary contributions of our study.

Targeted cyber-attacks are a subset of devoted assaults that are directed at a particular person, business, or organization with the goal of achieving a specific goal, such as stealing confidential information from a back-end database or disrupting system functionality. Targeted cyber-attacks have a trait of discrimination and are not random in nature. It means that attackers who carry out targeted assaults distinguish between their targets and watch for the right moment to carry out their attack strategy. Targeted cyber-attacks usually require several stages to achieve the goal. A successful targeted cyberattack method often consists of acquiring information, infecting targets, exploiting systems, stealing data, and keeping control. These actions are all crucial components of focused cyberattacks. Targeted cyberattacks require the achievement of all the aforementioned phases in order to be implemented. Targeted

cyberattacks vary from conventional assaults in that they are more sophisticated and frequently have significant intrusive motivations. Attackers take more time selecting their targets, looking for security holes, and creating unique malware. Targeted cyber-attacks are usually implemented by professionals instead of simply using attack tools. Another idea, known as an "advanced persistent threat, " must be brought up when discussing targeted cyber-attacks. A targeted cyberattack may be thought of as a subclass of APT. APTs are typically targeted cyberattacks using more sophisticated attack techniques. APT is implemented using a variety of different attack routes. It has been around for a while without being found in a real network setting. In general, the terms "targeted cyber-attacks" and "APTs" are equivalent when discussing advanced, complex, and targeted network attacks. There aren't many differences between APTs and targeted cyberattacks. The Kill Chain model may be used to explain how targeted cyberattacks are carried out.

Antivirus software and HIDS are the representations of the host-based detection technique. Malicious programmes are often found by monitoring system calls, network access, file operations, process formation, and memory modifications. Malicious programmes may be identified by static and dynamic analysis, and APT assaults can be stopped. The pattern-matching-based HIDS approach can successfully recognise known assaults. However, undetected assaults cannot be stopped, and the rule base has to be continually updated. Then a method of behavior-based detection is suggested. Researchers have suggested a number of detection methods based on the observation of anomalous behavior in recent years as data mining and machine learning technology have advanced.

Network-based detection technique: Malware's pattern of command and control channel exhibits certain regularity (e. g., attack payload signature, network communication sequence characteristics, and created domain name). Network traffic produced by targeted cyberattacks is distinct from that produced in typical office settings. As a result, using the network detection approach, it is possible to determine the attack load of the attacking process.

Multisource data fusion can be used to achieve the correlation between events. Results of anomalous detection indicate event correlation for raw input data. Through complementarity, data fusion may lessen data duplication and cooperative information gathering. Since the formats of the original heterogeneous data are inconsistent, the characteristics must be extracted. The comprehensiveness connection of data from several sources is used to provide the overall anomalous outcomes. The primary approach of heterogeneous event fusion used nowadays is to create a global feature vector by extracting features from various security data. A worldwide abnormality that cannot be conveyed by a single data source might be better reflected by the interaction between many data sources.

In comparison to single source data, multisource heterogeneous data requires a distinct approach to anomaly detection. There are three different types of multisource heterogeneous data anomaly detection methods currently available. The global anomaly is determined following an analysis of the results of the application of anomaly detection algorithms to various data sources. Second, several data sets are combined to create a single data source that has the same data pattern. In this approach, the classic single-source data anomaly detection problem is created out of the multisource data anomaly detection problem. Thirdly, more data sources are used in the anomaly detection process to reinforce and support the findings.

Alert-Intrusion Correlation

The three sorts of security mechanisms offered by computer security—authentification, authorization, and auditing—are designed to safeguard a system. To protect the systems against assault, these three procedures are necessary. However, an extra layer of security is required if these techniques' conception and execution are flawed.

IDS have been suggested in order to add another line of defence. With both commercially sponsored and open source components being extensively used, intrusion detection technology is becoming more and more popular in workplace networks. It has certain flaws, though, including the propensity for alert flooding, contextual issues brought on by assaults that are likely to produce numerous related alerts, false alerts, and scalability. Correlation is suggested as a solution to these flaws. However, it is unclear if the possible additional leverage from the variety of devices necessitates the security administrator's choice of which reports apply to the same or to separate occurrences. This problem inspired the researcher to investigate the relationship between alerts generated by intrusion detection sensors and diverse log resources.

The term "alert" refers to an alarm that an intrusion detection system (IDS) generates to inform interested parties of a noteworthy incident. A low level entity evaluated by IDS is an incident. For the purposes of intrusion detection and response, intrusion correlation refers to the interpretation, integration, and analysis of information from all accessible sources concerning the target system activity. Intrusion event correlation and intrusion alert correlation are the two forms of intrusion correlation. The primary distinction between these two forms of intrusion correlations is that intrusion alert correlations identify abuse or abnormalities whereas intrusion event correlations evaluate neutral events. This connection is also mentioned in the IDMEF standard, which specifies that if an analyzer finds an event that matches a rule, it notifies one or more of its managers. A warning message may be associated with a single detected event or a number of detected events, depending on the analyzer. While intrusion warning correlation is crucial for security administrators and intrusion event correlation is vital for forensic investigation, this study will focus on the latter.

There are several sources from which an alert might be generated, and it can result in different stages of an assault. A multi-step procedure called alert correlation takes as input alerts from one or more IDS and outputs a high-level description of the malicious behavior on the network. Data must be gathered from a variety of sources (such as firewall, web server logs, IDS from various manufacturers, and so on) in order to obtain effective identification. The most evident

advantage of correlation of warnings generated by diverse log resources is the reduction in the number of alerts that a security officer must deal with.

By organizing several warnings that are a part of an ongoing assault, a correlation engine can minimize alert volume while only evaluating a single log resource. This technique is known as alert threading. A correlation engine should be able to tell whether reports from different log resources pertain to the same occurrence in the situation of heterogeneous log resources. The degree to which one or more characteristics or measurements on the same set of elements exhibit a propensity to fluctuate simultaneously is known as correlation. Correlation can improve detection abilities and provide a fuller picture of threats that a single sensor or device may only be able to monitor in part without losing the security-relevant data. Correlation can also take use of the supplementary coverage provided by various log resources.

Reports from various log resources using various analysis methods may support one another and so increase the confidence in the detection.

An alert will be compared to all other alert threads that have comparable properties or features (such as source IP address, destination IP address, or ports), and if a match is found, alerts with a high degree of feature similarity will be correlated; if none are, a new thread will be formed. To group alerts that are a part of the same ongoing assault, low-level events are aggregated in the first phase utilizing the ideas of attack threads. If there is an attribute overlap, the alerts are clustered, which means that only attributes that are present in both alerts are taken into account. The goal of this phase is to offer a higher-level perspective of the security condition of the system by combining warnings that indicate various attack actions.

Analysis

IDS that use signatures or rules often rely on professionals to create the rules and programme the system. Since Denning's (1987) first study, other intrusion detection methods have been suggested and put into practice. These programmes may be broadly divided into two categories: anomaly-based detection and misuse or signature-based detection. With signature-based detection, established attack signature patterns are saved and utilized to identify new attacks with similar signatures. The key drawback is that it is unable to identify new assaults with unidentified signatures. SNORT primarily concentrates on lone wolf attacks as opposed to group assaults. While our algorithms identify the possible risks based on the cooperation patterns of the attackers, it logs categorize visits as potential threats based on visit kinds and behaviors. Human variables are taken into account while identifying these patterns. SNORT can only analyze a visit at a time, whereas our models can aggregate data from many time periods to get more accurate references. The information gathered by our models might be added to SNORT, a real-time intrusion detection system, to improve its capacity to identify additional possible threats.

For Alert-Intrusion Correlation, All alert correlation mechanisms should be used in the alert correlation procedure in accordance with the capability requirements. The method of recognising known and unknown attacks, as well as multi-step attacks, should be further improved, as these capability criteria will help to solve the problem of a significant number of false alerts.

2. Conclusion

In order to analyze the alert records, researchers are now focusing on employing data mining and social network algorithms. These publications outline numerous of the authors' original intrusion detection methods. They also make hidden patterns visible that would not be seen by merely looking at system communications.

SNORT is one of the most frequently utilized IDS systems globally. According to the website www.snort. org, SNORT is being used to safeguard systems by over 400, 000 registered users. SNORT is the de facto industry standard and a reliable, proven system technology. In past research, SNORT log data was not used to compare various anomaly detection methods. To analyze SNORT log data, there are currently a number of software programmes available, however these programmes are only visualization tools and do not really contain any data mining methods.

3. Future Directions

We will need to use research from several fields, such as new technologies, data mining techniques, attacker psychologies, typical user behaviors, etc., in order to identify assaults more precisely and to construct a strong detection system. One strategy will never be sufficient to handle all scenarios. Despite the complexity of incursion patterns and the unpredictability of human behavior, we will work to identify attacks quickly and cheaply.

HeteMSD still need further characteristics to allow for the integration of the human expert, who in addition to being a simple observer also possesses the expertise necessary to improve the early analyses that the framework has suggested. Anomaly detection based on multisource data fusion needs further development. The proposed expansion of the correlation approach will be the subject of further study. Last but not least, security investigation is anticipated to focus heavily on security knowledge reasoning in the near future.

It is necessary to conduct further study on the intrusion alert correlation technique in order to detect unknown attacks utilizing a mix of anomaly and abuse detection approaches.

References

[1] Y. Li, W. Dai, J. Bai, X. Gan, J. Wang, and X. Wang, “An intelligence-driven security-aware defense mechanism for advanced persistent threats, ” *IEEE Transactions on Information Forensics and Security*, vol.14, no.3, pp.646–661, 2019.

[2] J. Navarro, A. Deruyver, and P. Parrend, “A systematic survey on multi-step attack detection, ” *Computers & Security*, vol.76, pp.214–249, 2018.

[3] Y. Liu, M. Zhang, D. Li et al., "Towards a timely causality analysis for enterprise security, " in *Proceedings of the Network and Distributed System Security Symposium*, 2018.
[4] R. Zuech, T. M. Khoshgoftaar, and R. Wald, "Intrusion detection and Big Heterogeneous Data: a Survey, " *Journal of Big Data*, vol.2, no.1, pp.1–41, 2015.
[5] J. Navarro, V. Legrand, S. Lagraa et al., "HuMa: A multi-layer framework for threat analysis in a heterogeneous log environment, " *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol.10723, pp.144–159, 2018.
[6] P. Bhatt, E. T. Yano, and P. Gustavsson, "Towards a framework to detect multi-stage advanced persistent threats attacks, " in *Proceedings of the 8th IEEE International Symposium on Service Oriented System Engineering, SOSE 2014*, pp.390–395, IEEE, UK, April 2014.
[7] Cheung, S., Lindqvist, U., Fong, M. W. (2003). Modeling multistep cyber attacks for scenario recognition. In *DARPA information survivability conference and exposition, 2003. Proceedings* (vol.1, pp.284–292). IEEE.
[8] S. Mathew, D. Britt, R. Giomundo, S. Upadhyaya, S. Sudit, Real-time Multistage Attack Awareness Through Enhanced Intrusion Alert Clustering, In Situation Management Workshop (SIMA 2005), MILCOM 2005, Atlantic City, NJ, October, 2005
[9] Mathew, Giomundo, Uoadhyaya, Sudit, Slotz, Understanding multistage attacks by attack-track based visualization of heterogeneous event streams, Proceedings of the 3rd international workshop on Visualization for computer security, Virginia, USA, 2006.
[10] Lingyu Wang, Anyi Liu, Sushil Jajodia. Using attack graphs for correlating, hypothesizing, and predicting network intrusion alerts. Computer Communications, Vol.29, No.15, 2006, pages 2917-2933.
[11] Ourston, et all, Applications of Hidden Markov Models to Detecting Multi-stage Network Attacks, Proceedings of the 36th Hawaii International Conference on System Sciences, 2003.
[12] Siraj, Maarof, Zaiton, Hashim, Intelligent Alert Clustering Model for Network Intrusion Analysis, ICSRS Publication, 2009.
[13] Siraj, Vaughn, Multilevel Alert Clustering for Intrusion Detection Sensor Data, Fuzzy Information Processing Society, USA, 2005.
[14] C. Warrender, S. Forrests, and B. Pearlmutter. Detecting intrusions using system call: Alternative data models. In proceedings of the 1999 IEEE Symposium on Security and Privacy, May 1999.
[15] S. Kullback, R. A. Leibler: On information and sufficiency, Annals of Mathematical Statistics, 22 (1): 79–86, 1951.
[16] Kim, S. J., & Hong, S. (2011). Study on the development of early warning model for cyber attack. In *2011 International Conference on Information Science and Applications (ICISA)* (pp.1–8). IEEE.
[17] Liu, Z., Wang, C., Chen, S. (2008). Correlating multi-step attack and constructing attack scenarios based on attack pattern modeling. In *International conference on information security and assurance, 2008. ISA 2008* (pp.214–219). IEEE.
[18] Miles, W. (2001). Hack proofing sun solaris 8—protect your solaris network from attack (1st ed., pp.83–85, 257). New York: Syngress.
[19] Namayanja, J. M., & Janeja, V. P. (2013). Discovery of persistent threat structures through temporal and geo-spatial characterization in evolving networks. In *IEEE Intelligence and Security Informatics (ISI).*
[20] Rehman, R. U. (2003). Intrusion detection systems with Snort: Advanced IDS techniques using Snort, Apache, MySQL, PHP, and ACID. Prentice Hall PTR.
[21] Valdes, A., & Skinner, K. (2001). Probabilistic alert correlation. In *Recent advances in intrusion detection* (pp.54–68). Springer.
[22] Youssef, A., & Emam, A. (2012). Network intrusion detection using data mining and network behaviour analysis. *International Journal of Computer Science & Information Technology, 3.6*, 87–98.
[23] Dokas, Kumar, Lazarevic, Srivastava, Tan, Data Mining for Network Intrusion Detection, USA, 2004.
[24] Zhai, Y., Ning, P., & Xu, J. (2005). Integrating IDS alert correlation and OS-level dependency tracking. North Carolina: North Carolina State University.
[25] Julisch, K. (2001). Mining Alarm Clusters to Improve Alarm Handling Efficiency. Proceedings of the 17th Annual Conference on Computer Security Applications. New Orleans, LA