Investigating Identification Techniques of Attacks in Intrusion Detection Systems Using Data Mining Algorithms

Seyed Amir Agah

MSc. Computer Software Engineering, Arak Branch, Islamic Azad University, Arak, Iran

Summary
According to extraordinary growth of network-based services, intrusion detection has been introduced as an important and efficient issue in the direction of network's security promotion. In order for coping with the infiltrators of networks and computerized systems, different methods have been codified which are called intrusion detection method the aim of which is that illegal usage and abuse and damaging the systems and computerized networks is recognized by both internal users and external attackers. Generally, intrusion detection methods are divided in to two main categories of abuse detection and abnormal behavior detection. In abuse detection method, known intrusion patterns are used for recognizing the intrusions but in abnormal behavior detection methods, normal behavior of the users has the main criterion. As a result, any inconsistent behavior with that is considered as an effort for permeating the system. In this paper, intrusion detection systems and their performance are considered and finally we will become familiar with intrusion detection methods including data mining algorithms.

Key words:
Intrusion detection, computerized networks, security, data mining algorithms

1. Introduction

In today's world, computer and attached computerized networks to virtual world play an important role in communication and information transference. Meanwhile, opportunist individuals have proceeded to the violation of computerized systems by accessing the important information of specific centers or other individuals' information and with the intention of applying the penetration or pressure or even disassembling systems' discipline. Hacker ،Cracker and Intruder are the words which are nowadays proposed in computerized circles more or less and penetrate other systems and endanger their security. Therefore, the necessity of retaining information security and efficiency in computerized networks which are in relation with the outside world is quite obvious, and since in technical terms creating computerized networks (hardware and software) without weak points and security failure is practically impossible, intrusion detection in computerized systems' researches is followed with a specific importance. Security is a challenging domain for computers and remained networks. Also, protection solutions including firewall, identity authentication and encryption are not usually enough for network security. Therefore, intrusion detection systems will be a very important defensive mechanism for vulnerable points of computerized networks that face a large volume of data in network traffic. Due to this reason, extraction of some patterns for wrong behaviors of users in intrusion detection systems is a time-consuming and difficult work. On the other hand, some data in intrusion detection systems make disturbance for intrusion detection action that recently many researchers have concentrated on intrusion detection system based on data mining techniques. Major problem in intrusion detection system is discovery precision of new attacks. Therefore, unsupervised systems should be used. On the other hand, intrusion in the system should be diagnosed in real time. In spite of the fact that intrusion detection system in offline situation is also beneficial for removing poor security of the network, data mining techniques including its algorithms can lead the hidden information to be discovered from the entrance of network data (Gupta et al., 2016). Data mining techniques, due to enjoyment of automation capability and performance improvement have extensively been used in intrusion detection; however, using some techniques for intrusion detection has some problems and limitations including which we can refer to high complexity, instability and low accuracy for less attacks that we can use collective categorizations for removing such problems, Gaidhan et al, 2014. Intrusion detection system is a software which supervises the activities of network or system in order to find destructive activities that in this regard, there are multiple techniques for intrusion detection which are for creating more security for the network which are statistic (Stolfo et al, 2015). Intrusion detection system discovers suspicious patterns of network traffic through analyzing traffic informational packets that major problem of the existent models is recognizing new attacks, low accuracy, recognition time and compatibility. Security researches in computerized networks have shown that environmental factors and conditions in computerized networks are effective in the quality of intrusion detection that the aim of this research is investigating intrusion detection in computerized networks using data mining algorithms.

Manuscript received May 5, 2017
Manuscript revised May 20, 2017

- Security:

Security in real world is a vital issue for all human beings. In pre-history era, security meant survival protection principles such as security versus others or animals' attack and also food supply security. Other needs such as security versus natural events or diseases were not commonly proposed for pre-history humans. Following civilization progress, security range even went beyond and involved wider dimensions including having a place for comfort and safe life and nowadays, personal properties concept has also been added to security definition. Most of what we do in real world is accompanied by danger; however, many of our activities are accompanied by less danger. For example, when we go traveling with an unfamiliar person or when we enter an unfamiliar city or country, we know the fact that there are some threats for our corporal security. Existent threats around us will be serious only when we are in an unprotected place and face a person who can abuse our situation. If we pay enough attention to our surrounding dangers, we will succeed to find a secured place or find a solution. For example, we can accompany a person who can direct us toward a secured place or take a taxi. Some affairs are accompanied by financial or psychological dangers, but they have no corporal danger. When we invest (in any forms of purchasing the land, stocks or even activity in commerce or working in the market), we expect this capital to be returned to us the sooner. As we know, some investments will be returned sooner or later, whereas some investments are not as such and lead to loss. For example, when we communicate with a new person, we hope this new relation to have some benefits for us. We also accept the danger that this problem might have no necessary advantage. In some fields, accessing a level of security that we expect is not possible. For example, we are always willing to have a long life and a healthy body; but statistical GPA of life span shows that this problem doesn't apply to many individuals. Some of us die in young ages. Some of us, during the lifetime, grapple with different diseases and some others of us survive for many long years and spend a healthy life. We compensate our inability in determination of destiny with insurance in order for it to protect us against negative financial effects, events and diseases. Absolute security whether in real life or in cyberspace is impossible; however, adequately suitable security is almost achievable in all environmental conditions.

Security as an important element in relation with information technology

During past decades especially in the past five years, the world has been the arena of impressive changes, which have changed many previous relations and equations. Such changes which are possible through extensive control of ICT, have begun with computer application as the automation and productivity increase device and now have actually changed social and individual life of human by evolution of its control in creating collaborative synergistic space. According to many pundits, as the genesis of writing has had a tremendous effect on human destiny that has impelled the historians to divide human life story on this planet to history and prehistory era, entrance to the acquired virtual space of modern information and communication technology has also figured out a new period of human civilization so that modern revolution has changed the manner of thought, production, consumption, commerce, management, communication, war and even religiosity and lovemaking. This great change has been accompanied by many requirements and consequences the most important of which is the genesis of modern concepts of virtual security or security in cyberspace. Following the occurred change in predicating computerized network expression from a small network of teamwork to global and widespread networks (internet) and according to increasing growth of the interactions and exchanges on computerized networks, the need for protection and electronic security systems in order for exchanges guarantee and creating legal commitment for the individuals involved in the equation is so vital. There are some systems including the laws, methods, standards and devices that are even more undertaking than common contracts and traditional methods and additionally guarantee the security and privacy of sensitive exchanged information more than ever. Information security in virtual spaces has always been considered as one of the basic infrastructures and requirements in developmental and pervasive control of ICT. Even though, absolute security is not accessible neither in real world nor in cyberspace, creating a level of security which is adequate and in proportion to needs and conducted investment is possible in almost all environmental conditions. Only in such an optimum level; real individuals, organizations, private companies and governmental organs play their expected role as efficient groups of this interactive and synergistic network in addition to trust to different parties who are all involved in an electronic exchange and presumably have never seen and known each other. Insurance of informational capitals and infrastructure equipment of the country, in addition to widespread dimensions of national security, is the key of countless online new commercial and non-commercial chances. Actuality is that security challenge ahead the country isn't inaccessibility to technology or lack of security products rather it is policy making, culturalization, and suitable efficiency of the existent sources and their compatibility in a manner that it provides unique need of the network and digital environment of the country. In this regard, it is important to note that information security architecture is a process of current processes in information technology architecture in different levels

including national and organizational levels that in this process, necessary devices will be used based on need.
Another important obtained point of other countries' experiences is that information security is a multi-sectoral problem and there is the need for widespread cooperations in this regard. Such cooperations should be considered both in national and international level. Determination of roles, duties and responsibilities is among the important points which should be defined in such cooperation (Saraswati et al. 2016) and (Ashfaq et al. 2017).
Generally, information technology security category refers to the following issues:

- computer security

Security in technical terms in cars, soft wares, data and networks

- Cyber Security:

Information technology security dependent on governments' policy; This term is usually used by means of governmental institutions and national policy makers in documents, laws and research projects and it is more or less synonymous with "internet security". Both two phrases refer to the aspects of network security and policymaking principles of the networks such as privacy definition, cyber-crimes, business and global communications. The difference of these two terms is not so much rather security of computers, networks and data has to a large extent been mixed with routine concepts of security in cyberspace (Elhag et al. 2015).
Security threats of the network
In computerized networks, any effort in order for eradication, disclosure, change, non-activation, robbery or gaining illegal accessibilities or illegal usage of a finance is called attack. Nowadays, unlike the past, attacks preparation by possessing abundant and available devices doesn't need much knowledge and the number of attacks against information security has considerably increased. In order to estimate the amount of security, we assume that aggressive accesses to the posted information through telecommunication channel and details related to encryption function except key bits (Gai et al. 2016).
In a computerized network, attacks are the result of active services' link, used protocols and open ports. Information security experts by concentrating on three above principles should cause a secure and robust network against kinds of attacks. Attacks in computerized networks are divided in to two categories of active and non-active. Active attacks change the system or network but non-active attacks search for collecting the information from systems. Active attacks affect the accessibility, integration and accuracy of the data, whereas, non-active attacks disturb the confidentiality (Saraswati et al. 2016).
Intrusion: it is predicated to a series of illegal actions which endanger accuracy and confidentiality or accessibility to a source. Intrusions can be divided to two internal and external categories. External intrusions are those kinds of intrusions which are accomplished by allowed or non-allowed individuals from outside the network toward internal network and internal intrusions are accomplished by allowed individuals in system and internal network from inside the network itself and they penetrate the systems and computerized networks through software defects, breaking passwords, eavesdropping of network traffic and weal points of designing in networks, services or network computers (Costa et al. 2015).
Permeability: computerized permeability is in fact the description of actions which lead to violation to security methods by means of a software system that such actions are accomplished by means of access to valid states through the transits. Also, a state is exposed to the danger of a state in which hackers attack sequence from valid state transits which results in permeability of valid states. Also a permeability state is in fact a type of all impermeable states and generally, permeability might determine many permeable states among which only one state is chosen (Gautam and Om, 2017).
Permeability in computerized systems is a weakness which exists in automatic security routines of the system, management controlling devices, internet controlling devices that hackers can lead to interruption of sensitive performance of the system by means of this weakness (Ashfaq et al. 2017).
Permeability ways in computerized networks: they include the attacks in a computerized network which become possible through three main factors including: active services, used protocols and open ports and according to unknown nature of users in computerized networks including the internet, all service providers will face attacks' challenge (Joffe, 2016).
Attacks to computerized networks include: attack to IP address, TCP protocol, espionage attacks, information forging, Applets, Cookies, practical programs, eavesdropping any of which are as the following:
Web forging: in this method, initially a version of website is copied and then the hackers manipulate the stored version without any change in the appearance. Hacker uploads the falsified page and somehow attracts user's attention to that and user clicks on the intended link and hacker attains his own goal (Devi et al. 2017).
Attack through TCP protocol: in this method, hacker disconnects internet service providers and users and IP introduces itself to the service providers as the user and any information exchange occurs between service provider and hacker. The advantage of this method through IP attack is that hackers attack only once and refrain from coping with security systems of password and

this method is the most common type of attack to internet service providers (Elhag et al. 2015).

- Attack to IP address: in this attack, hackers in the direction of network plan access accede IP service provider through different techniques, then pose themselves between service provider and user and steal the information by transmitting fake packages. In fact, in this method, hackers introduce themselves as the receiver for service provider and as the service provider for user and they are able to transfer their own packages with correct number as the intermediate between user and server (Saraswati et al. 2016).
- Espionage attacks: in this method, hackers through uncoordinated relations with TCP protocol replace serial number of sent packages of service providers with serial number of the next packages and copy the packages for themselves, then hackers again send their intended package, but this time, they send their own determined serial number for service providers and in this way, in case of user and service provider's unawareness, information is diminished or raised by the service providers (Adebowale et al. 2017).
- Information forging of IP address: hacker can introduce itself as host or hacker through routing. Hacker forges service provider's address based on user's address and then makes a new address for user and in this way, hacker interrupts user's connection and connects itself to service provider by forged address of the user (Joffe, 2016).
- Email forging: this online method is very simple and easy. Because, it is enough that hackers have adequate information about programming and email sending (Gaidhane et al. 2014).
- Applets: Applets are in fact java codes that can be dangerous. Browser pages are directly loaded on memory, namely java codes are automatically applied following the entrance to a browser web page. This capability causes the hackers to be able to write destructive codes and upload in web pages (Levitt and Gihan, 2017).
- Cookies: Cookies are small files that dynamic web pages can create in users' computers and maximum length of these files is 4 KB. Such compact and apparently poor information can be so beneficial for hacker (Devi et al. 2017).
- Attacks to passwords: in this method, hackers take control of all needed sections in order for finding passwords including confidential, commercial, security information password and even E-mail password (Petersen, 2015).
- Eavesdropping: although the hackers can hear the data by installing the Sniffer software, they are not permitted to change those data (Kevric et al., 2016). It is possible to eavesdrop on the network when the Ethernet protocol is designed based on the Subscription management in the computer networks. All the machines use the common wire in the local network. Therefore, it is possible for the machines to eavesdrop the traffics of the network. However, the Ethernet provides a filtration for each machine, so they can take their own traffic related to the specific address. The next level is about the eavesdrop program which is able to remove the filters and put the Ethernet in a free mood. The Ethernet hardware can receive the transmission pockets from the networks (Dash et al., 2017; Chand et al., 2016). Therefore, the process of eavesdropping is possible in the following networks: wireless, WAN and LAN networks (Nair et al., 2017).
- DNS systems: Because of the lack of security, cover and the problems related to the DNS settings, if the hacker can change the unreal IP for the users, it is able to show the address for the fake sites and users. He can points to the user whether the site is true or forged. Then the emails are sent to the wrong address. Numerous dangers threat this system such as cash poisoning: this process contains of entering the wrong information in the cash of server, which relates to the name of the system (DNS). We can send the forged reply using the spoof address based on the applicant information. There is an exception here: if the forged reply gets sooner than the main reply, it places in the cash. After that, the cash would be poisoned by means of the forged information. Numerous users are in danger until the expiration of forged information (Eesa et al. 2015; Kim et al. 2014).
- DOS: hackers send to much traffic to the IP address through the destructive bot. therefore, servers cannot answer.
- DDOS: hackers are able to send verities of demands to the related IP address.
- Running specific program and controlling the system: the attacker can control the program using specific instruction. This process destroys the data, so the CGI Scripts would able to control the system.
- Ping method: hackers can find the IP address of different sites using this method. We can determine whether the IP address is active or not by sending four packets to the mentioned address and achieve the response. We can also obtain different information in this process.
- The weak points in the design of networks system: operating systems are applied programs. Those weak points, which threat the software, also threat the operating system. In other words, the attacker can implement these weak points in order to control the system (Elhag et al. 2015). The attackers begin to attack after choosing the best method of attacking. The attacker first recognizes the goal, and then he uses appropriate scenarios in accordance with the weak points in the software. He can obtain some information as the following: password, which contains of personal information, number of children and dictionary attacks. In this type of intrusion, hackers test specific words in order

to enter the system. They also test possible combinations of characters in this step. This kind of method is useful for those codified words with few letters. After finding the network architecture and applicants, they replace their IP address in another place by forging the IP address.

- Certain signature attack: in this kind of attack, the attacker is able to access the messages and digital signature, which is provided by the owner of signature. Actually, this kind of attack is possible to execute. Therefore, each method of sign production must be safe in this method. The attacker is aware of the general key of the sign's owner (Levitt and Gihan, 2017).
- Certain plaintext attack: this is a kind of attack in order to analysis password. The attacker has sufficient samples of certain plaintext and the codified version. Classic cryptographies are vulnerable in contrast with the certain plaintext attacks.
- Collision attack: In the cryptographies, the collision attack takes place in the related function. Here, the purpose is finding two optional entrances and a summary of produced message. There are two types of collision attacks as the following: Classical collision attack and Chosen-prefix collision attack. The attacker does not have any control on data in the classical collision attack. Data are chosen by means of the algorithm. The chosen-prefix collision attack is more powerful than the classical one. The attacker can choose two different documents and then adds some values to all documents.
- Man-in-the-middle attack: it is a kind of attack that is called "the active eavesdropping". The attacker provides independent connections among the victims and distributes their messages. The Man-in-the-middle attack is appropriate for the encrypted communications and interaction protocols of the keys. In this type of attack, when two users are interchanging their general keys, the attacker puts himself between the two users. Both users suppose they have a secure connection in this kind of attack, but the attacker have been heard everything (Saraswati et al. 2016).

- Chosen plaintext attack

Chosen plaintext attack (CPA) is an attack model for cryptanalysis where the attacker is assumed to have access only to a set of intended ciphertexts and related encrypted message. In this attack, the attacker has access to the encryption hardware. It aims to gain more information which can decrease the security of cryptographic schemes. In the worst case of chosen plaintext attack, the secret key of scheme can be disclosed. In some CPAs, only a small portion of the text needs to be chosen by the attacker. Such attacks are known as textual injection attacks (Elhag et al., 2015).

- Replay attack (playback attack)

Replay (playback) attack is a kind of attacks on computer networks performed through retransmission of valid data intercepted by the attacker. These attacks are carried out in order for the purposes such as retransmission of the intercepted credentials to the server and illegal authentication (Devi et al., 2017).

- Malwares

Malwares are computer programs designed and applied by the attackers to disrupt the computer performance, gather the important information, or access to private computer systems. The malware can be appeared in the form of code, script, active content, and other software. The malware is a general term used for various forms of bothersome software. It includes:

- Computer virus
- Ransomware
- Computer worm
- Trojan
- Rootkits
- Back door
- Key logger
- Dialer
- Adware (advertising software)
- BHO malicious
- Rogue security software and etc.

Intrusion detection and its methods: Network-based computer systems play a significant role in communication, which are faced with the challenge of intrusive hackers. Hence, besides the intrusion prevention methods including user authentication such as passwords, using firewalls and data protection like the encryption, the intrusion detection is another filter for protecting the computer cables. Its purpose is to detect the unauthorized use, misuse and damage to computer systems and networks by both internal users and external attackers (Dash et al., 2017).

Generally, there are two approaches to implement the intrusion detection test:

1. Misuse detection: the function of this technique is detection of the existing attacks and definition of a pattern for analysis and search engine motor of a series of events corresponded with a predetermined pattern.

2. Anomaly detection: the function of this technique is detection of system's normal performance and indexing of the normal system behavior of analysis engine and search for abnormal activity (Gai et al., 2016).

Intrusion detection systems: these systems are created as hardware and software systems and each one has its own advantages and disadvantages applied by internal and external users. Speed and accuracy are the advantages of the hardware systems and the lack of security breach by

hackers is another ability of these systems. The detection methods used in intrusion detection systems are divided into two categories:

1) Method for detection of abnormal behavior
2) Signature-based detection method or detection of abuse (Joffe, 2016).

Method for detection of abnormal behavior: in this method, a view of normal behavior is created. An anomaly may indicate an intrusion. The approaches like neural networks, machine learning techniques and even biological immune systems are used in order to create the views of normal behavior. For detecting the abnormal behavior, the normal behaviors should be detected and their specific pattern should be found. The behaviors following these patterns are normal and the events deviating from these patterns more than the usual statistics are known as abnormal behavior. Unusual intrusions are difficult to detect because there is no stable pattern for monitoring. The event which is occurred very higher or lower than two standard deviations often is assumed abnormal. For example, a user logs in and logs out twenty times in a day- instead of usual one or two times, or a computer is used at midnight, while it was not supposed to be turned on after office hours. Each of these cases can be considered as an abnormal behavior. The techniques and criteria used in detecting the abnormal behavior are as followings:

- Detection of threshold level: the number of log in and out of system or the time of system usage are among the characteristics of system's or user's behavior which can indicate the abnormal behavior of system originating from an intrusion. This level is extremely stable and explorative (Kenkre et al., 2015).
- Statistical criteria: from the parametric aspect, all collected characteristics are considered based on a specific pattern and in a nonparametric state, they are compared according to the experimental values. NIDS can be named among the famous IDS used for statistical measuring to detect the intrusion (Gai et al., 2016).
- Legalistic criteria: it is like nonparametric statistical criteria, so that the observed data is defined acceptably based on the specified applied patterns. But it differs from patterns determined as a law and is not numerical (Eesa et al., 2015).
- Clustering methods: the cluster's aim of analysis is to divide the data into separate partitions. Its difference with classification is that the detection of new patterns is gained from the pre-defined patterns. However, clustering is discovering and documentation of set of unknown facts such as geographical location. Clustering consists of three stages of displaying the available data, determination and calculation of approximate data i.e. what is known as the similarity of objects (Joffe, 2016).

2. Architecture of intrusion detection systems

The various architectures of intrusion detection systems include:

- Host-based intrusion detection system (HIDS)
- Network-based intrusion detection system (NIDS)
- Distributed intrusion detection system (DIDS)

Host-based intrusion detection system (HIDS)

This system is responsible to detect and identify the illegal activities on host computer. It can detect the attacks and threats on critical systems (such as access to files, Trojan horses, etc.) which cannot be detected by network-based intrusion detection systems. HIDS protects only the hosts placed on it and their network interface card (NIC) works on a regular mode by default. The regular mode can be useful in some cases because all NICs do not have the capability of irregular mode. HIDSs are informed of all types of additional local information of the host which should be monitored with security implementations (including system calls, change system files, and system connections). This provides appropriate data at the time of combining with network communications (Kim et al., 2014; Devi et al., 2017).

Network-based intrusion detection system (NIDS)

The term NIDS originates from the fact that it monitors all over the network from where it is placed. The detection and identification of unauthorized intrusions before arriving to critical systems is the responsibility of NIDS. NIDSs usually consist of two sections of monitor (sensoe) and factor. These two sections often are installed behind the firewall and other access points to detect any unauthorized activity. The network factors can be alternative network infrastructure in order for searching the network traffic. The installation of network factors and monitors has the advantage that eliminates any attack at the initial stage. In addition, the sequences of investigating one or more hosts can be useful for searching the attack signs (Saraswati et al., 2016).

Distributed intrusion detection system (DIDS)

These systems consist of several NIDS or HIDS or a combination of both with a central management station. Each IDS that exist in the grid transmit its report to the central management center. Reviewing the received reports is the duty of the central station. The central station, also, has a duty to update the database of the detection rule of each IDS on the network. Figure 6 indicated a Distributed Intrusion Detection System. NIDS 1 and 2 protect Public servers and NIDS 3 and 4 are

responsible for the task of protecting the internal network. The network between NIDS and the central management system can be private or use the existing infrastructure for sending data. Additional security will achieve when we use the existing network for sending management data through encryption and virtual private network technology (VPN) which is most recommended. Converting the network traffic to a readable text and analyzing the network performance are possible to detect the bottlenecks. Due to the high volume of data in intrusion detection systems data mining algorithms have been used to extract the patterns and behavior categories of data. Particularly, by using the existing classification methods, we can achieve the graph model which is related to those characteristics. The data model of the data collection in intrusion detection system is flat. This means that a dataset includes a number of records which each record has different features. If the graph has been extracted from the flat model, then we can use the concepts of graph analysis. For example, K-Means algorithm which is a clustering algorithm can be very useful in intrusion detection systems and knowledge discovery. Here, the flat model turns to graph model. In the resulting graph model, features form the graph nodes, and communication method of features illustrates the edges in the graph.

Data Mining: it is equivalent to terms such as knowledge extraction, data collection, data verification, and even data dredging which actually describes knowledge discovery in databases. To put it simply, it means extracting knowledge from the enormous volume of data. Also, we can consider the knowledge discovery from data, knowledge extraction, pattern analysis, and data dredging as a synonym of data mining. Data mining is a section of the major knowledge discovery which has a lot of application in decision making.

Data mining includes a set of methods and technics to extract new, hidden and unexpected information or detecting a related pattern in a huge amount of raw data. Dividing the existing data into several groups is the goal of clustering. Data of different groups should be as different as possible together and data in the same group must be very similar. Unlike the classification, groups are not predetermined in clustering and it is not clear that according to which characteristics grouping is done. Therefore, after clustering, an expert should interpret the clusters. After examining the clusters, some of the parameters that are considered irrelevant should be deleted, and again clustering should be done. After that the data were divided into several reasonable and justifiable groups, we can use this classification to obtain information from data or do another classification. Kohnen algorithm and means-K algorithm are the most important algorithms, which can be used for clustering.

Conclusion

According to the enhancement of using the computers and detection of intrusion in the computer networks, the amount of security has been developed in the computer networks. It is necessary to consider to the security of network by means of different platforms. This method has some advantageous for the secure vulnerable in all of the computer systems. Therefore, intrusion is defined as compatible, accessible and valid key in the computer sources. Intrusion discovery systems (IDSs) play a vital role in in discovering the anomalies and attacks in the network. Results show that data mining techniques and usage of its algorithms for detecting the hidden and related data are effective in less running time. Moreover, the main four elements of data classification, high level of human interaction, lack of labeled data, and effectiveness of service attack rejection using data mining algorithms that most important one is K-Means (a clustering algorithm) can be very useful in intrusion detection systems and knowledge mining and discovery, and flat model turns into graph model.

References

[1] Adebowale, A., Olutayo, A., Sunday, I., Ogbonna, A. C., & Oluwabukola, O. (2017). Scalable Unsupervised Ensemble Algorithm For Effective Insider Threat Detection. Universal Journal of computer and Technology (UJCT), 1(3), 76-83.

[2] Ashfaq, R. A. R., Wang, X. Z., Huang, J. Z., Abbas, H., & He, Y. L. (2017). Fuzziness based semi-supervised learning approach for intrusion detection system. Information Sciences, 378, 484-497.

[3] Chand, N., Mishra, P., Krishna, C. R., Pilli, E. S., & Govil, M. C. (2016, April). A comparative analysis of SVM and its stacking with other classification algorithm for intrusion detection. In Advances in Computing, Communication, & Automation (ICACCA)(Spring), International Conference on (pp. 1-6). IEEE.

[4] Costa, K. A., Pereira, L. A., Nakamura, R. Y., Pereira, C. R., Papa, J. P., & Falcão, A. X. (2015). A nature-inspired approach to speed up optimum-path forest clustering and its application to intrusion detection in computer networks. Information Sciences, 294, 95-108.

[5] Dash, S., Mishra, R. K., Das, R. K., & Panda, M. (2017). Comparison of AIS based Data Mining Algorithms for Intrusion Detection. International Journal of Computer Science and Information Security, 15(1), 619.

[6] Dash, S., Mishra, R. K., Das, R. K., & Panda, M. (2017). Comparison of AIS based Data Mining Algorithms for Intrusion Detection. International Journal of Computer Science and Information Security, 15(1), 619.

[7] Devi, R., Jha, R. K., Gupta, A., Jain, S., & Kumar, P. (2017). Implementation of Intrusion Detection System using Adaptive Neuro-Fuzzy Inference System for 5G wireless communication network. AEU-International Journal of Electronics and Communications, 74, 94-106.

[8] Eesa, A. S., Orman, Z., & Brifcani, A. M. A. (2015). A novel feature-selection approach based on the cuttlefish

optimization algorithm for intrusion detection systems. Expert Systems with Applications, 42(5), 2670-2679.
[9] Elhag, S., Fernández, A., Bawakid, A., Alshomrani, S., & Herrera, F. (2015). On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems. Expert Systems with Applications, 42(1), 193-202.
[10] Gai, K., Qiu, M., Tao, L., & Zhu, Y. (2016). Intrusion detection techniques for mobile cloud computing in heterogeneous 5G. Security and Communication Networks, 9(16), 3049-3058.
[11] Gaidhane, R., Vaidya, C., & Raghuwanshi, M. (2014). Survey: Learning Techniques for Intrusion Detection System (IDS).
[12] Gautam, S. K., & Om, H. (2017). Comparative Analysis of Classification Techniques in Network Based Intrusion Detection Systems. In Proceedings of the First International Conference on Intelligent Computing and Communication (pp. 591-601). Springer Singapore.
[13] Gupta, D., Singhal, S., Malik, S., & Singh, A. (2016, May). Network intrusion detection system using various data mining techniques. In Research Advances in Integrated Navigation Systems (RAINS), International Conference on (pp. 1-6). IEEE.
[14] Joffe, R. L. (2016). U.S. Patent No. 9,356,942. Washington, DC: U.S. Patent and Trademark Office.
[15] Kenkre, P. S., Pai, A., & Colaco, L. (2015). Real time intrusion detection and prevention system. In Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014 (pp. 405-411). Springer International Publishing.
[16] Kevric, J., Jukic, S., & Subasi, A. (2016). An effective combining classifier approach using tree algorithms for network intrusion detection. Neural Computing and Applications, 1-8.
[17] Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.
[18] Levitt, Karl, and Gihan Dias. "Towards Detecting Intrusions in a Networked Environment." (2017).
[19] Nair, R., Nayak, C., Watkins, L., Fairbanks, K. D., Memon, K., Wang, P., & Robinson, W. H. (2017). The Resource Usage Viewpoint of Industrial Control System Security: An Inference-Based Intrusion Detection System. In Cybersecurity for Industry 4.0 (pp. 195-223). Springer International Publishing.
[20] Petersen, R. (2015). Data Mining for Network Intrusion Detection: A comparison of data mining algorithms and an analysis of relevant features for detecting cyber-attacks.
[21] Saraswati, A., Hagenbuchner, M., & Zhou, Z. Q. (2016, December). High Resolution SOM Approach to Improving Anomaly Detection in Intrusion Detection Systems. In Australasian Joint Conference on Artificial Intelligence (pp. 191-199). Springer International Publishing.
[22] Saraswati, A., Hagenbuchner, M., & Zhou, Z. Q. (2016, December). High Resolution SOM Approach to Improving Anomaly Detection in Intrusion Detection Systems. In Australasian Joint Conference on Artificial Intelligence (pp. 191-199). Springer International Publishing.
[23] Stolfo, S. J., Malkin, T., Keromytis, A. D., Misra, V., Locasto, M., & Parekh, J. (2015). U.S. Patent Application No. 14/846,188.