International Journal of Innovation in Computer Science and Information Technology
Vol.3, NO.4 , P: 82 - 90
Received : 13 August 2021
Accepted : 12 October 2021

Architecture of Cyber Assessment within the Organization

Mohammad Nakhaei*
Master of Information Technology Engineering, Department of Computer, Zanjan Branch, Islamic Azad University, Zanjan, Iran
Modernm24@gmail.com
Dr. Naser Modiri
Faculty of Computer Engineering, Faculty of Engineering and Basic Sciences, Zanjan Branch, Islamic Azad University, Zanjan, Iran
Nassermodiri@yahoo.com

Abstract

Nowadays the widespread use of computer networks has made things easier for most people in most governmental and non-governmental organizations and institutions, on the other hand computer networks development has created dangers and threats, and everyone has access to the Internet, which is a very large database of information. Ignoring these threats and security problems has sometimes lead to a lot of damage, such as the loss of valuable information of owners and customers. Traditionally, techniques such as user authentication, data encryption, malware detection systems, and firewalls have been used to protect computer security, but today, cyber-attack detection systems have become very common to prevent such threats. The purpose of cyber-attack detection systems is to detect various types of malicious traffic of networks and computers that are not detected by firewalls.

Key words: Cyber, intrusion, firewall, defense, attack

Introduction

If we want to compare firewall systems and cyber-attack detection with anti-theft systems, firewall takes the role of door and window locks. These types of locks prevent some theft, but experienced thieves can bypass these locks and break into the house.

Therefore, a combination of advanced locks and warning systems are used in most cases. Cyber-attack detection systems act as warning systems, as if they add another layer of security protection to the network. In this case, the cyber-attack detection system has to check the traffic of the whole network and then deliver the filtered traffic to the firewall. In this case, the cyber-attack detection system should check the packages very quickly and not cause a bottleneck. In other cases, the cyber-attack detection system only checks part of the traffic and does not need to monitor the entire network. In terms of overall network security, the first case is the safest, because the traffic is checked and filtered before reaching the firewall, and the attacker can’t cause the firewall to crash.

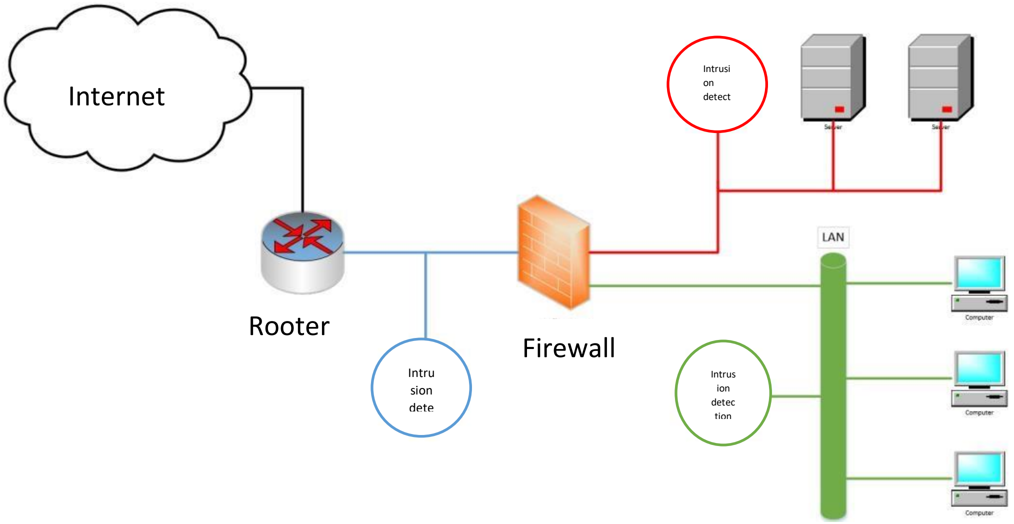

Figure 1 shows the location of the cyber-attack detection system in the network.

As can be seen in the figure, the cyber-attack detection system can be located in different places. In the first case, the cyber-attack detection system can be placed in front of the firewall.

Cyber-attack detection systems use special analytical methods to detect attacks, identify their sources, and alert the network administrator, and their overall purpose is to monitor attempts to breach security. There are still many opportunities to detect and neutralize attacks despite significant progress and much work in this area. [1]

Security breaches include external intrusions, i.e. attacks carried out from outside the organization, and internal intrusions, i.e. attacks within the organization. Attacks on information systems are divided into several groups as follows:

1. DoS: The attacker tries to prevent users from legally accessing information systems in such attacks. In a DDoS attack the attacker tries to prevent users from legally accessing information systems through distributed systems.

2. R2L: The attacker tries to access the victim's system without having an account in these attacks, such as repeatedly guessing the password.

3. U2R: The attacker has access to the victim's system but tries to score great user points, such as a buffer overflow.

4. Probe: The attacker tries to get information from his target host [2] such as using the nmap program

Researchers have made great strides in developing robust cyber-attack detection systems. Unfortunately, in most cases these systems can't work well against attackers who are always trying to carry out new attacks with different distributions [3].

D) The importance and necessity of conducting research (including differences and existing research gaps, the need for the subject, its possible theoretical and practical benefits, as well as possibly new research materials, methods or processes) used in this research:

Vital national infrastructure is always vulnerable to cyber-attacks. Protecting them is important for any organization. Cyber defense plays a very important role in vital national infrastructure. The key to increasing cyber security is reducing the level of vulnerability. Although threat awareness is important all attacks become more severe as vulnerability declines. [4]

Cyber Security Lifecycle Like other IT processes, cybersecurity often follows a life cycle model of predicting, protecting, detecting, and responding.

Threats to computer security are growing rapidly. However, phishing and malware are known to be the most important threats. The purpose of phishing is to obtain information from fraudulent individuals or entities. Attackers design web pages similar to real web pages and direct users to these pages through ways such as clicking on links and using other phishing methods to get what they want [5].

According to the annual publication of the Dutch National Cyber Security Center, the current state of cyber security in the country, cyber-attacks have been phishing in 91% of the cases. In this regard, SANS data show that 95% of all attacks on corporate networks are the result of successful phishing. Governmental organizations invest millions in protecting their internal systems and infrastructure, but they train their employees on mobile security in a short, low-cost period, and this shows that large investments are made in defensive technologies but there is little investment in human education [6]. Malware includes the three main groups of virus, Trojan horse and worm.

Another major concern about cybersecurity is spam, which is defined as unsolicited e-mail messages. In addition to reading these messages is time consuming, a malicious program may run automatically when reading the message.

How seriously security issues should be taken. The following are examples of security threats that will be examined. According to a report by Troy Hunt, a US website security expert and regional director at Microsoft [7] About 773 million emails are leaked. This intrusion is due to the violation of people's passwords and is not due to the intrusion and hacking of a large database. If you do not use a secure password method and manage all the information online using a password, you may have a security breach and should take this threat seriously.

Also [8] two white hat hackers discovered while browsing websites that there are many websites that are used to deceive victims, including journalists to gather information from them and hack their computers in the future. These websites have been able to access the personal information of Israeli officials.

In recent years, countries such as Russia have somehow pervaded different countries by spreading false information in different societies. This influence has progressed to such an extent that it has interfered in the elections of some countries and has caused some problems for democracy. To counter this type of interference, an association named Integrity Initiative has started working to be effective in providing security in 2018 in European countries [9].

Redelinghuys et al. (2018) proposed the use of neural networks and support vector machines to detect cyber-attacks. In this experiment, both the support vector machine and the neural network were trained with normal data and attack patterns. The data used in this paper is DARPA, compiled for KDD competitions by MIT's Linkton Laboratory. The way through which these algorithms work is that they learn normal behavior, then if they see behavior that is very different from normal behavior they mark it as an attack. Finally, the performance of these two algorithms are compared with each other [10].

Murray et al. (2019) presented a framework that uses the Bayesian network to detect adaptive cyber-attack. The Bayesian network is known as a graphical modeling tool and is used to model problems where there is uncertainty. In this model, known signature attacks are taught through the Bayesian network, and if the new behavior observed is in accordance with those known signatures, it is introduced as an attack. The main challenge in this system is the change of signatures over time so the system must be retrained. This method is called the adaptive model because of adding new signatures to the signature database and re-training [11].

In the disorder-based cyber-attack detection method proposed by Ma et al. (2019), each data is assigned to a cluster if its distance from that cluster is less than a fixed value. The input of this algorithm is an unlabeled dataset. This algorithm tries to identify malicious packets from the data it receives. This unsupervised disorder-based method of detection considers two assumptions about the data. The first assumption is that the number of normal data is significantly greater than the number of malicious data. The second assumption is that the malicious data themselves are qualitatively different from normal data. The simple basic idea is that since malicious data is small and different from normal data, it appears as out-of-bounds data and is thus identifiable. The problems of this proposed method include the difficulty of finding the mentioned constant value, low detection rate and inaccuracy of the second hypothesis. The assumption that malicious data is always different from normal data is not always true. This is because the attacker is constantly trying to bring his attack pattern as close to normal data as possible, thus attacks are less likely to detect [12].

Another disorder-based model was proposed by Ahmad et al. (2019). In this method, first, using genetic algorithm, appropriate features are extracted from the packets. The modified support vector machine algorithm is then applied for classification. Modified support vector machine is made of both supervised and unsupervised support vector machines (one-class support vector machines). In this context, the result of using a support vector machine with monitoring is high performance, and the use of one-class support vector machine makes it possible to detect new attacks. In this method, there is also a part for data processing in which packets are filtered before being delivered to the support vector machine. This model has been compared with cyber-attack detection systems of real world networks and has a better performance than the systems available at that time [13].

In the method proposed by Villalonga et al. (2020), the combination of SNORT is used as the signature-based part, and NETAD and PHAD are used for the disorder-based part. SNORT is an open source network management system that can be used in a variety of ways: detecting cyberattacks, eavesdropping and recording exchanged packets. No creative method is used in this system [14].

Sanislav et al. (2017) proposed a system using multilayer neural networks, radial basis performance networks, and collective learning. In this system, if one of the methods has a positive answer and the other has a negative answer about the destructiveness of the packet, a decision is made on which class will be selected as the final answer regarding to the weight of each method. The proposed model has a better result than using either of the two algorithms alone. In addition, the multilevel neural network collective learning function is better than the radial base function for detecting normal behavior, but the opposite is true for detecting attacks. A simple forward neural network is called a multilayer network. A multilayer neural network is a network of nodes used for classification. In each layer there are neurons whose output is always forward. If there is only one layer, it is called perceptron. Multilayer networks use a variety of learning techniques, the most popular of which is backwards. In this case, the network output is compared with the correct answer to calculate the error size. Then the error is returned with different techniques throughout the network. Using this algorithm, the weight of each edge is determined in such a way that the error is minimized [15].

Gong et al. (2021) used a self-organizing map for the disorder-based part of their proposed hybrid model. For the signature-based part, the J.48 decision tree was used to categorize destructive behaviors [16].

The self-organizing map is a neural network proposed by Alromaihi et al. (2018) to analyze and visualize large-sized data. The self-organizing map is based on unsupervised learning that maps nonlinear statistical relationships between large-scale inputs to a two-dimensional network. This two-dimensional network is called the output space. The self-organizing map efficiently places similar patterns in adjacent locations in the output space, and provides high-dimensional data visualization. The simulation results on KDD99 data show that the result of the hybrid solution is better than using each of the methods independently [17].

In the hybrid method presented by Akyildiz et al. (2019) signature-based detection is applied after disorder-based detection. In the disorder-based part, an artificial immune system is used for detection. Behaviors that are known to be destructive are then categorized using a self-organizing map. That is, it is determined what class each attack belongs to. The main purpose of this system of applying disorder-based methods is to identify new attacks. During learning, both parts of the system are taught separately. In learning the artificial immune system only normal samples are used, and in learning self-organizing maps only attack samples are used. The reason for this is that the task of the self-organizing map is to cluster similar attacks here, and as a result to extract similar features of attacks belonging to a group. This method claims to have a low false alarm and a high detection rate for DOS and U [1] R attacks. However, due to the fact that the self-organizing map is used only to categorize attacks, so it will not have an effect on improving system performance [18].

Mendhurwar et al. (2019) presented a hybrid model in which after signature-based detection, the disorder-based detection method is used. In the signature-based part, random forests are applied and the out-of-range data detection algorithm is used with the help of random forests in the disorder-based part. Random forest is a set of unpruned trees. This algorithm is more accurate than other machine learning algorithms, especially for large data sets. The random forest produces a large number of classification trees, and each tree is made with random data with a different alternative to the original data. After the forest is created, a new data to be classified is placed in all the trees in the forest and goes through the tree to the end. Each tree makes a decision about its data class, and finally declares the forest of the class with the most votes as that data class. Since in a random forest algorithm, each tree is made using bootstrap samples, there is no need for cross-validation. Experiments on this method indicate that it performs better than using only one of the disorder-based or signature-based parts [19].

Meira et al. (2020) proposed a hybrid cyber-attack detection method based on two well-known machine learning algorithms, random forests and -kmeans. In this method, after the signature-based detection, a disorder-based detection is applied. In the signature-based part network communications are divided into two categories of normal and attack using the random forest based on the training set label. In the disorder-based part of the k-means algorithm, it divides the network data into K clusters based on the similarity of their properties. Some of these groups are known as attacks [20].

In the method presented by Cimini et al. (2020), signature-based detection is made first, followed by disorder-based detection. In the signature-based method, the decision tree is used. In the learning process, this tree is created using educational data. Then there is the disorder-based part, in which the support vector machine is used. The support vector machine algorithm is applied to all leaves of the tree that have a normal label. Thus, a range is considered for each subset of normal data. When new data is entered into the system for classification, it first goes through the tree; if this tree labels an attack on the data, the cyber-attack detection system sends an alert to the system administrator. If the tree recognizes the data label as normal, the data is compared with the range specified by the support vector machine for the same sheet in which it is located. If this data is seen in the normal data range, the data label is normal, otherwise the attack is unknown. In this model, one-class support vector machine is used. One-class support vector machine takes only the data associated with a label as input and specifies its range. Experiments on this model have shown that it performs better than applying either method alone, or applying both methods in parallel [21].

To improve productivity, Redelinghuys et al. (2018) proposed a way to separate important data in which the main data set is processed before being given to the classification algorithm. This process is independent of the learning algorithm used for the cyber-attack detection system. In this research, a new criterion is introduced that can help to find the ones from the whole data that makes it possible to categorize the packets better with the help of that data. Experiments on labeled data show that by reducing the number of samples and using the support vector machine classification algorithms, the nearest neighbor, K, results in better detection of network attacks, detection rate and accuracy compared to the state in which these algorithms are constructed using the original data set [22].

Intrusion

Intrusion [1] refers to an operation that attempts to bypass the system security mechanism to gain unauthorized access to a network or computer system. This operation is carried out by foreign and domestic intruders.

Cyber-attack Detection Systems

A cyber-attack detection system is a program that tries to identify intruder activities by analyzing the current network traffic or demand analysis, and if it detects incoming traffic to a network or machine is not allowed by the user but arises from the activities of an intruder, appropriately warns or reacts to the network administrator.

Data Mining

Data mining is the process of finding knowledge of large amounts of data stored in databases, data warehouses or other repositories.

Rule-based Model

The rule-based model is a type of supervised learning that results in ***if-then*** rules. The section after ***if*** specifies conditions and ***then*** section specifies the final answer.

Conclusion

The main goal of this project is to evaluate the architectural layers of the organization's cyber evaluation that can distinguish normal packers from abnormal ones. The main innovation in the project is the use of lazy model algorithms and rule-based model, which has not been used for cyber-attack detection systems so far and the use of all the available algorithms in the classification methods in WEKA and Rapidminer software is present and the extraction of 5 data samples from the raw data that are the best answer for different models and related algorithms. Extraction of 5 data samples took a lot of time and all the different algorithms in the classification models were simulated and implemented with different data sets, and finally we proposed 5 initial data samples. Finding the best data set requires repeated tests of each algorithm with different datasets, modeling and evaluation, which finally succeed in providing 5 different data samples in terms of differences in the type of attributes that provide the best answer for the algorithms.

Resources

1. Sharpe, R., Van Lopik, K., Neal, A., Goodall, P., Conway, P. P., & West, A. A. (2019). An industrial evaluation of an Industry 4.0 reference architecture demonstrating the need for the inclusion of security and human components. Computers in Industry, 108, 37-44.
2. Salazar, L. A. C., Ryashentseva, D., Lüder, A., & Vogel-Heuser, B. (2019). Cyber-physical production systems architecture based on multi-agent's design pattern—comparison of selected approaches mapping four agent patterns. The International Journal of Advanced Manufacturing Technology, 105(9), 4005-4034.
3. Augusto-Gonzalez, J., Collen, A., Evangelatos, S., Anagnostopoulos, M., Spathoulas, G., Giannoutakis, K. M., ... & Nijdam, N. A. (2019, September). From internet of threats to internet of things: A cyber security architecture for smart homes. In 2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) (pp. 1-6). IEEE.
4. Givehchi, O., Landsdorf, K., Simoens, P., & Colombo, A. W. (2017). Interoperability for industrial cyber-physical systems: An approach for legacy systems. IEEE Transactions on Industrial Informatics, 13(6), 3370-3378.
5. Aazam, M., Zeadally, S., & Harras, K. A. (2018). Fog computing architecture, evaluation, and future research directions. IEEE Communications Magazine, 56(5), 46-52.
6. Sengan, S., Subramaniyaswamy, V., Nair, S. K., Indragandhi, V., Manikandan, J., & Ravi, L. (2020). Enhancing cyber–physical systems with hybrid smart city cyber security architecture for secure public data-smart network. Future generation computer systems, 112, 724-737.
7. Alguliyev, R., Imamverdiyev, Y., & Sukhostat, L. (2018). Cyber-physical systems and their security issues. Computers in Industry, 100, 212-223.

8. Hofer, F. (2018, October). Architecture, technologies and challenges for cyber-physical systems in industry 4.0: A systematic mapping study. In Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (pp. 1-10).

9. Jin, D., Li, Z., Hannon, C., Chen, C., Wang, J., Shahidehpour, M., & Lee, C. W. (2017). Toward a cyber-resilient and secure microgrid using software-defined networking. IEEE Transactions on Smart Grid, 8(5), 2494-2504.

10. Redelinghuys, A. J. H., Basson, A. H., & Kruger, K. (2019). A six-layer architecture for the digital twin: a manufacturing case study implementation. Journal of Intelligent Manufacturing, 1-20.

11. Murray, L., Budenske, J., Gangopadhyay, S., & Finstad, R. K. (2018, May). Cyber resilience and integrity self-awareness of mobile autonomous systems. In Open Architecture/Open Business Model Net-Centric Systems and Defense Transformation 2018 (Vol. 10651, p. 106510E). International Society for Optics and Photonics.

12. Ma, S., Zhang, Y., Lv, J., Yang, H., & Wu, J. (2019). Energy-cyber-physical system enabled management for energy-intensive manufacturing industries. Journal of cleaner production, 226, 892-903.

13. Ahmad, A., Babar, M., Din, S., Khalid, S., Ullah, M. M., Paul, A., ... & Min-Allah, N. (2019). Socio-cyber network: The potential of cyber-physical system to define human behaviors using big data analytics. Future Generation Computer Systems, 92, 868-878.

14. Villalonga, A., Beruvides, G., Castaño, F., & Haber, R. E. (2020). Cloud-based industrial cyber–physical system for data-driven reasoning: A review and use case on an industry 4.0 pilot line. IEEE Transactions on Industrial Informatics, 16(9), 5975-5984.

15. Sanislav, T., Zeadally, S., & Mois, G. D. (2017). A cloud-integrated, multilayered, agent-based cyber-physical system architecture. Computer, 50(4), 27-37.

16. Gong, S., & Lee, C. (2021). Cyber Threat Intelligence Framework for Incident Response in an Energy Cloud Platform. Electronics, 10(3), 239.

17. Alromaihi, S., Elmedany, W., & Balakrishna, C. (2018, August). Cyber security challenges of deploying IoT in smart cities for healthcare applications. In 2018 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW) (pp. 140-145). IEEE.

18. Akyildiz, I. F., & Kak, A. (2019). The Internet of Space Things/CubeSats: A ubiquitous cyber-physical system for the connected world. Computer Networks, 150, 134-149.

19. Mendhurwar, S., & Mishra, R. (2019). Integration of social and IoT technologies: architectural framework for digital transformation and cyber security challenges. Enterprise Information Systems, 1-20.

20. Meira, J., Andrade, R., Praça, I., Carneiro, J., Bolón-Canedo, V., Alonso-Betanzos, A., & Marreiros, G. (2020). Performance evaluation of unsupervised techniques in cyber-attack anomaly detection. Journal of Ambient Intelligence and Humanized Computing, 11(11), 4477-4489.

21. Cimini, C., Pirola, F., Pinto, R., & Cavalieri, S. (2020). A human-in-the-loop manufacturing control architecture for the next generation of production systems. Journal of manufacturing systems, 54, 258-271.

22. Redelinghuys, A., Basson, A., & Kruger, K. (2018, June). A six-layer digital

twin architecture for a manufacturing cell. In International Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing (pp. 412-423). Springer, Cham.

23. Vigneswaran, R. K., Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2018, July). Evaluating shallow and deep neural networks for network intrusion detection systems in cyber security. In 2018 9th International conference on computing, communication and networking technologies (ICCCNT) (pp. 1-6). IEEE.