Journal of Engineering and Applied Sciences 12 (16): 4241-4245, 2017
ISSN: 1816-949X

Intrusion Detection System Based on Machine Learning in Cloud Computing

[1]Mohammed Hasan Ali, [1]Mohamad Fadli Zolkipli, [2]Mustafa Musa Jaber and
[3]Mohammed Abdulameer Mohammed
[1]Faculty of Computer Systems and Software Engineering,
Universiti Malaysia Pahang, Pahang, Malaysia
[2]Nabu Research Academy, Selangor, Malaysia
[3]Universiti Utara Malaysia, Kedah, Malaysia

Abstract: Detection of attacks in the computers and networks continues to be a relevant and challenging area of researchers. Intrusion-detection system is an essential technology in network security. Currently, intrusion detection still faces some challenges like large amounts of data to process, low detection rates and high rates of false alarms, especially in cloud environment which more vulnerable to attacks. This study includes an overview of intrusion-detection systems and introduces the reader to some fundamental concepts of IDS methodology to work in cloud computing also discuss the primary intrusion-detection techniques and propose a new classifier algorithm fast learning network to work based on the intrusion-detection system.

Key words: ELM, SVM, ANN, IDS, FLN, network

INTRODUCTION

In last year's cloud computing has rapidly emerged widely in computing system because it's provide many services for the end user and its can reach these resources across networks anytime anywhere. Pew Institute published a research about, "the future of cloud computing" and depicted that about 71% of technology stakeholders and critics believe that by the year 2020, most people will work in internet-based applications. Therefore, it can be seen that the future of cloud computing technology is bright and will be widely used in the world (Oktay and Sahingoz, 2013).

The new security challenges emerged during moving from local computing paradigm to cloud computing paradigm because of the distributed nature of cloud computing, many researchers mention the security is biggest challenge (Kim and Ahn, 2011; Mehmood *et al*., 2013) because cloud computing have to provide a high quality service and protect the data. Cloud computing suffers from form different attacks such as Denial of Service (DOS), Distributed Denial of Service (DDOS), flooding, etc. There are many tools as firewall and Intrusion Detection System (IDS) are effective solutions to resist them (Modi *et al*., 2013).

In this study, it is aimed to introduce the intrusion detection models and why the normal (host, network) IDS is not suitable for cloud computing. Also aimed for introduce some of the machine learning and artificial intelligence algorithms that used to builds effective intrusion detection.

MATERIALS AND METHODS

Cloud computing: Cloud computing is known as a new style of computing that virtualized resources are provided. In cloud computing users use a variety of devices like, laptop, smartphone. Cloud provides services to its users in different ways, Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) these different way to services provide to the users control the complete virtual machines and also users can create applications in cloud or execute provided via internet (Armbrust *et al*., 2009). Cloud computing has rapidly emerged as a widely accepted paradigm in computing systems in which users can request some computing capabilities and services and he can reach these resources across networks anytime, anywhere. Cloud computing refers to the provision of resources on demand via a computer network as in Fig. 1.

The open services and structure of cloud computing, it involves multi domains, multi tenancies and multi mesh distributed which are more vulnerable and prone to security risks (Patel *et al*., 2013). The traditional network security measures like firewall is good to stop the outside attacks but the attacks from inside the network as well as

Corresponding Author: Mohammed Hasan Ali, Faculty of Computer Systems and Software Engineering,
Universiti Malaysia Pahang, Pahang, Malaysia

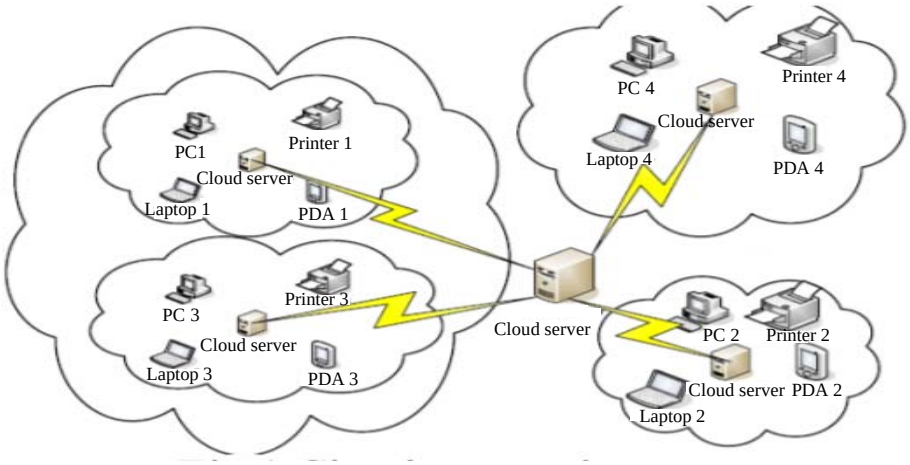

Fig. 1: Provision of resources in cloud computing (Madhavi, 2012)

some complicated outside attacks cannot be tackled effective by using such mechanisms Also cloud computing more attractive for potential intruders because of the fully distributed and open structure and the problem becomes more critical when a cloud is abused by an insider intruder (Patel *et al*., 2013) so to mitigate the risk of these attacks should be incorporated the IDS in the infrastructure of the cloud computing (Modi *et al*., 2013).

RESULTS AND DISCUSSION

Intrusion detection system: Intrusion detection is the process of monitoring computers or networks for illegal entry, activity or file modification (Whitman and Mattord, 2012; Modi *et al*., 2013). As shown in Fig. 2 is the infrastructure of the IDS, shown how the IDS monitoring the network and in same time connected to network admin to send alarms if detect any incidents happen.

ID is the identification of attempted or ongoing attacks on computer system or network. Issues in ID research include data collection, behavior classification, data reduction, reporting and response (Frank, 1994). The quality of the feature construction and feature selection algorithm is one of the most important factors that affect the effectiveness of the IDS (Nguyen *et al*., 2012). Achieving a reduction of the number of relevant traffic features without negative effect on classification accuracy is a goal that largely improves the overall effectiveness of the IDS (Kukar, 2012). Many techniques have been used in many IDSs to perform these important tasks like Artificial Intelligence algorithms, Machine learning methods and Hybrid techniques that use the combination two or more of techniques. It is advantageous, since each

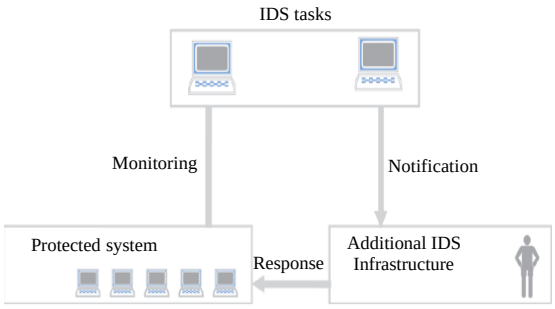

Fig. 2: Intrusion detection system infrastructures

technique has some advantages and drawbacks (Modi *et al*., 2013). IDS monitor the activities of the user and track the network traffic then it determines whether the activity is malicious or not if the activity is malicious IDS generates an alarm. IDSs use various techniques like anomalies or signatures of attack to detect the attack and success of IDS also depends upon these techniques (Kumar and Gohil, 2015; Bhuyan *et al*., 2014). Today, researchers mostly concentrate on anomaly based network intrusion detection because it can detect known as well as unknown attacks. According to the placement of IDS in the network, the normal IDS are two types: Host and Network based on the placed of the IDS. Many of examples of IDS have been applied to a cloud environment by several works as in Table 1.

Some researchers mention some limitations in the traditional IDS to work in the cloud environment. Baysa talk about the modern malwares can easily thwart traditional HIDS by using obfuscation and encryption techniques (Baysa *et al*., 2013). When (Mishra *et al*., 2016) published a survey about IDS in cloud computing,

Table 1: Several works based IDS

Researchers	Description
Chouhan (2015)	This research explained how can use data mining approaches as both anomaly and misuse detection and most of the present techniques focus on anomaly detection systems. Also they observe that most techniques have been validated using the KDD99
Madhavi (2012)	This research explain the features of IDS that provide a view of unusual activities and how they IDS becomes crucial part in cloud computing
Raghav (2013)	In this work researchers mention because of distributed structure of cloud, the traditional IDS is not flexible for providing security and they showed how the security concerns in cloud are the main obstacles in cloud adoption, especially DOS and DDOS attacks. In the end they proposed a hybrid mechanism to enhance the detection rate
Shelke *et al.* (2012)	In this work the researchers, explain the different between the traditional IDS and multi-threaded NIDS that they proposed because the traditional IDS has problem about deal with big amount of data like in cloud environment
Vieira *et al.* (2010)	They have proposed an IDS service at cloud middleware layer which has an audit system designed to cover attacks that NIDS and HIDS cannot detect. The architecture of IDS service includes the node, service, event auditor and storage

they mention the traditional HIDS and NIDS designed to work on normal environment with normal flow of data and hence are limited in capability to detect attacks in virtualized environment and they proposed some future challenges like the placed of IDS in cloud computing. Signature matching techniques require regular maintenance as small variations in attacks in attack patterns can evade the security tool. Hence, the next level of defense integrates the conventional NIDS tool with conventional anomaly detection approach based on machine learning (Modi *et al*., 2013). Because the normal anomaly activated after the snort which make its only forwarded the benign traffic that passes from the snort and also its cannot verification of alerts of snort but only its can analysis for the benign traffic which can improve performance but does not help to reduce false alarms (Mishra *et al*., 2016) and mishar mention some IDS based machine learning like. Based on that our paper proposes a new NIDS based on machine learning algorithms to improve the IDS work. And mishar mention some IDS based machine learning algorithms such as Back Propagation and Artificial Neural Network (BP, ANN) and the disadvantage of these algorithms but they didn't talk about a new algorithms form machine learning like Extreme Learning Machine (ELM) (Huang *et al*., 2004) and Fast Learning Network (FLN) (Li *et al*., 2014).

Machine learning: There are many types of machine learning methods have been utilized in IDS to discover anomalies in system by classified the behavior by learning the normal network traffic and detecting traffic that differs from the normal pattern based on established metric (Fossaceca *et al*., 2011). The reason of using of Artificial Neural Network (ANN) for IDS because it's able to classify data as be normal or not and also able to generalization data from incomplete data (Modi *et al*., 2013).

Machine learning methods like SVM, ANN widely used for IDS but these methods generally suffer from some points very important for IDS, like the long training time, require parameter training or do not perform well in multi-class classification. Cannady (1998) proposed misuse detection based on a three layers of ANN but the IDS accuracy low. Gong *et al*. (2005) proposed also IDS based on Multi-Layer Perceptron (MLP) of ANN and in the end they showed that more hidden layers increase detection accuracy of IDS. The learning process relatively slow in ANN because the algorithm training samples would be used to define the free parameters of ANN which make ANN take much more time to training (Kiranyaz *et al*., 2009). However, IDS based ANN requires more training samples and time for effective learning (Modi *et al*., 2013). In Huang *et al*. (2004) based on Single hidden Layer Feed Forward Neural Network (SLFN) proposed a new artificial neural network classifier knows Extreme Learning Network (ELM), it's avoids several ANN disadvantages when worked based IDS. Jaiganesh *et al*. (2013) proposed a survey about classification techniques based IDS, they select the most popular classification algorithm, Support Vector Machine (SVM) and ELM in the end they found both ANN and SVM face some limitations like slow learning speed, poor scalability when compare with ELM which provides good generalization performance and faster learning speed.

There are several researchers of IDS based on the ELM (Fossaceca *et al*., 2011) tried to explore the efficacy of combining the learning decisions of multiple classifiers to formulate a single decision that is more accurate than any of the individual classifiers. The motivation for using an ensemble of classifiers is that prior research demonstrated that individual classifiers have varied ability to detect specific classes in a multi-class learning problem. By introduced the novel Multiple Adaptive Reduced Kernel Extreme Learning Machine (MARK-ELM) to work based IDS, MARK-ELM suitable for processing multi-class NIDS. Approaches have displayed good detection performance for some classes of attack but poor performance for others because the It's depends on unbalance dataset (KDD99). The proposed approach showed good detection performance with a high rate of

false positives which is huge challenge for network operators. Also the researchers during testing mode didn't depend on the data set testing mode to evaluate the results.

Raman Singh and Harish Kumar research explained that the common challenges for IDSs are large amounts of data to process, low detection rates and high rates of false alarms. They used the online sequential extreme learning machine to design the IDS based anomaly by analyzing the network traffic. For performance evaluation proposed technique the standard NSL-KDD and Kyoto university benchmark datasets are used to test the proposed IDS. Both datasets feature that used in this research extracted from KDD Cup 99 data set. So the algorithm has not been validated on large data set such as KDD 99 and further validation has to be performed. Concept of data profiling that used in this research is heuristic and not model based, heuristic is vague and general because it's created by the designer based on observed patterns and for example, used the heuristics in detection when you might not be sure, a virus is there but you can look for specific key attributes of a virus. Heuristic is an approach to problem solving, learning or discovery that employs a practical method not guaranteed to be optimal or perfect. I Proposed a primitive but without giving details and analysis for the current solutions that provide same hybrid in other fields and didn't mention the ELM limitations (Ali and Zolkipli, 2016).

On another hand, ELM still has some insufficiencies ELM tends more hidden neurons than conventional tuning-based learning algorithms in many applications which would make a trained ELM need longer time for responding to unknown testing samples (Huang *et al*., 2004). In, Li *et al*. (2014) proposed a new algorithm fast learning network (FLN) based on thought of ELM. The FLN is a Double Parallel Forward Neural Network (DPFNN) (Wang *et al*., 2011) which is essentially a parallel connection of a multilayer FFNN and an SLFN. The re-coded external information from the hidden nodes, along with the external information itself directly from the input nodes is fed into the output nodes of the DFNN's. Input weights as well as hidden layer biases are generated in a random manner for FLN's but where an analytical approach based on a least squares method is used to determine the weights of values for the connection between the output layer and the input layer and the weights of values for connecting the output node and the input. If a comparison is made between relating methods FLN is capable of reaching a good general high speed performance with impressive stability in most scenarios, whilst running with a smaller number of hidden units. With all this benefits of FLN this study proposes to design IDS based on FLN.

CONCLUSION

This research motivated by the shortcomings of prior approaches of intrusion-detection system work. In this study, we propose a new machine learning algorithm FLN to work based on IDS to approve the speed and the accuracy to reduce the impact of false alarms rate. Which still a big challenge of an intrusion-detection system for the future work there are some limitations of the current intrusion detection and the fast learning network algorithm, the unbalance data set of the intrusion-detection and random select of the parameters to the algorithm that not provide the optimal solution for the system.

REFERENCES

Ali, M.H. and M.F. Zolkipli, 2016. Review on hybrid extreme learning machine and genetic algorithm to work as intrusion detection system in cloud computing. ARPN. J. Eng. Appl. Sci., 11: 460-464.

Armbrust, M., A.D. Joseph, R.H. Katz and D.A. Patterson, 2009. Above the clouds: A Berkeley view of cloud computing. MCs Thesis, EECS Department, University of California, Berkeley, California.

Baysa, D., R.M. Low and M. Stamp, 2013. Structural entropy and metamorphic malware. J. Comput. Virol. Hacking Tech., 9: 179-192.

Bhuyan, M.H., D.K. Bhattacharyya and J.K. Kalita, 2014. Network anomaly detection: methods, systems and tools. IEEE. Commun. Surv. Tutorials, 16: 303-336.

Cannady, J., 1998. Artificial neural networks for misuse detection. Proceedings of the Conference on National Information Systems Security, October 6-9, 1998, Hyatt Regency, Crystal City, Arlington, Virginia, pp: 368-381.

Chouhan, P., 2015. A survey: Analysis of current approaches in anomaly detection. Intl. J. Comput. Appl., 111: 32-36.

Fossaceca, J.M., T.A. Mazzuchi and S. Sarkani, 2011. MARK-ELM: Application of a novel multiple kernel learning framework for improving the robustness of network intrusion detection. Expert Syst. Appl., 42: 4062-4080.

Frank, J., 1994. Artificial intelligence and intrusion detection: Current and future directions. Proceedings of the 17th Conference on National Computer Security Conference Vol. 10, October 11-14, 1994, Baltimore Convention Center, Baltimore, Maryland, pp: 1-12.

Gong, R.H., M. Zulkernine and P. Abolmaesumi, 2005. A software implementation of a genetic algorithm based approach to network intrusion detection. Proceedings of the 6th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks, May 23-25, 2005, Towson, Maryland, USA., pp: 246-253.

Huang, G.B., Q.Y. Zhu and C.K. Siew, 2004. Extreme learning machine: A new learning scheme of feedforward neural networks. Proceedings of the IEEE International Joint Conference on Neural Networks, July 25-29, 2004, IEEE, Budapest, Hungary, ISBN:0-7803-8359-1, pp: 985-990.

Jaiganesh, V., S. Mangayarkarasi and P. Sumathi, 2013. Intrusion detection systems: A survey and analysis of classification techniques. Intl. J. Adv. Res. Comput. Commun. Eng., 2: 1629-1635.

Kim, K.J. and S.J. Ahn, 2011. Proceedings of the International Conference on IT Convergence and Security 2011. Vol. 120, Springer, Netherlands, ISBN:978-94-007-2910-0, Pages: 642.

Kiranyaz, S., T. Ince, A. Yildirim and M. Gabbouj, 2009. Evolutionary artificial neural networks by multi-dimensional particle swarm optimization. Neural Networks, 22: 1448-1462.

Kukar, M., 2012. Transductive Reliability Estimation for Individual Classifications in Machine Learning and Data Mining. In: Reliable Knowledge Discovery, Dai, H., J.N.K. Liu and E. Smirnov (Eds.). Springer, Berlin, Germany, pp: 3-27.

Kumar, U. and B.N. Gohil, 2015. A survey on intrusion detection systems for cloud computing environment. Intl. J. Comput. Appl., 109: 6-15.

Li, G., P. Niu, X. Duan and X. Zhang, 2014. Fast learning network: A novel artificial neural network with a fast learning speed. Neural Comput. Appl., 24: 1683-1695.

Madhavi, M., 2012. An approach for intrusion detection system in cloud computing. Intl. J. Comput. Sci. Inf. Technol., 3: 5219-5222.

Mehmood, Y., U. Habiba, M.A. Shibli and R. Masood, 2013. Intrusion detection system in cloud computing: Challenges and opportunities. Proceedings of the 2nd National Conference on Information Assurance (NCIA), December 11-12, 2013, IEEE, Rawalpindi, Pakistan, ISBN:978-1-4799-1286-5, pp: 59-66.

Mishra, P., E.S. Pilli, V. Varadharajan and U. Tupakula, 2016. Author's Accepted Manuscript. J. Netw. Comput. Appl., 77: 18-47.

Modi, C., D. Patel, B. Borisaniya, H. Patel and A. Patel et al., 2013. A survey of intrusion detection techniques in cloud. J. Netw. Comput. Appl., 36: 42-57.

Nguyen, H.T., K. Franke and S. Petrovic, 2012. Feature Extraction Methods for Intrusion Detection Systems. In: Threats, Countermeasures and Advances in Applied Information Security, Manish, G. (Ed.). Information Science Reference Publisher, New York, USA., pp: 23-52.

Oktay, U. and O.K. Sahingoz, 2013. Attack types and intrusion detection systems in cloud computing. Proceedings of the 6th International Conference on Information Security & Cryptology, September 20-21, 2013, Gazi University, Ankara, Turkey, pp: 71-76.

Patel, A., M. Taghavi, K. Bakhtiyari and J.C. Junior, 2013. An intrusion detection and prevention system in cloud computing: A systematic review. J. Netw. Comput. Appl., 36: 25-41.

Raghav, I., 2013. Intrusion detection and prevention in cloud environment: A systematic review. Intl. J. Comput. Appl., 68: 7-11.

Shelke, M.P.K., M.S. Sontakke and A.D. Gawande, 2012. Intrusion detection system for cloud computing. Int. J. Sci. Technol. Res., 1: 67-71.

Vieira, K., A. Schulter, C.B. Westphall and C.M. Westphall, 2010. Intrusion Detection for grid and cloud computing. IT Prof., 2: 38-43.

Wang, J., W. Wu, Z. Li and L. Li, 2011. Convergence of gradient method for double parallel feedforward neural network. Intl. J. Numer. Anal. Model., 8: 484-495.

Whitman, M.E. and H.J. Mattord, 2012. Principles of Information Security. Course Technology Publisher, Australia, Pages: 617.