Classifier System in Cloud Environment to Detect Denial of Service Attack

Wafa' Slaibi Alsharafat
Prince Hussein Bin Abdullah
Faculty of Information Technology
Al al-Bayt University
Mafraq, Jordan

ABSTRACT

Cloud Computing is a modernize computer services in 21 century. This system has some of similarities with distributed systems, through using network environment features. Therefore the security requirement is one of most critical issues in this type of environment. Because of vast number of online users over network with times, the possibility for attacker or malicious actions to perform destruction actions becomes profitable and worthy. One of the major security encounters and challenges in cloud environment is the detection of possible attempts of attacks. For detecting these malicious activities especially Denial of Service (DoS) attack, this paper will propose Learning Classifier System for Intrusion Detection System (LCS-IDS) to detect DoS in cloud environment attacks by taking advantage of learning from attacks themselves and simulate possible DoS attacks through Genetic Algorithm, generator, to rise detection rate compared with other systems in this field.

Keywords

Cloud computing, Intrusion Detection, Denial of Service, Learning Classifier System.

1. INTRODUCTION

Cloud Computing is paradigm of distributed environment that consists from nodes, servers. Cloud environment provides hardware, software services and vast array of applications that may be used by any potential users via computer networks as; Internet or Intranet, private networks [1]. National Institute of Standards and Technology (NIST) specifies the definition of cloud computing as: "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."[19].

Cloud computing is service oriented architecture that has been implemented to overcome information technology overhead for users and system by improve system's flexibility and reduced total cost of ownership and provide high security level since network environment suffers from different type of attacks which they are[2]:

1- Denial of Service (DoS)
2- Probe
3- Remote to User attacks (R2L)
4- User to Root (U2R)

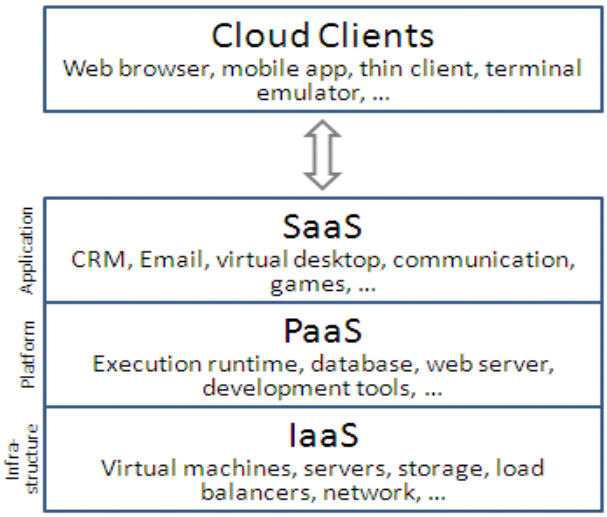
Cloud environment concerns about, in order, 1- Confidentiality, integrity, and availability (CIA). For Availability, the challenge appears as Denial of Service (DoS) attack, a main threat that faces to availability by making services or resources or systems unavailable by flooding systems with numerous useless traffic [3].

Delivering model in cloud environment states in three main types;

1- Infrastructure as a Service (IaaS)
2- Platform as a Service (PaaS)
3- Software as a Service (SaaS).

IaaS provides services as storage capacity, database management and hardware and computer capabilities based on needs and demands. In SaaS is a highly scalable internet applications are hosted by cloud environment and presented as services for users. While PaaS offers the needed platform for designing, developing, building and testing applications. Figure 1 summarizes main service for each model [4,6].

Figure 1.Cloud Computing models [20]



Cloud computing paradigm started in the mid of 2007 and it is developed rapidly to satisfy diffusion and infusion of IT in systems and to take advantages of using new technology. There are different features that encourage systems to adapt pervasive computing, cloud computing as follows [1, 5]:

1. Elasticity and Scalability: The cloud capabilities will be reformed based on users demands and requirements. In addition cloud environment resources as memory capacity or computing powers will be scaled up or down according to the users' needs and requirement.

2. Ease of use: users can access directly to the services via web browser without the need for extra resources.
3. Device and Location Independent: Users can access to Cloud services regardless their geographic locations and their devices type.
4. Deploying custom application using PaaS: Cloud environment enables user to outsourcing their application by using PaaS instead of in-house development.
5. Reduced cost: there is no need to purchase one-time usage or rarely computing tasks because these equipement can be outsourced by a third-party. Further, cloud focuses on operational expenses rather than capital expenses.
6. Multi-Tenancy: concerns about sharing resources and cost amongst potential users by applying isolation and virtualization.
7. Reliability: this feature satisfied by storing redundant copies of data in multi-locations and this will facilitate disaster recovery and will make data more reliability and availability. For that cause, cloud computing more suitable for disaster recovery is suitable for business stability.
8. Maintenance: Cloud applications can be accessed remotely so no need to install application on user's local devices.
9. Performance: As a centralization feature of cloud environment this makes monitoring facilities easier, and architectural independent using web services as the interface with entire system.

Intrusion Detection System (IDS) has developed as fundamental component of Information infrastructure. Heady in [2] specified that IDS specified to detect any unauthorized activities that will violate CIA of the entire system resources. Network-IDS (NIDS) is one of common and main technique for monitoring network traffics and sort network attacks into:

1. Denial of Service (DoS
2. User to Root (U2R)
3. User attacks (R2L)
4. Probe.

New IDS focused on cloud computing, with it, on network environment, a local network intrusion detection engine receives network packet and compares it with the various pattern of DoS. Afterwards, the generated alerts are processed by IDS to determine security level of cloud environment. In this paper, section 2 will presents related work for detecting DoS attacks in cloud environment, and section 3 presents the proposed Learning Classifier Intrusion Detection System (LCS-IDS). Finally, the expectation and conclusion for applying proposed method and future work to develop system with high detection rate.

2. RELATED WORK

All studies concerns about intrusion detection in network environment as in cloud environment which can be expressed by Detection rate. Cloud computing security and intrusions detection have been noticed and highly interest as hot research topic. So, different methods and techniques have been proposed and developed to improve and raise detection rate. Researchers in [7] proposed a Management Unit which acts as a mini IDS that assigned to each user according to their actions. So, This IDS can be distributed amongst multiple controller nodes which may contain IDS instances from multiple users. In [8] Distributed Intrusion Detection System (DIDS) is developed and classified as a signature based detection system. Here, IDS observes network packet if the packet matches with existing bad packet pattern then analyze these packet for serious attacks. If there was matching, an alert will be broadcasted to rest of IDSs and cooperative operation module. By voting, the bad packet will be added for block rule. This mechanism classified as prevention mechanism for bad packets. This mechanism works at network layer of OSI. But there is no specific mechanism at this layer to identify the bad packet which in this case causing Distributed Denial of Service (DDOS) attack.

Queuing theory model was suggested by Singh te.al.in [9] where the proposed model used for detecting DoS attack. Queuing theory model depends on two models; first model detect sudden changes in specific parameters of incoming packets that reside in the queue While the second model, considered as a signal generation module which implemented for further and additional processing.

Flooding based DoS attack detection has been introduced in [10]. Researchers have implemented covariance-matrix statistical approach which based on studying and monitoring network traffic features and their correlativity changes then perform a comparison between covariance matrixes for normal traffic with new incoming traffic based on predetermined thresholds. A a result, we can detect either the incoming traffic is attack or not by decision space. Also, this approach gain improved accuracy and efficiency through simulation experiments. Lastly, covariance-matrix had three spaces: covariance-matrix, captured traffic space and decision space.

In [11] authors suggested voting IDS to specifies the degree of attack effects and dangerous that will be shared via set up IDS client tools. As a result, the most dangerous attack determined through voting. Snort environment is an environment performs voting IDS. In [17] authors had designed eXtended Classifier System (XCS) and artificial intelligent network with inner modification(ANNXCS) for classifier generation to achieve better detection rate XCS take into account action set and use previous action set determine earlier. This method designed and tested by using KDD dataset and achieves 98.8%. Also Researchers in [18] presented an integrated neural network IDS. This Model consists of three phases: Phase-1 clustering and Selecting, Phase-2 aims for training while Phase-3 specified to learn and detect. This model obtain overall average detection rate of 96.6% of intrusion detection in network environment.

3. PROPOSED LEARNING CLASSIFIER INTRUSION DETECTION SYSTEM

Learning Classifier System (LCS) was invited in 1995 by Holland [12]. LCS is a machine learning paradigms that use rules which denoted as a classifiers. So, LCS is a rule-based system will be able to automatically build a set of rules to detect DoS attack. In addition, LCS considered as Reinforcement Learning (RL) that depends on an explicit set of rules for all system states. RL represents the state as a collection of features or characteristics called "attributes". According to the a generalization capability of LCS, LCS don't need the explicit enumeration or list While LCSs do depends on explicit rules according [13] by using # to indicates to don't care, less significant, input variable. LCS will be adapted in this work with entire elements as shown in Figure 1 and these elements are:

3.1 Message Input

Message is a rule where each message consists from a set of input variables that represent network traffic features from environment. Network feature attributes will be presented in rule using different encoding system as Real or Binary encoding for inputs. Here, Real encoding is preferable to encode significant features while in significant features encoded using (#) as shown in Figure 2.

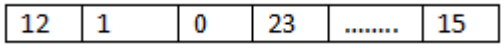


Fig. 2: Real encoding

3.2 Effector

Effector responsible for firing the action of selected rules to the environment. This fired action specifies if the incoming network packet is DoS or not.

3.3 Reward

Rewards (or penalty) processed on the rule that fired an action to the environment. Reward (or penalty) will be accomplished by adding a specific value to the rule fitness. So, fitness value considered as a parameter to give correct prediction if the incoming packet is an attack or not.

3.4 LCS execution

LCS executes any incoming input from environment as shown in Figure 3.

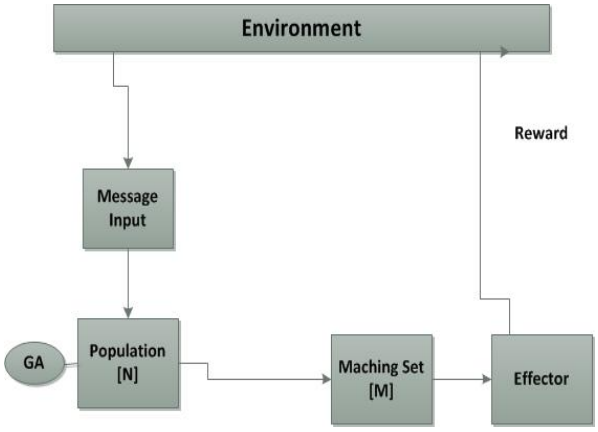


Fig. 3: Learning Classifier System (LCS)

LCS is a classifier system that used to determine if the incoming network packet classified as an attack or not. So LCS will choose the best action according to given its input. In this sense, right detection decision will be rewarded otherwise penalty.

To sum it all, by detector the system will receive an input from environment which later will be compared with the conditions part for every classifier in population (classifier pool) [N]. Classifiers that only match the incoming input will be placed in a set called Match Set [M].

In this case, [M] represent Knowledge base to deal with input packets. Each classifier in [N] associated with a fitness value which indicates the "usefulness" of the rule that used to fire an action for the environment.

By applying bidding mechanism, a specific rule will be selected from rules in [M. This selected rule becomes the system's fired action. And this cycle performed with every input packet.

A bidding Mechanism will be implemented for rewarding the activated rule or any related rule at any stage that helps in taking suitable and correct action which fired the action to the environment. This Mechanism achieved by sharing fitness amongst activated rules.

3.5 Classifier generator, Genetic Algorithm (GA)

The main role of GA in LCS is classifier production role. At the start, the initial classifiers are produced by applying and assign fitness value that signifies rules efficiency for determine whether the incoming network packet is attack or not (DoS, PROBE, U2R, R2L). GA perform its function by selecting a pair of classifiers from classifiers pool, [N], based on their fitness value. Then crossover and mutation process will take a place to produce new pair of classifier which may be placed as a new classifiers in classifiers pool.

Holland in [14] was firstly proposed GA to test the adaptation of genetic systems. GA has several forms as; Steady State Genetic Algorithm (SSGA) and Simple Genetic Algorithm (SGA). SGA differ than SSGA by replace all the current generation with new generation while SSGA replace the poorest classifiers and keep the best of current classifiers. The main steps of SSGA and related modification summarized as follows:

1. Selection: Selection is the start step to produce new generation. Parents Selection, two classifiers, depends on the quality of each parents that measured by classifiers fitness. In addition, there are different selection mechanisms as Roulette Wheel, Tournament, Stochastic selection and etc. [14]. After conducting a set of experiments, Roulette Wheel methods gain better result compared with other methods.

2. Crossover: Crossover is a second step that takes a place after selecting parents. Here two classifiers will be combined where each one take some features from each other. Even more, double crossover points will be implemented and every crossover operation determined by crossover probability (Pc).

3. Mutation: this process concerns about changing the value of each feature in condition part according to probability (Pm). Meanwhile,

random search is conducted to feed population with new "genetic material".

4. Replacement: The new produced classifiers will be replaced instead of old classifier according to their fitness value without exceed the population size.

So, GA plays a significant role in LCS by acting as a classifier generator. And LCS rule base can be described according to ecology of rules as: "each individual rule evolves in the context of the external environment and the other rules in the classifier system." [15].

4. CONCLUSION AND FUTURE WORK

To represent the performance of LCS-IDS) model, KDDcup'99[16] is implanted as a dataset for training and testing the feasibility of the proposed model that contains network packets, record, that represent DoS attacks to train classifier system for different network request to rise detection rate. The proposed classifier system will handle vast data in cloud environment by using single IDS node and being at a central point. Proposed LCS-IDS would like to work efficiently for analyzing data concurrently. Cloud environment is a wide distributed environment that can be an attracted environment for possible intrusions. Cloud environment faces a highly repetitive network access either from service providers, attacker or users. By using GA, there will be chance to learn from incoming network packets to generate different possibilities for attackers requests. This needs resulted according to attacker's method by changing their request frequently to evade any attempts to detect their flooding requests that prevent achieving their goal of making online systems out of service. According to efficient usage of classifiers systems in different application by achieving wanted results, Initial results give good expectations that proposed LCS-IDS will achieve expected detection rate compared with other methods. In addition, we will run our proposed approach and compare it with the other methods used to detect DoS attacks as [17, 18,]. By comparing results in order to gain better Detection rate. As a result, a set of enhancement will be suggested in future by changing type of GA, or modifying GA operators. At the initial stages of results, LCS-IDS shows a good detection rate compared with other methods as shown in Figure 4. After that, a set of modifications has been scheduled to improve detection rate.

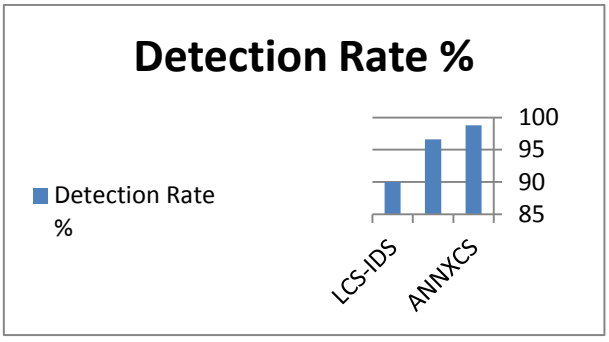


Fig. 4: Detection Rate

5. REFERENCES

[1] Shaikh, R., Sasikumar, M. 2012. Security Issues in Cloud Computing: A survey. International Journal of Computer Applications (0975 – 8887) Volume 44– No19. April 2012.

[2] Heady,R. , George, L., Maccabe, A., and Servilla, M. 1990. The Architecture of a Network Level Intrusion Detection System, Technical report, University of New Mexico.

[3] Bouzida Y., Cuppens F., Gombault S. 2006. Detecting and Reacting against Distributed Denial of Service Attacks. IEEE International Conference on Communication. Volume 5.

[4] Meiko Jensen, Jorg Schwenk, Nils Gruschka, Luigi Lo Iacon.2009. On technical Security Issues in Cloud Computing. Proc. of IEEE International Conference on Cloud Computing (CLOUD-II, 2009), pp. 109-116, India, 2009.

[5] Olive, Ch. 2011. Cloud Computing Characteristics Are Key. © General Physics Corporation 2011.

[6] Ram,S., Velmurugan,Thirukumaran. 2012. Effective Analysis of Cloud Based Intrusion Detection System. International Journal of Computer Applications & Information Technology Vol. I, Issue II, September, 2012 (ISSN: 2278- 7720).

[7] Dhage, S. N., et al. 2011. Intrusion Detection System in Cloud Computing Environment. In International Conference and Workshop on Emerging Trends in Technology (ICWET 2011) ' TCET, Mumbai. India, pp. 235-239, 2011.

[8] Chi-Chun Lo, Chun-Chieh Huang,Joy Ku. 2010. A Cooperative Intrusion Detection System Framework for Cloud Computing Networks. 39th International Conference on Parallel Processing Workshops,2010,pp280-284.

[9] Singh, N., S. Ghrera, and P. 2010 . Chaudhuri, Denial of Service Attack: Analysis of Network Traffic Anormaly using Queuing Theory. Arxiv preprint arXiv:1006.2807, 2010.

[10] Yeung, D.S., S. Jin, and X. Wang. 2007. Covariance-matrix modeling and detecting various flooding attacks. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2007. 37(2): p. 157-169.

[11] Lo, C.C., C.C. Huang, and J. Ku. 2010. A cooperative intrusion detection system framework for cloud computing Networks. 2010: IEEE.

[12] Holland, J. H. 1975. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press, Ann Arbor, MI.

[13] Lanzi, P.-L. 2002. Learning Classifier Systems from a Reinforcement Learning Perspective. Journal of Soft Computing, 6(3-4):162–170.

[14] Goldberg D.E.1989. Genetic Algorithms: in search Optimization and Machine Learning, Adison Wesley, 1989.

[15] Forrest, S. & Miller, J. 1991. Emergent Behavior in Classifier Systems. Physica D 42: 213-217.

[16] KDD-CUP 1999 Data, http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.

[17] Alsharafat.W. 2010. Applying Artificial Neural Network and eXtended Classifier System for Network Intrusion

Detection (ANNXCS-NID). The International Arab journal of Information Technology. Volume 10 No.3.

[18] Alrashdan.W. Naoum. R., ALsharafat.W., Alkhazaaleh. M. 2010. An Integrated Model of Intrusion Detection Based on Hybrid Neural Network and Support Vector Machine. International Journal of Computer Science and Network Security. Vol. 10 No. 11 pp. 10-13

[19] Olive.Ch. 2011. Cloud Computing Characteristics Are Key. General Physics Corporation (White Paper), www.gpworldwide.com.

[20] https://images.search.yahoo.com/search/images;_ylt=AwrB8pqkQGNWFjYAT_WJzbkF?p=cloud+computing.