ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

Detection of Masquerade Attacks by using DDSGA: Data-Driven Semi-Global Alignment Approach with CIDS Framework

Shital Rajabhau Ghuge[1], Prof. Sandip Satav[2]

P.G. Student, Department of Computer Engineering, JSCOE, Hadapsar, Pune, Maharashtra, India[1]

Associate Professor, Department of Computer Engineering, JSCOE, Hadapsar, Pune, Maharashtra, India[2]

ABSTRACT: Masquerade attack is one of the dangerous attack. Masquerade attackers expressanauthorized user to make use of the user services and advantages. The semi-global alignment algorithm (SGA) is one of the most operative and speciallymethods to find out these attack but it has not strsetch the accuracy and executions required by large scale, multiuser systems. To increase all the effectiveness and the execution of this algorithm, we suggest the DataDriven Semi-Global Alignment, DDSGA approach. From the security operative view point, DDSGA increase the scoring systems by adopting different alignment arguments for each user. more ever, it allow little change in user command series by allowing little becoming different in the low-level showing of the command to ability to perform a task . It makes suitable changes in the client using technique by updating the pattern of a user according to its current using technique. To faultless the runtime positioned, DDSGA to make as small the alignment overhead and parallelizes the find out and the update. After representing the DDSGA phases, we represent the experimental results. This result is show that DDSGA get the high hit ratio of 88.4 percent with a low false positive rate.

KEYWORDS: Masquerade detection, sequence alignment, security, intrusion detection, attacks

I. INTRODUCTION

Intrusion in network is the case when there was a breakthrough into the network by external or internal bad guys who can have different goals but mainly to have access to protected data or denying legitimate users from accessing their resources. The technology grows and many techniques are put in place to control the security on network; at the same time attackers keep innovating new techniques to succeed their mission. Attackers have their technique to disguise themselves (spoofing) so that they can cross the security perimeters like firewall.An intrusion detection system (IDS) can take a form of software or hardware platform used to monitor the state of network or system for unusual activities (behavior) .Its main purpose is not to prevent attack like firewall does but to detect breakthrough if any and alert the security team about it. Whatever form it takes, IDS have four components: the sensor of events as it is generated from source, the analyzer to check on characteristics or behavior of the event, the log for future references, the last part is the case when IDS is operating in active mode .The source of event depends on what type of IDS deployed; if that was network-based IDS (NIDS), the source reflects any promiscuous network port of the hardware and this last one captures packet flow events. But if it was host-based IDS (HIDS) the source is log of events already collected and stored somewhere on a host. Distinction of these sources and associated events is a key to understanding functionality of IDS. So, we distinguish three categories of intrusion detection systems: network-based, host-based and hybrid one. Host-based intrusion detection system (HIDS) is always in a form of software which is an intelligent system installed on a host and can be configured to operate as standalone or in distributed mode just like other systems. HIDS mainly uses operating system audit event logs to watch pattern of activities and attempt to discover unusual activities on single computer. Since HIDS relies on log, it means that the scope of attack detection is the activities initiated by local users. HIDS creates Database of state of every installed program on the host where it is running and

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

continue to watch its behavior continuously. Network-based intrusion detection system (NIDS) is a type of IDS which relies solely on the network traffic as raw data in order to detect suspicious activities. It can be either software or hardware; an example of hardware IDS is CISCO ISA 5000 series. This last one is capable of logging suspected packet flow. Furthermore, the administrator can access and view it. If it is deployed as software the good example is the use of packet analyzer systems installed on a host in LAN, so that all traffic flows must have a copy sent to that host. The technique is known as mirroring. Both using network traffic flow, they can detect flows which have deviated from expected metrics or rules on the entire network segment or subnet but they cannot interpret behavior inside any other machine.In Global Alignment approach , it checks complete part of commands it take more time to calculate or to find out test gap penalties.In Local alignment approach, it checks more similar part of commands so in that we can't get proper result of test gap penalties. But in semi-global alignment approach, firstly it checks similar part of strings and then go for deep inspection of string.A masquerade attack is an attack that uses a fake identity, such as a network identity, to gain unauthorized access to personal computerInformation through legitimate access identification. .This attack is very dangerous to organization because attacker is a legal user it has already have all privileges or rights.so, there is need to detect this type of attack .so, this system uses alignment approach to detect attack. There are some alignment approaches like, global, local, semi-global alignment approach to detect attack. Masquerader is one of the attackers who easily discovers the authorized user which uses the services and protection of the user. For this first SGA (Semi Global Alignment) algorithm can be used. It is most efficient and operative algorithm but the drawback of this algorithm is that it has not yet accuracy for the multiuser systems. For that view of point the DDSGA (Data Driven Semi Global Alignment) algorithm can be introduced. It is easily detect the attacks. It improves the effectiveness and efficiency more than the SGA algorithm. For the security system DDSGA improves the scoring system by adopting distinct alignment for each user. DDSGA minimizes the overhead of alignment and detects parallel and update it. After describing the DDSGA phases, we can got experimental results and this result shows the DDSGA can find the high hit ratio of 88.4% with low false positive ratio. DDSGA improves false positive rate and reduced Marion Townsend cost.

II.RELATED WORK

We studies some detection approaches for masquerade detection. The uniqueness approaches consider that command that have not been seen in the training data point out a masquerader. Again, the chance that a masquerader has issued a command is controversially related to the number of the users that use such command while performance of uniqueness is relatively poor. One-step Markova is based upon one-step transitions from a command to the next. This method false alarm rate is not satisfactory. Sconlau et al. toggled between a Markove model and simple not dependent. This approach accomplishes the good performance among the regard methods. The main idea about the compression approach is that new and old data of same user should compress at the same ratio. Masquerading user will compress data in different ratio. For binary data classification Support Vector Machine(SVM) indicate set of machine learning algorithm SVM can gives a large set of pattern but it result in high false alarm and low detection rate. Maxion and Townsend .applied a Naïve Bayes classifier widely used in text classification task and also classify user command data sequences into masquerader. An episode is introduced which is based on Naïve Bayes technique. According toNaïve Bays algorithm these episodes are Masquerade or normal. Which is used to the number of command in Masquerade block this technique improves the hit ratio but there is high false positive rates. So he does not update the user profile. In Naïve Bayes algorithm information on the probabilities of commands one user over the other usersHisham A. Kholidy and FabrizioBaiardi In this paper, they have presented the architecture of a distributed system for intrusion detection called CIDS, which employs multiple elementary detectors and combination of their alerts to make an accurate determination of intrusion. Then we presented an instantiation of this architecture with three elementary detectors and a manager with a graph-based and a Bayesian network based inference engine. They evaluated the system under a real-world web based e-commerce application and three classes of attacks. CIDS was found to bring down the incidence of missing alarms and false alarm with negligible impact on the performance. We are currently exploring how to set up the Rule Objects for different attack classes in an automated manner. The approach uses a feedback control loop to adjust the weights or probabilities in the rules. We are developing a larger set of test cases to carry out statistically large set of experiments to measure false positives and false negatives in CIDS. We are adding a timing

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

module to estimate the speed of propagation of attacks and augmenting the Response Engine to have a choice of responses which will be decided based on the timing information.[1]
Brian D. Davison and Haym HirshWe have presented a method that fulfills the requirements of an Ideal Online Learning Algorithm. Incremental Probabilistic Action Modeling has an average predictive accuracy at least as good as that previously reported with C4.5. It operates incrementally, will remember rare events such as typos, and does not retain a copy of the user's action history. IPAM can generate top-n predictions, and by weighing recent events more heavily than older events it is able to react to 'concept-drift'. Finally, its speed and simplicity make it a strong candidate for incorporation into the next adaptive interface.[2]
ChristopherIn this paper**,** SVMs provide a new approach to the problem of pattern recognition (together with regression estimation and linear operator inversion) with clear connections to the underlying statistical learning theory. SVM training always finds a global minimum. The simple geometric interpretation of SVM provides fertile ground for further investigation. An SVM is largely characterized by the choice of its kernel, and SVMs thus link the problems they are designed for with a large body of existing work on kernel based methods. I hope that this tutorial will encourage some to explore SVMs for themselves.[3]
Szymanski and Y. ZhangIn this paper, we show how to apply recursive data mining to solve the masquerade intrusion detection and author identification problems. Compared to the results from other researchers, our results are very promising. [4]
Subrat Kumar Dash, K. S. Reddy, and K. A. PujariWe propose, in this paper, a method that takes into consideration this aspect of user behavior while detecting masquerade attacks. Our scheme is based on the premise that the commands used by a legitimate user or an attacker may differ from the trained signature. But the deviation of the legitimate user is momentary whereas that of an attacker persists longer. By introducing this novel concept in the detection mechanism, the performance improves. We show this empirically using several benchmark datasets [5]
A. S. Sodiya, O. Folorunso, S. A. Onashoga, and P. O. OgundeyiIn this work, an improved semi-global alignment called Crosssemiglobal algorithm, is designed to improve the efficiency of masquerade detection. In the previous pair-wise algorithms, a fix value is always assumed as the gaps score. In Cross-semiglobal algorithm, the scoring function on which the algorithms based their scores is constructed from legitimate users' sequence of commands. This principle was implemented using platform independent C/C++ framework. The designed was tested using a systematically generated ASCII coded sequence audit data from Windows and UNIX operating systems as simulations for standard non-intrusive and intrusion data. The result shows a reduction in false positive rate from 7.7% using semi-global alignment to 5.4% using cross-semiglobal.

III. PROPOSED WORK

Security is important for every organization or system. From thesecurity competenceviewpoint, DDSGA prototypes more perfectlythe consistency of the behavior of distinct users byintroducing distinct parameters. Furthermore, it offers twoscoring systems that tolerate changes in the low-levelrepresentation of the commands functionality bycategorizing user commands and bring into lineinstructions in the same class without sinking the alignment groove. The scoring systems also tolerate both permutations of its commands and changes in the user behavior over time. All these features strongly reduce false positive and missing alarm rates and improve the detection hit ratio. In the testing using the SEA data set, the presentation of DDSGA is always improved than the one of SGA. From the computational viewpoint, the Top-Matching Based Overlying approach decreases the computational load of alignment by disintegrating the signature sequence into a smaller set of overlapped subsequences. Furthermore, the detection and the update processes can be parallelized with no loss of accuracy.

A] Algorithm:

Input V= Set of all user profile
Output: classification of user

1. User login

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

2. Observe each and every activity of user.
3. Build the user profile by using previous training data set
4. Now take current user command
5. Test current user command with previous data.
6. Test the gap penalties
7. If current data and previous data is same there is no penalties then that user is normal user

Otherwise that user is attacker

IV. ARCHITECTURAL VIEW

The architecture diagram of the system shown below helps us to understand the system.

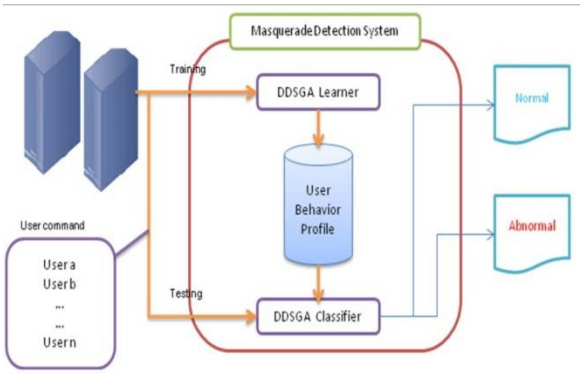


Figure 1:- System Architecture

- This paper is introducing the Data-Driven Semi-Global Alignment (DDSGA) approach. The DDSGA which recovers the discovery accuracy as well as the computational presentation of the Enhanced-SGA and of HSGAA which is also based upon SGA.
- The main idea of DDSGA is of considering the best alignment of the active session sequence to the recorded sequences of the same user. After discovering the misalignment areas, we label them as anomalous. Theseveral anomalous areas are a strong indicator of a masquerade attack.
- DDSGA can tolerate small mutations in the user sequences with small changes in the low level representation of user commands and it is decomposed into a configuration phase, a detection phase and an update one.
- The configuration phase, computes, and build user profile for each user, the alignment limits to be cast-off by the discovery as well as update phases.
- The detection phase aligns the user current session to the signature sequence. The computational performance of this phase is improved by two approaches. One is the Top-Matching Based Overlapping (TMBO) and second one is the parallelized approach.
- In the update phase, DDSGA extends both the user signatures and user lexicon list with the new patterns. It get done to reconfigure the system parameters.

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

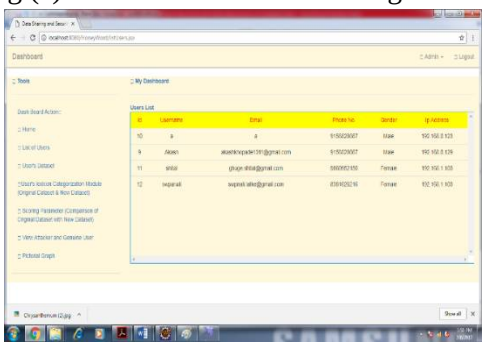
(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

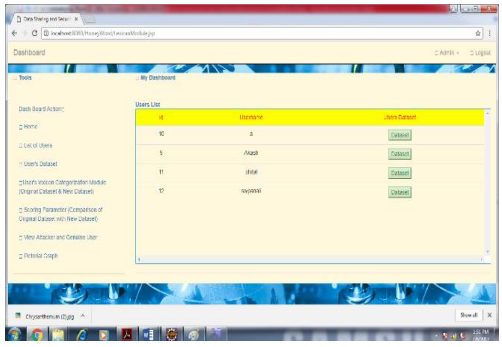
Vol. 6, Issue 7, July 2017

V. EXPERIMENTAL RESULTS

Following fig (a) shows user list with user registration details



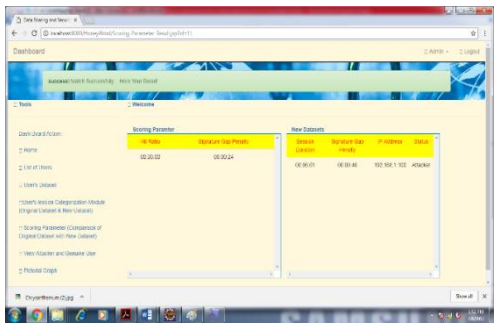
(a)



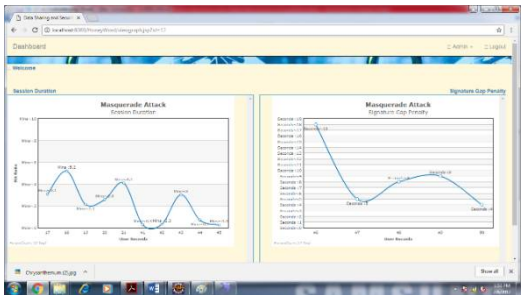
(b)

(b)

(c)



(d)



(e)

Fig (a). User list (b).User dataset (c). User list with dataset (d). Calculate scoring parameter (e). Result Analysis

ISSN(Online) : 2319-8753
ISSN (Print) : 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 7, July 2017

VI. CONCLUSION

To overcome drawback from SGA problem we have DDSGA model this model is more accurate. This keep the evenness by providing different parameter to different user and then it offers two level scoring system that tolerate means avoids modification in the small level commands functionality of user command and aligning commands in the same class but without reducing the alignment score. The recording systems also allow all to carry out of its instructions and modifications in the operatorperformance extra time. All features powerfullydamagewrong positive and lost alarm rates and increase the detection hit ratio. In the SEA data set, the performance of DDSGA is always improved as match to the one of SGA.

REFERENCES

[1] W. Dumouchel. (1999). "Computer intrusion detection based on Bayes Factors for comparing command transition probabilities".Technical report 91, National Institute of Statistical Sciences, [Online]. Available: www.niss.org/downloadabletechreports.
[2] W. Ju and Y. Vardi.(1999). "A hybrid high-order Markov chain model for computer intrusion detection".Nat. Inst. Statist.Sci. Research Triangle Park, NC, USA, Tech. Rep. 92.[Online]. Available: www.niss.org/downloadabletechreports.html
[3] Brian D. Davison and Haym Hirsh, "Predicting sequences of user actions," in Proc. Joint Workshop Predicting Future: AI Approaches Time Ser. Anal., 1998, pp. 5–12.
[4] T. Lane and C. E. Brodley, "Approaches to online learning and concept drift for user identification in computer security," in Proc4th Int. Conf. Knowl. Discovery Data Mining, New York, NY, USA, Aug. 1998, pp. 259–263.
[5] B. Christopher, "A tutorial on support vector machines for pattern recognition," Data Mining Knowl. Discovery, vol. 2, no. 2, pp. 121– 167, 1998.
[6] B. Szymanski and Y. Zhang, "Recursive data mining for masquerade detection and author identification," in Proc. IEEE 5th Syst., Man .Cybern. Inf. Assurance Workshop, West Point, NY, USA, Jun. 2004, pp. 424–431.
[7] S. K. Dash, K. S. Reddy, and A. K. Pujari, "Episode based masquerade detection," in Proc. 1st Int. Conf. Inf. Syst. Security, 2005, pp. 251–262.
[8] A. Sharma and K. K. Paliwal, "Detecting masquerades using a combination of Na€ıve Bayes and weighted RBF approach," J. Comput. Virology, vol. 3, no. 3, pp, 237–245, 2007.
[9] Subrat Kumar Dash, K. S. Reddy, and K. A. Pujari, "Adaptive Naive Bayes method for masquerade detection", Security Commun. Netw., vol. 4, no. 4, pp. 410–417, 2011.
[10] S. Malek and S. Salvatore, "Detecting masqueraders: A comparison of one-class bag-of-words user behavior modeling techniques," in Proc. 2nd Int. Workshop Managing Insider Security Threats, Morioka, Iwate, Japan. Jun. 2010, pp. 3–13.
[11] A. S. Sodiya, O. Folorunso, S. A. Onashoga, and P. O. Ogundeyi, "An improved semi-global alignment algorithm for masquerade detection," Int. J. Netw.Security, vo1. 12, no. 3, pp. 211–220, May 2011.
[12] T. F. Smith and M. S. Waterman, "Identification of common molecular subsequences," J. Molecular Biol., vol. 147, pp. 195–197, 1981.
[13] D. B. Christopher and T. D. Herbert, "Receiver operating characteristic curves and related decision measures: A tutorial," ChemometricsIntell. Lab. Syst., vol. 80, pp. 24–38, 2006.
[14] K. Wang and S. J. Stolfo, "One class training for masquerade detection," in Proc. IEEE 3rd Conf. Workshop Data Mining Comput.Secur., Florida, Nov. 2003.