2010 Second International Conference on Network Applications, Protocols and Services

Time Based Intrusion Detection on Fast Attack for Network Intrusion Detection System

Faizal, M.A.[1], Mohd Zaki M.[2], Shahrin Sahib[3], Robiah, Y.[4], Siti Rahayu[5], S., Asrul Hadi, Y.[6]
[12345]Faculty of Information and Communication Technology
Univerisiti Teknikal Malaysia Melaka,
Karung Berkunci No. 1752 Pejabat Pos Durian Tunggal, 76109 Melaka.

[6]Faculty of Information and Science Technology
Multimedia University
Jalan Ayer Keroh Lama
75450 Ayer Keroh, Melaka

{[1]faizalabdollah, [2]zaki.masud, [3]shahrinsahib, [4]robiah, [5]sitirahayu}@utem.edu.my, [6]asrulhadi.yaacob@mmu.edu.my

Abstract- In recent years network attack are easily launch since the tools to execute the attack are freely available on the Internet. Even the script kiddies can initiate a sophisticated attack with just a basic knowledge on network and software technology. To overcome this matter, Intrusion Detection System (IDS) has been used as a vital instrument in defending the network from this malicious activity. With the ability to analyze network traffic and recognize incoming and on-going network attack, majority of network administrator has turn to IDS to help them in detecting anomalies in network traffic. The gathering of information and analysis on the anomalies activity can be classified into fast and slow attack. Since fast attack activity make a connection in few second and uses a large amount of packet, detecting this early connection provide the administrator one step ahead in deflecting further damages towards the network infrastructure. This paper describes IDS that detects fast attack intrusion using time based detection method. The time based detection method calculates the statistic of the frequency event which occurs between one second time intervals for each connection made to a host thus providing the crucial information in detecting fast attack.

I. INTRODUCTION

In this information and communication technology age; the society cannot imagine living without the Internet and information systems. Nowadays the Internet plays an important role in stock market, access to weather forecast, E-medicine, E-commerce and even daily newspapers. The networking revolution has fully come of age in the last decade. As the network grows in size and complexity and computer services expands, vulnerabilities within local area and wide area network has become mammoth and causing lot of loop hole in security aspect [1]. The problems occur due to the increasing number of intrusion tools and exploiting scripts which can entice anyone to launch an attack on any vulnerable machines. An attack on network can be in 5 phases, which are Reconnaissance, Scanning, Gaining access, Maintaining Access and Covering tracks [2]. Identifying the first 2 activities will let the administrator to prevent the attack from doing further damage to the service offered by the network.

The attack can be launched in term of fast attack or slow attack. Fast attack can be defined as an attack that uses a large amount of packet or connection within a few second [3]. Meanwhile, slow attack can be defined as an attack that takes a few minutes or a few hours to complete [4]. Both of the attack gives a great impact to the network environment due to the security breach. Currently IDS is used as one of the defensive tools in strengthens the network security especially in detecting the first two phases of an attack either in form slow or fast attack. IDS acts as the monitoring tool to capture and analyze the network traffic for any anomalies activity.

This paper presents a novel methodology on detecting fast attack using time based detection technique for intrusion detection system. In this methodology the features capture from the network traffic is computed with respect to the time. The derived time based features from this methodology can help identify fast attack since the detection is based on the number of time the attacker made towards the host in second. Thus focusing in the time based features the early detection of the attack can be achieved and the security personnel can directly do the necessary action to stop further damages.

The rest of the paper is structured as follows. Section 2 discuses the related work on Intrusion detection system, Section 3 presents the methodologies and the technique use in time based intrusion detection for fast attack. Section 4 elaborates on the analysis and result. Finally, section 5 conclude and discuss the future directions of this work.

978-0-7695-4177-8/10 $26.00 © 2010 IEEE
DOI 10.1109/NETAPPS.2010.33

IEEE computer society

II. RELATED WORK

An intrusion detection system can be divided into two approaches which are behavior based (anomaly) and knowledge based (misuse) [5], [6]. The behavior based approach is also known as anomaly based system while knowledge based approach is known as misuse based system [7], [8].

The misuse or signature based IDS is a system which contains a number of attack description or signature that are matched against a stream of audit data looking for evidence of modeled attack [9]. The audit data can be gathered from network traffic or an application log. This method can be used to detect previous known attack and the profile of the attacker has to be manually revised when new attack types are discovered. Hence, unknown attacks in network intrusion pattern and characteristic might not be capture using this technique [10].

Meanwhile, the anomaly based system identifies the intrusion by identifying traffic or application which is presumed to be normal activity on the network or host [4], [25]. The anomaly based system builds a model of the normal behavior of the system and then looks for anomalous activity such as activities that do not confirm to the established model. Anything that does not correspond to the system profile is flagged as intrusive.

False alarms generated by both systems are major concern and it is identified as a key issues and the cause of delay to further implementation of reactive intrusion detection system [11]. Therefore, it is important to reduce the false alarm generated by both of the system [26]. Although false alarm is a major concern in developing the intrusion detection system especially the anomaly based intrusion detection system, yet the system has fully met the organizations' objective compared to the signature based system [12]. The false positive generated by the anomaly based system is still tolerable even though expected behavior is identified as anomalous while false negative is intolerable because they allow attack to go undetected [12]. Based on this motivation, anomaly based intrusion detection system is selected as an approach in detecting fast attack. Furthermore this research also managed to reduce the false alarm using the new model proposed by the logistic regression technique.

The success of an IDS depends on the decision upon a set of features that the system is going to use for detecting the attacker especially the fast attacks. This is because the mechanism of a fast attack requires only a few seconds and the technique used by the attacker to launch the attack is also different [13]. To the best of our knowledge, there is no comprehensive classification of features that intrusion detection system might use for detecting network based attacks especially fast attacks. Different researchers use different names for the same subset of feature while others use the same name but different types [14]. Furthermore, understanding the relationship as well as the influence of the features in detecting the fast attack is also necessary to avoid any redundant features selected for the intrusion detection system.

III. METHODOLOGY

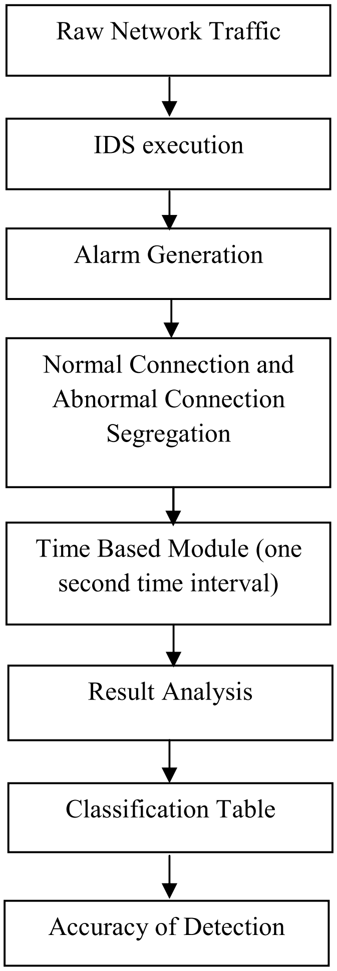

Figure 1: Time Base Detection Methodology

The relationship and the influence of the features have not been stated inside the network security related journals, articles and white-paper [14]. Therefore this research will introduce a set of minimum feature used to detect the fast attack and expose how the feature influences the detection of the attacker especially fast attacks. For this paper the researcher used the derived features that is based on time based traffic which is computed with respect to the past t seconds (***t*** is the size of the time windows interval, for example one second or one minute). For KDDCUP99[15], these features are designed to capture properties that mature over a two second temporal windows and for this research the mature time is considered to be one second and the next section will rectify why this duration of time is choose.

Previous work done by Hussain et al, [16] used 60 connections per second from source IP address as one of the criteria to identify the intrusion. The selection of 60 connections per second was purely based on the observation. The detailed process of the observation is not clearly stated in this research which is similar to Kanlayasiri et al [17] where the research identifies the portscan activity by looking at a host that make 20 connections per second.

We used TCPDUMP application to read the raw network traffic. In this research we only concentrated on the TCP protocol since TCP protocol is widely used protocol [18].

Due to huge amounts of network traffic, it is difficult to distinguish the normal and abnormal behavior of network traffic [19]. Therefore we used current intrusion detection system to distinguish between the normal and abnormal behavior of the network traffic. This technique has been applied by Caulkins et al, [20] in his research using Snort to distinguish the normal and abnormal behavior of the network traffic. Beside Snort, Bro also has been used as a tool to distinguish between the normal and abnormal connection. Zhang and Leckie, [21] using Bro inside their research in distinguishing the normal and abnormal connection. In our research, we used the Bro to distinguish between the normal and abnormal connection. Bro default configuration has been applied in this research.

The alarm generated by Bro will be captured and the attacker IP address will be identify. The IP address of the normal and attacker will be used to find the number of connection made by both of the IP address based on one second time internal. The one second time interval is selected based on the argument stated by Jung et al, [22] that detecting intrusion as quickly as possible may prevent major losses. The output from the time based module will be used for searching and comparing for each of the IP address based on the one second time interval. The output of the normal and abnormal connection rates within one second time interval will be combine together to feed inside the logistic regression technique to identify the accuracy of the detection. Then, the Classification Table based on the logistic regression model will be used to assess the detection model and accuracy of the detection. The classification table is the most appropriate test if the test objective is based on the classification [23].

Therefore the classification table is chosen as one of the test used to assess the model. Using the classification table, the percentage of the detection attack rate and detection normal rate can be calculated. Furthermore, error rate can be calculated in the classification table also. The error rate of the classification table can be divided into two categories which are false positive and false negative. False positive means that the number of errors in which a normal event is considered as an attack event. Meanwhile, false negative means the number of errors in which the attack event is predicted to be normal, but is in fact an attack.

IV. ANALYSIS AND RESULT

We test our approach using real time network traffic captured from of the agencies in Malaysia. The network traffic was captured using port mirror technique from the external router of the agencies as depicted in Figure 2. The network traffic will be divided into normal and abnormal network traffic using the IDS approach as stated in Figure 1. The output based on one second time interval will be collected of the normal and abnormal network traffic. There are 1045 connection has been declare as normal connection while 71 connection is an abnormal connection. Both of the connection will be combined and fed inside the logistic regression technique using the SPSS tool.

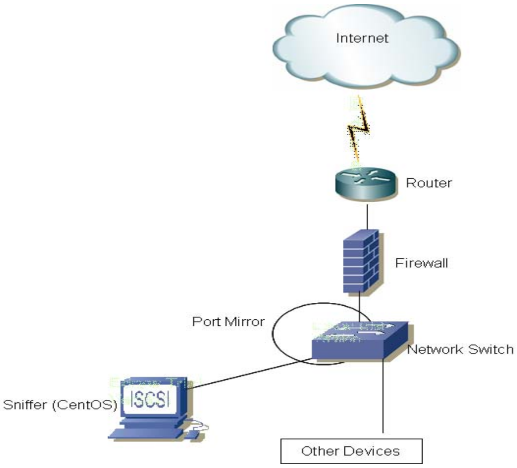

Figure 2 : Location of Data Collection

The logistic regression technique will produce the classification table to identify the accuracy of detection. The classification table produced by the logistic regression technique was based on the Null Model and Full Model. For the Classification table, there are two model involve which are null model and full model. The Null Model is a model which has only a constant value without any mechanism to distinguish between the attack and normal connection. Meanwhile the Full Model is model generated after the predictor is involved in the detection. The predictors include focus on the number of connection based on one second time interval. Table 1 and 2 show the result of the Null Model and Full Model.

TABLE 1
Null Model

Classified		Predicted	
		Normal	Attack
Observed	Normal	1045	0
	Attack	71	0

Detection Attack rate = 0%, False Negative = 6.4%, Detection Normal rate = 100%, False Positive = 0%, Overall Detection rate = 93.6%

Table1, shows that the model manages to predict 100% correct in classifying the normal however the false negative was also very high which is 6.4%. This indicates that the model makes a lot of false negative which is very dangerous to the organization because many attacks are not detected. Unfortunately, the model also does not have capabilities to detect the attack because using constant, the model assumes most of the data are normal.

TABLE 2
Full Model

Classified		Predicted	
		Normal	Attack
Observed	Normal	1043	2
	Attack	10	61

Detection Attack rate = 85.9%, False Negative = 0.9%, Detection Normal rate = 99.8%, False Positive = 3.2%, Overall Detection rate = 98.9%

After the predictor was included inside the model, the detection attack accuracy was high and reduces the false positive as depicted in Table 2**Error! Reference source not found.**. The model has capabilities to predict 85.9% correct in classifying the attack and only 3.2% false positive. The false negative generated from the full model also reduces to 0.9%. Although the attack detection rate is only 85.9% but it is still acceptable since the current intrusion detection system have 80% capabilities to detect the intrusion [24]. The model has better prediction and has capabilities to distinguish the difference between the attack and normal traffic. Furthermore, the overall percentage of the classification table for the null model was 93 %. The result of the overall percentage increase to 98.9 % after the full model is applied to the data. As a conclusion, the increase in of the correct percentage for the classification between the attack and normal indicate that the model is suitable, fits and good in predicting the normal and abnormal behavior.

V. CONCLUSION AND FUTURE WORK.

Before determining a network traffic is a potential threat to a network or not, there is a need for an IDS to have a method in differentiating whether it is malicious or not. Therefore, this research has introduced a new methodology to identify a fast attack intrusion using time based detection. The method used to identifies anomalies based on the number of connection made in 1 second. The approach is then tested on real network traffic data and the result is then evaluated by using the Classification Table based on the logistic regression model. From the test and analysis it is shown that the model is suitable for predicting the normal and abnormal behavior.

For further validation, the methodology will be implemented on a different set of real network traffic. In view of the fact that this research only concentrate on the TCP connection only, in the near future the researcher are planning to investigate use other protocol and other flag to recognize the fast attack intrusion activity. Inspecting other protocol and flag it may help to detect fast attack intrusion activities that launch using UDP or ICMP protocol. Finally the approach introduce in this research will be implemented on a production network for accessing the performance on the anomalies detection using time based detection.

ACKNOWLEDGMENT

This research has been supported by University Technical Malaysia Melaka and Malaysia Government under FRGS Fund.

REFERENCES

[1] Papadogiannakis, A., Polychronakis, M. & P. Markatos, E., (2010). Improving the Accuracy of Network Intrusion Detection System Under. Load Using Selective Packet Discarding. European Conference on Computer System, Paris, France.

[2] EC-Council. (2009) ,CEH Training Module.

[3] Ertoz, L., Eilertson, E., Lazarevic, A., Tan, P., Docas, P., Kumar, V. & Srivastava, J. (2003). Detection of Novel Network Attacks Using Data Mining. In ICDM workshops on Data Mining for Computer Security. Melbourne.

[4] Wenke Lee. (1999). A Data Mining Framework for Constructing Feature and Model for Intrusion Detection System. PhD thesis University of Columbia.

[5] Cuppen, F. & Miege, A. (2002). Alert Correlation in a Cooperative Intrusion Detection Framewok. In Proceeding of the 2002 IEEE Symposium on Security and Privacy. IEEE, 2002.

[6] Cabrera, J.B.D., Ravichandran, B & Mehra R.K. (2000). Statistical Traffic Modelling for Network Intrusion Detection. In Proceeding of the IEEE Conference.

[7] Yeophantong, T, Pakdeepinit, P., Moemeng, P & Daengdej, J. (2005). Network Traffic Classification Using Dynamic State Classifier. In Proceeding of IEEE Conference.

[8] Farah J., Mantaceur Z. & Mohamed BA. (2007). A Framework for an Adaptive Intrusion Detection System using Bayesion Network. Proceeding of the Intelligence and Security Informatics, IEEE, 2007.

[9] Wang Y., Huang GX. & Peng DG. (2006). Model of Network Intrusion Detection System Based on BP Algorithm. Proceeding of IEEE Conference on Industrial Electronics and Applications, IEEE, 2006.

[10] Sekar, R., Gupta, A., Frullo, J., Shanbhag, T., Tiwari, A., Yang, H. & Zhou, S. (2002). Spesification-based Anomaly Detection: A New Approach for Detecting Network Intrusions. In Proceeding of CCS ACM Conference.

[11] Karl Levitt. (2002). Intrusion Detection: Current Capabilities and Future Direction. Proceeding of IEEE Conference of the 18th Annual Computer Security Application, IEEE, 2002.

[12] Garuba, M., Liu, C. & Fraites, D. (2008). Intrusion Techniques: Comparative Study of Network Intrusion Detection Systems. In Proceeding of Fifth International Conference on Information Technology: New Generation, IEEE, 2008.

[13] Robertson S., Siegel EV., Miller M. & Stolfo SJ. (2003). Surveillance Detection in High Bandwidth Environment. In Proceeding of IEEE Conference on the DARPA information Survivability and Exposition, IEEE, 2003.

[14] Onut IV and Ghorbani AA. (2006). Toward a Feature Classification Scheme for Network Intrusion Detection. In Proceeding of the 4th annual Communication Network and Services Research Conference, IEEE, 2006.

[15] KDDCUP99 dataset. http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[16] Hussain A., Heidermann, J. and Papadopoulos, C. "A Framework for Classifying Denial of Service Attacks". In Proceeding of 2003 ACM SIGCOMM, Germany, 2003.

[17] [20] Kanlayasiri, U., Sanguanpong, S. & Jaratmanachot, W. "A Rule Based Approach for Port Scanning". In Proceeding of Electrical Engineering Conference. Thailand, 2000.

[18] Gavrilis, D & Dermatas, E. (2005). Real-Time Detection of Distributed Denial of Service Attack Using RBF Network and Statistical Feature. International Journal of Computer Network, Vol 48, pp 235-245.

[19] Jalili R., Fatemeh IM., Morteza A., Hamid RS. (2005). Detection of Distributed Denial of Service Attacks Using Statistical Pre-processor and Unsupervised Neural Network. ISPEC, Springer-Verlag Berlin Heidelberg, 2005.

[20] Caulkins BD., Joohan L., Morgan CW. (2006). Bootstrapping Methodology for the Session-Based Anomaly Notification Detector (SAND). ACM, Melbourne 2006.

[21] Zhang, D & Leckie, C. (2006). An Evaluation Technique for Network Intrusion Detection Systems. In Proceeding of the First International Conference on Scalable Information Systems, Hong-Kong, June 2006.

[22] Jung, J., Paxson, V., Berger, W. & Balakrishnan, H. (2004). Fast Portscan Detection Using Sequential Hypothesis Testing. In Proceeding of the 2004 IEEE Symposium on Security and Privacy (S & P' 04).

[23] Hosmer D.W and Stanley, L. (2000). Applied Logistic Regression Second Edition. USA. John Wiley and Son Inc.

[24] Cavusoglu, H., Mishra B. K. & Raghunathan, S. (2004). The Effect of Internet Security Breach Announcements on Market Value of Breached Firms and Internet Securiy Developers. International Journal of Electronic Commerce, Vol 8, pp,4.

[25] Zhenging, H, Zhitang, L. & Junqi, W., (2008). A Novel Network Intrusion Detection System (NIDS) Based on Signatures Search of Data Mining. Proceeding of the 1st International Conference on Forensic Application and Techniques in Telecommunications, Informations and Multimedia, Adelaide, Australia.

[26] Zhenwei Yu and Jeffrey J. P Tsai. (2008). An Adaptive Automatically Tuning Intrusion Detection System. ACM Transactions on Autonomous and Adaptive System, Vol 3, No. 3.