International Journal on Emerging Technologies 1(1): 71-74(2010) ISSN : 0975-8364

Firewall anamolies detection and removal techniques – a survey

Rupali Chaure and Shishir K. Shandilya***

**PG Deptt. of CS and Engg. NRI Institute of Information Science and Technology Bhopal, (MP) INDIA*

(Received 5 Jan., 2010, Accepted 26 March., 2010)

ABSTRACT : Due to the rapid growth in the field of Internet, the related security mechanisms are the key area of research. Firewalls serve the solution for secured Internet experience. Latest firewalls are fully-equipped for providing hi-end security to the network. However, due to the continuous growth of security threats, the firewall mechanisms and policies are compulsorily needed to get updated. The manual processing for detecting anomalies in firewall is complex and often error-prone. Any minor change in the rule set of firewall leads to the requirement of rigorous analysis for maintaining the consistency and efficiency of firewall mechanism. The presented paper covers the advancements of various approaches proposed by researchers in this field. This paper also discusses the policies for creating, modifying and sustaining the rule sets of firewall in such a way that makes the rule set optimal and free from known anomalies.

I. INTRODUCTION

A firewall is a system acting as an interface between a network and one or more external networks. It helps implementing the security policy of any network by deciding which packets to let pass through and which to block, based on the set of rules defined by the network administrator. Any error in defining the rules may compromise the system security by letting undesired traffic pass through or blocking the desired traffic. The rules when defined manually often results in a set that contains conflicting, redundant or overshadowed rules, which creates anomalies in the firewall policy. A network firewall protects a computer network from unauthorized access. Network firewalls may be hardware devices, software programs, or they may be a combination of the two. Network firewalls guard an internal computer network (home, school, business intranet) against malicious access from the outside. Network firewall may also be configured to limit access to the outside network of internal users. If passwords provide a 'door' to cover the 'doorway' into your 'house', then firewalls provide 'shutters' to cover the 'windows'. A firewall does absolutely nothing to protect the windows you leave open - that's the job of the programs, which provide the services at those windows.

The firewall is ideally a separate computer, which exists between a network and the Internet. It can be a purpose-built device - some of them are available as small black boxes which look like network hubs. This computer can be any old 486, with a highly secure operating system that provides an inbuilt firewall. None of the network computers should be able to access the Internet or can be accessed from the Internet without going through the firewall.

II. FIREWALL RULES

Whenever a packet is tested by the Firewall, it means that the header of the incoming or outgoing packet is tested against all the rules one by one, which are stored in the Firewall rule set. The rules in the Firewall rule set consists all the header information like source and destination address, source and destination port address and the corresponding action to performed i.e. whether to accept or deny any packet which matches all the other fields of any rule in the rule set. The rules are stored in the rule set in the following format,

```
<order> <prtcl> <S_ip> <S_port> <D_ip> <D_port> <action>
```

Here all the terms have respective meanings with properly defined domains. Order is the number at which the rule is stored in the rule set, prtcl is the type of the protocol specified in the packet's header, `s_ip` and `s_port` are the source machines' IP address and port number respectively. Similarly `D_ip` and `D_port` are the IP address and port number of the destination. In the last action field defines the resulting action to be performed on the packet which matches all the previous fields. The action field can be either ACCEPT or DENY. These rule sets of any firewall defines the Security Policy of that organization. The security policy of any organization is very dynamic i.e. it can be altered anytime whenever the administrator wants to modify the rules. So such frequent changes are the reason for the inconsistencies in the rule set.

III. FIREWALL ANOMALIES

As the rule set is very large it becomes difficult to check all the rules for any redundancy. Hence the updating of rule set may generate erroneous set of rules which are unable to perform their intended job i.e. protection from unauthorized access to the network or from the network. These errors in the rule set are called anomalies that have to be detected and removed from rule set for the efficient working of any firewall. Till date, five types of anomalies are discovered and studied namely, Shadowing Anomalies, Correlation Anomalies, Generalization Anomalies, Redundancy Anomalies, and Irrelevance Anomalies.

A. Shadowing anomaly

Two rules are said to have shadowing anomaly, whenever the rule which comes first in rule set matches all the packets and the second rule which is positioned after

the first rule in rule set does not get chance to match any packet because the previous rule has matched all the packets. It is a very critical problem since the rule coming later to the previous rule will never get activated. Hence the traffic to be blocked will be allowed or the traffic to be permitted can be blocked.

B. Correlation anomaly

Two rules are said to have correlation anomaly if both of them matches some common packets i. e. the rule one matches some packets, which are also matched by the rule second. The problem here is that the action performed by both the rules is different. Hence in order to get the proper action such correlated rules must be detected and should be specified with proper action to be performed.

C. Generalization anomaly

Two rules which are in order one of them is said to be in generalization of another if the first rules matches all the packets which can be also matched by the second rule but the action performed is different in both the rules. In this case if the order is reversed then the corresponding action will also be changed. The rule, which comes later in the rule list, is shadowed by the previous rule and also it has no effect on incoming packets. The super set rule is called General rule and the subset rule is called Specific rule. If such generalization relation exists between two rules then the super set rule should be placed after the subset rule in the rule list.

D. Redundancy anomaly

Two rules are said to be redundant if both of them matches some packets and the action performed is also the same. So there is no effect on the firewall policy if one of redundant rules will be removed from the rule set. It is very necessary to search and remove the redundant rules from the rule set because they increase the search time, space required to store the rule set and thus decrease the efficiency of the firewall. The firewall administrator should detect and remove such redundant rules to increase the performance of the firewall.

E. Irrelevance anomaly

Any rule is said to be irrelevant if for a given time interval it does not matches any of the packets either incoming or outgoing. Thus if any type of the packets do not match a rule then it is irrelevant i.e. there is no need to put that rule in the rule set. Till now all the above four anomalies are detected and removed successfully but irrelevant anomaly is still not completely defined in any automated software implementation yet.

The size of the rule set varies according to the type of the organization. Generally the rule set is very large because different administrators come and modify the policy rules according to their requirements and so is the reason of occurrence of anomalies. Because of the large size of the rule set it is difficult to detect anomalies by manually checking the rules one by one. So there is different software implemented to perform the job of anomaly detection and removal automatically.

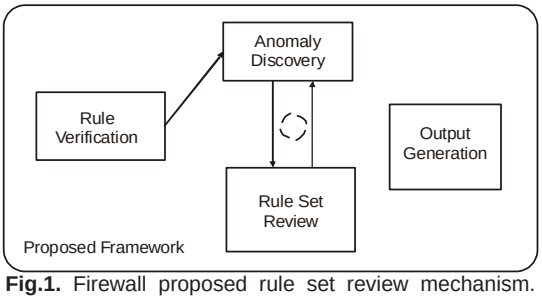

Fig.1. Firewall proposed rule set review mechanism.

IV. EVOLUTION

The endless growth of internet in today's commercial and technical scenario finds the need to secure the data which should be protected against unauthorized access. Firewalls perform this job of protecting any network. A lot of research work has been done in the field of Firewalls. The main problem that arises in firewalls is that anomalies are generated during updating the rules in the rule set. So the main interest of research is the detection and removal of firewall anomalies. There are a number of approaches for this, which varies to each other in some implementation context.

Anomaly free editing in firewall policy rules which includes insertion, deletion and modification in rule set with the help of a tool developed Firewall Policy Advisor (FPA). All this work was carried out on a single firewall environment. All the anomalies of single and multi firewall environment were detected and a set of algorithms defined to automate this process by creating a policy tree of rules in the rule set of a firewall [1-4].

Chotipat and Chomsiri introduced a method of analyzing packets from the filtering rule list by using the concept of Relational Algebra in 2004 [5,6]. They mapped the firewall rules onto relations. Then by performing various relational algebra operations like select, project, join, set difference etc. some anomalies have been discovered and removed. A raining 2D box model was also represented which shows a simulation of packets by rectangular boxes, which fall, like rain.

An open source application implemented which validates large computer networks. It explains the generic network rules and automatically detects the selected anomalies. This work is based on the concept that in every network there are some global variables that can be profitably used for detecting network anomalies irrespective of the type of networks, its users and the equipments used in the networking. This work describes that how the firewall anomalies can be detected and removed by performing the analysis of network behavior using all types of signatures of attacks [7].

The process of protecting a network with the help of a firewall designed by the software called Firmatto. It designs a new anomaly free set of rules based on the present firewall rule set and places a new firewall with this new rule set in between the previous one and the outside network. The drawback of this approach was that it does not provide any user interactions and its performance degrades for large rule sets [8].

The rule ordering and optimization is done by use of Direct Acyclic Graphs and also a proof was given to show that this problem is NP- hard. A study of simulation results was presented to show that by performing rule optimization and reordering we could increase the firewall efficiency along with maintaining policy integrity [9].

In addition to these static approaches this task of anomaly detection and removal by using maximum entropy method. They developed a behavioral based method to detect and remove anomalies in network traffic. A comparison is done on current network traffic against baseline distribution, which gives a multidimensional view of network so that the administrator can easily detect anomalies that cause abrupt changes in the network traffic and it also provides the information about the detected anomalies [10].

A method for detection and removal of Firewall policy anomalies within a set of given firewall rules by using the relationship between the rule attributes [11]. They further extended their work [12] and again performed the task of anomaly detection and removal, but this time on both firewalls and network intrusion detectors; so that the network security policy will be prevented from unknown attacks as well as it can undergo the process of detection and removal of firewall misconfigurations.

A new algorithm for merging of rules to decrease the firewall rule set and so increasing the firewall efficiency [13].

Katic and Pale [14] in 2006 have presented a similar system to reorder and merging firewall rules but on LINUX firewall rules. This software is called FIRO, which was used with IP tables in LINUX, which can be adopted for any other operating system.

For large number of rules in the rule set a better approach was introduced for rule ordering that performs the processing of each packet individually. They checked the hit rate of each rule in the rule list and frequently matching rules are placed at higher priorities, the delay introduced due to mismatching rules is then considered as a limiting factor to perform a firewall at its best. A simple algorithm was presented for Access Control Lists' optimization and it was checked against several simulated rule sets generated by an in house numerical model which generates Access Control Lists and traffic flows based on given parameter set. This simulation results in an increase of 23% firewall efficiency [15].

Acharya *et. al.*, have performed a traffic aware firewall optimization. Their work was based on traffic characteristics, which affects abruptly the firewall performance. So they implemented a technique to automatically update the firewall rule set based on any dynamic changes in the traffic flow. Also they have showed that by altering firewall rule order it could increase the performance and reduce the expected cost to lowest. They presented a better simulation framework in Aug. 2006 [17]. Now they performed a study and analysis of different firewalls. In this they invented some new methodologies to perform inspection and analysis on both multi dimension firewall rules and traffic logs. It explains the importance of traffic information in the process of firewall optimization. Simulation software was implemented to perform a study, analysis, and evaluation of multidimensional list-based firewalls.

Because of today's large organizations they have a large set of policy rule. And so it has become more difficult to manually check and correct any errors that occur due to existence of any anomalies. A method to deduce firewall policy rules by mining its network traffic log with Association Rule Mining and mining firewall log using frequency was introduced in 2006 [18]. This process can detect so many hidden and undetectable anomalies also it identifies decaying rules and dominants and treats them accordingly. It results in an anomaly free firewall rule set which is based on the dynamic network traffic logs' mining.

Osman *et. al.*, [19] has shown that in high speed networks with the help of their algorithms M-CUSUM (Multiple channel Cumulative Sum) we can detect anomalies online from a high speed network. The algorithm M-CUSUM is based on the calculation of the counter value of each bucket in the proposed reversible sketch i.e. K- ary Hash Table functions.

A new concept of use of genetic algorithms was to improve the firewall performance that has a large rule set of about thousands rules. To perform firewall policy rule set optimization is proven to be NP- hard so an optimal solution was given using binary integer program with branch-n-bound methods. As an alternative approach the use of genetic algorithms Meta heuristic adaptive search algorithm was explained, and the experimental results show that it effectively increases firewall performance.

Capretta (2007) has performed conflict detection in Coq language. They defined all the anomaly rules in Coq and then performed anomaly detection process. A proof was also shown for the correctness of the Coq algorithm, which detects all the possible anomalies in the given rule set [20].

Haakon *et. al.*, carried out an analysis about all the simulation work done and presented a comparison among all the anomaly detectors. The result shows that any of the detectors designed till now are not effective so some central requirements are defined like problem formulation, training, testing and validation [21].

V. RECENT SCENARIO

Along with automatic detection of firewall misconfigurations a new feature of dynamic routing information. It provides the complete view of the network that helps in defining optimization to improve the scalability of designed software [22].

Recently a new method of representation of policy rule set by using Direct Acyclic Graphs so that the firewall performance increases by a sorting algorithm which optimizes and reorder the policy rules to achieve a better rule set. This gives an optimal ordered rule set which gives better performance in 98% cases [23].

An open source Linux based firewall software for packet filtering which gives a complete graphical user interface for policy rule insertion, removal and to keep the rule set consistent [24].

VI. CHALLENGES

In case of firewalls whenever the rule set is static and well organized till then only we can expect the firewall to be free of anomalies. But as the size of rule set of a firewall is comparatively large for large organizations, so the rule set manipulation is a difficult task. And from the first configuration of the firewall it is updated every time a new administrator takes over. Depending on the organization if there is a frequent change in the management then security policies are also applied accordingly. Because of any of the above reasons if the policy rules are changed and depending on that the firewall rule set is also modified then it generates several anomalies whether we are inserting or deleting or updating any of the rule in the rule set. So the main challenge is not to establish a firewall but maintain it and performing desired security actions is most challenging job. Research techniques defined till now are able to detect various types of anomalies in the rule set, reordering of rule set, optimization of rule set statically and based on the behavior of the network traffic. The rule set if contains any rule which is irrelevant i.e. it does not match any of the incoming or outgoing packets along specified time duration then that rule should be identified and removed from the rule set. Such an irrelevant rule makes extra time and space overhead which decreases the firewall performance. The irrelevant anomaly is not yet properly detected and no research technique yet has described the removal method for it.

VII. CONCLUSIONS AND FUTURE DIRECTIONS

Most of the papers discussed are intended to perform the anomaly detection and removal by using different techniques. All of them consider that the rules are written in predicate like language. The policy rules have very simple attribute like fields but in some cases some firewalls define the rules with time parameters defined within the rules, and the actions performed are restricted to be only accept and deny. One more observation was carried out about the anomalies that almost no paper includes irrelevant anomaly as important one, but we observe that due to the effects of it the rule size is increased enormously.

REFERENCES

[1] Ehab S. Al-Shaer and H. Hamed. " Management and translation of filtering security policies". *In IEEE International Conference on Communications,* (ICC '03), (2003).

[2] E. Al-Shaer and H. Hamed. "Firewall Policy Advisor for Anomaly Detection and Rule Editing." IEEE/IFIP Integrated Management Conference (IM'2003), March(2003)

[3] E.S. Al-Shaer and H.H. Hamed. "Discovery of policy anomalies in distributed firewalls". *In IEEE Infocom*(2004).

[4] Al-Shaer and H. Hamed, "Conflict classification and Analysis of Distributed Firewall policies", *IEEE J SEL AREA COMM*, (2005)

[5] Chotipat Pornavalai and Thawatchai Chomsiri."Firewall Rules Analysis", *International Technical Conference on Circuits/ Systems, Computers & Comm.* (ITC-CSCC 2004), JULY(2004).

[6] Thawatchai Chomsiri, Chotipat Pornavalai: Firewall Rules Analysis, International Conference on Security & Management, SAM 2006, Las Vegas, Nevada, USA, June 26-29(2006).

[7] Deri Luca and Suin Stefano and Maselli Gaia (2003) Design and implementation of an anomaly detection system: An empirical approach. In Proceedings of Terena TNC.

[8] Y. Bartal, A.J. Mayer, K. Nissim, A. Wool, Firmato: A novel firewall management toolkit, in: Proceedings of the IEEE Symposium on Security and Privacy, (1999).

[9] Errin W. Fulp. "Optimization of network firewall policies using ordered sets and directed acyclical graphs". Technical report, Computer Science Department, Wake Forest University, (2004).

[10] Yu Gu, Andrew McCallum and Don Towsley. "Detecting Anomalies in Network Traffic Using Maximum Entropy Estimation", Tech. rep., Department of Computer Science, UMASS, Amherst, (2005).

[11] F. Cuppens, N. Cuppens, and J. Garc´ýa. "Detection and removal of firewall misconfiguration". *In International conference on Communication, Network and Information Security* (CNIS2005), Phoenix, AZ, USA, IASTED(2005).

[12] Cuppens, F., Cuppens-Boulahia, N., and Garcia-Alfaro, J. (2005). "Misconfiguration Management of Network Security Components". *In Proceedings of the 7th International Symposium on System and Information Security*, Sao Paulo, Brazil.

[13] Muhammad Abedin, Syeda Nessa, Latifur Khan, Bhavani Thuraisingham. "Detection and Resolution of Anomalies in Firewall Policy Rules". In Proc. 20th IFIP WG 11.3 Working Conference on Data and Applications Security (DBSec 2006), Springer-Verlag, July 2006, SAP Labs, Sophia Antipolis, France(2006).

[14] T. Katic, P. Pale. "Optimization of Firewall Rules". Information Technology Interfaces, (2007).

[15] V. Grout, J. Davies and J. McGinn "An Argument for simple embedded ACL optimization", *Computer Communications*, **30**(2).

[16] S. Acharya, J. Wang, Z. Ge, T.F. Znati, and A. Greenberg. Traffic-aware firewall optimization strategies. *In Proceedings of the International Conference on Communications*, (2006).

[17] Subrata Acharya , Jia Wang , Zihui Ge , Taieb Znati , Albert Greenberg, "Simulation Study of Firewalls to Aid Improved Performance", *In Proceedings of the 39th annual Symposium on Simulation*, (2006).

[18] K. Golnabi, R.K. Min, L. Khan, and E. Al. Shaer, "Analysis of firewall policy rules using data mining techniques", *IEEE NOMS 2006, Vancouver, Canada,* April (2006).

[19] Osman, S., Vaton, S. and Gravey, A. (2007). A novel approach for anomaly detection over high-speed networks. In, Proceedings of EC2ND.

[20] V. Capretta, B. Stepien, A. Felty and S. Matwin, "Formal Correctness of Conflict Detection for Firewalls", FMSE'07, ACM, Virginia, USA, Nov. (2007).

[21] Haakon Ringberg, Matthew Roughan , Jennifer Rexford, "The need for simulation in evaluating anomaly detectors", ACM SIGCOMM *Computer Communication Review*, **38**(1)**:** (2008).

[22] Ricardo M. Oliveira, Sihyung Lee, and Hyong S. Kim, "Automatic Detection of Firewall Misconfigurations using Firewall and Network Routing Policies", [PFARM'09]. *IEEE DSN Workshop on Proactive Failure Avoidance, Recovery, and Maintenance (PFARM)*, Lisbon, Portugal, Jun. (2009).

[23] Ashish Tapdiya, Errin W. Fulp, "Towards Optimal Firewall Rule Ordering Utilizing Directed Acyclical Graphs," icccn, pp.1-6, 2009. *Proceedings of 18th International Conference on Computer Communications and Networks*, (2009).

[24] A Multi Agent framework for anomalies detection on distributed Firewalls using data mining techniques in 2009 by Kamel Karoui, Fakher Ben Ftima, Henda Ben Ghezala (2009).