Anomaly Detection: Firewalls Capabilities and Limitations

Sultan Alsehibani and Sultan Almuhammadi
Information and Computer Science Department
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia
Email: (g201038620, muhamadi)@kfupm.edu.sa

***Abstract*—Firewalls are the most deployed basic security devices that are used to protect private networks from unauthorized accesses and intrusions. Firewall's security protection depends mainly on the quality of the firewall's configured policies. However, as firewalls policies grow in size, the interactions between policies of the same firewall or different firewalls become complex, which makes it difficult to design and manage firewalls policies in large scale systems. This paper identifies and compares recent firewall anomaly management frameworks, tools, and algorithms. It compares the anomaly management approaches in terms of visual representation, need for manual interference, existence of implementation, features, and limitations. It also classifies these approaches as single or distributed architectures, and the modes of these approaches as real-time or offline. Useful recommendations are provided as a result of this study.**

***Keywords*—Firewalls, Large Scale Systems, Anomalies, Detection, Resolution.**

I. Introduction

Due to the rapid growth and development of big data applications and other internet services, network security becomes an important concern. Since information security is important to all organizations and business, security of networks and computing environments is gaining the attention of researchers around the world. Firewalls are some of the main common means of dealing with network attacks in both large and small networks. Firewalls help in securing networks by preventing packets from entering the network using predefined rules and policies.

Usually firewalls are deployed on the edges of the network which provides security to the private network. However, distributed firewalls occur when every machine in the network is set into a firewall to filter packets from and into itself. Distributed firewalls are gaining more attention as a good security measure.

As firewalls evolve to match the growing scale of today's systems for big data applications, it is important to address the limitations of typical firewalls. These include firewall anomalies as they affect the security of the network, and the efficiency of the firewall. In [1], the authors classified firewall anomalies into two types inter-firewall anomalies and intra-firewall anomalies. Inter-firewalls anomalies are anomalies that occur between different firewall device, while intra-firewall anomalies are anomalies within the same firewall device.

An example of a firewall policy is illustrated in Figure I. As networks grow, firewall policies are extended to include hundreds of rules. Thus, making the task of managing the rules a daunting task. However, multiple tools and frameworks have been developed to detect and resolve firewall anomalies. In this paper, we provide an overview of different recent firewall anomaly management tools and frameworks. We will also highlight the features and limitations of each tool, framework, or approach.

Rule	Proto	SIP	SP	DIP	DP	Action
1	TCP	10.1.2.*	*	192.168.*.*	80	Deny
2	UDP	10.1.*	*	192.168.1.*	25	Allow
3	UDP	*	*	192.168.2.1	25	Deny
4	*	10.1.1.*	*	*	*	Allow
5	*	*	*	*	*	Deny

Figure 1. Example of Firewall Policy

Today's systems complexity makes it hard to depend only on firewalls to ensure security of large systems. Sophisticated intruders may bypass firewall rules easily. With the rapid increase of big data applications, it is important to deploy machine learning techniques and other dynamic anomaly detection approaches besides firewalls.

This paper reviews recent firewall anomaly detection tools and frameworks. It shows their limitations of detecting anomalies especially in large scale systems. Section II provides a literature review of the work done in firewall anomaly management. Section III discuses the firewall anomaly classifications. Section IV presents the different types of firewall anomalies. Section V discusses firewall detection tools, frameworks, and approaches focusing on their features and limitations in big data applications. Section VI summarizes firewall resolution tools and highlights their features and limitations. Finally, we discuss the results of this study and provide useful recommendations in Section VII.

II. Literature review

Designing and analyzing firewall policies requires tremendous effort to ensure the correctness of these policies. Performing this task properly involves solving a number of firewall problems, such as firewall verification, equivalence, and completeness problems. First, the firewall verification problem aims to ensure that a firewall accepts or discards a given set of packets. Another problem is the firewall equivalence problem, which shows that two walls accept or reject the same set of packets. Furthermore, the firewall completeness is a problem,

which ensures that a firewall accepts or rejects every possible packet.

Elmallah et al. [2] showed 13 different problems that need to be solved to logically analyze firewall policies. Moreover, the authors evaluated the hardness of these 13 problems and proved them to be NP-hard problems by reduction from the 3-SAT problem. Moreover, even after ensuring the correctness of firewall policies, firewall anomalies might occur.

Anomalies can be classified into two categories: (a) intra-firewall anomalies, which are the conflicts within the rules of the same firewall, and (b) inter-firewall anomalies, which are the conflicts that occur between different firewall devices within the same network. We observe that some tools only detect intra-firewall anomalies, whiles others detect anomalies in both categories.

Hu et al. [3] introduced a tool to identify intra-firewall anomalies and derive effective anomaly resolutions using a rule-based approach. The tool provides a grid-based visual representation, which establishes a cognitive sense regarding the detected anomalies.

Anomaly management framework for firewalls has been developed in [4], [5], [6]. The authors used a segmentation-based technique to effectively identify intra-firewall anomalies. The approach identifies firewall anomalies by dividing each packet in a network space into disjoint segments then identifying relationships among different segments.

Abbes et al. [7] proposed a quick method for managing a firewall configuration file. The authors represented the set of filtering rules using a firewall anomaly tree (FAT). The FAT modification by the system administrator, automatically reveals emerged anomalies and helps the administrator to find the adequate position for a new added filtering rule. The authors showed implementations, and computer experiments for all the presented algorithms.

Inter-firewall anomalies have been addressed by the following papers. An automated approach proposed in [8] for checking that distributed firewalls react according to a security policy specified in a high level declarative language. If errors are detected, some useful feedback is returned in order to correct the firewall configurations. The authors showed that the approach can be implemented using a Satisfiability Modulo Theories (SMT) solver named Yices.

Sultana et al. [9] introduced an approach to eliminate inter-firewall redundancies in cross domain firewall systems while maintaining the privacy of each firewall system. The authors also showed an experimental evaluation of the provided approach.

Three algorithms that detect firewall anomalies in Cisco devices are given in [10]. The authors suggested different possible resolution strategies for firewall anomalies detected by the implemented system.

Firewall anomalies are not only restricted to networks based on IPv4, they also extend to IPv6 networks. In [11], the authors investigated a model checking techniques for automated policy anomaly detection. The authors showed that with a few adoptions existing approaches can be extended to support the IPv6 protocol with its specialties like the tremendously larger address space or extension headers. The performance is evaluated empirically by measurements with a prototype implementation.

III. Firewalls anomalies classifications

Classification of firewall policies was first introduced in [1] and later extended in [12] When there is a conflict between the firewall rules, meaning a packet matches two rules or more, it is said that a firewall policy anomalies. Firewall anomalies do not affect the policy by incorrect packet filtering but it also wastes the time needed to scan the firewall policies and it wastes space that is used to store Access Control List (ACL). The most implemented match policy in firewall is the first match policy, where a device scans the firewall policies sequentially and in order. Therefore, as the number of firewall anomalies increase the time needed to scan the policies will increase. In general firewall anomalies can be classified as intra-firewall anomalies, and inter-firewall anomalies.

A. Intra-Firewall Anomalies

Intra-Firewall Anomalies can be described as the conflicts that occur within the same firewall device. In other words, if a packet matches two different rules within the same firewall policy, then it is an intra-firewall anomaly. However, a single firewall device nowadays can have multiple interfaces. Therefore, the classification of intra-firewall anomalies can be further broken down to intra-intreface and inter-interface anomalies based on the presence or absence of multiple interfaces.

1) Intra-Interface Anomalies: If a conflict arises within the same interface then an intra-interface anomaly is present. In other words if a packet matches more than one rule in an ACL then an intra-interface is present. In addition, intra-interface anomalies can be present if a packet matches more than one rule in different ACLs assigned to the same interface.

2) Inter-Interface Anomalies: If there exist a conflict between different ACLs assigned to different interfaces within the same firewall then an Inter-interface anomaly is present. For example, if a packet can be allowed to access the network in one ACL assigned to one interface however the same packet will be denied in the other interface within the same firewall.

B. Inter-Firewall Anomalies

If there is a conflict between different rule in different firewalls within the same network then an inter-firewall anomaly is present. An example of an inter-firewall anomaly is when a packet is routed through two different firewalls on its path to its destination and one firewall allows the packet and the other one denies it. Inter-firewall anomalies can happen with firewalls within the same sub-network or between devices in different sub-networks. This type of anomalies is more common in large scale systems.

IV. Firewall anomalies

There are different types of firewall anomalies namely shadowing, correlation, redundancy, generalization, and irrelevance anomalies. This section briefly defines these anomalies.

- **Shadowing Anomaly** Shadowing occurs when all the rules that match a later rule, match a preceding rule. Thus, the later rule will never be a match.
- **Correlation Anomaly** Correlation occurs when two rules with different actions share a set of packets.
- **Redundancy Anomaly** Redundancy occurs when two rules with the same action match a set of packets. Thus, if one of the rules was remove the security policy will not be affected.
- **Generalization Anomaly** Generalization occurs when a preceding rule matches packets that match a later rule and the rules perform different actions.
- **Irrelevance Anomaly** Irrelevance occurs when a rule in a firewall does not match any packet at a give time interval. This can happen if both the source address and the destination address fields of the rule do not match any domain.

V. Firewall Anomalies Detection Approaches

A lot of work has been done to analyze firewall's policies and to detect firewall's anomalies. This section provides a summary of the main features of recent firewall anomalies detection approaches.

Al-Shaer et al. [1] proposed a tool to detect firewall anomalies called Policy Advisor. The firewall Policy Advisor uses a state machine approach to identify anomalies in firewall rules. The Policy Advisor linearly search the firewall rules comparing every two rules to each other. The main limitations of this rule are it works in small networks as it compares rules in pairs, and it requires manual insertion of rules into the tool. Due to this limitation, the Policy Advisor approach is not recommended for large scale systems.

A tool called Fireman [13] is developed to detect firewalls anomalies. Fireman overcomes the pair-wise limitation of the firewall Policy Advisor [1] by analyzing the all firewall rules together instead of pairs. The main limitation of Fireman is it cannot handle dynamic firewalls as it adapts a static analysis approach. This approach clearly does not fit big data applications where dynamic analysis is required.

Souayeh et al. [8] focused on checking whether a distributed firewall configuration conforms to a given security policy. Security policies can be specified in an expressive enough declaration language. After that, a priority-based approach is used to determine whether a distributed firewall configuration conforms to a security policy. This approach evaluates the security policy against six conformity theorems. Moreover, it shows that these conformity theorems can be evaluated automatically using a satisfiability solver modulo theories (Yices). Furthermore, the priority-based approach ensures avoidance of firewall anomalies. The main limitations of this approach are the security policy must be specified prior to analyzing firewall policies, and the approach cannot handle dynamic networks, and hence, it is not recommended for big data applications.

Hu et al. developed a tool named FAME [3]. The authors identified two approaches to represent firewall anomalies: packet space representation, and grid representation of anomalies. The paper introduced a rule-based techniques which utilizes a binary decision diagram data structure (BDD) to represent rules. Then, the list of rules are converted into a set of disjoint network packet spaces by performing various set operations. The main limitations of this approach are it operates on single firewalls, and it cannot handle dynamic firewalls.

Access Control Lists (ACL) Scan [10] is a tool that detects firewall anomalies from security devices in real-time by scanning ACLs. The current solution works with Cisco devices however, it can be extended to work with other vendors devices. The ACL Scan tool needs at least one IP address of each configured device in the network . After that a recursive algorithm is used to find all the links between the devices. The intra-firewall policy is determined by comparing each firewall rule the preceding rules. For inter-firewall policies, first the intra-firewall anomaly is used. then the algorithm searches for remaining interfaces of other devices. finally, the algorithm checks for interfaces in the entire network that is in the same sub network in a similar way to the second step. Moreover, Ameya's work detect the following anomalies in firewall's policies shadowing, exact match, and correlation anomalies. Although the algorithm can be extended to any firewall device the current implementation of the tool is restricted to Cisco devices

In [4], [5], [6], the authors adapted a rule-based segmentation technique to identify relationships among the firewall rules. The system after that notifies the administrator of malicious activity. Using packet space segmentation technique firewall anomalies are detected. After that, the risk of anomalies is assessed. Risk assessment is measured using an upper bound and lower bound threshold values. Based upon the risk, the firewall rules can be re-ordered. Vanikalyani et al. [4] showed that 63 percent of the conflicts has been resolved. However, after using the inter-firewall optimization in [5], 92 percent of the conflicts were resolved. Furthermore, this approach detects the generalization, redundancy, correlation, and specialization anomalies. Nagpure et al. [6] showed no significant improvement to the rule-based segmentation technique.

Sultana et al. [9] concentrated on resolving cross domain firewalls anomalies without violating the privacy of each firewall. The authors suggested that the firewall configurations are turned in an equivalent set of firewall rules that are not the actual firewall rules. Afterwards, each firewall can compare the set of firewall rules with the other based on a suggested decision tree modulation. Sultana's work can detect redundancy anomalies only. However, detecting redundancy anomalies is extended to cover cross domain firewalls.

Tree structure approach is used in [7] to detect firewall anomalies. This approach provides the administrator with the ability to verify if the placement of a new rules introduces any conflicts.

Table I summarizes our review on recent anomaly detection approaches. For each approach, we list and compare the following attributes:

- **Type:** Intra-firewall anomalies / Inter-firewall anomalies
- **Mode:** Offline / Real-time
- **Representation:** Visualization of anomalies

Table I. MAIN FEATURES OF FIREWALL ANOMALY DETECTION APPROACHES

Approach	Type	Mode	Repres.	Arch.	Manual	Implem.	Features	Limitations
Policy Advisor[1]	Intra Inter	Offline	-	D	Yes	Yes	State Machine detection Pairwise anomaly detection Search is linear	Small networks only Manual insertion of rules
Fireman [13]	Intra Inter	Offline	-	D	No	Yes	Compares all rules in FC Static Analysis of firewall policies	No dynamic networks
Priority [8]	Intra Inter	Offline		D	No	No	Security policy in high level language Verification of security policy Conflict avoidance	Must specify security policy first No dynamic networks
FAME [3]	Intra	Offline	Grid	S	No	Yes	Visualization	No distributed firewalls No dynamic networks
ACL Scan [10]	Intra Inter	Real-time	-	D	No	Yes	Dynamic networks implementation Real-time implementation Multiple interfaces devices	Cisco devices only
Segmentation [5], [4], [6]	Intra	Offline	-	S	No	No	Calculates the risk level for each segment	No distributed firewalls no dynamic networks
FAT [7]	Intra	Offline	Tree	S	Yes	Yes	Verification of rule placement	No distributed networks Static analysis

- **Architecture:** Single firewall / Distributed firewalls
- **Manual Interference:** Boolean to indicate if manual interference by the system administrator is needed.
- **Implementation:** Boolean to indicate if the approach is implemented.
- **Features:** Lists the features of the approach
- **Limitations:** Lists the limitations if the approach.

VI. FIREWALL ANOMALIES RESOLUTION

As the work continues on analyzing and detecting anomalies within firewall policies. It is important to find resolutions for the detected anomalies. The following section provides a summary of the work done in resolving firewall anomalies.

FAME [3] resolves anomalies by assigning a conflict constraint on each conflicting segment. A conflict constraint would be either to accept or deny a packet belonging to that conflicting segment group, which eliminates the need of rule reordering.

ACL Scan [10] resolves anomalies by either deleting or modifying one of the rules. For Exact match the system will notify the system administrator to delete one of the rules. For correlation anomaly warning, three solutions are available deleting one of the existing rules, modifying one of the rules, or adding a specialized rule for permitting/denying the traffic under consideration in both the devices.

Finally, FAT [7] does not directly offer a firewall resolution approach. However, by inspecting the leaf nodes in FAT anomalies can be prevented. Table II summarizes the features and limitations of anomaly resolution approaches.

Table II. FEATURES AND LIMITATIONS OF ANOMALY RESOLUTION

Approach	Features	Limitations
FAME [3]	Conflict constraint Rule reordering is not needed	No distributed firewalls No dynamic networks
ACL Scan [10]	Delete exact match Three options for correlated rules	Cisco devices only
FAT [7]	Verification of rule placement Anomaly prevention	No distributed networks Static analysis

VII. DISCUSSION AND RECOMMENDATIONS

In this section, we discuss the recent approaches in firewall anomaly detection and provide our recommendation on each approach. Moreover, we present recommendations for future anomaly detection tools and a methodology to study this subject more effectively.

A. Recommendations on Existing Approaches

In the review of recent algorithms and approaches to detect and resolve firewall anomalies, we find that there is no generic approach to detect and resolve firewall anomalies in real-time for distributed firewalls. However, some of the reviewed approaches are recommended for certain scenarios. These approaches can be further improved to create more generic tools.

The ACL scan approach is a very recommended tool since it detects and resolves firewall anomalies in real-time for distributed firewalls for Cisco devices. However, it has no visual representation and it works only with Cisco devices. Thus, we recommend the extension of the ACL scan approach to be a generic approach that works with devices from other vendors and provide it with visual representation.

The FAME is recommended for single firewall architecture. It detects and resolve intra-firewall anomalies and provides grid visualization. However, it needs to be extended to work in real-time with dynamic network and distributed firewalls.

Like FAME, FAT is recommended for single firewall. It provides verification and simulation of rule placement. However, it requires manual interference which makes it inconvenient. Thus we recommend automating the tool for convenience and simplicity. Moreover, it needs to be extended to work in real-time with dynamic network and distributed firewalls.

B. Recommendations for Future Tools and Methodologies

We observe that all aforementioned approaches rely on rule matching and statistical techniques. The main challenge facing these anomaly detection approaches remains unsolved since attackers usually change their intrusion techniques. These

approaches simply do not learn from their mistakes and need continuous adjustment or reconfiguration. By "their mistakes" we mean their false negatives and false positives, which are essential feedback information for any anomaly detection system to improve its performance. We suggest that future anomaly detection for large scale systems should provide a combination of rule matching and machine learning techniques that are based on analysis of dynamic datasets of network traffic.

Moreover, we recommend a methodology to study anomaly detection for big data application. We suggest that any study should include a combination of theoretical analysis, and empirical evaluation. The theoretical study includes the analysis of network anomalies as well as statistical profiling, clustering, and behavioral approaches using graph modeling and machine learning techniques. While, the empirical evaluation includes an experiment on large datasets, typically provided by concerned organizations. Other alternatives to this methodology would be either a pure theoretical study or a practical experimental study. We suggest avoiding these alternatives since the former may not reveal the real problems faced by network security engineers, while the latter is ineffective due to the various facets of the problem.

VIII. Conclusion

Firewall and network security require proper maintenance and management to provide adequate level of security for big data applications. Installing the firewalls and configuring them does not necessarily make the system secure. One important reason is that managing firewalls rules can be very complex, especially in large scale systems. Furthermore, network vulnerability may result due to firewall rule anomalies. We reviewed and discussed recent firewall anomaly management systems and gave useful recommendations and propose a methodology to study anomaly detection in today's large systems.

Future work includes extending the ACL scan approach to a generic approach such that the syntax will not be exclusive for Cisco devices only. FAME and FAT need to be extended to work in real-time with dynamic network and distributed firewalls. Furthermore, graph modeling and machine learning techniques should be used in anomaly detection systems whenever big data is involved. To sum up, this review of firewalls capabilities and limitations opens great research directions for anomaly detection in large scale systems.

References

[1] E. S. Al-Shaer and H. H. Hamed, "Discovery of policy anomalies in distributed firewalls," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4. IEEE, 2004, pp. 2605–2616.

[2] E. S. Elmallah and M. G. Gouda, "Hardness of firewall analysis," in *Networked Systems*. Springer, 2014, pp. 153–168.

[3] H. Hu, G.-J. Ahn, and K. Kulkarni, "Detecting and resolving firewall policy anomalies," *IEEE Transactions on Dependable and Secure Computing*, vol. 9, no. 3, pp. 318–331, 2012.

[4] G. Vanikalyani, P. Avinash, and P. Pandarinath, "Cross-domain search for policy anomalies in firewall," *International Journal of Computer Applications*, vol. 104, no. 6, 2014.

[5] S. Kachare and P. Deshmukh, "Firewall policy anomaly management with optimizing rule order," *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*, vol. 4, no. 2, pp. 201–205, 2015.

[6] R. Nagpure, P. J. Dhuri, J. K. Patil, M. P. Kini, and A. J. Patil, "Detection and resolution of firewall policy anomalies," *The International Journal of Science and Technoledge*, vol. 3, no. 2, p. 59, 2015.

[7] T. Abbes, A. Bouhoula, and M. Rusinowitch, "Detection of firewall configuration errors with updatable tree," *International Journal of Information Security*, pp. 1–17, 2016.

[8] N. B. Y. B. Souayeh and A. Bouhoula, "Formal checking of multiple firewalls," *arXiv preprint arXiv:1207.3691*, 2012.

[9] A. R. Sultana and A. Kavarthapu, "Resolving cross domain firewall policy anomalies," *International Journal of Computer Applications*, vol. 124, no. 14, 2015.

[10] A. Hanamsagar, B. Borate, N. Jane, A. Wasvand, and S. Darade, "Detection of firewall policy anomalies in real-time distributed network security appliances," *International Journal of Computer Applications*, vol. 116, no. 23, 2015.

[11] C. Lorenz and B. Schnor, "Policy anomaly detection for distributed ipv6 firewalls," in *e-Business and Telecommunications (ICETE), 2015 12th International Joint Conference on*, vol. 4. SCITEPRESS, 2015, pp. 210–219.

[12] A. Hanamsagar, N. Jane, B. Borate, A. Wasvand, and S. Darade, "Firewall anomaly management: A survey," *International Journal of Computer Applications*, vol. 105, no. 18, 2014.

[13] L. Yuan, H. Chen, J. Mai, C.-N. Chuah, Z. Su, and P. Mohapatra, "Fireman: A toolkit for firewall modeling and analysis," in *2006 IEEE Symposium on Security and Privacy (S&P'06)*. IEEE, 2006, pp. 15–pp.