IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received April 6, 2021, accepted April 18, 2021, date of publication May 3, 2021, date of current version May 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3077146

HSViz: Hierarchy Simplified Visualizations for Firewall Policy Analysis

HYUNJUNG LEE[1], SURYEON LEE[2], KYOUNGGON KIM[3], (Member, IEEE), AND HUY KANG KIM[4], (Member, IEEE)

[1]Korea Securities Computer Corporation, Seoul 07329, South Korea
[2]Seoul Women's University, Seoul 01797, South Korea
[3]Department of Forensic Sciences, Naif Arab University for Security Sciences, Riyadh 14812, Saudi Arabia
[4]School of Cybersecurity, Korea University, Seoul 02841, South Korea

Corresponding author: Huy Kang Kim (cenda@korea.ac.kr)

ABSTRACT Most of the companies have firewalls in order to protect their internal networks and assets from the attacker of the cyber space. Firewall policies should be maintained and organized with high importance. However, considering the length of time needed in analyzing the highly complex policies and the risks of disabling firewall that may arise in case of a false policy setting. It is extremely hard to securely optimize the performance of firewalls. This paper is to suggest a visualization tool that shows the status and the types of policies applied throughout the firewalls so that such difficulties related to the maintenance of firewall policies can be resolved. The proposed tool is designed in six different angles; (1) Hierarchy-view, (2) Anomaly-view, (3) Distributed-view, (4) ANYPolicy-view, (5) SearchResult-view, and (6) Top and Bottom Used-view. The core of the overall function is to facilitate the easy identification of the policy interrelationships. The visualization tool has been tested by being applied across approximately 24 different firewall policies. The processing speed of each function and abuse detection rate were all reviewed positively. By the help of the tool, identifying the services, performance improvement, and visibility of the policy relations, which thereby will lead to better safety in preserving the assets intact. A video of the proposed visualization tool can be found on the web site: https://youtu.be/43OfHN8dteU

INDEX TERMS Firewall policy visualization, policy analysis, data visualization, rule anomaly detection.

I. INTRODUCTION

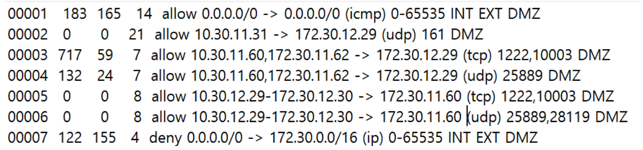
Firewall is regarded as the most basic security system that has to be equipped alongside other infrastructure such as servers and etcetera. Cisco, Palo-Alto Networks, Fortinet, Check-Point, and Symantec are the most representative firewall manufacturers. Companies no longer question the need to introduce a firewall and this implies the fact that the level of understanding with regards to cyber security has been matured compared to the previous days. In other words, this also means that firewall has been settled as the most basic and essential security product for all corporations.

However, as the same firewall is operated for an extended period of time, the problems related to the firewall goes accumulated on and on. The greatest problem of all is the increasing level of complexity in firewall policies when they are operated without being managed. What this entails are as follows: First is difficulty in identifying relevant services related to firewall policies. The more the policies are added, the more difficult it becomes to know the services to receive and the personnel to contact for a help when it comes to a malfunctioning firewall or a need for a firewall replacement. Second is degradation in firewall performance. A firewall may become overloaded with unnecessary, redundant, misused, and unused policies. As the number of policy lines increases, more references have to be made when the firewall operates, which in turn leads to low performance. Third is inability to gain visibility over authorized services by the firewall. It becomes difficult to discern impermissible policies and vulnerable services as the policies become too much complicated while being intertwined in so many different lines.

To solve this problem, there have been many calls for the optimization of firewall policies since far long ago. Nevertheless, there is a risk of causing firewall failures by deleting a policy mistakenly or by applying an irrelevant policy. So it is not easy to initiate a policy change from a firewall operator's

The associate editor coordinating the review of this manuscript and approving it for publication was Ilsun You.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/



```
00001  183  165  14  allow 0.0.0.0/0 -> 0.0.0.0/0 (icmp) 0-65535 INT EXT DMZ
00002  0    0    21  allow 10.30.11.31 -> 172.30.12.29 (udp) 161 DMZ
00003  717  59   7   allow 10.30.11.60,172.30.11.62 -> 172.30.12.29 (tcp) 1222,10003 DMZ
00004  132  24   7   allow 10.30.11.60,172.30.11.62 -> 172.30.12.29 (udp) 25889 DMZ
00005  0    0    8   allow 10.30.12.29-172.30.12.30 -> 172.30.11.60 (tcp) 1222,10003 DMZ
00006  0    0    8   allow 10.30.12.29-172.30.12.30 -> 172.30.11.60 (udp) 25889,28119 DMZ
00007  122  155  4   deny 0.0.0.0/0 -> 172.30.0.0/16 (ip) 0-65535 INT EXT DMZ
```

FIGURE 1. **Firewall policies in text mode.**

perspective. There have been many attempts to visualize the firewall policies to find a way to cope with their growing complexities. This paper aims to solve the aforementioned problems associated with policy complexities through the means of firewall policy visualization system. In particular, HSViz (Hierarchy Simplified Visualizations for Firewall Policy Analysis) provide better visibility and help users be more intuitive in discerning the status of firewalls in concern. The contributions of our paper are summarized as follows.

- Hierarchical views based on IP octets
- Anomaly policy case views under single and distributed firewall environments
- Six signatures of ANY allow policies
- Search result and unused policy views

The rest of the paper is organized as follows. Section II explores the studies undertaken in relation to firewall policy visualization, and Section III and IV explain the detailed functions of the HSViz, a firewall policy visualization tool. Section V explores how the tool is used. Section VI identifies the results of the firewall policy tests in practice. Section VII presents conclusions and future tasks.

II. BACKGROUND

A. MOTIVATION FOR RESEARCH

Firewall is a device that allows or denies packet that passes through the firewall based on the policy with predetermined source and destination. As the services that passes through the firewall become more diversified, the number of lines of the firewall policy increases and the interrelationship becomes more complicated. Figure 1 illustrates the firewall policy screen that has been set actually. Since it provides an extensive range of information in text format, it is difficult to assess the policy at a glance and analysis of the relationship between the policies is not easy.

Although there are currently programs provided by each firewall manufacturer for inquiries on the policies, it is difficult to make inquiries on all the information that operator wants easily and promptly or confirm their association relationship at a glance. In addition, firewall policies are confirmed in the method of extracting values that match specific factor value for each line through options such as 'inc' and 'grep' in the entire range of policies at the time of policy inquiry, although there are differences between firewall manufacturers. Such confirmation method is useful in the event of quickly confirming desired fragmental information including confirmation or making inquiries on the existence of the policy by extracting the firewall policy through the line directly under the system console by experts capable of catching the relevant information proficiently with familiarity in the firewall operation. However, intuitive policy inquiry method has limitations in easily recognizing policies that must not be allowed for security purpose by assessing the mutual relationship of firewall policies or difficulties in assessing the misuses of firewall policies that can occur in accordance with diversified conditions with the cognitive viewpoint of human.

If the number of policies set for the firewall is small, the work for confirmation of the simple value of the policy by the operator can be easy. However, as the number of the text lines of the firewall policy increases, the level of complexity increases and its limitations become clearer. Limitations in the cognition of human exist in the execution of the task for finding vulnerable policies by human, which can induce mistakes of missing the important aspects. As such, means of visualization is presented in order to reduce the mistakes that human can induce and enable recognition of information containing better meaning at a glance by supplementing such problems.

B. NEED FOR VISUALIZATION

Human is not capable of processing all the information inputted from outside simultaneously. The phenomenon of selectively ignoring other information in order to allocate the limited cognitive processing resources of human only to information one is interested in is referred to as selective attention [1]. The visual system of human moves that attention of human to domain that is relatively more pronounced in comparison to those in the surroundings or domain or subject that is preferred or targeted through unconscious or conscious actions. The invisible gorilla is a representative experiment conducted by Chabris and Simons [2] who are cognitive psychologists. It experimented on the phenomenon of majority of people not realizing the appearance of gorilla among the people as they are concentrating on the movement of the ball even if they are seeing the gorilla visually. Based on such theory, the human brain has difficulty in discerning the desired important information if it receives information in text format with no visual prominence or change. Therefore, there is a need for the method of assisting the selective attention capabilities of human in order to quickly convey important information. For such purpose, text can be expressed prominently or visualized data such as diagrams, etc. can be used, which is the reason for the need for visualization. As there is a saying that 'A single picture speaks more loudly than thousands of logs', visualized data such as diagram can deliver large quantities of information at once. Moreover, greater quantities of information can be conveyed more effectively since it is possible to contain information by using different shapes, colors and size allocation, etc. Colin Ware [3] asserted that visualization is important as the vision of human is in itself a highly powerful and precise pattern analyzer and that data in enormously broad range is delivered to the cognition system of human through the operation of the eyes and visual cortex of the cerebellum as an enormous parallel processor.

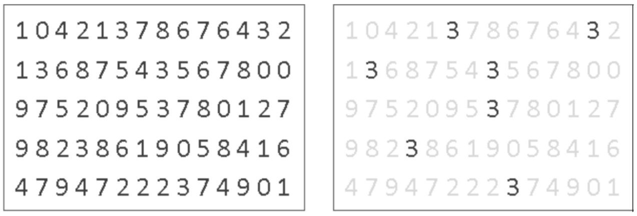


FIGURE 2. Visualization effect.

Figure 2 illustrates an example of the selective attention capabilities of human. Although it is difficult to find '3' in the left diagram, it can be found easily in the right diagram the provided difference in prominence through expression with color with different intensity [4].

Visualization in the area of security is a process of production of the contents of texts such as log, etc. through processing thereof, and defines who text data will be expressed visually. Since visualized data can deliver large quantities of information at once, there is a particular need to quickly recognize important information in the area of computer security for which data increases more explosively that any other areas. As atypical data are being produced explosively, there is increase in the interest on analysis and visualization of complex information. We can obtain answers for the problems that could not be resolved through visualization and induce new questions as well as execute investigation and discovery, and provide support for decision making. In addition, it is possible to deliver information, enhance efficiency and obtain inspiration. This paper presents diversified visualization techniques that can assist with quick and effective cognition of key information from large scale atypical data on the basis of the fact that it is difficult to recognize meaningful information from the firewall policy text line.

C. FIREWALL POLICY OVERVIEW

Firewall composes a barrier between a trusted network and the outside network to control network traffics that passes based on an already defined security principle. Firewall either allows or blocks the access to a starting point or a destination point based on the already defined policies. Firewall policies typically consist of the following elements as shown in Table 1: Order, Protocol Type, Source IP address (SIP), Source Port (SPort), Destination IP (DIP), Destination Port (DPort), Action: [5]–[9] The firewall policy is distinguished by the value of each element. Source Port is not used as firewall policy components because it is randomly assigned. Order represents the order in which the firewall references, and the lower the number, the higher the priority.

In this paper, we name firewalls as FW, firewall A as FW [A], and firewall B as FW [B]. Each line of policies set up in the firewall is labeled Rule x, Rule y, and the firewall policy components are labeled Rx [order], Rx [action].

The relationship between firewall policies can be divided into 5 main categories. Inclusively Matched (IM), Exactly Matched (EM), Partially Match (PM), Complete Disjoint (CD), and Correlated (C) as shown in Figure 3. Inclusively Matched (IM) is a firewall policy that is part of another firewall policy. If Rule x (Rx) and Rule y (Ry) are IM, the entire Rx belongs to part of Ry. Exactly Matched (EM) is when the two firewall policies are completely matched. If firewall policy Rx, Ry is EM, the components of Rx and Ry policy are fully matched. Partially Match (PM) is when the two firewall policies overlap. If the firewall policy Rx, Ry is PM, then some parts of Rx and Ry overlap each other. Completely Disjoint (CD) is a case where the two firewall policies are completely different from each other without being nested together. A Correlated (C) relationship is when two firewall policies overlap each other, and the policy components of the overlapping areas differ.

TABLE 1. Policy example of firewall A (FW [A]).

Order	Protocol	SIP	DIP	DPort	Action
1	TCP	140.192.37.2	161.120.33.40-46	25	deny
2	TCP	140.192.37.0-10	161.120.33.42	20	deny
3	TCP	140.192.37.0-10	161.120.33.41	25	allow
4	TCP	140.192.37.1	161.120.33.41	25	deny
5	TCP	140.192.37.3	161.120.33.43	21	allow
6	TCP	140.192.37.0-10	161.120.33.43	21	deny
7	TCP	140.192.37.1	161.120.33.42	20	deny
8	TCP	140.192.37.0-15	162.120.33.43	50	allow
9	TCP	140.192.37.0-10	162.120.33.43	50	deny
10	TCP	161.120.33.5-7	140.192.37.0-10	21,80	deny
11	TCP	161.120.33.0-10	140.192.37.4-7	25	allow
12	TCP	161.120.33.5-7	140.192.37.4-7	21,80	allow
13	TCP	161.120.33.5-7	140.192.37.1-2	25	deny

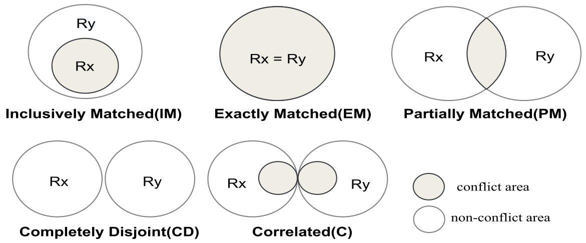


FIGURE 3. Firewall policy relation [10].

There are four categories of firewall misuse policies, as shown in Figure 4. If the SIP, DIP, and DPort of the two policies are mutually inclusive, and the action and policy order are different, it is classified as Shadowing or Generalization. Redundancy is when the action is the same and the SIP, DIP, and DPort are inter-inclusive. If the actions are the same and have a partially inclusive relationship between policies, it is classified as Correlation.

D. RELATED WORK

The area of network and service management, and information visualization underwent rapid and extensive growth during the latter part of the '80's and early '90's [11]. Research related to firewall visualization among various types of data sources related to the area of network security can be divided largely into research on the analysis of firewall event logs and on analysis of policies set for the firewall. An extensive range of visual analysis research was conducted on the methods of analyzing firewall policies, which are the firewall setting data, as well as for area of visualizing the firewall policy regulations for detecting abnormal signs, and firewall policy

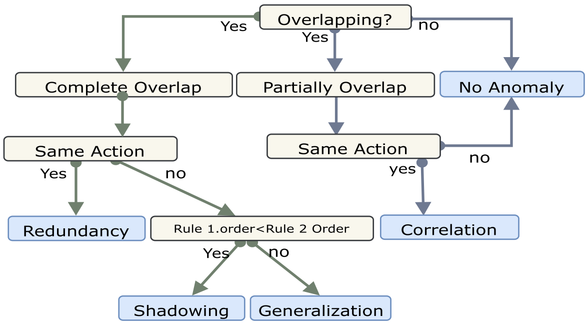


FIGURE 4. Firewall anomaly detected policy classification [8].

itself or abnormal regulations including presentation of the solutions thereof.

1) DETECTION OF ANOMALY POLICY

Al-Shaer *et al.* [5]–[10] pursued studies on the detection of anomaly policy. They presented the algorithm for detection by categorizing anomaly policies into four types, shadowing, generalization, redundancy, and correlation and detected abnormal sign policies by using Policy Advisor. Tung *et al.* [12] proposed firewall policy visualization tool referred to as PolicyVis that visualizes the misused policies under single multiple-firewall environment. PolicyVis expresses visualization by using three types of fields among the elements that compose firewall policy, namely, Source IP, Destination IP and Destination Port. X-axis and Y-axis was set as the destination and source, respectively, in 2D graph formation with expression of port information as points. Misuse policies are divided into four types for visualization. Illustrated blocking and granting policies that have been entered abnormally. It is expressed in 3-stage hierarchy structure with the bottom hierarchy displaying the physical topology map of the entire network, middle hierarchy generating and displaying logical firewall topology, and the top hierarchy displaying blocking and granting policies that have been applied abnormally, which is the most important aspect of the firewall policy. Khan *et al.* [13] defined misused cases by composing the policies in tree format. They visualized the misuse cases through 2D box model and realized the tool for detection of abnormal policies in tree format referred to as PolicyVisor.

2) VISUALIZATION OF FIREWALL POLICIES

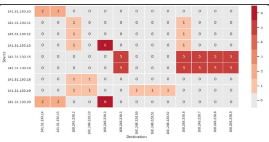
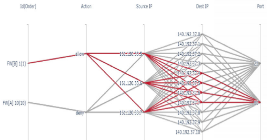
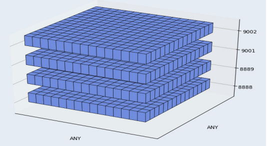
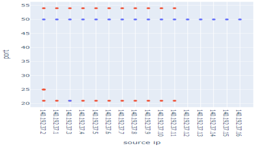
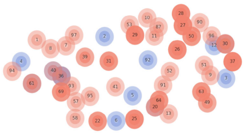
Kim *et al.* [14] realized visualization of firewall policy in the 6-dimensional space in 3D format through the tool referred to as FRuVATS. It is possible to express all the sources and IP in all ranges, and displayed the identified results by analyzing policies for which collision occurred, identifying activated and inactivate domains, and by using visualization model. Starting point, destination and order were illustrated in the X,Y and Z-axes, respectively, and granting, refusal, activation and inactivation policies were distinguished and expressed by using different colors. It has the advantage of being able to express the entire IP address space. Kim *et al.* [15] used the tool referred to as Firewall Policy Checker (FPC) to assist the firewall administrator to execute firewall policy inspection easily and detect hazardous services and illegal servers such as telnet and TFP, etc. Moreover, mutual correlation among them were expressed by using expression of starting point, destination and port, and lines in the format of globe. In addition, this tool visually expressed collision domains of the selected regulations once the 2 regulations related to the abnormal regulations of the firewall are selected. PortVis [16] visualizes the activity of more than 65,000 application ports. The tool provides a time series visualization to investigate port activity over time. Keim *et al.* [17] attempted to analyze packet-level network data flows in Radial Traffic Analyzer. Traffic generated by network hosts is represented radially in HistoMap. Mansmann *et al.* [18], [19] proposed a tool called VISUAL FIREWALL EDITOR that expressed the current status, entities and groups of firewall policy access control list (ACL) by using Sunburst Chart. It is characterized by the ability to confirm the host entity included with the network and group entities as the reference, thereby visualizing the current status of hit count, which is the number of the use of policies, by using the sunburst chart as well. Created Voids, a tool proposed by Morrisy *et al.* [20] distinguishes and expresses the allowed and not allowed domains in the policy by utilizing parallel coordinate graph by analyzing the firewall policies.

3) RESOLVE FIREWALL ANOMALY POLICIES

Hu *et al.* [21] expressed the space derived from the firewall regulations in 2D and defined the overlapping space with misuse policies by dividing the separated space into multiple dimensions. A solution for the misused policies is presented by recombining the space divided into multiple dimensions. For this purposed, policy analysis tool called FAME was created to identify the abnormal policies and expressed the abnormal policies by using regulation-based partitioning technique. Saâdaoui [22] proposed management framework for effective deduction of solution through the technique for identification of misuse policies of firewall and regulation-based partitioning technique. It extracts whether it is an actually erroneous composition or intended composition under the dispersion environment by using FDD decision-making diagram. It analyzed the route and extracted the erroneously composed policies through regulations, and presented efficient solutions for the policies for which error exists.

HSViz's six visualization models have differentiated features from other papers. (1) Hierarchy-view represents policy ranges based on destination IP octets at the user's choice. (2) Anomaly-view and (3) Distributed-view represents in parallel coordinate charts, making it easier to identify anomaly detected policies. We applied the four classification criteria presented in the Ehab Al-Shaer's paper [6]. (4) AnyPolicy-view is characterized by extracting and representing only Any allowed policies. The 3D representation method is similar to the FRuVATS [14] form. (5) Search-view simplifies the user

TABLE 2. Six functions of HSViz.

No	Function	View	Needs for Visualization	Forwarding information	Use case
1	Hierarchy-view		The service that goes through the firewall, and the trend of source, destination, number of ports, etc. existing in the firewall policy.	Services via firewall Allowed port count	Firewall failover, Firewall relocation Risk check with port allowed count
2,3	Anomaly, Distributed-view		Unused, redundant and unnecessary policies that exist in firewall policy. (also, exist between firewall policies)	Policy anomalies in single or distributed firewalls	Optimize policy Improve performance
4	AnyPolicy-view		The excessively allowed policy existing in the firewall policy by SIP, DIP, and DPort.	Excessively allowed policy	Identify and remove vulnerabilities Security audit tool
5	SearchResult-view		The results of firewall policy search based on source, DIP, DPort.	Search result	Check for vulnerable services Verify assets
6	Top&Bottom Used-view		Unused policy and frequently used policy	Policy used count Use of frequency	Identify unused frequently used policies

input value to three, indicating only the distribution status of the policy. It is similar to the expression of Policyviz [12]. (6) Top and Bottom Used-view is a feature that expresses the highest and lowest frequency of use of firewall policies.

III. VISUALIZATION TOOL OVERVIEW

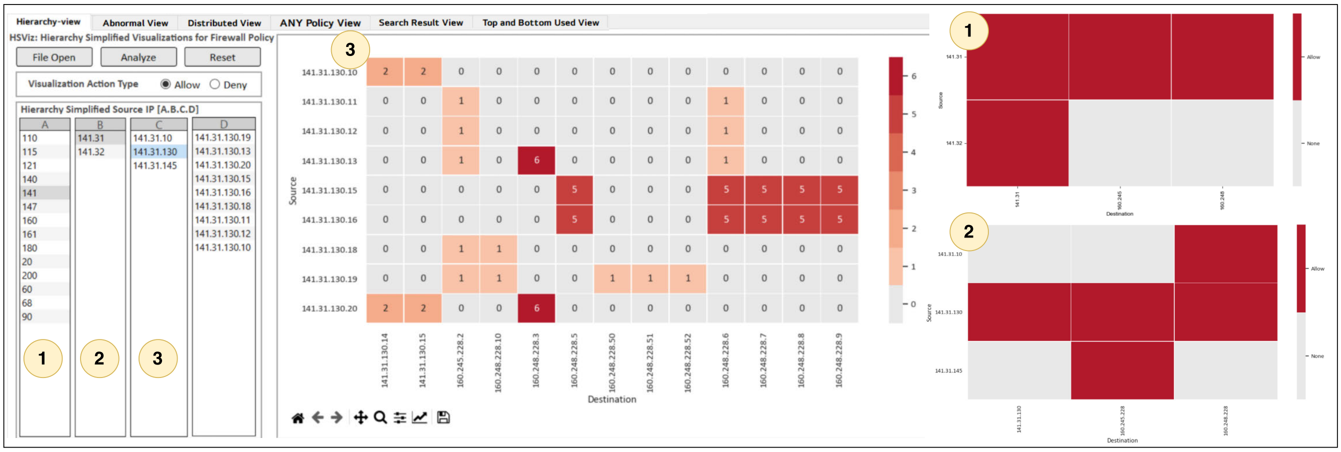
HSViz provides a total of six functions. Table 2 summarizes HSViz full functionality. (1) Hierarchy-view, which separates the layers based on the octets of the SIP and visualizes them in a segmented grid; (2) Anomaly-view, which visualizes four situations of policy anomalies in the form of a parallel coordinates plot; (3) Distributed-view, which visualizes a policy anomaly in a distributed firewall environment in the form of a parallel coordinates plot; (4) ANYPolicy-view, which provides a three-dimensional representation of six excessive allow policies in 3D graphs, (5) SearchResult-view, which visualizes the search results of the policy in the form of a 2D grid, and (6) Top and Bottom Used-view, which describes the top-level and unused policy distribution status. Each of these six functions can be used to analyze the firewall policy in a simple and intuitive way to facilitate recognition of the correlation of objects serviced by the firewall and effectively visualize anomaly policies.

A. SIMPLIFICATION

HSViz provides users can select an octet-based separation of the IP addresses to check the details and to see the deny status of the selected area based on an overall view. It simplifies the complex IP address configuration for the user to check the details with greater ease. The visual information provided to the user from the results of visualization in a grid form was designed in such a way that would make it intuitive for the user to grasp. To be more specific, there are less than five elements presented in the results, with a designated color for each case in which there is a policy mapped with the DPort information, and there is variation in the intensity of the color used depending on the number of ports allowed access.

B. INTUITIVE DISCOVERY OF POLICY ANOMALIES

Single firewall and distributed firewall policies were analyzed to extract policy anomalies, which were then visualized in the form of a parallel coordinates plot. It was made possible to identify anomalies with minimal information during visualization. The information expressed in the parallel coordinates plot was designed to minimize the visualized information that the user would have to process, with no more than five types of information displayed, which were ID (Order), Action, Source IP address (SIP), Destination IP address(DIP), and Destination Port (DPort). By doing so, the user can quickly recognize and understand the information, and thus the system is simple and intuitive in terms of information recognition and processing. Anomaly policies can be easily identified because the policy order, SIP, DIP, and DPort differences are distinguished by lines using five major pieces of information that distinguish policies as the variables on the parallel coordinates group.



A	B	C	D
110	141.31	141.31.10	141.31.130.19
115	141.32	141.31.130	141.31.130.13
121		141.31.145	141.31.130.20
140			141.31.130.15
141			141.31.130.16
147			141.31.130.18
160			141.31.130.11
161			141.31.130.12
180			141.31.130.10
20			
200			
60			
68			
90			

Source \ Destination	141.31.130.14	141.31.130.15	160.245.228.2	160.248.228.10	160.248.228.3	160.248.228.5	160.248.228.50	160.248.228.51	160.248.228.52	160.248.228.6	160.248.228.7	160.248.228.8	160.248.228.9
141.31.130.10	2	2	0	0	0	0	0	0	0	0	0	0	0
141.31.130.11	0	0	1	0	0	0	0	0	0	1	0	0	0
141.31.130.12	0	0	1	0	0	0	0	0	0	1	0	0	0
141.31.130.13	0	0	1	0	6	0	0	0	0	1	0	0	0
141.31.130.15	0	0	0	0	0	5	0	0	0	5	5	5	5
141.31.130.16	0	0	0	0	0	5	0	0	0	5	5	5	5
141.31.130.18	0	0	1	1	0	0	0	0	0	0	0	0	0
141.31.130.19	0	0	1	1	0	0	1	1	1	0	0	0	0
141.31.130.20	2	2	0	0	6	0	0	0	0	0	0	0	0

FIGURE 5. Main panels in hierarchy-view.

C. ANY ALLOW POLICY VISUALIZATION

The use of the ANY allow policy is presented by visualizing for each case via a 3D graph. We show a total of six visualizations of the number of cases where the ANY allow policy of the firewall policy can be used. With this function, firewall operators and decision-makers can achieve a quick, at-a-glance view of the existence of the ANY allow policy, providing good evidence to support decision-making, such as policy optimization or elimination measures.

D. SEARCHING FOR SIP, DIP, AND DPort

We can check the SIP, DIP, and DPort used in the policy. When searching SIP, the DIP and DPort assigned to the policy from which SIP is used are visualized. When searching a SIP, visualize DIP and DPort. When searching DIP, visualize the SIP and DPort. In addition, when searching a DPort, we visualize the SIP and DIP.

E. UNUSED AND MOST FREQUENTLY USED POLICIES

We visualized the use of the policy with count variable information provided by the firewall policy. The network node graph was used to visualize the status of unused policies and the distribution status of the most frequently used policies.

IV. VISUALIZATION MODEL

A. HIERARCHY-VIEW: SIMPLIFICATION

The reason that the task of visualizing the firewall policy is complicated and difficult is that there is a large number and wide variety of factors to be expressed including the SIP, DIP, ports, and protocol type of the firewall. The total number of cases where the IP address can be expressed is 256^4 for a total of about 4.3 billion as shown in Figure 6.

If the number of factors such as source, destination, port, and protocol is multiplied, then the total number of cases increases in proportion to the multiple of the factor value. When expressing a firewall policy with a large number of cases in consideration of factors making up the policy, it becomes increasingly complex as the number of policies increases, and there are limitations of expression. In order to address the issue of complexity of expression, an attempt was made to visualize the SIP by dividing it into octets and classifying them into domains A, B, C and D in the proposed visualization system. By narrowing down the scope of expression by selecting specific IP addresses in the A, B, and C layers based on the SIP, it is possible to check the details of the policies specific to the selected IP addresses, and thus the user can select the scope of expression and check the details.

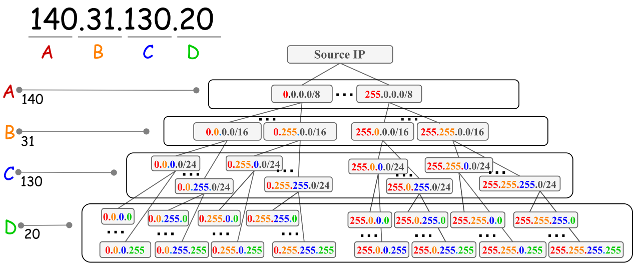


FIGURE 6. Octet-based IP address layer separation.

The form of a grid for each layer depicted when Octet A, Octet B and Octet C are selected from all SIP existing in the policy as a way to check the presence of a policy for each band. The Y-axis represents the source address and the X-axis the destination address, and it is possible to check for the existence of a policy based on the source address selected for each layer based on the octet of the IP address as shown in Figure 8.

Figure 5 shows that users can select a band among the listed objects to see it in detail. For instance, if users selects the 141.31 band, all numbers of Octet C among the source objects whose SIP starts with 141.31 will be analyzed, and as a result, numbers that include Octet C such as 141.31.10.x, 141.31.130.x, and 141.31.145.x will be listed. At the same time, the corresponding IP addresses will be represented on the Y-axis, and the policies corresponding to the IP addresses will output the DIP up to Octet C on the X-axis. Thus, a grid will be created with the Y-axis representing the SIP and the X-axis the DIP, with both the x- and y-axes covering octets A, B and C.

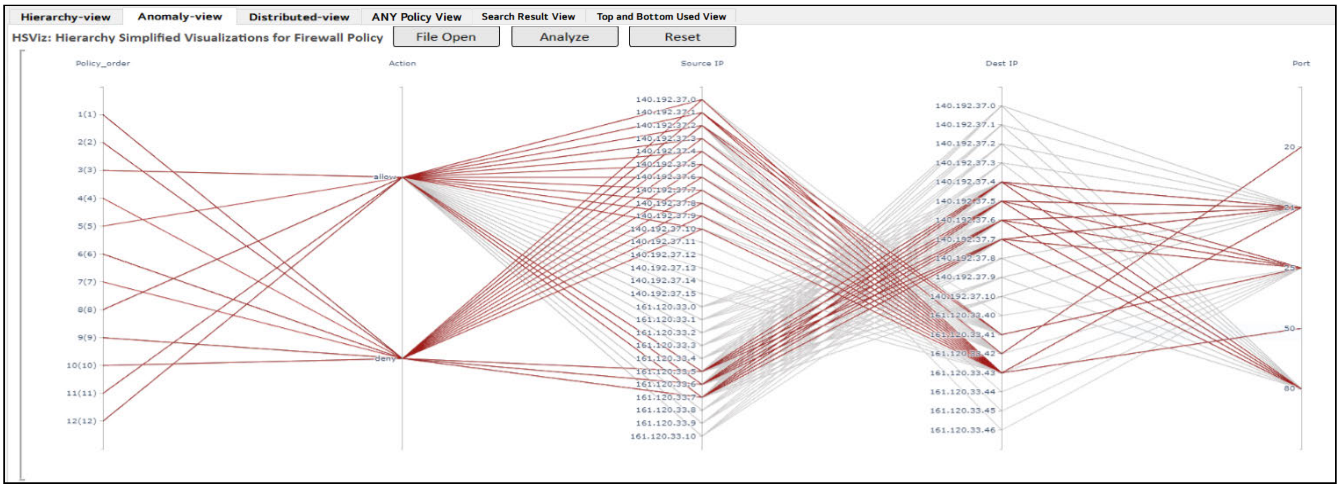


FIGURE 7. Full screen for detected anomaly policies in anomaly-view.

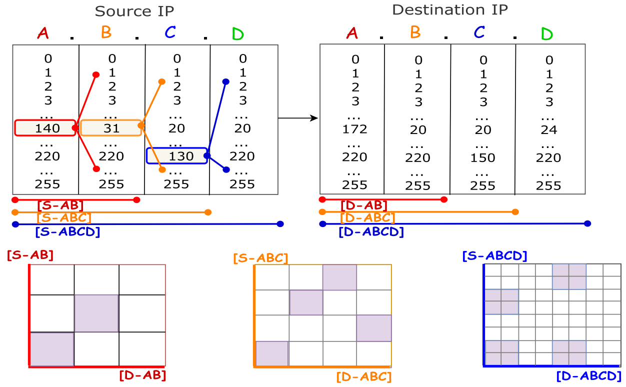


FIGURE 8. A concept of the hierarchy firewall policy visualization.

In the same way, if the user selects Octet C to be analyzed in detail, all the numbers of the subordinate Octet D under the selected Octet C will be listed. At the same time, a grid will be created with the SIP starting with the number up to the selected Octet C on the Y-axis and all DIP matching the SIP on the X-axis. For example, selecting 141.31.130 in Domain C will have all DIP that start as 141.31.130 such as 141.31.130.19, 141.31.130.13, and 141.31.130.20 will be output on Domain D, and a grid will be created with the details of the corresponding IP addresses on the Y-axis and the DIP mapped to the IP addresses from the Y-axis in the policy on the X-axis.

If there are policies with the Y-axis and the X-axis as the destination and source, the areas where such policies exist will be colored to distinguish them. In the grid at the bottom, a number is added to the grid to indicate the number of ports in the allow or deny policy, and the intensity of the color increases with more allowed ports in the form of a HEATMAP. In figure 5, the area in red color means that there is an allow policy in the matched source and destination bands. The advantage of this view is that it is possible for the user to intuitively check the overall policy status of the selected object, and the firewall administrator can recognize at a glance the overall and detailed views of the related services where there exists an allow policy in the SIP to be checked. Since the color will darker with more allowed ports, the administrator can confirm the safety level of the policy by checking the allowed ports.

B. ANOMALY-VIEW: VISUALIZING POLICY ANOMALIES

Among the policy components, the elements that can be used as an indicator for classifying the policies were listed in a parallel coordinates plot and expressed in the form of a parallel graph. The correlation between policies was analyzed by using the Order, Action, SIP, DIP, and DPort of the policies as variables. Generally, indicators that determine whether a policy is misused are the order of policies and the inclusive relationship between policies. By expressing the correlation between individual policies with a parallel graph, it became possible to intuitively check policy anomalies.

Typical anomaly detected policies are classified into four categories: (i) Shadowing, (ii) Correlation, (iii) Generalization, and (iv) Redundancy Anomaly [8]. When analysis is started by loading the policy to be analyzed, policies with four of the anomaly cases described above are output in a text format and in the form of a parallel coordinates plot at the same time. The correlations of the policies derived from anomaly detected policies are listed laterally and expressed in a parallel form on the graph. Policies marked in dark red are anomaly detected policies, and whether there is policy matching with respect to the policy order, action, source, destination, and port and the order of policies are shown. The data in Figure 7 were analyzed using the example data shown in Table 1, and it shows the results of expressing all the anomaly detected policies. By expressing the firewall policy in a parallel coordinates plot, it was easy to understand the correlation between the policies under comparison. It is easy to check whether there is a difference in the order of the policies, whether the actions are different, or whether there are differences in the source and DPort and so on, making it easy to grasp the anomaly used relationship. Figure 7 shows

TABLE 3. Firewall policy for shadowing anomaly.

Order	Type	SIP	DIP	Port	Action
3	TCP	140.192.37.0-10	161.120.33.41	25	allow
4	TCP	140.192.37.1	161.120.33.41	25	deny

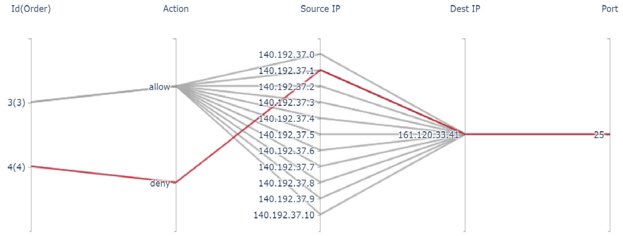


FIGURE 9. Shadowing anomaly graph in anomaly-view.

TABLE 4. Firewall policy for correlation anomaly.

Order	Type	SIP	DIP	Port	Action
1	TCP	140.192.37.2	161.120.33.40-46	25	deny
3	TCP	140.192.37.0-10	161.120.33.41	25	allow

the result of expressing four possible anomaly use cases that can occur in a single firewall in a parallel coordinates plot.

1) SHADOWING ANOMALY

Shadowing is one of the typical cases of anomaly detected policies that need to be addressed to ensure efficient performance of firewalls. In the case of policies with the same source, destination, and port but with different actions, a subordinate policy can become deactivated by the super-ordinate policy. That is, the subordinate policy exists but is not actually reflected and thus becomes unnecessary.

$$\begin{aligned} Rx[order] &< Ry[order], \quad RxEMRy, \ Rx[action] \\ &\neq Ry[action] \\ Rx[order] &< Ry[order], \quad RxIMRy, \ Rx[action] \\ &\neq Ry[action] \end{aligned}$$

If there is a deny policy in the super-ordinate level, the allow policies in the subordinate levels will be shadowed, and the firewall administrator should be aware of such shadowing. Rule 4 is shadowed by Rule 3, which has the identical source, destination and port but different actions as shown in Table 3 and Figure 9.

2) CORRELATION ANOMALY

Correlation refers to a case in which a super-ordinate policy and a subordinate policy have a mutually inclusive relationship, but the actions are different.

$$Rx \ C \ Ry, \quad Rx[action] \neq Ry[action]$$

In Table 4 and Figure 10, since the policy corresponding to the DIP of Rule 3 from the SIP of Rule 1 will be denied due to the super-ordinate policy, the Allow Rule from 140.192.37.2 to 161.120.33.41, which is part of Rule 3, will not be activated.

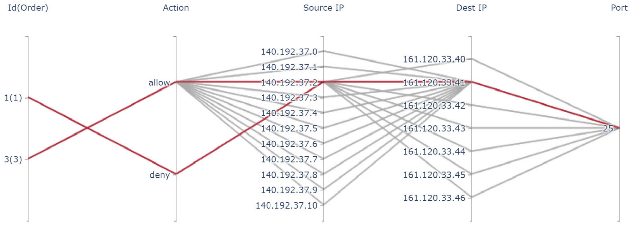


FIGURE 10. Correlation anomaly graph in anomaly-view.

TABLE 5. Firewall policy for generalization anomaly.

Order	Type	SIP	DIP	Port	Action
5	TCP	140.192.37.3	161.120.33.43	21	allow
6	TCP	140.192.37.0-10	161.120.33.43	21	deny

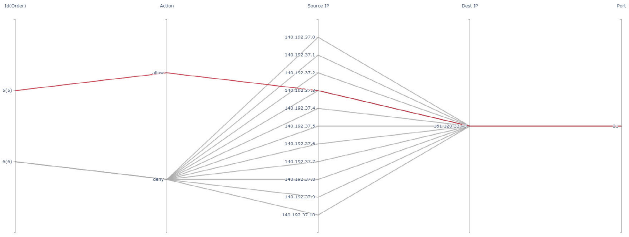


FIGURE 11. Generalization anomaly graph in anomaly-view.

3) GENERALIZATION ANOMALY

Generalization refers to a case in which the subordinate rule includes the super-ordinate rule, but the actions are different. It seems similar to shadowing, but the two are strictly different. Shadowing is a case where the super-ordinate includes or matches a subordinate policy, whereas generalization is a case where a subordinate policy includes the super-ordinate policy.

$$Rx[order] < Ry[order], \quad RyIMRx, \ Rx[action] \neq Ry[action]$$

In Table 5 and Figure 11, Rule 6, a subordinate rule with a broad range, is applied as a deny rule, but Rule 5, which has a higher priority and narrower range, is allowed, and thus this case falls under the category of generalization anomaly. Rule 6 is a generalization of Rule 5, and this is more of a warning rather than a misuse.

4) REDUNDANCY ANOMALY

Redundancy is also a case of anomaly detected policies that must be addressed to optimize firewall performance along with shadowing. Redundancy increases the processing time by increasing the size of the filtering rules of the firewall. In Table 6 and Figure 13, it occurs when all the elements of the subordinate rule are included in the super-ordinate rule or match the elements thereof and the actions also match.

$$\begin{aligned} Rx[order] &< Ry[order], \quad RxEMRy, \ Rx[action] \\ &= Ry[action] \\ Rx[order] &< Ry[order], \quad RxIMRy, \ Rx[action] \\ &= Ry[action] \end{aligned}$$

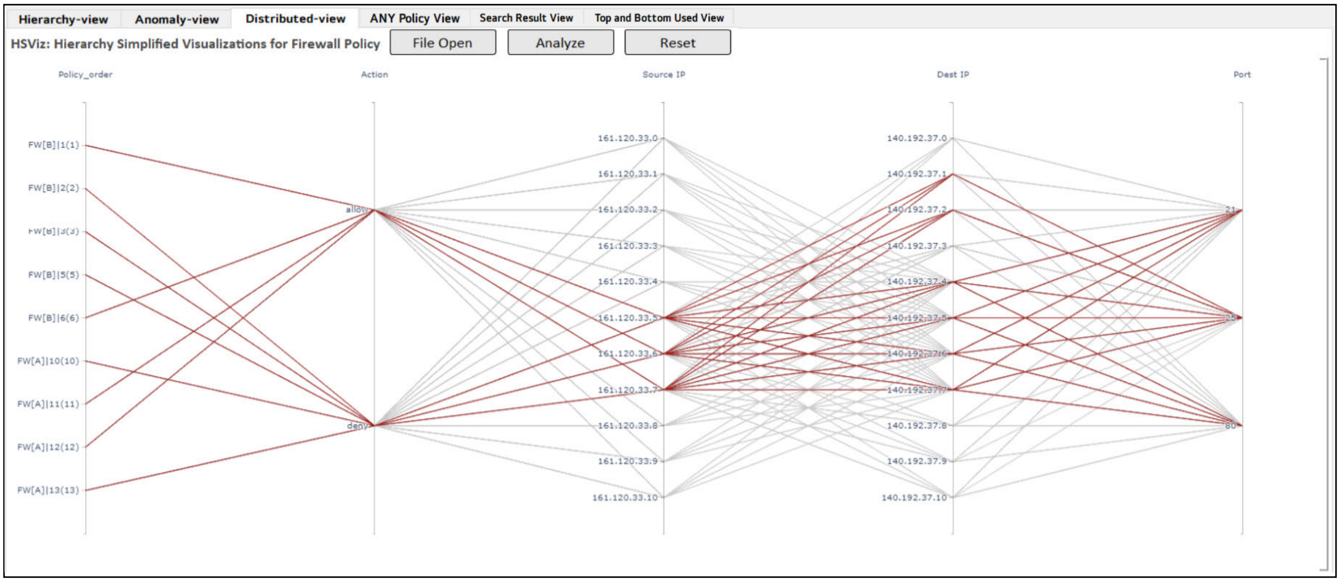


FIGURE 12. Distributed-view: distributed firewall policy anomaly view.

TABLE 6. Firewall policy for redundancy anomaly.

Order	Type	SIP	DIP	Port	Action
2	TCP	140.192.37.0-10	161.120.33.42	20	deny
7	TCP	140.192.37.1	161.120.33.42	20	deny

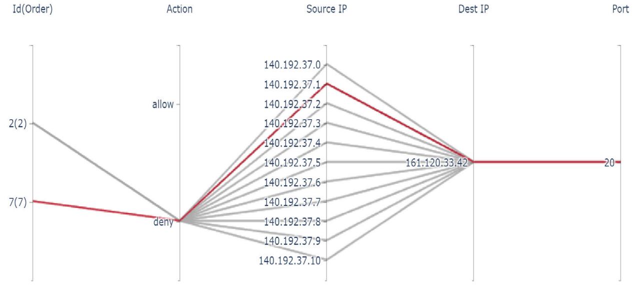


FIGURE 13. Redundancy anomaly graph in anomaly-view.

In the case below, Rx is said to be redundant to Ry.

$$Rx[order] < Ry[order], \quad RyIMRx,\ Rx[action] = Ry[action]$$

Rule 7 is processed to be redundant by Rule 2, the source, destination, ports and actions of which are identical to those of Rule 7.

C. DISTRIBUTED-VIEW: DISTRIBUTED FIREWALL VISUALIZATION

There are cases where multiple firewalls are set up and operated in each section in a configuration where the infrastructure environment is separated into three tiers or in an environment where a large network is separated. In such a large-scale environment, firewall administrators need to check for any anomaly detected policies of a single firewall from the perspective of network performance management when providing services, in addition to checking the policies of associated firewalls in the entire network. Anomaly policies in a distributed firewall environment are classified into four types: (i) Shadowing, (ii) Spuriousness, (iii) Redundancy Anomaly, and (iv) Correlation [8]. Each case of anomaly policies in multiple firewall environment was analyzed and visualized based on the policies presented in Table 1 and 7 examples. Figure 12 shows the result of analyzing policy anomalies in a multiple firewall environment.

When Distributed-view on HSViz loads the policy files of two firewalls, the policy files are analyzed by the designated algorithm, and a parallel graph similar to the one presented in Anomaly-view gets displayed. The firewall name and related policies detected as anomaly policies are displayed in the Order. Figure 12 shows the results of extracting all anomaly policies detected in the cases of misuse in a multiple firewall environment after loading and analyzing firewall policies. Referring to Figure 14, we assume that the closest firewall to the flow source subdomain (FW [A], Table 1) is called the most-upstream firewall, while the closest firewall to the flow destination sub-domain (FW [B], Table 7) is called the most-downstream firewall.

TABLE 7. Firewall B's rule (FW [B]).

Order	Type	SIP	DIP	Port	Action
1	TCP	161.120.33.5-7	140.192.37.4-7	80	allow
2	TCP	161.120.33.5-7	140.129.37.5	21	deny
3	TCP	161.120.33.0-10	140.192.37.4-7	21	deny
4	TCP	161.120.33.6	140.192.37.4-7	23	allow
5	TCP	161.120.33.7	140.192.37.1-10	25	deny
6	TCP	161.120.33.5-7	140.192.37.2-6	25	allow

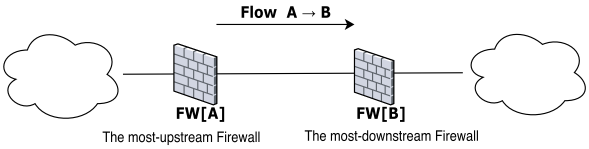


FIGURE 14. Upstream and downstream firewall concept [8].

TABLE 8. Firewall policy for shadowing anomaly.

FW	Order	Type	SIP	DIP	Port	Action
A	10	TCP	161.120.33.5-7	140.192.37.0-10	21,80	deny
B	1	TCP	161.120.33.5-7	140.192.37.4-7	80	allow

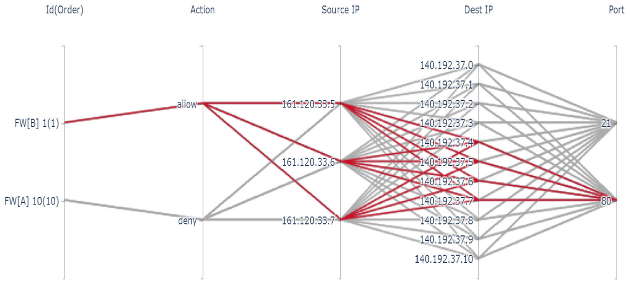


FIGURE 15. Shadowing anomaly graph in distributed-view.

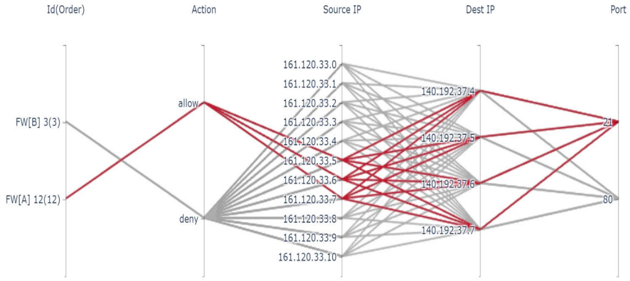


FIGURE 16. Spuriousness anomaly graph in distributed-view.

1) SHADOWING ANOMALY

In a distributed firewall environment, shadowing occurs when the rule that is denied in the upstream firewall, FW [A], is allowed in the downstream firewall, FW [B]. The following are cases where shadowing occurs.

$$\begin{aligned} RdEMRu, &\quad Ru[action] = deny,\ Rd[action] = allow \\ RdIMRu, &\quad Ru[action] = deny,\ Rd[action] = allow \\ RuIMRd, &\quad Ru[action] = deny,\ Rd[action] = allow \\ RuIMRd, &\quad Ru[action] = allow,\ Rd[action] = allow \end{aligned}$$

An examination of the interrelationship between the policies expressed in a parallel coordinates plot, it can be seen that rule 1 of FW [B] is blocked by rule 10 of FW [A], as shown in red (see Table 8, Figure 15).

2) SPURIOUSNESS ANOMALY

Spuriousness occurs when a rule denied by a downstream firewall is allowed by an upstream firewall.

$$\begin{aligned} RuEMRd, &\quad Ru[action] = allow,\ Rd[action] = deny \\ RuIMRd, &\quad Ru[action] = allow,\ Rd[action] = deny \\ RdIMRu, &\quad Ru[action] = allow,\ Rd[action] = deny \\ RdIMRu, &\quad Ru[action] = allow,\ Rd[action] = allow \\ RuIMRd, &\quad Ru[action] = deny,\ Rd[action] = deny \end{aligned}$$

In Table 9, Rule 3 of downstream FW [B] denies Rule 3 allowed by upstream FW [A], thereby causing Rule 12 to become partially block. This can be expressed as a parallel coordinates plot graph as shown in Figure 16.

TABLE 9. Firewall policy for spuriousness anomaly.

FW	Order	Type	SIP	DIP	Port	Action
A	12	TCP	161.120.33.5-7	140.192.37.4-7	21,80	allow
B	3	TCP	161.120.33.0-9	140.192.37.4-7	21	deny

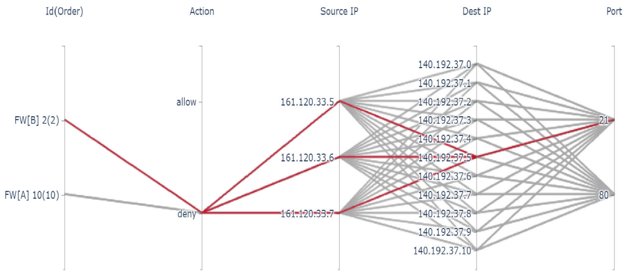


FIGURE 17. Redundancy anomaly graph in distributed-view.

TABLE 10. Firewall policy for redundancy anomaly.

FW	Order	Type	SIP	DIP	Port	Action
A	10	TCP	161.120.33.5-7	140.192.37.0-10	21,80	deny
B	2	TCP	161.120.33.5-7	140.129.37.5	21	deny

TABLE 11. Firewall policy for correlation anomaly.

FW	Order	Type	SIP	DIP	DPort	Action
A	13	TCP	161.120.33.5-7	140.192.37.1-2	25	deny
B	5	TCP	161.120.33.7	140.192.37.1-10	25	deny

3) REDUNDANCY ANOMALY

Redundancy occurs when a Rule denied by an upstream firewall is denied by a downstream firewall.

$$\begin{aligned} RdEMRu, &\quad Ru[action] = deny,\ Rd[action] = deny \\ RdIMRu, &\quad Ru[action] = deny,\ Rd[action] = deny \end{aligned}$$

In Table 10, Rule 2 of downstream FW [B] can be completely mapped to Rule 10 of upstream FW [A], with some of the actions being identical, and this is a case of redundancy anomaly. This can be expressed as a parallel coordinates plot graph as shown in Figure 17.

4) CORRELATION ANOMALY

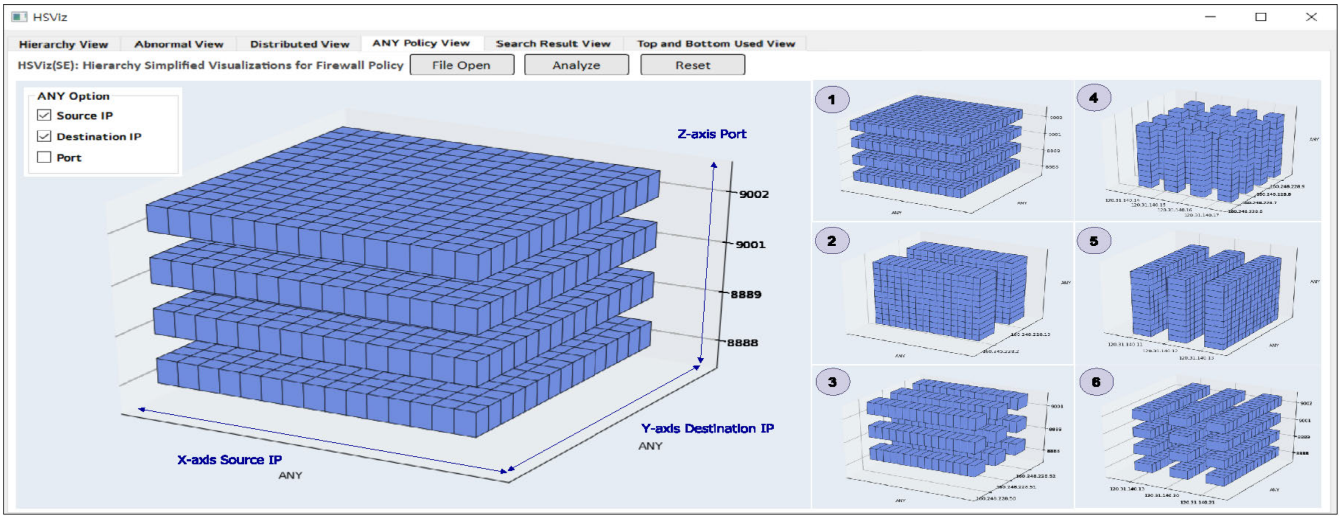
In a multiple firewall environment, correlation anomaly occurs when the policies between firewalls are mutually inclusive and actions are either identical or different as shown in Table 11 and Figure 19.

$$\begin{aligned} RdCRu, &\quad Ru[action] = allow,\ Rd[action] = allow \\ RdCRu, &\quad Ru[action] = deny,\ Rd[action] = deny \\ RdCRu, &\quad Ru[action] = allow,\ Rd[action] = deny \\ RdCRu, &\quad Ru[action] = deny,\ Rd[action] = allow \end{aligned}$$

Because Rule 13 of upstream FW [A] and Rule 5 of FW [B] are mutually inclusive and allowed, they fall under the category of correlation anomaly.

D. ANYPOLICY-VIEW: 3D GRAPH VISUALIZATION

The firewall policy is allowed as ANY, which means that all IPs from 0.0.0.0 to 255.255.255, are allowed, rather than IPs that are components of the firewall policy, such as SIP,



HSViz(SE): Hierarchy Simplified Visualizations for Firewall Policy

ANY Option
☑ Source IP
☑ Destination IP
□ Port

FIGURE 18. Any policy-view: any allowed firewall policy view.

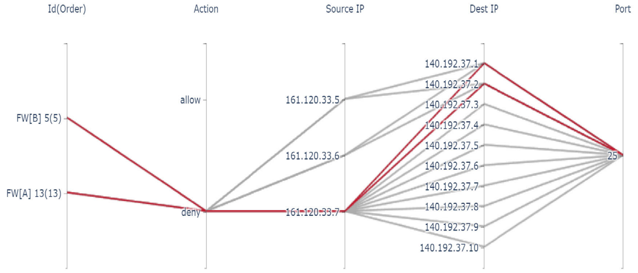


FIGURE 19. Correlation anomaly graph in distributed-view.

TABLE 12. Any allowed policy cases.

Action	SIP	DIP	DPort	Visualization
Allow	**ANY**	**ANY**	**ANY**	
	ANY	**ANY**	Specific	Figure 20
	ANY	Specific	**ANY**	Figure 21
	ANY	Specific	Specific	Figure 22
	Specific	**ANY**	**ANY**	Figure 23
	Specific	Specific	**ANY**	Figure 24
	Specific	**ANY**	Specific	Figure 25

DIP, or DPort, that are subdivided into host or network bands. It also means that for ports, all ports in the range of 0 to 65535 are granted. That is, if the source or destination is permitted to ANY, the IP address range granted to that source or destination is all IP addresses from 0.0.0.0 to 255.255.255.255. The number of cases in which ANY policies can be used is as shown in Table 12.

The existence of an ANY-allowed policy is that it allows policies to all sources, destinations or ports. It provides an attacker with an access path to an internal service that the administrator does not know and expose a vulnerability in the service that should not be allowed to access. To assist in the recognition and decision-making of the need for the elimination of ANY policies, the following 3D graphs have been used to visualize the status of ANY allow policies. Figure 18 shows the full UI representation of the tool, the X-axis means the source, the Y-axis means the destination, and the Z-axis means the DPort. Once the policy file for analysis has been entered and loaded, an options window has been set up on the right to select the location of the ANY Allow policy we want to analyze. After loading the policy file, visualize the existence of the ANY policy existing at the SIP, DIP and DPort in 3D graph.

TABLE 13. SIP and DIP: ANY, DPort: Specific.

ID	Type	SIP	DIP	DPort	Action
93	tcp, udp	0.0.0.0	0.0.0.0	8888, 8889, 9001, 9002	allow

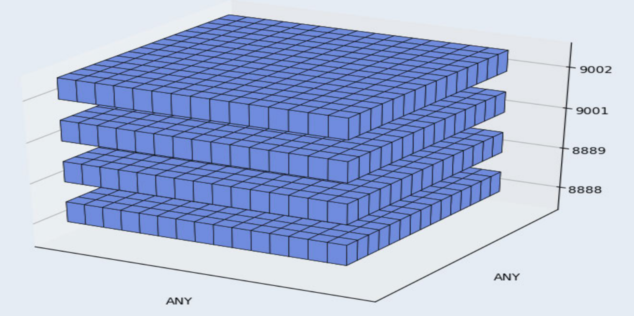


FIGURE 20. Any policy view - SIP, DIP: ANY, DPort: Specific.

1) SIP: ANY, DIP: ANY, DPort: SPECIFIC

The SIP and DIP is given as ANY and the DPort is specified in Table 13; Then, the results of the visualization can be expressed as shown in Figure 20.

2) SIP: ANY, DIP: SPECIFIC, DPort: ANY

The SIP and DPort is given as ANY, and only the DIP is specified in Table 14; Then it is expressed as shown in Figure 21.

3) SIP: ANY, DIP: SPECIFIC, DPort: SPECIFIC

The SIP is given as ANY, and DIP and DPort are specified in Table 15; Then, it is expressed as shown in Figure 22.

TABLE 14. SIP and DPort: ANY, DIP: Specific.

ID	Type	SIP	DIP	DPort	Action
94	tcp, udp	0.0.0.0	160.245.228.2	0 - 65535	allow
94	tcp, udp	0.0.0.0	160.245.228.10	0 - 65535	allow

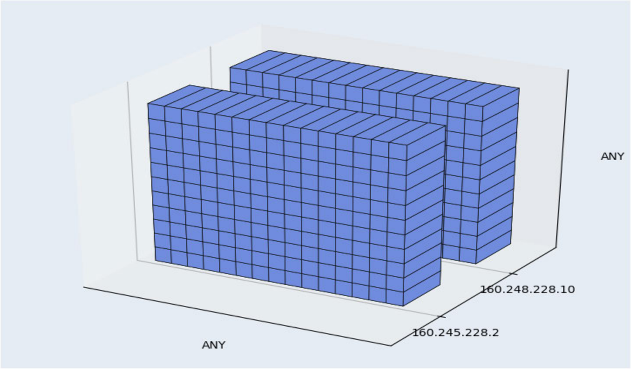


FIGURE 21. Any policy view - SIP and DPort: ANY, DIP: Specific.

TABLE 15. SIP: ANY, DIP and DPort: Specific.

ID	Type	SIP	DIP	Port	Action
90	tcp, udp	0.0.0.0	160.248.228.50	8888, 8889, 9001	allow
90	tcp, udp	0.0.0.0	160.248.228.51	8888, 8889, 9001	allow
90	tcp, udp	0.0.0.0	160.248.228.52	8888, 8889, 9001	allow

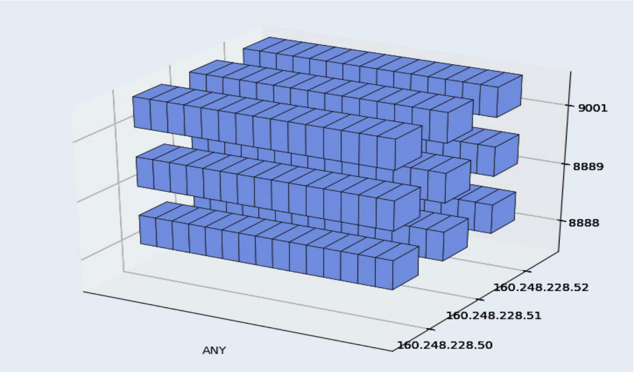


FIGURE 22. Any policy view - SIP, DIP: ANY, DPort: Specific.

TABLE 16. ANY policy (DIP, DPort: ANY).

ID	Type	SIP	DIP	DPort	Action
95	tcp, udp	120.31.140.11-13	0.0.0.0	0-65535	allow

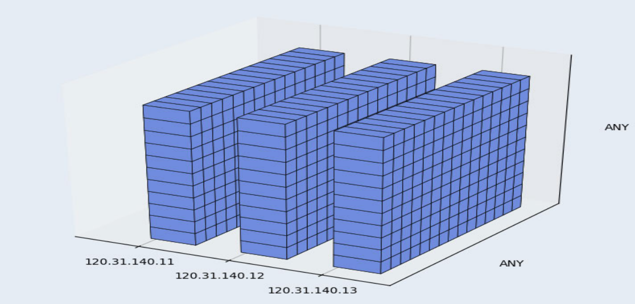


FIGURE 23. Any policy view - SIP: Specific, DIP, DPort: ANY.

4) SIP: SPECIFIC, DIP: ANY, DPort: ANY

The SIP is specified, and the DIP and DPort is allowed as ANY in Table 16; Then, it is expressed as shown in Figure 23.

TABLE 17. ANY policy (DPort: ANY).

ID	Type	SIP	DIP	DPort	Action
92	tcp	120.31.140.14-17	160.248.228.6-9	0-65535	allow

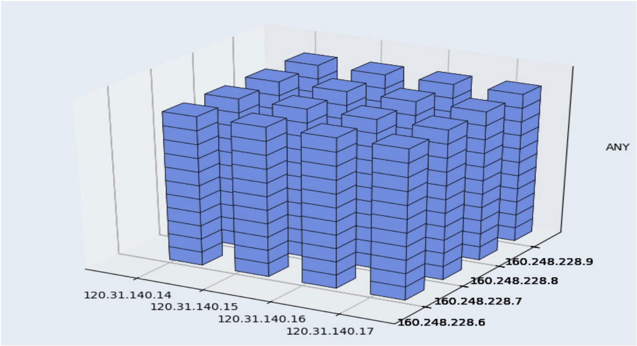


FIGURE 24. Any policy view - SIP, DIP: Specific, DPort: ANY.

TABLE 18. ANY policy (DIP: ANY).

ID	Type	SIP	DIP	DPort	Action
91	120.131.140.13	0.0.0.0	8888, 8889, 9001, 9002	tcp,udp	allow
91	120.131.140.20	0.0.0.0	8888, 8889, 9001, 9002	tcp,udp	allow
91	120.131.140.21	0.0.0.0	8888, 8889, 9001, 9002	tcp,udp	allow

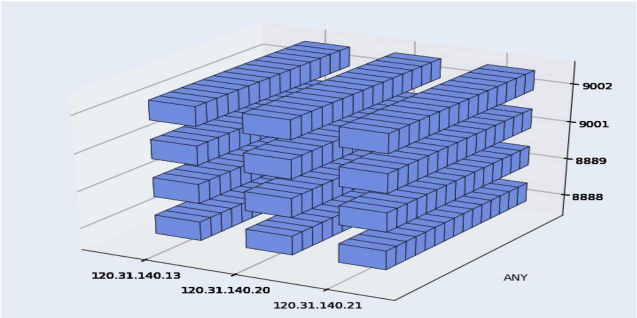


FIGURE 25. Any policy view - DIP: ANY, SIP, DPort: Specific.

5) SIP: SPECIFIC, DIP: SPECIFIC, DPort: ANY

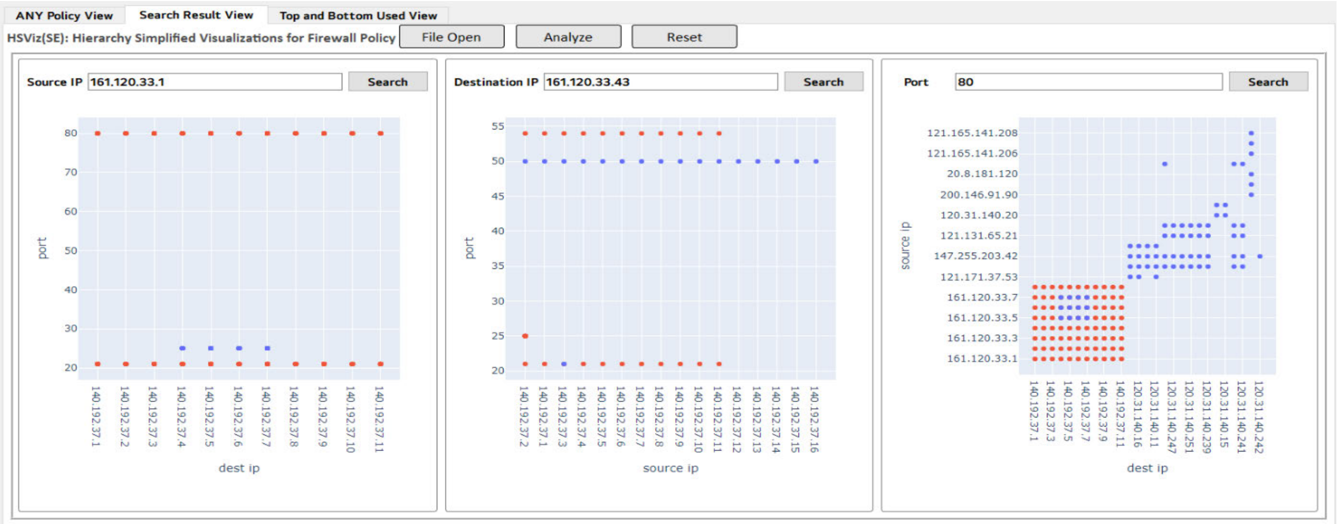
The SIP and DIP are specified, and the DPort is ANY in Table 17; Then it is expressed as show in Figure 24.

6) SIP: SPECIFIC, DIP: ANY, DPort: SPECIFIC

The SIP, DPort is specified and the DIP is allowed ANY in Table 18; Then it is expressed as shown in Figure 25.

E. SEARCH RESULT VIEW

The main factors determining firewall policy identification are SIP, DIP, DPort, protocol, and action information. Among this information, search for each SIP, DIP, or DPort value, and output a policy mapped to that IP or DPort information. When searching for SIP in firewall policy, the 2D graph shows the DIP mapped to SIP as X-axis and DPort information as Y-axis. When searching for DIP, the 2D graph visualizes the point mapped to the DIP with the SIP as the X-axis and DPort information as the Y-axis. This is useful when checking the existence of a policy that is allowed or denied in a particular SIP or DIP. In particular, the ability to visualize SIP and DIP as DPort information can also be used as a function to extract



HSViz(SE): Hierarchy Simplified Visualizations for Firewall Policy

Source IP 161.120.33.1

Destination IP 161.120.33.43

Port 80

FIGURE 26. **Search result view: Firewall policy search result visualization.**

TABLE 19. **Firewall policy search result (SIP: 161.120.33.1).**

Name	SIP	DIP	DPort	Type	Action
FW [A]	161.120.33.1-8	140.192.37.1-11	21,80	tcp	deny
FW [A]	161.120.33.1-11	140.192.37.4-7	25	tcp	allow

TABLE 20. **Firewall policy search result (DIP: 161.120.33.43).**

Name	SIP	DIP	DPort	Type	Action
FW [A]	140.192.37.2	161.120.33.40-46	25	tcp	deny
FW [A]	140.192.37.3	161.120.33.43	21	tcp	allow
FW [A]	140.192.37.1-11	161.120.33.43	21	tcp	deny
FW [A]	140.192.37.1-16	161.120.33.43	50	upd	allow
FW [A]	140.192.37.1-11	161.120.33.43	54	udp	deny

TABLE 21. **Firewall policy search result (DPort: 80).**

Name	SIP	DIP IP	DPort	Type	Action
FW[A]	161.120.33.1-8	140.192.37.1-11	21,80	tcp	deny
FW[A]	161.120.33.5-7	140.192.37.4-7	21,80	tcp	allow
FW[A]	121.171.37.53	120.31.140.238	80,8080,443	tcp	allow
FW[A]	60.192.135.85	120.31.140.238	80,8080,8009	tcp	allow
FW[A]	147.255.203.42	120.31.140.238	80,8080,8009	tcp	allow
FW[A]	68.73.77.253	120.31.140.238	80,8080,443	tcp	allow
FW[A]	121.171.37.53	120.31.140.16	80,8010,443	tcp	allow
FW[A]	60.192.135.85	120.31.140.16	80,8010,443	tcp	allow
		....			

policies that allow vulnerable ports to be a threat that should not to be exposed.

The full UI consists of three screens as shown in Figure 26. The left area visualizes the results of the policy that are matched by inputting SIP, the middle area by DIP, and the right area by entering Port to visualize the results. For the policy to be matched by inputting the values of one of the SIP, DIP, or DPort from the user, a point is expressed in the grid box based on information other than the values entered. Red for dots means deny and blue means allow.

In Figure 26, the box in the left visualizes the results of a policy extraction that contains the SIP entered from the user in the overall firewall policy. Table 19 shows that only policies containing the SIP 161.120.33.1 entered by the user have been extracted. It is represented as a dot on a graph consisting of the DIP X-axis and the DPort Y-axis.

In figure 26, the box in the middle visualizes the results of a policy extraction that contains the DIP entered from the user in the overall firewall policy. Table 20 shows that only policies containing the SIP 161.120.33.43 entered by the user have been extracted. It is represented as a dot on a graph consisting of the SIP X-axis and the DPort Y-axis.

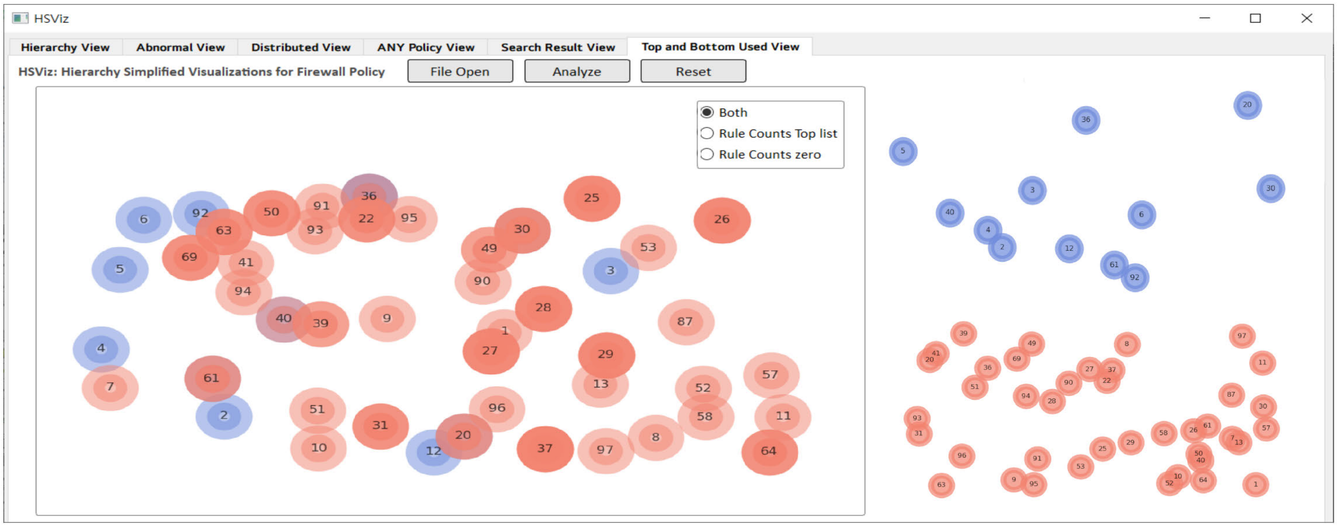
Figure 26, the box to the right visualizes the results of a policy extraction that contains the DPort entered by the user in the overall firewall policy. Table 21 shows that only policies containing the DPort 80 entered by the user have been extracted. It is represented as a dot on a graph consisting of the DIP X-axis and the SIP Y-axis.

F. TOP AND BOTTOM USED VIEW

The firewall system provides additional information other than the set policy value, which is typically the number of times the policy is used, so that the use and frequency of the policy can be checked. The count value is not a component of the firewall policy, but is provided in the settings as additional information about the policy to show the status of its use. Top and Bottom Used View uses count values to provide visualization of the most used and unused policies. To make it easier to grasp the distribution status at a glance, a network diagram graph is used. In Figure 27, unused policies are represented in red and most frequently used policies in blue. If an agent who identifies a unused policy through the tool tries to remove it, the firewall log should be checked together to ensure that the policy extracted from the unused policy is not really used. Frequency parent policies and unused policies can be identified together or separately.

V. USAGE SCENARIO

Those who will benefit the most from the tool proposed in this paper are firewall administrators. Problems associated with



● Both
○ Rule Counts Top list
○ Rule Counts zero

FIGURE 27. Top and bottom used-view: Most frequently used and unused rule view.

the operation of multiple firewalls or firewalls with a large number of Rule lines resulting in great complexity can be solved using the visualization tool proposed in this paper. The difficulty of checking the services associated with the firewall policy, which was mentioned as a problem in the introduction section of this paper, and the issue of visualizing the services allowed under the firewall policy can be addressed by checking the associated services and excessively open policies and services using the Hierarchy-view function. On the other hand, the problem of degraded firewall performance can be solved by detecting and rectifying anomaly policies applied to single and multiple firewalls by using the Anomaly-view and Distributed-view functions.

A. USE TO IDENTIFY ASSOCIATED SERVICES PROVIDED VIA FIREWALL TO RESPOND FAILURES

In case of a firewall failure or a need for physical relocation of the firewall or a patch that needs to be restarted, the firewall administrator needs to have a good grasp of the impact on the services provided by the firewall in question. The degree of impact on the service should be determined by checking with the staff that deal with the services provided via the firewall. Generally speaking, it is true that the larger the network infrastructure, the more difficult it is for the firewall administrator to know all the services that are being provided. Using the proposed tool, it is possible to intuitively grasp the destination of the related services where policies exist in the form of a grid by subdividing the services into octets A, B, C and D based on the SIP of the policies. This makes it easier to determine which services are provided via the firewall in question and understand the flow of processes. In other words, it is difficult for the firewall administrator to grasp the flow of all the processes occurring via the firewall, and the tool proposed in this study can help alleviate the challenges faced by firewall administrators.

B. FIREWALL PERFORMANCE IMPROVEMENT BY DETECTING ANOMALY POLICIES

The higher the number of applied policies and the higher the policy complexity, the lower the firewall performance. Therefore, the task of improving the performance of firewalls via services that provide fast speed is of high importance. Through the Abnormal-view, Distributed-view, we can improve policies by detecting anomalies of the firewall policies in operation. It is quite useful to identify the existence of clearly visualized anomaly detected policies and to use it when it is necessary to recognize the need for policy cleanup or to persuade other co-workers involved in the policy improvement task. In addition, policy analysis on distributed firewall environments is possible, making it easy to visualize and verify anomaly detected policies in large network environments.

C. CHECKING FOR EXCESSIVELY ALLOWED POLICIES AND PORTS

A policy allowing traffic into the company's internal system that has been applied to the firewall located in the incoming Internet traffic area may become the initial site of penetration by external attackers. If there is excessive application of an unmanaged policy that allows access, attackers will be able to enter the internal system. Policies set to allow access to all sources, destinations, ports, etc. must be managed with extra care. Policies that open up services excessively can be checked based on the number of open ports shown on the heatmap, which gets created when Octet C of the Hierarchy-view of the proposed tool is selected. The higher the value indicating the number of ports, the more open ports there are. The firewall administrator can check such visualized results as well as the domains on the heatmap indicated in a dark color as a way to examine the safety level of the policies associated with those domains. Additionally,

TABLE 22. HSViz performance test result: (1) hierarchy-view, (2) anomaly-view.

Name	Line(c)	Allow(c)	Deny(c)	Processing time in Hierarchy-view				Detected counts and processing time in Anomaly-view						
				Parsing(s)	A-B(s)	B-C(s)	C-D(s)	S(c)	C(c)	G(c)	R(c)	Sum(c)	Detect(s)	View(s)
FW[1]	24	24	0	11.11	2.74	1.76	2.08	1	0	0	0	1.46	2.28	1
FW[2]	50	49	1	13.30	1.07	2.14	1.08	0	0	0	2	1.92	2.55	2
FW[3]	53	52	1	15.50	10.32	1.56	1.74	2	0	1	3	40.91	3.22	6
FW[4]	55	54	1	11.99	59.61	0.94	1.68	7	3	1	3	1.89	2.80	14
FW[5]	61	61	0	6.17	4.21	0.68	1.72	2	0	0	0	16.89	2.44	2
FW[6]	61	60	1	9.93	4.63	2.32	1.57	6	0	0	2	9.25	801.32	8
FW[7]	72	71	1	6.36	6.04	3.89	3.47	0	0	1	0	4.25	2.52	1
FW[8]	75	73	2	31.66	100.04	1.55	1.13	1	0	1	0	32.40	2.49	2
FW[9]	75	74	1	9.92	56.18	1.62	1.32	1	0	1	1	2.39	2.71	3
FW[10]	77	77	1	7.55	15.90	1.74	1.45	3	0	0	2	14.83	2.50	5
FW[11]	89	89	0	7.34	1.95	2.33	4.02	2	0	0	0	1.25	2.76	2
FW[12]	94	93	1	35.66	7.69	1.55	1.64	2	0	1	0	7.23	2.98	3
FW[13]	116	115	1	7.38	3.81	3.55	34.77	2	0	0	2	79.58	2.44	4
FW[14]	130	128	2	11.47	87.89	1.55	1.03	22	0	1	4	39.34	3.45	27
FW[15]	135	134	2	8.71	2.08	2.03	49.43	1	0	1	3	51.90	2.53	5
FW[16]	174	174	1	17.46	41.11	19.70	81.07	4	0	0	2	77.64	2.69	6
FW[17]	195	194	1	22.23	6.30	1.54	4.41	13	0	0	0	33.73	56.08	13
FW[18]	210	210	0	16.87	20.23	1.57	1.85	16	0	0	0	31.51	3.56	16
FW[19]	568	563	5	28.04	1.59	1.01	1.48	61	1	1	8	520.84	92.60	71
FW[20]	679	647	32	50.15	1.34	1.67	0.46	30	0	1	4	68.98	3.83	35
Sum	2993	2942	54	328.78	434.74	54.69	197.41	176	4	10	36	1038.20	997.74	226
Average	149.65	147.1	2.7	16.44	21.74	2.73	9.87	8.8	0.2	0.5	1.8	51.91	49.89	11.3

the ANYPolicy-view feature allows us to identify excessively allowed services and ports.

D. DETECT SECURITY VULNERABILITIES IN FIREWALL POLICIES

Firewall operators should periodically identify vulnerable services and take steps to eliminate vulnerabilities. The tool proposed in this paper allows for integrated verification by extracting only the current status of the port on which the vulnerability exists. For example, a remote desktop connection (3389 port) service, SSH, TELNET service, etc. can be found on a web server in the DMZ section, and other policies that are mistakenly allowed.

E. ASSET IDENTIFICATION

With 80 port being serviced, the distribution of web services can be checked through firewall policies. If you want to check which servers provide Web services, you can check the allow status of ports 80, 443. Also, you can see the distribution of servers that provide DB services (3306, 1433, 1521, etc.).

F. USE AS A FIREWALL POLICY AUDIT TOOL

Allow policies to a company's internal system applied to a firewall located in an Internet interface section could be the first path to entry for cyber attackers. If the unmanaged firewall allow policy is over-applied, attackers will be able to enter the internal system. The potential problem with the existence of the ANY policy is that there is a threat that exposes vulnerabilities in unidentified internal services and provides an access path to internal services that the administrator does not know. Checking the existence of a highly management-critical ANY policy is an item that is often identified during security audits. If an entity cannot receive all policies for security reasons, or there is no time to analyze them, auditors can use the tools to simply check the status of the ANY policy's application. Operators can also check the overall status of the ANY policy with the resulting screen provided by this tool, making it easy to use for removal.

G. USE TO COMMUNICATE WITH DECISION MAKERS

The task of applying changes to firewall policies to services in operation should be careful after understanding their impact. Many operations operate without improving firewall policies on the grounds that the task of modifying firewall policies may affect service safety. In this case, accumulated firewall policies may affect stable operations. This tool can help improve firewall performance by identifying and removing unused policies. Furthermore, the use of visualization-result images as a basis for persuading decision makers in the process of removing ANY policies for security enhancement can greatly help drive and proceed with the task of security enhancement.

VI. EVALUATION

This chapter aims to prove the efficiency of the proposed tool by applying arbitrary firewall policy data that is actually in use to the aforementioned algorithms and performing a performance analysis.

A. DATASET AND TEST ENVIRONMENT

The test was carried out using a laptop with the following specification: Intel Core i5-8250U @ 1.6GHz 1.80GHz, 16G RAM, SSD 256GB. The test environment was created based on Python 3.7, Mysql, and pymysql libraries. Matplolib and seaborn libraries were used for the heatmap function on Hierarchy-view. We used the Plotly.Parcoords library as for the parallel coordinates plot graph on Anomaly-view and Distributed-view, we used matflolib to generate an ANY Policy View 3D Graph and a potly library to generate a Search Results View 2D Graph. We also used the networkx library to implement Top and Bottom Used-view Network node graphs.

TABLE 23. HSViz performance test result: (3) distributed-view.

Name	Line(c)	Allow(c)	Deny(c)	Parsing(s)	Shadowing(c)	Spuriousness(c)	Redundancy(c)	Correlation(c)	Total(c)	View(s)
FW[21]	1027	1027	0	120.9	11	12	0	14	37	6.8
FW[22]	242	242	0	42.4						
FW[23]	480	480	0	15.0	2	1	0	1	4	2.7
FW[24]	214	212	2	12.4						
Average	491	490	1	47.7	6.5	6.5	0	7.5	20.5	4.8

TABLE 24. HSViz performance test result: (4) ANYPolicy-view, (5) SearchResult-view, (6) top and bottom used-view.

Name	line(c)	ANY Policy View(s)						Search Result View(s)			Top & Bottom View(s)		
		SIP	DIP	DPort	SIP, DIP	DIP,DPort	SIP, DPort	SIP	DIP	DPort	Both	Top	Zero
FW[1]	24	0.96	0.66	0.57	0.46	0.58	0.64	2.69	3.01	1.13	0.84	0.26	0.70
FW[2]	50	6.26	4.21	1.19	1.07	1.66	1.04	3.26	3.21	2.79	1.69	0.50	1.00
FW[3]	53	2.34	9.25	0.94	0.75	0.69	0.45	3.02	1.52	1.98	2.99	1.13	2.99
FW[4]	55	2.99	12.66	1.03	0.69	0.95	0.66	4.13	5.25	3.22	1.55	0.56	1.04
FW[5]	61	0.64	62.43	1.18	0.47	0.42	0.68	2.83	1.17	1.37	0.89	0.28	1.03
FW[6]	61	0.95	6.16	0.97	0.49	0.50	0.89	2.79	2.53	1.57	2.95	1.06	2.10
FW[7]	72	23.38	0.71	0.68	0.60	0.52	0.53	3.27	1.13	1.47	1.24	0.16	1.14
FW[8]	75	1.55	1.17	1.06	0.80	0.70	0.65	5.06	2.14	2.81	1.64	0.82	1.03
FW[9]	75	16.78	6.34	2.16	2.57	1.22	1.09	2.16	3.11	1.79	1.34	0.35	1.13
FW[10]	77	18.72	0.52	0.66	0.59	0.48	0.66	3.68	1.42	1.69	1.35	0.24	1.36
FW[11]	89	18.45	15.86	0.68	0.40	0.59	0.81	3.69	1.40	1.78	1.28	0.34	1.24
FW[12]	94	12.19	0.68	0.45	0.47	0.52	0.61	3.11	1.17	1.62	1.59	0.15	1.52
FW[13]	116	16.45	0.46	0.68	0.60	0.44	0.48	3.26	1.47	2.12	1.94	0.25	1.74
FW[14]	130	1.24	0.79	0.93	0.63	1.09	0.68	4.67	2.05	2.27	3.14	1.49	1.44
FW[15]	135	35.32	39.82	0.73	0.02	0.45	0.52	3.45	1.44	1.91	2.43	0.19	2.28
FW[16]	174	16.72	0.57	0.64	0.64	0.56	0.51	3.41	1.25	1.90	2.79	0.39	2.35
FW[17]	195	17.81	1.70	0.74	0.50	0.46	0.59	3.24	1.33	1.20	3.39	0.10	2.77
FW[18]	210	0.90	0.77	0.75	0.49	0.41	0.60	3.35	1.28	2.15	5.52	1.20	2.65
FW[19]	568	33.57	72.70	2.25	0.52	0.37	0.54	4.70	6.25	35.16	11.26	2.01	9.52
FW[20]	679	12.69	1.00	0.43	0.76	0.65	0.85	3.27	2.84	1.69	2.99	0.99	1.37
Sum	2993	239.90	238.47	18.70	13.53	13.27	13.49	69.04	44.99	71.60	52.77	12.46	40.39
Average	149.65	12.00	11.92	0.93	0.68	0.66	0.67	3.45	2.25	3.58	2.64	0.62	2.02

B. PERFORMANCE

Hierarchy-view and Anomaly-view were tested using a total of 20 firewall policy files as listed in Table 22, (c) stands for count, and (s) stands for the time taken to abbreviate second. For Hierarchy-view, the processing speed per section and parsing was measured based on allowed policies, and in the case of Anomaly-view and Distributed-view, the policy anomalies were detected for each case in single and multiple firewall policy files, respectively, and anomaly detected policy numbers were counted. As a result, the parsing rate of the Hierarchy-view was 16.4 seconds on average. The processing rate of A-B, B-C and C-D visualization was 21.7, 2.7, 9.9 seconds on average, showing the highest A-B interval processing rate. The total number of policy anomalies in Anomaly-view were found in the order of Shadowing, Redundancy, Generalization, and Correlation. Table 23 shows that the Distributed-view function was tested with 2 pairs of distributed firewall policies. For distributed-view, the policy parsing time averaged 47.7 seconds and visualization time averaged 4.8 seconds.

Table 24 specifies the total number of lines in the 20 firewall policies. AnyPolicy-view measures the rate of visualization processing for six cases where the ANY policy may exist based on the allow policy. SearchResult-view measured the speed from SIP, DIP, and DPort to each input value and then to the resulting visualization. Top and Bottom Used-view also measured the top 10 policies, unused policies, and the rate at which both cases were visualized. As a result, the average rate of visualization is up to 12.00 seconds, and at least 0.66 seconds. The average processing speed of the visualization processing of SearchResult-view was identified by a maximum of 3.58 seconds and a minimum of 2.25 seconds. The average processing speed of visualizations for Top and Bottom Used-view was identified at 2.64 seconds, with the most frequently used policies averaging 0.62 seconds and unused policies averaging 2.02 seconds.

VII. CONCLUSION AND FUTURE WORK

A. CONCLUSION

In this paper, the newly designed visualization tool, HSViz, has been proposed to help the intuitive recognition of firewall policies and to effectively detect firewall anomaly policies in single or distributed firewall environments. HSViz provides six functions that intuitively visualize firewall policies. The six features are:Hierarchy-view, Anomaly-view, Distributed-view, AnyPolicy-view, SearchResult-view, and Top and Bottom Used-view. We can use the HSViz tool for various purposes such as identifying related services, improving firewall performance through detection of anomaly policies, checking excessively allowed services, detecting security vulnerabilities, identifying assets, using it as a security audit tool, and persuading decision makers. Test results of HSViz visualization effects with 24 firewall policies confirmed that effective recognition of firewall policies and detection anomalies are possible.

B. FUTURE WORK

Whereas the results of visualizations provided by HSViz are obvious advantages in the effective recognition and anomaly detection of firewall policies, there is a problem that as firewall policies become more complex and population increases. Also, there is a need to overcome the limitations of throughput. It is necessary to improve its functionality by reflecting the asset importance of the source and destination server present in the firewall policy so that it can be recognized as a risk in the firewall policy.

REFERENCES

[1] W. A. Johnston and V. J. Dark, "Selective attention," *Annu. Rev. Psychol.*, vol. 37, no. 1, pp. 43–75, 1986.

[2] C. Chabris and D. J. Simons, *The Invisible Gorilla: And Other Ways Our Intuitions Deceive US*. Harmony, 2010.

[3] C. Ware, *Information Visualization: Perception for Design*. San Mateo, CA, USA: Morgan Kaufmann, 2019.

[4] R. Marty, *Applied Security Visualization*. Reading, MA, USA: Addison-Wesley, 2008.

[5] E. Al-Shaer, "Managing firewall and network-edge security policies," in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp.*, vol. 1, Apr. 2004, p. 926.

[6] E. Al-Shaer, H. Hamed, R. Boutaba, and M. Hasan, "Conflict classification and analysis of distributed firewall policies," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 10, pp. 2069–2084, Oct. 2005.

[7] E. S. Al-Shaer and H. H. Hamed, "Firewall policy advisor for anomaly discovery and rule editing," in *Proc. IFIP/IEEE 8th Int. Symp. Integr. Netw. Manage.* Springer, 2003, pp. 17–30.

[8] E. S. Al-Shaer and H. H. Hamed, "Discovery of policy anomalies in distributed firewalls," in *Proc. IEEE INFOCOM*, vol. 4, Mar. 2004, pp. 2605–2616.

[9] M. Q. Ali, E. Al-Shaer, and T. Samak, "Firewall policy reconnaissance: Techniques and analysis," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 2, pp. 296–308, Feb. 2014.

[10] E. Al-Shaer and H. Hamed, "Design and implementation of firewall policy advisor tools," DePaul Univ., CTI, Chicago, IL, USA, Tech. Rep., 2002.

[11] V. T. Guimaraes, C. M. D. S. Freitas, R. Sadre, L. M. R. Tarouco, and L. Z. Granville, "A survey on information visualization for network and service management," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 285–323, 1st Quart., 2016.

[12] T. Tran, E. S. Al-Shaer, and R. Boutaba, "PolicyVis: Firewall security policy visualization and inspection," in *Proc. LISA*, vol. 7, 2007, pp. 1–16.

[13] B. Khan, M. K. Khan, M. Mahmud, and K. S. Alghathbar, "Security analysis of firewall rule sets in computer networks," in *Proc. 4th Int. Conf. Emerg. Secur. Inf., Syst. Technol.*, Jul. 2010, pp. 51–56.

[14] H. Kim, S. Ko, D. S. Kim, and H. K. Kim, "Firewall ruleset visualization analysis tool based on segmentation," in *Proc. IEEE Symp. Vis. Cyber Secur. (VizSec)*, Oct. 2017, pp. 1–8.

[15] U.-H. Kim, J.-M. Kang, J.-S. Lee, H.-S. Kim, and S.-Y. Jung, "Practical firewall policy inspection using anomaly detection and its visualization," *Multimedia Tools Appl.*, vol. 71, no. 2, pp. 627–641, Jul. 2014.

[16] J. McPherson, K.-L. Ma, P. Krystosk, T. Bartoletti, and M. Christensen, "PortVis: A tool for port-based detection of security events," in *Proc. ACM workshop Vis. Data Mining Comput. Secur. (VizSEC/DMSEC)*, 2004, pp. 73–81.

[17] D. Keim, F. Mansmann, J. Schneidewind, and T. Schreck, "Monitoring network traffic with radial traffic analyzer," in *Proc. IEEE Symp. Vis. Analytics Technol.*, Oct. 2006, pp. 123–128.

[18] F. Mansmann, D. A. Keim, S. C. North, B. Rexroad, and D. Sheleheda, "Visual analysis of network traffic for resource planning, interactive monitoring, and interpretation of security threats," *IEEE Trans. Vis. Comput. Graphics*, vol. 13, no. 6, pp. 1105–1112, Nov. 2007.

[19] F. Mansmann, T. Göbel, and W. Cheswick, "Visual analysis of complex firewall configurations," in *Proc. 9th Int. Symp. Vis. Cyber Secur. (VizSec)*, 2012, pp. 1–8.

[20] S. P. Morrissey and G. Grinstein, "Visualizing firewall configurations using created voids," in *Proc. 6th Int. Workshop Vis. Cyber Secur.*, Oct. 2009, pp. 75–79.

[21] H. Hu, G.-J. Ahn, and K. Kulkarni, "FAME: A firewall anomaly management environment," in *Proc. 3rd ACM Workshop Assurable Usable Secur. Configuration SafeConfig*, 2010, pp. 17–26.

[22] A. Saâdaoui, N. Ben Youssef Ben Souayeh, and A. Bouhoula, "FARE: FDD-based firewall anomalies resolution tool," *J. Comput. Sci.*, vol. 23, pp. 181–191, Nov. 2017.

HYUNJUNG LEE received the B.S. degree in computer engineering from the Seoul Women's University, Seoul, in 2005, and the M.S. degree in cybersecurity from Sungkyunkwan University, Seoul, in 2009. She is currently pursuing the Ph.D. degree with the School of Cybersecurity, Korea University, under the supervision of H. K. Kim.

Since 2011, she has been working as a Network and Security System Operator at KOSCOM, which developing and operating the core IT systems in Korean capital market and financial investment industry. Before joining KOSCOM, she was a Security System Operator at NFSOFT, from 2008 to 2010, and performed penetration testing in various industries when she worked for A3 Security and SK Infosec, from 2005 to 2007. She has authored one security book and has translated two security books. Her research interests include network security, vulnerability analysis, data analysis, and visualization.

SURYEON LEE is currently pursuing the B.S. degree in computer engineering from Seoul Women's University, Seoul.

She developed a Web-based shooting game site and a location-sharing platform for students. She also performed a project to analyze media and SNS responses using text mining techniques. Her research interests include data analysis, visualization, and data privacy.

KYOUNGGON KIM (Member, IEEE) received the B.S. degree in computer science from Soongsil University, in 2008, and the M.S. and Ph.D. degrees in information security from Korea University, in 2015 and 2020, respectively. He is currently an Assistant Professor with the Department of Forensic Sciences, Naif Arab University for Security and Sciences (NAUSS). He has performed penetration testing for over 130 clients in various industries when he worked for Deloitte, PwC, and Boutique Consulting Firms for over 15 years. He has authored a book on Internet hacking and security and has translated numerous security books. His research interests include cybercrime and network forensics, vulnerability analysis, smart city security, and CPS and the IoT security. He was awarded sixth place at DefCon CTF, in 2007, and a first prize from the First Hacking Defense Contest hosted by the Korea Information Security Agency.

HUY KANG KIM (Member, IEEE) received the B.S. degree in industrial management, the M.S. degree in industrial engineering, and the Ph.D. degree in industrial and systems engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 1998, 2000, and 2009, respectively. He is currently a Professor with the School of Cybersecurity, Korea University. He founded A3 Security Consulting, the first information security consulting company in South Korea, in 1999. Before joining Korea University, he was a Technical Director (TD) and the Head of the Information Security Department of NCSOFT, from 2004 to 2010, one of the most famous MMORPG companies in the world. His research interest includes solving many security problems in online games based on the user behavior analysis.