Jurnal Teknik Informatika (JUTIF)
Vol. 4, No. 3, June 2023, pp. 601-610

DOI: https://doi.org/10.52436/1.jutif.2023.4.3.1013
p-ISSN: 2723-3863
e-ISSN: 2723-3871

ANALYZING SURICATA ALERT DETECTION PERFORMANCE ISSUES BASED ON ACTIVE INDICATOR OF COMPROMISE RULES

Didit Hari Kuncoro Raharjo*[1], Muhammad Salman[2]

[1,2]Department of Electrical Engineering, Faculty of engineering, Universitas Indonesia, Indonesia
Email: [1]didit.hari@ui.ac.id, [2]muhammad.salman@ui.ac.id

(Article received: April 27, 2023; Revision: May 14, 2023; published: June 26, 2023)

Abstract

Many studies have been related to the Intrusion Detection System (IDS) performance analysis. Still, most focus on inspection performance on high-capacity networks with packet drop percentage as a performance parameter. Few studies are related to performance analysis in the form of detection accuracy based on the number of rules activated. This research will analyze the performance of IDS Suricata based on the number of active rules in the form of Indicator of Compromise (IoC), including IPRep, HTTP, DNS, MD5, and JA3. The analysis method focuses on the detection accuracy of varying the number of active rules up to 1 million, expressed in 5 scenarios. In scenarios 1 to 4, where IoC rules are tested separately, the reduction in detection accuracy performance starts to occur when the number of active rules is at 100,000 and continues to decrease when the number reaches 1 million. However, in scenario 5, where the IoC rules are tested together, the percentage of rules detection accuracy decreases when the number of active rules from each IoC is less than 10,000. The percentage decrease in detection accuracy performance in scenario five can occur with an average reduction of 19.64%. Even further in scenario 5, when the total number of rules reaches 1,000,000 or 200,000 from each IoC, IDS Suricata fails to detect all rules (detection percentage is 0%). This research show that the higher number of rules activated, the decrease in the Suricata IDS performance in terms of detection accuracy.

Keywords: *Detection, IoC, Performance, Suricata.*

1. INTRODUCTION

Implementing a security perimeter in the form of an Intrusion Detection System (IDS) via an anomaly monitoring mechanism for data traffic on the network and data on endpoints is one of the methods for defending against cyber threats [1-3]. IDS (either Network IDS or Host IDS) is a tool that provides alerts if a data context matches the parameters (rules) based on the features or characteristics of data packet anomalies. Depending on the kind of rule, the IDS-applied rules will seek for matching data in the header or information content; if found, a warning message will be displayed [4, 5]. Information in the form of indicators of compromise (IoC) may be used to construct IDS rules. IoC is one of the tactical components of Cyber Threat Intelligence (CTI) reports that can be applied to security perimeter devices (IDS, Firewalls, etc.) as a parameter indicating the incidence of cyber-attacks. In addition to IPRep, Hash File, DNS (Domain Name System), URL (Uniform Resource Locator), and SSL fingerprint (JA3) are often used forms of IoC [6-9].

Suricata is an open-source IDS tool that applies a Signature-Based detection mechanism and is built on the network side (NIDS), with the capability to inspect data packets utilizing multi-threading to detect huge quantities of data packets [1, 10-13]. Suricata will inspect data packets by decoding packet metadata and then carrying out a matching procedure with the rules through Suricata's detection engine. With the concept of the IDS signature, it is only reasonable for the blue team to think that it must create as many comprehensive rules as possible to identify system threats. Logically, the greater the completeness of the rules used to IDS, the greater the probability of detection.

Speaking of NIDS, there are open-source detection engines other than Suricata, including Snort and Zeek (formerly Bro). The performance test research conducted by [1, 10, 12-18] shows that Suricata is an IDS superior to competitors. However, most of the performance tests that have been conducted focus on IDS inspection capabilities based on network throughput, represented in terms of packet drop percentage.

Iyengar conducted a performance test between Snort and Suricata, focusing on RAM usage, number of packet drops, and CPU usage, showing Suricata to be better at detecting high-speed networks [1].

Qinwen Hu et al. evaluated the efficacy of Snort and Suricata in terms of drop rate and precision [10]. By evaluating detection on a 100Gbps data stream for drop rate performance, it is determined that Snort has a higher drop rate than Suricata. At the same time, it was evaluating accuracy using the Pytbull framework, where both demonstrate 100 percent precision.

Wong et al. tested Suricata on SCADA by implementing the ENIP (EtherNet Industrial Protocol) protocol [12]. Performance testing focused on RAM usage, packet drop, and CPU utilization at network throughput (18Mbps), where standard rules (18,000 rules) plus 30 ENIP-related rules were used. The tests show that the addition of ENIP rules does not affect the performance of Suricata, and the CPU utilization and packet drop limits are still good enough to run on SCADA networks.

Waleed et al. tested three varieties of intrusion detection systems, including Snort (versions 2 and 3), Suricata, and Zeek [13]. Suricata is preferable to Snort and Zeek based on the percentage of dropped packets based on network throughput during performance testing.

Murphy compared the efficacy of Snort and Suricata in his dissertation by measuring detection accuracy, RAM and CPU efficiency, and the number of packet drops [14]. Tests indicate that Suricata is superior in terms of detection accuracy, while Snort is superior in terms of RAM usage efficiency. Regarding packet drop, the results of the two IDSs were similar.

Brumen et al. tested the efficacy of Suricata and Snort on two operating systems (Windows and Linux) in terms of packet loss and device utilization (CPU and RAM) [15]. Even though Suricata consumed more device resources (CPU and RAM), Snort's percentage of dropped packets was significantly higher. In addition, in terms of operating system utilization, Linux is superior to Windows because it has fewer packet drops.

Ernawati et al. evaluated the performance of three types of intrusion detection systems (Suricata, Portsentry, and Port Attack Scan Detector/PSAD) based on accuracy parameters, resource utilization (CPU, RAM, Disk), and detection speed [16]. Suricata and PSAD have the highest performance, particularly regarding detection accuracy (100%) and resource consumption.

A performance comparison between Suricata and Snort3 conducted by Hover revealed that Suricata had a more significant number of detections than Snort3. They were using the Pytbull framework for testing based on default rules. However, the research does not indicate how many alerts should be noticed, thus, it cannot state whether the number of rules discovered is 100 percent ideal [17].

Jian Guo et al. conducted a performance evaluation of Suricata, focusing on network throughput that might exceed 20Gbps after tweaking to reduce packet loss. This research did not, however, figure out the percentage of detection based on the number of rules [19].

From the above research, all tested how much packet drop occurs based on network throughput. The hypothesis that was raised was that the smaller the percentage of packet drop, the greater all the data was successfully inspected. However, the hypothesis regarding how many percent of accuracy is generated (testing between the number of attacks compared to the rules that are owned) in doing detection needs to be more visible. Some research test accuracy, but the framework only shows which IDS generates more alerts, not how many rules should be detected.

Lukaseder et al. showed that the percentage of packet drops does not affect the detection accuracy of Suricata [18]. Although Suricata has a higher percentage of packet drops than Snort (in testing over a 7 Gbps network throughput), the accuracy in detecting the number of attacks per minute produced by Suricata is higher than Snort.

Raharjo et al. tested the detection accuracy of Suricata based on the number of rules activated [20]. When the number of rules (IP Reputation/IPRep) reached one million, the percentage of rule detection reduced to 16.24%. It should be noted that this research was limited to using only one parameter from the IoC (IPRep) and only using the default configuration from Suricata (no tuning configuration). The results of this research may break the previous logical assumption, where the greater the completeness of the rules used to IDS, the greater the probability of detection.

Based on these facts, this research is expected to determine whether the amount and variety of rules affect the accuracy of IDS in performing detection. It will build on previous research [20] where additional performance evaluation is required with IoC types besides IPRep and tuning configuration. This research will evaluate the performance of IDS Suricata in identifying rules for five forms of IoC (IPrep, HTTP, DNS, MD5 Hash, SSL JA3) in five distinct scenarios.

2. RESEARCH METHOD

2.1. Research Stages

The research consists of two phases: the design and testing phases, as shown on Figure 1. The design phase includes the specification of the evaluation testing environment, the design of dataset generation, the evaluation scenario design, and performance parameters. In the testing phase, the outcomes of the design phase will be implemented to produce a generated dataset, rules parsed from dataset pcap, and the detection percentage for each scenario.

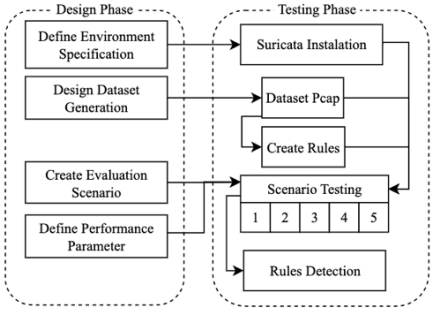

Figure 1. Research stage.

2.2. Environment Specification

The specifications of the devices (on virtualization environment) utilized in this research are listed in Table 1:

Table 1. Environment specification

Item	Specification
Operating System	Ubuntu Focal 20.04.4 LTS
Processor	Intel(R) Xeon(R) Silver 4114 CPU @ 2.20GHz; cores=16; threads=16
RAM	16 GB
Harddisk	200 GB

Suricata is an IDS with multi-threading capabilities, so CPU power, number of threads, and RAM (for tuning requirements), will be highly influential. The specifications listed in Table 1 will be one of the important benchmarks that affect the research conclusion.

2.3. Design of Dataset Generation

The research will utilize Pcap as a dataset. The utilized Pcap is a data packet from network traffic/traffic that contains IoC parameters. IDS Suricata will detect this Pcap to assess the performance accuracy. To generate the desired Pcap dataset, the python-scapy program modifies network traffic data packets to generate the required Pcap data set [21]. Figure 2 explains the process design for generating data sets with python-scapy:

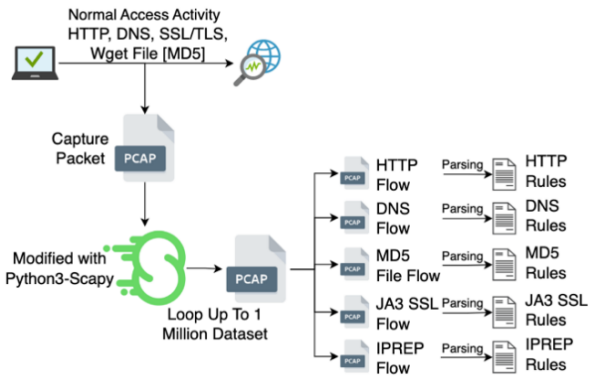

Figure 2. Design of dataset traffic generation

Based on figure 2, the generation process is executed by altering the parameters of data packets. Initial normal data packets are captured via application access activities (HTTP, HTTPS, DNS), downloading files to get hash values (wget data), and scanning processes (multiple IPs). The regular data packet is modified to build a data packet containing one million IoC parameters.

2.4. Evaluation Scenario

Scenarios are designed in such a way as to cover possible conditions to be used in the detection process using IDS.

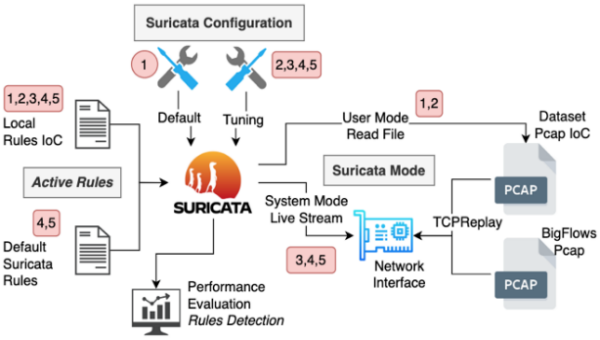

Figure 3. Evaluation Scenario

Figure 3 above is the scenario flow used in this research, where the performance testing process will be separated into five types of test scenarios (the numbers indicate the scenario flow) using the created Pcap dataset and IOC rules. The comparison between the scenarios represented in Figure 3 can be clarified in Table 2 below:

Table 2. Comparison evaluation scenario

Parameter	Scenario				
	1	2	3	4	5
Suricata Configuration					
Default	✓	×	×	×	×
Tuning	×	✓	✓	✓	✓
Suricata Run Mode					
User Mode	✓	✓	×	×	×
System Mode	×	×	✓	✓	✓
Dataset Testing					
Tested separately	✓	✓	✓	✓	×
Tested concurrently	×	×	×	×	✓
Pcap BigFlows	×	×	✓	✓	✓
Activation of suricata.rules	×	×	×	✓	✓

The following is an explanation of the parameters in table 2:

1.) Suricata Configuration

There are two types of Suricata configuration implementation, namely default and tuning. In default configuration, the use of multi-threads will automatically utilize half of the device's processors/cores when Suricata is installed. While the tuning configuration will change the parameters so that Suricata works more effectively, which in this study will be based on research findings presented at the 2016 Suricata Conference (SURICON) [22].

Table 3. Comparison: default and tuning configuration

Parameter	Default	Tuning
Defrag memcap	32mb	2.000mb
Defrag hash-size	65.536	1.000.000
Flow memcap	128mb	4.000mb
Flow hash-size	65.536	10.000.000
Flow prealloc	10.000	10.000.000
Stream memcap	64mb	1.000mb
Reassembly memcap	256mb	2.000mb
Reassembly depth	1mb	2mb
Host memcap	32mb	2.000mb
Set cpu-affinity	*no*	*yes*
worker-cpu-set prio default	*medium*	*high*
AF-Packet Ring Size	2.048	20.480
AF-Packet block-size	32.768	393.216

Tuning settings of IDS Suricata were restricted to modifying memcap, hash size, and CPU affinity in this research. Table 3 compares the default Suricata configuration to the tuning that was deployed in this research:

2.) Suricata Run Mode

There are two types of Suricata run mode implementation, namely user mode (read file) and system mode (live monitoring). User mode will directly read the Pcap file (option -r), without going through the device's network interface where Suricata is installed. System mode will read directly streamed data packets from the network interface (option -i). The TCPReplay application will stream the pcap dataset directly to the network interface.

3.) Dataset Testing

There are two forms of implementation of using datasets and IOC rules in this research, ie tested separately or concurrently. In tested mode separately, the IoC rules activated in Suricata adjust the dataset being tested, with one type of IoC per test. While in other mode will test all Pcap datasets concurrently, including the activation of IoC rules on Suricata and all five types of IoC datasets.

4.) Pcap BigFlows

For the test to be close to actual conditions, the test scenario in this research will also stream data containing normal activities in parallel. The dataset used is BigFlows.pcap which contains general daily traffic samples from TCP, UDP, File Transfer, HTTP/browsing access, Chat, and others [23].

5.) Activation of suricata.rules

Suricata has access to the rules provided by Emerging Threads, both free/open source (community rules) and paid (pro), which are then called suricata.rules [24]. This research used the free version of suricata.rules (using the suricata-update function) in December 2022 and found 29,192 active rules.

2.5. Performance Parameter

The performance parameter for Suricata is the percentage value of detection rules. This value derives from the following:

$$\frac{\sum \text{Deteted alert IoC}}{\sum \text{Active rules IoC}} \text{ x } 100\% \quad (1)$$

Formula (1) is the ratio of the number of alerts Suricata can detect compared to the number of alerts that Suricata should have detected. According to formula (1), the performance of IDS Suricata is evaluated as 100 percent if it can detect all IoC parameters from the tested dataset following the planned scenario.

3. HASIL DAN PEMBAHASAN

3.1. Dataset Generation Results

The generation of datasets produces five pcap datasets that have data traffic containing IoC, as detailed in Table 4.

Table 4. Dataset details

IoC Dataset	Size (byte)	Number of lines
HTTP	2.170.666.690	7.000.000
DNS	230.000.022	2.000.000
Hash MD5	1.327.186.464	10.000.000
SSL JA3	10.583.000.024	27.000.000
IPRep	344.000.024	4.000.000

The pcap generated in Table 4 is a pcap containing stream packet data IoC with variations reaching 1 million. The number of lines of each IoC can be different, depending on the number of packets transmitted in 1 type of IoC. For example, in http IoC, there are 7 million lines, which means that in 1 http IoC data packet stream, there are 7 data packets. To minimize detection bias in scenarios 4 and 5 between local rules IoC and suricata.rules, the generated dataset is modified so that it is not identified by suricata.rules.

3.2. Rules Creation

Local IoC rules will be extracted from the pcap dataset created. Specifically for IPRep and MD5, the number of active rules from this category is only one, not a million. But its rules will refer to a file containing a million IP Reputation or MD5 lists configured on Suricata (suricata.yaml). The properties of the rules were successfully retrieved from the pcap dataset and the suricata.rules utilized in the test are in table 5.

Table 5. Rules details

Rules Name	Size (byte)	Number of Rules / List	SID Range
local-dns.rules	124.555.556	1.000.000	1.000.001 - 2.000.000
local-http.rules	178.222.229	1.000.000	3.000.001 - 4.000.000
local-md5.rules	119	1.000.000	4.000.001
local-ja3.rules	126.888.896	1.000.000	4.000.002 - 5.000.001
local-ip.rules	97	1.000.000	5.000.002
suricata.rules	21.478.472	29.192	2.000.001 - 3.000.000

Each rule must have a unique Signature Identification (SID) (as described in Table 5 in the SID range) so as not to conflict with the identity of other rules.

3.3. Scenario Result

Each element of the rules matrix will be tested ten times, after which the average detection % will be

calculated using the formula (1). For each case, the result and analysis are described as follows:

1.) Scenario 1

Scenario 1 shows the conditions for utilizing Suricata with its default settings to read a dataset pcap file. Each dataset and rule from IoC will be evaluated independently, and the percentage value of rule detection will be determined for each number of activated rules. Figure 4 depicts the outcomes of testing scenario one graphically

Figure 4 demonstrates that until the number of rules reaches 10,000, IDS Suricata can identify all rules (100 %) with a 100 percent detection rate. However, as the number of rules climbed to 1,000,000, the detection percentage for HTTP, MD5, and JA3 rules decreased significantly, reaching beyond 50 %.

NUMBER OF ACTIVE RULES	RULES DETECTION (%)				
	IP-SC1	HTTP-SC1	DNS-SC1	MD5-SC1	JA3-SC1
1.000	100,00	100,00	100,00	100,00	100,00
2.000	100,00	100,00	100,00	100,00	100,00
3.000	100,00	100,00	100,00	100,00	100,00
4.000	100,00	100,00	100,00	100,00	100,00
5.000	100,00	100,00	100,00	100,00	100,00
6.000	100,00	100,00	100,00	100,00	100,00
7.000	100,00	100,00	100,00	100,00	100,00
8.000	100,00	100,00	100,00	100,00	100,00
9.000	100,00	100,00	100,00	100,00	100,00
10.000	100,00	100,00	100,00	100,00	100,00
100.000	100,00	38,50	100,00	32,38	16,20
200.000	100,00	37,66	100,00	16,19	8,10
300.000	81,60	37,46	100,00	10,79	5,40
400.000	61,20	37,46	100,00	8,10	4,05
500.000	48,96	37,27	100,00	6,48	3,24
600.000	40,80	37,04	100,00	5,40	2,70
700.000	34,97	37,00	100,00	4,63	2,31
800.000	30,60	36,95	100,00	4,05	2,02
900.000	27,20	36,78	100,00	3,60	1,80
1.000.000	24,48	36,56	100,00	3,24	1,62

Figure 4. Rules detection accurary on scenario 1

Only after the total number of IPRep rules reached 300,000 did the percentage fall from 80 % to 24 % occur. As for DNS rules, the system is stable and capable of detecting the entirety (100 %) until the number of rules reaches one million.

2.) Scenario 2

NUMBER OF ACTIVE RULES	RULES DETECTION (%)				
	IP-SC2	HTTP-SC2	DNS-SC2	MD5-SC2	JA3-SC2
1.000	100,00	100,00	100,00	100,00	100,00
2.000	100,00	100,00	100,00	100,00	100,00
3.000	100,00	100,00	100,00	100,00	100,00
4.000	100,00	100,00	100,00	100,00	100,00
5.000	100,00	100,00	100,00	100,00	100,00
6.000	100,00	100,00	100,00	100,00	100,00
7.000	100,00	100,00	100,00	100,00	100,00
8.000	100,00	100,00	100,00	100,00	100,00
9.000	100,00	100,00	100,00	100,00	100,00
10.000	100,00	100,00	100,00	100,00	100,00
100.000	100,00	100,00	100,00	100,00	100,00
200.000	100,00	100,00	100,00	100,00	100,00
300.000	100,00	100,00	100,00	100,00	100,00
400.000	100,00	100,00	100,00	100,00	100,00
500.000	100,00	100,00	100,00	100,00	100,00
600.000	100,00	100,00	100,00	100,00	86,87
700.000	100,00	100,00	100,00	100,00	74,46
800.000	100,00	100,00	100,00	100,00	65,15
900.000	100,00	100,00	100,00	100,00	57,92
1.000.000	100,00	100,00	100,00	100,00	52,12

Figure 5. Rules detection accuracy on scenario 2

Scenario 2 presents the identical testing conditions as Scenario 1 (user mode) but with a Suricata tuning configuration. Figure 5 depicts the results of testing scenario two graphically:

Figure 5 demonstrates that Suricata's tuning configuration has led to a significant change. For HTTP, DNS, MD5, and stable IPRep rules, it can identify the entirety (100 %) till the rule count reaches one million. Even while there is still a performance decline for the JA3 rules when compared to scenario 1, a new performance decline occurs when the number of rules reaches 600,000, and it remains above 50 % to detect rules exceeding 1,000,000.

3.) Scenario 3

Scenario 3 spcifies the same test process conditions as Scenario 2 (tuning configuration), except that Suricata is executed in system mode (live monitoring) on the network interface utilizing TCP Replay. The findings of testing scenario three are depicted graphically in Figure 6.

NUMBER OF ACTIVE RULES	RULES DETECTION (%)				
	IP-SC3	HTTP-SC3	DNS-SC3	MD5-SC3	JA3-SC3
1.000	100,00	100,00	100,00	100,00	100,00
2.000	100,00	100,00	100,00	100,00	100,00
3.000	100,00	100,00	100,00	100,00	100,00
4.000	100,00	100,00	100,00	100,00	100,00
5.000	100,00	100,00	100,00	100,00	100,00
6.000	100,00	100,00	100,00	100,00	100,00
7.000	100,00	100,00	100,00	100,00	100,00
8.000	100,00	100,00	100,00	100,00	100,00
9.000	100,00	100,00	100,00	100,00	100,00
10.000	100,00	100,00	100,00	100,00	100,00
100.000	100,00	100,00	84,94	100,00	100,00
200.000	100,00	100,00	89,88	100,00	100,00
300.000	100,00	97,33	85,71	100,00	96,71
400.000	100,00	94,63	86,04	100,00	95,23
500.000	100,00	97,56	82,18	100,00	80,04
600.000	100,00	80,42	80,54	100,00	44,97
700.000	100,00	56,55	79,94	100,00	15,36
800.000	100,00	36,30	84,67	100,00	0,00
900.000	100,00	23,59	85,55	100,00	0,00
1.000.000	99,81	20,20	70,06	100,00	0,00

Figure 6. Rules detection accuracy on scenario 3

Figure 6 illustrates that under live monitoring and tuning situations, IDS Suricata can identify all rules (100 %) until the number of rules approaches 10,000. This condition is superior to previous experiments that lacked tweaking, as just fifty percent of the server's processor is utilized [20]. IPrep and MD5 rules continue to be able to identify (> 99.8 %) until the number of rules reaches one million. After the previous two cases were perfectly recognized, the percentage of correctly detected DNS rules begins to fall when the number of rules reaches 100,000 and continues to decrease to 70 % when it reaches 1,000,000.

The performance of HTTP rules began to degrade when the number of rules hit 300,000, then declined by 56 % when the number of rules reached 700,000 and continued to decline substantially until just 20% of rules reached 1,000,000. As for the JA3 rules, when compared to scenario 2, the number of rules at 300,000 shows a slight decrease in performance, which decreases dramatically

beginning at the number of rules at 700,000 (15 %) and continues to decrease until it is no longer detected (0%) when the number of rules reaches 1,000,000.

4.) Scenario 4

Scenario 4 describes the same test process conditions as Scenario 3 (live monitoring and tuning configuration) but adds the number of active rules, excluding local IoC rules, via the suricata.rules file. Figure 7 is a graphical representation of scenario four testing outcomes.

Figure 7 illustrates that under live monitoring, tuning, and the addition of suricata.rules, IDS Suricata can detect all rules (100 %) until the number of rules reaches 10,000. IPrep rules can detect very well, despite a little drop (>99.4 %), until the number of rules approaches 1,000,000. The percentage of DNS rules begins to decline when the number of rules exceeds 100,000 and continues to decrease to 47 % when the number of rules reaches one million. The percentage of MD5 rules begins to fall slightly when the number of rules reaches 400,000 and continues to decline to 82 % when the number of rules reaches 1,000,000.

NUMBER OF ACTIVE RULES	RULES DETECTION (%)				
	IP-SC4	HTTP-SC4	DNS-SC4	MD5-SC4	JA3-SC4
1.000	100,00	100,00	100,00	100,00	100,00
2.000	100,00	100,00	100,00	100,00	100,00
3.000	100,00	100,00	100,00	100,00	100,00
4.000	100,00	100,00	100,00	100,00	100,00
5.000	100,00	100,00	100,00	100,00	100,00
6.000	100,00	100,00	100,00	100,00	100,00
7.000	100,00	100,00	100,00	100,00	100,00
8.000	100,00	100,00	100,00	100,00	100,00
9.000	100,00	100,00	100,00	100,00	100,00
10.000	100,00	100,00	100,00	100,00	100,00
100.000	100,00	96,48	83,75	100,00	94,83
200.000	100,00	89,32	90,21	100,00	85,73
300.000	100,00	92,74	95,36	100,00	48,30
400.000	100,00	78,44	87,57	99,76	6,56
500.000	100,00	53,93	83,18	99,94	0,00
600.000	100,00	30,90	82,02	99,96	0,00
700.000	100,00	23,59	83,76	99,54	0,00
800.000	100,00	15,66	87,16	97,81	0,00
900.000	99,95	3,95	78,66	91,02	0,00
1.000.000	99,45	0,00	47,44	82,68	0,00

Figure 7. Rules detection accuracy on scenario 4

The performance of HTTP rules begins to diminish when the number of rules exceeds 100,000, then decreases by 53 % when the number of rules reaches 500,000, and continues to drop until it is unable to detect (0%) when the number of rules reaches 1,000,000. As for the JA3 rules, the performance began to degrade at 100,000 rules, declined significantly at 400,000 rules (by 6 %), and continued to decrease until it was no longer detected (0%) at 500,000 rules.

5.) Scenario 5

Scenario 5 covers the identical test process conditions as Scenario 4 but simultaneously activates all five types of local IoC rules and suricata.rules. Figure 8 on the following page depicts the results of testing scenario five graphically.

Figure 8 illustrates that under live monitoring, tuning, and all IoC rules and suricata.rules are enabled, the number of detected rules decreases by less than 100 percent when the number of active rules for each IoC is less than 10,000. As the number of IoC rules increases, the frequency of all rule types decreases.

Initial testing of the IPRep and DNS rules separately in scenarios 1 through 4 revealed a pretty good detection rate; however, when tested with other rules in scenario 5, the detection rate decreased significantly.

NUMBER OF ACTIVE RULES	RULES		RULES DETECTION (%)				
	@IOC	TOTAL	IP-SC5	HTTP-SC5	DNS-SC5	MD5-SC5	JA3-SC5
	1.000	5.000	38,71	96,46	42,48	97,37	100,00
	2.000	10.000	42,91	93,73	61,18	94,86	100,00
	3.000	15.000	39,92	94,27	49,89	99,38	100,00
	4.000	20.000	43,63	83,05	51,76	94,95	100,00
	5.000	25.000	58,21	88,18	65,37	98,93	100,00
	6.000	30.000	56,40	89,45	64,75	97,08	100,00
	7.000	35.000	48,42	76,97	57,44	89,38	100,00
	8.000	40.000	50,12	84,45	60,60	96,10	100,00
	9.000	45.000	54,03	81,86	60,07	93,80	99,99
	10.000	50.000	47,92	87,57	56,68	96,55	99,86
	20.000	100.000	34,03	91,83	43,02	96,42	100,00
	30.000	150.000	36,85	94,18	50,07	98,93	99,98
	40.000	200.000	27,71	95,56	36,27	97,58	99,85
	50.000	250.000	26,75	91,53	34,82	94,19	97,47
	60.000	300.000	9,83	82,68	20,37	97,77	98,18
	70.000	350.000	11,19	74,78	20,84	90,39	90,93
	80.000	400.000	14,20	75,82	22,28	91,72	93,38
	90.000	450.000	8,74	66,82	15,27	80,97	94,21
	100.000	500.000	6,27	45,38	11,53	67,83	90,35
	120.000	600.000	1,94	4,43	1,71	8,72	70,27
	140.000	700.000	0,00	3,70	0,00	10,65	37,17
	160.000	800.000	0,00	0,00	0,00	1,08	26,05
	180.000	900.000	0,00	0,00	0,00	0,00	7,63
	200.000	1.000.000	0,00	0,00	0,00	0,00	0,00

Figure 8. Rules detection accuracy on scenario 5

Compared to the preceding case, the percentage of MD5 and HTTP rules from the initial 1,000 rules has decreased by less than one hundred percent (100 %). For the JA3 rules, the proportion remains at 100 percent until the number of rules reaches 40,000 but then begins to decrease as the number of JA3 rules increases. When there are 200,000 rules for each IoC (for a total of 1,000,000 active IoC rules), the percentage of detection for the five rules is 0 % or not detected.

3.4. Gap Detection

The most frequently utilized scenarios for using Suricata IDS as an anomaly detection process are scenarios 4 and 5, which involve live monitoring using suricata.rules and local rules, out of the five scenarios conducted. Comparing the results of Scenario 4 and Scenario 5 and the type of IoC, there is a substantial difference in the proportion of detection rules between the two scenarios.

1.) Rules IPRep

Figure 9 depicts that the potential for undetectable IPRep rules is highly valued when executed in conjunction with other IoC in Scenario 5, ranging from 41 to 100 % with an average of 59.46%.

2.) Rules HTTP

As depicted in Figure 10, the potential for undetected HTTP rules is relatively significant when run in conjunction with other IoC in Scenario 5, ranging from 3 to 100 %, with an average of 23.22 %.

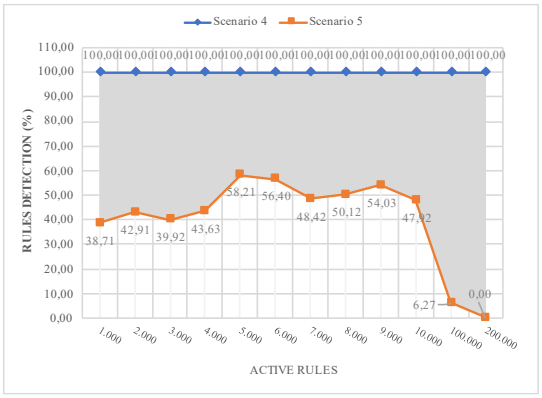

Figure 9. Gap IPRep rules detection between scenario 4 & 5

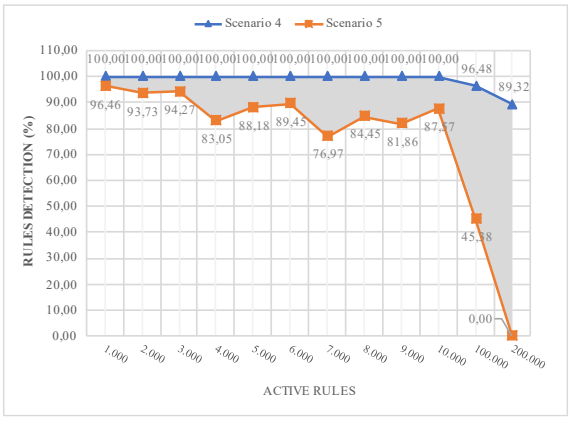

Figure 10. Gap HTTP rules detection between scenario 4 & 5

3.) Rules DNS

Figure 11 depicts that the chance for undetected DNS rules is highly valued when run in conjunction with other IoC in Scenario 5, ranging from 34 to 100% with an average of 51.52 %.

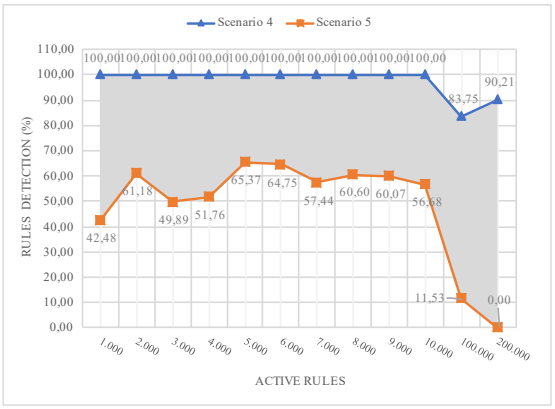

Figure 11. Gap DNS rules detection between scenario 4 & 5

4.) Rules MD5

Figure 12 depicts that the potential for undetectable MD5 rules is relatively significant when executed in conjunction with other IoC in Scenario 5, ranging from 1 to 100 % with an average of 14.48 %.

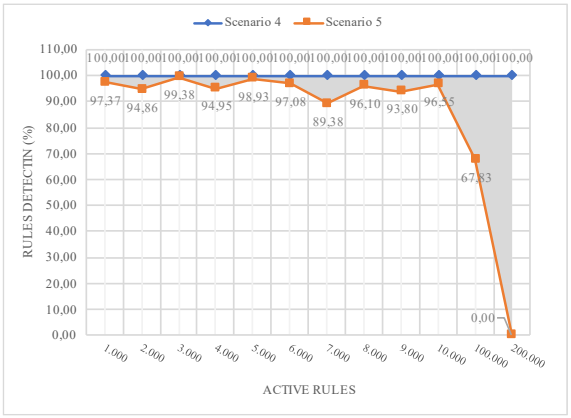

Figure 12. Gap MD5 rules detection between scenario 4 & 5

5.) Rules JA3

Figure 13 depicts that the potential for undetectable JA3 rules is rather low when executed in conjunction with other IoC in Scenario 5, ranging from 0 % to 100 %, with an average of 9.15 %.

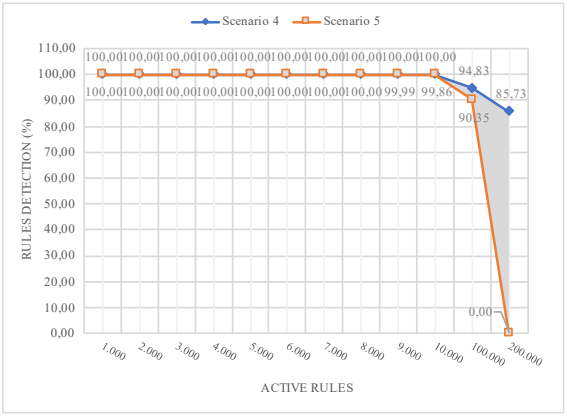

Figure 13. Gap JA3 rules detection between scenario 4 & 5

NUMBER OF ACTIVE RULES	IPREP RULES DETECTION (%)						HTTP RULES DETECTION (%)						DNS RULES DETECTION (%)						MD5 RULES DETECTION (%)						JA3 RULES DETECTION (%)					
	SC1	SC2	SC3	SC4	SC5	GAP [4-5]	SC1	SC2	SC3	SC4	SC5	GAP [4-5]	SC1	SC2	SC3	SC4	SC5	GAP [4-5]	SC1	SC2	SC3	SC4	SC5	GAP [4-5]	SC1	SC2	SC3	SC4	SC5	GAP [4-5]
1.000	100,00	100,00	100,00	100,00	38,71	61,29	100,00	100,00	100,00	100,00	96,46	3,54	100,00	100,00	100,00	100,00	42,48	57,52	100,00	100,00	100,00	100,00	97,37	2,63	100,00	100,00	100,00	100,00	100,00	0,00
2.000	100,00	100,00	100,00	100,00	42,91	57,09	100,00	100,00	100,00	100,00	93,73	6,27	100,00	100,00	100,00	100,00	61,18	38,82	100,00	100,00	100,00	100,00	94,86	5,14	100,00	100,00	100,00	100,00	100,00	0,00
3.000	100,00	100,00	100,00	100,00	39,92	60,08	100,00	100,00	100,00	100,00	94,27	5,73	100,00	100,00	100,00	100,00	49,89	50,11	100,00	100,00	100,00	100,00	99,38	0,62	100,00	100,00	100,00	100,00	100,00	0,00
4.000	100,00	100,00	100,00	100,00	43,63	56,37	100,00	100,00	100,00	100,00	83,05	16,95	100,00	100,00	100,00	100,00	51,76	48,24	100,00	100,00	100,00	100,00	94,95	5,05	100,00	100,00	100,00	100,00	100,00	0,00
5.000	100,00	100,00	100,00	100,00	58,21	41,79	100,00	100,00	100,00	100,00	88,18	11,82	100,00	100,00	100,00	100,00	65,37	34,63	100,00	100,00	100,00	100,00	98,93	1,07	100,00	100,00	100,00	100,00	100,00	0,00
6.000	100,00	100,00	100,00	100,00	56,40	43,61	100,00	100,00	100,00	100,00	89,45	10,55	100,00	100,00	100,00	100,00	64,75	35,25	100,00	100,00	100,00	100,00	97,08	2,92	100,00	100,00	100,00	100,00	100,00	0,00
7.000	100,00	100,00	100,00	100,00	48,42	51,58	100,00	100,00	100,00	100,00	76,97	23,03	100,00	100,00	100,00	100,00	57,44	42,56	100,00	100,00	100,00	100,00	89,38	10,62	100,00	100,00	100,00	100,00	100,00	0,00
8.000	100,00	100,00	100,00	100,00	50,12	49,88	100,00	100,00	100,00	100,00	84,45	15,55	100,00	100,00	100,00	100,00	60,60	39,40	100,00	100,00	100,00	100,00	96,10	3,90	100,00	100,00	100,00	100,00	100,00	0,00
9.000	100,00	100,00	100,00	100,00	54,03	45,97	100,00	100,00	100,00	100,00	81,86	18,14	100,00	100,00	100,00	100,00	60,07	39,93	100,00	100,00	100,00	100,00	93,80	6,20	100,00	100,00	100,00	100,00	99,99	0,01
10.000	100,00	100,00	100,00	100,00	47,92	52,08	100,00	100,00	100,00	100,00	87,57	12,43	100,00	100,00	100,00	100,00	56,68	43,32	100,00	100,00	100,00	100,00	96,55	3,45	100,00	100,00	100,00	100,00	99,86	0,14
100.000	100,00	100,00	100,00	100,00	6,27	93,73	38,50	100,00	100,00	96,48	45,38	51,10	100,00	100,00	84,94	83,75	11,53	72,22	32,38	100,00	100,00	100,00	67,83	32,17	16,20	100,00	100,00	94,83	90,35	4,49
200.000	100,00	100,00	100,00	100,00	0,00	100,00	37,66	100,00	100,00	89,32	0,00	89,32	100,00	100,00	89,88	90,21	0,00	90,21	16,19	100,00	100,00	100,00	0,00	100,00	8,10	100,00	100,00	85,73	0,00	85,73
300.000	81,60	100,00	100,00	100,00			37,46	100,00	97,33	92,74			100,00	100,00	85,71	95,36			10,79	100,00	100,00	100,00			5,40	100,00	96,71	48,30		
400.000	61,20	100,00	100,00	100,00			37,46	100,00	94,63	78,44			100,00	100,00	86,04	87,57			8,10	100,00	100,00	99,76			4,05	100,00	95,23	6,56		
500.000	48,96	100,00	100,00	100,00			37,27	100,00	97,56	53,93			100,00	100,00	82,18	83,18			6,48	100,00	100,00	99,94			3,24	100,00	80,04	0,00		
600.000	40,80	100,00	100,00	100,00			37,04	100,00	80,42	30,90			100,00	100,00	80,54	82,02			5,40	100,00	100,00	99,96			2,70	86,87	44,97	0,00		
700.000	34,97	100,00	100,00	100,00			37,00	100,00	56,55	23,59			100,00	100,00	79,94	83,76			4,63	100,00	100,00	99,54			2,31	74,46	15,36	0,00		
800.000	30,60	100,00	100,00	100,00			36,95	100,00	36,30	15,66			100,00	100,00	84,67	87,16			4,05	100,00	100,00	97,81			2,02	65,15	0,00	0,00		
900.000	27,20	100,00	100,00	99,95			36,78	100,00	23,59	3,95			100,00	100,00	85,55	78,66			3,60	100,00	100,00	91,02			1,80	57,92	0,00	0,00		
1.000.000	24,48	100,00	99,81	99,45			36,56	100,00	20,20	0,00			100,00	100,00	70,06	47,44			3,24	100,00	100,00	82,68			1,62	52,12	0,00	0,00		

Figure 14. Recapitulation of analysis results on scenarios 1 - 5

Figure 14 summarizes the performance analysis from scenarios 1–5, grouped by IoC type. There is a trend of decreasing detection accuracy percentage as the number of active rules increases. When the IoC rules are activated variably (simultaneously as in scenario 5), at the point where the number of rules from each IoC is 200,000, Suricata's rules detection accuracy is 0 % for all types of IoC rules.

4. DISCUSSION

This research show that tuning the Suricata IDS configuration is essential for achieving the best performance in anomaly detection. Compared to the default setting, performance testing demonstrates a significant improvement. This is readily apparent when comparing the outcomes of scenarios 1 and 2.

With the specified device specifications (refer to Table 1), IDS Suricata can still detect all rules (100%) in scenarios 1 through 4, where the rules testing process is conducted independently until the number of rules reaches 10,000. A significant performance decline occurred when the number of rules exceeded 100,000. However, in scenario 5, when the number of active rules for each IoC is less than 10,000 (or the total number of active rules is less than 50,000), the decrease in the number of detected rules is even less than 100 percent. And the number of detected rules will continue to decrease until they are no longer detected by Suricata when the number of active rules from each IoC is only 20 % of the one million active rules in scenarios 1–4. Even with only 1,000 rules for each IoC (or total IoC rules of 5,000), only 3 out of 5 types can be detected above 96 % (HTTP, MD5, JA3). Meanwhile, IPRep and DNS were only detected in less than 43 %.

It should be noted that the specifications of the Suricata IDS device are also an essential point in detecting anomalies. The specifications used in this research are device specifications in the medium or silver range. The selection of these specifications is based on the types of devices that are often used and or are still affordable to be used by organizations, so the results of this study are generally expected to represent the results of IDS Suricata performance tests.

Research conducted by Red Piranha shows the ability of Suricata lossless detection on network capacity (~0% packet drop) at 60Gbps [25]. The Suricata device specifications has 72 threads and 128 GB RAM, significantly higher than the device in this research (16 threads and 16 GB RAM). Red Piranha and other research conducted by [1, 10, 12-18], focuses on analyzing detection performance on high-speed networks (the parameter is packet drop percentage). In contrast, this research focuses on analyzing the performance in detection accuracy based on the number of active rules (parameter is detection percentage). The interesting thing to investigate further is whether the results of IDS performance analysis with a packet drop percentage value close to 0 % will automatically have a 100% detection accuracy value when the number of active rules reaches one million or more.

However, the risk of performance degradation can be reduced by providing high IDS device specifications, as was done by Purzynski et al. [22] and Jakimoski et al. [25]. But, using high specifications will indirectly pose a risk from a budget perspective which will automatically affect overall operations, bearing in mind that cyber security cannot be left only to IDS devices.

In genuine cases of implementing IDS, scenario 5 is an ideal condition often implemented in anomaly detection. Suricata that has been installed will be tuned, applying the default rules from Suricata (community edition or pro). Also, if the team from security engineering or cyber threat intelligence finds information about a 0-day attack or new malware, it will add local rules that come from extraction IoC. With the findings from this research, the security team, apart from having to prove whether a 0-day attack or new malware anomalies had entered the system or not, also had to double-check whether the IDS failed to detect due to decreased performance (even though the rules had been defined).

5. CONCLUSION

This research findings indicate that the accuracy detection performance of IDS Suricata would decline as the number of activated rules increases. This conclusion is reinforced based on the test result data, where there is an indication of a decrease in detection accuracy as the number of active rules increases. In other words, the application and variation of many rules criteria may result in detection failure. The risk of cyberattacks on owned systems is not limited to those undetected due to unknown information (zero-day attacks), but also includes those previously known and/or defined in the rules but were not identified due to performance issues. This implies that there is an increased risk of cyber attacks.

The results of this research will encourage and strengthen the argument that a defence mechanism is needed to continuously check to ensure that incidents do not occur in the system, which is currently a trend called cyber threat hunting.

6. FUTURE RESEARCH

The next research could include assessing performance on devices with greater spesification, surh as threads and RAM capacity. In addition, with the release of Snort3 in 2021 as the successor to Snort2, it will be possible to compare the performance of rules detection between Suricata and Snort3, both of which have multi-threading capabilities, with an emphasis on comparisons of detection accuracy based on the number of rules.

REFERENCES

[1] N. Iyengar, "Evaluation of Network Based IDS and Deployment of multi-sensor IDS," *arXiv preprint arXiv:2007.11654*, 2020.

[2] P. Alekar, "Survey on Intrusion Detection System (IDS)," *International Journal of Technology Research and Management*, vol. 5, no. 7, pp. 1-5, 2018.

[3] H. Hindy *et al.*, "A taxonomy and survey of intrusion detection system design techniques, network threats and datasets," Faculty of Engineering, Electronic and Electrical Engineering, University of Strathclyde Institutional, 2018. [Online]. Available: https://strathprints.strath.ac.uk/id/eprint/64653

[4] K. Sengaphay, S. Saiyod, and N. Benjamas, "Creating snort-IDS rules for detection behavior using multi-sensors in private cloud," in *Information Science and Applications (ICISA) 2016*: Springer, 2016, pp. 589-601.

[5] H. Hindy *et al.*, "A taxonomy and survey of intrusion detection system design techniques, network threats and datasets," 2018.

[6] M. Bertovič, "Utilization of Threat Intelligence in Information Security," Computing and Information Center, Czech Technical University in Prague., 2017.

[7] H. Almohannadi, I. Awan, J. Al Hamar, A. Cullen, J. P. Disso, and L. Armitage, "Cyber threat intelligence from honeypot data using elasticsearch," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, 2018: IEEE, pp. 900-906.

[8] C. Pace, "The threat intelligence handbook: A practical guide for security teams to unlocking the power of intelligence," *Annapolis, CyberEdge Group*, 2018.

[9] J. Althouse. "TLS Fingerprinting with JA3 and JA3S - Salesforce Engineering Blog." https://engineering.salesforce.com/tls-fingerprinting-with-ja3-and-ja3s-247362855967/ (accessed 11 Jan, 2023).

[10] Q. Hu, S.-Y. Yu, and M. R. Asghar, "Analysing performance issues of open-source intrusion detection systems in high-speed networks," *Journal of Information Security and Applications*, vol. 51, p. 102426, 2020.

[11] W. Park and S. Ahn, "Performance Comparison and Detection Analysis in Snort and Suricata Environment," *Wireless Personal Communications*, vol. 94, no. 2, 2017.

[12] K. Wong, C. Dillabaugh, N. Seddigh, and B. Nandy, "Enhancing Suricata intrusion detection system for cyber security in SCADA networks," in *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017: IEEE, pp. 1-5.

[13] A. Waleed, A. F. Jamali, and A. Masood, "Which open-source IDS? Snort, Suricata or Zeek," *Computer Networks*, vol. 213, p. 109116, 2022.

[14] B. R. Murphy, "Comparing the performance of intrusion detection systems: Snort and Suricata," Colorado Technical University, 2019.

[15] B. Brumen and J. Legvart, "Performance analysis of two open source intrusion detection systems," in *2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2016: IEEE, pp. 1387-1392.

[16] T. Ernawati, M. F. Fachrozi, and D. D. Syaputri, "Analysis of Intrusion Detection System Performance for the Port Scan Attack Detector, Portsentry, and Suricata," (in English), *IOP Conference Series. Materials Science and Engineering*, vol. 662, no. 5, 2019, doi: https://doi.org/10.1088/1757-899X/662/5/052013.

[17] C. Hoover, "Comparative Study of Snort 3 and Suricata Intrusion Detection Systems," Bachelor of Science, Computer Science and Computer Engineering, University of Arkansas, 2022. [Online]. Available: https://scholarworks.uark.edu/csceuht/105

[18] T. Lukaseder, J. Fiedler, and F. Kargl, "Performance evaluation in high-speed networks by the example of intrusion detection," *arXiv preprint arXiv:1805.11407*, 2018.

[19] J. Guo, H. Guo, and Z. Zhang, "Research on High Performance Intrusion Prevention System Based on Suricata," *Highlights in Science, Engineering and Technology*, vol. 7, pp. 238-245, 2022.

[20] D. H. K. Raharjo, A. Nurmala, R. D. Pambudi, and R. F. Sari, "Performance Evaluation of Intrusion Detection System Performance for Traffic Anomaly Detection Based on Active IP Reputation Rules," in *2022 3rd International Conference on Electrical Engineering and Informatics (ICon EEI)*, 2022: IEEE, pp. 75-79.

[21] R. Rohith, M. Moharir, and G. Shobha, "SCAPY-A powerful interactive packet manipulation program," in *2018 international conference on networking, embedded and wireless systems (ICNEWS)*,

2018: IEEE, pp. 1-5.

[22] M. Purzynski and P. Manev, "Suricata Extreme Performance Tuning," presented at the Suricon 2016, 2016. [Online]. Available: https://suricon.net/suricon-2016-washington-dc/.

[23] F. Klassen. "TCPReplay Sample Captures." https://tcpreplay.appneta.com/wiki/captures.html (accessed December, 2022).

[24] P. Inc. "Proofpoint Emerging Threats Rules." https://rules.emergingthreats.net/ (accessed December, 2022).

[25] K. Jakimoski and N. V. Singhai, "Improvement of hardware firewall's data rates by optimizing suricata performances," in *2019 27th Telecommunications Forum (TELFOR)*, 2019: IEEE, pp. 1-4.