SRC - Network monitoring

Universidade de Aveiro

Bruno Lemos, Tiago Marques



SRC - Network monitoring

Departamento de eletrónica, telecomunicações e Informática

Universidade de Aveiro

Bruno Lemos, Tiago Marques (98221) blemos@ua.pt, (98459) tamarques@ua.pt

Contents

1	Nor	n-anomalous behavior analysis and description	1									
	1.1	Statistical Analysis	1									
		1.1.1 Flows Analysis	1									
	1.2	Up/Down Bytes Analysis	2									
	1.3	Country Statistical Analysis	4									
		1.3.1 Flows Analysis	4									
		1.3.2 Up/Down Bytes Analysis	5									
2	Anomalous behavior detection, description, and possible causes											
	2.1	Botnet Detection through Traffic Analysis	7									
		2.1.1 Methodology and Results	7									
		2.1.2 Conclusion	9									
	2.2	Attack on Servers	9									
		2.2.1 Methodology and Results	10									
		2.2.2 Conclusion	12									
	2.3	Exfiltration	12									
		2.3.1 Methodology	12									
		2.3.2 Conclusion	15									
	2.4	Country Statistics	15									
		2.4.1 Methodology	16									
3	SIE	M Rules	20									
	3.1	Introduction	20									
	3.2	Increased Upload Volumes	20									
	3.3	Anomalous Communication with Countries	21									
		3.3.1 Block for Unusual Country Communications	21									
		3.3.2 Alert for Abnormally High Communication Volume	21									
	3.4	Internal Communication within the Network	22									
	3.5	Verification of Distribution Protocols	23									
	3.6	Increasing Server Requests	23									
	3 7	Conclusion	24									

List of Figures

1.1	Average flows to the internet and within the network per user	1
1.2	% of Tcp and Udp flows	2
1.3	Up/Down Bytes per Protocol Flow	3
1.4	Up/Down Bytes per Protocol Flow	4
1.5	Up Bytes per Flow to Countries	5
1.6	Down Bytes per Flow to Countries	6
2.1	Number of flows per server in normal and attack data	8
2.2	Timeline flow of suspect with IP 192.168.100.114	G
2.3	Comparison of flows to servers	11
2.4	Flow timeline of the device with IP 192.168.100.176	12
2.5	Flow timeline of the device with IP 192.168.100.188	12
2.6	Comparison of flows to servers	13
2.7	Up Bytes to Internet	14
2.8	Up Bytes to Internet	15
2.9	Number of Flows to Countries in 'normal' dataset and anomalus	
	$dataset \dots \dots$	17
2.10	Timeline flow of suspect with IP 192.168.100.114	18
2.11	IPs communicating with Russia and China	19
3.1	SIEM rule - 2x more upload volume table	20
3.2	SIEM rule - 2x more upload volume	21
3.3	SIEM rule - Country	22
3.4	SIEM rule - Internal Communication	23
3.5	SIEM rule - Destribution protocol	23
3.6	SIEM rule - Server Requests	2.4

Chapter 1

Non-anomalous behavior analysis and description

1.1 Statistical Analysis

This method helps in understanding the behavior of network traffic, identifying trends, and detecting anomalies or performance issues.

1.1.1 Flows Analysis

From the provided data, the dataset consists of 70.9% of flows originating from the network and destined for the internet, while 29.1% of flows occur within the internal network.

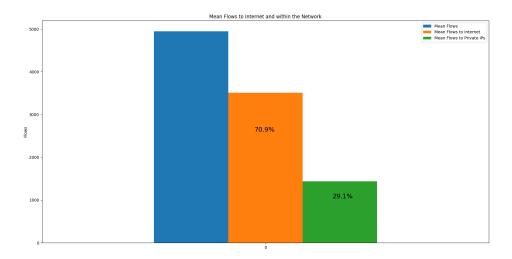


Figure 1.1: Average flows to the internet and within the network per user

Additionally, out of all the flows, 12.08% are classified as UDP (User Datagram Protocol) flows, while the remaining 87.92% are TCP (Transmission Con-

trol Protocol) flows.

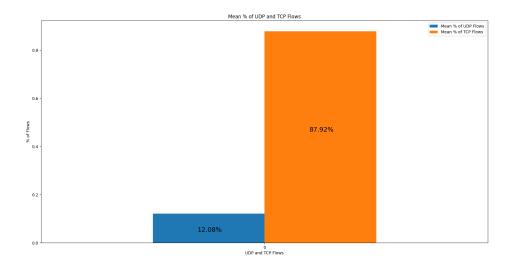


Figure 1.2: % of Tcp and Udp flows

$1.2 \quad Up/Down \ Bytes \ Analysis$

For the transfer of data to and from the network, we looked up for the up/down bytes per flow in the TCP and UDP protocols. We get the following data:

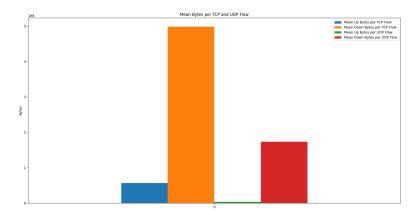


Figure 1.3: Up/Down Bytes per Protocol Flow

Having less upstream (upload) bytes per flow compared to downstream (download) bytes per flow typically indicates an asymmetry in data transfer. When there is a significant difference between upstream and downstream bytes per flow, it suggests there is more data being received by the client or user compared to the data being sent from the client or user. In our case the down bytes per flow are at least 85% larger than the up bytes per flow of the same protocol.

1.3 Country Statistical Analysis

1.3.1 Flows Analysis

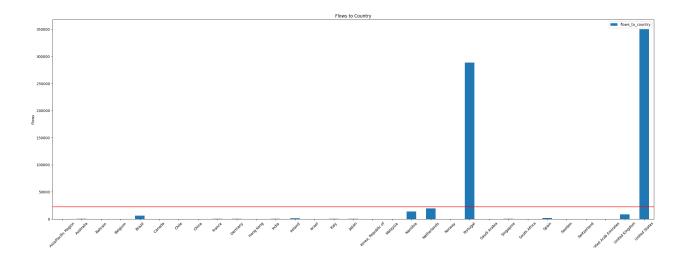


Figure 1.4: Up/Down Bytes per Protocol Flow

From the graph above we can conclude that:

- **High Traffic to USA:** The significant number of flows to the USA, well above the mean, suggests that this country is a popular destination for network traffic from the network. It indicates a substantial volume of communication or data transfer between your network and this country.
- Potential Data Centers: The high flows to the USA may indicate the presence of data centers, hosting providers, or popular online services located in the United States that attract traffic from the network.
- Smaller Flow Counts: The flows to Brazil, Netherlands, Namibia and the UK (between others), though slightly below the mean, indicate a significant level of network traffic to these countries. It could indicate business partnerships or user/employee activity in these regions.
- Regional Service Providers: The high flows to Portugal may be attributed to the utilization of regional service providers, cloud services, or content delivery networks located in Portugal. It indicates that the company relies on local infrastructure or services to support its network operations.

1.3.2 Up/Down Bytes Analysis

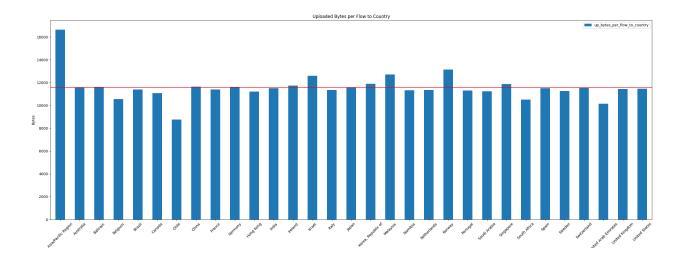


Figure 1.5: Up Bytes per Flow to Countries

From the graph above we can conclude that:

- Unique Upload Patterns: The Asia/Pacific region, compared to other countries, exhibits a noticeable deviation in terms of up bytes per flow. This indicates that there may be specific factors influencing upload activity in that region that differ from the rest of the countries in the dataset. Another possibility for this occurrence is because the Asia/Pacific is a region in itself and not a country.
- Consistent Upload Activity: The similarity in up bytes per flow among the majority of countries indicates a consistent level of upload activity from your network to these destinations. It suggests that data transfer or communication patterns are relatively balanced across these regions.

The same can be said about the down bytes per flow to Countries

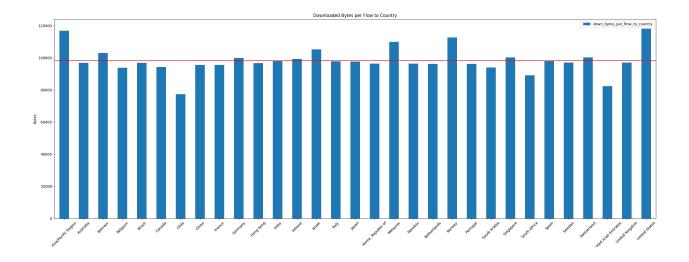


Figure 1.6: Down Bytes per Flow to Countries

Chapter 2

Anomalous behavior detection, description, and possible causes

2.1 Botnet Detection through Traffic Analysis

In the context of our cybersecurity analysis, one of the significant threats we aim to identify is the presence of botnets within the network. Botnets are networks of compromised devices, typically used for malicious purposes such as launching DDoS attacks, spreading malware, or conducting data exfiltration. To detect possible botnet activities, we analyze the traffic data and compare the communication patterns of normal and potentially compromised devices.

2.1.1 Methodology and Results

Identifying Servers with High Traffic

To identify potential botnet activities, we focus on analyzing flows between devices in the network. We get the following data:

- 1. Identifying Servers in Normal Data: We analyze the data_normal dataset and identify the devices that receive traffic from the devices inside the network. The flows are filtered based on both the source (src_ip) and destination (dst_ip) addresses, specifically looking for private IP addresses within the range of 192.168.100.* as destination addresses. The number of flows to each identified server is calculated and stored in the servers_normal dataframe.
- Identifying Servers in Attack Data: We analyze the data_attack dataset and filter the devices on the same criteria, which is, both the source (src_ip) and destination (dst_ip) addresses are private and belong to the network 192.168.100.0/24.
- 3. Comparing Normal and Attack Data: The servers_normal and servers_attack dataframes are merged based on the destination IP address (dst_ip). This allows us to see the differences between normal and attack data to each server. The resulting dataframe, servers_attack, contains the number of flows for each server in both normal (number of flows in normal data) and attack (number of flows in attack data) data.

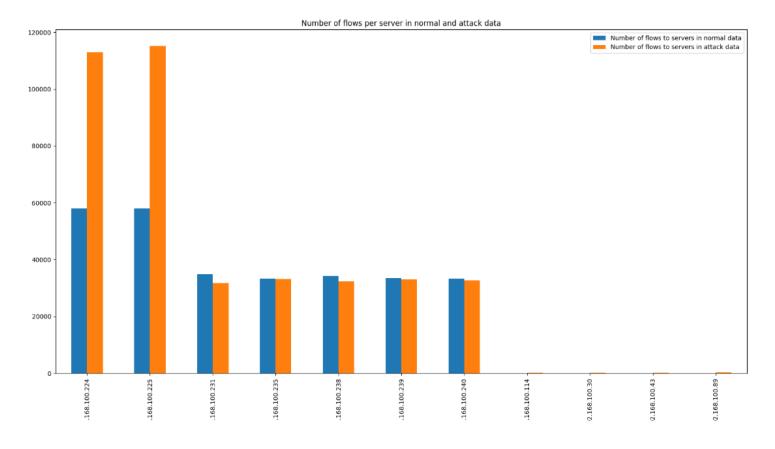


Figure 2.1: Number of flows per server in normal and attack data

Identifying Potential Botnet Suspects

When observing the figure above, we find 4 machines that in the normal dataset did not have any type of communication between private IPs. These machines have the IPs 192.168.100.(114/30/43/89). We look in more detail at the order of flows/timeline of the machines in question and notice that they all have a similar behavior.

The graph below demonstrates the behavior of the machine with the IP 192.168.100.114 with the other machines of suspicious IPs.



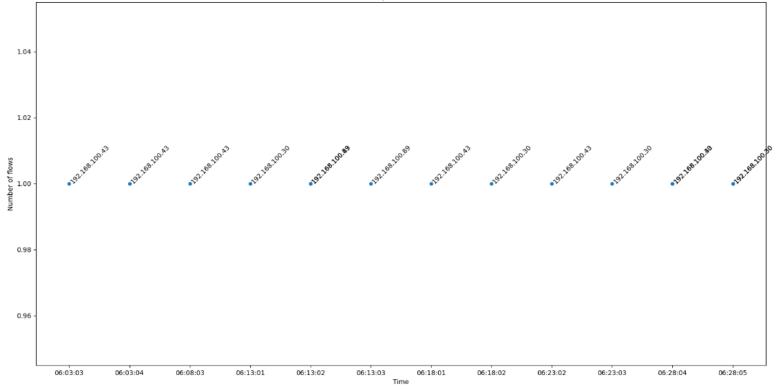


Figure 2.2: Timeline flow of suspect with IP 192.168.100.114

2.1.2 Conclusion

When looking at the graph of figure 2.2 we can see that this machine has periodic flows every 5 minutes with the remaining machines with suspicious IPs. This temporal precision and the longevity of this behavior (starting at 06:03:03 and ending at 15:40:37) exposes this machine a bot on the network. And as the other devices have the same very similar behavior they can also be considered bots.

2.2 Attack on Servers

Based on Figure 2.1, there is a large increase in the number of flows for servers with IPs 192.168.100.(224/225). With this in mind we look in more detail at the machines within the network that have flows with the servers.

2.2.1 Methodology and Results

- Identifying Machines Communicating with Servers: Initially we extract from the 'normal' dataset the ips that communicate with the servers in the network and the number of flows that each one of them has with each server.
- Attack data analysis: Similarly, we analyze the network flows in the attack data. We filter the flows to only include those directed towards the identified servers. Again, we count the number of flows from each source IP address to identify any deviations from the normal behavior.
- Comparison and detection: We compare the flow counts from the normal data analysis and attack data analysis. By calculating the percentage rise in flow counts for each source IP, we can identify significant increases in traffic to the servers. We focus on source IPs with a rise in flow counts exceeding a certain threshold, such as 200%. Additionally, we ensure that these source IPs have a minimum count of flows to the servers in the attack data.

By applying these criteria, we obtain the following graph and we can see that there is a large increase in flows to servers on machines with IPs 192.168.100.(176/188).

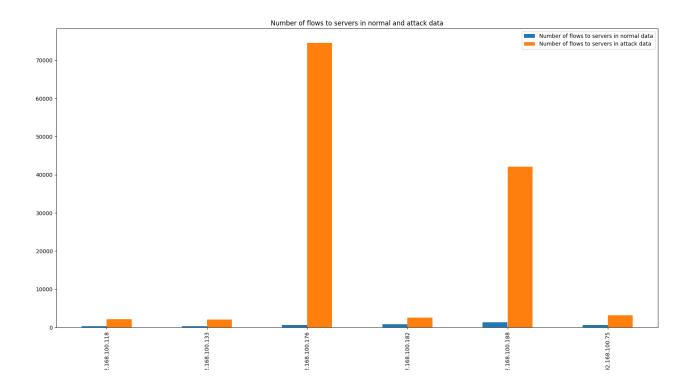


Figure 2.3: Comparison of flows to servers

• **Detailed analysis:** From the identified suspicious source IPs, we extract all their traffic flows in the attack data. This allows us to perform a more detailed analysis of the network activity associated with these IPs. From these suspicius IPs we got the following timeline:

	timestamp	src_ip	dst_ip	proto	port	up_bytes	down_bytes	dst_country
0								
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								United States
24								
25								
26								
27								
28								
29	07:54:43	192.168.100.176	192.168.100.225	udp	53	160	349	

	timestamp		dst_ip	proto		up_bytes	down_bytes	dst_country
0								
- 1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29	10:16:05	192.168.100.188	192.168.100.235	tcp	443	18740	167682	

Figure 2.4: Flow timeline of the device with IP 192.168.100.176

Figure 2.5: Flow timeline of the device with IP 192.168.100.188

2.2.2 Conclusion

Looking at the figures above we can see that there is a large number of flows to the network servers (192.168.100.224 and 192.168.100.225). Followed by flows to public ips where it looks like there is a small download of a small file or data. This behavior profile indicates the establishment of a C&C infrastructure where the compromised devices receive instructions from the attackers and execute them. Therefore, we can assume that these machines are compromised.

2.3 Exfiltration

In this chapter, we investigate potential exfiltration activities. These attacks typically involve infiltrating the target's defenses, gaining access to valuable information, and transferring it to an external location controlled by the attacker. Due to the uploading of data that this type of attack performs, we can detect it through a large increase in the number of bytes uploaded to the internet.

2.3.1 Methodology

Analysis on Normal and Attack Data

Just looking at the total number of uploaded bytes is not enough, as a machine can have a large increase in the number of up bytes from one dataset to the next, but this could have been a consequence of that machine being more active on the network.

With this in mind, for each dataset, we calculate the total number of bytes sent to the internet and the number of flows to the internet for each machine. Then the number of uploaded bytes per flow is calculated, this approach provides a per-machine perspective and can offer insights into individual host behavior within the network.

Comparison of Upstream Flow Statistics

We filter for large increases in the number of up bytes per flow from the 'normal' dataset to the anomalous dataset and focus on machines with a twofold increase. Machines with IPs that fit this criteria are considered suspicious.

By applying the criterion, the following graph is obtained:

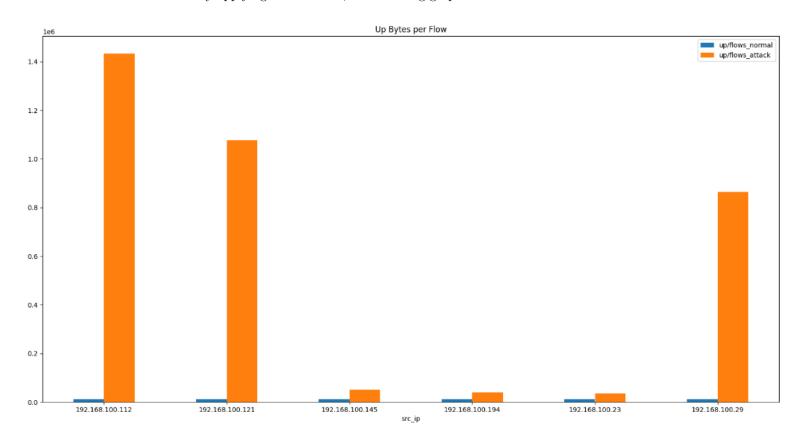


Figure 2.6: Comparison of flows to servers

Looking at the image, 3 machines immediately stand out, which have the IPs 192.168.100.(112/121/29), due to the massive increase in the number of up bytes per flow from one dataset to the other.

Up Bytes to the Internet

We proceeded to look for how many up bytes these machines sent to the internet in the anomalous dataset and found that 3 machines that stand out in the previous figure sent more than 2 Gb each to the internet.

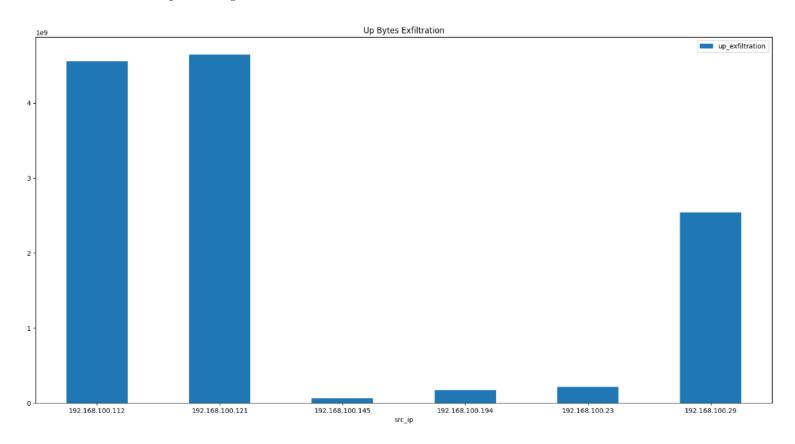


Figure 2.7: Up Bytes to Internet

Focusing on the 3 machines that have large increases in the 2 previous graphs, we can say with some certainty that these machines are uploading data/files to the internet.

To where?

	src_ip	dst_ip	counts	up_bytes	org	dst_country
32	192.168.100.112	142.250.184.184	27	4520510113	AS15169 GOOGLE	United States
182	192.168.100.121	13.107.42.39	30	4598712591	AS8068 MICROSOFT-CORP-MSN-AS-BLOCK	United States
371	192.168.100.29	13.107.42.27	18	2508936132	AS8068 MICROSOFT-CORP-MSN-AS-BLOCK	United States

Figure 2.8: Up Bytes to Internet

We researched where the data from the 3 machines suspected of uploading data/files in the anolamo dataset were being sent. We noticed that, in all the machines, a large part, almost all, of the bytes to be sent to the internet were being sent to a specific IP.

2.3.2 Conclusion

The presence of three machines in the network exhibiting the following behaviors strongly suggests an exfiltration attack:

- Rise in uploaded bytes per flow: Each of the three machines is experiencing a significant increase in uploaded bytes per flow. This indicates a sudden surge in data transfer from these machines.
- More than 2GB uploaded: Each machine has uploaded over 2GB of data, which is a substantial amount of information being transferred.
- Specific IP destinations: Although each machine is sending data to a different IP address, the fact that they are individually targeting specific IPs can still be suspicious. Attackers may distribute the exfiltration traffic across multiple IPs to evade detection or to send data to different locations for different purposes.

While the specific IP destinations may differ, the combination of the significant rise in uploaded bytes per flow, the large data transfer volumes, and the focused targeting by each machine can indicate an exfiltration attack.

2.4 Country Statistics

In this chapter, we analyze the network data to derive statistics related to countries. The goal is to understand the distribution of flows, up bytes, and down bytes across different countries and identify any significant differences between normal and attack data.

2.4.1 Methodology

To perform the country statistics analysis, we follow the following approach:

- 1. **Data Preparation**: We filter out flows between private IP addresses to public IP addresses associated to a country and separate the data into two categories: normal and attack data.
- 2. **Data Normal Analysis**: We calculate the number of flows, up bytes, and down bytes for each country using the normal data. This allows us to understand the distribution of network activity to other countries in the absence of any attacks.
- 3. **Data Attack Analysis**: Similarly, we analyze the attack data to calculate the number of flows, up bytes, and down bytes for each country. This helps us identify the impact of attacks on different countries.
- 4. Comparison and Difference Calculation: We merge the normal and attack data analysis results to compare the statistics between the two datasets. By calculating the differences in the percentage of flows, up bytes per flow, and down bytes per flow, we aim to identify countries that exhibit significant changes during attack scenarios.
- 5. Threshold Criteria: We define a set of threshold criterias to select countries that show substantial variations. Specifically, we consider countries that have a 99% or higher increase in the percentage of flows, a 99% or higher increase in up bytes per flow or a 99% or higher increase in down bytes per flow. Additionally, we only consider countries with more than 100 flows to ensure statistical significance.
- 6. **Percentage Rise Calculation**: For the selected countries, we calculate the percentage rise in the percentage of flows, up bytes per flow, and down bytes per flow. This metric helps us understand the relative increase in network activity during attack scenarios.

By applying the methodology described above, we obtain the following graphs:

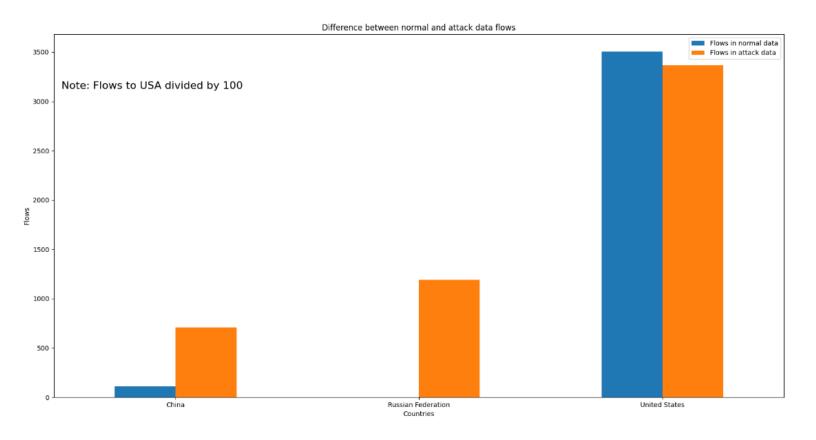


Figure 2.9: Number of Flows to Countries in 'normal' dataset and anomalus dataset $\,$

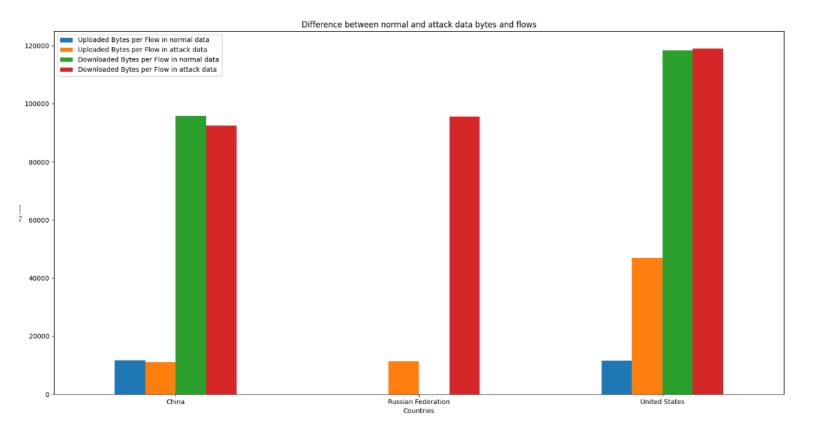


Figure 2.10: Timeline flow of suspect with IP 192.168.100.114

Results and further analysis

The graphs illustrate a significant disparity in data flow between normal conditions and potential attack scenarios. In Russia, there is a substantial increase in the number of flows, from having 0 to getting over 1000 flows. This notable surge in network activity raises suspicions regarding potential security breaches. Similarly, China also experiences a rise in flows, although to a lesser extent. Conversely, the data flow values in the United States remain relatively consistent.

Furthermore, it is crucial to examine the upload and download bytes to further assess the situation. In China, the upload and download bytes align closely with normal data flow patterns, indicating a relatively stable data transfer environment. However, in Russia, a significant and conspicuous disparity is observed. Previously, there were no bytes per flow, but now there is a alot of flows in both upload and download bytes. This abnormal surge in data transfer

activity further heightens concerns about potential malicious activities.

In the United States, a suspicious trend is observed in the upload bytes, which have risen approximately five times more than the normal data flow. Conversely, the number of download bytes remains within the expected range. Such a discrepancy raises suspicions regarding unusual data flow patterns and necessitates thorough investigation.

The increase in up bytes per stream to the US is explained by the previously detected exfiltration attack where the attacker sent the illegally obtained data to machines in the United States.

Looking in more detail at the private IPs on our network that communicate with Russia, we also find that they communicate in greater volume with China, which may explain the increase in the number of flows to this country. Therefore we can consider the machines with the IPs 192.168.100.155 and 192.168.100.77 as malicious.

	dst_country	avg_flows_country	src_ip	flows_country
537	China	3.565657	192.168.100.155	265
555	China	3.565657	192.168.100.77	335
1741	Russian Federation	6.015152	192.168.100.155	555
1742	Russian Federation	6.015152	192.168.100.77	636

Figure 2.11: IPs communicating with Russia and China

Chapter 3

SIEM Rules

3.1 Introduction

Security Information and Event Management (SIEM) systems play a crucial role in monitoring and protecting networks against potential threats. SIEM rules are essential components of this system, as they help identify and respond to suspicious activities or events. In this chapter, we will discuss several SIEM rules related to network security and their significance in maintaining a secure network environment.

3.2 Increased Upload Volumes

One of the critical indicators of a potential security breach is a sudden increase in upload volumes. To detect such incidents, SIEM rules can be configured to generate an alert if the upload volume surpasses a certain threshold. In this case, the threshold is set to 2 times the volume of the average user. This rule helps identify abnormal data transfer activities that may indicate unauthorized access or data exfiltration attempts. Additionally, blocking source IPs with more than 1GB upload can prevent unauthorized access and data exfiltration attempts, enhancing overall security.

src_ip	upload_volume
192.168.100.112	4567543333
192.168.100.121	4660201006
192.168.100.194	183953929
192.168.100.23	229869028
192.168.100.29	2549419509
192.168.100.49	122441194
192.168.100.75	108338449

Figure 3.1: SIEM rule - 2x more upload volume table

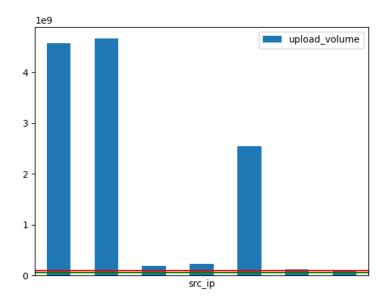


Figure 3.2: SIEM rule - 2x more upload volume

3.3 Anomalous Communication with Countries

3.3.1 Block for Unusual Country Communications

ISEM blocks communication with countries present in the attack data but absent in the normal data and have more than 100 flows to them. This rule potentially stops malicious intent or covert activities from those countries.

3.3.2 Alert for Abnormally High Communication Volume

ISEM generates an alert if the number of flows to a specific country exceeds four times the 'expected' number of flows to any country. This means that a country that normally has 500 flows to it, if in one day it has more than 2000 flows (4x the expected) an alert is issued. This rule helps identify potential threats or unauthorized data transfers.

```
src ips that communicate with Russian Federation: BLOCKED
        192.168.100.155
        192.168.100.77
src ips that communicate with a country that is not in normal data: ALERT
        192.168.100.155
        192.168.100.77
        192.168.100.38
        192.168.100.47
src ips that communicate with a country but have 4x more flows than the average to that country and flows contry > 100:
            dst_country
                         avg_flows_country
                                                          flows_country
                                                    src_ip
  83
                Internal
                               1979.898990
                                           192.168.100.176
                                                                   74580
  96
                Internal
                               1979.898990
                                           192.168.100.188
                                                                   42077
  537
                  China
                                  3.565657
                                            192.168.100.155
                                                                     265
  555
                  China
                                  3.565657
                                                                     335
 1741
       Russian Federation
                                  6.015152
                                            192.168.100.155
       Russian Federation
                                  6.015152
                                                                     636
```

Figure 3.3: SIEM rule - Country

3.4 Internal Communication within the Network

In order to ensure a secure internal network environment, ISEM has implemented strict rules regarding internal communication. The general policy is to restrict communication within the internal network to only allow communication between servers or between a end-user device and a server. Any other communication attempts are blocked, and the source IP address is denied access.

This rule helps to minimize the potential for unauthorized access or data breaches within the internal network. By limiting communication to authorized servers, ISEM reduces the risk of internal threats and enhances the overall security posture of the network.

Should any communication attempts be detected from a non-server device within the internal network, ISEM promptly blocks the source IP address, preventing further communication and mitigating any potential risks. This proactive approach helps to maintain the integrity and confidentiality of sensitive data within the internal network.

src_ip	dst_ip	dst_country
192.168.100.114	192.168.100.43	Internal
192.168.100.89	192.168.100.30	Internal
192.168.100.43	192.168.100.30	Internal
192.168.100.30	192.168.100.43	Internal

Figure 3.4: SIEM rule - Internal Communication

3.5 Verification of Distribution Protocols

The verification of distribution protocols entails monitoring and comparing the distribution patterns of network traffic for each user. If there is a change of 40% or more from the established normal distribution, an alert is triggered for the source IP. This process ensures the integrity and security of network traffic by promptly addressing any deviations.



Figure 3.5: SIEM rule - Destribution protocol

3.6 Increasing Server Requests

It is essential to monitor the flow of requests from each source IP address to the servers. By tracking the number of requests sent to the servers by each source IP, we can establish a baseline or average for comparison.

If the number of requests from a specific source IP address to the servers exceeds a certain threshold, set at 5 times the average value, it indicates a significant increase in traffic from that specific source. This sudden surge in requests can overwhelm the server's resources and impact its overall performance.

To address the issue of excessive requests, we block the IP address in question.

```
        src_ip
        flows_to_servers

        83
        192.168.100.176
        74580

        96
        192.168.100.188
        42077
```

Figure 3.6: SIEM rule - Server Requests

3.7 Conclusion

SIEM rules are essential components of a comprehensive network security strategy. By configuring rules that monitor and detect anomalous activities, organizations can proactively respond to potential security threats. The SIEM rules discussed in this chapter provide a starting point for establishing an effective network security monitoring system. However, it's crucial to regularly review and update these rules to adapt to evolving threat landscapes and ensure the ongoing protection of network resources.