SIEM rules to detect anomalous network behaviors

Segurança em Redes de Comunicações

Universidade de Aveiro

Paulo Gil 76361, Diogo Correia 90327

2022/2023

Departamento de Eletrónica, Telecomunicações e Informática

Contents

Introduction

This project aims to conduct an analysis of the network traffic of a company that wants to implement a reliable Cybersecurity system. We were given two files, *data8.parquet* with the data describing the typical behavior in the network, and *test8.parquet* with possible illicit activities within the network. Both files represent one full day of measures.

We began to make queries to those files and create plots to better visualize the different behaviors. Figure 1.1 describes the number of flows over time, and shows the increased number of flows in test8.

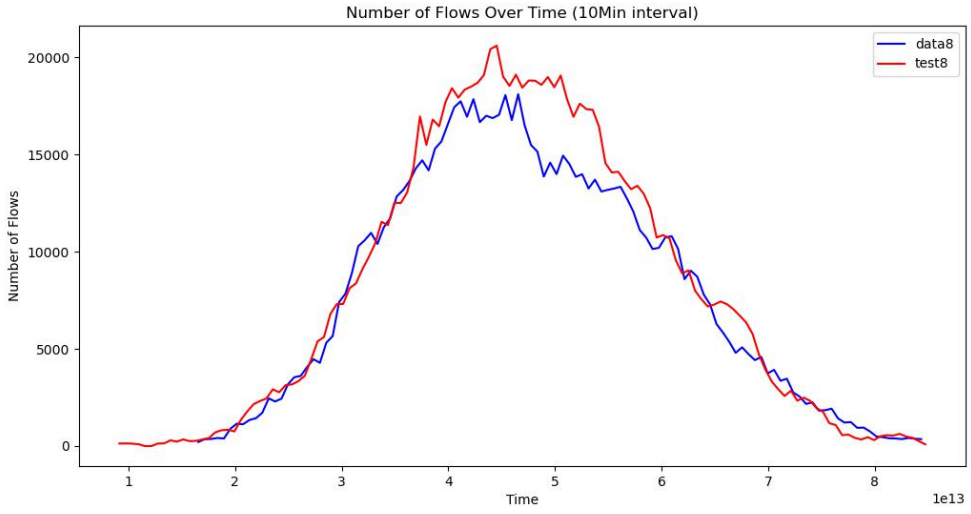

Figure 1.1: Number of Flows Over Time

This discrepancy called for a further analysis, described throughout this report, and finally the definition of some SIEM rules for detection and alert of anomalous behavior.

Analysis on normal behavior

An overall analysis of the normal flow for each country is depicted in Figure 2.1.

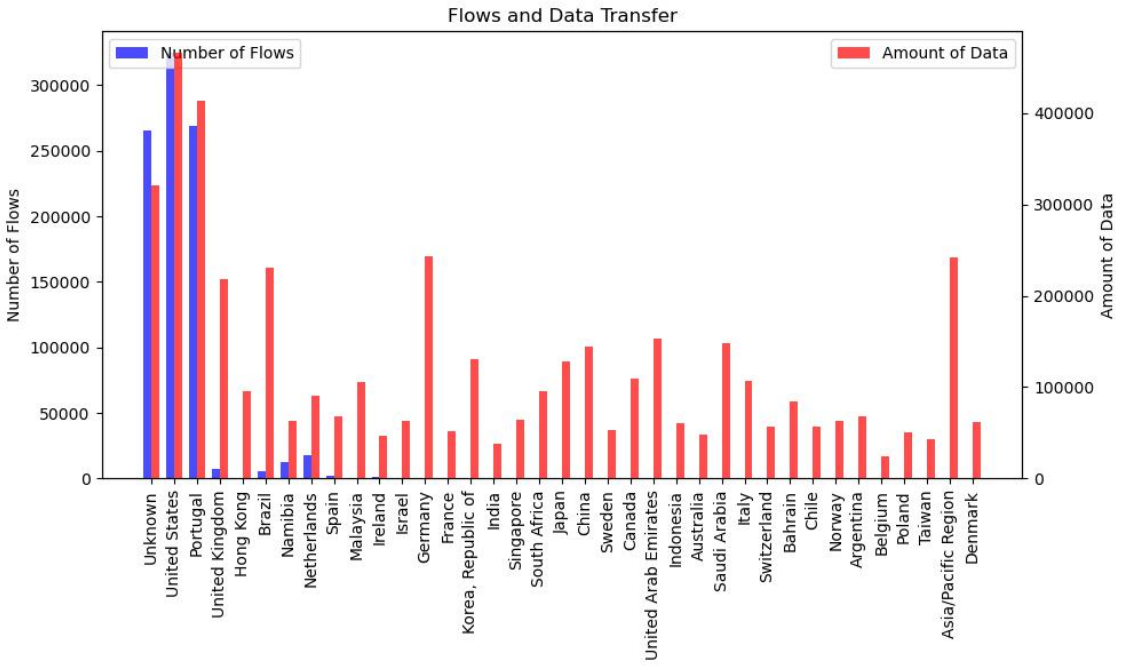

Figure 2.1: Number of Flows and Data Transfered for All Countries in data8

From these flows, the most significant addresses in terms of data byes transfered are depicted in Figure 2.2, first for up_bytes and then for down_bytes.

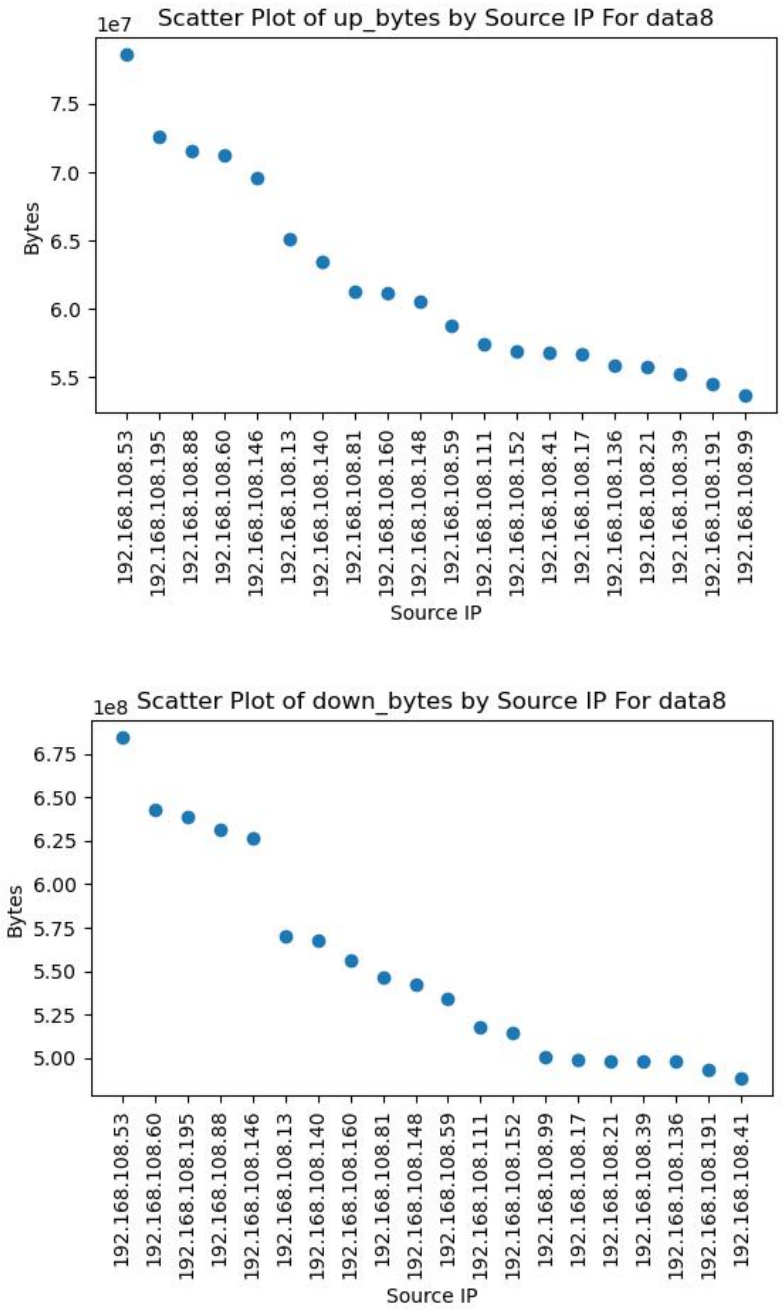

Figure 2.2: Amount of bytes, in total, sent upstream and downstream by IP address

By analysing the flows of the internal machines, we concluded that 192.168.108.226 and 192.168.108.234 are DNS servers because they only receive requests through UDP on port 53 and 192.168.108.240 is either a Web or a Mail server for the same reason as before, but this time using TCP on port 443.

This analysis was made by filtering all communications between private addresses and then counting the number of data flows, taking into consideration the address, the protocol and the port. The result is as seen in Table 2.1

dst_ip	**proto**	**port**	**count**
192.168.108.226	udp	53	197
192.168.108.234	udp	53	196
192.168.108.240	tcp	443	197

Table 2.1: Network Data

Analysis on anomalous behavior

3.1 Suspicious Flows

The data from *test8* gave us an insight on new countries to where data was being sent, apart from the ones registered in *data8*. We grouped the number of flows and the amount of data transfered in total for each new country and we obtained the result depicted in Figure 3.1.

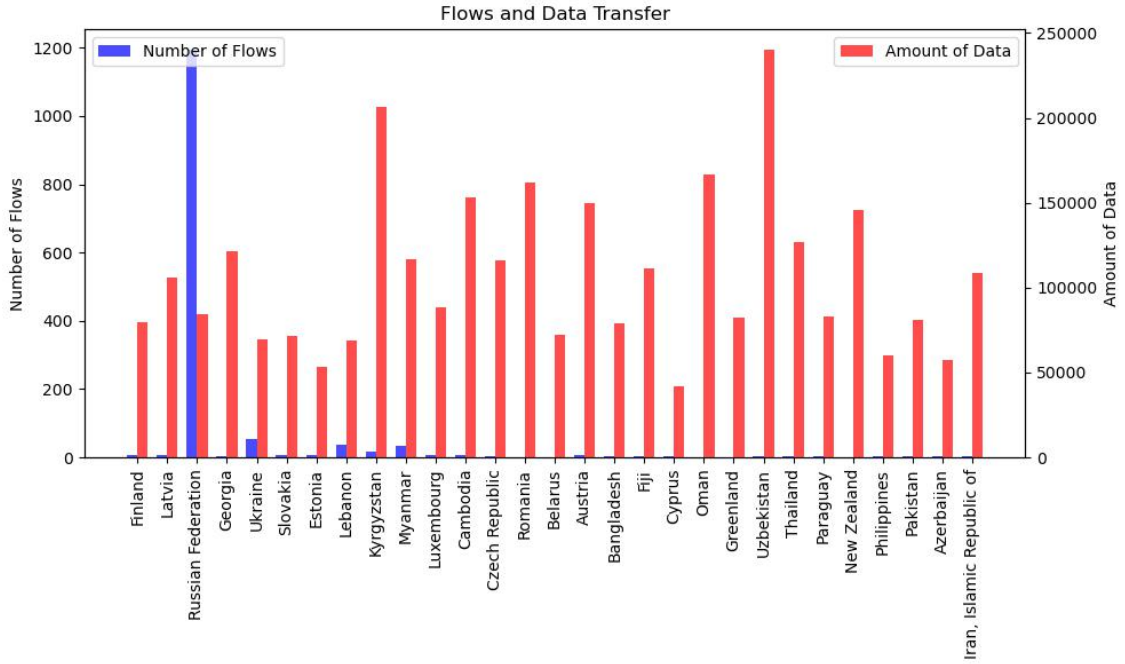

Figure 3.1: Number of Flows and Data Transfered for New Countries

A quick analysis tells us that the country with more flows was, by far, the Russian Federation (1194), but the amount of data transfered wasn't so significant (84 108 bytes). On the other hand, Kyrgyzstan and Uzbekistan, with way fewer flows (17 and 3), transfered more than double the

amount of Russia (206 170 bytes and 240 097 bytes, respectively). The whole data for all the suspicious countries is presented in Table 3.1.

Because of the distinctive behavior portrayed for the Russian Federation, we gathered those communications and concluded that they were all TCP in port 443.

country	**data_flows**	**data_bytes**
Finland	8	79 464
Latvia	6	106 175
Russian Federation	1 194	84 108
Georgia	2	121 333
Ukraine	53	69 493
Slovakia	7	71 732
Estonia	7	52 994
Lebanon	36	69 066
Kyrgyzstan	17	206 170
Myanmar	35	116 810
Luxembourg	6	88 583
Cambodia	7	153 278
Czech Republic	4	116 117
Romania	1	162 252
Belarus	1	72 277
Austria	6	149 597
Bangladesh	3	78 801
Fiji	3	111 133
Cyprus	2	41 648
Oman	1	166 806
Greenland	1	82 619
Uzbekistan	3	240 097
Thailand	3	127 156
Paraguay	4	82 870
New Zealand	1	145 889
Philippines	4	60 182
Pakistan	2	81 157
Azerbaijan	3	57 486
Iran, Islamic Republic of	2	108 822

Table 3.1: New Countries Data

3.2 Exfiltration

When analysing the flow of data upstream, we detected what might have been an issue with data exfiltration. We first calculated the amount of bytes being sent upstream on both contexts and concluded that, in the illicit scenario, the average amount of data being sent upstream with each communication was 20.3k bytes, two times more than the 9.8k bytes being sent in the normal scenario.

This led us filtering all the IP addresses that sent more than the average amount, in total, and we obtained the results in Figure 3.2, for the top 20 addresses.

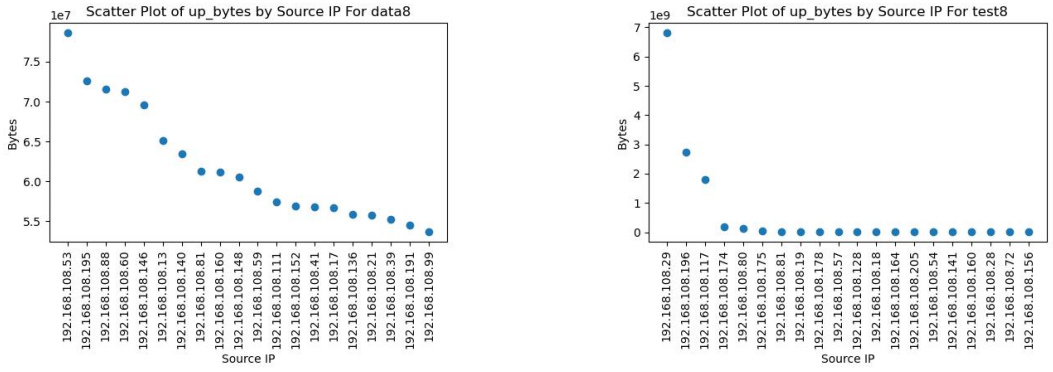

Figure 3.2: Amount of bytes, in total, sent upstream by IP address

It's worth noting that the y axis for *test8* is two orders orders of magnitude larger that in *data8*. Applying the same limit for both plots would render the plot for *data8* unreadable.

With these suspicious addresses in mind, and considering the context they're in, we proceeded to try and understand if these were really cases of exfiltration.

We took the 5 addresses that stood out the most, 192.168.108.29, 192.168.108.196, 192.168.108.117, 192.168.108.174, 192.168.108.80 and, for each one, checked the destination IP address for the largest upstream transactions to understand who the target was, using the service WHOIS.

For the address 192.168.108.29 (Table 3.2), the destination address was 142.250.184.246, which was traced back to Google LLC (GOGL), so it was much likely an exfiltration from their service Google Drive.

For the addresses 192.168.108.196 and 192.168.108.117 (Table 3.3), the destination addresses were, respectively, 13.107.42.52 and 13.107.42.29, which we traced back to Microsoft Corporation (MSFT) so, much likely an exfiltration from OneDrive.

For the addresses 192.168.108.174 and 192.168.108.80 (Table 3.4), the destination addresses were, respectively, 104.244.42.1 and 104.244.42.129, which we traced back to Twitter Inc. (TWITT). Contrary to the exfiltration from Google Drive and OneDrive, this traffic didn't move as much data (as we can also see in Figure 3.2, on the right) but the up_bytes were always significantly greater than the down_bytes.

timestamp	**src_ip**	**dst_ip**	**proto**	**port**	**up_bytes**	**down_bytes**
3277856	192.168.108.29	142.250.184.246	udp	443	151624747	1683607
3397894	192.168.108.29	142.250.184.246	udp	443	108415643	1234857
3517680	192.168.108.29	142.250.184.246	udp	443	217389365	2445431
3637548	192.168.108.29	142.250.184.246	udp	443	197393735	1913871

Table 3.2: Some of the traffic to Google from 192.168.108.29

timestamp	src_ip	dst_ip	proto	port	up_bytes	down_bytes
3767772	192.168.108.196	13.107.42.52	tcp	443	405232314	4711674
3887737	192.168.108.196	13.107.42.52	tcp	443	157392732	2383406
4007676	192.168.108.196	13.107.42.52	tcp	443	240847227	2258830
4127517	192.168.108.196	13.107.42.52	tcp	443	297176841	3203779

Table 3.3: Some of the traffic to Microsoft from 192.168.108.196

timestamp	src_ip	dst_ip	proto	port	up_bytes	down_bytes
4295337	192.168.108.174	104.244.42.1	tcp	443	343713	13784
4307493	192.168.108.174	104.244.42.1	tcp	443	559514	18057
4319630	192.168.108.174	104.244.42.1	tcp	443	323388	8235
4331674	192.168.108.174	104.244.42.1	tcp	443	594208	24473

Table 3.4: Some of the traffic to Twitter from 192.168.108.174

3.3 Botnets

Botnets are characterized by devices that become puppets and perform undesired actions behind the curtains. One of the issues can be privilege escalation within a network or DDoS attacks, and the communication between private addresses is expected.

To detect this, we analysed the communication between local machines on test8, like we did for data8 to identify regular services like DNS and Web servers.

This time, apart from the 3 services, we found 4 other communication flows (Table 3.5) that, because they weren't in the regular communications data file, we can assume that they are not regular services from the corporate but, in fact, in the context of this specific analysis, they could be botnets.

To gain further confidence, we calculated the ratio between upload and download on the flows involving the suspicious addresses, and they all presented a ratio close to 1, as seen in Table 3.6.

Because the communication between the infected devices is usually the sharing of information and instructions or coordinating attacks, the traffic in both directions tend to be balanced, hence the ratio of 1.

dst_ip	proto	port	count
192.168.108.110	tcp	443	3
192.168.108.181	tcp	443	3
192.168.108.226	udp	53	196
192.168.108.234	udp	53	197
192.168.108.240	tcp	443	197
192.168.108.76	tcp	443	3
192.168.108.93	tcp	443	3

Table 3.5: Possible Botnets

src_ip	dst_ip	up_bytes	down_bytes	ratio	diff
192.168.108.76	192.168.108.110	296089	296190	0.999659	0.000341
192.168.108.181	192.168.108.76	126155	125847	1.002447	0.002447
192.168.108.93	192.168.108.181	331529	329218	1.007020	0.007020
192.168.108.76	192.168.108.181	325142	322765	1.007364	0.007364
192.168.108.110	192.168.108.93	272824	270753	1.007649	0.007649
192.168.108.110	192.168.108.181	294838	297449	0.991222	0.008778
192.168.108.93	192.168.108.76	367209	363998	1.008821	0.008821
192.168.108.76	192.168.108.93	302867	299569	1.011009	0.011009
192.168.108.93	192.168.108.110	342834	347594	0.986306	0.013694
192.168.108.181	192.168.108.110	64801	65928	0.982906	0.017094
192.168.108.181	192.168.108.93	88819	86728	1.024110	0.024110
192.168.108.110	192.168.108.76	302124	310490	0.973055	0.026945

Table 3.6: Botnet Data Transfer Statistics

3.4 Command & Control

Command & Control attacks can sometimes be disguised with DNS encapsulation, the so called DNS Tunneling. Knowing the DNS servers of the corporate (Table 2.1), we gathered the flows that targeted their addresses, and retrieved the source addresses with more requests, as depicted in Figure 3.3.

The addresses 192.168.108.27, 192.168.108.31 and 192.168.108.59 clearly stood out from the rest, with an abnormal amount of requests to a DNS server, which right away puts them under suspicion.

Considering the context of the analysis, it is fair to assume that these were target of C&C attacks.

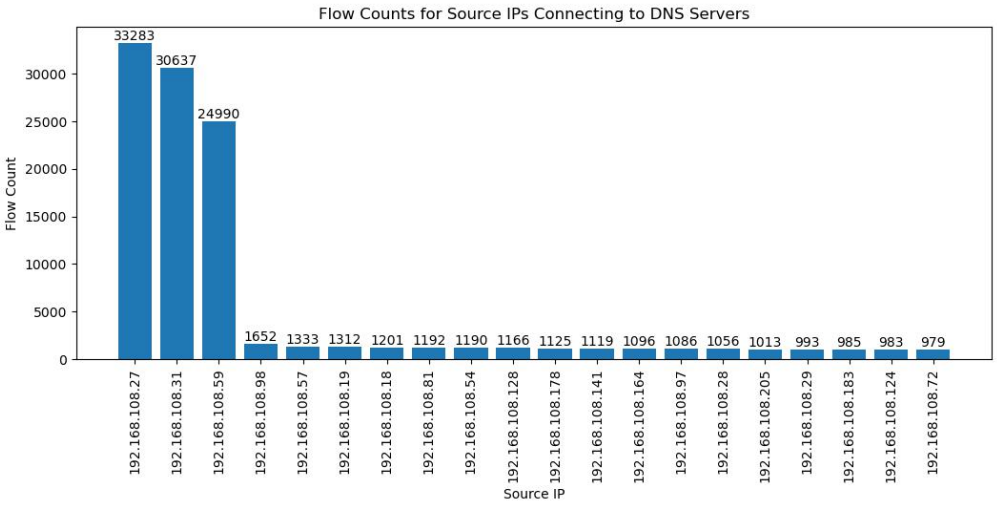

Figure 3.3: Number of Flows connecting to DNS Servers

SIEM rules

4.1 Suspicious DNS Traffic

This rule targets UDP DNS traffic, which is commonly used for exfiltration or data exfiltration attempts. By identifying cases where the amount of up_bytes is unusually high for DNS requests, you can potentially detect suspicious activities.

4.1.1 Conditions

- Protocol is UDP
- Destination port is 53
- Upstream bytes is significantly higher than the average for DNS traffic

4.1.2 Actions

- Generate an alert for potential exfiltration activity

4.2 Unusual Traffic on Port 443

This rule focuses on traffic on port 443, which is commonly used for secure web communication. By monitoring the amount of outgoing traffic (down_bytes) on this port, it is possible identify cases where the volume exceeds normal patterns, potentially indicating exfiltration attempts.

4.2.1 Conditions

- Protocol is TCP or UDP
- Destination port is 443 (HTTPS)
- Downstream bytes (outgoing traffic) is significantly higher than the average for port 443

4.2.2 Actions

- Generate an alert for potential exfiltration activity

4.3 Suspicious Twitter Exfiltration

It is possible to identify potential anomalies indicating Twitter exfiltration attempts. A significantly higher number of flows to a Twitter IP address compared to other destinations suggests abnormal behavior.

4.3.1 Conditions

- Protocol is TCP or UDP
- Destination IP is associated with Twitter's IP ranges
- Destination port is 443 (HTTPS)
- Number of flows to the specific Twitter IP address is significantly higher than the average number of flows to other IP addresses

4.3.2 Actions

- Generate an alert for potential Twitter exfiltration activity

4.4 Communication with New Countries

This rule focuses on identifying communications with IP addresses located in countries that have not been previously communicated with. By maintaining a database or list of known IP ranges for each country, it is possible to compare the destination IP address of network traffic with those ranges to determine if it falls within the IP range of a previously uncommunicated country.

4.4.1 Conditions

- Protocol is TCP or UDP
- Destination IP is outside the known IP ranges of previously communicated countries

4.4.2 Actions

- Generate an alert for potential communication with new countries

4.5 Command and Control (C&C) Communication with Abnormal DNS Requests

This rule monitors the number of DNS requests made from your network within a specified time window. A high volume of DNS requests can be an indicator of C&C communication, as botnets and other malicious infrastructures often use DNS as a covert communication channel.

4.5.1 Conditions

- Protocol is TCP or UDP
- Abnormally high number of DNS requests within a specified time window

4.5.2 Actions

- Generate an alert for potential Command and Control (C&C) communication with abnormal DNS requests

4.6 Botnet Detection with Unusual Communications with Internal Machines

This rule detects potential botnet activity by analyzing suspicious network behavior which can be an unusually high number of connections from a single source IP within a time window or unusual traffic patterns, such as sudden surges in volume or packet counts. Botnets exhibit distinct behavior that deviates from normal traffic. By monitoring this behaviours, it is possible to detect botnet activity.

4.6.1 Conditions

- Protocol is TCP or UDP
- High number of connections to multiple internal machines within a specified time window

4.6.2 Actions

- Generate an alert for potential botnet communication with unusual communications with internal machines

Attachments

This report should be accompanied by a folder with all the code used to perform queries and make plots from the data files.