Techniques for Anomaly Detection in IPv4 & IPv6 Network Flows

Bianca I. Colón Rosado
Grace Rodríguez Gómez
Advisor: Dr. Humberto Ortiz-Zuazaga
University of Puerto Rico, Río Piedras Campus

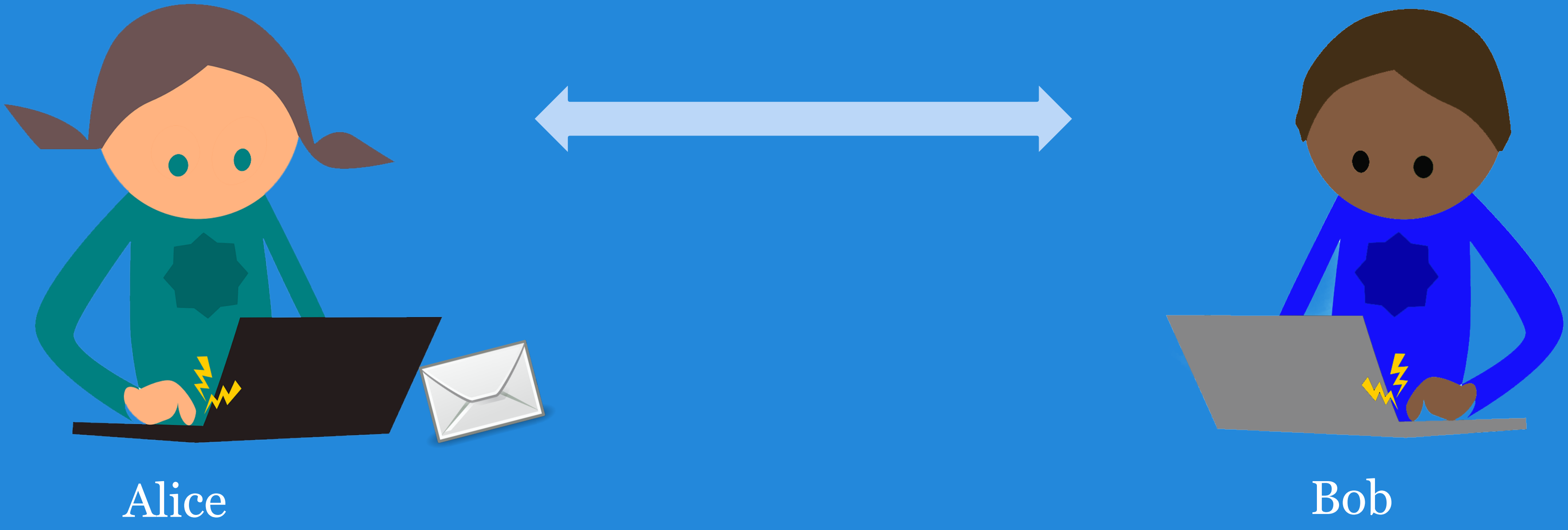
Flows

A summary of a sequence of packets sent from a source computer to a destination, which may be another host, a multicast group, or a broadcast domain.

A flow could consist of source IP, destination IP, source port, destination port, packets, bytes, flags, source time, etc.

Flow example

Source IP Address:	sPort	Packets	bytes	Destination IP Address	dPort
203.13.173.243	53	2	96	128.3.45.10	2124



Analysing Flow Data: Anomaly Detection

- Anomaly detection is a method that searches for unusual and out of the ordinary activity in traffic flow packets. In this research, however, we are classifying a flow anomaly those packets with an inexplicable amount of data (bytes).

  - One way to analyse flow data is with anomaly detection.

IPv4 vs IPv6

- **IPv4**
  - Internet Protocol version 4 (IPv4) is the fourth version in the development of the Internet Protocol (IP) Internet, and routes most traffic on the Internet.
- **IPv4 Address Structure:**
  - IPv4 uses 32-bit (four-byte) addresses, which limits the address space to 4,294,967,296 (2^{32}) addresses. As addresses were assigned to users, the number of unassigned addresses decreased.
- **Example:**
  - 136.145.181.112

IPv4 vs **IPv6**

- **IPv6:**
  - Internet Protocol version 6 (IPv6) is the latest revision of the Internet Protocol (IP) and the first version of the protocol to be widely deployed. IPv6 was developed by the Internet Engineering Task Force (IETF) to deal with the long- anticipated problem of IPv4 address exhaustion.

- **IPv6 Address Structure:**
  - IPv6 uses a 128-bit address, allowing 2^{128}, or approximately 3.4×10^{38} addresses. This 128 bits are divided into eight 16-bits blocks. Each block is then converted into 4-digit Hexadecimal numbers separated by colon symbols.
  - **Example:**

2607:2000:1000:1160:225:90ff:fe8e:b4a0

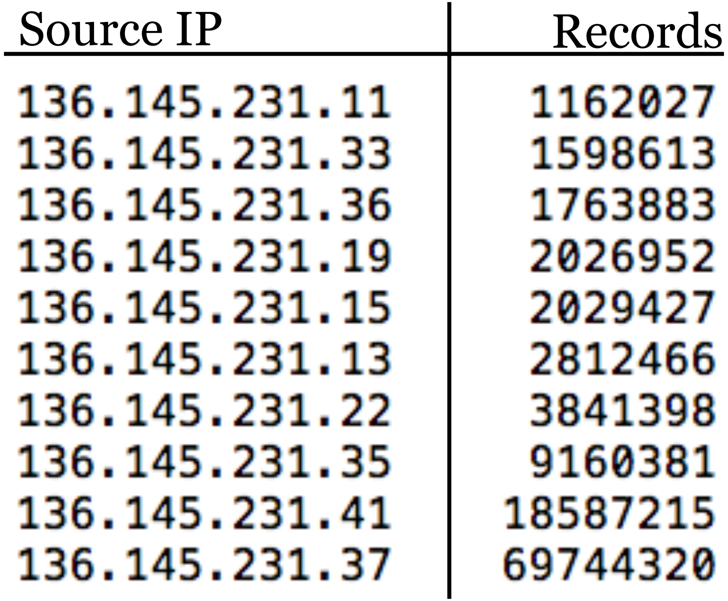
SiLK

System for Internet-Level Knowledge, is a collection of traffic analysis tools developed by the CERT Network Situational Awareness Team (CERT NetSA) to facilitate security analysis of large networks.

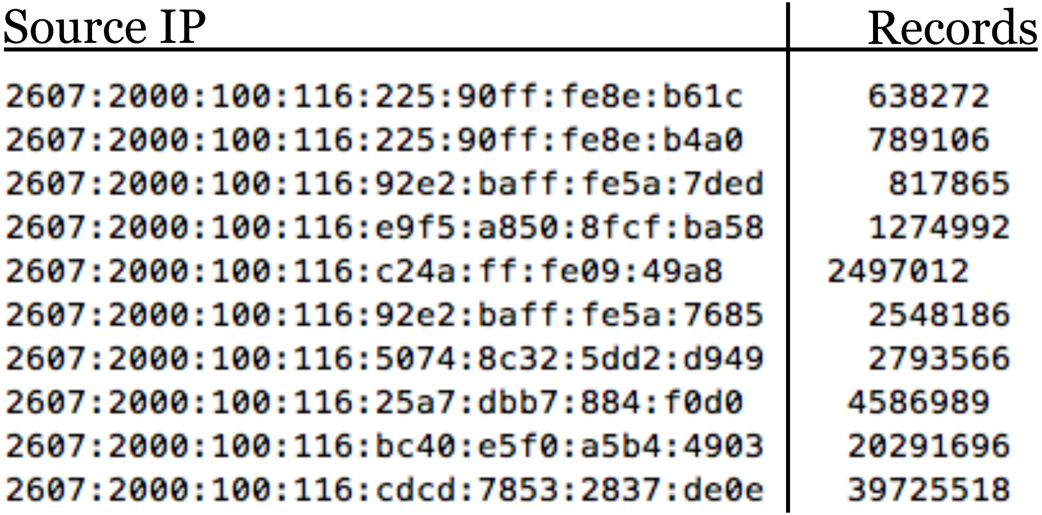
The SiLK Tool Suite supports the efficient collection, storage, and analysis of network flow data, enabling network security analysts to rapidly query large historical traffic data sets.

Top 10 Flows Program

This program returns the IP’s who has more activities in the DMZ network.



Source IP	Records
136.145.231.11	1162027
136.145.231.33	1598613
136.145.231.36	1763883
136.145.231.19	2026952
136.145.231.15	2029427
136.145.231.13	2812466
136.145.231.22	3841398
136.145.231.35	9160381
136.145.231.41	18587215
136.145.231.37	69744320

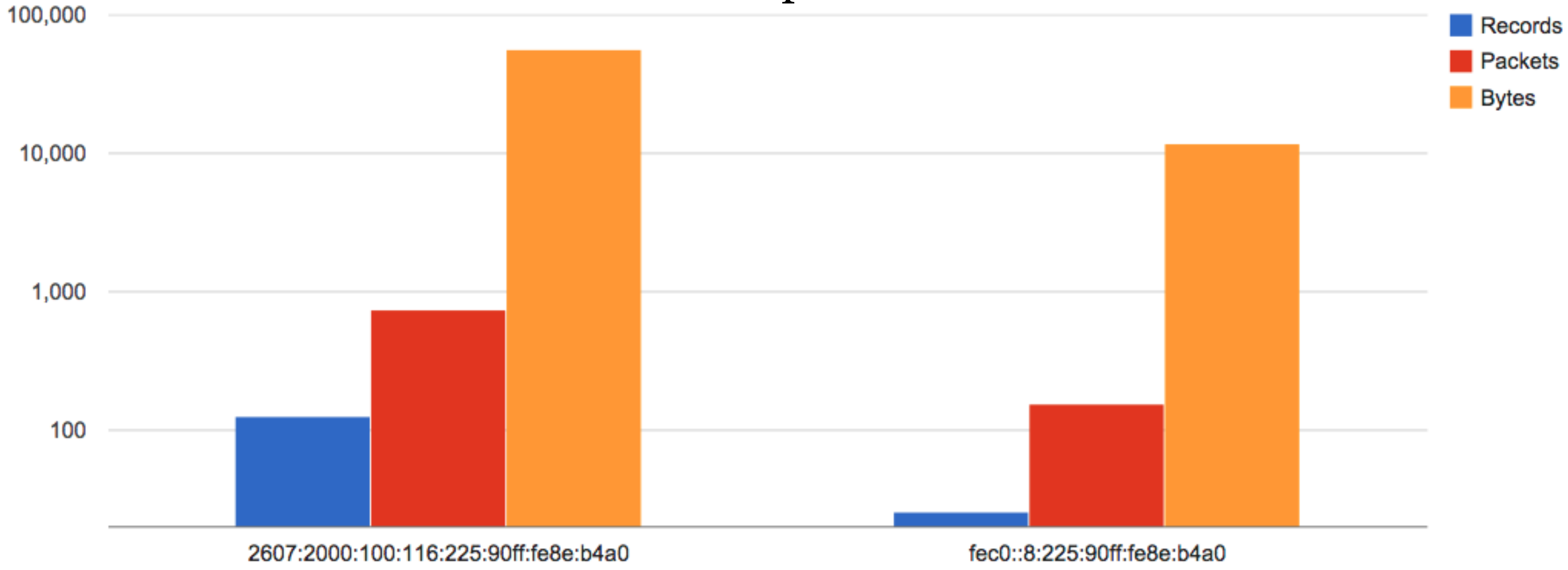


Source IP	Records
2607:2000:100:116:225:90ff:fe8e:b61c	638272
2607:2000:100:116:225:90ff:fe8e:b4a0	789106
2607:2000:100:116:92e2:baff:fe5a:7ded	817865
2607:2000:100:116:e9f5:a850:8fcf:ba58	1274992
2607:2000:100:116:c24a:ff:fe09:49a8	2497012
2607:2000:100:116:92e2:baff:fe5a:7685	2548186
2607:2000:100:116:5074:8c32:5dd2:d949	2793566
2607:2000:100:116:25a7:dbb7:884:f0d0	4586989
2607:2000:100:116:bc40:e5f0:a5b4:4903	20291696
2607:2000:100:116:cdcd:7853:2837:de0e	39725518

FLOWBAT

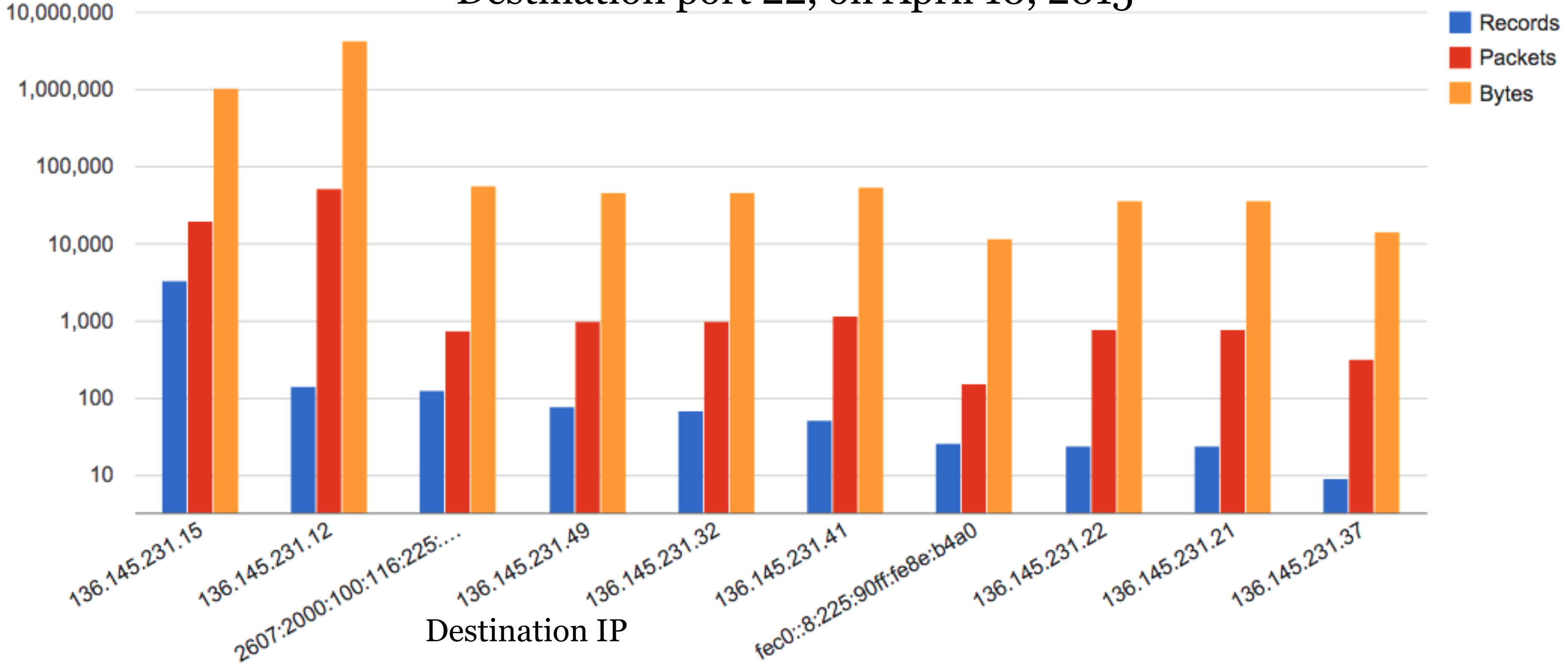
- Flow Basic Analysis Tool
- FlowBAT is a graphical flow-based analysis tool.
- Utilizing the power and versatility of network flow records, FlowBAT can help provide visibility for network administrators and network security practitioners.

IPv6 in port 22



Destination IP

Destination port 22, on April 10, 2015



Converting IPv6 address to numbers in the range of 0 to 1

- To be able to display IPv6 flow data in a 3D graph, we wrote a program in Python that read flow data from a file and convert the addresses into x and y coordinates ranging from 0 to 1.

- The file had all the IPv6 flows that were captured from the network of the University of Puerto Rico, Río Piedras Campus in April 10, 2015.

- The flow data from the file that were used was the destination port as **Z**, IP Source Address as **X**, and IP Destination Address as **Y**.

Converting IPv6 address to numbers in the range of 0 to 1

Here is an example of an IPv6 address:

2001:dc0:2001:0:4608**::**25

Adding zeros in the four colons we get:

2001:dc0:2001:0:4608**:0:0:**25

which is the same as 2001dc02001046080025

This is then multiplied by i where 0 <= i < 8

$$2001 * 216^{7} + dc0 * 216^{6} + 2001 * 216^{5} + 0 * 216^{4}$$

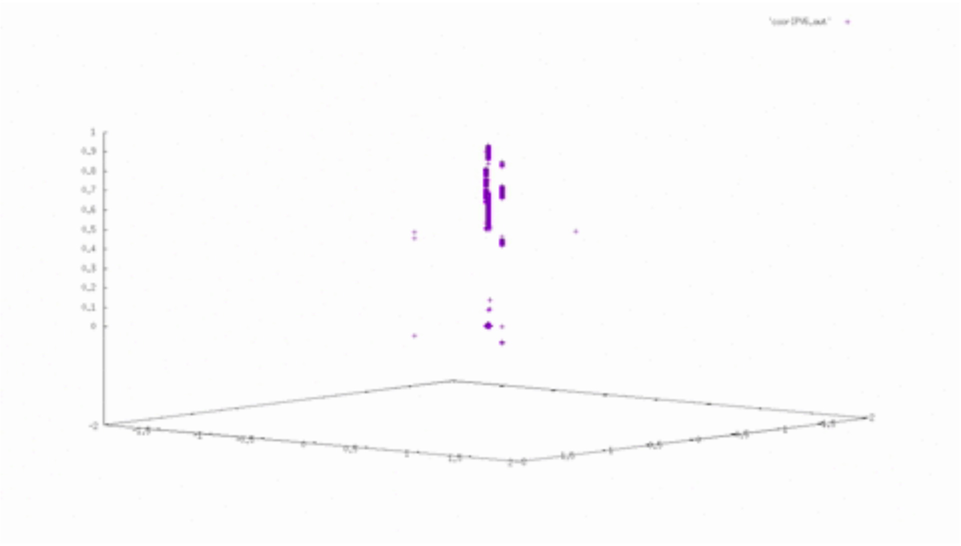
$$+ 4608 * 216^{3} + 0 * 216^{2} + 0 * 216^{1} + 25 * 216^{0}$$

The total of is 4.254076705501262e+37

Which divided by equals **0.1250**

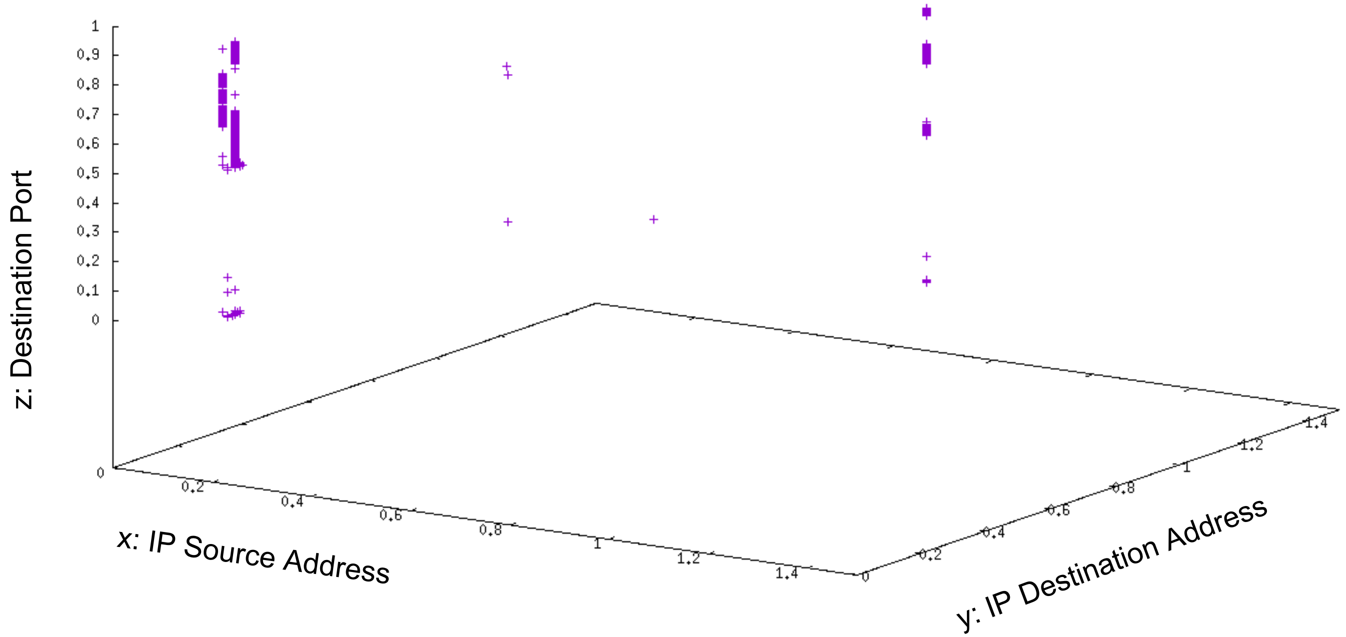
Cube IPv6

IPv6 Flow Data of the UPR-RP from April 10, 2015

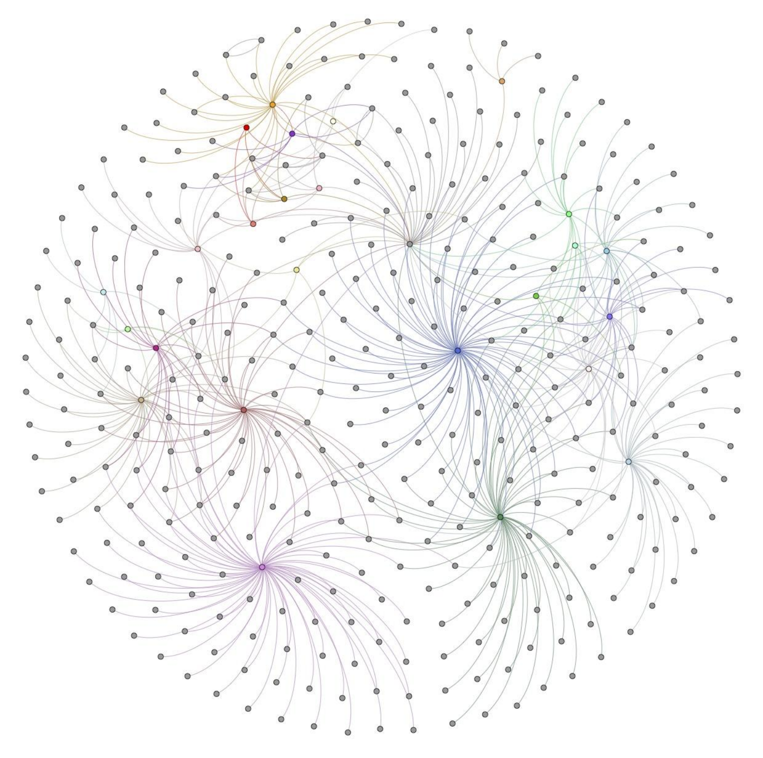


Cube IPv6

IPv6 Flow Data of the UPR-RP from April 10, 2015



Force directed layout of IPv6 connectivity graph



Conclusion

- Using SiLK we can saw IPv4 and IPv6 flows. And we can graph IPv6 addresses
- Its a lot easier to analyse flow data displaying the data with visualisation methods such as graphs.
- After implementing the program with the IPv6 data, we realised there can be a big amount of flows for only one day. Therefor, it would be almost impossible to try and analyse each flow one by one. It is then more practical to write a code that reads the data and organises it depending on its data type.

Future Work

- In the future, we will explore new approaches to find new techniques. Implement these techniques for anomaly detection to our collection of flows from UPR's network, and compare results with the results of current techniques.
    - Implement Benford's law

- Make an implementation that displays the data in real-time in the graph.

Acknowledgments

- This work is supported by the scholarship Academics and Training for the Advancement of Cybersecurity Knowledge in Puerto Rico (ATACK-PR) supported by the National Science Foundation under Grant No. DUE-1438838.

- Our Research Advisor, Dr. Humberto Ortiz-Zuazaga

- Dr. José Ortiz Ubarri

References

Paper: http://ccom.uprrp.edu/~humberto/research/anomaly-detection.pdf

SiLK: https://tools.netsa.cert.org/silk/

Book: http://tools.netsa.cert.org/silk/analysis-handbook.pdf

FlowBAT: http://www.flowbat.com/

Gnuplot: http://www.gnuplot.info/

Gephi: http://gephi.github.io/

TOA: https://github.com/cslab-uprrp/toa

Information: http://tools.ietf.org

http://www.tutorialspoint.com/