DC4: Exploratory Data Analysis
Group: Adriana Beltran Andrade, Catherine Park, Casey Perez, Rachel Yan, Elaine Ye

BankWorld's Bank of Money: Cyber Threat Analysis

Introduction:

Bank of Money (BOM) is an important organization in BankWorld, spanning from coast to coast in the single landmass on the planet. With their great size and importance, any potential issues that may compromise the BOM is a threat to anyone that needs to rely on it. Since the BOM is essential to financing in BankWorld, any cyber threat would hurt all the nations' economies. Therefore, any unusual activity must be looked into immediately and taken seriously. Using information provided from their Firewall and IDS logs, this report aims to narrow down on potential security threats/events and understand the overall trends of the network. We hope to narrow down specific periods of unusual activity and understand what may have caused them as well as considering possible ways to address these issues.

Hypothesis:

With the given context and data, we hypothesize that there are sinister activities taking place within the Bank of Money organization. We suspect that there are certain individual(s) who are imitating or mishandling workstation IP addresses in the BOM in order to gain access to private information that could compromise the entire organization. With everyone in BankWorld relying on the BOM, it is important to maintain a secure network and ensure that sensitive information is not released or used for nefarious purposes. Those seeking to gain unauthorized access to the BOM network may intend to cause economic mayhem across all of the countries or may be looking out for their own financial gain.

Methodology:

While considering the situations and the data given, we initially recognized the extensive web of activity connecting the networks in the BOM. Therefore, we relied heavily on analyzing the connections between the internal IP addresses in order to get a sense of what may be happening in the network. We sought to identify any trends that we could find over the two days in both the Firewall and IDS data as well as any unusual drops/jumps in internet activity. Furthermore, potentially negative classifications and requests were traced to discover specific sets of IP addresses that may not belong in the network or were being mishandled. Suspicious activity that we identified was also further investigated in order to figure out who or what may be behind these events as well as ways to mitigate the issues.

Data Visualizations & Analysis:

Firewall & IDS General Trends:

While trying to understand what occurs on a normal day for the BOM, we sought to graph the general activity over the day. Although we were limited to two specific days, April 6th and 7th, we used this data to get a better understanding of when activity may peak or drop as well as how high the activity level usually stays.

To examine firewall activities, we drew line graphs to show total instances of different firewall syslog priority for both days (Figures 1 & 2). On both days, most of the firewall activities had the syslog priority of "Info," which is about informational messages [3]. The second most common activity had the syslog priority of "critical," which indicates hard device errors [3]. Notably, the activities with "info" had rapid increases and decreases at the beginning and the end of daily records but remain consistent at other time periods. In addition, compared to the record on April 7th that had activities with warnings within an hour, firewall activities on April 6th had warnings and errors for several hours spanning from the evening of the 5th to midnight of the 6th. These unusual trends may require further investigation.

The IDS, or Intrusion Detection System, data helped to identify the network traffic that would occur within that day. Activity was identified specifically as either generic protocol, potential violations/threats, or miscellaneous. As shown in both Figures 3 and 4, Generic Protocol Command Decode occurred in the highest amounts over both days at around 750 to 875 instances in a given 24-hour period, indicating that systems and connections were running under typical protocol. Miscellaneous activity was also relatively consistent on April 6th, staying at around 625 to 750 total instances over almost 18 hours. On April 7th, there were lower instances of miscellaneous activity at around 250 to 325 instances; however, this activity also remained relatively consistent over the day. For both of these activity classifications, it was noted that both began to drop at around 18:00.

For other activity classifications, it was noted that an uptick in unusual activities would usually occur around the same time that generic protocol began to peak. This could possibly indicate someone attempting to distract people from their regular activities or trying to cover their tracks. Potentially nefarious activities — attempted information leaks, potential corporate privacy violations, and potentially bad traffic — all occurred on April 6th, while only privacy violations occurred on the 7th. All negative activities happened within a shorter time span of 6 hours. Known attempted information leaks occurred the most on April 6th, with about 200 to 250 instances, while other activities stayed below 125 instances. On April 7th, however, only privacy violations were noted; these occurred over a 3-hour time period. Privacy violation activity peaked at around 15:00 and dropped immediately. This unusual activity early on in the morning may indicate an individual(s) attempting to cover their activity or traces at network peaks. Additionally, the difference in the trends between April 6th and 7th potentially tells us

that the majority of attempted nefarious activities happened early in the morning of the 6th. However, with the lack of alerts on the 7th about information leaks, it is likely that the suspects succeeded in obtaining the information they needed the next day, after failing on the 6th.

Unusual Activity:

Given the high volume of informational messages in firewall data and the unusual internet patterns on April 6th, we decided to only look at firewall activities that do not have Syslog Priority of "Info" on the 6th. To understand what IP addresses were involved and in what directions these IP addresses were connected, we created an interactive network graph with arrows indicating the direction from Source IP to Destination IP and the width of arrows indicating the amount of traffic (Figure 5). Unsurprisingly, the node that receives the most traffic is 10.32.0.1, the firewall interface to the internet. The nodes that pass large amounts of activities to the firewall are IP addresses 10.32.5.x, which are internet websites accessible to Bank of Money employees. To get a closer look at connections among IP addresses, we zoomed in on one cluster of the network graph (Figure 6). In this cluster are many IP addresses with the format 172.23.x.x; these are individual computers located in offices or cubicles throughout the BOM offices which have visited internal websites. While the network graph does not specifically detect suspicious events, it shows the overall network of the firewall at Bank of Money where individuals visit websites and the websites report traffic to the firewall. This type of graph would also be helpful to examine the network involved for any suspicious event in the future.

Narrowing down again from the Firewall data to the IDS data, we can pinpoint other IP addresses of interest. In order to best represent potential threats, the IDS data was specifically used for threat/negative activity classifications. In Figure 7**,** we can see that multiple workstation IP addresses are attempting to access the IP address 172.23.0.10: the DNS Server which is essential to running the BOM network operations. Each IP address has been noted as a potential privacy violation along with their total number of connections/instances, possibly indicating that potential suspects are attempting to gain access and disrupt an important network. There were also a handful of IP addresses that have been shown to be potentially bad traffic, and one specific IP address stands out with its amount of attempted information leaks (about 160 total connections/instances). The IP address 172.23.231.69 is noted to have an unusually high amount of connections to 172.23.0.1, which is the firewall interface of BOM. Attempting to leak potentially sensitive information from the firewall may compromise the safety of the whole network and disrupt regional activities.

While on April 6th, the network was limited to fewer IP addresses with large amounts of connections, in Figure 8**,** on April 7th, we can see that there were a greater number of IP addresses involved in potential privacy violations at smaller amounts. Similar to earlier, the privacy violations were mostly attempting to connect to 172.23.0.10, which is the DNS server. This suggests that the suspects may have failed to obtain what they needed the previous day

and therefore were trying again on the 7th. With the lack of IP addresses that were alerted to be attempting to leak information, this may signal that they succeeded on the 7th. The lack of alerts from the detection system indicates that security was unable to stop any nefarious leaks. All of these unusual activities and privacy violations point towards compromising the Bank of Money.

Workstation IP Spoofing:

While analyzing the workstations' traffic data from the firewall data, we came across a pattern of disallowed IP addresses that, after further analysis, appear to be spoofed IP addresses. We assume these are coming from workstations inside the network as most of the directions are outbound, if not empty. Workstations have IP addresses of 172.23.X.X, and the IP addresses in question are 172.28.29.X, which are very similar and hard to distinguish due to having only one number in the middle of the IP be different to break the pattern. There are 30 different IP addresses following the questionable format, and we believe there are multiple entities controlling these attacks and working together since the IP addresses used have only different final digits (i.e. the two digits preceding them are .29), creating a pattern within these unknown IP addresses.

These IP addresses are active from 18:00 to 00:00 on April 6, 2012 to April 7, 2012, with the traffic volume dropping drastically but not completely at 00:00 hours. This is the same time period where the attempted data link was detected by the IDS the night prior. These spoofed IP addresses appear to be trying to communicate with permitted websites and the firewall for the corporate headquarter databases, which have the most number of destination requests from hours 18:00 to 23:00.

None of these spoofed IP addresses were detected by the IDS logs. All of their requests were built with 40 less teardowns, which may still indicate an on-going data security issue. Because of the odd hours of activity and most of their requests going towards the data center, we have a strong suspicion of malicious external activity rather than employees trying to bypass the firewall spoofing their IP addresses to visit websites. Due to the similar time frame of this kind of activity during consecutive days, we believe it was the same group of individuals who attempted the data leak the night prior.

Mitigating Cybersecurity Threats:

From the given patterns of malicious activity inside the BOM network over the provided two-day period in April, we have formulated some precautionary measures for the BOM to consider implementing. In order to make network access more secure overall, we suggest that those working to combat cyber threats in the BOM should quickly identify trends in the intrusion detection system to narrow down suspicious IP addresses. From there, they should determine whether or not those IP addresses exist internally, or if they're a product of IP

spoofing in order to imitate the network's workstations. Once this has been determined, they can take action against the threats as needed.

Another alternative to increase security across the entire network is introducing multi-factor authentication to each internal IP address. Compared to the traditional static authentication, this would mitigate the amount of external IP addresses infiltrating the BOM network, as they would not have access to an SSH key for the BOM server. For IP addresses not being handled by one specific worker, extra precautions must be taken by having the cybersecurity team keep a close eye on the activities the addresses undergo.

In addition, BOM could develop an external app for users to access their banking information and conduct business through. In this case, app shielding and real-time fraud detection platforms should be used to ward off external malware infiltration [2].

Finally, having a site-to-site VPN would allow for secure access from each office location of the BOM to the BOM network. This would prevent those not working internally for the BOM from accessing most of their private data and information. Limiting remote access VPNs to trusted individuals of the BOM would minimize remote entry into the BOM network, ensuring that any remote malicious activity could be traced to a specific individual(s). In addition, using VPN provides two advantages over alternative approaches by reducing cost of maintenance and allows for large scale application [1].

Conclusion:

Based on our observations and analysis, we conclude that the Bank of Money organization might have become a target of cyber crime. We believe that the suspect(s) was trying to gain illegal access to sensitive banking information by compromising the network security. It is possible that, following an unsuccessful attempt to information leak on April 6th, the suspect succeeded on their second try. The suspect also attempted a cover-up by creating distraction during the peak of traffic on April 7th. Using clues from the firewall data, we were able to pinpoint a set of questionable workstation IP addresses that might have been impersonated and used by the same individual who tried to infiltrate the network. We hope that our findings will provide valuable knowledge for future direction in the context of cybersecurity and precautions that the BOM should consider.

Data References:

1. BOM Network Description File
2. Firewall Data Files
3. IDS Data Files

Citations:

1. Zhang, Z., Zhang, Y.-Q., Chu, X., & Li, B. (2004). An Overview of Virtual Private Network (VPN): IP VPN and Optical VPN. Photonic Network Communications, 7(3), 213–225.
2. Attack the Hack: How Banks Can Beat Modern Malware. (n.d.). OneSpan. Retrieved April 5, 2021, from https://www.onespan.com/blog/attack-hack-how-banks-can-beat-modern-malware
3. Wikipedia contributors. (2021, February 19). Syslog. In Wikipedia, The Free Encyclopedia. Retrieved 16:54, April 5, 2021, from https://en.wikipedia.org/w/index.php?title=Syslog&oldid=1007727842

Appendix:

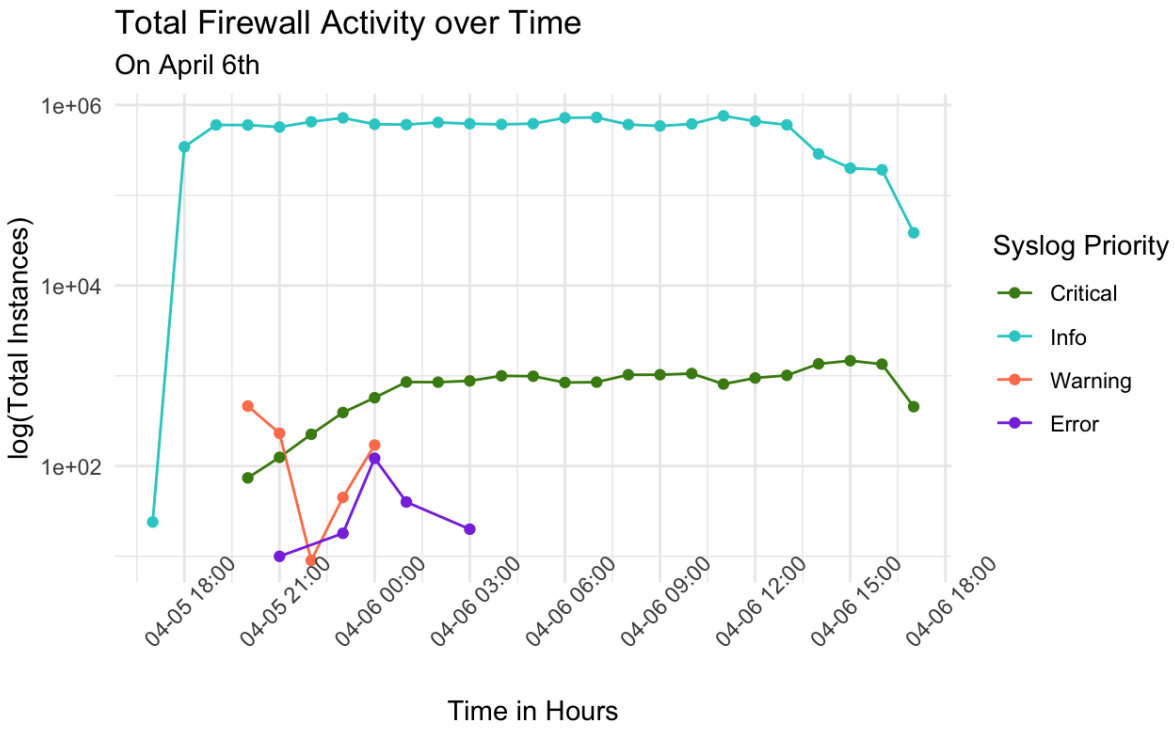

Figure 1. Total Firewall Activity over Time on April 6th.

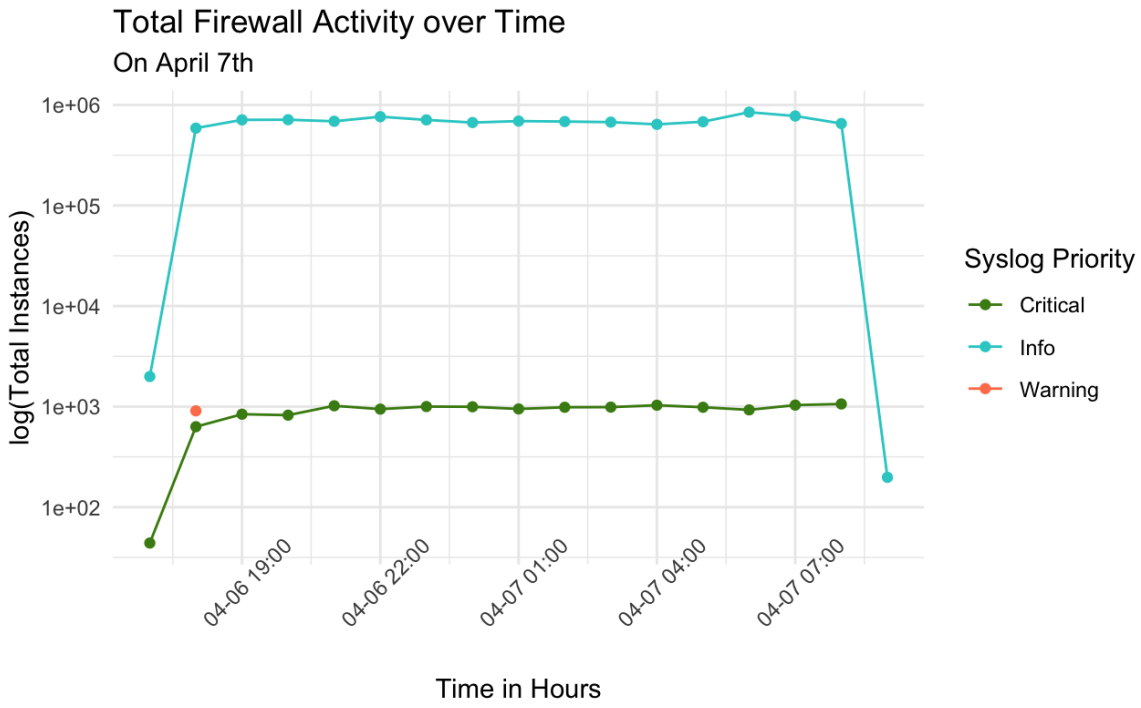

Figure 2. Total Firewall Activity over Time on April 7th.

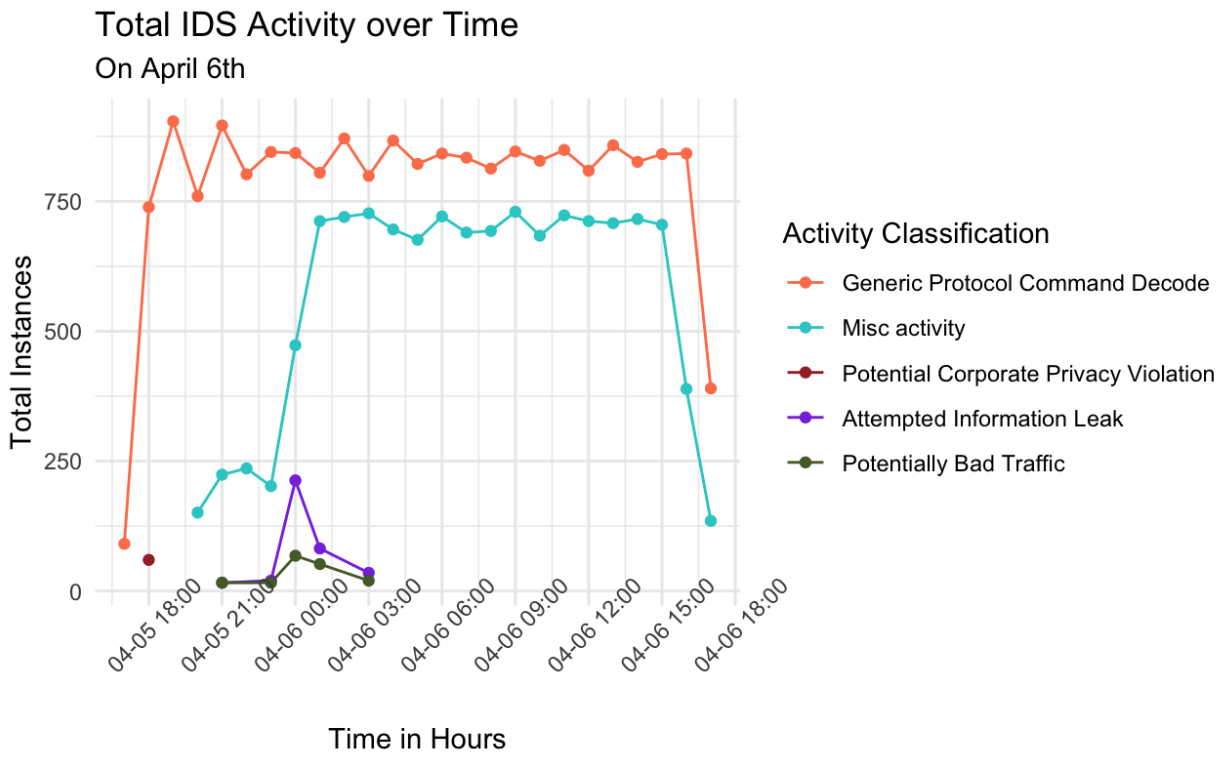

Figure 3. Total IDS Activity over Time on April 6th.

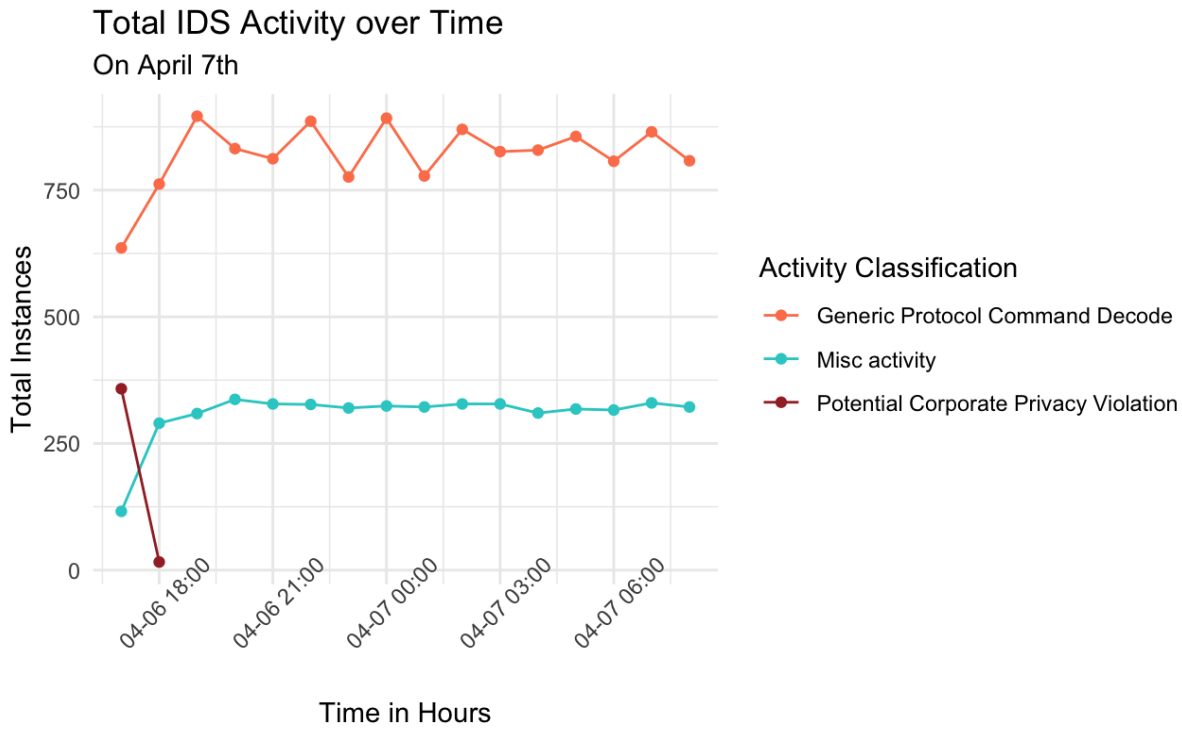

Figure 4. Total IDS Activity over Time on April 7th.

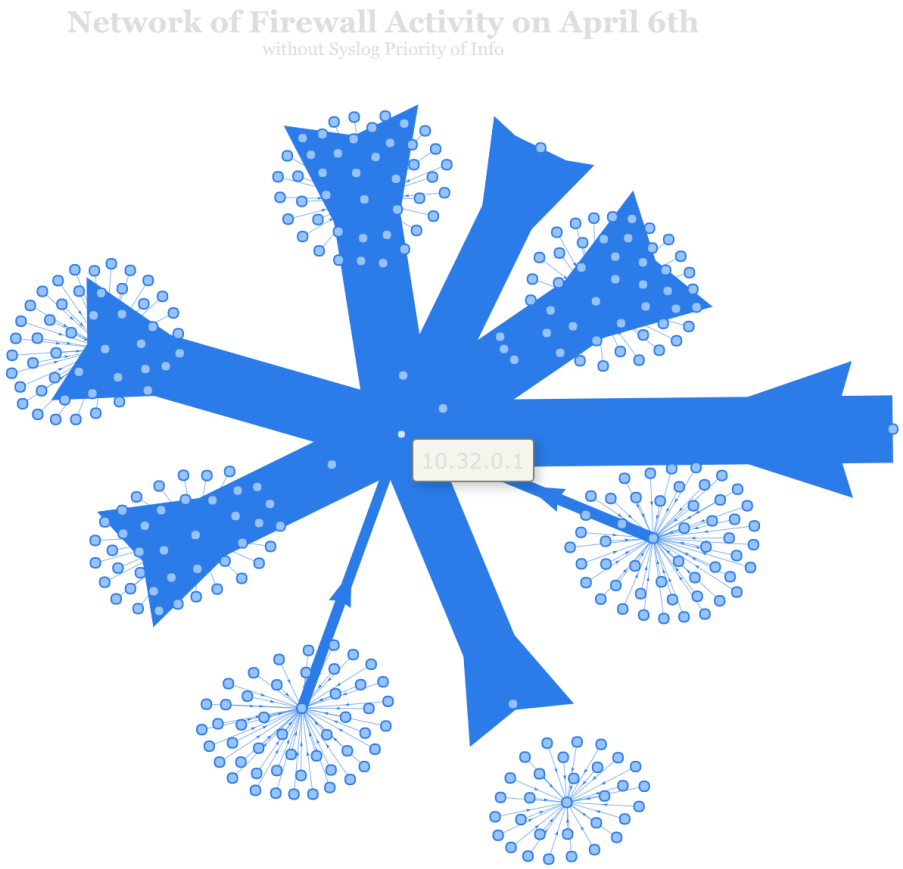

Figure 5. Network Graph of Firewall Activity on April 6th without Syslog Priority of Info.

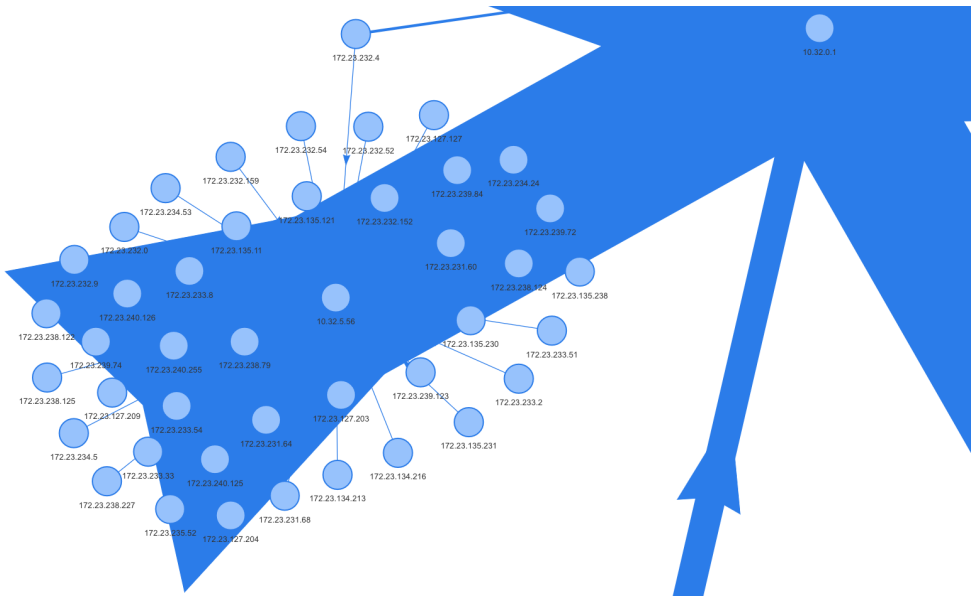

Figure 6. Zoom-in to one cluster of the Network Graph.

(Figures 1 to 6 can be found on "IDS_firewall_trends.rmd/html")

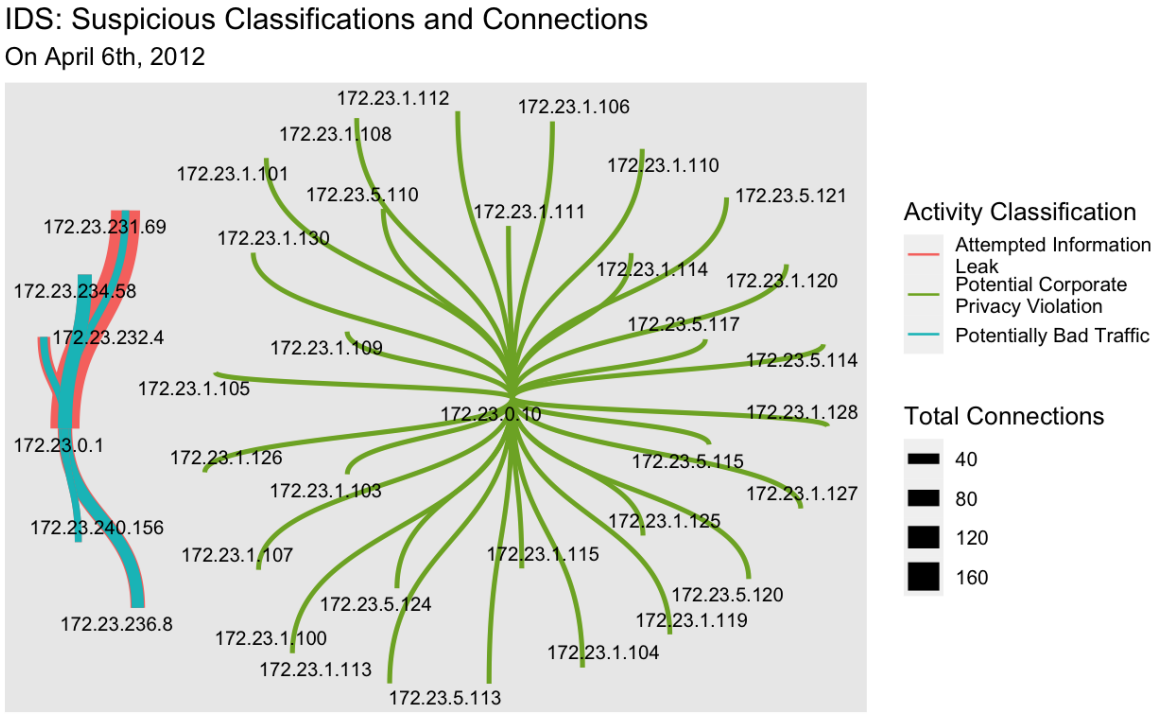

Figure 7. IDS Connections & Suspicious Activities on April 6th, 2012.

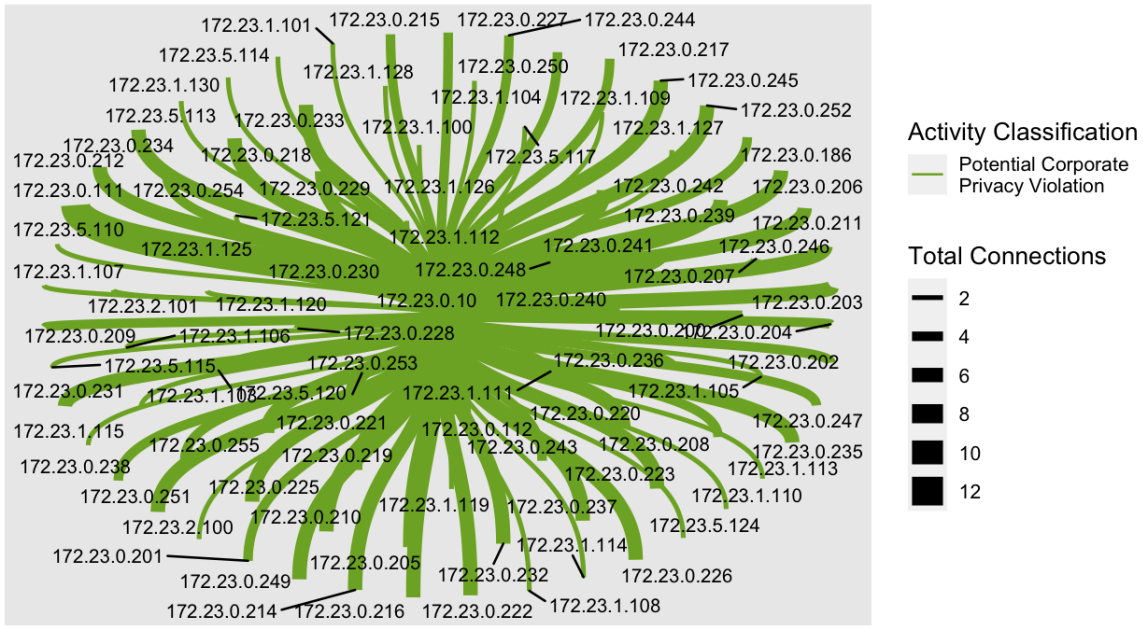

Figure 8. IDS Connections & Suspicious Activities on April 7th, 2012.