Analyzing Advanced Persistent Threats

Kyle Voight

CSC 490-002 | Independent Study

Dr. Leune

December 19, 2020

Table of Contents:

1: Introduction

An Advanced Persistent Threat (APT) is a hacking group, typically backed up and funded by a nation state, who use sophisticated hacking techniques to gain unauthorized access to computer systems and networks for malicious purposes. APTs are commonly heard of in the news due to how many of them exist, and although they are all hacking, each known group has been seen to attack for specific reasons, such as politically or economically. These APTs are labelled with names and numbers. For example, APT41 is Double Dragon, and Dragon is one possible name applied to APTs coming out of China. For this paper, I will be looking at the APT groups that are focused on attacking places that contain intellectual property, such as our academic institution, Adelphi University. APTs are difficult to detect if one does not know what they are looking for, an Indicator of Compromise (IoC). Through resources such as Cisco's Talos Intelligence or FireEye Mandiant, known APT IoCs have been published for the world to see and use to compare against their own networks to determine if they could have been intruded. These IoCs can come in the form of an IPv4/IPv6 address, a malware packet, a domain name, a file hash, and so on. Even though these IoCs have been published, it is possible that they are no longer used by these APTs since they are now compromised, so that must be kept in mind.

With Dr. Leune and I's first meeting, we discussed the processes that must be done to attempt to find an APT on our system. The scope first had to be minimized since it was still a broad idea to search for one. By looking at how Adelphi's network architecture is set up and gathers information of users and entities entering the system, we decided to focus on the netflow and metadata gathered by the firewall. This means that the scope for IoCs will be towards domain names and IP addresses instead of the deep packet analysis (malware packets and file hashes) since those are entirely encrypted and we cannot look at the details of it. Through the use of Adelphi University's live network system, we will investigate and attempt to detect an (or multiple) APT that recently intruded the system. We will use the netflow traffic of every Adelphi campus combined (Manhattan, Poughkeepsie, Happauge, Garden City) to detect an APT that is focused on intellectual property, and then when it is found, it will be studied and learned about. After studying it completely, we will attempt to recreate how the APT did its attack to hopefully help the university in the future with any case relating to this.

2: Initial Methodology

1. Gather notes from resources (FireEye Mandiant) and create a list of APTs that are focused on intellectual property, primarily against academic institutions.
    a. Learn/note their IoCs based on history to refer to when looking for a detection on the live network.
2. Come up with a plan of what I want to hope to accomplish with gathering the APT information.
3. Based on an APT IoC and the knowledge of the university's network infrastructure, try to determine where it will be best to locate the attempted APT access to the university.
4. Determine where I should begin looking, such as the netflow (metadata)
    a. If this does not work, go to deep packet inspection, or vice versa.
5. Locate the APT on the live network from all of the university's netflow combined and note what it was that allowed me to locate it.
6. Determine if the APT successfully entered our system or not, and then begin to learn from it now that I (should) know which APT It is based on cross checking with my notes.
    a. Gather more background information about the APT such as what country it originates from or historical moments from this APT.
7. With all of this knowledge about the attempted attack, I will then figure out why it was done, and what information the APT was attempting to go for or have successfully gotten (SSN, Student IDs, documents from a department, etc.).
8. I will try to use tools such as vulnerability scanners and vulnerability detection methods that I have learned from previous courses to determine more details about the APT going after Adelphi.
9. I will hypothesize how the attack was done and attempt to recreate the attack to further the understanding of this APT and threat detection for Adelphi to use in the future.

3: Initial Feedback

After going through with the initial methodology, I successfully figured out where to look. Based on the layout of the school's structure, we came to the determination that I will focus on the netflow and metadata exclusively since deep packet analysis is encrypted and we cannot look at the specific details behind that information without much more intense work.

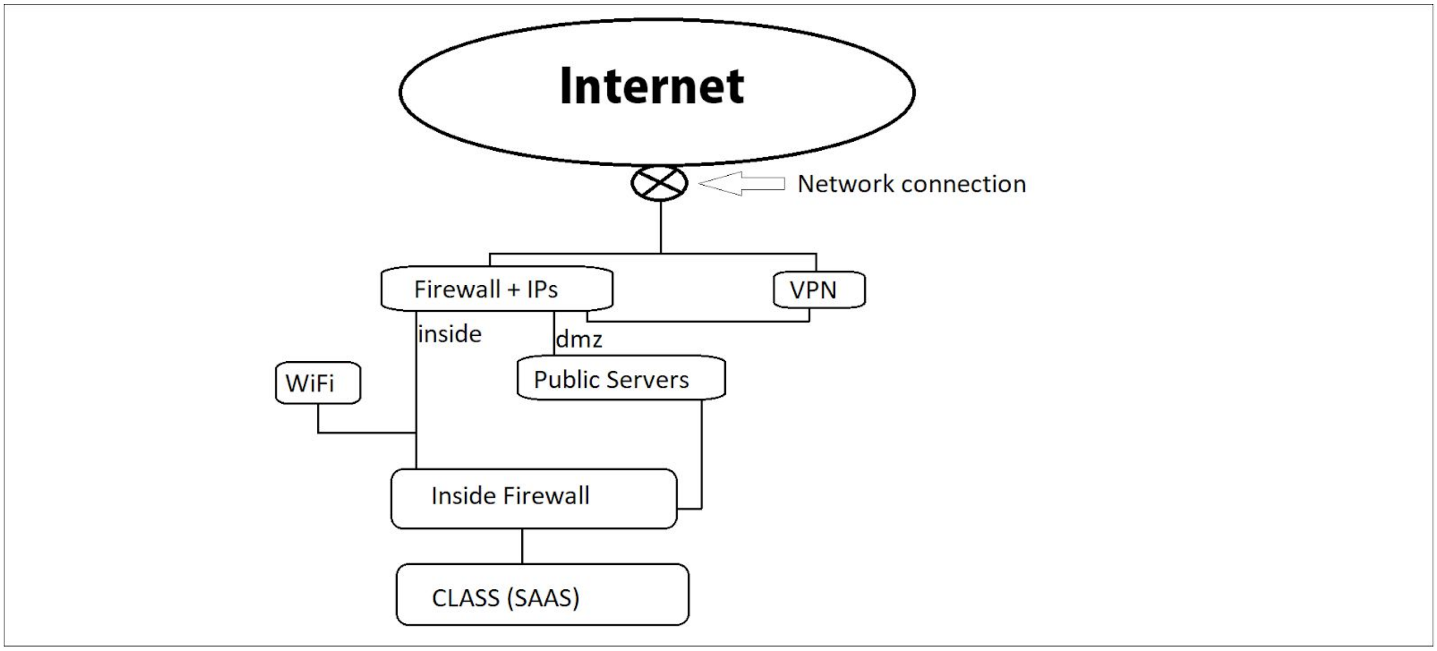


Figure 1. A rough sketch of Adelphi University's infrastructure layout for its live network.

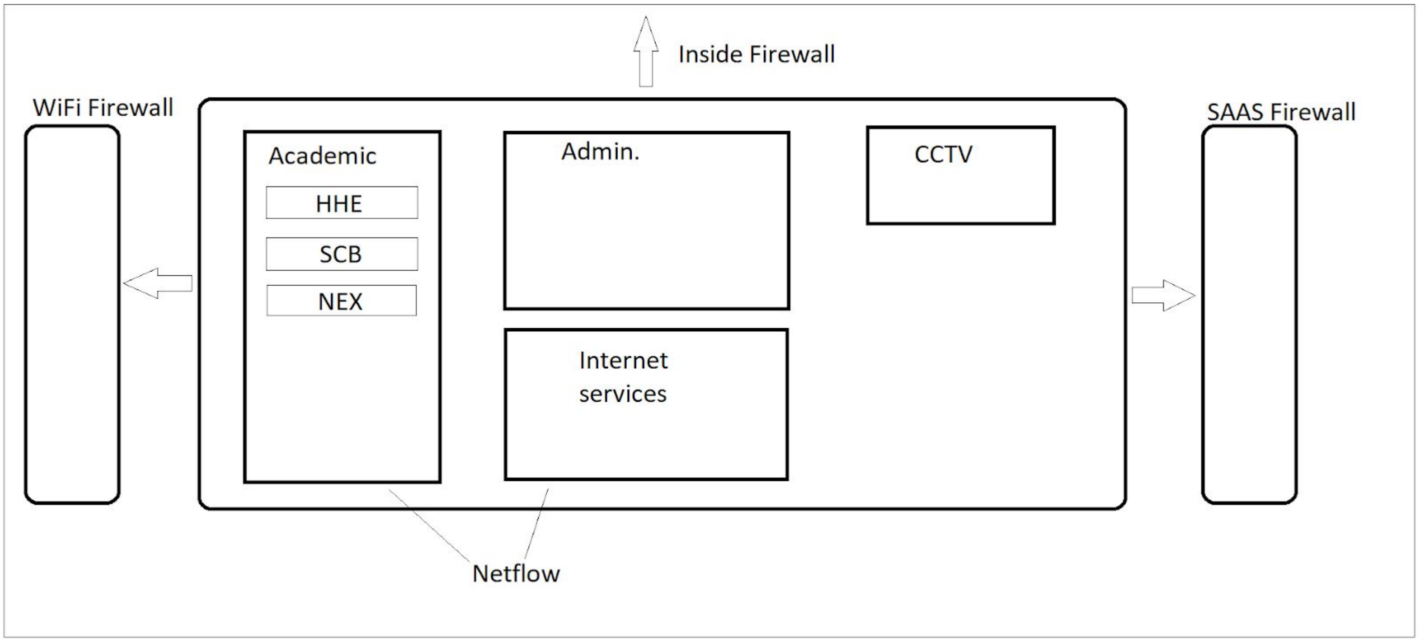


Figure 2. A rough sketch of the inside of the architecture of the system showing how it is laid out and how netflow is gathered from numerous parts of the network.

I created my first sample list of IP addresses and domains from a set of APTs from FireEye (*Advanced Persistent Threat Groups (APT Groups))* to look at, which are:

- APT1 - A Chinese group also known as PLA UNIT 61398 China
- APT3: A Chinese group also known as Buckeye
- APT18: Dynamite Panda (China)
- APT28: Fancy Bear (Russia)
- APT40: Periscope Group (China)
- APT41: Double Dragon (China)

The list consists of 115 IP addresses, IP address ranges, and domain names from these groups. Dr. Leune and I conducted a brief test of only the 79 IP addresses and ranges to look back 7 days and see if the system shows any hits of these IoCs. The first issue was that the ranges were too large and were resulting in too long of a search since there were 3 million IP addresses in these ranges. We ran "ipcalc 233.166.0.0/15" for one range and that alone came back with 131,070 hosts. We scrapped those, especially since these were all in the APT1 range and the IoC list was outdated, and then ran it again. After the revised search, we came back with nothing surprising, there were no matches.

4: First Hypothesis

Due to this finding, we began to theorize why there was no information found and what could possibly be happening. We came to the conclusion that since there were no hits this could be because of 3 things. The first possibility is that the first search was too small of a sample size focused only on IP addresses and not enough to locate any hit or detection on any part of the system. The second possibility is that the memory is volatile, and the third being that because of the COVID-19 pandemic, it is negatively affecting how busy the network traffic is during an average weekday, meaning that there is less room for possible hits and it will be more difficult to find these kinds of attacks from possibly occurring.

5: Revision of Plans

We decided that due to our new theory, there would have to be adjustments made to account for this possibility. I first thought that the domain names for the first list would need to be used and I also considered expanding gathering IoCs to another group of APTs that weren't focused around academic institutions, but more general intellectual property. I then discovered that there was a group that could work well with this new focus. Based out of Asia is a group named Calypso which was a good start for moving towards more general intellectual property attackers (*Calypso APT: new group attacking state institutions* 2019). After looking into this group and its technical details behind how they attack, a new thought crossed my mind, which was to consider looking towards the malware, RATs (Remote Access Trojans), malware droppers, and more of these malicious variants that APTs use for intrusion. I took this idea as well and and looked at a common RAT named Poison Ivy, which is still commonly found 8 years after being first discovered (Fireye site) and used it as a test to see if anything of a general RAT would pop up on our system as well, thinking that if this RAT pops up, I can research which APTs have been known to use it or if there was anything in recent history that involves Poison Ivy with intellectual property theft. This step was good since it meant that even though I was still looking for APTs, there was a somewhat general search going on as well. So my new plan was to gather many more IP addresses from less academically focused APTs, domain names to compare against the systems DNS logs (which is good because the domain names didn't have to be necessarily exactly accurate as the system would pick up on keywords), and Poison Ivy IoCs as well. We also discussed and agreed to reach out to some companies and businesses to attempt to get assistance with their lists of IoCs and any information that could be useful on threat hunting or demos. Those companies were FireEye, PresicionSec, RecordedFuture, and InfoSec, but none of them got back to me with any useful information.

6: Second (new) Methodology

1. Gather new sets of IoCs for both IP addresses and domain names from APTs focused more generally on intellectual property rather than schools and universities.
2. Gather IoCs from recent attacks of common malware and RATs to use as a form of looking both more generally towards whether Adelphi is being attacked, and also if anything interesting occurs on the system.
    a. Take note of any repeating IoCs from older lists to see if those have not been retired since that would mean it's potentially still in active use, leading to a greater interest.

3. Reach out to Adelphi University's senior information security engineer, Don Becker, and work with them to utilize these sets of newly found IoCs to detect intrusion on the network.
4. If there is any detection(s), research what is associated with this detection, such as if it is a common occurrence based on a specific APT, or if there is anything of interest to take away from it.
5. If there are no detections, theorize and conclude why this is happening and what can be taken away from all of this research.

7: New Feedback

I created two new lists containing more RATs and malware IoCs after following the first steps in the methodology above. These two lists contained the Calypso APT and Poison Ivy information that I was planning on getting, as well as a North Korean APT named Hidden Cobra and thanks to Cisco's Talos Intelligence, a list of pieces of malware, worms, droppers, and RATs. Talos posts weekly updates about IoCs of known malware from the previous weeks such as:

- Doc.Malware.Emotet-9774982-0
- Win.Packed.njRAT-9775005-1
- Win.Packed.Razy-9775377-1
- Win.Packed.Gh0stRAT-9776529-0
- Win.Dropper.Tofsee-9775522-0
- Win.Dropper.Remcos-9775269-0
- Win.Packed.Dridex-9776370-1

These malwares are then listed with their domain and IP address IoCs. I took a look at three separate updates for the week of October 9th, then October 30th, and November 6th to try to get a couple of different points in time to see if any of the different dates could lead to a trend or noticeable hits based off of certain periods in time out of interest (Largent, 2020). Dr. Leune and I created a Python script to allow for the domain names to be formatted properly so that it could be read by Adelphi's DNS logs (see Appendix). Instead of checking the IP logs this time of 600+ IP addresses, Dr. Leune insisted I hold onto them and continued compiling a large list of metadata IoCs and formatted domains so that I can use it with Don Becker after we met to begin the new search.

8: Plan After Meeting with Don Becker

After being shown some of the schools systems and how it is gathered and logged via things like Splunk, Mr. Becker asked for me to compile all that I had and send it over for him to evaluate so that I could analyze the results that we get back. I created two more independent lists before combining them, one which involved another weekly update from Talos of the week of November 13th to November 20th. In my second list, I evaluated 4 threats and their IoCs, 2 of which were malware and 2 of which were much more general APTs, which were:

- Emotet - Trojan found in 2014 used to steal sensitive and private information typically spread through email spam and phishing (*Emotet Malware – An introduction to the banking Trojan*).
- Upatre - Family of backdoor trojans used for Windows systems to run additional malware on the system that has been affected by this one (*Trojan.Upatre*).
- APT19 - Deep Panda (China) - Focused on the private sector and military. Using this as an attempt to see if they attempted to get into Adelphi at all which would be surprising (*Deep Panda - Intelligence Team Report Ver. 1.0*).
- APT27 - Emissary Panda (China) - Focused on aerospace, defense, and technology sectors. Using this for a similar reason to APT19s reason (*09/19/2019 - Emissary Panda APT: Recent infrastructure and RAT analysis* 2019).
- APT41 - Double Dragon (China) - This was seen earlier in the project but new IoCs were found (*Double Dragon APT41, a dual espionage and cyber crime operation*).

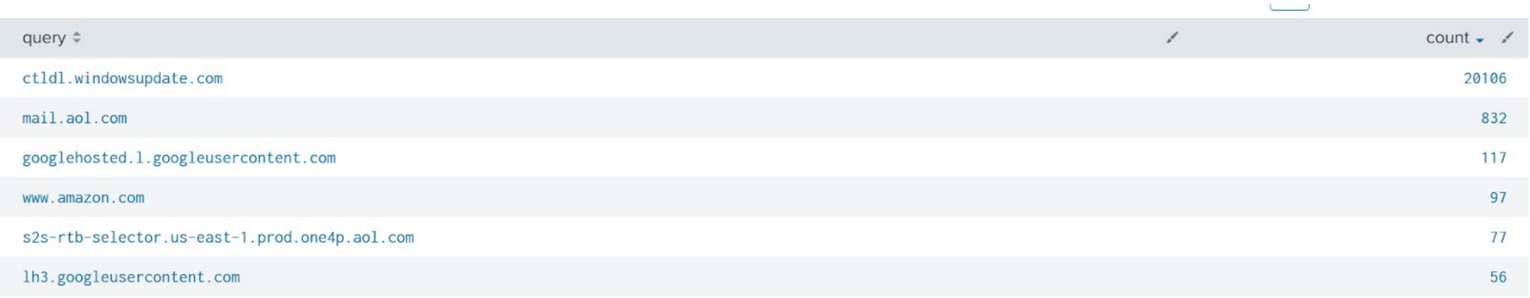
After gathering all of this information, I came up with a final plan for the remainder of the semester.

1. Have Mr. Becker run all of my netflow data against the systems logs gathered from all of these groups.
    a. If he requires any changes to be made with my data so that the search could be more efficient, I will do that.

2. If there are any results, analyze the information to look for anything interesting or surprising from the logs.
3. If there are no results, continue considering why this is the case.
4. Come up with final conclusions about the results.

9: Information Gathered and Results

I put together all of my lists into compiled .txt files for Mr. Becker used and sent him a full IP address list, full domain list, and another full domain list but formatted for the logs to read properly. He ran those and immediately got some odd results. There appeared to be a set of false positives since there was a large number of hits that seemed like it would be impossible if it was malicious.



query	count
ctldl.windowsupdate.com	20106
mail.aol.com	832
googlehosted.l.googleusercontent.com	117
www.amazon.com	97
s2s-rtb-selector.us-east-1.prod.one4p.aol.com	77
lh3.googleusercontent.com	56

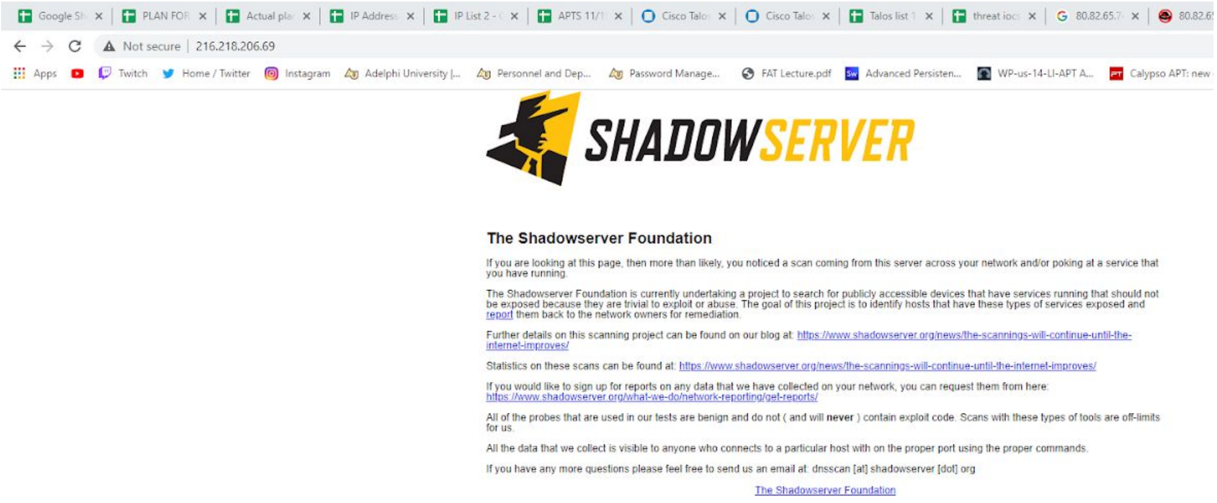
Figure 3. Snapshot of false positive hits from initial list given to Mr. Becker.

As seen from the image, there were well known domains that were in the list. This led to looking at the IoC list from Talos, which showed the domains for one of the malware threats from October 30th to November 6th, Win.Dropper.Tofsee-9786165-0, to contain amazon.com (*Vulnerability Information*). Realizing my mistake, I had to go through the domains and flush all of the ones that are known/not malicious to make sure this would be reduced and hopefully not happen again. I then was asked to put all of my IP addresses in CIDR notation (1.127.0.49 => 1.127.0.49/32), so I did both of those tasks and delivered the final updated list containing 229 domain names, and 558 IP addresses, making the active search 787 pieces of metadata I am searching for. We agreed to look back a week to see if anything interesting would show up. The result was 2 .csv tables. The first table contained a total of 319 logs for the domain hits.

- Name: “icanhazip.com” - This was found in the Win.Packed.Phorpiex-9785125-1 malware IoC, but due to it being denied, nothing of interest was there (Largent, 2020).
    - Total hits: 12
- Name: “ip-api.com” - This was found in the Win.Packed.Razy-9775377-1 IoC malware IoC, but it is most likely because ip-api.com is the main domain and there are documents stored on there which could be potentially malicious, but only the domain is being hit., so this isn’t interesting (Largent, 2020).
    - Total hits: 292
- Name: “pastebin.com” - Similar to ip-api.com and it’s domain name but not the actual link. This was in the Win.Packed.Dridex-9776370-1 but once again nothing of interest was there (Largent, 2020).
    - Total hits: 16

The second table contained the IP traffic from 12:00:00AM on November 30th, 2020 to 11:59:59PM on December 10th, 2020. This contained 1,109,835 logs containing the receive time, source address, destination address, source port, destination port, IP protocol, action, and session end reason. Out of all of these, this is what I found interesting.

- For most of the source addresses, they were rarely of interest as a lot of them were “10.12.10.27”, “10.20.12.30”, or along the lines of “10.” something. However, there were a few of the IP addresses I had listed here, but none of them were from APTs. The ones that were hitting here were from the Talos Intelligence lists, one of the most common being “80.82.65.74”. This appeared on the “Win.Malware.Chthonic-9785809-0” IoC list and when searching about the IP address on Google, found that it is a reported address based out of the Netherlands (Largent, 2020). Since it kept receiving “policy-deny” whenever it would show up, it’s not a problem. Another source address that I found interesting was “216.218.206.69”. When I went to look it up, I came across the page that was displayed below, stating that it is a test to see if the owner's network they are trying to enter has any vulnerabilities. These were denied though with no surprise.



The Shadowserver Foundation

If you are looking at this page, then more than likely, you noticed a scan coming from this server across your network and/or poking at a service that you have running.

The Shadowserver Foundation is currently undertaking a project to search for publicly accessible devices that have services running that should not be exposed because they are trivial to exploit or abuse. The goal of this project is to identify hosts that have these types of services exposed and report them back to the network owners for remediation.

Further details on this scanning project can be found on our blog at: https://www.shadowserver.org/news/the-scannings-will-continue-until-the-internet-improves/

Statistics on these scans can be found at: https://www.shadowserver.org/news/the-scannings-will-continue-until-the-internet-improves/

If you would like to sign up for reports on any data that we have collected on your network, you can request them from here: https://www.shadowserver.org/what-we-do/network-reporting/get-reports/

All of the probes that are used in our tests are benign and do not (and will **never**) contain exploit code. Scans with these types of tools are off-limits for us.

All the data that we collect is visible to anyone who connects to a particular host with on the proper port using the proper commands.

If you have any more questions please feel free to send us an email at: dnsscan [at] shadowserver [dot] org

The Shadowserver Foundation

Figure 4. The Shadowserver Foundation appears after entering the IP address in the URL and states that it is undertaking a project to search for accessible devices and let the network owners know for remediation.

- For the destination addresses, most of them were uneventful. They usually involved the IP net range of Google like “172.217.12.132” and “172.217.10.110” (*IP Address Details*), or Adelphi’s IP net range like “192.160.132.0/24”, or “192.147.12.145” (*IP Address Details*). There were a couple of destination addresses that I did find interesting though, such as “66.242.128.13” which is an IoC for North Korea’s Hidden Cobra APT but is a low risk based off of my information and online (*X-Force IP Report 66.242.128.13*). The log showed that it was denied access though so that isn’t of use. There was also one address which was “67.195.204.77” which is a yahoo.com host name which I found interesting. Outside of that there was nothing too interesting that showed up from the past week.

10: Conclusion

After reviewing everything from the past semester and seeing how my work was affected based on multiple factors, I believe that there are a few things that can be taken from this. The first conclusion is that although the sample size that was used was a good size and there was a good plan laid out, especially with having malware threats and general APTs used to cover the possibility of something unexpected occurring for one of those after being tested against the system, it still may have not been enough. The IoCs used in the list could not have been used anymore since they were published, and the publications varied from either a couple months ago to a few years ago, so that was definitely something I would have adjusted by looking for earlier publications. I think that the Talos Intelligence usage wasn't as reliable or as good enough as I was hoping it would be either so if I had more time, I would have most likely gone with leaning back towards the intellectual APTs IoC searching rather than the malware threats since even if we found one from the malware, it would be incredibly difficult to determine where it came from or who it was sent from. I would have also adjusted my time management differently had I done this again, since some of the parts of this project I didn't expect to take as long as they would, such as the searching with Mr. Becker. The second conclusion about this project is that even though nothing of interest was found, what can be taken from this is that the COVID-19 pandemic has definitely affected how these groups like APTs are potentially attacking institutions and businesses due to it. The degree of traffic on campus has been significantly reduced to a very small level, especially since at this part of the semester there are less people who are on campus at Adelphi's Garden City campus due to residential students having to move back home for finals. So since the traffic has been minimized, APT groups would most likely not make an attack or attemp;t to gain access anywhere since it would be more obvious, and this could mean that they are planning to make another form of attacking or intruding in some other form so businesses and companies such as FireEye and even Adelphi could take into consideration that an APT could end up somewhere completely unexpected in the near future if the COVID-19 pandemic persists.

11: Appendix

Python script used to format the domain names for Adelphi’s DNS logs to use:

```
domains=[“domainsgohere.txt”]
output=[]

for domain in domains:
    out=""
    for f in domain.split("."):
        out+="({size}){domain}".format(size=len(f), domain=f)
    output.append(out+"(0)")

for domain in output:
    print(domain)
```

References

- 09/19/2019 - Emissary Panda APT: Recent infrastructure and RAT analysis. (2019, September 19). Retrieved December 20, 2020, from https://meltx0r.github.io/tech/2019/09/19/emissary-panda-apt.html
- Advanced Persistent Threat Groups (APT Groups). (n.d.). Retrieved December 20, 2020, from https://www.fireeye.com/current-threats/apt-groups.html
- Alert (TA17-318A) HIDDEN COBRA – North Korean Remote Administration Tool: FALLCHILL. (2017, November 14). Retrieved December 20, 2020, from https://us-cert.cisa.gov/ncas/alerts/TA17-318A?es_p=5470315
- AlienVault - Open Threat Exchange. (n.d.). Retrieved December 20, 2020, from https://otx.alienvault.com/
- Calypso APT: New group attacking state institutions. (2019, October 31). Retrieved December 20, 2020, from https://www.ptsecurity.com/ww-en/analytics/calypso-apt-2019/
- *Deep Panda - Intelligence Team Report Ver. 1.0* (Rep.). (n.d.). Retrieved http://cybercampaigns.net/wp-content/uploads/2013/06/Deep-Panda.pdf
- *Double Dragon APT41, a dual espionage and cyber crime operation* (Rep.). (n.d.). Retrieved https://content.fireeye.com/apt-41/rpt-apt41
- Emotet Malware – An introduction to the banking Trojan. (n.d.). Retrieved December 20, 2020, from https://www.malwarebytes.com/emotet
- F. (2014). *POISON IVY: Assessing Damage and Extracting Intelligence* (Publication). Retrieved https://www.fireeye.com/content/dam/fireeye-www/global/en/current-threats/pdfs/rpt-poison-ivy.pdf
- Indicators of compromise for malware used by APT28. (n.d.). Retrieved December 20, 2020, from https://www.ncsc.gov.uk/news/indicators-of-compromise-for-malware-used-by-apt28
- IP Address Details. (n.d.). Retrieved December 20, 2020, from https://ipinfo.io/172.217.12.132
- IP Address Details. (n.d.). Retrieved December 20, 2020, from https://ipinfo.io/AS20306/192.160.131.0/24

- Largent, W. (2020, November 13). Threat Roundup for November 6 to November 13. Retrieved December 20, 2020, from https://blog.talosintelligence.com/2020/11/threat-roundup-1106-1113.html
- Largent, W. (2020, November 20). Threat Roundup for November 13 to November 20. Retrieved December 20, 2020, from https://blog.talosintelligence.com/2020/11/threat-roundup-1113-1120.html
- Largent, W. (2020, November 6). Threat Roundup for October 30 to November 6. Retrieved December 20, 2020, from https://blog.talosintelligence.com/2020/11/threat-roundup-1030-1106.html
- Largent, W. (2020, October 16). Threat Roundup for October 9 to October 16. Retrieved December 20, 2020, from https://blog.talosintelligence.com/2020/10/threat-roundup-1009-1016.html
- Plan, F., Fraser, N., O'Leary, J., Cannon, V., & Read, B. (2019, March 04). APT40: Examining a China-Nexus Espionage Actor. Retrieved December 20, 2020, from https://www.fireeye.com/blog/threat-research/2019/03/apt40-examining-a-china-nexus-espionage-actor.html
- Rewterz Threat Alert – Emotet – IoCs. (2020, October 07). Retrieved December 20, 2020, from https://www.rewterz.com/rewterz-news/rewterz-threat-alert-emotet-iocs-5
- Trojan.Upatre. (n.d.). Retrieved December 20, 2020, from https://blog.malwarebytes.com/detections/trojan-upatre/
- Vulnerability Information. (n.d.). Retrieved December 20, 2020, from https://talosintelligence.com/
- X-Force IP Report 66.242.128.13. (n.d.). Retrieved December 20, 2020, from https://exchange.xforce.ibmcloud.com/ip/66.242.128.13