THE CYBER ESPIONAGE BLUEPRINT:

Understanding Commonalities In Targeted Malware Campaigns

Alex Cox, Principal Research Analyst, RSA FirstWatch

RSA® FirstWatch
Intelligence Report

Table of Contents

Introduction 3

Typical Cyber Espionage Attack Sequence 4

- Exploit 4
  - Watering Hole (Strategic Web Compromise) 4
  - Spear Phish 4
- Initial Foothold 5
- Second-Stage Download / Tools 5

The Cyber Espionage Malware Library 6

- Sample Detection 6
  - Timeline of Submission 6
  - Detection Statistics 7
  - Top Tier Malware Detection 7
  - Takeaways 7
- Host Change Analysis 8
  - Filenames 8
  - Directories 8
  - Takeaways 9
- Network Analysis 9
  - Ports 9
  - Protocol Analysis 9
  - Command and Control 9
  - Takeaways 10

Summary 12

- The Cyber Espionage Attacker Blueprint 12

Table of Figures

Figure 1: Malware submissions collected by RSA FirstWatch by month 6

Figure 2: Results of Detections by Multiple AV vendors 7

Figure 3: Samples Detected by Symantec versus other vendors 8

Figure 4: Breakdown of directories where malware was seen on a host 8

Figure 5: Samples Detected by Microsoft versus other vendors 9

Figure 6: Breakdown of files with specific names versus generic names 9

Figure 7: Breakdown of directories where malware was seen on a host 10

Figure 8: Port Analysis 11

Introduction

"It's 0800 hours, half a world away. Joe sits in a non-descript building in one of dozens of cheap industrial desks that grid the room. He taps on his keyboard, bringing his computer monitor to life. He composes an email, asking his recipient if he received the invoice spreadsheet that he had sent yesterday. His recipient responds that he hasn't, asks about Joe's family, and laments about the yet again rising price of gas. Joe attaches and resends the document, echoing the gas woes. It's a scene that is repeated daily at businesses all across the interconnected world. Joe, however, is different. In fact, Joe isn't even his real name; it's a pseudonym he's using at the moment to ensure his target opens the attachment that he has sent. He is a product of a military assembly line, trained to hack into corporations and pilfer secrets. It's early in his career, and he largely follows a script of what to do, but with some luck and the tutelage of the senior agents in his division, he hopes to move to a tier two job soon. The command console of Joe's Remote Access Trojan (RAT) indicates that a connection has been made from his target's network, which lets him know that his target has opened the attachment and he has established a foothold in yet another Fortune 100 company. It's the start of a good day."

The RSA FirstWatch threat and intelligence research team analyzed over two thousand cyber espionage malware samples that spanned sixty different Trojan families. Key findings from this research:

- 54% percent of cyber espionage malware sample files used random or nonsensical filenames
- 68% percent of cyber espionage malware samples used standard ports to communicate
- 67% percent of cyber espionage malware samples were installed in the user profile directory

Though this scene is fictional, it is intended to reflect activity going on around the world today. Scenarios similar to this are being taking place daily, by both ally and adversary. The world has seen a marked increase in the public reporting of such events. Stolen information ranging from national security secrets to corporate intellectual property is being reported as part of large scale, Cyber Espionage campaigns. These Cyber Espionage campaigns are launched by skilled and militarized attackers whose goals and directives are clearly defined as part of their operational orders. Repeatedly, when these events are detected and investigated, the focus is on the who, what, when, where and why of a specific attacker. Often these efforts are seen as part of an incident response campaign. In this report, the RSA FirstWatch team has elected to focus on the totality of these types of attacks rather than on any specific campaign or attacker. The resulting "blueprint" that has been derived reflects the common methodologies, specifically related to malware most often used in Cyber Espionage campaigns. This includes those attacks associated with nation-states and cyber terrorism organizations.

This attacker "blueprint" was generated over the course of 12 months by the RSA FirstWatch intelligence team after having spent the last year collecting approximately 2400 samples that span 60 different families of Trojans (including first-stage Remote Access Tool (RAT) and second stage backdoors) used in Cyber Espionage campaigns. The malware analyzed was collected from a variety of sources including current events (e.g. news stories and reports), data mining, and both public and private industry information-sharing groups. All samples analyzed related to this report have been used in targeted, Cyber Espionage attacks and have been forensically matched for accuracy. **The RSA FirstWatch team believes that through understanding the basic Cyber Espionage attacker "blueprint", and commonalities noted between many advanced campaigns, organizations have the ability to craft effective best practices for detection and response at both the host and network level. Through doing so, we believe that the playing field can be leveled.**

Typical Cyber Espionage Attack Sequence

While the initial infection vectors of our malware collection are not always apparent, history and experience show that Cyber Espionage-related malware attacks typically occur in the following sequence:

- Waterholing / Spear Phishing Initial foothold
- Second Stage Download & Tools

This order of succession may vary; however, this is the most common sequence of events associated with these attacks. Over the course of the last several years, there have been frequent examples of this type of attack campaigns reported in the news. Some noteworthy examples include Aurora[i], Ghostnet[ii], Elderwood[iii], VOHO[iv], Facebook[v] and Red October[vi].

Exploit

Watering Hole (Strategic Web Compromise)

The watering hole analogy is used to describe what's less commonly known as a "strategic web compromise". In this method of infection, the attacker compromises a website that is of interest to the target and installs some sort of exploit system that will infect visiting machines with their malware of choice. This method is less surgical than others, but the wide net that is cast often can snag targets of opportunity that can be later exploited for further gain.

Spear Phish

A much more directed and stealthy attack involves what is known as a "spear phish". Spear phishing is an e-mail spoofing fraud attack that targets a specific organization. These attacks are driven by the desire to access (in an unauthorized manner),sensitive or confidential data. These attacks are not typically random. When stalking their targets these attackers tend to focus on profit, intellectual property and/or intelligence (i.e. military).

Many times advanced reconnaissance efforts are undertaking to locate potential victims in the target environment. Attackers craft falsified e-mails that appear to be originating from an otherwise trusted source. Many times cleverly orchestrated social engineering campaigns accompany the delivery of these falsified e-mails. Typically, these e-mails invite the victim(s) to do one of the following things:

- Click on an attached document that has be embedded with a Trojan
- Click on an attached file containing other forms of malware
- Visit a link that redirects the victim to an exploit site

Because of the personalized nature of these attacks, they are often times very successful. An excellent breakdown of the methodology used in these attacks can be found in this paper by TrendMicro: http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-spear-phishing-email-most-favored-apt-attack-bait.pdf

[i] http://www.cio.com/article/732122/_Aurora_Cyber_Attackers_Were_Really_Running_Counter_Intelligence

[ii] http://www.infowar-monitor.net/2009/09/tracking-ghostnet-investigating-a-cyber-espionage-network/

[iii] http://www.symantec.com/connect/blogs/elderwood-project

[iv] http://blogs.rsa.com/wp-content/uploads/VOHO_WP_FINAL_READY-FOR-Publication-09242012_AC.pdf

[v] http://blogs.wsj.com/digits/2013/02/15/facebook-we-were-hacked-but-dont-panic/?KEYWORDS=facebook

[vi] http://www.securelist.com/en/analysis/204792272/Red_October_Detailed_Malware_Description_5_Second_Stage_of_Attack

Initial Foothold

Once a target has been successfully compromised, an initial malware package is installed by the exploit. These malware packages are known as a “first stage” and are commonly used to download additional tools and malware and perform data stealing tasks,

Depending on the Tactics, Techniques, and Procedures (TTP) of the attacker, this first-stage malware can be a commodity remote access Trojan such as Gh0st[1], Poison Ivy[2] or Spynet[3].

In other cases, a custom first-stage (unseen in the wild) malware package is installed, largely depending on the sophistication of the attacker group.

Second-Stage Download / Tools

In most cases, once the initial foothold has been established, the attacker will download what is known as a second-stage malware backdoor. A backdoor enables the attacker to circumvent formal authentication processes on an infected host. Once a host has been infected, one or more backdoors may be implemented in order to facilitate easy access to the host in question. In addition to facilitating unfettered access to the host, backdoors are commonly used to download malicious code and content (e.g. password crackers etc.). Many times attackers will leverage Trojans or worms to deliver and implement the backdoor into a susceptible host.

As mentioned previously, these backdoors are implemented in order to establish permanence in the host or environment in question in the event that the first-stage malware attack is detected and removed by defenders. Once the foothold has been established, the attackers will seek to move laterally onto other hosts within the environment. In doing so the attackers will seek to upload and download files while extracting data from the enterprise environment. The tools selected and chosen by attackers vary and are largely dependent on the capabilities of the Remote Access Trojans (RAT) that are in use within the environment.

[1] http://blog.trendmicro.com/trendlabs-security-intelligence targeted-attack-in-taiwan-uses-infamous-gh0st-rat/

[2] http://www.poisonivy-rat.com/

[3] https://www.youtube.com/watch?v=HmdVV3ClisA

The Cyber Espionage Malware Library

The malware library of Cyber Espionage campaigns analyzed by RSA FirstWatch is composed of approximately 2400 samples that span 60 different families of both first-stage RAT, and second-stage backdoor malware used by Cyber Espionage campaigns. RSA FirstWatch did not include second-stage (e.g. sysinternals etc.) tools in this malware collection, as they are often benign tools that are used for malicious purposes.

While this collection is certainly not all-inclusive, the RSA FirstWatch team believes it is a good cross-section of the current malware landscape associated with Cyber Espionage campaigns.

For RSA customers using either RSA Security Analytics or RSA NetWitness, a detailed detection feed can be enabled by subscribing to the RSA FirstWatch APT feeds in RSA Live. The feed consists of command and control servers compiled entirely from the sample set analyzed in this paper.

In order to provide a clear procedural understanding of our analysis, the RSA FirstWatch team has assembled the following order of succession:

- Sample Detection
  - Timeline of submission
  - Detection statistics
  - Top-Tier AV results
- Host Change Analysis
  - Filename
  - Directory
- Network Analysis
  - Ports
  - Command and control servers
  - Network protocol and communication method

Sample Detection

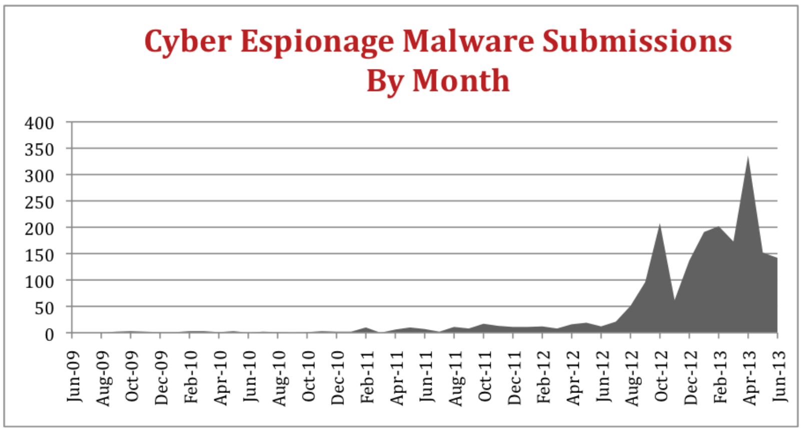
One of the first things the RSA FirstWatch team did as part of this sample collection effort was to search for detection using the VirusTotal search engine. VirusTotal is a service provided by Google, which enables the user to analyze files and Uniform Resource Locators (URL) in order to detect various types of malware including Viruses, Worms, Trojans and Rootkits.

Additionally, information security researchers use Virus Total to detect false positives in samples submitted for analysis that give the appearance of being malicious. Using VirusTotal in this manner (e.g. searching for existing hashes) as opposed to uploading malicious code and content for scanning is desirable for several reasons. It allows the information security researcher to determine how visible the malware sample is to the outside world, either as a victim performing an investigation of a compromise, or as an information security or antivirus firm conducting organic research.

Timeline of Submission

The RSA FirstWatch team found that when we looked at submissions over time, the submission of these malware samples gave weight to the theory that Cyber Espionage attacks are an ever-increasing incline.

Figure 1: Cyber Espionage Malware submissions to VirusTotal by month



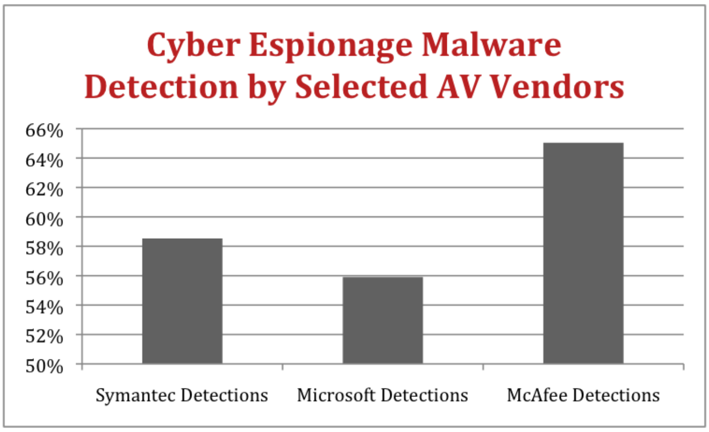
The submissions collected by the RSA FirstWatch team are demonstrated over time are depicted in Figure 1: Submissions by Month beginning in September 2012. Our analysis covers a years worth of malware collection, however we are able to get a wider view of the malware timeframe by observing submission times in VirusTotal. The number of submitted cyber espionage malware submitted increased almost 900% in 2013 compared to the previous years combined. While it appears to be a astounding increase in the number of attacks, it is also possible that defenders have become more aware of these attacks and are simply reporting them more. In either case, the prevalence of these types of attacks is not in question in the current threat landscape.

Detection Statistics

Of the approximately 2400 samples that observed and assessed for detection, the vast majority were detected typically with 16 to 40+ vendors having signatures that matched the particular sample. However, it should be noted that this is the number that have signatures that match today, not at the time of a particular cyber espionage attack.

Though the detection statistics across the antivirus industry are telling, most large corporations use one of a few "top-tier" antivirus vendors. In this case, the RSA FirstWatch team profiled Symantec, McAfee and Microsoft. In almost all cases, the detection rates stood between 55%-65% detection across these three vendors, with McAfee detecting the highest percentage. (Note: the detection rates are given to provide context for our research, not as a comment about the relative efficacy of a particular antivirus)

Figure 2: Results of Detections by Multiple AV vendors



Takeaways

This data supports the observations made by the RSA FirstWatch team; specifically those dealing with the information security industry's comprehension of Cyber Espionage attack trends and frequency. However, it should be noted that malicious code and content associated with Cyber Espionage malware is typically obfuscated in a custom fashion. Put another way, this means that if an existing sample of malicious code and content has been detected in the past, the next use of the same malicious code and content may not be. Antivirus detection typically rise on a linear curve, with detection increasing as the time from the initial infection passes (and defenders discover and submit the sample to their antivirus/antimalware vendors). As a result, it may not be accurate to assert that antivirus and antimalware no longer work with the same degree of efficacy they once did but rather that they do not work well for undefined threats until some time has passed.

Additionally, this data supports the assumption that top tier antivirus/antimalware vendors have a 50/50 chance of detecting advanced threats and attacks.

Host Change Analysis

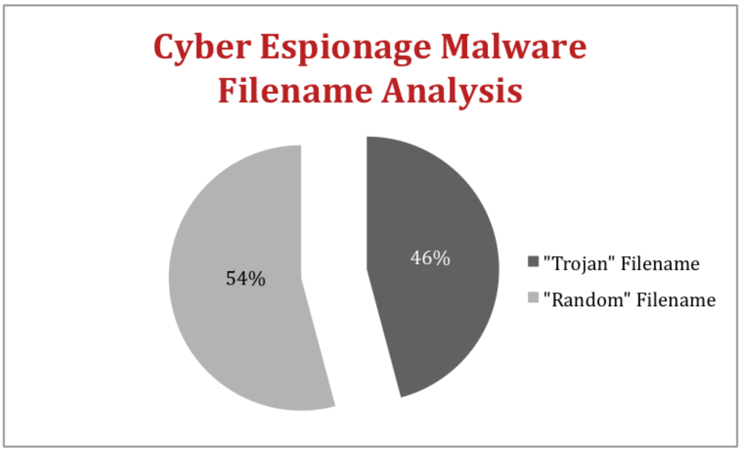
Filenames

When examining changes to the file system, we focused on identifying one of two characteristics about the installed malware.

- Does it seek to mimic an existing process or program (a “Trojan” filename)?
- Does it use a random or nonsensical filename (a “random” filename)?

What we found was an almost 50/50 split:

Figure 3: Breakdown of files with specific names versus random or nonsensical names



Amongst our samples of Cyber Espionage campaigns, “Trojan” names followed a common naming convention, typically seeking to mimic either productivity software, or an operating system file.

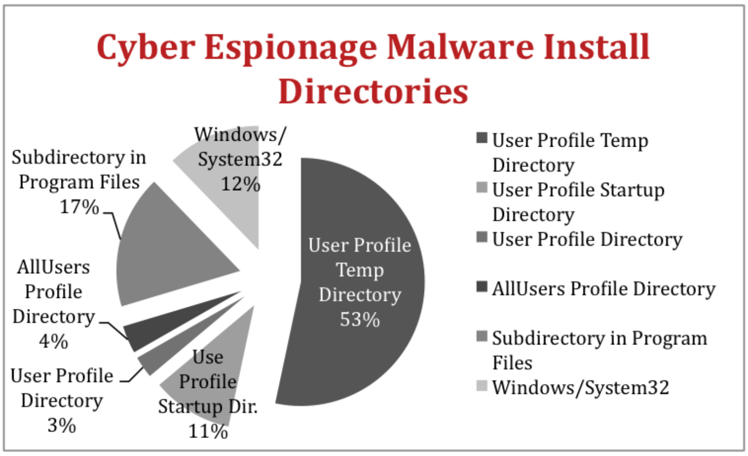
Some of the more common filenames found were: AcroRd32.exe, adobe_sl.exe, AdobeRe.exe cfmon.exe chrome.exe, crsss.exe, current.ext, svohost.exe and scvhoct.exe.

Directories

Malware install directories were split between three categories:

- User Profile Directory
- Windows/System32 Directory
- Program Files Directory

Figure 4: Breakdown of directories where malware was seen on a host



This behavior can largely be explained due to the privilege level of the user during the time of installation. RSA FirstWatch theorizes that this is true unless the exploit in question allows the attacker to run code at the administrator privilege level.

During our research, we noted that a large percentage of the malware analyzed ran directly out of the user-profile temp directory. Additionally, in some cases malware would run out of the user profile start directory, which would also provide malware survivability after reboot. Furthermore, malicious code and content running out of the windows directory and its subdirectories were likely installed with administrator privilege.

Takeaways

This data suggests that a common process running out of an atypical location on the disk may clue the defender to the presence of a compromise. Additionally, the data suggests that it is good practice to investigate random filenames. Furthermore, locations that provide "autorun" functionality after a reboot of the system should be investigated.

Network Analysis

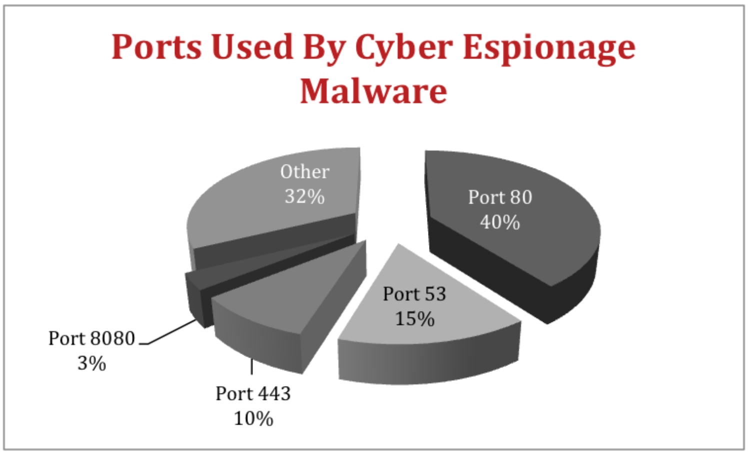
The third part of the research analysis process conducted by the RSA FirstWatch team was to investigate the communication methodology used by the sample set. During this phase of the analysis, the RSA FirstWatch team focused on a number of concepts:

- Ports used for communication
- Common command and control (C2) methodology
- Protocol analysis

Ports

In order for an infected machine to properly communicate with the attacker for control, it must have a clear path into and out of the network. We found that the use of "allowed paths" (network ports that are typically required for business purposes) for malicious purposes was rampant among the collected samples.

Figure 5: Port Analysis

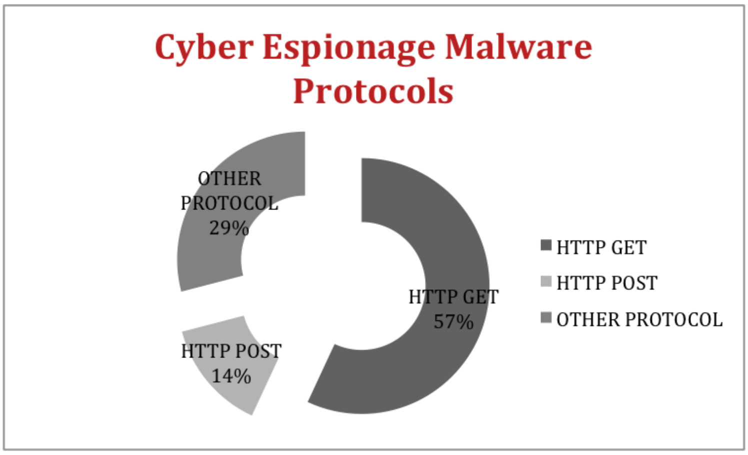


The vast majority of Cyber Espionage malware analyzed during this study used ports 80\8080, 53, 443 (HTTP, DNS, and HTTPS) as expected. The traffic was predominantly TCP-based, with a few UDP outliers. In total, 68% of cyber espionage malware samples used standard ports to communicate.

Protocol Analysis

The majority of samples used a valid HTTP protocol to communicate, with an easily identifiable GET or POST request.

Figure 6: Protocol Analysis



Often the HTTP traffic would reveal identifying information as part of the header request, in the example below the username and computer name are revealed as part of the User-Agent header:

```
GET /default.html HTTP/1.1
User-Agent: 5.1 23:08 SandboxPC\\\\admin
Host: xxxxxxxx.xxxxxpc.net
Cache-Control: no-cache
```

Defenders should recognize that the majority of the samples evaluated used a valid HTTP communication sequence for command and control traffic making them “blend-in” with day to day network communication at most large enterprises. The ability to holistically identify outliers in HTTP communication and unusual variation in header structure and contents is critical in detecting this traffic.

Command and Control

Tracking command and control points for malware is one of the most important concepts to grasp in the Cyber Espionage realm as endpoints are often reused across campaigns. Tracking these endpoints can pay dividends to defenders.

When analyzing Cyber Espionage malware we identified 1268 unique command and control points, both by domain and IP. A basic breakdown is as follows:

Domains

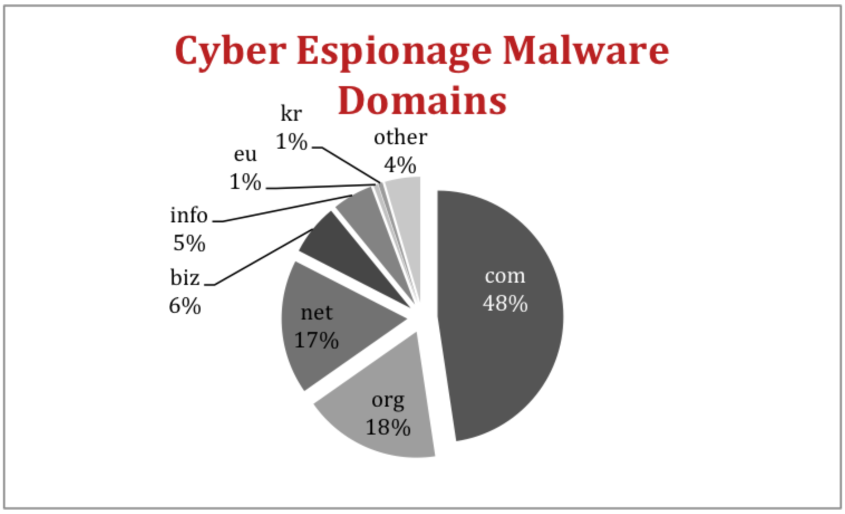


Figure 7: A breakdown of domains used for command and control

As expected, .com, .org and .net are highly represented, with .biz and .info coming in behind them. The high percentage of .org domains is due to a large number of dynamic dns providers such as:

- 3322.org
- authorizeddns.org
- cable-modem.org
- change.org
- changeip.org
- dyndns.org
- hopto.org
- myftp.org
- mypop3.org
- no-ip.org
- onmypc.org
- zapto.org

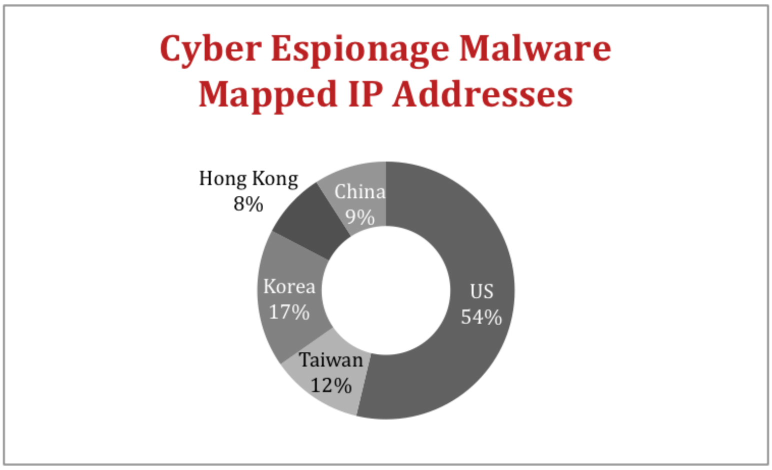
All of these domains were highly utilized in the sample set. Most of the outlier domains were centered around various country codes.

IP Addresses

When conducting the initial analysis the RSA FirstWatch team made certain assumptions about the dataset and associated outcomes of the data once analyzed. When examined in more detail through the mapping of IP addresses to geographic locations, the RSA FirstWatch team noted that 5 countries surpassed all others seen in the data set. Figure 8 below depicts the results of this analysis.

It is important to note that though there is ample information suggesting that certain nation states are more active in Cyber Espionage campaigns than are others, the IP addresses (and associated hosts) may be located in nation states where the attacker is not located. The data generated during this analysis revealed that the United States was number one in terms of IP addresses mapped to a nation state followed shortly thereafter by a number of nations located in Asia. This would suggest that the United States is the 'hub' of activity from an IP address to host perspective.

Figure 8: Mapping IP Addresses to Nation States



Takeaways

The data observed within this portion of the analysis conducted by the RSA FirstWatch team's research suggests that defenders should pay very close attention to the allowed paths in and out of their networks. Additionally, the data provides an insight into the analysis of HTTP header information that suggests that compromises, by the presence of identifying information in the header fields, as well as inconsistencies in header information and construction, occur at an alarming rate.

Furthermore, the analysis suggests that domain research related to dynamic domain providers continues to be abused. This was clearly observed and noted in the *.org space. It should also be noted that atypical domains remain an excellent point of initial as well as continued investigation. Moreover, the analysis suggests that though the traffic bound for some Asian nations is of concern, that a majority of the C2 traffic associated with the sample set was located in the United States of America.

Summary

Through examining more than two thousand samples of malicious code and content related to Cyber Espionage campaigns that represented over sixty different Trojan families, the RSA FirstWatch team was able to identify a common attack blueprint of Cyber Espionage campaigns. This blueprint reveals a number of key points related to the campaigns as well as the tools, techniques and procedures (TTP) associated with the threat actors behind them – points that defenders should bear in mind when performing both network and host based investigations related to Cyber Espionage attacks.

The Cyber Espionage Attacker Blueprint

- Cyber Espionage campaigns typically compromise a target through spear phishing or waterholing.
- Attackers often gain a foothold using a commodity remote access Trojan. However, more advanced attacks use custom malware. Once the initial foothold is established, the attacker will download a second malware backdoor.
- Cyber espionage campaigns commonly try to blend in to go unnoticed. In fact, 67% of the cyber espionage malware samples were installed in the user profile directory and 68% used standard ports to communicate.
- In more than half of the examples (54%) the attacker will use random or nonsensical filenames that provide "autorun" capabilities after reboot.
- The malware often uses dynamic domain providers to facilitate communication.

In order to defend against these types of attacks, RSA recommends the following to security analysts:

- **Focus on Configuration Management:** Configuration management on the host is of critical importance, as a "known good" state allows the defender to zero in on processes that don't fit the norm. These will often be masqueraded as benign processes, but will usually have telltale clues that they aren't legitimate.
- **Know Your Network:** Knowledge of the network is critical for quick detection of intrusion. This goes far past netflow, and ideally requires full session data and protocol detection to be effective.
- **Look at common processes** running out of an atypical location
- **Investigate Random Filenames**
- **Investigate Locations that provide "autorun" capability after reboot**
- **Pay close attention to the allowed paths in and out of your network**
- **Analyze HTTP header information:** This can reveal compromises, both by the presence of identifying information in the header fields, as well as inconsistencies in header information and construction.
- **Review Atypical Domains**

About RSA

RSA, The Security Division of EMC, is the premier provider of intelligence-driven security solutions. RSA helps the world's leading organizations solve their most complex and sensitive security challenges: managing organizational risk, safeguarding mobile access and collaboration, preventing online fraud, and defending against advanced threats.

Combining agile controls for identity assurance, fraud detection, and data protection, robust Security Analytics and industry-leading GRC capabilities, and expert consulting and advisory services, RSA brings visibility and trust to millions of user identities, the data they create, the transactions they perform, and the IT infrastructure they rely on. For more information, please visit www.RSA.com and www.EMC.com.

DRAGON WP 0713