Most Ransomware Isn't As Complex As You Might Think
Yes, we should be able to detect most of it

Dr. Engin Kirda – engin@lastline.com
co-founder and chief architect, Lastline Labs

Ransomware has been widely touted as a highly dangerous, sophisticated and destructive breed of malware. And some of it certainly is. But based on recent academic research into constraints, commonalities and advancements across 15 ransomware families, many of the ransomware families in the wild today are not necessarily as sophisticated or scary as most believe. While certainly even simple programs can extort innocent people who aren't able to separate real from fake cyber threats or protected by advanced security technology, what's important to note is that the many ransomware today doesn't fall into the hardest-to-catch camp that some more advanced threats do. In essence, most ransomware today is a blunt instrument for making a quick profit rather than an advanced surgical tool.

This white paper will explore the topic and examine additional analysis by the Lastline Labs team when taking a closer look at the behaviors of some of the ransomware samples studied in the newly published academic paper "Cutting the Gordian Knot: A Look Under the Hood of Ransomware Attacks" co-authored by Amin Kharraz (Northeastern University), William Robertson (Northeastern University), Davide Balzarotti (Institut Eurecom), Leyla Bilge (Symantec Research Labs) and myself (Northeastern University, Lastline Labs).

In addition, I'll suggest some of the implications for security professionals as they look to protect organizations and individuals from future ransomware attacks.

Historically speaking, the idea of ransomware is not novel. Even back in 1989, a malware instance called PC Cyborg encrypted a victim's files on the hard drive, and asked the user to pay a ransom to retrieve the contents of the files again. Although the concept of ransomware is not novel, what is interesting is that ransomware attacks have been increasing in popularity (i.e., as much as 165% in 2015 compared to previous years). In fact, there have been many recent notable attacks such as the attack against Sony's infrastructure, and the Cryptolocker ransomware that managed to infect approximately 250,000 computers around the world, including an entire police department that needed to pay a ransom to decrypt their documents. Given the significant and clear growth in ransomware attacks, it is very important to develop effective protection techniques against this type of malware. However, designing effective defense mechanisms is not practically possible without having an insightful understanding of these attacks and how they have evolved over the years.

Currently, many of the recent ransomware reports focus on the advancements in ransomware attacks and their levels of sophistication, rather than providing some insights about effective defense techniques that should be adopted against this threat. The general public is typically left with the impression that the ransomware problem is not solvable, and that this specific type of malware is among the nastiest variant of them all. The aim of my Black Hat talk and this white paper is to set the ransomware problem in perspective. Ransomware, clearly, is a problem, and there is no denying that it continues to cause damage on the Internet. However, I believe that compared to other types of malware out there, ransomware has functionality inherent in its nature that behavior-based systems can often use against it. For example, ransomware often aims to delete or encrypt large numbers of files, has to search for specific types of artifacts on the victim's

machine (e.g., documents, pictures, etc.), and often has to iterate through network drives or directories to discover targets. All these behaviors can be used against ransomware to detect it. In 2014, I conducted a study with my Ph.D. student **Amin Kharraz** which resulted in a research paper that I co-authored with **Dr. Leyla Bilge**, **Dr. Wil Robertson**, and **Dr. Davide Balzarotti**. The paper was published at DIMVA 2015 in Milan, Italy ("Cutting the Gordion Knot: A look Under the Hood of Ransomware Attacks". The aim of this white paper is to provide insights into some popular ransomware functionality, and discuss how we can create and design effective detection mechanisms in the near future.

For the study, we used a collection of ransomware samples that were categorized in 15 different families. Our data set covers the majority of the ransomware families that have been observed in the wild between 2006 and 2014. We used multiple sources including manual and automatic crawling of public malware repositories, and repositories that were available to us through Anubis and the Lastline Knowledge Base. The analysis of the data set confirms the folk wisdom that ransomware attacks, in general, have been increasing in numbers and sophistication. However, our analysis also shows that although a majority of the malware samples use some sort of evasion and stealth against detection systems, many of the attacks are not very sophisticated in nature (e.g., such as simply locking the victim's computer desktop, or using superficial approaches to target the victim's resources).

In the study, with my co-authors, we analyzed 1,359 ransomware samples. Our study shows that much of the ransomware file system activity is abnormal, and that it can accurately be monitored by dynamic analysis and detection systems. For my Blackhat talk and this white paper, I dug into Lastline's Knowledge Base to retrieve some more examples and data.

Complexity and Sophistication – what does that even mean?

Some readers might find the claim that most ransomware is not complex somewhat provocative and incorrect. A common way among security professionals to determine the "sophistication" and "complexity" of a malware sample is to analyze the code and look for evidence of evasion techniques that aim to bypass automated detection and analysis mechanisms. For most ransomware samples, this is indeed true, and there is often evidence of stealth and evasion against signature-based detection systems (e.g., the use of packing and obfuscation). In some cases, as we also highlight in this white paper, there are also evasion attempts against behavioral-based dynamic analysis systems (i.e., see Figure 1). However, note that such evasion attempts against detection systems are not unique to ransomware, and are commonly seen in all malware families. Hence, when we talk about complexity and sophistication of ransomware in this whitepaper, we refer to the complexity of *the attack* that the ransomware launches rather than the complexity and sophistication of the code that aims to bypass the detection phase. This is because the evasion part of the code can easily be automated (e.g., there are many packers that can be used to evade signature-based detectors and make the reverse engineering difficult -- such as Themida). However, the concrete attack launched by the ransomware (e.g., deletion, encryption, locking, etc.) needs to be specifically created by the attacker, and as we argue in this whitepaper, these attack behaviors can often reveal a ransomware attack attempt and allow us to detect malicious code that exhibits this behavior.

Evasion	Possibly stalling against analysis environment (loop)
Evasion	Self-modifying code at runtime

Figure 1. Excerpt from a dynamic analysis report for a Cryptolocker variant that attempts to thwart dynamic analysis. Note that this type of behavior is common for other types of malware as well, and is not unique to ransomware.

A Historical Look at Ransomware

Our data set contained a total of 1,359 active ransomware samples that we compiled from public and private sources that were observed in the wild between 2006 and 2014. To obtain accurate labels for these samples, we cross-checked the malware samples by automatically submitting the list of MD5 hashes to VirusTotal. In our labeling, we were conservative, and consider a malware to be ransomware if at least three AV engines recognize it as belonging to this category.

To obtain the family names, we parsed the naming schemes of the AV vendors that are commonly used to assign malware labels. In 77% of samples, AV engines followed the same labeling scheme and our naming policy was mainly based on the popularity of the family name in the community (e.g., Gpcode, Reveton). The remaining 23% of the samples were labeled in an inconsistent way among the different antivirus software, and in this case, we simply selected the most common label among the list of the top 39 AV engines. As one would expect, the data set shows a rapid emergence of new families between 2012 and 2014, as well as a significant growth in the number of new samples in each family.

We observed that 61.22% of the samples (57 variants) only targeted the desktop of compromised computers, without touching the documents on the file system. We also observed the emergence of other malicious activities, such as changing the browser setting or performing multiple infections to install other malware, in 3.23% of the samples. Despite the fact that the number of samples performing additional malicious activities (e.g., stealing private information) is not alarmingly high right now, this activity has been increasing. However, this type of activity is not unique to ransomware, and is classic malicious behavior.

Encryption Mechanisms

About 5% of the samples among four families in the samples we looked at employed some encryption mechanism during our experiments. As most would expect, current ransomware samples use both customized and standard encryption techniques during attacks. A common approach is to use the strong cryptographic functionality that Windows platforms provide (e.g., CryptoAPI). After all, these libraries are widely available, and they are easy to use. One of the goals of crypto-style ransomware is to make it difficult to recover a victim's data without the correct key. As is widely known, modern crypto-based ransomware families like Cryptolocker (see Figure 2) and Cryptowall do a good job at making the victim's data difficult to retrieve and use standard libraries. This was not always the case, but the attackers have learned from their mistakes and their creations have evolved over the years. These families make use of the Windows *CryptEncrypt* function with a handle to the encryption key and a pointer to a buffer that contains the plaintext to be encrypted. In these families, the plaintext in the buffer is directly overwritten with the encrypted data created by this function, hence, making the recovery very difficult. Clearly, ransomware families such as Cryptolocker and Cryptowall have received much attention because of their use of advanced encryption techniques, and are widely considered to be sophisticated because of their correct deployment of encryption. However, many other ransomware families also exist that do a bad job in cryptography. For example, a Gpcode variant generates a static key during the attack that can be recovered during dynamic analysis by comparing encrypted files with the original ones. Similarly, TeslaCrypt was claiming to use asymmetric RSA-2048 to encrypt files, but was using symmetric AES instead, and its encryption has been successfully broken.

If a standard crypto API is used, it is typically easier to detect this behavior during dynamic analysis. At the end of the day, the ransomware is depending on existing, standard functionality on a system that creates an opportunity for close monitoring. While it is true that a system that has been attacked would be difficult to recover if strong cryptographic APIs are used, at the same

time, this specific functionality and behavior of ransomware allows to construct dynamic analysis and defense systems that can look for this behavior. Hence, what the community generally considers to be a sophisticated ransomware behavior can be a double-edged sword for the attackers.

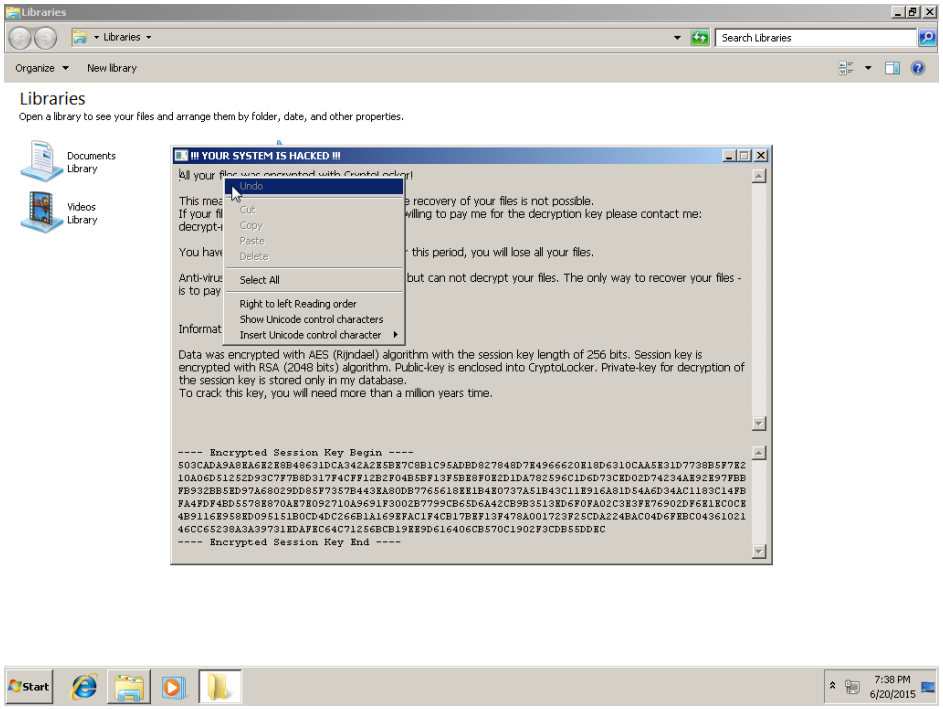

Figure 2. A ransom dialog screenshot captured automatically that a Cryptolocker variant shows during automated dynamic analysis.

Deletion Mechanism

More than 35% of samples among the five common ransomware families in the data set did not perform any encryption, but deleted a victim's files if the ransom was not paid. Most ransomware families we observed deleted files in a very straight-forward way: In NTFS, each file has a Master File Table (MFT) entry that reflects the changes on the corresponding file or folder. When a ransomware attack occurs, the malware lists the non-system files and initiates a delete operation for each of them. The MFT entry for each file is updated by changing the status flag value of the file from 0x01 to 0x00. Consequently, when a file is deleted in a typical ransomware attack, the MFT entry is updated, but the content of the file is not deleted immediately. Hence, our study suggests that monitoring the changes in the MFT table can be an effective venue for detecting ransomware during dynamic analysis and detection. Also, in many ransomware attacks where files are deleted, there is a good chance that the deleted files can be recovered because they have not been securely wiped from disk.

Locking Mechanisms

A classic ransomware behavior is to lock the desktop of the computer under attack so that it becomes inaccessible to the victim. This is typically done by creating a new desktop and making it persistent. Many ransomware instances simply use *CreateDesktop* to create a fresh desktop environment and eliminate unnecessary processes. The new desktop is created via a DESKTOP SWITCHDESKTOP access mode that enables the *SwitchDesktop* function to activate the new desktop, and receive input from the victim. The desktop is assigned to a thread using the *SetThreadDesktop* function. A significant number of samples in our data set (61%) use very similar approaches to establish a persistent desktop lock. Other samples (8 variants) in families such as Urausy, Reveton, and Winlock employed another straight-forward approach to lock the desktop. In these families, the lock banner is simply downloaded as a HTML page with corresponding images based on the victim's geographical location, and it is then displayed in full screen in a IE window with hidden controls. The banner plays a local law enforcement warning in the language used in the victim's geographical location. The warning typically says that the operating system is locked due to infringement against certain laws (e.g., distributing copyrighted materials or visiting illegal sites) in that location.

The Achilles' Heel for Ransomware: Looking for Files to Attack and Having to Inform the Victim that the Attack has Occurred

Clearly, the fact that ransomware has to inform the victim that the attack has taken place is, at the same time, a weakness that is inherent in its nature. This is because such locking mechanisms and fake warning messages can be used as telltale signs for behavioral dynamic analysis and detection systems that the sample under analysis is a ransomware instance. For example, Figure 3. shows excerpts from a dynamic analysis run that looks for specific signs that ransomware might exhibit such as iterating over large numbers of files over many directories and network drives, and displaying a document or a modal dialog to the user while maintaining background activity.

File	Searching for files across mounted drives
File	Searching for files iterating over directories

Stealth	Displaying documents while maintaining background activity
Stealth	Displaying images while maintaining background activity

Figure 3. Excerpts from a dynamic analysis report for Cryptolocker that lists activities that are telltale signs for ransomware. Once the malware successfully executes, ransomware-like behaviors such as searching over many files are difficult to hide.

Conclusion

Drawing from these observations, it is apparent that many ransomware variants are relatively straight-forward when compared to other more advanced malware that aims to remain completely stealthy. In fact, only a small fraction actually irrecoverably deletes the files it threatens to if the target victim refuses or is unable to pay. Additionally, a sizeable portion of the ransomware families studied don't use encryption. Thus, most ransomware is something the security

community has the tools and means to hold at bay until the attackers wielding it, perhaps inevitably, dial up the sophistication. Additionally, victims of ransomware may consider these findings before paying the ransom without seeking professional help first.

About the Author

Dr. Engin Kirda is chief architect at global breach protection provider Lastline - which he co-founded in 2011 - as well as a full professor of computer science at Northeastern University in Boston. He has co-authored more than 100 published research papers. Before Northeastern, he held faculty positions at Institut Eurecom in the French Riviera and the Technical University of Vienna where he co-founded the Secure Systems Lab that is now distributed across multiple institutions in Europe and the U.S. Engin's recent research has focused on malware analysis and detection, web application security and practical aspects of social networking security - including the de-anonymization of social network users. He has served on program committees of numerous well-known international conferences and workshops. In the past, Engin has consulted the European Commission on emerging threats, and gave a Congressional Briefing in Washington D.C. on advanced malware attacks and cyber-security. He also spoke at SXSW Interactive 2015 about "Malware in the Wild."

About Lastline Labs

Lastline Labs is the research and development arm of Lastline Inc., where some of the best minds in the academic community collaborate to advance novel technologies relevant to cyber security. Our team analyzes new security threats, vulnerabilities and hacking techniques as well as the evolution, proliferation and impact of advanced malware. Headquartered in Santa Barbara, California, our team consists of over 30 expert scientists and engineers throughout the Americas, Europe and Asia. Lastline Labs enables Lastline Inc. to continue to develop cutting-edge technology to defend against evolving threats and shape the future of security.

Notable Presentations

Lastline Labs
6950 Hollister Avenue, Suite 101, Goleta CA 93117
Lastline.com +1 (805) 456-7076