RANSOMWARE:
A Cyber Risk Prioritization Model

A Master Thesis

Submitted to the Faculty

of

American Military University

by

DANIEL WOLFFORD

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

December 2020

American Military University

Charles Town, WV

COPYRIGHT

DEDICATION

I dedicate this thesis project to the United States Air Force, Space Command, U.S. Cyber, and the global hacker community. I also want to recognize the integrity, service, and excellence offered to me by family, friends, coworkers, and colleagues throughout this academic enterprise.

ACKNOWLEDGMENTS

Big thanks to Dr. Harry Cooper for inspiring me to hold myself to a higher standard. Thank you to David Berns, Matt Brill, Rosangela Carreon, Dave & Julie Dudley, Andrei Ivasko, Sean Kaufman, Robert Statica, Rei & Sy Stern, Meg VanDeventer, and Danielle Wolfford for encouraging me to pursue the creation of “DYRWOLF”, the world’s first and only ransomware prevention and intelligence software company. With a grateful heart, I acknowledge the unconditional love of my immediate family: Debra, Bryan, Xawn, Danielle, Michael, William, Victoria, and James.

ABSTRACT OF THE THESIS

RANSOMWARE:

A Cyber Risk Prioritization Model

This project is a Master's Capstone Thesis based on the research, analysis, reporting, and mathematical prioritization of ransomware. Publicly published, open source intelligence reports on more than 200 different ransomware attacks were collected and analyzed over the past 16 weeks to create a new cyber risk management model. A review by the Institutional Review Board (IRB) was not required for this project because there are no human subjects. Starting with the literature review, we compare historical and modern examples of ransomware intelligence from multiple sources. Literature generally focuses on one of three primary subjects: the attacker, the victim, or the malware. Our model combines all three types with information on many different types of ransomware including technical analysis of the code, the criminals themselves, their victims, affected sectors, tactics, commands, payment type, ransom amount, contact information, and the answer to questions that can be used to calculate the likelihood and impact of an attack. We also go into detail about the data collection technique, sources, formatting, statistical analysis, and limitations. Finally, we analyze the data and discuss the results in terms of the likelihood of a successful ransomware attack, impact of a successful attack, risk to critical sectors, trends, and predictions.

Keywords: cybersecurity, exploit, prioritization, ransomware, risk, threat

TABLE OF CONTENTS

Introduction

Ransomware is a form of extortion that takes place in Cyberspace, the fifth domain of warfare. Essentially, attackers gain unauthorized access to a computer system through various means, encrypt the victim's files, then demand a payment for the decryption code. Ransomware attacks were not very common up until a few years ago when insurance companies decided it would be cheaper to pay the ransom than to manually recover the files, thus creating a booming new business model. Before the age of ransomware, most companies thought it was good enough to have onsite back-ups, if at all. As ransomware attacks became more prominent, many companies responded by upgrading their back-up strategies to include offsite & offline copies of important data. Predictably, attackers responded by adding new tactics such as data theft, also known as exfiltration. Now if victims refuse to pay the ransom, attackers call them out publicly on various websites and threaten to expose their stolen data.

The purpose of this thesis project is to develop and implement a novel cyber risk management model focused solely on ransomware threat intelligence. The plan is to research, analyze, and report as much information with context as possible over the next 13 weeks, so that many different ransomware families can be mathematically prioritized based on the likelihood of a successful attack and the potential impact. The literature review will explore and discuss factors related to the rise and abundance of ransomware attacks and attempt to uncover the identities and motivations of attackers. Individuals, small businesses, multinational corporations, and entire cities are getting hit with ransomware and it is getting measurably worse. Hopefully, this project will help the community understand the problem, come up with better solutions, prevent attacks, and ultimately put cyber criminals out of business.

Literature Review

Before we dive directly into ransomware analysis and prioritization, we should first discuss the history of how we got here from the early days of malicious software, also known by the portmanteau “malware”. Chen & Robert’s (2004) research on the evolution of viruses and worms indicated the first known program to replicate itself and spread to multiple nodes on a network happened way back in 1971. The program, dubbed “Creeper”, was developed by Bob Thomas and probably not meant to be malicious, but it connected to remote computers, used their resources without consent, and dropped a note to “catch me if you can” (Meltzer & Phillips, 2009). Programs that self-replicate over a network are called worms and worms are a type of malware.

Another type of malware is called a virus. Chen & Robert (2004) wrote how the first usage of the word “virus” to describe a self-replicating computer program, dubbed “Brain”, was written by a student named Fred Cohen in 1983. The virus even tried to hide itself unlike Creeper which was clearly overt. Worms and viruses became so much of a nuisance to computers and networks by 1987, companies like G Data Software and McAfee began to offer anti-virus solutions (Chen & Robert, 2004). Viruses sound similar to worms but have two main differences: 1) Viruses need some type of host executable to spread, while worms spread all on their own 2) Viruses typically do not spread to other computers without human involvement such as moving a USB drive, while worms spread quickly from one computer to another across the network without human involvement.

The first known description of a virus being combined with cryptography to be used exclusively for extortion was written in 1996 by Young and Yung for the Institute of Electrical and Electronics Engineers (IEEE). They proposed a malicious use of public-key cryptography to target computers with what they called a “cryptovirological attack” (Young & Yung, 1996). Little did they know that their idea would be the greatest unsolved cyber threat more than 20 years later and counting.

The most abundant source of literature focused on ransomware is considered to be open source intelligence (OSINT) and can be divided into reports based on evidence that was observed first hand by legitimate cyber security analysts and/or reports based on other reports. Occasionally, a security vendor’s marketing department notifies various news agencies on reports from criminals where a particular victim did not pay their ransom. This may be considered a benefit to the victim’s customers assuming their data was stolen. Victims would at least be notified of the theft before their data was leaked by the attackers. There are some in the security community that view this practice as a form of victim shaming.

We acknowledge that more often than not, there is value from reports on reports when the author adds context or makes connections that may have otherwise been overlooked. For example, Sophos commissioned an independent study of 5,000 IT managers across 26 countries, then released The State of Ransomware 2020, in May (Bourne, 2020). Ransomware is not reported nearly as often as it occurs due to a combination of factors including embarrassment, privacy, and victim shaming (Cartwright & Cartwright, 2019). Results of the survey indicate a large majority ransomware attacks are simply not reported by mainstream media.

Out of the two types of OSINT reporting, information that comes directly from the source such as Alert (AA20-049A) Ransomware Impacting Pipeline Operations written and published by the Cybersecurity and Infrastructure Security Agency (CISA, 2020), Navigating the MAZE: Tactics, Techniques and Procedures Associated With MAZE Ransomware Incidents published by Fire Eye (Kennelly, Goody, & Shilko, 2020), A deep dive into Phobos ransomware by

Malwarebytes Labs (Doniec, 2019), EKANS Ransomware and ICS Operations by Dragos (Dragos, 2020), Thanos Ransomware: Destructive Variant Targeting State-Run Organizations in the Middle East and North Africa by Palo Alto (Falcone, 2020) are preferred. Reports like these contain detailed malware analysis and actionable intelligence that can be used to detect ongoing attacks at other organizations and even prevent attacks if companies are proactive.

Another valuable source of information we are using for the project comes from academia in the form of peer-reviewed journals and research. These papers are not as easy to find and process as OSINT, but at least we can rely on the fact they are not published by marketing departments. Articles like Collier's article about the attack on NHS (Collier, 2017), Riddel's analysis of the attack on Garmin (Riddel, 2020), Salvi's Ransomware-as-a-Service (Salvi, 2019), and Arabo's Detecting Ransomware Using Process Behavior Analysis (Arabo et al., 2020) discuss details about various ransomware families, attack methodology, economic impact, and cyber criminals perpetrating the attack.

In terms of subject matter, literature about ransomware is typically focused on one of three subjects: the attacker, the victim, or the ransomware. The goal of our thesis project is to create a new risk management model that helps analysts quickly process multiple sources and prioritize threats by extracting the most relevant data points in order to answer specific questions about the threat. Authors may use this model to help shape their works, so they do not skip over important information. With the combination of better reporting and better data, decision makers can focus their attention and security budgets on threats that are most likely to occur and/or have the highest impact on their organizations. To the best of our knowledge, no other researchers have made an effort to combine all three aspects to prioritize ransomware in terms of risk.

Malware Focused

The most references to ransomware found by keyword searching are focused on malware analysis. Researchers like Mauro Conti et al. (2018) and Harshada Salvi (2019) presented their findings on ransomware and associated economics. Both papers covered various aspects of ransomware including functionality, distribution models, encryption algorithms, and ransom amounts typically paid in bitcoin. Harshada Salvi (2019) was focused on five different families of Ransomware-as-a-Service (Frozr Locker, Jokeroo, Philadelphia, Satan, Stampado) and how entrepreneurial attackers host, license, and distribute ransomware using a revenue-sharing subscription model. Salvi also detailed pricing, customization options, encryption algorithms, file extensions, tactics, techniques, and defensive measures. Conti et al. (2018) described 20 different types of ransomware with great breadth and depth on initial attack vectors, family evolutions, ransom payment addresses, transactions, statistics, and timelines. It is also worth noting Conti et al. (2018) provided more information than usual about the attackers associated with each ransomware family although it was not their primary focus.

Abdullahi Arabo et al. (2020) and Doniec (2019) both analyzed ransomware process behavior. Arabo et al. (2020) used ten different machine learning classifiers to determine if they could distinguish ransomware from benign software. In their lab environment, they collected data for seven different families of ransomware (Cerber, Krotten, NoMoreRansom, ViraLock, WannaCry, WinLocker, and Petya) by monitoring RAM, network connections, threads, API calls, DLL loading, command line info, executable path, disk usage, and file read/writes. However, they did not mention details about pricing or encryption algorithms like Salvi (2019) or Conti et al. (2018). Doniec (2019) focused on a single family known as Phobos, which has strong ties to another ransomware known as Dharma or CrySis. Blogs like this are more appealing to average readers with screenshots and visualizations when compared to more academic resources like Arabo et al., Conti et al., and Hull et al.

Hull et al. (2019) addressed the challenge of categorizing ransomware by developing a classifier, a Windows based Application Programming Interface (API) scraper, a safe environment for testing, a predictive model based on behavior (dubbed “RanDep”), and like Conti et al. (2018) analyzing many different families of ransomware (Cerber, Chimera, CTB-Locker, Donald Trump, Jigsaw, Petya, Reveton, Satana, TeslaCrypt, TorrentLocker, WannaCry, CryptoLocker, Odin, Shade, Locky, Spora, CryptorBit, and CryptoWall). Their model divided ransomware attacks into eight stages (fingerprint, propagate, communicate, map, encrypt, lock, delete, threaten) and unlike most other authors, they interviewed victims to understand how ransomware works in the real world (Hull et al., 2019). As mentioned by the authors, a major improvement would be realized if they add a machine learning capability to the model in future work.

Interviews carried out by Hull et al. (2019) demonstrated that a wide range of people have experienced ransomware firsthand from students to experts. Universities are targeted more than other organizations due to their large presence online, also known as their digital footprint. Most victims had minimal security such as basic firewalls and antivirus. Email was identified as the most common attack vector. Endpoint security, user education, stricter policies, access control, and better staffing were recommended as solutions. This particular research paper is unique in that they do not offer specific mitigation strategies other than what was mentioned by the people that were interviewed. The authors’ efforts were focused more on analyzing the ransomware and discussing their model. It is a well written and deeply sourced resource for ransomware research.

The focus on ransomware behavior for detection is not unique. Security companies such as Sophos, Coveware, and Mandiant have devoted a large portion of their commercial services to offer this capability to their clients. Last year, Sophos director of engineering, Mark Loman (2019), wrote about the characteristics and behavior of 11 families of ransomware (WannaCry, Matrix, GandCrab, SamSam, Dharma, BitPaymer, Ryuk, LockerGoga, MegaCortex, RobbinHood, and Sodinokibi). Coveware publishes blog posts and articles about ransomware trends throughout the year in addition to their posts about victims mentioned later (Coveware, 2020). Mandiant, acquired by FireEye, publishes articles about attackers and malware, but never victims. Their most relevant white paper focused on ransomware protection and containment strategies (McWhirt, 2020). Similar to the experts interviewed by Hull et al. (2019), Mandiant recommended hardening endpoints, implementing better policies, limiting remote users, tightening access control, and reducing the digital footprint (McWhirt, 2020).

Attacker Focused

Hernandez-Castro et al. (2020) and Conti et al. (2018) both published papers on the economic analysis of ransomware. Conti focused more on the malware, while Hernandez-Castro focused more on the attackers. Both papers explained how ransom prices are set by the attacker and are typically paid in bitcoin. Hernandez-Castro et al. (2020) went further and described ransomware in terms of revenue and profitability. They painted a picture that suggests cyber criminals are running a long-term business with strategy more than they are attacking victims for quick gains. While Doniec (2019) and Salvi (2019) offered defensive measures, Hernandez-Castro et al. (2020) offered recommendations on how criminals can increase profits by determining a victim’s willingness to pay and implementing price discrimination, but they did not take data exfiltration into consideration which is a relatively new tactic.

Occasionally, government organizations such as the FBI and CISA (Cybersecurity & Infrastructure Security Agency) publish articles focused on what they call “threat actors". A threat actor is a person deemed to be a threat, acting on behalf of a foreign nation. CISA Alert AA20-049A, released February 2020, alerted critical infrastructure & asset owner operators to threat actors targeting pipeline operations with ransomware. Government articles get straight to the facts unlike academic authors (Arabo et al., Conti et al., & Salvi) who tend to add superfluous commentary. CISA’s alert is basically a list of bullet points that describe the technical details of the attack and the threat actor’s tactics. Like the majority of all ransomware attacks, the attack started with a spearphishing link delivered via email and the attacker used ransomware to attack the victim’s Windows network (CISA, 2020). The alert ended with a list of mitigation factors like Doniec (2019) Salvi (2019), and others which included network segmentation, multi-factor authentication, data backups, email security, updates, patches, antivirus, and limiting remote access to resources (CISA, 2020).

Cartwright and Cartwright wrote about attackers building trust through their reputation to decrypt files if victims pay the ransom (2019). They specifically mentioned Hernandez-Castro’s et al. work on the victim’s willingness to pay and the ransomware family known as “Locky”. Cartwright and Cartwright proposed and proved four research questions to determine the long- and short-term effects of an attacker’s reputation on profitability. They concluded that it is always in the attacker’s interest to return files to victims. It is interesting to read research on attacker’s strategy from a financial point of view because ransomware is guided by their profit more than political views as with hacktivism or nation-state interests as with threat actors. Several other papers have been written about the economics of ransomware because whether we like it or not, this is a billion dollar business. From this perspective, there may be advantages in fighting cyber criminals by treating them like competitors and work to put them out of business. Our proposed model could be used to rate each attacker like how companies collect business intelligence about their competitors before picking a location to open a new store.

Victim Focused

News articles and reports that are focused on the victims tend to come across as marketing more than analysis. Riddel (2020) reported the attack and its effects on Garmin by the cyber criminal group known as Evil Corp and similar attacks affecting EasyJet and Equifax. Recorded Future's Insikt Group (2020) used their internal tools to create a detailed report on victims and whom they refer to as ransomware operators. They specifically mention SUNY Erie Community College, University of Utah, Cygilant, Argentina's immigration agency, K-Electric of Pakistan, Ontario's College of Nursing, Equinix, and Quebec's Department of Justice as victims and criminals using Sodinokibi, Maze, DarkSide, Avaddon, Egregor, MountLocker, Emotet as operators. Coveware (2020), the most prominent ransomware negotiator, also reported on victims and payments with limited data on ransomware families and attack vector frequency. All three of these reports lack analysis of the ransomware code.

The most dangerous type of ransomware attack is the one that targets hospitals, clinics, and healthcare. Human life can, has, and will be impacted. It seems unconscionable, but these attacks happen every single day as far back as 2016, described by Justin Pope in *Innovations in Clinical Neuroscience*. In a unique question/answer format, Pope wrote about the most notable ransomware victims of his time. Hollywood Presbyterian paid $17k in bitcoin, Methodist Hospital was hit by Locky but did not pay the ransom, Medstar Health was attacked, and Kansas Heart Hospital paid an undisclosed ransom amount (Pope, 2016). Like most articles, Pope ends with a list of steps to prevent ransomware such as data backups, training, education, updates, and patches. More recent information about health and medical organizations as victims was provided by the Insikt Group (2020), Coveware (2020), Hull et al. (2019), and the Blueliv Team (2020) showing how the problem is still present and getting worse. This concludes the literature review.

Research Plan

Nearly 100 open source articles per day are written and published about ransomware. Articles with malware analysis directly from the analyst are set aside for processing. On average, about 10 articles per day contain enough information for me to extract for this project. We are currently tracking 200 unique ransomware cases involving 46 different families of ransomware. In each article, look for the name of the ransomware, the root family, the attacker, source URL, origin of the attacker, programming language that the ransomware was written in, encryption types, file extensions, ransom note left for victims, the intended target, the target's sector, initial attack vector, tactics, techniques, and procedures (TTPs), Mitre ATT&CK codes, ransom payment type, ransom amount in dollars, attacker's contact information, indicators of compromise (IoCs). Then we look for the answer to ten yes or no questions to determine the likelihood and impact of a successful attack.

The first five questions are used to determine likelihood: is the attacker a group, is the attack advanced, is the target vulnerable, is the track automated, and does the attack use an unpatched exploit? The next five questions help to determine the impact: is the target in a critical sector, is there a free decryptor, is the ransom expensive, is data stolen, and does the attack cause human death? Adding the answers together and multiplying likelihood times the impact results in a threat score that can be used to mathematically prioritize threats from 0 to 100. The plan is to use this data to understand trends, similarities, differences, and which ransomware families have the highest and lowest threat scores.

Limitations

This project is mainly limited by time because the author is but a single person collecting 16 weeks' worth of data on a highly complicated type of malware. Ideally, data should be collected and analyzed over a longer period to track how ransomware changes from year to year. Every time a new tactic or technique is discovered, ransomware operators are quick to add it to their arsenal and every time defenders learn to detect a new tactic or technique, attackers pivot to new ones. Another limitation is the Traffic Light Protocol (TLP) and access to closed source intelligence. Articles with timely, valuable information are typically labeled TLP:Red which means they are not to be shared outside of trusted members of a closed group. We do not have access to these articles or any other closed source such as law enforcement or classified military intelligence.

Significance

Traditional threat intelligence reports typically focus on the attacker, the victim, or the malware. It is rare to find a threat intelligence report that contains detailed information about all three. We believe this project is significant to the cyber security community because our model combines all three of these subjects with additional sections for likelihood, impact, and risk. Using our unique cyber risk model for ransomware, a major product or outcome is a comprehensive report with threat intelligence on malware, attackers, victims, and mathematical prioritization of risk based on likelihood and impact of a successful attack.

Analysis of this data is useful to organizations defending against ransomware attacks because within the right context, it contains actionable intelligence. A defender can compare our data to their organization, to see which threats are most likely to be a problem so they can focus their effort on these threats rather than ransomware that are unlikely to affect them and/or have a low impact. No sector is safe, and attackers are willing to hit even the most vulnerable target. Cities under attack lose the ability to provide services to citizens, hospitals are not able to care for patients, and banks are unable to process transactions.

For the same reason as defenders, decision makers can use this project's threat model to help allocate their security budget to invest in people, processes, and technology to prevent or detect ransomware attacks that are most likely to occur and would have the most damaging impact. This saves organization money, time, and effort because they avoid spending resources on threats that are unlikely to occur or that would have a low impact on their organization. If done correctly, this new threat model will be used by the Department of Homeland Security to protect our nation’s critical sectors from ransomware attacks, which in turn could save American lives.

Problem, Purpose, and Goal

Ransomware is increasing in the number of successful attacks and size of impact. Entire cities were attacked in the last two years causing government services to go offline (Peters, 2019) and several hospitals have been attacked, causing people to die prematurely (O’Neill, 2020). The problem is getting worse because so many victims pay the ransom, making this a very profitable business model. Insurance companies pay ransoms too because in most cases, it is cheaper than recovery. Making matters worse, cyber criminals gangs often team up with other gangs, offer affiliate marketing, promote ransomware as a service, and continuously update their tactics to evade prevention and put pressure on victims to pay. So far, a solution has not been identified.

The purpose of this project is to understand ransomware and come up with a solution. We discuss what the community is currently doing with regards to ransomware intelligence and what is missing, so that our solution is more robust. We want to put a dent in the ransomware business by using our model to help organizations prioritize threats, which enables them to reallocate their security budget and focus on threats with the highest risk. This is more than simply throwing money at the problem. Making attacks more expensive increases the cost of success, which makes this business less profitable. The end goal is for the community to use our model and reduce ransomware profits so much that cyber criminals go out of business. We can measure our success based on how many organizations use the model compared to how many organizations are attacked and pay ransoms.

Definitions

Attacker: person or group that is responsible for malicious cyber attacks.

Cyber: an artificial domain of warfare consisting of computers, networks, data, and all related devices.

Decryptor: software used to convert encrypted files or data back into a readable format.

Encrypted: files or data that are converted into an unreadable format.

Exploit: (noun) a tactic, technique, or code used to attack a vulnerable cyber target.

Exploit: (verb) the act of attacking a vulnerable cyber target.

MITRE ATT&CK: a framework created by a not-for-profit organization using federal funding to identify and categorize cyber exploits.

Patch: code that fixes a vulnerability.

Ransom: money in the form of bitcoin requested by attackers to decrypt a target’s encrypted files or data.

Ransomware: Malicious software used by attackers to encrypt or otherwise hold files or data hostage until a ransom is paid by the target.

Target: an organization that attackers intended to exploit with cyber weapons such ransomware, also known as the victim.

Threat Intelligence: actionable and timely information with context about threats that can be used to prevent or detect vulnerabilities, exploits, and attackers.

Vulnerability: a flaw in their code or procedures that might allow attackers to exploit a target

Limitations & Assumptions

Threat intelligence reports are typically released to the public by security vendors for marketing purposes. There are many private and classified sources, but we did not have access to them for this project. Our research is limited to data that was reported publicly. We also assume their intelligence is accurate and based on facts. Since we did not have access to the original data, we were not able to validate their findings independently. We must trust security vendors to report the truth. In cases where we do have access to original data, we must assume the data has not been altered because we do not have access to the hardware it was copied from and unable to verify the complete chain of custody. This is one reason why attribution of an attack to a particular person or group is so difficult.

Theoretical Framework

Our framework is based on the Department of Homeland Security's Risk Management Fundamentals and Doctrine (Beers, 2011, p. 19-21). By combining qualitative risk analysis and quantitative data analysis, we can calculate the likelihood and impact of each individual ransomware attack. By analyzing multiple attacks over a period, we are able perform quantitative data analysis to determine statistics, trends, outliers, and make predictions.

Research Design

This project is a combination of qualitative risk analysis and quantitative data analysis. Guided by the Department of Homeland Security Risk Management Fundamentals and Doctrine, qualitative risk analysis of ransomware is achieved by calculating the likelihood and impact of each individual attack (Beers, 2011, p. 19-21). By analyzing multiple attacks over a period of time, it is then possible to perform quantitative data analysis to determine any number of differences and trends.

Our research questions are divided into two types: informational and significant. Informational questions typically do not change the outcome of our qualitative analysis, but they are used in our quantitative analysis. They are as follows: What is the date of the attack? What is the name of the ransomware? What family does this ransomware come from? Who is the attacker? What is our source? Who is the target? What sector is the target in? What was the initial attack vector? What are the attackers tactics, techniques, and procedures? What are the associated MITRE ATT&CK codes? What is the attacker's suspected country of origin? What language is the ransomware written in? What type of encryption was used? What was the file extension? What was the ransom note? What was the payment type? What was the payment amount in dollars? What is the attacker's contact information? What are the indicators of compromise?

The next set of questions are significant because they affect our qualitative analysis used to calculate the likelihood and impact of a successful ransomware attack and our quantitative analysis used to determine differences and trends. They are simply binary which reduces mistakes and decreases the amount of time it takes to find the answer. They are as follows: Are the attackers a group of people? Is the attack advanced? Is the target vulnerable? Is the attack automated? Does the attack exploit an unpatched vulnerability? Is the target in a critical sector? Is there a decryptor? Is the ransom expensive? Did the attack kill anyone? Did the attackers also steal data?

Finding answers to these questions for each attack is time consuming because intelligence reports typically focus on one aspect such as the ransomware or the victim. Researchers must fill in informational gaps by seeking additional sources. A web tool called INOREADER is the best way we found to process many reports. This tool allowed us to add hundreds of sources to display thousands of reports in a standardized and easy to read format. Every day for the past 15 weeks, we used INOREADER to look for new reports of ransomware attacks. Middle clicking the report headline brings up the original source in a new tab which enabled us to process more than 200 intelligence reports in such a short amount of time. This process could easily be continued daily to maintain a ransomware threat database.

All of our sources for ransomware intelligence reports are open to the public including individual blogs, corporate sponsored analysis, news articles, and government reporting. News articles are the most frequent and typically contain the least amount of usable data. Corporate analysis gives away enough to entice users to pay for extra content. Government reports are timely, trusted, and contain actionable intelligence, but they do not reveal their sources for national security reasons. Individual blogs contain the most information but may not always be timely or trusted. Therefore it is important to combine information from multiple reports to answer as many questions as possible.

Research questions

Here are the ten significant research questions in null hypothesis form:

- Single attackers do not increase likelihood.
- Simple attacks do not increase likelihood.
- Targets that are not vulnerable do not increase likelihood.
- Manual attacks do not increase likelihood.
- Exploiting a patched vulnerability does not increase likelihood.
- Attacking a non-critical target does not increase impact.
- Having a free decryptor does not increase impact.
- Attacks with ransoms under $1m do not increase impact.
- Attacks that do not harm human life do not increase impact.
- Attacks that do not steal data do not increase impact.

Process

The following are all the steps necessary to reproduce these data from start to finish as a daily process for cyber threat intelligence analysts and researchers:

1) Create an account on https://inoreader.com

2) Subscribe to this bundle https://www.inoreader.com/bundle/0014cd640df8

3) Search for articles with the keyword “ransomware”

4) Open each article in a new window

5) Create a new Spreadsheet with Google Drive or Microsoft Office

6) Add the following columns:

    a) DATE

    b) NAME

    c) FAMILY

    d) ATTACKER

    e) SOURCE

    f) TARGET

    g) SECTOR

    h) INITIAL VECTOR

    i) TTPs

    j) ATT&CK CODE

    k) ORIGIN

    l) LANG

    m) ENCRYPTION

    n) EXTENSIONS

    o) NOTE

    p) PAYMENT TYPE

    q) DOLLAR AMOUNT

    r) CONTACT INFO

    s) IOCS
    t) GROUP or RaaS?
    u) ADVANCED?
    v) VULNERABLE?
    w) AUTOMATED?
    x) UNPATCHED?
    y) CRITICAL SECTOR?
    z) NO DECRYPTOR?
    aa) EXPENSIVE?
    bb) DEATHS?
    cc) EXFILTRATION?
    dd) LIKELIHOOD
    ee) IMPACT
    ff) RISK

7) Columns t:cc are yes/no questions. Input 1 for yes, 0 for no.

8) Columns dd:ff are calculations:

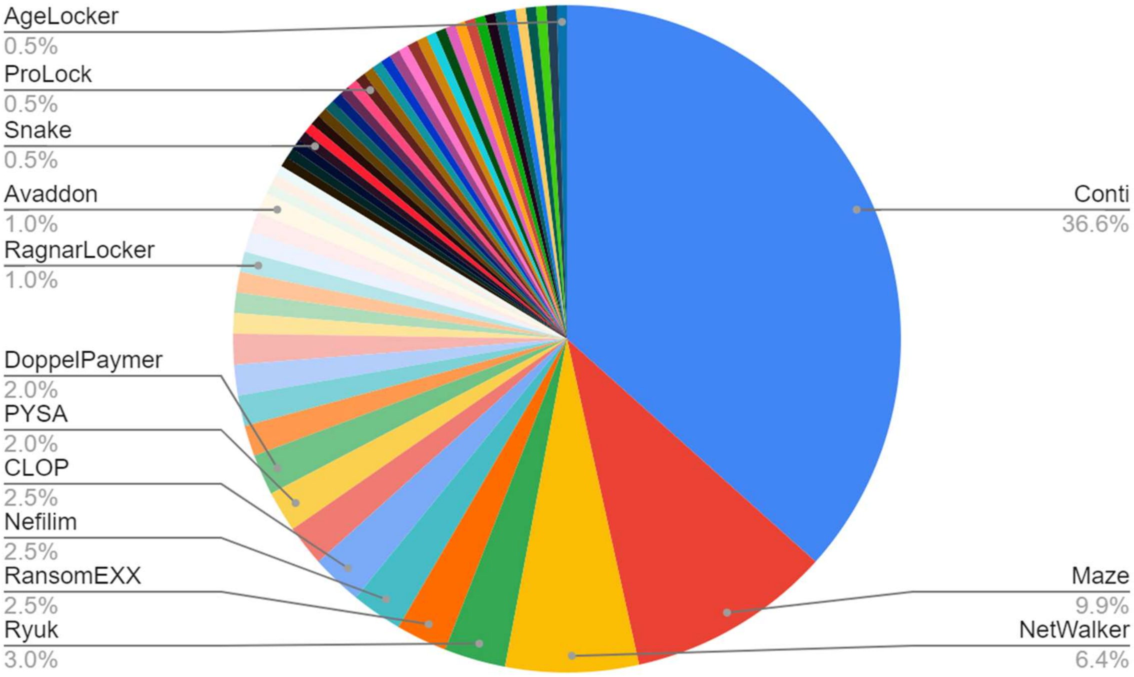
    a) LIKELIHOOD =SUM(T#:X#)
    b) IMPACT =SUM(Y#:AC#)
    c) RISK =MULTIPLY(AD#,AE#)

9) For each article in step 4, read entire article and input relevant data into spreadsheet

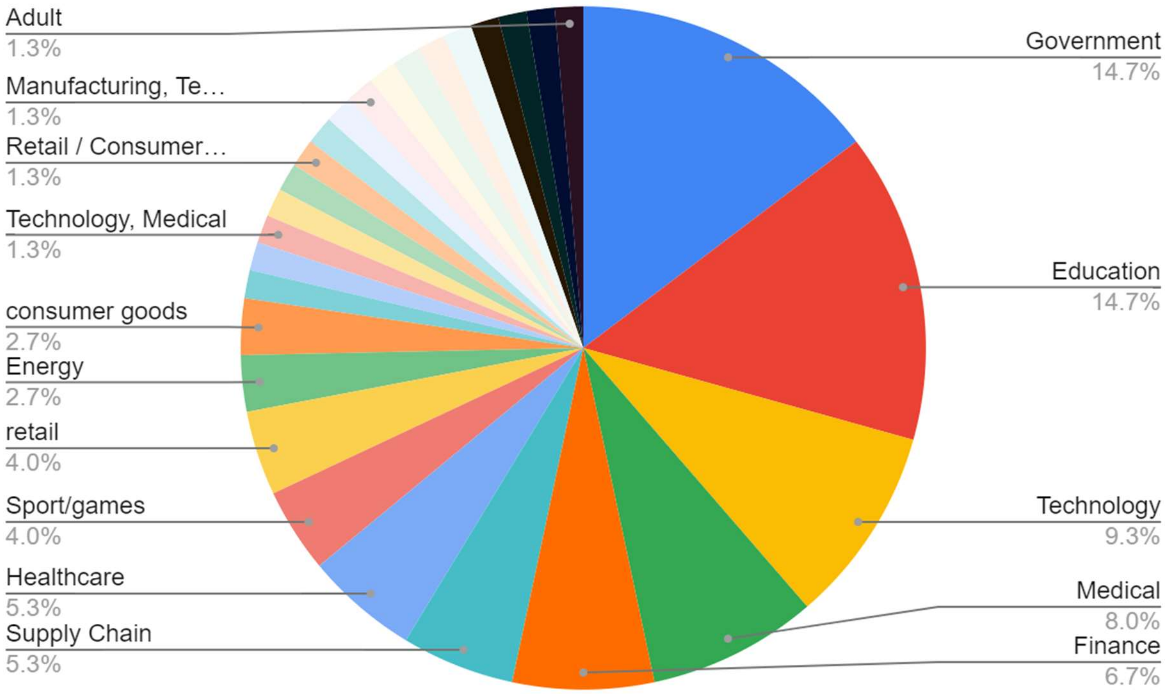
10) Use pivot tables to compile relevant charts and trends

Data Analysis

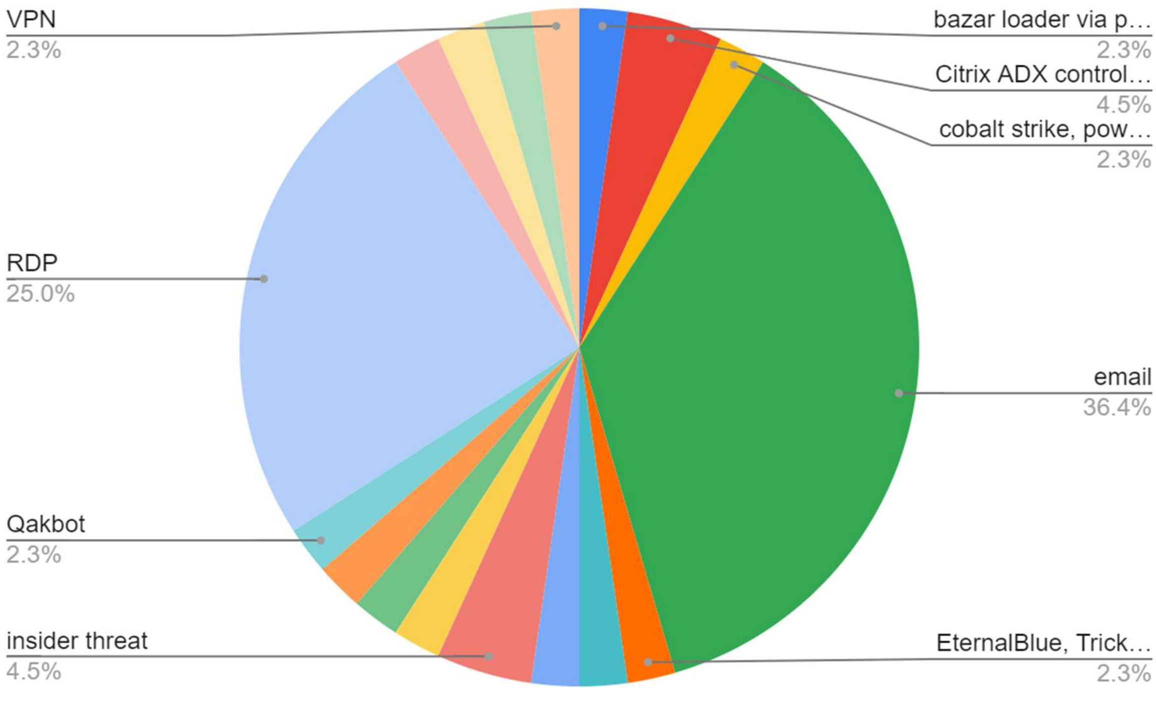
From August to November, we collected data on 57 different families of ransomware. The top five most frequently reported families were: Conti, Maze, NetWalker, Ryuk, and Clop. Be aware, these are the most reported groups we found in our INOREADER bundle of sources. Different bundles will have different results.



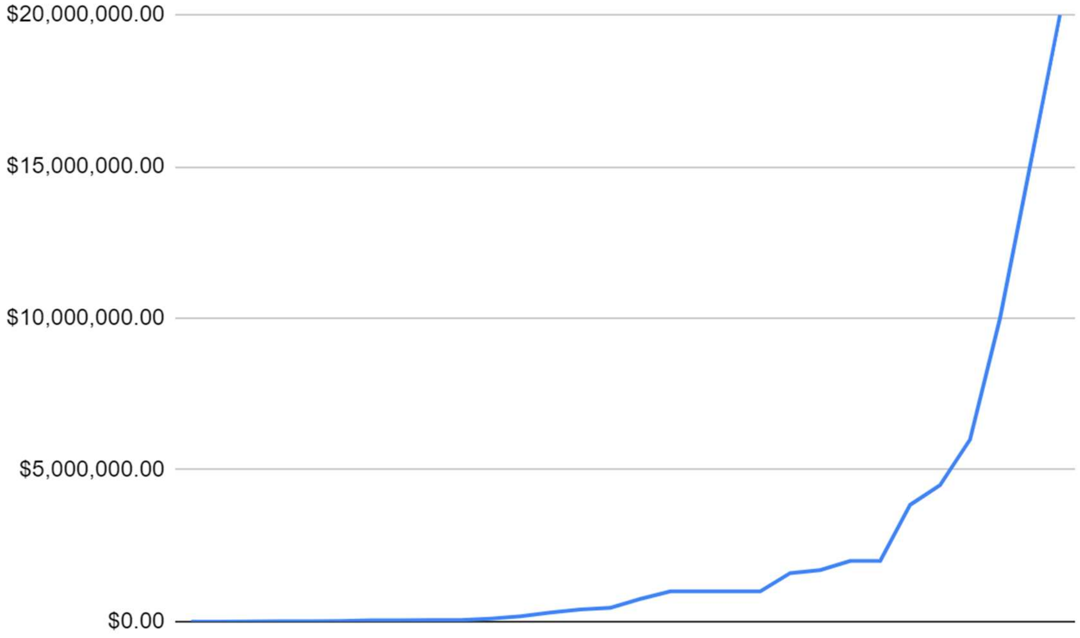
Using all available data from open source intelligence reports, we extracted each target's sector. The top five most frequently target sectors were government, education, technology, medical, and finance. The medical industry is interesting because there is an ongoing pandemic and some ransomware groups have stated they will not target hospitals or refund the ransom if they attack one by mistake.



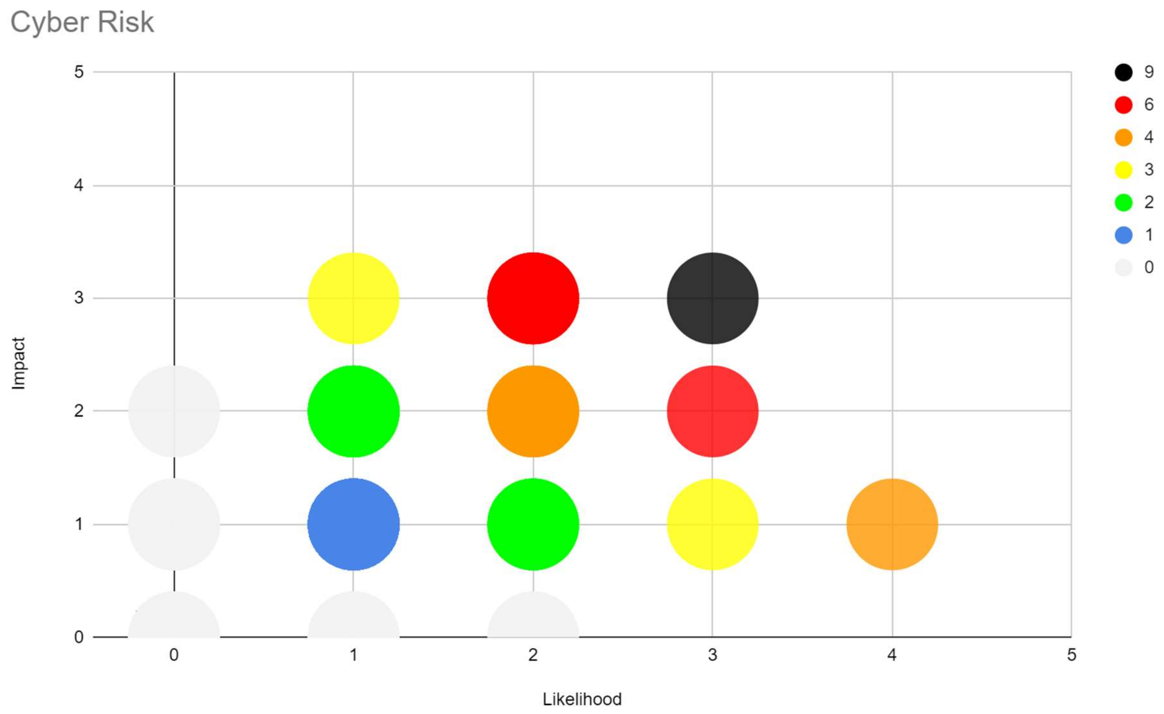
Looking at initial attack vectors, we see that email and RDP are the two most common ways attackers can infect victims with ransomware. This makes sense because attackers typically target the weakest link of the system which are humans, also known as social engineering. Attackers can trick people to open phishing emails, then steal valid credentials for RDP.



By far the most common payment type was bitcoin while the ransom amounts ranged anywhere from $300 to $20 million. Keep in mind these data are not complete universally as many victims pay in private. Our data is from public sources.



For each report, we tried to find the answer to our ten binary (yes or no) research questions, divided. For each yes answer, the data point was marked 1, while no answers were marked 0. Simple addition of the first five answers allows us to calculate total likelihood, while adding the answers to the second set of questions returns the total impact. Multiplying these two values together returns the total risk score for each row of data. Our data ranges from a scale of 0 to 9 and is color coded by order of severity: grey, blue, green, yellow, orange, red, black. The following bubble chart displays these data visually.



Conclusion

The results of this research project indicated the ransomware family known as “Clop” was the highest risk to most organizations, the most common initial attack vector was email, the average ransom amount was $2 million paid in bitcoin, and the most frequently reported target sector was a tie between government and education. It is possible and reasonably simple to process and analyze cyber threat intelligence reports using this new risk model. However, it is labor intensive for an analyst to read and extract all relevant data and to find the most correct answer to each research question. Future work could automate the process by using natural language processing to parse reports.

Resources

Arabo, A., Dijoux, R., Poulain, T., Chevalier, G. (2020). Detecting Ransomware Using Process Behavior Analysis. *Procedia Computer Science, 168,* 289-296.

Beers, R. (2011). *Risk Management Fundamentals.* DHS Office of Risk Management and Analysis. Retrieved online from https://www.dhs.gov/xlibrary/assets/rma-risk-management-fundamentals.pdf

Blueliv Team (2020). *Threat intelligence vs. the rise in sophisticated ransomware.* Leap in Value. Retrieved online from https://www.blueliv.com/cyber-security-and-cyber-threat-intelligence-blog-blueliv/threat-intelligence-vs-the-rise-in-sophisticated-ransomware/

Bourne, V. (2020). *The State of Ransomware 2020.* Sophos. Retrieved online from https://news.sophos.com/en-us/2020/05/12/the-state-of-ransomware-2020/

Cartwright, A., Cartwright, E. (2019). Ransomware and Reputation. *Games 10* (26), 1-14.

Chen, T., Robert, J. (2004). *The Evolution of Viruses and Worms.* ResearchGate. Retrieved online https://www.researchgate.net/publication/228869267_The_Evolution_of_Viruses_and_Worms

CISA (2020). *Ransomware Impacting Pipeline Operations.* Department of Homeland Security. Retrieved online from https://us-cert.cisa.gov/ncas/alerts/aa20-049a

Collier, R. (2017). NHS ransomware attack spreads worldwide. Canadian Medical Association Journal (CMAJ), 189(22), E786–E787. https://doi.org/10.1503/cmaj.1095434

Conti, M., Gangwal, A., Ruj, S. (2018). *On the Economic Significance of Ransomware Campaigns: A Bitcoin Transactions Perspective*. University of Padua. Retrieved online from https://arxiv.org/pdf/1804.01341.pdf

Coveware (2020). *Ransomware Demands continue to rise as Data Exfiltration becomes common, and Maze subdues.* Coveware. Retrieved online from https://www.coveware.com/blog/q3-2020-ransomware-marketplace-report

Doniec, A. (2019). *A deep dive into Phobos ransomware.* Malwarebytes Labs. Retrieved online https://blog.malwarebytes.com/threat-analysis/2019/07/a-deep-dive-into-phobos-ransomware/

Dragos (2020). *EKANS Ransomware and ICS Operations.* Dragos Inc. Retrieved online from https://www.dragos.com/blog/industry-news/ekans-ransomware-and-ics-operations/

Falcone, R. (2020). *Thanos Ransomware: Destructive Variant Targeting State-Run Organizations in the Middle East and North Africa.* Palo Alto Networks. Retrieved online from https://unit42.paloaltonetworks.com/thanos-ransomware/

Hernandez-Castro, J., Cartwright, A., Cartwright, E. (2020). An economic analysis of ransomware and its welfare consequences. *Royal Society Open Science 7*(3).

Hull, G., John, H., Arief, B. (2019). Ransomware deployment methods and analysis: views from a predictive model and human responses. *Crime Science 8* (2), 1-22.

Insikt Group (2020). *Cyber Threat Analysis: Q3 Malware Trends.* Recorded Future. Retrieved online from https://go.recordedfuture.com/hubfs/reports/cta-2020-1105.pdf

Kennelly, J., Goody, K., Shilko, J. (2020). *Navigating the MAZE: Tactics, Techniques and Procedures Associated With MAZE Ransomware Incidents.* FireEye. Retrieved online https://www.fireeye.com/blog/threat-research/2020/05/tactics-techniques-procedures-associated-with-maze-ransomware-incidents.html

Loman, M. (2019). *How Ransomware Attacks.* Sophos Labs. Retrieved online from https://www.sophos.com/en-us/medialibrary/PDFs/technical-papers/sophoslabs-ransomware-behavior-report.pdf

McWhirt, M. (2019). *Ransomware Protection and Containment Strategies: Practical Guidance for Endpoint Protection, Hardening, and Containment.* FireEye. Retrieved online from https://www.fireeye.com/blog/threat-research/2019/09/ransomware-protection-and-containment-strategies.html

Meltzer, T., Phillips, S. (2009). *From the first email to the first YouTube video: a definitive internet history.* The Guardian. Retrieved online from https://www.theguardian.com/technology/2009/oct/23/internet-history

O'Neill, P. (2020). *A patient has died after ransomware hackers hit a German hospital.* MIT Technology Review. Retrieved online from https://www.technologyreview.com/2020/09/18/1008582/a-patient-has-died-after-ransomware-hackers-hit-a-german-hospital/

Peters, J. (2019). *22 Texas towns hit by coordinated ransomware attack.* The Verge. Retrieved online from https://www.theverge.com/2019/8/20/20823139/texas-towns-ransomware-attack-cities-fbi-threat-computers-offline

Pope, J. (2016). Ransomware: Minimizing the Risks. *Innovations in Clinical Neuroscience 13*(11-12), 37-40.

Salvi, H. (2019). RAAS: Ransomware-as-a-Service. *International Journal of Computer Sciences and Engineering, 7*(6), 586-590.

Young, A., Yung, M. (1996). *Cryptovirology: extortion-based security threats and countermeasures.* Proceedings 1996 IEEE Symposium on Security and Privacy. Retrieved online from https://ieeexplore.ieee.org/document/502676