PowerShell Malware Analysis Using a Novel Malware Rating System

David Arnold, Charlotte David, and Jafar Saniie

Embedded Computing and Signal Processing (ECASP) Research Laboratory (http://ecasp.ece.iit.edu)
Department of Electrical and Computer Engineering
Illinois Institute of Technology, Chicago IL, U.S.A.

***Abstract* - Recent high-profile cyberattacks highlight an increased use of social engineering attacks and ransomware by hackers worldwide. These attacks target human operators directly, bypassing many of the cyber-safeguards developed through years of malware analysis. In response to these challenges, many organizations have turned to white-hat hackers and penetration testing to identify potential weaknesses and reinforce cyber-safety protocols. To assist in the threat evaluation process, malware rating systems are often used to highlight the danger and potential damage malware may cause. Current malware rating systems focus on assigning a danger score for malware based on its ability to move throughout the network, damage system resources, and evade detection. Due to the increased reliance on social engineering, a new malware rating system is proposed that incorporates malware deceitfulness as a means to trick human operators. The novel rating system will score malware based on its Stealth, Ease of Creation, Deceitfulness, Versatility, Instantaneousness, and Persistence. This system provides operators with insight into each key characteristics as opposed to a single value. To showcase the malware evaluation process, different PowerShell Reverse Bind Shell malwares are rated based on the proposed criteria.**

I. INTRODUCTION

Malware analysis and evaluation is an important function of cyber defense and penetration testing, providing insight into network vulnerabilities. To assist in this process, malware rating systems are used to identify software that pose the greatest risk to network integrity. Under the current system, a danger score is derived from the malware's ability to propagate through the network and enact damage against network resources [1]. However, as threats continue to evolve to meet today's advanced security architectures the application of a single overall metric can omit malware that may be specialized in its design. As an alternative, a model that presents data on a wider number of characteristics allows operators to observe deficiencies in their security architecture that specialized malware may be able to penetrate. A novel malware rating system is presented that provides a score related to six primary malware characteristics including Stealth, Ease of Creation, Deceitfulness, Versatility, Instantaneousness, and Persistence.

Initially, malware began popping up prior to the 1990s and started out as an exploration of loop-holes present in the MS-DOS systems and often resulted in temporary system crashes or increased consumption of system resources [2-4]. In the 1990s, malware began to evolve into more malicious purposes, often targeting the Windows Operating System through simple mail and macro worms. This era also saw the first anti-virus software. By the early 2000s, the widespread adoption of the Internet launched a golden age of worms and viruses. New users would be easily tricked through email attachments, free downloads, and open network sharing. As cybersecurity awareness grew, malware needed to shift away from simple deployment methods and saw the introduction of rootkits and ransomware. This saw the beginning of early anti-virus evasion protocols and lateral network propagation. Enhanced cybersecurity awareness among network administrators has greatly decreased the number of potential attack vectors for low-skilled hackers. As a result, human operators have become the weakest link in organizational cybersecurity profiles. Social Engineering attacks circumvent cyber-defense protocols by relying on human error to download malicious payloads known as malware. In most cases, ransomware is deployed to encrypt the target's hard drive and demand payment for system recovery. According to Verizon, approximately 10% of all breaches were categorized as ransomware, with a median loss of $11,150 [5]. Further, the FBI's Internet Crime Complaint Center (IC3) received 2,084 ransomware complaints through the first half of 2021, with a predicted loss of $16.8 Million [6]. Compared to the same time from 2020, there was a 62% increase in reporting and a 20% increase in the reported loss over the same time frame from 2020. In May 2021, ransomware successfully activated within the networks of Colonial Pipeline, halting operations and crippling the oil supply chains for several days along the United States East Coast [7-8]. This incident highlights how even a simple social engineering attack can disrupt operations without requiring in-depth knowledge of network and operating system security.

As part of the recovery and post-attack analysis process, the Cyber Kill Chain was developed by Lockheed Martin to analyze successful cyberattacks and prepare for future intrusions [9-11]. The Cyber Kill Chain is composed of seven stages and includes Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control, and Actions on Objectives. Upon discovering a breach, cyber defense teams will create the Cyber Kill Chain of the attack in order to better analyze the hacker's techniques and objectives, providing important feedback and knowledge to prevent future attacks. During the Reconnaissance stage, the attacker scans the

network and researches the target to identify potential vulnerabilities and high-priority targets. Next, the weaponization stage will involve the creation of malware packages for exploiting backdoors, propagating through the network, and establishing persistence. Weaponized software is then delivered to the target via email, USB, or any other entrance vector. At this stage, vulnerabilities are exploited to allow malicious software to be installed on the network and provide access to critical services and resources. The malware will then install itself within the target and begin to search for other victim computers, establish persistence, and provide access to the host machine. When remote access is provided to the hacker, they will attempt to gain command and control over primary systems. Finally, the attacker will begin to complete their objectives, which may include espionage or resource destruction. An example of a successful Cyber Kill Chain that exploits the Kerberos Authentication protocol can be seen in Figure 1. Since the novel malware rating system includes the ability to successfully deliver malware to the target system (Stealth and Deceitfulness) and to establish full command and control (Persistence) it considers many of the components of the Cyber Kill Chain and is useful for post-attack scenarios.

Through this paper, the novel malware rating system will be presented. First, a description of each metric will be provided, these metrics include Stealth, Ease of Creation, Deceitfulness, Versatility, Instantaneousness, and Persistence. Next, a simple PowerShell Reverse Bind Shell and its variations will be presented for an example application of the rating system. Finally, the novel malware rating system will be applied to the PowerShell Malwares and the resulting ratings will be discussed.

II. NOVEL MALWARE RATING SYSTEM

The proposed malware rating system is composed of six primary characteristics that are important to the functionality and successful execution of the observed malware. Each characteristic is evaluated and rated on a 3-point scale, with a value of 3 denoting high proficiency in the characteristic and value of 1 indicating weak proficiency. Table 1 describes the six characteristics, namely Stealth, Ease of Creation, Deceitfulness, Versatility, Instantaneousness, and Persistence. Stealth refers to the malware's ability to evade standard anti-virus (AV) software. Ease of creation refers to whether the adversary needs to have in-depth knowledge in order to craft the malware or to utilize tools for generating the malicious code. Deceitful malware is proficient at tricking employees and operators into running the executable, thus launching the attack. Versatility measures whether the malware can be easily modified to target different subsystems, this can include the ability to run across multiple Operating Systems. Malware capable of providing necessary system access or cause the intended damage without operator intervention is denoted as having high instantaneousness. Finally, Persistence entails the malware's ability to automatically run upon system restart or its ability to obtain further footholds within the network. Each of these criteria will now be explained in further detail along with the requirements for each rating level.

TABLE 1
NOVEL MALWARE RATING SYSTEM CHARACTERISTICS

Criterion	Description
Stealth	Is the malware detected by standard Anti-Virus (AV) software?
Ease of Creation	Does the malware require intimate knowledge of toolkits or target characteristics for creation?
Deceitfulness	How likely is the malware to be run by an employee/operator?
Versatility	Can the malware be modified to target different subsystems or Operating Systems?
Instantaneousness	Will the malware provide necessary system access/cause intended damage without operator intervention?
Persistence	Upon startup, does the malware automatically create a foothold for future restarts?

A. *Stealth*

The Stealth characteristic refers to how well the malware avoids detection by standard anti-virus (AV) software commonly deployed to identify malicious executables. To determine this metric, the suspected malware is uploaded to virustotal.com and will be run through a battery of commonly used AV tools [12]. As an added benefit, virustotal.com provides the accumulated malware to the developers of the anti-virus software. After receiving the results from the toolkit, the malware will be assigned a rating based on how well it remained undetected. Malware that is detected by less than 30% of AV software is awarded a Stealth rating of 3, while

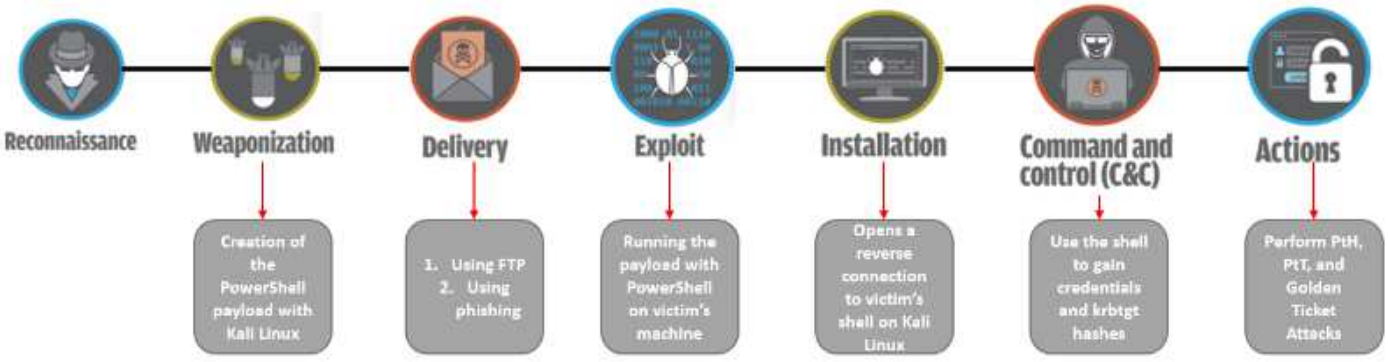

Fig. 1. Cyber Kill Chain for the successful acquisition of Kerberos hashes to compromise the Kerberos authentication mechanisms.

malware that is detected by between 30% and 60% is awarded a rating of 2. Finally, malware that is detected by over 60% of AV software is given a Stealth rating of 1, which is the lowest within the novel rating system. Generally, zero-day exploits and other novel malware will score the highest on this metric. However, certain encryption methods and reconfigurable code may be a new threat to the viability of AV software. The rating requirements for the Stealth metric are displayed in Table 2.

TABLE 2
STEALTH METRIC RATING REQUIREMENTS

Stealth	
Rating	Requirement
3	Detected by <30% of Anti-Virus Software.
2	Detected by between 30% and 60% of Anti-Virus Software.
1	Detected by >60% of Anti-Virus Software.

B. Ease of Creation

Due to the commercialization of malware commercialization, many tools now exist to greatly simplify the knowledge required to successfully generate functioning malware. Tools, such as the Msfvenom tools, allow hackers to generate shellcode payloads that utilize known exploits and vulnerabilities [13]. As a result, the expertise required to launch attacks has been greatly reduced. The Ease of Creation metric allows the proposed rating system to account for the range in expertise among the hacking population. To achieve the highest rating of 3, the malware requires little knowledge of the underlying tools or exploits and will be generated using very few lines of code. Attacks launched using malware with an Ease of Creation rating of 3 are most likely to compromise systems that have few safeguards in place or be launched by a script kiddie. Next, a value of 2 indicates that the malware needed to be modified during the creation process. For instance, changing the file type to adjust for file share requirements would require a deeper understanding of the target. Finally, the lowest rating refers to malware that is manually generated or requires extensive knowledge of the target application. Zero-day exploits will most commonly fall within this category as they require extensive research and development for successful implementation. A summary of the rating requirements for the Ease of Creation metric are provided in Table 3.

TABLE 3
EASE OF CREATION METRIC RATING REQUIREMENT

Ease of Creation	
Rating	Requirement
3	Required information and toolkits are widely available. Creation can be completed within a few lines of code.
2	Instructions are available for creation but require an understanding of the target system and the toolkit for creation.
1	Little information is available regarding the malware. Creation requires deep understanding of the target system.

C. Deceitfulness

Most malware attacks, especially ransomware, operate by tricking the user into executing a malicious payload. Tactics for tricking users has evolved over the years to include malicious email attachments, compromised USB drives, and even hidden within document Macros. The Deceitful characteristic represents how well the malware executes these tactics in order for the exploit to be activated. Within the novel rating system, malware that can trick even the most cyber-aware employees will receive a rating of 3. Attacks that score highly on this metric will hide behind seemingly benign applications. Attacks such as supply chain and watering-hole attacks are examples of highly scoring malware as they will be downloaded from a trusted source. Average techniques, such as phishing emails, will be identified by informed employees and missed by uninformed employees that may not be familiar with cybersecurity practices. Finally, the lowest scoring malware will often require action on behalf of the hacker. This can include uploading the malware after bypassing weak-authentication on an FTP server. Since the malware will most likely not be seen by the employee, it will require additional exploits for execution. The rating requirements for the Deceitfulness metric are displayed in Table 4.

TABLE 4
DECEITFULNESS METRIC RATING REQUIREMENTS

Deceitfulness	
Rating	Requirement
3	Even the most diligent employees will run this malware.
2	Informed employees will correctly identify the software as malware while uninformed employees may still run the malware.
1	Few or no employees will run this malware.

D. Versatility

Next, the Versatility metric refers to how well the malware can be modified for other purposes. For malware to receive the highest rating, the malware should be easily reconfigurable or even capable of attacking multiple different platforms without modification. For example, modular malware that can interchange infiltration techniques and other payload components would achieve a high score in Versatility. A score of 2 will represent a malware that can be adjusted, but may require extensive retooling to achieve the desired effects. The lowest score is awarded to malware that is very target-specific, especially when the target's architecture has few similarities to commonly used operating systems. Malware designed for achieving precise control over specific targets will normally achieve the lowest score of 1, especially when the target's architecture has few similarities to commonly used operating systems. A summary of the rating requirements for the Versatility metric are provided in Table 5.

TABLE 5
VERSATILITY METRIC RATING REQUIREMENTS

Versatility	
Rating	Requirement
3	The malware can easily be modified to attack new targets or caused other intended damages.
2	Some modification is required for the malware to target other systems.
1	It is impossible, or extensive modifications are required, for the malware to be used for other purposes.

E. Instantaneousness

Instantaneous malware is capable of achieving the attacker's objectives with limited input, such as through pre-written scrips included in the payload. Malware can score a rating of 3 by achieving command and control (C2) or causing

the intended damage without needing the attacker to intervene. Attacks intended to circumvent physical network segmentation, such as the Stuxnet virus, will score highly as they are capable of identifying the target, deploying according to a given deadline, and damage the target without any outside interaction. A moderate score of 2 will be awarded to malware that does automated reconnaissance or retrieves limited credentials upon activation. The lowest rating of 1 is allocated to malware that achieves limited objectives for the hacker. For instance, simple remote shells require extensive action after activation in order to achieve C2 or cause damage. The rating requirements for the Instantaneous metric are displayed in Table 6.

TABLE 6
INSTANTANEOUSNESS METRIC RATING REQUIREMENTS

Instantaneousness	
Rating	Requirement
3	Root access or intended damage is caused without any operator intervention.
2	Limited operator intervention is required to escalate permissions or cause damage.
1	Several steps or procedures are required to obtain the desired access.

F. Persistence

The final metric for the novel malware rating system is Persistence, which measures whether the malware is capable of establishing a foothold within the system. To achieve a rating of 3, the malware should automatically search for mechanisms for surviving a system restart and re-establish communication with the attacker. For instance, the malware could attempt to create a new system service or inject itself into an existing service that is set to start upon restart. A lower rating of 2 will require the attacker to make modifications to the host and may already include some pre-written code within the payload to assist in this activity. The lowest score will be reserved for malware that is not suitable for persistent execution. Attacks intended to commit damage will often not require persistence, and simple programs for establishing a remote shell will be included in this lowest classification. The rating requirements for the Persistence metric are displayed in Table 7.

TABLE 7
PERSISTENCE METRIC RATING REQUIREMENTS

Persistence	
Rating	Requirement
3	The malware automatically establishes a foothold for startup upon system restart.
2	Some steps are required by the operator are required to ensure access after system restart.
1	Extensive activity is required for the malware to remain active after system restart.

III. POWERSHELL MALWARE

To test the proposed malware rating system, a set of PowerShell Reverse Bind Shell Malware was generated and evaluated. The purpose of these reverse bind shells is to allow an intruder to create a backdoor shell connection to the victim's computer. Upon execution, the payload will attempt to establish a connection with a port listener on the attacker's host machine. If the payload is successful, the attacker is rewarded with a command shell that can then be used to further explore the target machine to achieve command and control or to inflict damage on the target. A common reverse bind shell scenario is portrayed in Figure 2, with the attacker creating a port listener on port 5555 which the victim's computer will attempt to establish a connection to. During this exercise the victim's machine is a Windows 2019 server and the attacker will attempt to gain a PowerShell shell. Access to PowerShell command line interface will grant the attacker with access to additional features and higher privileges, potentially including access to valuable Kerberos ticket hashes.

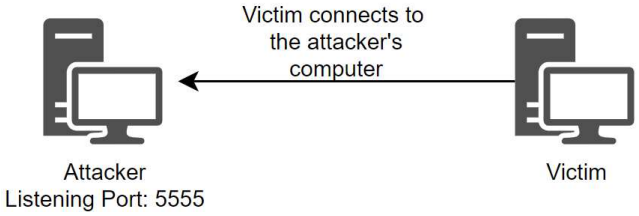

Fig. 2. Remote Bind Shell Attack

Three implementations of the reverse bind shell will be examined using the novel malware rating system. Additionally, it is assumed that they will be transmitted to the victim machine via a phishing attack. The first malware is a simple executable generated by msfvenom, a popular component of the Metasploit toolkit. Msfvenom simplifies the malware creation process by generating, encoding, and outputting the desired payload based on a library of known exploits and shellcodes. For this exercise, the windows/shell/reverse_tcp payload was selected and encoded using the x86/shikata_ga_nai encoder, providing a layer of obfuscation to prevent detection by the AV tools. Upon execution, the malware will attempt to connect back to the attacker's machine, which will be handled by the Metasploit exploit handler. The next malware was the Invoke-Shellcode.ps1 shellcode written by Matt Graeber [14-15]. This script allows the attacker to inject a reverse bind shell into the already running PowerShell terminal, bypassing many AV systems by residing within memory rather than the disk. However, transferring the shellcode to the target and tricking the user to running the script can be difficult as it requires a PowerShell terminal for maximum effectiveness. The final malware is composed of a malicious Macro hidden within a Microsoft Office document, based on the implementation of Matt Nelson [16]. After the document is opened, the payload is executed when the user accepts the Macros that are present. Since accepting the document macros is often required when downloading Microsoft Office documents, employees naturally accept them whenever they download and open a new document. As a result, this malware easily bypasses employee detection while also successfully deploying its payload.

TABLE 8
NOVEL MALWARE RATING SCORES FOR THREE REVERSE BIND SHELL MALWARES

	Stealth	Ease of Creation	Deceitfulness	Versatility	Instantaneousness	Persistence
Windows/shell/reverse_tcp	53/74 = 71% => 1	3	2	1	3	1
Invoke-Shellcode.ps1	30/74 = 41% => 2	2	1	3	1	2
Macro Payload	0/74 = 0% => 3	1	3	3	2	2

IV. APPLICATION OF THE NOVEL RATING SYSTEM

After generation, each malware was then rated according to the proposed malware rating system. Table 8 includes the assigned rating for the three malwares and Figure 3 presents these ratings visually. Each malware and their ratings will now be discussed.

First, the windows/shell/reverse_tcp malware was analyzed and rated. The simple executable was identified by 71% of the anti-virus software on virustotal.com, equating to the lowest score of 1 for the Stealth rating. This was expected as msfvenom is a popular tool and the payload is well known. Since the windows/shell/reverse_tcp shellcode was generated using a single line within the msfvenom tool, it is awarded the maximum rating of 3 for the Ease of Creation characteristics. For the Deceitfulness metric, a moderate score of 2 was awarded, however this relies heavily on the executable name provided by the attacker and the strength of the phishing email. Upon execution, the shellcode will immediately award a PowerShell interface to the attacker, awarding the malware with an Instantaneous rating of 3. The simplicity of the shellcode meant that it does not take any action to establish Persistence and receiving a score of 1.

Second, the Invoke-Shellcode.ps1 malware was evaluated for each of the proposed characteristics. After submission to virustotal.com, 41% of the AV software successfully identified the software as dangerous, awarding a score of 2. The malware was moderately easy to create and execute as the code is readily available. However, to fully take advantage of the malware, an attacker needs to understand how to prevent it from being written to disk, decreasing the Ease of Creation score to 2. To maximize the effectiveness of the malware, Invoke-Shellcode.ps1 should be executed via the PowerShell command line, which is very unlikely even with the most advanced phishing techniques, giving the malware a Deceitfulness rating of 1. Unlike the simplicity of the windows/shell/reverse_tcp shellcode, the Invoke-Shellcode.ps1 script is part of the larger Powersploit toolkit providing greater flexibility and versatility, receiving a Versatility score of 3. As noted with Deceitfulness, the malware requires some additional actions to activate, awarding a value of 1 for Instantaneousness. Finally, the malware can be injected into other processes, increasing persistence past closing out the PowerShell. Since this requires some action on behalf of the attacker, it receives a Persistence score of 2.

Finally, the Microsoft Office Macro payload was analyzed using the novel rating system. Unlike the other two submissions, the Macro payload successfully evaded all AV software, awarding the highest score of 3 for Stealth. This was expected as the payload lives within the Microsoft Office document, potentially obscuring it from the virus detection tools. Due to the complexity of the generation process and the required expertise, the Macro payload received an Ease of Creation rating of 3. Since malicious software resides within a commonly used document type, it is very likely to avoid detection by even cyber-aware employees the system awards a Deceitfulness rating of 3. In regards to the malware's Versatility, the payload can be easily swapped out for different applications. Additionally, this implementation utilized the Invoke-Shellcode.ps1 script, providing the same Versatility score of 3. To properly execute, the payload does require the user accept the document's Macros, delaying the Instantaneousness of the malware to a score of 2. Finally, a Persistence score of 2 was awarded, as it utilized the same script as Invoke-Shellcode.ps1.

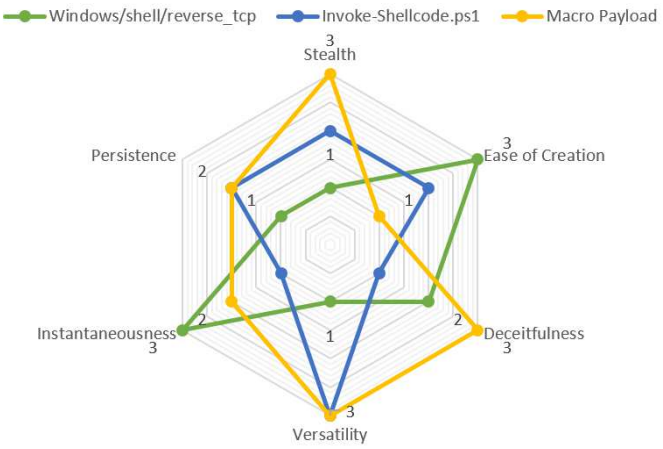

Fig. 3. Radar Chart for all three reverse bind shell malwares.

V. CONCLUSION

Altogether, a novel malware rating system was proposed to account for the increasing prevalence and impact of ransomware and social engineering attacks. The proposed rating system evaluates malware based on Stealth, Ease of Creation, Deceitfulness, Versatility, Instantaneousness, and Persistence. Three PowerShell Reverse Bind Shell Malwares were generation as an example application of the malware rating system. In the future, the novel malware rating system can be applied to an even larger set of malware. Additionally, the categories can be extended from a maximum rating of three to five in order to provide greater granularity during analysis.

ACKNOWLEDGEMENTS

This material is based upon work supported under an Integrated University Program Graduate Fellowship. Any opinions, findings, conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Department of Energy Office of Nuclear Energy.

REFERENCES

[1] R. J. Bagnall and G. French, "The Malware Rating System (MRS)," Veridian, Oakton, 2001.

[2] M. N. Alenezi and H. Alabulrazzaq, "Evolution of Malware Threats and Techniques: A Review," *International Journal of Communication Networks and Information Security,* vol. 12, no. 3, pp. 326-337, 2020.

[3] A. P. Namanya, A. Cullen, I. U. Awan and J. P. Disso, "The World of Malware: An Overview," in *2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud)* , 2018.

[4] H. Oz, A. Aris, A. Levi and A. S. Uluagac, "A Survey on Ransomware: Evolution, Taxonomy, and Defense Solutions," ACM, 2021.

[5] Verizon, "DBIR 2021 Data Breach Investigations Report," Verizon, 2021.

[6] "Alert (AA21-243A) Ransomware Awareness for Holidays and Weekends," U.S. Cybersecurity & Infrastructure Security Agency (CISA) , 31 August 2021. [Online]. Available: https://www.cisa.gov/uscert/ncas/alerts/aa21-243a. [Accessed 10 January 2021].

[7] U.S. Government Accountability Office , "Colonial Pipeline Cyberattack Highlights Need for Better Federal and Private-Sector Preparedness (infographic)," U.S. Government Accountability Office, 18 May 2021. [Online]. Available: https://www.gao.gov/blog/colonial-pipeline-cyberattack-highlights-need-better-federal-and-private-sector-preparedness-infographic. [Accessed 15 July 2021].

[8] The White House, "FACT SHEET: The Biden-Harris Administration Has Launched an All-of-Government Effort to Address Colonial Pipeline Incident," The White House, 11 May 2021. [Online]. Available: https://www.whitehouse.gov/briefing-room/statements-releases/2021/05/11/fact-sheet-the-biden-harris-administration-has-launched-an-all-of-government-effort-to-address-colonial-pipeline-incident/. [Accessed 15 July 2021].

[9] Lockheed Martin, "Cyber Kill Chain," Lockheed Martin, [Online]. Available: https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html. [Accessed 3 July 2020].

[10] I.-C. Mihai, S. Pruna and I.-D. Barbu, "Cyber Kill Chain Analysis," *International Journal on Information Security & Cybercrime,* pp. 37-42, 20214.

[11] M. J. Assante and R. M. Lee, "The Industrial Control System Cyber Kill Chain," SANS Institute InfoSec Reading Room, 2015.

[12] Virustotal, "Virustotal," [Online]. Available: https://www.virustotal.com/gui/home/upload. [Accessed 20 July 2020].

[13] Offensive Security, "Msfvenom Using the Msfvenom Command Line Interface," offensive Security, [Online]. Available: https://www.offensive-security.com/metasploit-unleashed/msfvenom/. [Accessed 15 July 2020].

[14] M. Graeber, "Invoke-Shellcode," 14 December 2016. [Online]. Available: https://github.com/PowerShellMafia/PowerSploit/blob/master/CodeExecution/Invoke-Shellcode.ps1. [Accessed 10 July 2020].

[15] Uneedsec, "Hacking with Powershell, Powersploit, and Invoke-Shellcode," Uneedsec, 15 April 2017. [Online]. Available: https://doxsec.wordpress.com/2017/04/15/hacking-with-powershell-powersploit-and-invoke-shellcode/. [Accessed 10 July 2020].

[16] M. Nelson, "Maintaining Access with Normal.dotm," Enigma0x3, January 2014. [Online]. Available: https://enigma0x3.net/2014/01/. [Accessed 10 July 2020].